Digitalization, Skilled labor and the Productivity of Firms[1]

Jóannes Jacobsen, Jan Rose Skaksen and Anders Sørensen, CEBR, Copenhagen Business School

1. Introduction

In the literature on information technology and productivity of firms, two main conclusions are that information technology plays an important role in increasing the productivity of firms, and there is a strong complementarity between skilled labor and information technology (see e.g. Doms et. al, 1997, Bresnahan et. al, 2002 and Bartel et al, 2005). In this paper we use a rich Danish dataset and consider how internal digitalization of firms affects their productivity, where digitalization is defined as the implementation of automated (ICT-based) information sharing within inventory control, distribution and production control. It turns out that digitalization increases the productivity of firms, if the skill level in the firm is sufficiently high. If the skill level is not sufficiently high, digitalization may actually decrease firms' productivity.

One big challenge when analyzing how ICT affects productivity is that there may be reverse causality between productivity and ICT. Further, since the data on internal digitalization comes from a survey, whereas all other data are very detailed register data, the risk of measurement errors is particularly high with respect to the level of digitalization. Therefore, to control for possible endogeneity problems, resulting from reverse causality and measurement errors, an important contribution of our paper is to instrument the digitalization of a firm. It turns out that the internationalization of the industry and region cluster to which the firm belongs is an explanatory variable of digitalization. One possible explanation for this is that exporting firms tend to be more productive than non-exporting firms (see e.g. Bernard and Jensen, 2004), and a higher degree of competition from more productive firms tends to induce firms to digitalize. More specifically, we find that if the export intensity of the other firms in the same industry and region increases by 10 percentage points, the probability that a firm is digitalized increases with approximately 3 percent.

To summarize the other main results of the paper, we find that there is a strong complementarity between skilled labor and digitalization, and digitalization only increases the productivity in firms where the skill level is sufficiently high. For the average firm, the estimated effect of digitalization is positive but insignificant. However, there is a big and statistically significant interaction effect between digitalization and the skill level of the firm. Hence, in firms where the skill level is high, the productivity effect of digitalization is much larger than in firms where the skill level is low. The estimated coefficient implies that among the 25 percent of the firms where the skill level is highest, digitalization increases productivity with approximately 4 percent, when estimating the effect by ordinary least square. When using the instrumental variable approach, the effect seems to be larger.

[1] This paper is part of the AIM project (www.aimprojekt.dk). We are very grateful to Industriens Fond for financial support.

Our paper is closely related to a large literature on ICT and productivity. Within this literature, most studies find that ICT increases productivity, but there is a lot of variation, see e.g. Stiroh (2004) or Draca et. al. (2006). Further, in some studies, it is found that there is a complementarity between organizational change and ICT (see e.g. Bresnahan et. al., 2002). This is related to a big literature on the complementarity between technological change and skilled labor, see e.g. Bond and Van Reenen (2007) for a survey. This literature mainly considers how technological change affects the demand for skilled labor relative to unskilled labor. In our paper, we also consider the complementarity between skills and technological change, but we do it by estimating how technological change – digitalization – affects the production function of the firm.

The rest of the paper is organized as follows. In section 2, we introduce the framework and discuss the hypotheses of the paper. In section 3, we present the data, and, in section 4, we report the main results of the paper. In section 5, we discuss some threats to the identification of our main parameters of interest, and, finally, in section 6, we conclude.

2. Framework and Hypotheses

We consider how the productivity of firms depends on the internal digitalization and the skills of workers. By digitalization we refer to the implementation of automated information sharing regarding received and/or placed orders within one or more of three internal functions: *inventory control, distribution, and production control*. This definition of internal digitalization follows from the EU harmonized question on "automated information sharing within the firm".

There is a big literature documenting that there is a complementarity between different types of skills and the use of information technology (see e.g. Bond and Van Reenen, 2007). Most studies find that information technology requires more skilled labor, but in studies using more detailed classifications of different types of information technology, it is found that the relationship between technology adoption and workforce skills varies across technologies. In Dunne and Troske (2004), it is e.g. found that technologies more closely related to engineering and design tasks are associated with more skilled workforces, but this is not the case for technologies more closely related to production activities.

The complementarity between skilled labor and information technology implies that some firms may even realize a lower productivity, if they introduce information technology without the right skills being present in the firm.

Our point of departure is the following Cobb-Douglas production function:

$$Y_i = e^{a_i} K_i^{\alpha_i} L_i^{\beta_i} e^{s_i \gamma_i} \tag{1}$$

where Y_i is value added, K_i is capital, L_i is (raw) labor, and s_i is human capital measured as the share of employed workers being skilled. We will assume that firms can chose between two different technologies, or production functions: a non-digitalized and a digitalized. The set of parameters of the production function (a_i, α_i, β_i, γ_i) is assumed to take one (base) combination of values in non-digitalized firms (a, α, β, γ), and another combination in digitalized firms ($a+a^d$, $\alpha+\alpha^d$, $\beta+\beta^d$, $\gamma+\gamma^d$). If digitalized and non-digitalized firms basically produce according to the same technology, then (a^d, α^d, β^d, γ^d) = ($0,0,0,0$). If digitalization only affects the total factor productivity, without affecting the marginal productivity of production factors (i.e. there are no complementarities between digitalization and specific production factors), then a^d is non-zero, whereas (α^d, β^d, γ^d) = (*0,0,0*). If digitalization affects total factor productivity as well as the marginal product of human capital (or strictly speaking the semi-elasticity of human capital), then a^d and γ^d are non-zeros, whereas (α^d, β^d) = (*0,0*). Obviously, it may also be the case that digitalization affects the productivity of capital and or labor.

The choice of technology is costly, implying that, in the short run, only a subset of firms has made the long run optimal choice. Therefore, we may identify the parameters by estimating the following production function:

$$y_i = a + \alpha k_i + \beta l_i + \gamma s_i + a^d D_i + \alpha^d k_i D_i + \beta^d l_i D_i + \gamma^d s_i D_i + \epsilon_i \quad (2)$$

where y_i, k_i and l_i are the logs of value added, employment and capital, respectively. D_i is a dummy indicating whether or not a firm is digitalized.

Until now, we have implicitly assumed that digitalization explains the productivity of firms. However, it may also be the case that the productivity of a firm affects the incentive to digitalize, and since the level of digitalization is self-reported, there may be measurement errors. To handle these problems of endogeneity, it is necessary to introduce exogenous instruments to explain digitalization. In this respect, our hypothesis is that the incentive to digitalize increases if the degree of competition increases in the business environment of the firm. There is a big literature showing that exporting firms are more productive than non-exporting firms, see e.g Bernard and Jensen (2004). Hence, in a business environment, with a lot of exporting, competitors tend to be more productive, and this in turn intensifies competition. Therefore, as an instrument explaining the digitalization of a firm, we use the export intensity of the other firms in the same region and industry. By doing this, we implicitly assume that the competitive pressure in the cluster only affects the productivity of the firm through digitalization. If this is not the case, we may still have problems with reverse causality. In section 5, we further analyze the strength of the instrument.

We use a two stage least square procedure. In the first stage, we estimate a probit model where the probability of being digitalized is explained by the export intensity of the firms in the same region and industry as well as the other explanatory variables in the production function (L_i, K_i and s_i). Then, in the second stage, we use the predicted probabilities of digitalization as the dummy variable (D_i) in equation 2.

3. Data

All our data are collected by Statistics Denmark. The data on internal digitalization come from a survey conducted on the use of ICT by Danish firms. In order to select firms to be included in the survey, Statistics Denmark stratified all Danish firms outside the agricultural sector according to size and industry. Strata with larger firms were overrepresented, but firms were sampled at random within each selected strata. In 2008, 4,257 firms answered the questionnaire, and in 2009, 4,303 firms. Of the 8,560 obtained answers, 1,681 are repeated observations. These survey data are merged with register data on e.g. number of employees, educational composition of employees and size of capital stock.

The impact of digitalization probably varies a lot across industries. Therefore, we chose to focus on the manufacturing sector. When we further exclude observations for which there are some missing information (e.g. on the digitalization of one of the three processes), we end up with 2252 observations.

As stated above, internal digitalization refers to the implementation of automated information sharing regarding received and/or placed orders within: inventory control, distribution and production control.[2] The answers to the questions on whether a firm is digitalized are qualitative, and should be interpreted as indicators on the level of digitalization. In the main part of the paper, we use the following definition of a digitalized firm:

1. *A firm is digitalized (i.e. D_i = 1) if it has digitalized both inventory control, distribution and production control.*
2. *A firm is non-digitalized (i.e. D_i = 0) if at least one of the three – inventory control, distribution or production control – are non-digitalized.*

With our data, it is possible to apply other definitions of digitalized firms. One possibility is that firms are classified as digitalized if either inventory control, distribution or production control is digitalized. Another possibility is that the degree of digitalization is measured as the number of digitalized processes. However, since the digitalization of the three processes seems to be strongly interdependent, we use the binary definition of digitalization.

The distribution of firms with respect to the digitalization of the three types of processes is shown in Table 1. It is seen that 942 of the firms (42 percent) are digitalized according to our definition. There are 693 firms (31 percent), which have not digitalized any of the processes. The rest (27 percent) have digitalized one or two processes.

Table 1. Distribution of digitalized functions across firm

	Number of firms
Control of Production	1349
Control of inventory	1367
Distribution Management	1056
All three functions digitalized	942
No functions digitalized	693
Total number of firms	2252

Another important variable in our analysis is the *share of educated labor*. This variable is measured as the share of full time employees having a vocational or further education. Hence, we use a relatively broad measure of education implying that 1 minus this share is the share of unskilled labor in the firm. In table 2 we report descriptive statistics on the variables entering the production function, and it is e.g. seen that the average *share of educated labor* is 63 percent.

[2] In the survey there is also asked to digitalization of accounting, but since this is not linked as closely to the production process as inventory control, distribution and production control, we do not include information on the digitalization of accounting.

Table 2. Summary statistics of production function variable

	Observations	Mean	Standard deviation	Min	Max
Log (value added)	2252	17.5051	1.3868	13.720	23.3011
Digitalization	2252	0.4183	…	0	1
Digitalization x Educated labor	2252	0.0059	0.0865	-0.6281	.3719
Educated labor	2252	0.6281	0.1575	0	1
Log (capital)	2252	9.8564	1.8653	1.0986	16.5674
Log (employment)	2252	4.3742	1.2588	1.0986	9.2858

Further, in Table 3 we report the correlations between the variables of the production function. The correlations between capital, labor and value added are relatively high, because they are all measures of the size of the firm. More interesting, we see that the correlation between digitalization and value added is relatively high. This mainly reflects that bigger firms tend to be more digitalized than smaller firms.

Table 3. Correlation matrix between production function variables

	Log (value added)	Digitalization	Digitalization x Educated labor	Educated labor	Log (capital)	Log (employment)
Log (value added)	1					
Digitalization	0.4281	1				
Digitalization x Educated labor	0.1151	0.0811	1			
Educated labor	0.1217	0.0766	0.5517	1		
Log (capital)	0.8064	0.3667	0.0211	-0.0272	1	
Log (employment)	0.9593	0.4400	0.0676	0.0665	0.7980	1

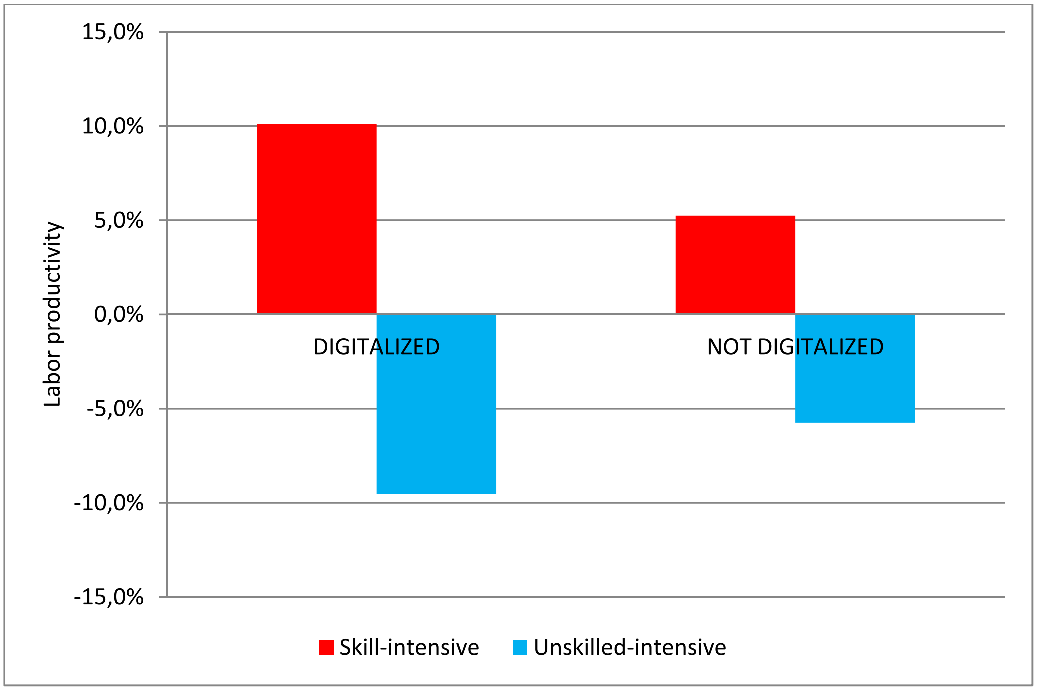
4. Results

This section presents the empirical results of the study. First, we present the descriptive statistics, and afterwards, the empirical result of the regression analysis.

Descriptive Statistics

We divide firms into two types: the first type consists of firms that have digitalized the three functions; the second type consists of firms that have not. Within the two firm types, the average productivity level of skill-intensive firms is compared to the average productivity level of unskilled-intensive firms. In Figure 1, the results are presented as deviations from the average productivity level for the firm type. In the calculations, firms are considered to be skill-intensive, if their skill share exceeds the average share of their firm type, whereas firms are considered to be unskilled-intensive, if their skill share is below the average share.

Figure 1: Labor Productivity for Types of Firms Divided after Digitalization Activities. (Deviation from Mean Within type)



NOTE: Skill-intensive firms are firms where the share of skilled labor is above the average; unskilled-intensive firms are firms where the share of skilled labor is below the average.

Two observations are evident from the figure. First, skill-intensive firms have higher labor productivity than unskilled-intensive firms. This result is independent of whether or not firms are digitalized. The second observation is that skills seem to be more important in firms that are digitalized that in non-digitalized firms. This is seen by the larger variation in labor productivity between skill-intensive and unskilled-intensive digitalized firms compared to that of non-digitalized firms. This suggests that digitalization is indeed a skill-intensive activity.

There are of course also a number of conclusions we cannot draw from Figure 1. We do not control for other firm characteristics. This implies, for example, that we may pick up spurious correlations. Because of this issue, firm characteristics have to be taken into account. We will include background firm characteristics in a regression analysis in the following sub-section. Further, there is the concern of endogeneity of the applied measure of digitalization. To deal with this issue, we apply an instrumental variable approach. The purpose is to mitigate problems of measurement errors as well as causality between digitalization and productivity. The suggested instrument is discussed in greater detail below.

OLS regressions

In this section, we turn to the regression analysis, which, first of all, focuses on estimating equation 2. In Table 4, we report the results when using ordinary least square.

Table 4: Productivity Effects of Digitalization

Dependent variable: log (value added)	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	OLS	OLS, Non-Digit	OLS, Digit
Const	12.2828***	12.3436***	12.3228***	12.2669***	12.3499***
	(0.0853)	(0.0870)	(0.0993)	(0.1056)	(0.1499)
Digitalization	0.0040	0.0022	0.0021		
	(0.0199)	(0.0198)	(0.0199)		
Share of educated labor	0.4517***	0.3663***	0.3713***	0.3324***	0.7133***
	(0.0668)	(0.0682)	(0.0684)	(0.0722)	(0.1544)
Log (employment)	0.9408***	0.9410***	0.9369***	0.9256***	0.9548***
	(0.0137)	(0.0138)	(0.0162)	(0.0161)	(0.0244)
Log (capital)	0.0847***	0.0842***	0.0877***	0.0926***	0.0689***
	(0.0091)	(0.0092)	(0.0107)	(0.0109)	(0.0165)
Digitalization x Educated labor (demeaned)		0.3202***	0.3110***		
		(0.1244)	(0.1257)		
Digitalization x Log (employment) (demeaned)			0.0112		
			(0.0246)		
Digitalization x Log (capital) (demeaned)			-0.0094		
			(0.0167)		
Year fixed effects	yes	yes	yes	yes	yes
Region fixed effects	yes	yes	yes	yes	yes
Industry fixed effects	yes	yes	yes	yes	yes
Observations	2252	2252	2252	1310	942
R-squared	0.9371	0.9374	0.9374	0.9283	0.9293

Note: Standard errors are in parentheses. ***, ** and * indicate significanse at 1%, 5% and 10% level, respectively

The first row reports results when it is assumed that digitalization only affects production through the constant term in the production function, but not through the productivity of the other production factors, i.e. (α^d, β^d, γ^d) = (0,0,0). In this case, the estimated effect of digitalization is positive but insignificant. The size of the coefficient indicates that the productivity of digitalized firms is 0.4 percent higher than the productivity of non-digitalized firms. The other estimated coefficients are highly significant. It is for example seen that an increase in the share of skilled labor of 1 percentage point increases value added with 0.45 percent.

In the second row, digitalization may also affect the productivity of skilled labor, but it is still assumed that digitalization does not affect the productivity of raw labor and capital. The effect of digitalization on the average firm is still insignificant, but there is a large and significant interaction term between digitalization and skilled labor. This implies that there is a significant positive effect of digitalization on the productivity of firms, if the skilled workers constitute a sufficiently high share of the workers in the firm. If the share of skilled workers increases with 10 percentage points, the impact of digitalization increases with 3.2 percentage point. Hence, if the share of skilled labor in a firm is 10 percentage points higher than in the average firm (i.e. 72.81 percent), digitalization increases productivity with approximately 3.6 percent. In 25 percent of the firms, the share of skilled labor is above74.07 percent, and in these firms digitalization increases productivity by approximately 4 percent. In 10 percent of the firms, the share of skilled labor is above 83.33 percent, and in these firms digitalization increases productivity by approximately 7 percent.

In the third row, digitalization may affect the productivity of all inputs in the production function as well as the constant. In this case the only significant effect of digitalization enters through the productivity of skilled labor. Further, it is interesting to note that the estimated effect on the productivity of skilled labor is basically the same as in the second row, which indicates that the estimated effect is robust.

In the first three rows, it is assumed that the time dummy, industry dummies and region dummies have the same impact on the digitalized production function as on the non-digitalized production function. In row 4 and 5, we allow all parameters to differ by estimating two separate production functions. Hence, in row 4 we report the coefficient of a production function for non-digitalized firms and in row 5 in the case of digitalized firms. The parameters in the two production functions are very similar, except that the importance of skills is much bigger in digitalized firms than in non-digitalized firms. If the share of skilled labor increases with 10 percentage points, value added in non-digitalized firms increases with around 3 percent, whereas it increases with 7 percent in digitalized firms. Again, this illustrates the strong complementarity between digitalization and skills.

2SLS regressions

One major problem when interpreting the results in Table 4 is that the variables entering the production function are endogenous, and there may be reverse causality. It is implicitly assumed that digitalization causes an effect on productivity, but it could also be the case that the productivity of a firm has a causal effect on whether a firm digitalizes. One possibility is that firms having unusual low productivity are more concerned about introducing measures to improve their productivity. If this is the case, the correlation between productivity and digitalization tend to become negative. Another possibility is that productive firms have more resources to finance new technology, which tend to give rise to a positive correlation between digitalization and productivity.

An additional issue that can make the variable for digitalization endogenous is measurement problems. It is well known that measurement problems may be substantial in survey data; the data type that the applied digitalization variable stems from. In this particular case, firms may answer that they are not digitalized when they are, and the other way around. In other words, we may wrongly treat digitalized firms as non-digitalized firms, and non-digitalized firms as digitalized.

Instrument

In order to control for endogeneity problems, resulting from measurement errors or reverse causality between productivity and digitalization, we need an instrument for digitalization. We will assume that the incentive of the firm to introduce digitalization depends on the extent to which the firm participates in an internationalized cluster of firms. One reason for this may be that the competitive pressure in a more international environment increases the value of digitalization. Another may be that internationalized firms tend to be more productive because they use more up to date technology, and the use of technology in one firm may inspire other firms in the cluster to introduce the same technology.

More specifically, we assume that the export intensity of the other firms in the cluster to which the firm belongs affects the incentive to digitalize. The cluster is defined as firms in the same region and industry. With respect to regions, we use the same five as we have controlled for above. With respect to industries, we use a narrow definition with 184 industries. Even though the measure is developed using a narrow definition of industries, it is still based on a large number of firms for each region x industry combination. The reason is that the measure is determined for the full population of Danish manufacturing firms. Moreover, the measure is calculated for the period 1999-2008. Consequently, the calculation of the export intensities for the clusters (region x industry) is based on more than 40.000 unique firms over the 12 years, resulting in more than 200.000 observations.

We have developed the region x industry export intensity excluding exports of the single firm in the calculation of the export intensity of the cluster. In Table 5, we apply a probit model to estimate the probability that a firm is digitalized. Besides the explanatory variables in the production function, we include industry and region dummies to control for heterogeneity across these dimensions. We include the export intensity of the cluster, i.e., the total export of the cluster relative to total production value. It is seen that this export intensity has a large and significant impact on the probability that the firm is digitalized.

Table 5: Digitalization and internationalization

Dependent variable: Prob (Digitalization)	
	Probit
Industry export intensity	0.9458***
	(0.2112)
Share of educated labor	0.4970*
	(0.2774)
Log (employment)	0.4954***
	(0.0450)
Log (capital)	0.0427
	(0.0290)
Year fixed effects	yes
Region fixed effects	yes
Industry fixed effects	yes
Chi2-test for instrument	20.06
Observations	2223
R-squared	0.2149

Note: Standard errors are in parentheses. ***, ** and * indicate significanse at 1%, 5% and 10% level, respectively

The industry export intensity is highly significant, and it can be shown that the size of the coefficient implies that a 10 percentage point increase in the export intensity of the production in the region and industry cluster implies that the probability that the firm is digitalized becomes 3 percent higher.

Productivity effects from digitalization

In order to obtain consistent estimates for the parameters to digitalization and the interaction term between the skill-intensity and digitalization, we employ a two-step IV procedure suggested by Wooldridge (2002). The first step involves estimating a probit model for the digitalization choice, as presented in Table 5 above. Having estimated the probit model, we subsequently obtain the fitted values from the regression, $\hat{G}_i$. The second step of our two-step IV approach involves estimating equation (2) using 2SLS with $\hat{G}_i$and $\hat{G}_i(\text{education share demeaned})_i$ as instruments. As we control for all the determinants of digitalization, except for the instruments, this provides us with IV estimates for the productivity level. It should be stressed that we instrument for the digitalization choice, not the input of skill-intensity.[3] The IV results are reported in Table 6.

Table 6: Productivity Effects of Digitalization (2SLS)

Dependent variable: log (value added)		
	IV	IV
Const	12.3286***	12.5042***
	(0.1272)	(0.1353)
Digitalization	0.0693	0.1222
	(0.1420)	(0.1694)
Share of educated labor	0.4465***	0.2252***
	(0.0681)	(0.0878)
Log (employment)	0.9308***	0.9224***
	(0.0256)	(0.0280)
Log (capital)	0.0835***	0.0818***
	(0.0092)	(0.0093)
Digitalization x Educated labor (demeaned)		0.7881***
		(0.2663)
Instruments	Cluster export intensity	Cluster export intensity, Cluster export intensity x Educated labor demeaned
Additional variables		
Year fixed effects	yes	yes
Region fixed effects	yes	yes
Industry fixed effects	yes	yes
Observations	2.223	2.223
Centered R-squared	0.9364	0.9351

Note: Standard errors are in parentheses. ***, ** and * indicate significanse at 1%, 5% and 10% level, respectively

The results in Table 6 are qualitatively similar to the OLS-results reported in Table 4. The main difference is that the interaction term between digitalization and skilled labor is larger. This implies that, in the 25 percent of the firms where the skill level is highest, digitalization increases productivity with approximately 20 percent, when taking into account the big but insignificant effect of digitalization in the average firm.

5. Threats to identification

There are a number of channels that constitute threats to identification. Above, we implicitly assumed that the industry export intensity only affects one productivity enhancing process, namely, adaption of new technology as indicated by digitalization. It may, however, be the case that the industry export intensity also affects other productivity enhancing processes.

[3] The skill-intensity may not be of too much concern in relation to endogeneity problems. In supplementary regressions we have instrumented the skill-intensity of the firms with their skill intensity 8 years back, and the results remain basically identical. Therefore, endogeneity problems with respect to the skill-intensity may not be severe.

Firms in export intensive industries may well be forced to increase productivity to stand up to international competition in export markets. One mechanism to increase productivity is – as studied in the previous sections – to adopt new technology, e.g., digitalization. However, this may not be the only productivity enhancing process.

Another possible mechanism is *reduction of economic slack*: It may be the case that firms in industries with low export intensity to a lesser extent optimize their production process than firms in export intensive industries. A counter argument is that the most important determinant of economic slack in production is the export intensity of the firm itself; a measure that more precisely reflects the international competitive environment of the firm. In this respect we would expect that the export intensity of the firm should enter in explaining firm productivity. Another reason for including the firm export intensity is the concern that the export intensity of the cluster is highly correlated with, and, thereby, just picking up the variation in the export intensity of firms. It should, however, be stressed that when including export measures in the regressions, the equations are no longer clear-cut production functions.

A third productivity enhancing process may be knowledge production: Firms may increase productivity through knowledge production, and this activity may partly be induced by competitive pressure from international markets. To investigate the importance of this mechanism, we present models for knowledge production. More specifically, we investigate whether firms in industries with high export intensities ceteris paribus have a higher probability of having product or process innovation.

A final concern is related to the interpretation of the industry export intensity. As mentioned, we interpret the measure as an indication of the competitive pressure in more international environments. The question is whether it is rather a broader measure of internationalization, which is important. To investigate this, we include a measure of industry import intensity in addition to the industry export intensity. If it is competitive pressure in international markets that are more important, we expect the industry export intensity to be more important in explaining internationalization. If it is driven more generally by internationalization, we expect the industry import intensity to play a more important role.

Economic Slack

As discussed above, we include variables for the firm's own participation at export markets as explanatory variables. By doing this, the results of the first stage regression are reported in Table 7.

Table 7: Digitalization and internationalization of industry	
Dependent variable: Prob (Digitalization)	
	Probit
Industry export intensity - firm export intensity	0.7425***
	(0.2272)
Firm export intensity	1.0472***
	(0.2177)
Firm export dummy	0.1087
	(0.1161)
Share of educated labor	0.4501
	(0.2757)
Log (employment)	0.4772***
	(0.0453)
Log (capital)	0.0315
	(0.0287)
Year fixed effects	yes
Region fixed effects	yes
Industry fixed effects	yes
Chi2-test for instrument	10.68
Observations	2223
R-squared	0.2183

Note: Standard errors are in parentheses. ***, ** and * indicate significanse at 1%, 5% and 10% level, respectively

We measure the instrument as the difference between the cluster export intensity and the firms export intensity, but we also include the firm export intensity as an explanatory variable for the probability of digitalization. In other words, we split the export intensity variable in two, where the first part measures the difference between cluster export intensity and firm export intensity, and the second part measures the firm export intensity. In addition to this we also include a firm export dummy that measures whether or not a firm is exporting.

It is seen that the export intensity of the firm has a point estimate around 1 that is highly significant. However, the difference between cluster and firm export intensity is still important for the probability of digitalization with a value around 0.75 that is highly significant. The applied instrument “cluster export intensity-firm export intensity” varies between -0.9 and 0.9.

In Table 8, we report the results of the second stage regression.

Table 8: Productivity Effects of Digitalization (2SLS)

Dependent variable: log (value added)	IV	IV
Const	12.3012***	12.4567***
	(0.1242)	(0.1358)
Digitalization	0.0464	0.0516
	(0.1708)	(0.1693)
Share of educated labor	0.4467***	0.2338***
	(0.0682)	(0.0865)
Log (employment)	0.9343***	0.9328***
	(0.0282)	(0.0280)
Log (capital)	0.0838***	0.0825***
	(0.0092)	(0.0093)
Digitalization x Educated labor (demeaned)		0.7868***
		(0.2572)
Instruments	Cluster export intensity-firm export intensity	Cluster export intensity-firm export intensity, (Cluster export intensity-firm export intensity) x Educated labor demeaned
Additional variables	firm export intensity, export dummy	firm export intensity, export dummy
Year fixed effects	yes	yes
Region fixed effects	yes	yes
Industry fixed effects	yes	yes
Observations	2.223	2.223
Centered R-squared	0.9368	0.9363

Note: Standard errors are in parentheses. ***, ** and * indicate significanse at 1%, 5% and 10% level,

We see that the interaction term between digitalization and skilled labor enters with a parameter of 0.7868 that is highly significant. Moreover, it is almost exactly the same value as in Table 6, where the firm export measures were not included. Hence, our results indicate that it is not because of economic slac that the export intensity of the cluster affects productivity through the digitaliztion of firms.

Knowledge production

To investigate whether the export intensity of the cluster really instrument digitalization, and not just knowledge production, we exchange digitalization with measures for product and process innovation in the first stage regression. The results are reported in Table 9.

Table 9: Knowledge production and internationalization of industry			
Dependent variable:	Dig	Product	Process
	Probit	Probit	Probit
Industry export intensity - firm export intensity	0.7425***	0.1511	0.1578
	(0.2272)	(0.2146)	(0.2208)
Firm export intensity	1.0472***	0.7679***	0.1805
	(0.2177)	(0.2192)	(0.2129)
Firm export dummy	0.1087	0.1130	0.0215
	(0.1161)	(0.0907)	(0.1127)
Share of skilled labor	0.4501	0.8936***	-0.0517
	(0.2757)	(0.2375)	(0.2355)
Log (employment)	0.4772***	0.1831***	0.2432***
	(0.0453)	(0.0438)	(0.0517)
Log (capital)	0.0315	0.0887***	0.0199
	(0.0287)	(0.0301)	(0.0335)
Year fixed effects	yes	yes	yes
Region fixed effects	yes	yes	yes
Industry fixed effects	yes	yes	yes
Chi2-test for instrument	10.68	0.50	0.51
Observations	2223	2.261	2.238
R-squared	0.2183	0.2113	0.1188
Note: Standard errors are in parentheses. ***, ** and * indicate significanse at 1%, 5% and 10% level, respectively			

We see that the export intensity of the cluster does not enter significantly in explaining any of the innovation variables. This indicates that it is not because of knowledge production that the export intensity of the cluster affects the productivity through digitalization.

Industry Import Intensity

To investigate whether it is a broader measure of internationalization than the export intensity that is important for digitalization, we include the import intensity of the cluster as an explanatory variable. The import intensity is measured as import relative to production value. The results are reported in Table 9. We see that the import intensity is non-significant, whereas the export intensity remains highly significant.

Table 10: Digitalization and internationalization of industry		
Dependent variable: Prob (Digitalization)		
	Probit	Probit
Industry import intensity	0.2771	-0.4491
	(0.5023)	(0.5208)
Industry export intensity		0.9999***
		(0.2178)
Share of educated labor	0.4710*	0.4978*
	(0.2795)	(0.2773)
Log (employment)	0.5059***	0.4960***
	(0.0450)	(0.0451)
Log (capital)	0.0468	0.0452
	(0.0292)	(0.0291)
Year fixed effects	yes	yes
Region fixed effects	yes	yes
Industry fixed effects	yes	yes
Observations	2.223	2.223
R-squared	0.2076	0.2151
Note: Standard errors are in parentheses. ***, ** and * indicate significanse at 1%, 5% and 10% level, respectively		

5. Conclusion

The main conclusion of this paper is that digitalization increases the productivity of firms, if their skill level is sufficiently high. In the average firm, the share of skilled labor is 63 percent, and for this firm, the estimated effect of digitalization is positive but insignificant. The estimated value indicates that digitalization increases the productivity of the average firm with approximately 0.5 percent. In firms where the skill level is above the average level, the effect of digitalization is much larger. In the 25 percent of the firms, where the skill level is highest, digitalization increases productivity with approximately 4 percent if we use the estimates found by applying ordinary least square, and an even larger effect if we use the instrumental variable approach.

According to our definition of digitalization, 42 percent of the firms in our sample are digitalized. Our results indicate that, if the rest of the firms become digitalized, the productivity, of these firms, increases with approximately 0.5 percent, if there is no simultaneous change in the share of skilled labor in these firms. However, our results also illustrate the importance of skills. Due to technological progress, the number of firms being digitalized will probably increase, and our results show that this gives rise to a big increase in the productivity of skilled labor. A 10 percentage point increase in the share of skilled labor, increases the productivity of non-digitalized firms with approximately 3 percent, but it increases the productivity of digitalized firms with approximately 7 percent.

The main focus of the paper has been on how digitalization affects the productivity of firms, but we have also shown that the competitive environment of firms may have an important impact on the incentive to digitalize. In particular, we have shown that the internationalization of the region and industry cluster to which a firm belongs has a big and significant impact on whether or not a firm is digitalized. It turns out that a 10 percentage point increase in the share of the production in a cluster, which is exported, increases the probability of a firm being digitalized with approximately 3 percent.

In this paper, we have estimated the impact of digitalization by estimating production functions where digitalization is an explanatory variable. Another approach would be to estimate the total factor productivity (TFP) of the firms, and afterwards analyze to which extend digitalization affects TFP. We have done that, using the method of Levinsohn and Petrin (2003), and we found basically the exact same reults as reported above.

References

Bartel, A.P., C. Ichniowski and K.L. Shaw, 2005, How Does Information Technology Really affect Productivity? Plant-level Comparisons of Product Innovation, Process Improvement and Worker Skills, NBER working paper 11773.

Bernard, A. and J.B. Jensen, 2004, Exporting and Productivity in the USA, *Oxford Review of Economic Policy*, 20, 343-357.

Bond, S. and J. Van Reenen, 2007, Microeconometric Models of Investment and Employment, *Handbook of Econometrics, Volume 6A,* 4417-4498.

Bresnahan, T., E. Brynjolfsson and L.M. Hitt, 2002, Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence, *Quarterly Journal of Economics* CXVII, 339-376.

Doms, M., T. Dunne and K.R. Troske, 1997, Workers, Wages and Technology, *Quarterly Journal of Economics 112, 253-290.*

Draca, M., R. Sadun and J. Van Reenen, 2006, Productivity and ICT: A Review of the Evidence, CEP Discussion Paper No. 749.

Dunne, T. and K, Troske, Technology Adoption and Work Force Skill in U.S. Manufacturing Plants, IZA Discussion Paper no. 1427.

Levinsohn, J. and A. Petrin, 2003, Estimating Production Functions using Inputs to Control for Unobservables, *Review of Economic Studies* 70, 317-341.

Stiroh,K.J., 2004, Reassessing the Impact of IT in the Production Function: A Meta-analysis and Sensitivity Tests, Mimeo, New York Federal Reserve.

Wooldridge, J., 2002, Econometric Analysis of Cross Section and Panel Data, MIT Press.