Measuring intangibles: Using register-based data as additional source to survey data for measuring R&D, ICT and organisational capital

Marina Rybalka, Statistics Norway, marina.rybalka@ssb.no

Abstract

Intangible capital is becoming increasingly important in economic research, especially due to its contribution to productivity growth. While the core categories of the intangible capital are now widely accepted, their measurement still represents a significant challenge mainly due to the limited data availability. Recently, registry-based sources have been innovatively used within H2020-project GLOBALINTO[1] to develop measures of intangible capital using occupation-based measurement approach based on linked employer-employee dataset and occupation classification ISCO08. To operationalize the concepts of research and development (R&D), information and communication technology (ICT) and organisational capital (OC), the GLOBALINTO-team has applied measures of investments in R&D, ICT and OC that are based on wage costs related to specific skills.

The following paper presents a methodological exercise, which explores whether the GLOBALINTO's occupation-based measurement of R&D can be applied as additional measure to those based on survey data. Given that R&D survey does not cover the smallest firms (with 0-9 employees), this conceptualization might be very useful both, for extension of the existing statistics with respect to small firms), as well as data sources for economic research.

The analysis finds that the occupational-based measure of R&D is more generous in the sense of identifying R&D active firms, the larger firms are. For the small firms and firms in manufacturing and R&D services, the aggregated measures of occupational-based and survey-based R&D are reasonably comparable, while the largest differences are observed for the large firms and firms in ICT services. Based on these findings the main recommendations are that the occupational-based measure of R&D can be used as a complementary *data source to official R&D data to gain on information for the small firms but keeping in mind challenges with R&D definition for the ICT sector. Given that Norway is part of the European statistical system, the proposed methodology can be potentially used by other countries.*

Keywords: *measurement of intangibles, R&D, ICT, survey, register data*

---

[1] The project "New Intangibles for European Growth" (GLOBALINTO) is the EU-financed project with financial support from the Horizon 2020 (H2020) program "The mechanisms to promote smart, sustainable and inclusive growth", the grant 822259 (for more details see https://globalinto.eu/).

1. Introduction

Intangible capital is becoming increasingly important in economic research, especially due to its contribution to firm productivity growth. Corrado et al. (2006) proposed a definition of intangible capital that is now prevalent in economic literature and comprises three broader categories, which are: (1) computerized information, (2) innovative property and (3) economic competencies. While the core categories of the intangible capital are now widely accepted, their measurement still represents a significant challenge to statisticians and economists at large. This is mainly due to the limited data availability.

Several EU-financed projects, such as INTAN, COINVEST, INNODRIVE, and SPINTAN, have proposed new measures to evaluate the size and contribution of intangible capital by using existing national accounts data at sectoral level or developing registry-based methodologies to create microdata sets, which led to a number of empirical studies (Corrado et al., 2009a, 2009b; Corrado, et al., 2016; Corrado, et al., 2019; Ilmakunnas & Piekkola, 2014; Piekkola, 2011; Roth, 2010; Roth & Thum, 2013; van Ark et al., 2009).

Recently, registry-based sources have been innovatively used within H2020-project GLOBALINTO[2] to develop measures of intangible capital (Piekkola et al., 2021), using *occupation-based* measurement approaches on linked employer-employee datasets and International Standard Classification of Occupations (ISCO-08) (ILO, 2012). To operationalize the concepts of research and development (R&D), information and communication technology (ICT) and organisational capital (OC), the GLOBALINTO-team (including Norway) has applied novel measures of investments in R&D, ICT and OC that are based on firm-level wage costs related to specific skills and occupations. Detailed occupational data for Norway are available for the period 2008–2019, while e.g. Finland and Denmark have this information available since 1999.

The following paper presents a methodological exercise, which explores whether the GLOBALINTO proposed *occupation-based* measurement of R&D can be applied as an

[2] The project "New Intangibles for European Growth" (GLOBALINTO) is the EU-financed project with financial support from the Horizon 2020 (H2020) program. GLOBALINTO seeks to develop new measures of intangible assets in order to analyze the role of knowledge production and diffusion for innovation, productivity and growth (for more details see https://globalinto.eu/).

additional source for measuring R&D at the firm level. Given that R&D surveys do not cover the smallest firms (i.e. with 0-9 employees) while the majority of Norwegian firms are small, this conceptualization might be very useful both for extension of the existing statistics (e.g. using the register-based methodology as instruments for R&D in small firms or constructing population of R&D active firms), as well as data sources for economic research. Given that Norway is a part of the European statistical system, the proposed methodology can be potentially used by other European countries.

Detailed conceptualization of GLOBALINTOS's approach to measuring investments in R&D, ICT and OC is presented in section 2. Section 3 examines how comparable R&D measures reported in official statistics are with estimates based on the occupational data both with respect to identification of R&D active firms and the scope of their R&D activity. While section 4 presents discussion and main recommendations.

2. GLOBALINTOS's approach to measuring investments in R&D, ICT and OC based on occupational data and descriptive statistics

The measurement of intangibles in GLOBALINTO project is conducted in several steps and follows several statistically important principles, with the goal of fitting the analysis into the wider context of established statistical classifications. First, methodologically, the measurement of intangible investment is based on employee data, where in particular, the occupations serve as the main criteria to identify “innovative labour”. For each of the relevant intangible capital types, a list of relevant occupations was prepared. Following (Bloch et al., 2021), Box 1 lists the relevant occupations that are engaged in intangible investment according to the GLOBALINTO's methodology (in bold OC=organizational occupation, R&D=R&D occupation and ICT= ICT occupation).

Second, individuals' wages and man-hours have been aggregated from employee (individuals) to the employer (firms) level data by core occupational groups (R&D, organizational, ICT and other work), which allow generation of intangible investments and linkage to other firm-level data. The linked employee-employer dataset (LEED) is then used in the detailed firm-level analysis (e.g. Piekkola et al., 2021a and 2021b, and Raknerud and Rybalka, 2021).

Box 1 GLOBALINTO Intangibles Assets occupations (based on ISCO-08*).

1 Managers
11 Managing Directors and Chief Executives
12 Administrative and Commercial Managers
121 OC Business Services and Administration Managers
122 Sales, Marketing and Development Managers
1221 OC Sales and Marketing Managers
1222 OC Advertising and Public Relations Managers
1223 R&D Research and Development Managers
13 Production and Specialized Services Managers
131 OC Production Managers in Agriculture, Forestry and Fisheries
132 OC Manufacturing, Mining, Construction and Distribution Managers
133 ICT Information and Communications Technology Services Managers
134 OC Professional Services Managers
14 Hospitality, Retail and Other Services Managers

2 Professionals
21 Science and Engineering Professionals
211 R&D Physical and Earth Science Professionals
212 R&D Mathematicians, Actuaries and Statisticians
213 R&D Life Science Professionals
214 R&D Engineering Professionals (excluding Electrotechnology)
215 R&D Electrotechnology Engineers
2151 Electrical Engineers
2152 R&D Electronics Engineers R&D
2153 ICT Telecommunications Engineers
216 R&D Architects, Planners, Surveyors and Designers
22 Health Professionals
221 R&D Medical Doctors
222 R&D Nursing and Midwifery Professionals
223 Trad. and Complementary Medicine Professionals
224 Paramedical Practitioners
226 R&D Other Health Professionals
23 Teaching Professionals
24 Business and Administration Professionals
241 OC Finance Professionals
242 OC Administration Professionals
243 Sales, Marketing and Public Relations Professionals
25 ICT Information and Communications Technology Professionals
26 Legal, Social and Cultural Professionals

3 Technicians and Associate Professionals
31 Science and Engineering Associate Professionals
311 R&D Physical and Engineering Science Technicians
312 Mining, Manufacturing and Construction Supervisors
313 Process Control Technicians
314 R&D Life Science Technicians and Related Associate Professionals
315 Ship and Aircraft Controllers and Technicians
32 Health Associate Professionals
321 R&D Medical and Pharmaceutical Technicians
33 Business and Adm. Associate Professionals
34 Legal, Social, Cultural Associate Professionals
35 ICT Information and Communications Technicians

Note: *International Standard Classification of Occupations.

Figure 1 shows the evolution of employment shares (in full-time equivalents) for each type of producing intangible capital labour (i.e., R&D, OC and ICT) for Norway from 2008 to 2019. As demonstrated by the figure, the share of R&D employees in Norway was rapidly increasing until 2015, but then decreased until 2018, probably because many R&D suppliers to the oil industry were affected by low oil-prices after 2015, and then increased again in 2019. The share of ICT employees has been increasing through the whole period from 4.3% in 2008 to 5.3% in 2019, while the share of managers has stagnated at the level of 8% after an increase in the start of the period under investigation.

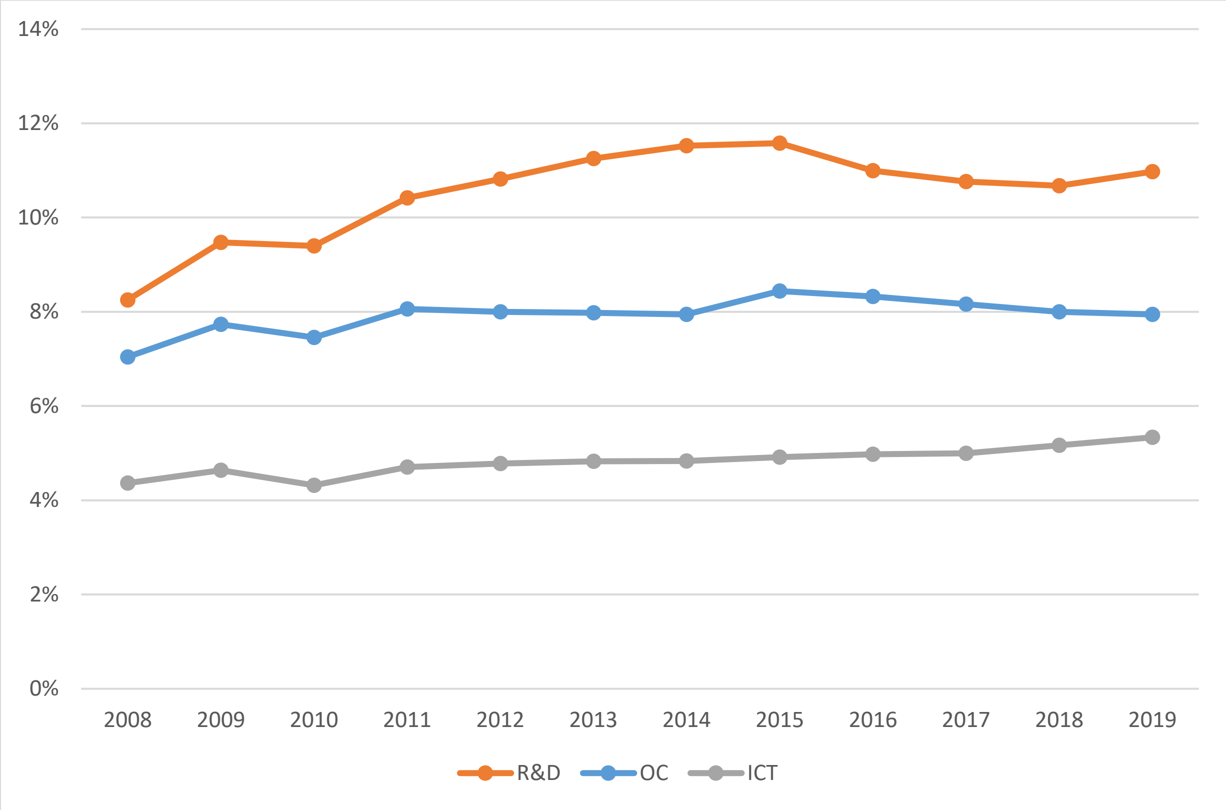

Figure 1. Employment shares within each intangible type for Norway, 2008-2019.

Table 1 reports the average shares for different types of labour over 2008-2019 by nine technology types used in GLOBALINTO (based on classifications of economic activities by technological intensity by Eurostat)[3]. Not surprisingly, R&D services are the most R&D intensive with respect to use of R&D employees (in full-time equivalents), ICT services are the most ICT intensive with respect to use of ICT employees, and management services are the most OC intensive with respect to use of OC employees. High-tech and medium-high-tech manufacturing industries are also highly R&D intensive. They also use relatively high shares of organisational employees, while knowledge intense services (KIS) use relatively high shares of ICT employees.

[3] With slightly modifies definition of sub-groups of Knowledge Intensive Services (KIS) that corresponded better to the three types of intangible assets studied at the microlevel by GLOBALINTO. For Eurostat's description of 2-digit NACE-codes (where NACE is a short name for "Nomenclature statistique des Activités économiques dans la Communauté Européenne", that is the industry standard classification used in the European Union) see: https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an3.pdf.

Table 1. Average employment shares within each intangible type for Norway by industry group.

Industry group	NACE-codes*	R&D	OC	ICT
High-tech	21, 26	**39.0 %**	**11.4 %**	**4.9 %**
Medium-high-tech	20, 27-30	**25.9 %**	**11.2 %**	1.6 %
Medium-low-tech	19, 22-25, 33	11.7 %	8.4 %	1.1 %
Low-tech	10-18, 31-32	5.0 %	8.2 %	1.3 %
KIS (Knowledge Intensive Services)	50-51, 58-66, 69-75, 78, 80, 84-93	7.1 %	6.9 %	**7.5 %**
ICT services	62-63	8.2 %	9.2 %	**63.6 %**
R&D services	71-72	**64.2 %**	8.5 %	3.2 %
Management services	69-70, 73	2.9 %	**25.4 %**	2.9 %
Low KIS	45-47,49,52-53,55-56,68, 77,79,81-82,94-96,97-99	2.6 %	5.7 %	1.2 %

* Statistical Classification of Economic Activities in the European Community

3. Comparison of R&D measures reported in official statistics with estimates based on the occupational data

Next exercise is to examine how comparable R&D estimates based on occupational data and GLOBALINTO's approach and R&D amounts reported in official statistics. For this purpose, I use R&D expenditures (both for intramural and purchased R&D) reported in R&D surveys and in the applications for R&D tax credits (TC), an additional register data source that is available in Norway and used by National Accounts at Statistics Norway in their estimations of R&D measures (Sørensen, 2016, in Norwegian). This analysis serves both as a valuable tool to validate the GLOBALINTO's methodology and as an exploration of whether this method could be used in estimating R&D expenditures both for purposes of statistics (e.g. by National Accounts) and/or economic analysis.

Sample description

As Table 2 shows, the number of firms in Norwegian R&D survey varies between 4 and 6 thousand, while tax credits applications give additional information for about 3-4 thousand firms with positive R&D, the majority of those are small firms (see e.g. figure 2.10 in Benedictow et al., 2018). The main advantage of using the register data on occupations is that they cover the whole population of firms and hence contain valuable information on small firms that are not covered by R&D survey. Though the tax credits applications also cover small firms, this additional data source for R&D is not available in many other countries. Moreover, it does not cover all small firms, but only those who applied and received R&D tax credits.

Table 2. Sample description by data source, 2008-2019.

	R&D survey				R&D survey+TC*				LEED** data used in GLOBALINTO	
Year	No.obs	R&D>0	Internal R&D>0	Purchased R&D>0	No.obs	R&D>0	Internal R&D>0	Purchased R&D>0	No.obs	Internal R&D>0
2008	5568	1633	1469	738	8335	4135	3984	1682	140791	14170
2009	4392	1426	1259	631	7015	3833	3686	1419	147621	14355
2010	6118	1634	1454	699	8531	3902	3744	1461	145696	13432
2011	4566	1399	1256	599	7426	4055	3933	1699	151622	13945
2012	5853	1705	1560	666	8888	4469	4331	1842	159801	14654
2013	4415	1503	1370	629	7960	4678	4544	1992	167237	15236
2014	4443	1916	1452	669	8407	5300	4955	2220	173728	15627
2015	5089	1946	1823	801	9381	5595	5461	2467	190045	17285
2016	4535	1931	1797	740	8582	5628	5506	1793	196853	17650
2017	5527	2297	2162	898	8505	5407	5300	5566	202722	18201
2018	4940	2121	2010	835	8038	5351	5260	5139	205073	18896
2019	5773	2365	2257	901	8491	5223	5141	5730	215394	19661

* The applications for R&D tax credits; ** The linked employer-employee data.

The main limitation of the occupation-based R&D measure is that it is based on the wage costs of R&D personal and, hence, does not cover all expenditures to internal R&D. However, from the tax credit applications I observe that about in average 65-70 per cent of the firm's internal R&D expenditures are the personal costs and this share is even higher for the smallest firms. This measure says nothing about purchased R&D. The usage of occupational data at the firm level has also an important requirement of availability of a link between individuals and the companies they are employed at.

Figure 2 shows development in aggregated R&D measures from R&D survey and R&D full-time equivalents (FTEs) from R&D survey versus those from occupation data used in GLOBALINTO by index (2008=100). While data in official statistics demonstrate a negative trend just after financial crises of 2008 and a positive trend for all measures since 2011, the GLOBALINTO's measure of R&D FTEs based on occupations has a different development. It was rapidly increasing until 2014, but then decreased after 2015 (probably because many R&D suppliers to the oil industry were affected by low oil-prices), and then increased again in 2018-2019. Such difference in the development of R&D FTEs from the two data sources might be due to different definitions, but also due to not including small firms in the R&D survey.

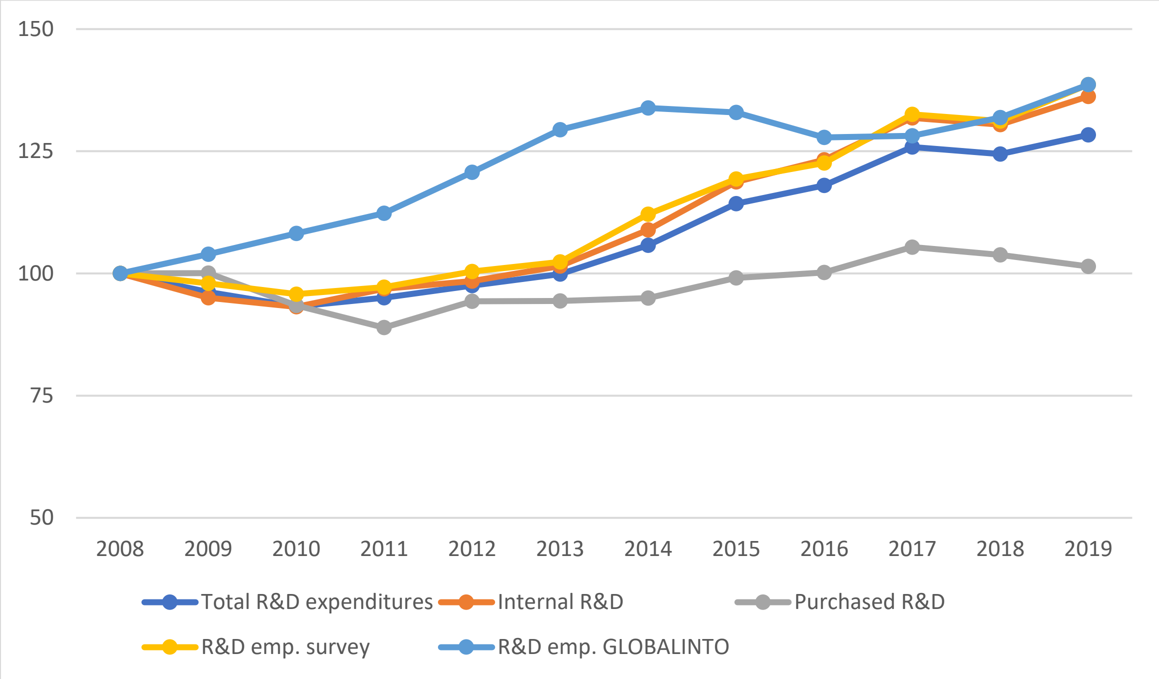

Figure 2. R&D expenditures in fixed prices (by type) and R&D full-time equivalents in GLOBALINTO vs survey (index 2008=100), 2008-2019.

To investigate these differences closer, I have merged the available datasets and studied, first, whether the same firm was defined as an R&D active firm in these datasets or not; and second, if definitions differed, what was the scope of misspecification by GLOBALINTO compared to the official data sources.

Comparison between GLOBALINTO and survey data by type of industry

Figure 3 reports share of companies with positive internal R&D by type of industry and data source. As it is higher selection of R&D active firms into R&D survey, the shares based on R&D survey are likely higher than the true shares in the whole economy, and hence, are assumed to be the *upper bound* of the true shares in the economy. Shares based on R&D survey+TC (tax credit applications) are calculated with respect to the *whole population* of the firms. As only part of R&D active firms is represented in both datasets, the shares based on the combined dataset are lower than the true shares in the whole economy, and hence, are assumed to be *lower bound* of the true share in the economy. Then shares based on GLOBALINTO measure are expected to be in between of these two.

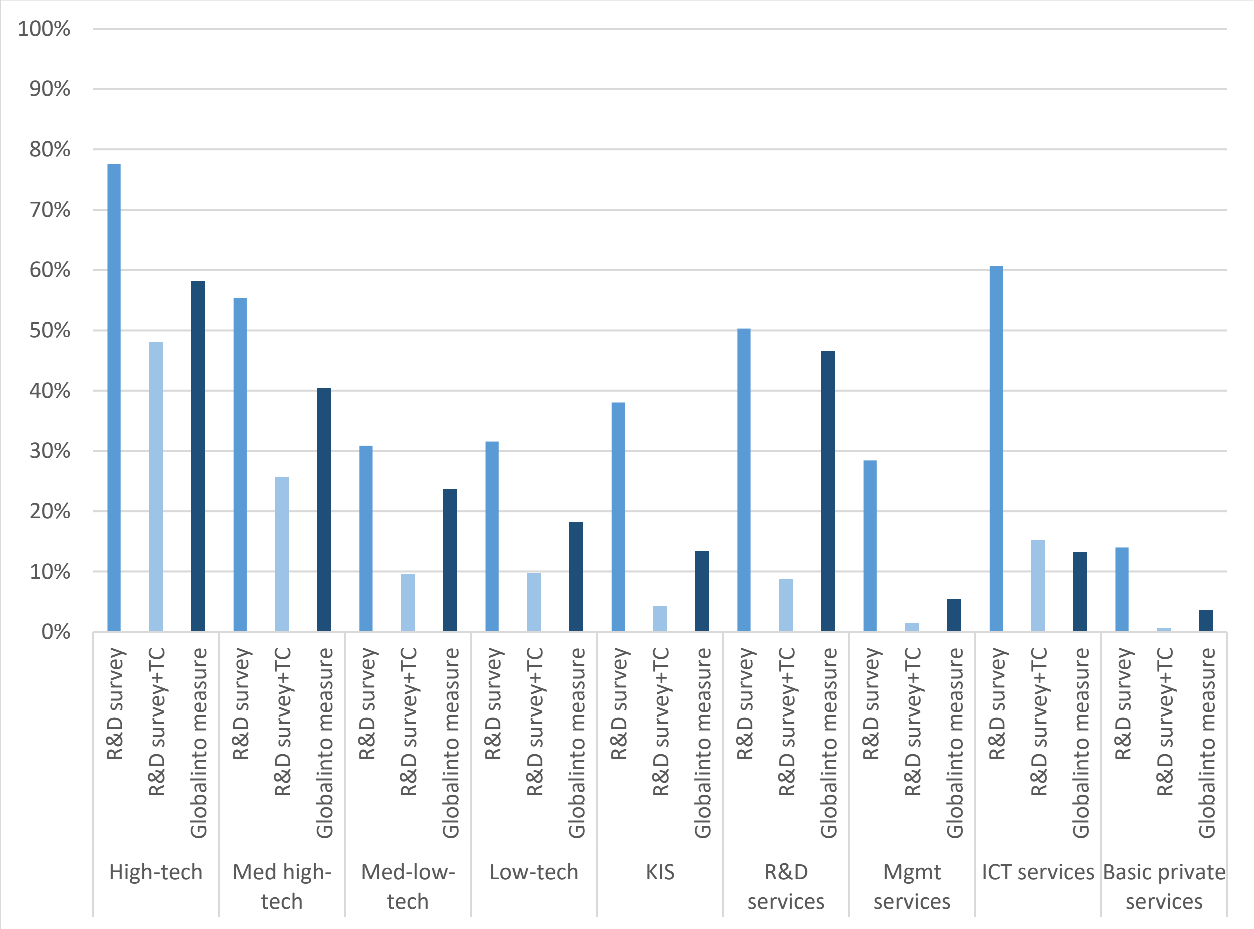

Figure 3. Share of companies with positive internal R&D by type of industry and data source.
Note: For definition of industry types see Table 1.

According to figure 3, the shares based on GLOBALINTO measure are indeed in between the other two measures for most industries, but not for the ICT services where the share based on GLOBALINTO measure is lowest. This observation raises a question whether the occupation-based definition of R&D personal in the ICT intensive firms should contain a part of ICT personal that might conduct R&D activities.[4]

From figure 3 I also observe that shares based on GLOBALINTO measure are closest to those based on R&D survey in high-tech manufacturing and in R&D services. In general, it seems that GLOBALINTO measures are more comparable with those based on the official R&D data sources in manufacturing than in services.

[4] To provide an answer on that question, one could e.g. run a regression on the firm being an R&D active firms on the LHS and different occupational codes for R&D and ICT personal to see which of the ICT-codes are important for firms in the ICT services of being defined an R&D-active firm. However, this type of the analysis is out of the scope of this paper.

Comparison between GLOBALINTO and survey data by firm size

Figure 4 reports corresponding shares with positive internal R&D by firm size (in total number of employees) and data source. Again, the shares based on GLOBALINTO measure are expected to be in between of the two shares based on the R&D survey (the upper bound) and on the R&D survey+TC (the lower bound).

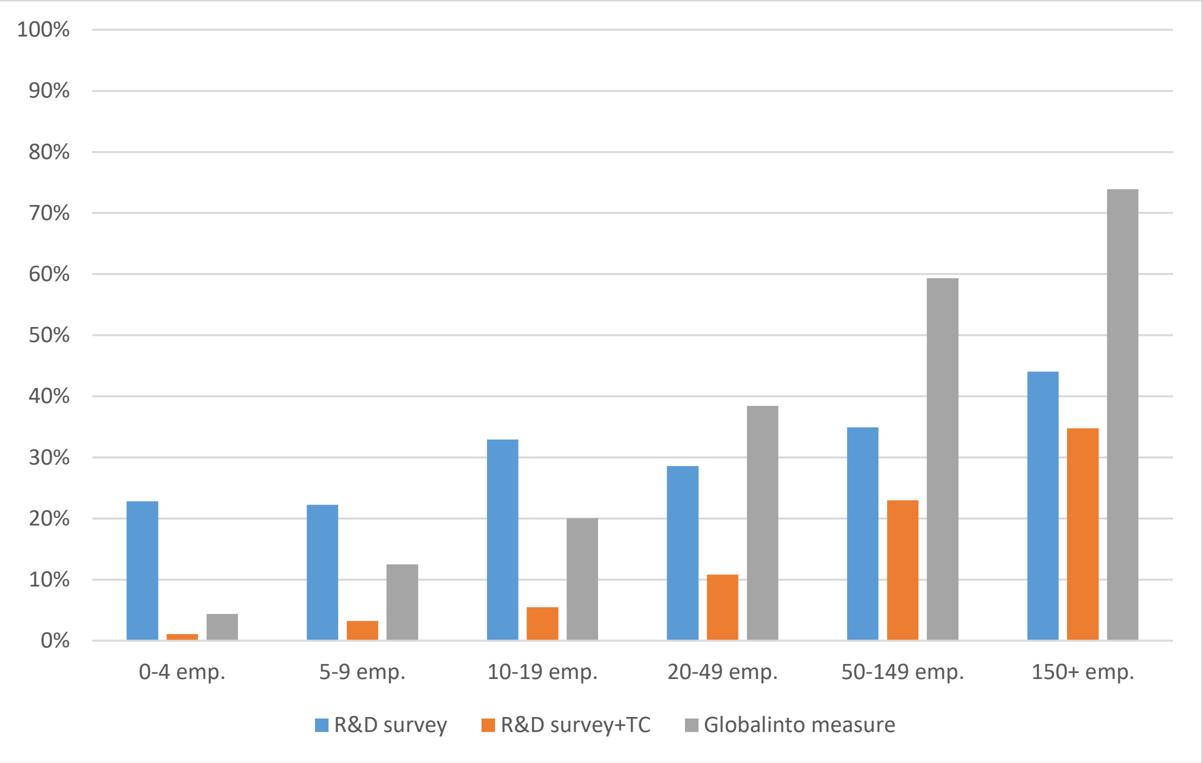

Figure 4. Share of companies with positive internal R&D by firm size and data source.

According to figure 4, the shares based on GLOBALINTO measure are in between the other two measures for the small firms (0-19 employees), but not for the medium and large firms. It means that GLOBALINTO measure is more generous in its definition of R&D active firms when it yields occupations in the larger firms and, hence, should be rather used in combination with the data from R&D survey (especially for the large firms). However, for the small firms the GLOBALINTO measure seems to work quite well.

Two types of misspecifications by firm size and industry group

As observed above, the GLOBALINTO measure is more generous in its definition of R&D active firms when it yields R&D related occupations. I check then, how often the GLOBALINTO method says that the firm is R&D-active (i.e. the wage costs to R&D

personal based on occupational codes are positive) while internal R&D in the R&D survey is zero. I call it the misspecification of type I. Figure 5 reports the misspecification-I rates by firm size and industry group.

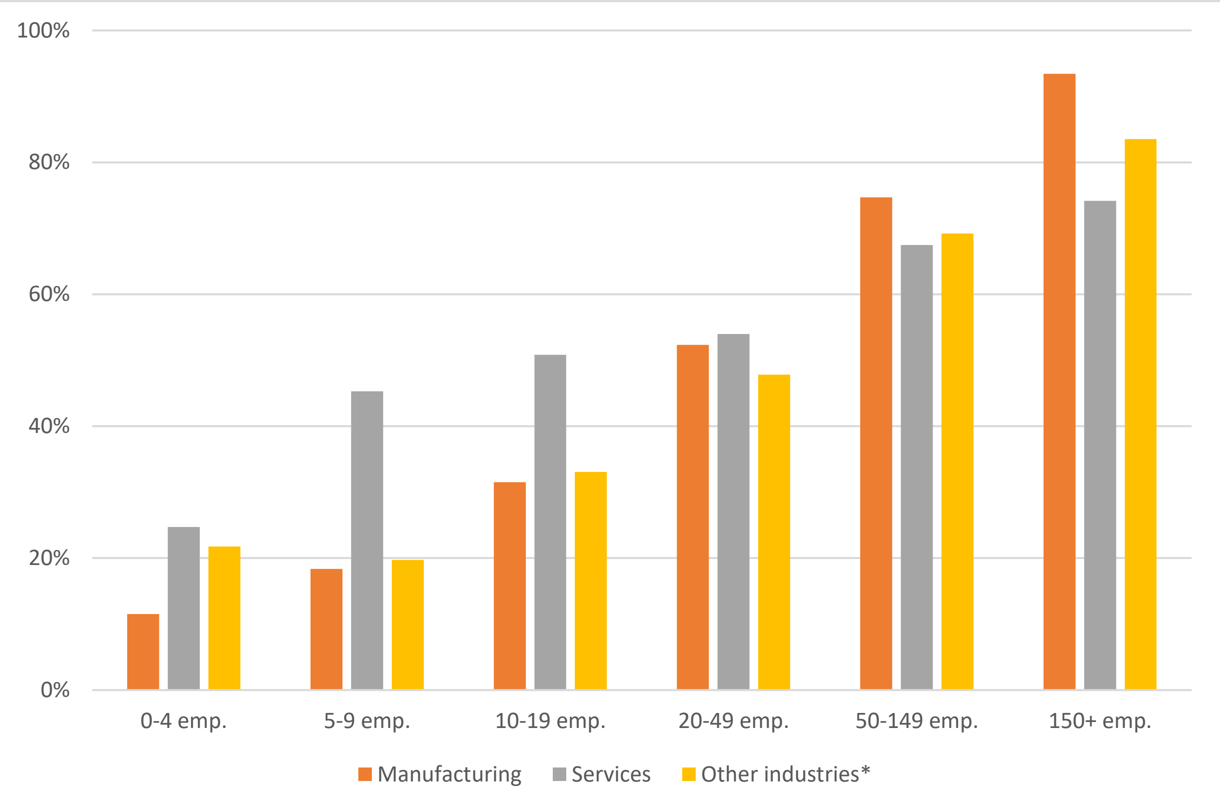

Figure 5. Misspecification-I rates (GLOBALINTO>0, internal R&D=0) by firm size and industry group.
Note: * Industry groups G (Wholesale and retail trade; repair of motor vehicles and motorcycles), H (Transportation and storage), I (Accommodation and food service activities), L (Real estate activities).

According to figure 5, the problem of misspecification of type I (GLOBALINTO>0, internal R&D=0) is less important for the small firms, but more serious the larger firms are. It is less common for small firms in manufacturing than in service industries. However, the opposite is observed for the largest firms (with more than 150 employees).

Further, I check how often the GLOBALINTO method says that the firm is R&D-inactive (i.e. the wage costs to R&D personal based on occupational codes are zero) while internal R&D in the R&D survey is positive. I call it the misspecification-II. The rates of the misspecification of type II by firm size and industry group are reported in figure 6. According to the reported numbers, type II of misspecification is much rarer than misspecification of type I, being equal in average to 7 per cent. The rates are higher for firms in services than in manufacturing regardless the firm size and are the lowest for large firms regardless of industry group.

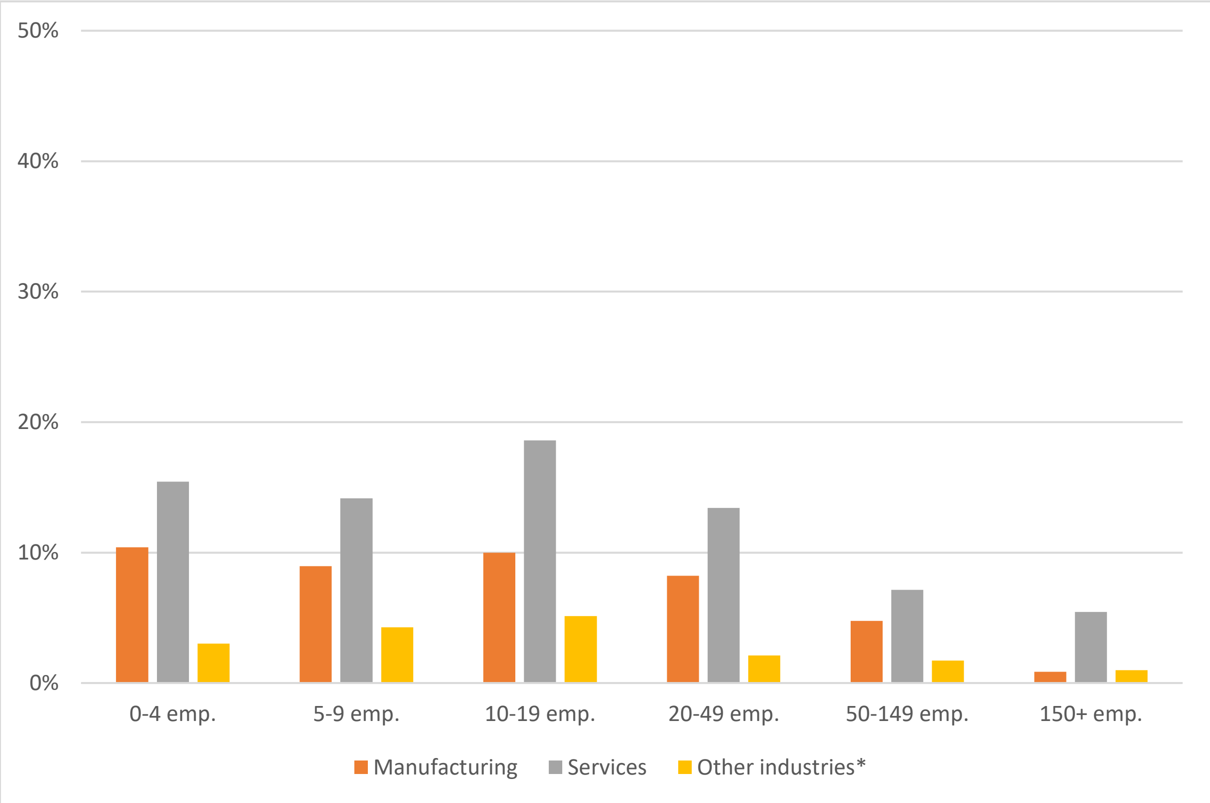

Figure 6. Misspecification-II rates (GLOBALINTO=0, internal R&D>0) by firm size and industry group. Note: * Industry groups G (Wholesale and retail trade; repair of motor vehicles and motorcycles), H (Transportation and storage), I (Accommodation and food service activities), L (Real estate activities).

Comparison of GLOBALINTO measure of R&D expenditures with survey data

Finally, I check the bias in the GLOBALINTO measure of R&D expenditures based on the wage costs for R&D personal from amount of internal R&D reported in the official data at the firm level. While GLOBALINTO assumes that only a part of work of R&D personal is dedicated to the innovative work (i.e. the assumed share is equal to 0.7), I take wage costs without any correction for the comparison here to explore an initial bias.

The main observations from figure 7 that reports these biases are:

- GLOBALINTO measure of R&D expenditures based on labour costs (without applying any correction and/or factor multiplier) is between 0.6-0.95 of observed internal R&D and between 0.6-0.7 of observed total R&D for the smallest firms (0-4 employees). This corresponds well to the shares reported in tax credit (TC) applications.

- If to apply correction of 0.7 for R&D labour costs (i.e. for innovative labour shares), the GLOBALINTO measure would be close to the observed R&D amounts for firms with 5-9 and 10-19 employees.
- The bias increases by firm size and is extremely high for the firms with more than 50 employees, the firms that are fully covered by an R&D survey (note that the bias is shown on the right axis).

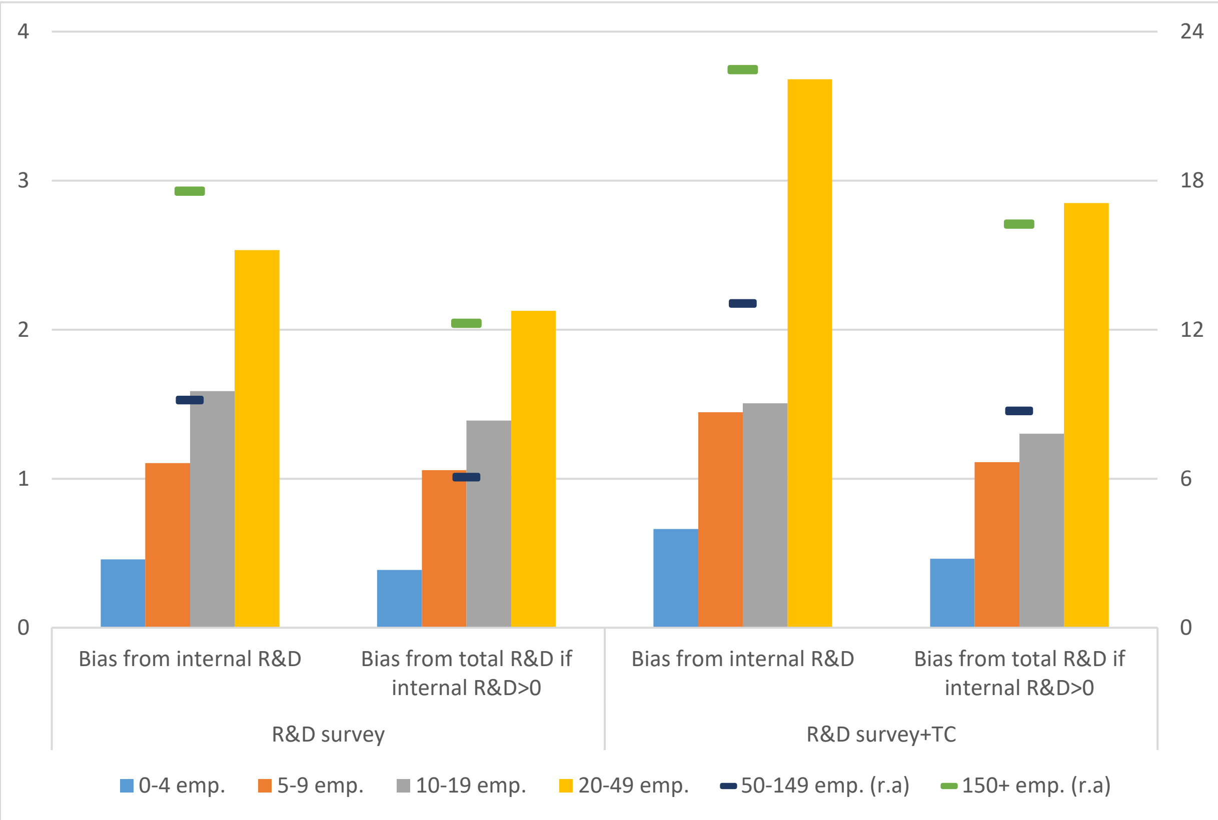

Figure 7. Bias of GLOBALINTO from other measures of R&D expenditures by data source and firm size.

4. Discussion and recommendations

The main advantage of using linked employer-employee data is full coverage of firms. Both private and public sectors are represented (and they use unified occupational codes) and all industries and size groups are covered. That makes it possible to conduct more advanced analyses than based only on the firms represented in the R&D survey.

The analysis finds that the occupational-based measure of R&D is more generous in the sense of identifying R&D active firms. This misspecification is larger the larger firms are. For the small firms and firms in manufacturing and R&D services the aggregated measures of occupational-based and survey-based R&D are reasonably comparable, while the largest differences are observed for the large firms and firms in ICT services. Based on these findings the main recommendations are that the occupational-based measure of R&D can be used as a complementary data source to official R&D data to gain on information for the small firms. However, one should keep in mind challenges with definition of R&D personal in the ICT sector. Given that Norway is part of the European statistical system, the proposed methodology can be potentially used by other countries.

References

Benedictow A., Bjøru, E. C., Eggen, F. W., Norberg-Schulz, M., Rybalka, M. & R. Røtnes (2018): Evaluation of SkatteFUNN [Report 18-2018], Samfunnsøkonomisk Analyse, Oslo.

Bloch, C., Piekkola, H., Derek, T. & Rybalka, M. (2021): Valuing intangible assets at the firm level – development of an occupation-based approach [H2020 GLOBALINTO D3.4].

Corrado, C., Hulten, C., & Sichel, D. (2006): Intangible Capital and Economic Growth [Finance and Economics Discussion Series, No. 2006–24]. Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.

Corrado, C., Hulten, C., & Sichel, D. (2009): Intangible Capital and U.S. Economic Growth. *Review of Income and Wealth*, 55(3), 661–685.

Corrado, C., Jaeger, K., & Jona Lasinio, C. (2016): Measuring Intangible Capital in the Public Sector: A Manual [EU 7 FP NO.612774].

Corrado, C., Haskel, J., & Jona-Lasinio, C. (2019): Productivity growth, capital reallocation and the financial crisis: Evidence from Europe and the US. *Journal of Macroeconomics*, 61.

Ilmakunnas, P., & Piekkola, H. (2014): Intangible investment in people and productivity, *Journal of Productivity Analysis*, 41(3), 443–456.

ILO (2012): *International Standard Classification of Occupations: ISCO-08*. Volume 1. International Labour Office, Geneva.

Piekkola, H. (2011): Intangible capital: The key to growth in Europe. *Intereconomics*, 46(4), 222–228.

Piekkola, H. (2020): Intangibles and innovation-labor-biased technical change. *Journal of Intellectual Capital*, 21, 649-669.

Piekkola, H., Bloch, C., Redek, T. & M. Rybalka (2021a): Productivity growth and the intangible divide, Chapter 5 in Muren D.M., M. Koman and T. Redek (eds.): *Achieving growth that matters*, Casnik Finance press, Ljubljana.

Piekkola, H., Bloch, C., Redek, T. & M. Rybalka (2021b): Intangibles from innovative work – their valuation and technological change [H2020 GLOBALINTO Working paper no. 5.3].

Raknerud, A. & M. Rybalka (2021): Innovation-driven productivity growth: the role of ICT capital and effective labour [H2020 GLOBALINTO Working paper no. 5.5], Statistics Norway, Oslo.

Roth, F. (2010): Measuring innovation—Intangible capital investment in the EU. *Intereconomics*, 45(5), 9–13.

Roth, F., & Thum, A.-E. (2013): Intangible capital and labor productivity growth: Panel evidence for the EU from 1998-2005. *Review of Income and Wealth*, 59(3), 486–508.

Sørensen K.Ø. (2016): Forskning og utvikling i Nasjonalregnskapet: Dokumentasjon av arbeidet med FoU i hovedrevesjonen 2014 [Notater 2016/32], Statistics Norway.

van Ark, B., Hao, J. X., Corrado, C., & Hulten, C. (2009): Measuring intangible capital and its contribution to economic growth in Europe [EIB Paper No. 3/2009]. European Investment Bank, Economics Department.