Firm human capital investment, wage inequality and employment

Authors:

Cecilia Jona-Lasinio and Francesco Venturini (UNIPG)

Deliverable: 5.4

Date: September 2022

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 101004776

Abstract

Using company-level data from three waves of the Continuing Vocational Training Survey (2005, 2010 and 2015), this paper provides an overview on European firms implementing training and the magnitude of their training effort. We illustrate in the descriptive statistics that there are significant differences in the wage and employment performance across companies in relation to the digital content of their production and training activities. We conduct a regression analysis documenting that a wage premium of 9% is associated with companies undertaking training and that an additional 8% is paid by firms arranging training for IT skills-intensive workers. The latter effect is pervasive across sectors and is not strictly related to industry exposure to the digital transformation. We seek to address both the simultaneity and the selectivity issue of our data implementing an instrumental variable regression and a propensity score matching procedure.

Keywords: training, IT upskilling, wage premium, European firms

Acknowledgments: The authors wish to thank Mikkel Barslund (KU Leuven), Klavs Ciprikis (ESRI), Uyen Nguyen (LISER), Fabrizio Pompei (UNIPG) and an anonymous UNTANGLED referee for valuable comments.

Please refer to this publication as follows:

Cecilia Jona-Lasinio & Francesco Venturini (2022). *Firm human capital investment, wage inequality and employment* (Deliverable 5.4).] Leuven: UNTANGLED project 1001004776 – H2020.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No. 1001004776.

Document control sheet

Project Number:	101004776
Project Acronym:	UNTANGLED
Work-Package:	WP5
Last Version:	Version 5
Issue Date:	12/09/2022

Classification

Draft	
Final	x
Confidential	
Restricted	
Public	x

Table of contents

Firm human capital investment, wage inequality and employment

1. Introduction

Training is seen as a tool for improving job opportunities and work conditions of employees, and for increasing company productivity (Becker, 1964). The need for training increases with the pace of technological change, which makes the formal education of younger workers obsolete and the experience of more tenured employees unfit to contribute to company performance (Bartel & Sicherman, 1998). In imperfect labour markets, namely when companies have monopsonistic power or employers are not mobile across companies and/or jobs, firms have incentives to bear the cost of training not only when it is designed to build firm-specific skills, but also when training targets more general competences as companies can capture part of the increased workers' productivity (Acemoglu & Pitschke, 1998a; 1998b; 1999). This type of inefficiency, associated with training, could be higher when the company is expected to innovate as workers are willing to accept lower wages today with the prospect of higher wages in the future (Acemoglu, 1997).

Digitalisation is one of the most important transformations affecting modern societies. A long stream of works has studied, both in theoretical and empirical terms, the interplay between this form of technical change and labour demand. In a pioneering study, Krueger (1993) finds that workers using the computer earn 15% more than non-user workers and that the expansion of computer use would explain around one third of the wage premium of educated workers in the last decades. Author et al. (1998) find a persistent skill upgrading in the US economy, especially in more computer-intensive industries. Acemoglu (1998) explains these trends as a consequence of the long-term increase in educated labour supply which, endogenously, stimulates the development and the adoption of new technologies, such as ICT, which are human capital-intensive (see Acemoglu, 2002 for a review of the early literature). A later generation of studies (Autor et al., 2003) point to the intangible nature of the latest wave of digital technologies which would displace workers performing routinised tasks, since they are repetitive and hence are codifiable in software (routine-biased technical change; see Goos et al., 2021 for a recent assessment).

Recent evidence indicates that 16% of European workers is exposed to skill-displacing technical change and that this effect mostly transits through an increasing task complexity (McGuinness et al., 2021). Technical change is found to mostly affect highly educated workers, stimulating the company provision of training and workplace learning, and ultimately promoting workforce upskilling. According to Cedefop (2016), 71% of European workers claim to need basic or moderate ICT skills to implement their job, whilst another 14% require advanced digital skills. However, there is wide variation in the requirement of digital skills, especially advanced across various types of productions, from 51% of workers in the ICT sector to 5% in the Accommodation sector.

Digitalisation has disruptive effects on production activities (Gal et al., 2019), forcing firms to change manual and highly routinised tasks, such as assembling, delivering, etc., as well as cognitive and non-routinised tasks, such as management, R&D, product design and customisation, etc. A recent report by the European Investment Bank (EIB, 2022) documents that companies using digital technologies are more likely to provide vocational training and that this investment increases with the complexity of the digital technologies adopted.

This study sheds light on the wage effect of employer-provided training in Europe. As long as training regenerates workers' competences, firms undertaking these measures should be able to pay higher wages compared to firms without training. On this basis, training can be seen as an intangible investment fuelling wage dispersion across companies. Our main goal is to ascertain (i) whether this process is related to the company exposure to the digitalisation process measured at industry level; and (ii) whether there is a differential effect between training targeted at digital skills-intensive jobs and training targeted at more general competences.

Using data for 112,000 European companies, collected from three waves of the EU Continuing Vocational Training Survey (2005, 2010 & 2015), we document wide gaps in wage (and occupational) levels among companies in relation to the digital content of their production and training activities. Specifically, we estimate a wage premium of 9% for companies undertaking training and an additional 8% for those firms arranging training for IT skills-intensive workers. We document that these results are robust to the procedure of estimation adopted, to selectivity and simultaneity issues, and to the set of control variables used.

Admittedly, the main caveat of our analysis is to use company-level data to infer the effect of training on workers' remuneration. In other words, we quantify the effect of training policies on the average wage paid by the firm, which covers both trainees and workers not engaged in training. This implies that the estimated impact is a net effect across workers and that, for instance, it may be affected by substitution effects (hires and fires). On the other hand, company-level data is less affected by selectivity issues than employee-level data. Indeed, skilled workers respond to wage differences and move across jobs and firms paying higher wages, increasing the company incentives to offer training in order to keep them. Our data allows us to exploit information on company access to public funds to identify the wage impact of training, as we extensively discuss in presenting our instrumental variables (IV) results.

Our work makes a threefold contribution to the literature. First, we provide novel evidence on the drivers of wage effect in Europe focusing on the role played by training in the digitising economy. Complementary evidence is offered by Brunello and Wruuck (2020) who review the main training policies pursued by European companies, identifying the main factors hindering investment in training (see also Cedefop, 2019). Second, we shed light on the differential effect on wages of IT training with respect to other forms of training. Prior works focusing on wage premia (O'Mahony et al., 2008) or wage polarisation (Michael et al., 2014) looked at the earlier diffusion of ICT. More recent studies look at the labour market effects of automation (Acemoglu & Restrepo, 2018) and the diffusion of Artificial Intelligence, AI (Webb, 2020; Acemoglu et al., 2022). Lastly, we complement with company-level evidence the stream of industry-level studies assessing the economic impact of training, defined as intangible investment, through growth accounting methods (O'Mahony, 2012; Squicciarini et al., 2015).

The paper is organised as follows. Section 2 describes data. Section 3 lays down the empirical model. Section 4 presents the econometric results and, finally, Section 5 concludes.

2. Data

The analysis is conducted using company-level data extracted from the waves of the EU Continuing Vocational Training Survey (CVTS) for the years 2005, 2010 and 2015. We use the version of the CVTS dataset releasing information on the sector of production at a 2-digit level (Nace Rev. 2 classification). The dataset provides information on nation-wide representative samples of companies with employees ranging from 10 to 999 units, from the following European countries: Bulgaria, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Italy, Latvia, Norway and United Kingdom.[1]

Since we are interested in the company response to the digital transformation, we classify firms in relation to the digitalisation of their production, using the taxonomy provided by OECD (Calvino et al., 2018, Table 3). We use the global classification of digitalised sectors built for the

[1] For info on the CVTS dataset see: https://ec.europa.eu/eurostat/web/microdata/continuing-vocational-training-survey. This dataset covers enterprises with 10 or more employed in the business sector for the years 2010 and 2015, and companies in the industry and service sectors for the year 2005. A larger version of the dataset includes microdata from 24 countries (Belgium, Bulgaria, Czechia, Denmark, Germany, Estonia, Spain, France, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden, United Kingdom and Norway) but with an industry breakdown to one-digit level only.

period 2001-03, which precedes the time interval covered by our analysis thus mitigating reverse causality problems as companies with training may self-select and move towards sectors involved in a more intensive digitalisation process. Another possibility is that firms operating in highly digitalised industries are structurally more productive, can afford to pay higher wages and invest more resources in training to keep up with the advances in digital technologies. The OECD categorisation reflects the intensity in the usage and exploitation of digital technologies at industry level along different dimensions: the share of ICT tangible and intangible (i.e. software) investment; the share of purchases of intermediate ICT goods and services; the per-worker stock of installed robots; the share of ICT specialists out of the total workforce; and the share of turnover from online sales. We consider as highly digitalised those sectors lying at the top quartile of the usage of the four types of ICT technologies described, and as lowly digitalised those at the bottom quartile (see Table 3, Calvino et al., 2018). The remaining industries (i.e., those at the second and third quartile) have an intermediate degree of digitalisation and are regarded as reference sectors in the regression analysis.[2]

3. Empirical model

We conduct a regression analysis, pooling together three different nation-representative samples of European companies for which information on continuing vocational training is available from the CVTS waves for the years 2005, 2010 and 2015. In our baseline specification (eq. (1)), we regress the average wage (in logs) against a variable, T, capturing the training policy implemented by the firm. In eq. (1), i denotes the firm, t years. T is mainly defined as a binary indicator reflecting whether the company has arranged vocational training activities for its employees. In robustness checks, we also consider three continuous proxies for the training effort of the company, namely the ratio of training costs to total labour expenses, the share of workers under training out of the total workforce and, finally, the average number of training hours per trainee. As we discuss more extensively below, such continuous measures of training are likely to be affected to a greater extent by reverse causality issues, making the binary

[2] The group of highly digitalised sectors (top quartile) includes: Computer, electronic and optical products (NACE rev. 2 category 26); Machinery and equipment n.e.c. (28); Transport equipment (29) Telecommunications (61); IT and other information services (62-63); Finance and insurance (64-66); Real estate (68); Legal and accounting activities, etc. (69-71); Advertising and market research; other business services (73-75); Administrative and support service activities (77-82). The group of lowly digitalised industries includes: Agriculture, forestry, fishing (01-03); Mining and quarrying (05-09); Food products, beverages and tobacco (10-12); Electricity, gas, steam and air conditioning (35); Water supply; sewerage, waste management (36-39); Construction (41-43); Transportation and storage (49-53); Accommodation and food service activities (55-56).

indicator our preferred measure for identifying the impact of training on wage dispersion. In our exploration of the wage effect of training, we also assess whether the way in which these activities are organised, i.e. whether training is internally managed by the company or is provided by external specialised trainers, such as private companies, education institutions and government agencies, has a differential impact on workers' remuneration. In this regard, we consider a set of dummies identifying companies with internal training only (T^I), companies with external training only (T^E), and companies pursuing both modes of training (T^B).[3]

Eq. (1) defines our baseline regression model. X_i is a vector of company characteristics whose effect may be confused with that of training. d_s, d_c, d_t denote industry-, country- and time-specific fixed effects. d_s should capture wage differences depending on the technology conditions of production. d_c should neutralise the effect associated with country-specific differences in training legislation, as well as in other relevant institutional (country-level) characteristics. d_t should capture the effect on wages of common technology shocks, business cycle, etc.[4] ε is the error term.

$$\ln w_i = \alpha + \beta T_i + \gamma X_i + d_s + d_c + d_t + \varepsilon_i \qquad (1)$$

In our main estimation we assume that T is exogenous with respect to the outcome variable and hence β can be regarded as an average treatment effect (ATE): β identifies, in essence, the average impact of training on wages on the total sample of firms, which includes both companies with training and companies without training. In a later step of the analysis, we relax this assumption and address the endogeneity issue of training variable in two respects. First, we identify the wage impact of training seeking to account for reverse causality and run an instrumental variable regression. Second, we account for selectivity issues and adopt a matching procedure to find out the wage impact of training only on the group of companies that adopt this policy (average treatment on the treated, ATT).

Next, we expand our specification (see eq. (2)) to assess whether the effect of training changes with given characteristics of the company (C). In this context, the parameter δ will identify the wage premium granted by companies with given characteristics, with respect to the main effect of training found for all other companies without these characteristics (β). Specifically, we

---

3 $T_i = T_i^I + T_i^E + T_i^B$.

4 The deterministic components of the model are also used to purge out the effect of price differences among industries, countries and years as wages are expressed in euro at current prices.

explore whether returns to training are related to how these activities are organised, i.e. internally, externally, or both:

$$\ln w_i = \alpha + \beta \; T_i + \delta \, T_i \times C_i + \gamma X_i + d_s + d_c + d_t + \varepsilon_i \qquad (2)$$

To mitigate omitted-variable bias, in identifying the wage effect of training, we enrich our regression models with the following control variables (X): (i) the average number of hours worked per employee (in logs); (ii) company size, as captured by a set of binary indicators for small-, medium- and large-sized firms (less than 50 employees, between 50 and 249 employees, and 250 and more employees, respectively); (iii) the share of male workers out of the total workforce; (iv) a dummy for companies having a contract agreement with social partners imposing the implementation of training; (v) a dummy variable for firms undertaking internal apprenticeship; (vi) and (vii) binary indicators identifying companies that access external training provided by education-sector institutions (schools, colleges, universities, etc.) or public training centres.

4. Empirical results

4.1. Summary statistics

Table 1 presents the proportion of companies undertaking training (without distinguishing its purposes) and those with a training programme focused on IT skill-intensive job positions (general IT and professional IT skills). Our sample consists of a pool of 112,000 companies, 65% of which undertake general training (73,000 firms).[5] This share increased from 54% in 2005 to 76% in 2015. A greater incidence can be found in highly digitalised industries, where the percentage of firms with training is 78% and denotes a rapid increase between 2005 and 2015. An upward trend can also be found in lowly digitalised industries (from 54 to 79%) in which, however, the proportion of companies with training remains smaller.

[5] We trim the sample excluding from the analysis companies at the extreme tails of wage distribution (below 1 and above 99% percentiles), thus mitigating problems related to censoring in employment data. Our main regression results that will be shown below are robust to the trimming procedure and also to excluding the smallest companies (those with less than 20 employees). All unreported results are available upon request.

Table 1. Proportion of firms with training

	All years		2005	2010	2015
	#	%	%	%	%
	(1)	(2)	(3)	(4)	(5)
Training					
All companies	73,070	65.0	53.7	60.8	75.7
High digital	12,669	77.8	64.8	75.5	84.5
Low digital	16,256	67.2	53.6	63.7	79.2
IT Training					
All companies	23,169	20.6	24.2	23.8	15.4
High digital	5,486	33.7	35.6	41.7	26.1
Low digital	4,280	17.7	22.4	19.6	12.3

Note: The figures consist in the absolute number of firms covered by analysis in all years (Column (1)) and the percentage of those with training (Column (2)). Columns (3)-(5) report the percentage of companies with training out of the total number covered by each wave of the CVTS.

The proportion of companies with IT training is one third of all companies with training programmes (21 vs. 65%). The need to train workers in IT skill-intensive positions seems to be partly explained by the company exposure to digitalisation, as the proportion of firms with IT training rises to 34% in highly digitalised industries. It should be noted, however, that the share of companies with IT training has dropped in all branches of the economy since 2005, revealing that the disruptive effect of digitalisation on workforce skills may have been more pronounced in the first half of the sample period.

Table 2. Proportion of firms with training by country (%)

	# Firms	Training	IT training	Training		IT training	
		All sectors	All sectors	Highly digitalised	Lowly digitalised	Highly digitalised	Lowly digitalised
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Bulgaria	5,648	26.2	7.1	40.4	33.4	14.8	6.2
Czech Rep.	15,326	81.5	19.7	89.6	82.6	34.0	16.8
Germany	8,063	66.3	33.1	76.3	64.6	41.9	30.5
Denmark	3,307	76.0	32.2	81.4	79.9	42.9	30.4
Estonia	4,136	68.1	16.5	74.9	70.1	25.1	15.7
Spain	22,993	71.3	25.0	82.7	70.0	41.9	19.8
Finland	2,930	77.5	25.4	85.9	76.1	37.8	23.0
France	3,861	83.7	13.4	86.9	84.3	21.0	10.5
Italy	35,155	57.4	15.1	72.6	64.5	27.5	13.1
Latvia	3,453	40.5	8.7	54.9	47.0	16.6	7.5
Norway	876	57.2	34.2	63.1	58.6	36.9	34.9
UK	6,651	66.8	36.3	71.3	69.7	44.2	36.7
TOTAL	112,399	65.0	20.6	77.8	67.2	33.7	17.7

Notes: The figures consist in the absolute number of firms covered by analysis per country (Column (1)), the percentage of those with training (Column (2)), with IT training (Column (3)), and those active in highly digitalised sectors (Column (4)) or in lowly digitalised sectors (Column (5)).

Table 2 reports the breakdown of our sample by country, showing for each European economy the percentage share of firms with training/IT training and their distribution across digital sectors. The highest proportion of firms with training can be found in France (84%); companies with IT training are prevalent in the UK, Norway, Germany and Denmark, where one third of all firms undertake programmes focused on advanced digital skills. Focusing on the major countries, it can be observed that, in highly digitalised sectors, the incidence of firms with training is below the European average (78%) in Italy and the United Kingdom; the latter country, however, excels in the highest proportion of firms with IT training which achieves 44% of the national sample.

Table 3. Total training cost relative to labour costs, by firm types and country (%)

	Firms with Training	Firms with IT training	Firms with Training		Firms with IT Training	
			Highly digitalised	Lowly digitalised	Highly digitalised	Lowly digitalised
Bulgaria	2.1	2.8	1.0	0.4	3.3	2.1
Czech Republic	1.2	1.6	1.3	0.9	1.9	1.2
Germany	1.4	1.6	1.4	0.8	2.1	1.5
Denmark	1.4	1.5	1.1	1.1	1.5	1.4
Estonia	1.6	2.0	2.0	0.8	2.8	1.9
Spain	1.5	1.8	1.5	0.9	1.9	1.5
Finland	1.3	1.7	1.4	1.1	2.0	1.9
France	2.0	2.0	1.8	1.7	2.0	2.1
Italy	1.4	1.6	1.2	0.8	1.8	1.4
Latvia	0.6	0.9	0.5	0.2	1.7	0.5
Norway	2.2	2.4	1.7	1.2	2.5	2.7
UK	1.9	1.9	1.6	1.3	2.1	2.1
TOTAL	1.5	1.7	1.4	0.9	2.0	1.5

Notes: The figures consist in the cost of training expressed as a percentage ratio of total labour costs, distinguishing across firms with training, with IT training, and those active in highly and lowly digitalised sectors.

Next, we quantify investment in training, expressing the cost of these activities as percentage ratio to total labour cost (Table 3). Due to data restrictions along all three waves of the CVTS survey, we can accurately quantify only costs for total training, without distinguishing company expenditure by type of training (IT skills vs the rest). As training costs, we consider both direct expenses for training and the implicit cost associated with the working hours lost by employees during training.[6] The relative incidence of training investment is 1.5%, and 1.7% if we restrict to firms with IT training. The latter group of companies invests more in training in almost all countries. If we consider all firms with training, the cost share of this investment looks relatively low in lowly digitalised sectors (0.9% of labour costs); if we consider firms with IT training, the investment share looks relatively high in highly digitalised sectors (2%).

We now provide a descriptive overview of the wage differences associated with training (Table 4). Firms with training pay wages one third higher than companies without training. This pattern is common to all countries. In particular, in all the major European economies (Germany, France, Italy, Spain), the size of the wage differential is substantial and similar (roughly 8,000 euro). As the right-hand side of Table 4 illustrates, firms with IT training pay even more

6 The implicit costs of training, defined as Personal Absence Cost (PAC), is computed as “Paid working time (in hours) spent on all CVT courses” multiplied by “Average labour cost per hour worked”. Figures in Table 3 use sampling weights reflecting the representativeness of surveyed companies on national universes.

than the reference group of companies without any type of training. Taken as a whole, the biggest continental economies in Europe denote the largest wage differentials (in absolute terms) with respect to the national control group.

Table 4. Wage differences for training firms

	Training			IT Training		
	No	Yes	t-statistic difference	No	Yes	t-statistic difference
Bulgaria	2,748	4,110	***	2,748	4,988	***
Czech Rep.	11,319	14,904	***	11,319	17,254	***
Germany	30,013	39,028	***	30,013	41,869	***
Denmark	45,816	52,331	***	45,816	54,275	***
Estonia	9,474	13,800	***	9,474	15,834	***
Spain	23,440	31,682	***	23,440	35,057	***
Finland	39,992	46,418	***	39,992	49,062	***
France	38,777	44,695	***	38,777	48,331	***
Italy	30,442	38,747	***	30,442	41,123	***
Latvia	3,491	5,734	***	3,491	8,174	***
Norway	43,309	48,125	***	43,309	50,180	***
UK	23,117	27,600	***	23,117	28,630	***
TOTAL	23,295	31,169	***	23,295	34,632	***

Notes: The stars denote the level of significance for the t-statistics on the mean difference between groups of firms for each country. ***,**, * significant at 1, 5 and 10%.

To complete our overview of the characteristics of firms with training, we briefly illustrate the differences in the employment levels between groups (Table 5). The table points to statistically large gaps between companies with and without training, with the former being 2-3 times larger than the control group. In Italy and France, this gap is particularly pronounced. Overall, IT training firms are considerably much larger than non-training firms.[7]

[7] Summary statistics on the set of covariates used as control variables in the regression analysis are reported in Table A.1 of the Appendix.

Table 5. Employment gap for training firms

	Training			IT Training		
	No	Yes	t-statistic difference	No	Yes	t-statistic difference
Bulgaria	49.3	127.7	***	49.3	154.3	***
Czech Republic	37.9	136.1	***	37.9	190.6	***
Germany	68.6	178.9	***	68.6	227.4	***
Denmark	44.7	113.4	***	44.5	129.5	***
Estonia	37.8	93.9	***	37.8	113.9	***
Spain	54.0	181.5	***	54.0	245.4	***
Finland	41.6	144.0	***	41.6	201.8	***
France	23.4	182.8	***	23.4	182.2	***
Italy	32.4	112.2	***	32.4	152.1	***
Latvia	59.2	118.2	***	59.2	186.3	***
Norway	49.8	93.1	***	49.8	105.1	***
UK	67.4	125.6	***	67.4	146.7	***
TOTAL	44.7	141.3	***	44.7	188.3	***

Notes: The stars denote the level of significance for the t-statistics on the mean difference between groups of firms for each country. ***,**, * significant at 1, 5 and 10%.

4.2. Baseline regression

Table 6 illustrates the findings of OLS regression for our baseline model.[8] Column (1) reports estimation results for our most conservative specification which includes only training - defined as dummy variable - and the full set of industry-, country- and time-effects. The coefficient of the explanatory variable indicates that, once the effects of all the deterministic components of the model have been accounted for, there is a wage difference of 19% between companies with and without training.

[8] All regressions use standard errors clustered at industry-by-country level.

Table 6. Impact of training on wages: baseline estimation

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	Dummy	0.189***	0.117***					
		(0.008)	(0.007)					
Training costs	Percentage			0.015***				
				(0.002)				
Trainees	Percentage				0.001***			
					(0.000)			
Avg training hours	Log					0.037***		
						(0.004)		
Internal training	Dummy						0.040***	0.041***
							(0.008)	(0.008)
External training	Dummy						0.091***	0.088***
							(0.008)	(0.008)
Int. & ext. training	Dummy						0.186***	0.181***
							(0.010)	(0.010)
Hours per worker	Log		0.543***	0.557***	0.542***	0.542***	0.540***	0.540***
			(0.044)	(0.044)	(0.045)	(0.043)	(0.043)	(0.043)
Medium-sized	Dummy		0.111***	0.132***	0.125***	0.120***	0.097***	0.096***
			(0.010)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)
Large-sized	Dummy		0.182***	0.210***	0.194***	0.194***	0.147***	0.145***
			(0.018)	(0.018)	(0.016)	(0.015)	(0.018)	(0.018)
Males	Percentage		0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Agreement	Dummy		0.043***	0.058***	0.044***	0.048***	0.036***	0.036***
			(0.007)	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)
Apprenticeship	Dummy		0.013*	0.020**	0.002	-0.001	0.008	0.007
			(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)
Government training centre	Dummy							0.017
								(0.014)
Education training centre	Dummy							0.036***
								(0.014)
Observations		112,399	112,399	112,399	71,662	71,509	112,399	112,399
R-squared		0.755	0.797	0.795	0.728	0.731	0.799	0.799

Notes: The dependent variable is the average wage (in logs). OLS estimates. The dependent variable is the log of wage. Standard errors clustered at industry-by-country level. Year-, Industry-, and Country-fixed effects are included in all regressions. ***,**, * significant at 1, 5 and 10%.

In Column (2), we introduce the set of control variables reflecting the structural characteristics of the company. As expected, wages are significantly higher in medium and large-sized companies, which have more funds to allocate to investments compared to smaller firms; wages are also higher in companies having a larger share of male workers and in those with a larger number of hours worked per employee. Column (2) allows us to check whether the coefficient of our key explanatory variable captures the effect of other idiosyncratic characteristics correlated with training activities, some of which may ultimately depend on the organisational capabilities of the company. For instance, firms may pursue wider training policies like those for initial apprenticeship, with the risk that the coefficient of the vocational training is upward biased. Companies may also undertake training programmes as these are imposed by the contract agreement between the organisations of employers and employees, implying that the coefficient of training does not specifically reflect the company decision to pursue this policy. Overall, accounting for all these factors does marginally influence our main estimates (0.117), despite control variables being significant and with the expected sign.

Next, we assess the sensitivity of these results to alternative training measures. As described above, we use the ratio of training expenses to total labour costs (Column (3)), the share of trainees out of the total workforce (Column (4)) and average number of training hours (Column (5)). These variables turn out to be largely significant and with a positive coefficient. Parameters in Columns (3)-(4) are semi-elasticities, therefore suggesting that a 1% increase in training costs or the share of workers under training is associated with a 1.5% and 0.1% wage increase, respectively. The positive coefficient associated with the share of trainees suggests that wage spillovers might occur when a larger number of workers is involved in training as these activities could enhance complementarities among trainees and workers without training, making them more productive and leading to higher wages. The coefficient of training in Column (5) is an elasticity indicating that a one percent increase in the number of training hours translates into a 0.037% wage premium. Since the average number of training hours (per trainee) is 24 per year, with 2.5 additional hours of training (roughly a 10% increase) the average salary would be expected to rise by 0.5%, i.e.150 euro as average for all workers.

Finally, we explore whether the wage premium associated with training varies in relation to how this activity is organised, i.e. whether it is internally managed by the company, it is outsourced to external bodies or the firm adopts both forms of training. We find that a 19% wage premium is associated with companies adopting a hybrid training policy (both internal and external training), whilst the wage premium associated with external or internal training only amounts to 9 and 4% respectively (Column (6)). The larger effect found for companies having training programmes organised both internally and externally may reveal that trainees become more productive, because they develop a larger or a more effective set of skills when they learn competences developed within the company combined with more general skills acquired through specialised centres or companies.

In Column (7), we refine the latter estimates by exploring whether the wage premium associated with the implementation of external training depends on the nature of the training centre. We therefore include two dummy variables capturing whether the training provider is a public training institution (i.e., financed or led by the government) or an education institution (schools, colleges, universities and other higher education institutions). Our estimation shows that only companies with training provided by an education institution pay wages higher than the reference group. Overall, this check does not change the main pattern of our results.

4.3. Extended regression: IT sectoral pattern and IT training

In this part of the work, we investigate in what respect the training policy pursued by the European companies is affected by digitalisation and whether the company response to such transformations, in terms of IT-related training content, helps explain the wage differentials existing across firms. This analysis is developed in Table 7 where, in Column (1), we report the main results of the previous set of regressions as reference (i.e., Column (2), Table 6). All estimations in Table 7 include the same set of controls used above but are not shown here for the sake of brevity. In Column (2), we include a binary indicator identifying firms undertaking training targeted to develop IT-related competences (general IT skills and professional IT skills). To discern the wage effect of this variable from the general tendency of a company to pay higher (or lower) wages in relation to its exposure to digitalisation, we include two dummies for those companies active in industries identified as highly digitalised or lowly digitalised (1st and 4th quartile of the ICT ranking developed by ICT). Firms active in these sectors are found to pay statistically higher wages than in the rest of the economy (i.e., medium-high and medium-low digitalised sectors, 2nd and 3rd quartile). In particular, a 53% higher wage is found for highly digitalised firms with respect to the reference group, and a 14% higher wage for firms in lowly digitalised sectors. These results conform to the evidence provided by Michaels et al. (2013) using industry-by-country data from the early 1970s, concerning the wage polarisation caused by the IT revolution. No less relevantly, we find that companies investing in IT skill-related programmes pay statistically higher wages. An 8.7% wage premium is associated with being employed by these firms, which adds to a 9.2% higher wage paid by firms undertaking other forms of training (see Column 2, Table 7).[9]

Next, we inspect whether the effect of training changes across sectors in relation to the digitalisation of their production. Accordingly, we run our regression model separately for companies active in highly digitalised and lowly digitalised sectors (Columns (3) and (4)). The coefficient size of general and IT training variables does not appear very different between these two types of industries compared to what we found for the pool of firms in Column (2).[10]

In Column (5), we explore whether the impact of IT training overlaps to how training is organised, and accordingly include a full set of dummies identifying firms with internal training provision, external training provision and both training modes. In this regression, the coefficient of IT training reduces only slightly and continues to be highly significant. As a last step (Column (6)), we interact IT training with the three variables capturing the organisational modes of training. We find that IT training companies with stand-alone external training provision, or those combining both internal and external training, pay statistically higher wages than companies without training (general reference group) and compared to companies with general training.

9 To exclude the possibility of our results being driven by companies in industries which have been leading both in wage increases and training expansion, we also have our model with industry-by-year dummies. These results are lined up to the main reference described in the main text.

10 This finding is confirmed by a formal Chow test conducted on one regression where training and IT training variables are allowed to differ between highly and lowly digitalised industries (unreported). The P-value of the Wald test on parameter equality between groups is 0.2 for IT training, and 0.5 for general training. In order to explore whether the impact of IT training changes with the sector exposure to digital transformation we have also interacted this variable with the dummies for highly digitalised and lowly digitalised industries. These interaction terms are always insignificant confirming the view that the wage premium associated with IT training companies is pervasive across sectors.

Table 7. Impact of training on wages: IT sectoral pattern and IT training

		(1) All sectors	(2) All sectors	(3) Highly digitalised sectors	(4) Lowly digitalised sectors	(5) All sectors	(6) All sectors
Training	Dummy	0.117***	0.092***	0.107***	0.082***		
		(0.007)	(0.007)	(0.017)	(0.012)		
IT training	Dummy		0.087***	0.060***	0.082***	0.072***	
			(0.009)	(0.012)	(0.018)	(0.008)	
Internal training	Dummy					0.023***	0.024***
						(0.008)	(0.008)
External training	Dummy					0.072***	0.068***
						(0.008)	(0.008)
Internal & external training	Dummy					0.155***	0.161***
						(0.008)	(0.009)
Internal training × IT training	Dummy						0.058***
							(0.008)
External training × IT training	Dummy						0.070***
							(0.012)
Int. & external training. × IT training	Dummy						0.090***
							(0.010)
High digitalised sector	Dummy		0.530***			0.512***	0.512***
			(0.058)			(0.058)	(0.058)
Low digitalised sector	Dummy		0.136***			0.134***	0.134***
			(0.028)			(0.028)	(0.028)
CONTROLS		YES	YES	YES	YES	YES	YES
Observations		112,399	112,399	16,276	24,203	112,399	112,399
R-squared		0.797	0.798	0.761	0.817	0.800	0.800

Notes: The dependent variable is the average wage (in logs). OLS estimates. Standard errors clustered at industry-by-country level. Year-, Industry, and Country-fixed effects are included in all regressions. All estimates include the control variables used in Table 6, namely hours per worker, size dummies, share of male workers, and the binary indicators for the companies with contract agreement for training and those with workers under apprenticeship. ***,**,*significant at 1, 5 and 10%.

4.4. Endogeneity issues

One concern with the OLS estimates shown above is that the sample of firms with training (treated) may not be random, but that both training and wage performance may depend on some unobservable characteristics or, worse, that the direction of causality runs in the opposite direction to what is assumed here, namely that companies with higher wages may be endowed with more productive workers that employers seek to keep through training and other activities promoting firm-specific human capital. All this would raise concerns about the consistency of our estimates due to selectivity and simultaneity issues. To address these concerns, we run two further types of regression. First, in the same spirit as Brunello et al. (2012), we run an IV-2SLS regression in which we instrument the training variables with a binary indicator identifying companies benefiting from public supports to implement training.[11] Specifically, the CVTS includes a question about whether the company accessed fiscal incentives, or direct funding disbursed by various public bodies (European Commission, the national government, regional authorities, etc.), to offer training to employees. We use this information to build two dummies to identify companies benefiting from tax discounts and those being awarded direct grants for training. The assumption behind this identification strategy is that the wage impact of these public incentives is entirely channelled by training. The plausibility of this assumption resides in the fact that (almost) all companies accessing public incentives implement training, thus ensuring that the condition of exclusion restrictions is satisfied. In our sample, around 22,000 companies (out of 73,000) benefit from some public support for training, but only 10% of them exploit both measures. This ensures that our external instruments capture two independent sources of variation, satisfying the relevance condition. Since the CVTS does not offer information on public incentives for IT training (but, as said, concerns training in general), we provide IT training with an interaction variable between the binary indicator identifying firms with fiscal incentives or direct funding for training, and the industry share of companies implementing IT training (excluding the firm itself). The instrumental variable should in this respect reflect the fact that companies operating in industries with a greater propensity to implement IT training are more likely to exploit public support to undertake training programmes targeted at IT workers.

Second, we infer whether the estimated wage impact of training is affected by selectivity issues, by estimating our model by means of the propensity score matching procedure developed by Imbens (2015). This method identifies the impact of the treatment variable (training) on the outcome variable (wages) based on the propensity scores yielded by a probit regression between the training dummy and the covariates used in our least-square regression. The difference in the outcome variable between treated and non-treated firms is computed on selected sub-groups of firms (blocks), identified as similar based on the normalised differences

[11] Similarly, Bloom et al. (2013) use variation in the user cost of R&D induced by regional fiscal incentives to research to quantify the impact of innovation on productivity growth of the US companies.

in the observed set of firm characteristics.[12] The average treatment is computed separately at single-block level and then averaged across blocks to recover the treatment for the full sample.

Table 8 compares the results of these two procedures with our main OLS estimates (Column (2) in Tables 6 and 7, respectively). Instrumental variable regressions in Columns (2) and (5) satisfy both the relevance and the orthogonality conditions, providing estimates for the training variable higher than those yielded by OLS regression. This finding may reflect the fact that companies accessing public support more effectively engage in general training and, eventually, are able to pay their workers higher wages. Another possible explanation is that the explanatory variable is affected by measurement errors and hence its wage effect is estimated with a strong attenuation bias with OLS. However, these problems do not seem to affect the estimation of the wage impact of IT training whose effect is estimated quite accurately with both procedures (OLS and IV).

An unaddressed issue of our instrument variables (IV) strategy is that there could be endogenous selection of the firms accessing public incentives and the risk is that our instrument may exacerbate, rather than attenuate, estimation bias in the wage impact of training. On this basis, we estimate a system of simultaneous equations including a 'zero-stage' in which we seek to explain the company access to public incentives through a variable capturing the financial difficulties of offering training. This information is available either for companies offering training or those without training and is modelled as a binary indicator. The ratio of this strategy follows, among others, Cin et al. (2017) who use R&D subsidy to infer the impact of the expanded research effort on company labour productivity. Table A.2 in the Appendix reports the results of the system estimation, in which the equations are modelled either as linear regressions or as probabilistic regressions in those cases in which the dependent variable is a binary indicator. The table shows that, although there is selectivity in those companies accessing public incentives to training, the wage impact estimated for training, when predicted by the variables of access to public incentives, is highly significant. Indeed, our proxy for the cost obstacle to training explains a significant portion of variation in the access of European companies to tax incentives and direct funding. Due to the error correlation found between the equation for training and the equation for public support, our system estimates yield lower

[12] The unavailability of information on the workforce's characteristics makes it impossible to fully exclude the presence of relevant omitted factors in our estimates. Ci et al. (2015) adopt an analysis procedure similar to that used in this study, based on the comparison between OLS and propensity score matching, to study the impact of the mid-carrier (on-the-job) training in Canada. Frölich (2007) uses the PSM procedure to study the distribution of wage effect of training using a similar set of covariates.

estimates for the coefficient of the training variable in the wage equation, mitigating thus the upward bias associated with single-equation regressions. Interestingly, system results are similar whether we use a linear regression or Probit regression model for explaining the endogenous dummy variables.

Furthermore, one may wonder whether the ability to access public funds for training is related to the managerial capabilities of the firm, and hence whether the wage effect of the instrumented variables does in reality capture management practices or other relevant company characteristics such as profitability, etc. Unfortunately, due to data unavailability, we cannot control for all these issues. However, as long as the firm's capability to access public incentives, induced by a higher management quality, is likely to materialise into more funds received, the robustness of training dummies can be assessed with a regression including, as control, the share of public provisions for training out of total labour costs. Therefore, in unreported regressions, we run our IV estimations and find that the coefficient of training dummies remains largely unchanged. This corroborates the view that public incentives operate through the extensive margin, i.e. they induce companies to implement training, and that the impact of our key variable is (probably) unrelated to the quality of management.

In implementing the PSM procedure, we distinguish the impact of the treatment variable on the overall sample (average treatment effect, ATE) and that exerted on the group of firms with training only (average treatment on treated, ATT). Both matching estimates (ATE) are comparable to the conditional mean estimates provided by OLS and IV regressions. It is well known, however, that the ATT offers a more accurate assessment of the wage effect of training as it focuses on the group of treated companies only. Overall, estimates in Table 8 confirm the robustness of our results, with the impact of training being highly significant and, for size, quite in line with OLS and IV estimates. Summing up, using our more conservative estimates, a wage premium of 12% would be associated with companies undertaking training, and an additional 8% with those arranging training for IT skill-intensive workers.

Table 8. Wage impact of training: IV-2SLS and PSM estimates

	(1) OLS	(2) IV-2nd stage	(3) PSM-ATE	(4) PSM-ATT	(5) OLS	(6) IV-2nd stage	(7) PSM-ATE	(8) PSM-ATT
Training	0.117***	0.184***	0.125***	0.129***				
	(0.007)	(0.031)	(0.004)	(0.005)				
IT training					0.087***	0.084***	0.077***	0.080***
					(0.009)	(0.021)	(0.003)	(0.003)
		1st stage				**1st stage**		
Fiscal incentives		0.299***				2.307***		
		(0.017)				(0.332)		
Direct funding		0.185***				2.121***		
		(0.013)				(0.181)		
F-test		338.7				94.7		
Hansen J p-value		[0.365]				[0.675]		
N. treated			73,070	73,070			23,166	23,166
N. controls			39,319	39,329			49,866	89,230
N. blocks			51	40			21	36
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES

Notes: The dependent variable is the average wage (in logs). The number of obs. is 112,339. OLS estimates are reported in Columns (1) and (5). IV-2SLS estimates are reported in Columns (2) and (6). Blocking propensity score matching estimates are reported in Columns (2)-(3) and (5)-(6). Standard errors are clustered at industry-by-country level. Year-, Industry-, and Country-fixed effects are included in all regressions. Estimates in Columns (1)-(4) rely on estimates in Column (2), Table 6. Estimates in Columns (5)-(8) rely on estimates in Column (2), Table 7. Controls included: hours per worker, size dummies, share of male workers, and the binary indicators for the companies with contract agreement for training and those with workers under apprenticeship. Estimates in Column (2) of this table use as instruments two dummy variables identifying companies benefiting from fiscal incentives or public funding for training. Estimates in Column (6) of this table use as instruments two dummy variables identifying companies benefiting from fiscal incentives or public funding for training, multiplied by the industry share of companies implementing IT training. The matching procedure considers only companies with common support. All estimates of the present table include the full set of control variables used in Tables 6 and 7. ***,**,*significant at 1, 5 and 10%.

5. Conclusions

This paper has investigated the company-level wage effect of training in selected European countries, by taking into account the different exposure to digitalisation and the digital content of training activities. Digital transformation forces firms to adopt measures for upgrading the skill structure of the workforce. This need is particularly strong for jobs based on digital skills (IT upskilling). Based on a large company-level dataset (112,000 companies), obtained by merging three different waves of the EU Continuing Vocational Training Survey, we have illustrated that there are significant differences in the wage performance related to training across European companies. According to our estimates, a wage premium of 9% is associated with companies undertaking training, defined in general terms, and an additional 8% is paid by firms providing IT training. Since information technologies are highly pervasive, are employed in a wide range of sectors and act as general-purpose technologies, the wage effect of the digital transformation channelled by the IT upskilling is broad-based and not strictly confined to those productions more exposed to digitalisation.

In the analysis we have addressed various issues potentially affecting our results. First, to mitigate omitted variable problems, we have accounted for a set of structural characteristics (size, contract training, etc.) and how training is organised by the firm (i.e., internally vs externally; education/public training centres vs private companies). Second, we have addressed the issues of simultaneity and selectivity, to account for the fact that companies with higher wages, and those with certain characteristics correlated with the outcome and treatment variables, are more likely to undertake training. Both robustness checks confirm the validity of our main results, namely that training is usually associated with a statistically significant wage premium, and that when targeted at IT skill-intensive jobs, it is particularly rewarding.

Our analysis offers useful insights for academics and policymakers interested in understanding the consequences of digitalisation and how to tackle its possible adverse effects. A wide literature has looked at the change in labour demand, wage levels and dispersion, as well as employment prospects associated with the diffusion and adoption of new digital technologies, but little is known about which company-level policy is more effective to tackle this process. This study helps to fill this important gap in our understanding. On aggregate, however, there is also the risk that wage differences across firms are likely to widen if falling-behind companies are not able to systematically organise policies for workplace learning and training, especially for some key job positions. Our findings complement recent evidence on the widening productivity gap between frontier and laggard companies in Europe and other advanced countries (Andrews et al., 2019), and on the fact that acting in the new technological fields may help reduce the distance from the most productive companies (Pompei & Venturini, 2022).

References

Acemoglu, D. (1998). Training and innovation in an imperfect labour Market. Review of Economic Studies 64(3), 445-464.

Acemoglu, D. (1998) Why do new technologies complement skills? Directed technical change and wage inequality. Quarterly Journal of Economics 113(4), 1055–1089.

Acemoglu, D. (2002). Technical change, inequality, and the labor market. Journal of Economic Literature, 40(1), 7–72.

Acemoglu, D. & Pischke, J. (1998a). Why do firms train? Theory and evidence. Quarterly Journal of Economics 113(1), 79-119.

Acemoglu, D. & Pischke, J. (1998b). The structure of wages and investment in general training. Journal of Political Economy 539–572.

Acemoglu, D., Pischke, J. (1999). Beyond Becker: Training in imperfect labor markets. Economic Jourrnal 109, F112-F142.

Acemoglu, D. & Restrepo, P., (2018) The race between man and machine: Implications of technology for growth, factor shares and employment. American Economic Review 108(6), 1488–1542.

Acemoglu, D., Autor., D., Hazel., D & Restrepo, P. (2022) Artificial Intelligence and jobs - Evidence from online vacancies. Journal of Labor Economics, 40(S1), pp. S293-S340.

Andrews, D., Criscuolo, C. & Gal., P.N., (2019). The best versus the rest: Divergence across Firms during the Global Productivity Slowdown. CEP Discussion Paper No 164. https://cep.lse.ac.uk/pubs/download/dp1645.pdf.

Autor, D.H., Katz, L.F. & Krueger, A.B. (1998) Computing inequality: Have computers changed the labor market? Quarterly Journal of Economics 113(4) 1169–1213.

Autor, D.H, Levy, F. & Murnane, R. J. (2003). The skill content of recent technological Change: An empirical exploration. Quarterly Journal of Economics 118(4), 1279–1333.

Bartel, A. & Sicherman, N. (1998). Technological change and the skill acquisition of young Workers. Journal of Labour Economics 16(4), 718-755.

Becker, G. (1964). Human capital. University of Chicago Press, Chicago.

Bloom, N, Schankerman, M., Van Reenen, J. (2013). Identifying Technology Spillovers and Product Market Rivalry. Econometrica, 81(4), 1347-1393.

Brunello, G., Comi, S.L., Sonedda, D. (2012). Training subsidies and the wage returns to continuing vocational training Evidence from Italian regions. Labour Economics 19, 361–372.

Brunello, G. & Wruuck, P. (2020). Employer-provided training in Europe: Determinants and obstacles. IZA Discussion Paper No. 12981.

Calvino, F., Criscuolo, C., Marcolini, L. & Squicciarini, M. (2018). A taxonomy of digital intensive sectors. OECD Science, Technology and Industry Working Papers, No. 2018/14, OECD Publishing, Paris, https://doi.org/10.1787/f404736a-en

Cedefop (2016). The great divide: Digitalisation and digital skill gaps in the EU workforce. #ESJsurvey Insights, No 9, Thessaloniki: Greece.

Cedefop (2019). Continuing vocational training in EU enterprises: developments and challenges ahead. Luxembourg: Publications Office. Cedefop research paper; No 73.

Cin, B.C., Kim, Y.J. & Vonortas, N.S. (2017). The impact of public R&D subsidy on small firm productivity: evidence from Korean SMEs. Small Bus Econ 48, 345–360).

Ci, W., Galdo, J. C., Voia, M. & Worswick, C., (2015). Wage Returns to mid-career investments in job training through employer-supported course enrollment: Evidence for Canada. IZA Discussion Papers 9007, Institute of Labor Economics (IZA).

EIB (2022). Digitalisation in Europe 2021-2022: Evidence from the EIB Investment Survey. European Investment Bank, Economics Department, https://www.eib.org/attachments/publications/digitalisation_in_europe_2021_2022_en.pdf.

Frölich, M. (2007). Propensity score matching without conditional independence assumption—with an application to the gender wage gap in the United Kingdom. The Econometrics Journal 10(2), 359–407.

Gal, P., Nicoletti, G., Renault, T., Sorbe, S. & Timiliotis, C. (2019). Digitalisation and productivity: In search of the holy grail – Firm-level empirical evidence from EU countries. OECD Economics Department Working Papers, No. 1533, OECD Publishing, Paris.

Goos, M., Rademakers, E., Röttger, R., (2021). Routine-biased technical change: Individual-level evidence from a plant closure. Research Policy 50(7) 104002.

Krueger, A. B. (1993). How computers have changed the wage structure: Evidence from microdata, 1984-1989. Quarterly Journal of Economics 108(1), 33–60.

Imbens, G.W. (2015). Matching methods in practice: Three examples. Journal of Human Resources, 50(2).

McGuinness, S., Pouliakas, K. & Redmond, P. (2021). Skills-displacing technological change and its impact on jobs: challenging technological alarmism?. Economics of Innovation and New Technology, 1-23.

Michaels, G., Natraj, A. & Van Reenen, J., (2014). Has ICT polarised skill demand? Evidence from eleven countries over 25 Years. Review of Economics and Statistics 96(1) 60–77.

O'Mahony, M., Robinson C. & Vecchi, M. (2008). The impact of ICT on the demand for skilled labour: A cross-country comparison. Labour Economics 15(6), 1435-1450.

O'Mahony, M. (2012). Human capital formation and continuous training: Evidence for EU countries. Review of Income and Wealth 58(3), 531-549.

Pompei, F. & Venturini, F. (2022). Firm level productivity and profitability effects of managerial and organisational capabilities and innovations. Untangled Research Papers No. 02/2022. https://projectuntangled.eu/untangled-research-papers/

Squicciarini, M., Marcolin, L. & Horvat, J., (2015). Estimating cross-country investment in training: An experimental methodology using PIAAC data. OECD Science, Technology and Industry Working Papers 2015/09.

Webb, M. (2020). The Impact of Artificial Intelligence on the labor market. Stanford University, mimeo.

APPENDIX A.1. Additional summary statistics

Table A.1 Summary statistics

		Mean	SD	Min	Max
Wage	Continuous	29,977.9	14,539.4	1,116.0	90,421.3
Training	Dummy	0.60	0.49	0.00	1.00
IT training	Dummy	0.20	0.40	0.00	1.00
Training costs	Percentage	0.96	1.90	0.00	98.54
Trainees	Percentage	47.21	32.20	0.00	100.00
Training hours per trainee	Continuous	24.00	48.71	0.00	2,000.00
Internal training	Dummy	0.07	0.26	0.00	1.00
External training	Dummy	0.26	0.44	0.00	1.00
Int. & external training	Dummy	0.26	0.44	0.00	1.00
Highly digitalised industry	Dummy	0.10	0.29	0.00	1.00
Lowly digitalised industry	Dummy	0.20	0.40	0.00	1.00
Hours per worker	Continuous	1,665.1	3,780.0	0.75	833,074
Small firm	Dummy	0.82	0.39	0.00	1.00
Medium firm	Dummy	0.16	0.36	0.00	1.00
Large firm	Dummy	0.03	0.16	0.00	1.00
Male workers	Percentage	63.63	27.17	0.00	100.00
Agreement	Dummy	0.19	0.39	0.00	1.00
Apprenticeship	Dummy	0.35	0.48	0.00	1.0
Government training	Dummy	0.05	0.22	0.00	1.00
Education training	Dummy	0.04	0.21	0.00	1.00
Fiscal incentives	Dummy	0.07	0.26	0.00	1.00
Direct funding	Dummy	0.08	0.28	0.00	1.00

Notes: Mean values are obtained using sampling weights.

Table A.2 IV- System estimation

	(1) 2 stage	(2) 1 stage	(3) 2 stage	(4) 1 stage	(5) 0 stage	(6) 2 stage	(7) 1 stage	(8) 0 stage
	Linear Wage	Linear Training	Linear Wage	Linear Training	Linear Tax incentives	Linear Wage	Linear Training	Linear Direct funding
Training	0.184***		0.119**			0.105**		
	(0.031)		(0.050)			(0.052)		
Tax incentives		0.298***		1.156***			0.291***	
		(0.017)		(0.147)			(0.017)	
Direct funding		0.186***		0.184***			2.589***	
		(0.013)		(0.014)			(0.498)	
Cost obstacles					0.048***			0.018***
					(0.008)			(0.004)
rho_23				-0.465***			-1.284***	
				(0.092)			(0.177)	
	Linear Wage	Probit Training	Linear Wage	Probit Training	Probit Tax incentives	Linear Wage	Probit Training	Probit Direct funding
Training	0.111***		0.089***			0.081***		
	(0.017)		(0.023)			(0.025)		
Tax incentives		1.882***		2.173***			1.871***	
		(0.107)		(0.119)			(0.109)	
Direct funding		1.163***		1.160***			0.810***	
		(0.064)		(0.064)			(0.158)	
Cost obstacles					0.319***			0.131***
					(0.021)			(0.020)
rho_23				-0.191***			0.177***	
				(0.043)			(0.064)	
CONTROLS				YES			YES	

Notes: The dependent variable is the average wage (in logs) in cols (1), (3) and (6), the training dummy in cols (2), (4) and (7); tax incentives and direct funding dummies in cols. (5) and (8). The number of obs. is 112,339. Standard errors are clustered at industry-by-country level. Year-, Industry-, and Country-fixed effects are included in all regressions. All estimates of the present table include the full set of control variables used in Tables 6 and 7. Rho_23 is the residual correlation between the equations for training and public incentives. The number of obs. is 112,339. ***,**,*significant at 1, 5 and 10%.

UNTANGLED Partners:

UNTANGLED is a three-year interdisciplinary Horizon 2020 research project that seeks to examine the interconnected trends of globalisation, demographic change and technological transformation, and their effects on labour markets in the European Union and beyond. By engaging a broad range of stakeholders, including companies and civil society organisations, we will develop practical policy proposals to help governments cushion the negative impacts of these trends and ensure their benefits are enjoyed fairly across regions and sectors.

Follow us on social media:

twitter.com/untangledEU

facebook.com/UntangledEU

linkedin.com/company/project-untangled-eu

www.projectuntangled.eu