DELIVERABLE

Productivity effects of process vs. product digitalization

Summary

Digitalization is considered an important driver of upcoming societal and economic transformations. However, holding both promises and challenges, its effects on the performance of individual firms are still underexplored. In this paper, we disentangle the phenomenon into two distinct factors: the digitalization of processes and the digitalization of product offering. We analyse the effects of the two digitalization factors on firm-level productivity. This analysis is based on a large European-wide unique dataset combining structured information from ORBIS and PATSTAT with web-scraped information on the firms involved in high-tech manufacturing. Building on a triangular structural equation model -- including a patenting equation and a productivity equation -- we find that digitalization boosts productivity both directly and indirectly. The direct effects occur through immediate effects on productivity, while the indirect effects occur through increased patenting. However, the positive effects occur largely for product digitalization, while process digitalization on average does not significantly contribute to productivity. Interestingly quantile regression estimates show that the effects of product and process digitalization show significantly contrasting patterns across the productivity distribution. While the effects of product digitalization are largest for highly productive firms, there are mildly positive effects of product digitalization for low-productivity firms.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870822

Deliverable Information

Deliverable number and name:	**Scientific publication**
Due date:	31st March 2022
Deliverable:	D5
Work Package:	WP2
Lead Partner for the Deliverable:	ISI
Author:	Torben Schubert, Sajad Ashouri, Matthias Deschryvere, Angela Jäger, Fabiana Visentin, Lukas Pukelis, Arash Hajikhani, Arho Suominen
Reviewers:	Scott Cunningham & Hugo Hollanders
Approved by:	Arho Suominen
Dissemination level:	Public
Version	27th April 2022

Disclaimer

This document contains a description of the **BIGPROD** project findings, work and products. Certain parts of it might be under partner Intellectual Property Rt (IPR) rules so, prior to using its content please contact the consortium coordinator for approval.

In case you believe that this document harms in any way IPR held by you as a person or as a representative of an entity, please do notify us immediately.

The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of BIGPROD consortium and can in no way be taken to reflect the views of the European Union.

The European Union is established in accordance with the Treaty on European Union (Maastricht). There are currently 27 Member States of the Union. It is based on the European Communities and the member states cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice and the Court of Auditors (http://europa.eu.int/)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870822.

Productivity effects of process vs. product digitalization

Torben Schubert[1,2,*], Sajad Ashouri[3,4], Matthias Deschryvere[3], Angela Jäger[1], Fabiana Visentin[5], Lukas Pukelis[6], Arash Hajikhani[3], Arho Suominen[3,4]

[1]Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe, Germany

[2]CIRCLE – Centre for Innovation Research, Lund, Sweden

[3]VTT Technical Research Centre of Finland, Helsinki, Finland

[4]Tampere University, Tampere, Finland

[5]UNU-Merit, Maastricht University, The Netherlands

[6] PPMI, Vilnius, Lithuania

*Corresponding author: torben.schubert@isi.fraunhofer.de

Abstract: Digitalization is considered an important driver of upcoming societal and economic transformations. However, holding both promises and challenges, its effects on the performance of individual firms are still underexplored. In this paper, we disentangle the phenomenon into two distinct factors: the digitalization of processes and the digitalization of product offerings. Based on a large Europe-wide unique dataset combining structured information from ORBIS and PATSTAT with web-scraped information on digitalization in firms involved in high-tech manufacturing, we analyse the effects of the two digitalization factors on firm-level productivity. This analysis is based on a large European-wide unique dataset combining structured information from ORBIS and PATSTAT with web-scraped information on the firms involved in high-tech manufacturing. Building on a triangular structural equation model -- including a patenting equation and a productivity equation --, we find that digitalization boosts productivity both directly and indirectly. The direct effects occur through immediate effects on productivity, while the indirect effects occur through increased patenting, which in turn affects productivity. However, the positive effects occur largely for product digitalization, while process digitalization, on average, does not significantly contribute to productivity. Interestingly, quantile regression estimates show that the effects of product and process digitalization have significantly contrasting patterns across the productivity distribution. Although, while the effects of product digitalization are largest for highly productive firms, contrary to insignificant average effects, there are mildly positive effects of product digitalization for low-productivity firms.

Keywords: digitalization, processes, products, productivity, patenting, innovation

1 Motivation

Recent years have brought a marked slowdown in productivity growth in most Western economies (Syverson 2019). While some authors contend that this slowdown is due to monetary policies such as quantitative easing or declining public investments (Summers 2014, 2015), an alternative technology-focused explanation has been proposed (Gordon 2014, 2015). In his secular stagnation hypothesis, he argued that productivity slowdown is due to depleting technological opportunities associated with existing technologies. The secular stagnation hypothesis remains, however, controversial given a lack of empirical evidence (Glaeser 2014). While there is evidence that capital-embodied technological change has lost relevance since the early 2000s (Schubert and Neuhäusler 2018), several authors have argued that a whole wave of new technologies is appearing at the horizon—in particular digital ones—which could become a driving force of future productivity growth (Mokyr 2014). However, some critical voices are doubting the universally positive effects of digital technologies on the firm level. Since the provision of digital products has very low or zero marginal costs, it creates severe downward pressures on prices at the firm level (van Ark 2016). Accordingly, some recent analyses show that the productivity effects of digital technologies, although positive overall, may only emerge after years of investment (Bäck et al. 2022) and only after organisational resistances have been overcome (Horvath et al. 2019, Brynjolfsson et al. 2019, Agrawal et al. 2021). Thus, it remains unclear how digital technologies will affect productivity at the firm level. In this paper, we intend to contribute to the emergent literature by providing empirical evidence on the digitalization–productivity relationship.

Furthermore, we intend to contribute to a better understanding of the mechanisms causing the productivity effects of digitalization. Some authors have argued that one mechanism by which the productivity effects -- be they good or bad -- unfold is the organisation-wide digitalization of routines and processes (Battistella et al. 2017, Eller et al. 2020, Parida et al. 2019, Annarelli et al. 2021). This view essentially suggests that digitalization affects firm performance by spurring deep transformation of an analogous firm into a digital one and thus stresses employing a process perspective for digital technologies. An alternative view suggests that the productivity effects of digitalization are not so much the result of a digital transformation of routines, capabilities or processes, but stem from changes in the product portfolio away from analogous products towards digital versions of them (van Ark 2016, Hatzius et al. 2016). Digital technologies may, on the one hand, improve firm-level production and innovation processes. On the other hand, digital technologies can reflect product innovation. Mirroring the well-established dichotomy of product and process innovation in the innovation literature (Hall 2009, Hall and Mohnen 2013), providing empirical evidence of separating the two dimensions of digitalization is the second goal of this paper.

For empirical analysis, we rely on a unique Europe-wide dataset for 2020, which combines structured data from Bureau Van Dijk's ORBIS data with patenting data from PATSTAT with web-scraped data on firms' efforts towards process and product digitalization. The final estimation sample contains information on more than 16,000 firms from high-tech and medium-high-tech manufacturing. We set up a two-equation model including an innovation equation and a productivity equation. In the innovation equation, we analyse how patenting is affected by digitalization in terms of capabilities and products. In the productivity equation, we analyse the final productivity outcomes of digitalization alongside patenting. This triangular model setup allows for the analysis of a rich set of structural direct and indirect pass-through effects by which the productivity effects of digitalization unfold. Our main results show that digitalization is an important driver of firm-level productivity in our sample. Its effects on productivity are both direct and indirect and increase patenting, which, in turn, affects productivity. However, the productivity effects result largely from product digitalization, i.e. genuine provision of digital products, whereas the effects of process digitalization are insignificant in both the innovation and the productivity equation. Simultaneously, complementary quantile regressions provide a nuanced picture and show that product and process digitalization have significantly differential effects across the productivity distribution. Even though the average firm does not gain from process digitalization, in particular, low-productivity firms do. Conversely, for product digitalization, firms in the lower parts of the productivity distribution do not experience any productivity gains. Here, it is largely high-productivity firms, which benefit from offering more digital products.

Our contribution is three-fold. On a conceptual level, we link the discussion on the productivity effects of digital technologies to a well-established distinction between product and process innovation (OECD 2018). We argue that this distinction was implicitly present in separate streams of the literature focusing on how companies digitalise their organisation (Parida et al. 2019, Annarrelli et al. 2021) and on how firms digitalise their product portfolios (Hatzius et al. 2016, van Ark 2016, Fredrich and Bouncken 2021). Yet, these literatures appear to have evolved in parallel with only limited cross-fertilisation, which may have caused confusion about the dual nature of the mechanisms by which digitalization affects productivity. On an empirical level, we contribute to a nascent yet rapidly increasing literature analysing the performance and productivity effects of digitalization. Despite measurement concerns (Ahmad and Schreyer 2016, Grömling 2016) and the competitive challenges associated with digital technologies (van Ark 2016), our results show that there are robust positive productivity effects. Interestingly, these positive productivity effects stem largely from product digitalization and less so from process digitalization, where only low-productivity firms gain. The findings reveal that in manufacturing industries, where firms' competition is subjected to technological development, digitalization efforts do not significantly boost productivity unless they result in offering digital products. On a practical level, our results underline the central importance of digital transformation at the firm level for competitive advantage.

2 Theory

2.1 Productivity effects of product and process innovation

In the 1950s, classical growth theory (Solow 1957) had already identified changes in productivity and, thus, technological progress as the only long-term driver of per-capita growth. This view was revolutionary since it discarded other explanations, such as capital accumulation and population growth, as drivers of growth. Solow showed that because both capital accumulation and population growth are subject to decreasing marginal returns, without technological progress, any economy will eventually converge to a steady-state without per capita growth. Since then, economics and many of its sub-disciplines, such as evolutionary economics (Winter and Nelson 1982, Dosi 1988), have been concerned with the analysis of technological progress. However, it still took three more decades until endogenous growth theory developed a formal framework establishing the link between technological progress and innovation as a deliberate investment activity (Romer 1990, Aghion and Howitt 1990).

Since then the theoretical insight that growth is the result of technological progress has also spurred efforts toward analysing the effects of investments in innovation at the firm level. At this level, the most natural reflection of technological progress and innovation is solidified in the innovation-productivity link. Hall and Mohnen (2013) highlight that progress indicators which may seem obvious at the macro-level of firms in the economy are quite complex at the micro-level, where there are many structured interdependencies between firms. Innovation carries both market and technological risks, leading to the risk of failing innovation activities (Eliasson 1991, Kerr et al. 2014, Grillitsch et al. 2019). This may be particularly important when innovation takes the form of winner-takes-all races for dominant designs (Utterback and Suarez 1993) or patents (Fudenberg et al. 1983). Thus, it is not a priori clear whether the link between innovation and productivity, which appears to be fairly stable at the macro level, is still as strong at the firm level.

In line with this view, Lindholm-Dahlstrand et al. (2019), for example, have argued that innovation at the micro level takes the form of risky experimentation, whose outcomes are difficult or even impossible to predict. What makes the effects of innovation at the macro level stable is that successful innovation experiments scale to the market and thereby diffuse through the economy. Some authors have also highlighted that even if a firm manages technological risks, it is not necessarily clear that it will also reap the benefits (Teece 1986). In line with this finding, Hall and Sena (2017) find that the productivity effects of innovation depend on appropriability choices, where formal intellectual property rights (IPR) protection appears to be associated with higher returns compared with informal protection mechanisms.

Because of the ambiguities of the effects of innovation on productivity at the firm level, a rich empirical literature has emerged where the most frequently used modelling framework is the Crépon-Duguet-Mairesse (CDM) model (Crépon et al. 1998). It models the productivity outcomes of innovation within three triangular blocks of equations reflecting decisions about R&D feeding into innovation outcomes, which, in turn, affect productivity (Mohnen and Hall 2013). A continuous stream in the CDM literature has, notwithstanding the risky and experimental character of innovation, nonetheless documented the overall relatively stable positive effects of innovation on productivity in various countries (Janz et al. 2003, Lööf and Heshmati 2004, Castellacci 2011, Baum et al. 2017, Edeh and Acedo 2021, compare also the special issue in Lööf et al. 2017). Similar results are found for single-equation models analysing the productivity effects of patents (Bloom and van Reenen 2002, Hall 2011).

However, despite the overall positive effects, findings are showing that there may be differences between the types of innovative activities (Jaumandreu and Mairesse 2017). A common differentiation is between product and process innovations, where product innovations aim to introduce new or improved products, process innovations aim to change production processes. For example, Hall and Mohnen (2013) review the theoretical mechanisms linking different types of innovation to productivity, claiming that while process innovations are often introduced to reduce costs of production, product innovations are more associated with growth in revenue resulting from the consumers' increased willingness to pay. Empirically, the results are somewhat ambiguous, particularly concerning process innovations. Hall (2011), for example, reported robustly positive effects from product innovations, while the effects from process innovations remain ambiguous.

Overall we thus conclude that although the link between innovation and productivity is neither empirically nor conceptually deterministic at the firm level it is nonetheless stochastically stable and positive. However, the effects may differ between product and process innovations. In the following sections, we will argue that the relationship between digitalization and productivity follows a similar pattern. There are digital processes reflected in general firm-level digital capabilities. We also discuss digitalising product offerings, which shares features of product innovation activities. Consequently, the effects of process digitalization and product digitalization may differ, just as the effects of product and process innovations do.

2.2 Process and product digitalization

The literature on the role and purpose of digitalization has mirrored the product and process view established in the innovation literature. Specifically, digitalization enables, on the one hand, the generation of new revenue streams by defining new business models (Sund et al. 2021) and promoting competitive advantages by exploiting value-producing opportunities (Hess et al. 2016, Matt et al.

2015, Vial 2021). This coincides with the product perspective. On the other hand, it triggers the innovation of business routines towards a more efficient and flexible performance (Sund et al. 2021) by using digital technologies as well as forming digital capabilities (Annarelli et al. 2021), emphasising the process perspective.

The literature has largely dealt with the firm-level effects of process digitalization. For example, it has been argued that absorbing new digital business resources to promote novel innovations, products, or services demands strengthening firms' capabilities in technical and social contexts (Selander et al. 2013). Employing diverse resources, deployment of digital technologies, managerial cognition in initiating changes and organising IT capabilities are the main challenges in creating new markets (Annarelli et al. 2021). Deployment of heterogeneous IT resources can facilitate exchanging and processing of information toward automated tasks, enhance the information flow, advance innovative ideas and facilitate the management of the innovation process (Drnevich & Croson 2013; Mishra et al. 2007; Nylén & Holmström 2015). In the digital era, there is an increased need for timely reconfiguration of resources and improvisational capabilities to spontaneously adopt appropriate strategies and optimal solutions to successfully apply digital transformation to the innovation process (Annarelli et al. 2021; Nylén & Holmström 2015). Firms which are adapted to operating in a complex, uncertain environment and thereby exhibit significant agility are well-positioned to leverage digital technologies for innovation (Del Giudice et al. 2021).

Despite all these aspects of process digitalization, such as the adoption of digital technologies, resources and infrastructure, product digitalization has not been sufficiently discussed in the literature (Wang 2021). While process digitalization focuses on capabilities, internal knowledge and organisational structures, digitalization can be expressed as a digital product, i.e. as an artifactual outcome (Urbinati et al. 2021). In manufacturing industries, competitive advantages are promoted based on organisation, technology and/or products (Björkdahl 2020). The promotion of new technologies is more centric in high-tech industries, and R&D inputs are used for product development (Hagedoorn & Cloodt 2003). In low-tech firms, innovation is less based on R&D expenditure, and the market advantage often lies with innovative and highly-tuned production processes.

The development of innovative digital products is a profound form of digital innovation. As a product innovation, the deployment of digital components in products increases production efficiency and functionality and simplifies editability and upgradability (Björkdahl 2020; Kollmann et al. 2021; Lanzolla et al. 2021). Integrating product development with artificial intelligence can also offer new functions and engage customer value creation, which brings up opportunities for new revenue streams. However, the development of digitally integrated products (digital products) may incur high costs. Moreover, hidden burdens, such as restructuring of the development process (e.g. testing a new product, which is crucial in manufacturing industries) (Björkdahl 2020) or the reusability of development platforms, should be considered (Lanzolla et al. 2021). Especially for low-tech firms, it may require re-assessing the entire business model of the manufacturer, as these investments need potential compensation from customers. However, it can be assumed that the digitalization of products provides valuable business potential for high-tech and medium-high tech firms.

Overall, we maintain that firms' digital transformation can direct firms in two important ways. On the process side, it can contribute to more transparent information flows inside the company, i.e. towards customers and suppliers. It can also lead to flexible organisational processes and an increased absorptive capacity regarding new knowledge, market development and customer needs. On the outcome side, it can lead to the promotion of digital products. Both aspects can potentially have significant effects on productivity for example by decreasing costs or by creating new demand.

However, while the digitalization literature has emphasized the potentially far-reaching firm-level effects, up to now there are very few empirical quantitative analyses of the effects of digitalization on firm-level outcomes. Moreover, to the best of our knowledge, the existing works analysing for example the effects of digital technologies on productivity (Bäck et al. 2022, Horvath et al. 2019, Agrawal et al. 2021), did either not explicitly distinguish between the product and process components of digitalization or did not treat them simultaneously. The knowledge on how product in contrast to process digitalization affects productivity is therefore very limited.

3 Data and identification

Focusing specifically, on the question of how product and process digitalization affect firm-level outcomes, innovation and productivity effects in particular, in this section, we will present the modeling strategy and the data. We start by presenting the structural econometric model in the next subsection. Then, in the next subsection, we discuss the data sources, the construction of the key explained and explanatory variables as well as the additional control variables. Finally, we conclude with an empirical description of the key features of the dataset.

3.1 Model and identification strategy

To identify the effects of firms' process digitalization and product digitalization on productivity, we set up a triangular structural equation model. Here, in the first step, we regress the patenting activity on our digitalization measures. In the second step, we regress productivity on the digitalization measures and patenting intensity. The equations take the following canonical form:

$$\log(patint_i) = x_i\beta + \theta \cdot digiproc_i + \psi \cdot digiprod_i + u_i \qquad (1)$$

$$\log(prod_i) = x_i\delta + \vartheta \cdot \log(patint_i) + \lambda \cdot digiproc_i + \varphi \cdot digiprod_i + v_i \quad (2),$$

where x_i is a vector of control variables, u_i and v_i are unobserved potentially correlated structural errors, $patint_i$ is the patent intensity, $prod_i$ our measure of labour productivity and $digiproc_i$ and $digiprod_i$ are our measures of process and product digitalization, respectively. This structure is conceptually similar to multi-equation CDM-type models that are used to trace the effects of innovation input over innovation output on productivity. Although it is not identical conceptually, it shares many of the estimation issues that arise in CDM settings.

One aspect concerns the mediation character of this triangular model. Even if the main interest is only in estimating the direct associations between digitalization and productivity, which are captured by λ and φ, respectively, there are also indirect effects passing through the patenting variable. The following total effects, being the sum of direct and pass-through indirect effects, can be defined.

$$TE(digicap) = \frac{\partial \log(patint)}{\partial digiproc} = \underbrace{\lambda}_{direct\ effect} + \underbrace{\vartheta \cdot \theta}_{indirect\ effect} \qquad (3)$$

$$TE(digiprod) = \frac{\partial \log(patint)}{\partial digiprod} = \underbrace{\varphi}_{direct\ effect} + \underbrace{\vartheta \cdot \psi}_{indirect\ effect} \qquad (4).$$

The second aspect concerns estimation. While under somewhat restrictive zero cross-equation correlation assumptions, it is possible to estimate Eqs. (1) and (2) consistently via single-equation ordinary least squares (OLS), any non-zero cross-equation correlations will cause an estimation bias. Several methods have been proposed in the CDM literature to solve this issue. Crépon et al. (1998)

used asymptotic least squares to estimate the reduced forms and then to infer the real coefficients using a minimum distance estimator. Castellacci (2011) relied on instrumental variable regression and then plugged first-step predictions in the second-step regression instead of the original values. Baum et al. (2017) proposed a generalised structural equation estimator wherein cross-correlations are accounted for by the addition of a latent variable. Thus, the identifying assumption in this model is that non-zero cross-correlations occur because of the presence of unobserved factors, which can be sufficiently accounted for by the inclusion of a latent firm-level variable estimated based on the correlation among the observables. While this assumption imposes restrictions on the sources of the cross-correlations (such as approximability through observable factors only), nonetheless, it is likely to contribute to alleviating the problems resulting from non-zero error cross-correlations.

Therefore, we followed the approach proposed by Baum et al. (2017). Although we have log-transformed continuous measures as explained variables, we can still partially adapt the estimation approach. Instead of having to account for limited dependent variable characteristics, Eqs. (1) and (2) can be principally estimated using linear regression. Instead of a generalised structural equation estimator, which typically causes heavy convergence problems, especially when latent variables are included, we estimate the latent factor in the first step and then use a seemingly unrelated regression estimator to estimate the whole model simultaneously. The main advantage of the seemingly unrelated regressions (SUR) estimator over single-equation OLS is that it allows us to estimate the significance of the total and indirect effects of digitalization, as defined in Eqs. (3) and (4).

A further concern may relate to the fact that digitalised and non-digitalised companies may differ substantially, which begs the question of whether observable effects would hold across these samples. To control for this potential source of estimation bias, one possibility is to rely on pre-regression matching to reduce estimation issues resulting from heterogeneous subpopulations. A particularly convenient method of obtaining robust heterogeneous samples relies on introducing regression weights which are based on entropy balancing. Entropy balancing determines regression weights such that the treatment and control groups are *a priori* similar in their characteristics. We define treatment in this context as a dummy that is equal to one if either the dichotomous product or process digitalization measures are unity. We then use the determined regression weights again in our most general SUR model including the latent factor variable. Despite the conceptual superiority of SUR estimation, as a point of reference, we also report the results from single-equation OLS estimation.

To test whether the effects differ across productivity, we estimate Eq. (2) not only by estimators that identify effects on the expected values but also by using quantile regression techniques where we observe effects at the distribution deciles.

3.2 Data construction and variables

3.2.1 Data sources

Estimating Eqs. (1) and (2) requires access to unique data sources comprising information on productivity, product and process digitalization, the firms innovation activities and other relevant control variables. For this end, we construct a unique European firm-level dataset that combines web-scraped data with administrative data. The administrative data obtained via the ORBIS database of Bureau Van Dijk contain information on productivity as well as key firm characteristics such as their size, age, sector and capital intensity. Besides, as the effect of patenting on productivity has been established, the productivity equation considers firms' patenting activity for a more accurate estimation of the impact of digitalization on productivity. The patent data acquired using PATSTAT reflects the information on the annual number of patent applications in the main patent jurisdictions, i.e. the United States Patent and Trademark Office (USPTO), the European Patent Office (EPO), the Japan Patent Office (JPO) and the World Intellectual Property Organization (WIPO).

We leveraged web-scraping to gather information on our main independent variables: process digitalization and product digitalization. In the literature, measuring different aspects of digitalization in firms has typically been conducted based on survey data (Björkdahl 2020, Blichfeldt and Follant 2021, Kohtamäki et al. 2020). Still, such approaches suffer from data coverage (Arora et al. 2020). Also, studying of firms' digital activities in reference surveys such as the Community Innovation Survey (CIS) is not yet well-established. For example, the Finland CIS survey does not tackle digitalization activities consistently in each update and it provides relatively low coverage for digitalization-related questions. For instance in the Finnish CIS of 2018, less than 10% of studied entities have reacted to such questions.

Instead, our analysis measures digitalization using a novel methodology through the companies' web pages. Websites provide valuable information on company behaviour (Gök, Waterworth and Shapira 2015; Kinne and Axenbeck 2020; Axenbeck and Breithaupt 2021), which is not exclusively limited to technical expertise or innovative outputs but also expresses the firms' processes, alliance network, human resources, etc. (Gök, Waterworth and Shapira 2015). Communicating the digitalization of firm processes throughout the website enables the use of web pages in developing extensive internal process measures on a large scale (Nathan & Rosso 2022). Moreover, utilizing web pages as a data source facilitates more frequent and updated data compared with conventional data sources (Arora et al. 2020). Furthermore, as firms' competition in manufacturing industries significantly relies on their technologies and product development, websites are useful sources of information reflecting firms' product development efforts. While the previous studies investigated product development efforts through patenting activities, it is worth mentioning that only less than 5% of companies participate in patenting activities (Li et al., 2018). Moreover, considering different firms' strategies or capabilities in patenting, focusing on companies' patenting efforts does not mirror a comprehensive view of firms' activities. Instead, examination of firms' products can lead to a broader understanding of the firms' activities. As the advantage of using web-mined data in capturing firms' digital processes has been discussed, websites are also the most practical sources to identify companies' products. Such inexpensive dissemination channels are valuable means for companies to present their technologies and signal their competitive advantage to their competitors. Therefore, exploring websites enables the investigation of firms' efforts toward developing digital products, which is referred to as product digitalization.

3.2.2 Sample construction and selection

The sample for analysis is constructed based on a set of data-cleaning steps. In our study, the firm population of interest is medium high-tech and high-tech firms[1] in the European Union (EU) as in the borders before Brexit. The relevant sectors were identified based on the Eurostat aggregation of manufacturing industries based on their technological intensity and using NACE revision 2 codings. The selection then identified whether the companies belong to any of the selected NACE codes using ORBIS. These included the following NACE 3-digit codes and identifiers:

- Manufacture of basic pharmaceutical products and pharmaceutical preparations (21)
- Manufacture of computer, electronic and optical products (26)
- Manufacture of air and spacecraft and related machinery (30.3)
- Manufacture of chemicals and chemical products (20)
- Manufacture of weapons and ammunition (25.4)
- Manufacture of electrical equipment (27)
- Manufacture of machinery and equipment n.e.c. (28)
- Manufacture of motor vehicles, trailers and semi-trailers (29)
- Manufacture of other transport equipment (30) excluding Building of ships and boats (30.1) and excluding Manufacture of air and spacecraft and related machinery (30.3)
- Manufacture of medical and dental instruments and supplies (32.5)

Consequently, using ORBIS initially resulted in 181,320 firms; however, web scraping was not possible for the whole sample size because of non-existing web pages. Thus, web-mined information

[1] The analysis follows the NACE code rev 2. classification and includes the high technology (codes 21 and 26) and medium high-technology sectors (codes 20, 27, 28, 29, 30) and two medium-tech sectors (codes 25, 32). For source for classification see EC: https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an3.pdf.

could be obtained for a sample of 96,751 firms. In the next step, only the active firms—defined as firms with ORBIS non-missing information in their 2019 turnover—were retained. This left us with a sample of 47,826 firms. Further data cleaning was conducted, which included removing non-manufacturing firms. We also removed observations from countries with very limited coverage, as well as those with the smallest firm sizes. Ultimately, considering the availability of productivity, digitalization and patenting activities, the final sample contained 16,162 observations with complete information, where the largest drop out occurred as a result of missings on productivity variable as provided by ORBIS.

In terms of the data's time structure, the coverage depends on the data sources and specific variables. At the time of scraping the year 2020 was almost but not fully complete, which explains why data coverage may still be improved. The web-scraped (cross-sectional) data were obtained from the firms' websites between December 2020 and August 2021.

3.2.3 Construction of variables

3.2.3.1 Dependent variables

Patenting activity: The patenting measure was constructed by considering all patent applications in the period 2015–2019. Patents were retrieved from PATSTAT and linked to each company. The patent intensity variable was constructed by dividing the total patent applications in 2015–2019 by the number of employees in 2019. To exclude the effect of duplication for patents in the same patent family, we selected those with the earliest filing date. In a robustness check, we also constructed a dummy variable for patenting activity. The dummy variable equals one if the firm has filed at least one patent in the period 2015–2019 and zero otherwise.

*Productivity: T*he productivity measure was constructed as the value-added in 2020 divided by the number of employees in 2020, available through ORBIS.

3.2.3.2 Explanatory digitalization variables

As the analysis tackles process digitalization and product digitalization in distinct ways, the indicators for digitalization use a similar methodology, but on separate sources of information. The first relies on all information reported on the website to capture digitalization at all possible levels; however, the latter focuses only on product-related data.

Investigation of the digital deployment in technology development has been attempted mainly on patent data, using specific patent technical codes, or searching for the term "digital" throughout the patent claims (Fredrich & Bouncken, 2021). While the first approach is not applicable to other than patenting activities, the latter one suffers from inadequate coverage where digital technologies such as artificial intelligence, computer design, cloud computing, and so forth do not contain the term "digital". Therefore, a suitable approach requires to simplify the complexity of text data and categorize similar topics and information communicated on the website. Kinne & Axenbeck (2020) applied topic modeling technique as a clustering method to reduce the dimensionality of textual information. Yet, utilizing topic modeling methods, when the model is trained on the study sample, hamper the extension of the methodology to future studies, since the future study must examine the difference in quality of the train and test data. Employing a separate, but reliable classification system provides a solid approach facilitating the expansion of the study to further samples. As part of the process, our analyses use a classification system to transform the information into a scientific body of knowledge, where the textual data were prepared and transformed using Microsoft Academic Graph (MAG) topics or Fields of Study (FOS) (Ashouri et al. 2021; Hajikhani et al. 2022).

MAG comprises over 120 million publications and associated bibliometric metadata, making it a large and heterogeneous database. The transformation process involved interlinking the web-scraped textual data from the company websites to Microsoft Academic FOS codes. Consequently, a quantitative representation of the text data and publications was created. Using the information of publications associated with FOS codes, the input text data infer an association with the vector containing the associated FOS codes and their similarity score (Hajikhani et al. 2022). Therefore using this approach enables the transformation of high-dimensional textual data into structured fields of study (FOS IDs) reflecting organised information on the input information.[2]

To construct a variable measuring product digitalization capabilities, the product description obtained via web-scraping is investigated. This approach utilises actual product descriptions presented on the website rather than conventional wisdom to determine digital products. This text is envisaged as a measure of firm-level technology and science-related activities. Consequently, this text is further mapped onto FOS codes, and a vector is generated which represents the product's embedded FOS (Ashouri et al. 2021). To identify the related digital knowledge embedded in the product, the presence of computer science associated FOS IDs for the firms' product is examined; then the products containing the related FOS IDs are scored as one, otherwise zero. The aggregation and average of such product-based binary scores at the firm level reveal how the firm offers digital products in its product portfolio. Equation (5) explains the product digitalization score, where $n_{digital}$ is the number of digital products of the firm and $n_{non-digital}$ is associated with the number of non-digital products.

[2] We share both the compiled model and the code in the Jupyter notebook format with detailed descriptions of the steps. The code can be accessed from Github at https://github.com/arash-hajikhani/Bigprod_FOS/blob/main/Text-to-FOS-Similarity.ipynb. In addition, the code for FOS similarity assessment can be accessed from Github at https://github.com/arash-hajikhani/Bigprod_FOS/blob/main/FOS_Similarity.ipynb

$$prod\_digi\_cont = \frac{n_{digital}}{n_{non-digital} + n_{digital}}. \quad (5)$$

To mitigate the risk of confounding sector biases, we normalise Eq. (5) concerning each sector by calculating a dummy equalling one if the firm-level measure is larger than the sector average. As a demonstration, Appendix 1 shows two company cases where FOS IDs have been generated for their products and the product digitalization score has been calculated.

The process digitalization measure utilises the companies' website FOS IDs by linking and classifying companies' website content with FOS codes (Ashouri et al. 2021). Once the activities of a company are offered at a higher level of granularity and harmonised, this allows for additional indicators with a more thematic orientation. This measure is constructed similar to a product digitalization score, but additionally considers the relative importance (or weight) of digital FOS IDs in comparison to all the FOS codes identified on the website. To calculate the product digitalization score, we only admit whether the product is digital or not, and the level of digital integration is not examined, as it depends on the nature and function of the product.

Technically, for the process digitalization measure, the weights of all digitalization-related FOS IDs obtained by MAG are summed; Eq. (6) shows the formula for the process digitalization score, where $x^{i}_{digital}$ is the similarity score for a digital FOS ID *i* concerning the other *n* digital FOS IDs found on that web page, and $x^{j}_{non-digital}$ is the similarity score of a non-digital FOS ID *j* found on the web page concerning the other *m* non-digital FOS IDs on the web page. The final value ranges from 0 to 1, where 0 represents non-digital firms and 1 is associated with fully digital companies:

$$proc\_digi\_cont = \frac{\sum_{i=0}^{n} x^{i}_{digital}}{\sum_{i=0}^{n} x^{i}_{digital} + \sum_{j=0}^{m} x^{j}_{non-digital}}.$$

Again, we create a dichotomous sector-cleaned version of this measure by creating a dummy equalling one if the firm-level measure is larger than the sector average.[3]

3.2.3.3 Controls

Servitization: Manufacturing industries have also witnessed the increasing growth of service-offering business models in recent years. Notably, digital transformation provides opportunities for capturing value from customers and, therefore, for offering new services. Furthermore, digitalization can simplify the process of service offerings, such as upgrading or editing, by diminishing the costs of human

[3] It should also be noted that the products descriptions and their associated information, and thus their transformed FOS IDs, reflect negligible overlap with the overall website FOS IDs[3]. However, this overlap turned out to be negligibly small, since the product information corpus was very small in comparison.

labour as well as the time of service. Thus, investigating the role of digitalization in productivity requires considering the impact of servitization and its interplay with digitalization (Kohtamäki et al. 2020). Similar to digitalization measures, although surveys are the traditional methodology for examining servitization, the servitization measure in empirical analyses is constructed using web-mined data. The measure of servitization employs a novel methodology to explore service offerings based on scraped text from company websites. The keywords identified throughout the company web pages are sources of information that companies use to communicate with their audience (Ashouri et al. 2021), which have been in the focus of previous studies for firms' innovation activities (Héroux-Vaillancourt et al. 2020, Li et al. 2018). To evaluate servitization, the corresponding measure proposes a dummy variable which obtains the value one when the company website covers the keywords "service" or "service + <other terms>", and equals zero otherwise.

Spillovers within industry: Firms' boundaries are increasingly porous, and spillovers from other firms should be considered valuable inputs when looking at firm productivity. To consider the impact of the innovative activities of other firms on the productivity of the focal firm and to measure the corresponding spillovers, we identify the neighbours participating in patenting activities. We construct a dummy variable that equals one when at least one firm is in the same country and within the same sector (4-digit NACE) patents and zero otherwise.

The econometric analysis also examines the key firm characteristics impacting productivity, including firm age, firm size (described by number employees and existence of multiple establishments as control variables), and firm's sector and country dummies. All the indicators were extracted from ORBIS.

3.3 Descriptive Statistics

We present the main descriptive results in **Error! Reference source not found.**. While the overall sample has a size of 45,795 firms, we note that for the final estimation, only 16,162 firms could be included. This drop in the observation numbers comes from two facts. First, in particular, the productivity measure and the product digitalization measure contained many missing values, which a priori reduced the effective sample size. Moreover, we deliberately restricted the estimation to an equal sample for both equations. Thus, while the patenting equation, i.e. Eq. (1), could technically be estimated based on a substantially larger sample, we decided to choose the largest sample with firms having information for all variables in all equations. Nonetheless, we do not expect that the reduction in sample size will cause substantial selection biases because the mean values for the variables appear to be very similar in both the estimation and the full sample.

Table 1: Descriptive statistics (full and estimation sample)

	Full sample	Estimation sample

Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Ln labour productivity	24,263	3.932	0.811	16,162	3.991	0.795
Ln patent intensity 2015–2019	40,445	0.003	0.065	16,162	0.002	0.055
Patent active 2015–2019	45,795	0.014	0.119	16,162	0.015	0.122
Product digitalization	32,199	0.337	0.473	16,162	0.330	0.470
Process digitalization	45,795	0.379	0.485	16,162	0.360	0.480
Servitization	45,735	0.167	0.373	16,162	0.246	0.431
Neighbours with patent	45,795	0.641	0.480	16,162	0.664	0.472
Ln employees	40,445	3.306	1.720	16,162	3.555	1.676
Ln firmage	45,694	3.058	0.745	16,162	3.144	0.696
Multiple establishments	45,191	0.735	0.441	16,162	0.751	0.433
NACE 21	45,795	0.036	0.187	16,162	0.036	0.187
NACE 26	45,795	0.153	0.360	16,162	0.136	0.343
NACE 20	45,795	0.140	0.347	16,162	0.145	0.353
NACE 27	45,795	0.137	0.343	16,162	0.132	0.339
NACE 28	45,795	0.387	0.487	16,162	0.426	0.494
NACE 29	45,795	0.071	0.256	16,162	0.070	0.255
NACE 30	45,795	0.009	0.094	16,162	0.009	0.092
AT	45,795	0.015	0.121	16,162	0.011	0.106
BE	45,795	0.012	0.107	16,162	0.015	0.122
BG	45,795	0.013	0.112	16,162	0.019	0.136
CZ	45,795	0.033	0.180	16,162	0.010	0.101
DE	45,795	0.164	0.370	16,162	0.036	0.186
ES	45,795	0.100	0.300	16,162	0.144	0.351
FI	45,795	0.025	0.157	16,162	0.000	0.000
FR	45,795	0.073	0.259	16,162	0.061	0.240
HU	45,795	0.018	0.133	16,162	0.000	0.000
IT	45,795	0.330	0.470	16,162	0.484	0.500
PL	45,795	0.061	0.240	16,162	0.074	0.262
PT	45,795	0.012	0.110	16,162	0.018	0.133
RO	45,795	0.015	0.121	16,162	0.016	0.126
SE	45,795	0.043	0.203	16,162	0.040	0.197
SK	45,795	0.013	0.114	16,162	0.000	0.000
UK	45,795	0.049	0.217	16,162	0.054	0.226

When inspecting the overall characteristics of the estimation sample, we see that the average firm in the sample was a relatively small firm with approximately 35 employees. However, there was great heterogeneity. The largest firm in the sample had more than 400,000 employees (not displayed). Across the considered high-tech and med-high tech manufacturing sectors (NACE 20, 21, 26, 27, 28, 29, 30), the distribution was fairly balanced, even though only a very small minority of firms belonged to NACE 30 (other vehicle construction). The average firm has a log labour productivity of 4, which corresponds to a value added of 76,000€ per employee. The share of patent active firms in the sample is 1.5%.

Focusing on the key explanatory variables, we see several very interesting patterns. For both the capability and the product digitalization measure, we see that roughly one-third of all firms are identified as more active than the industry average. Looking at our controls, approximately 25% of our firms offer services. The share of firms with neighbours with patenting is relatively high: 66%. This, however, is also because we were forced to apply a relatively coarse measure based on country-sector co-locations. Using more fine-grained regional geo-location information is technically possible but was out of the scope given the time frame of the project. The average firm was relatively young, being 3.5 years old; 75% of the firms in our sample had multiple establishments.

A final noteworthy point is that the country-wise distribution was substantially more skewed, with 48% of the firms based in Italy. This selectivity was due to missing data problems in ORBIS, which were substantially smaller for Italy than for other countries. While this does not allow for any claims to country-wise representativeness, we do however include country dummies to control for country-level heterogeneity. Moreover, since the overall sample is relatively large, we are at least principally able to identify country heterogeneity, in particular for countries with a decent number of observations such as Germany, Spain, France, the UK and Poland. Nonetheless, the skewed country-wise distribution should still be taken into account when interpreting our results.

4 Results

In this section, we present the main regression results of the effects of product and process digitalization on the firms' innovation activities and their realized productivity levels. In the second part of this section, we conclude with a series of robustness checks to show that the key results are not overly dependent on the specific modelling choices.

4.1 Main results

Table 2 reports our main regression results for the patenting equation (Eq. 1 of Section 3.1). It shows that ourTable 2 main expectation concerning a complementary relationship between patenting and digitalization is corroborated at least for product digitalization. In the most general specification in Column 3 including simultaneously the product and process digitalization measure, the coefficient of product digitalization is positive and significant at the 1% level, where the size of the coefficient indicates that a firm having a higher product digitalization score than the industry average has an approximately 0.27% (b=0.0027, p<0.01) higher patent intensity. In the same regression model, process digitalization does not appear to affect patenting intensity in any of the regressions (b=-0.0002, p>0.05). In The results are not dependent on the inclusion of the latent factor as suggested by Baum et al. (2017), which suggests that the endogeneity correction accounting for cross-equation error-correlations does not significantly affect the results.

Besides the main results on product and process digitalization, we find in addition that the servitization measures are significantly negative in all regressions implying that the firms offering also services have a significantly lower patenting intensity, which may be explained by the fact that services are more difficult to patent.

Table 2: Effect of digitalization on patenting activity

	(1) Ln patent intensity	(2) Ln patent intensity	(3) Ln patent intensity
Neighbours with patent	0.00114	0.00110	0.00067
	(1.00)	(0.97)	(0.56)
Servitization	−0.00339**	−0.00363***	−0.00368***
	(−3.24)	(−3.45)	(−3.49)
Ln employees	−0.00049	−0.00057	−0.00056
	(-1.49)	(−1.73)	(−1.68)
Ln firm age	−0.00036	−0.00029	−0.00029
	(−0.54)	(−0.44)	(−0.44)
Multiple establishments	−0.00020	−0.00024	−0.00027
	(−0.19)	(−0.23)	(−0.25)
Product digitalization		0.00274**	0.00273**
		(2.89)	(2.89)
Process digitalization		−0.00033	−0.00028
		(−0.36)	(−0.31)
Constant	0.01085*	0.01031*	−0.00044

	(2.47)	(2.34)	(-0.05)
Sector dummies	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes
Latent factor	No	No	Yes
N	16162.00000	16162.00000	16162.00000
r2	0.00426	0.00477	0.00488
P	0.00001	0.00000	0.00000

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results of the productivity regression are reported in Table 3. First, we note that, as expected, patent intensity affects productivity in all regressions. The coefficient is also quite stable, where the size indicates that a 1% increase in patenting is associated with a 0.35% increase in productivity. Most importantly, we find the pattern repeated that product digitalization leads to an increase in productivity (elasticity of 0.03%) while process digitalization does not.

Several further effects on the control variables are interesting to note. First, unlike the negative association with patent intensity, the coefficient of servitization in the productivity regression is positive and strongly significant. Firms offering services have 7% higher productivity. Moreover, we observe that spillovers also appear to play a role. Firms that have neighbouring firms with patents have approximately 3.4% higher productivity. Finally, we note that, in general, larger and older firms, as well as firms with multiple establishments, also enjoy higher productivity.

Table 3: Digitalization, patenting intensity and productivity

	(1) Ln productivity 2020	(2) Ln productivity 2020	(3) Ln productivity 2020
Ln patent intensity	0.36402^{***}	0.35719^{***}	0.35685^{***}
	(3.84)	(3.77)	(3.77)
Neighbours with patent	0.03488^{*}	0.03453^{*}	0.03317^{*}
	(2.55)	(2.52)	(2.33)
Servitization	0.07682^{***}	0.07440^{***}	0.07425^{***}
	(6.10)	(5.87)	(5.86)
Ln employees	0.03811^{***}	0.03689^{***}	0.03694^{***}
	(9.65)	(9.30)	(9.31)
Ln firm age	0.04054^{***}	0.04159^{***}	0.04159^{***}
	(5.05)	(5.18)	(5.18)
Multiple establishments	0.10265^{***}	0.10212^{***}	0.10204^{***}
	(8.00)	(7.96)	(7.95)
Product digitalization		0.03624^{**}	0.03623^{**}
		(3.18)	(3.18)
Process digitalization		0.00645	0.00657
		(0.58)	(0.60)
Constant	3.57120^{***}	3.56093^{***}	3.52760^{***}
	(67.57)	(67.20)	(31.61)
Sector dummies	Yes	Yes	Yes

Country dummies	Yes	Yes	Yes
Latent factor	No	No	Yes
N	16162.00000	16162.00000	16162.00000
r2	0.30986	0.31033	0.31033
P	0.00000	0.00000	0.00000

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As already explained in Section 3.1, the simultaneous equation estimation using SUR allows for the identification of the direct and indirect effects of digitalization on productivity. These results are presented in Table 4. Overall, the largest effect stems from the direct effect on productivity. The indirect effect is significant at the 5% level, but it has a 0.1% relatively small magnitude. When summed up, the total effect amounts to 3.7%, as compared to 3.6% from the direct effect. Again, all decompositions for the process digitalization measure are not significant.

Table 4: Direct, indirect and total associations between digitalization and productivity

	Direct	Indirect	Total
Product digitalization	0.036***	0.001*	0.037***
Process digitalization	0.006	−0.000	0.006

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.2 Differences across the productivity distribution

We now turn to an analysis of how the effects of product and process digitalization differ between the different quantiles of the productivity distribution. The main regression results are reported in Table 5, whereas a visual representation of the effects of the two digitalization dimensions is presented in Figure 1.

Table 5: Digitalization, patenting intensity and productivity at different locations of the productivity distribution

	(1) Ln productivity 2020	(2) Ln productivity 2020	(3) Ln productivity 2020	(4) Ln productivity 2020	(5) Ln productivity 2020
	10%	25%	50%	75%	90%
Ln patent intensity	0.22700	0.13749	0.38650***	0.39098***	0.28171
	(1.56)	(1.48)	(4.55)	(3.62)	(1.66)
Neighbours with patent	0.03956	0.05633***	0.03333*	0.01414	0.01647
	(1.78)	(3.97)	(2.57)	(0.86)	(0.64)
Servitization	0.05623**	0.05106***	0.05216***	0.06691***	0.06638**
	(2.89)	(4.11)	(4.59)	(4.63)	(2.92)
Product digitalization	0.00464	0.01429	0.03677***	0.04292***	0.05222*
	(0.27)	(1.28)	(3.59)	(3.30)	(2.56)
Process digitalization	0.02579	0.02634*	0.01156	-0.00087	-0.02475
	(1.52)	(2.43)	(1.17)	(-0.07)	(-1.25)
Ln employees	0.10091***	0.07735***	0.05005***	0.01459**	-0.02203**
	(16.58)	(19.89)	(14.05)	(3.22)	(-3.10)

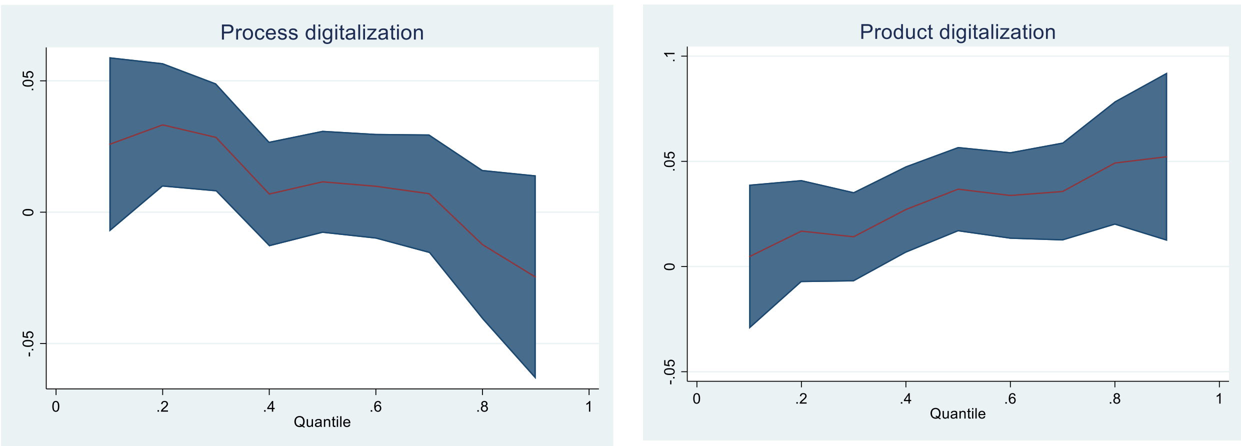
Ln firmage	0.09482***	0.05497***	0.02851***	0.00545	-0.00818
	(7.70)	(6.98)	(3.96)	(0.59)	(-0.57)
Multiple establishments	0.05736**	0.07028***	0.10287***	0.13620***	0.13390***
	(2.92)	(5.59)	(8.94)	(9.31)	(5.83)
Constant	2.39133***	3.04257***	3.40203***	4.05452***	4.46912***
	(15.58)	(31.03)	(37.89)	(35.53)	(24.94)
Sector dummies	Yes	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes	Yes
Latent factor	Yes	Yes	Yes	Yes	Yes
N	16162.00000	16162.00000	16162.00000	16162.00000	16162.00000
Pr2	0.2674	0.2471	0.1975	0.1524	0.1251

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We observe that the effects of product digitalization increase across the productivity distribution, where the effects at the 10% are with 0.0046 close to zero but steadily increase to 0.0522 at the 90%-quantile. Moreover, at the lower end of the productivity distribution, up to the 30% quantile, the effects are non-significant. Only above the 40% quantile of the productivity distribution did the effects turn significant, taking their highest value at the 90% quantile of the distribution with a coefficient of 0.052 as opposed to 0.036 at the median. For process digitalization, the effects are highest at the lower end of the productivity distribution, peaking at the 20% and 30% quantile, whereas they also turn significant at the 5% level. Beyond the 30% quantile, the coefficient continuously declines and even turns negative, even though the effects are not statistically significant. Overall, the results suggest that low-productivity firms seem to benefit from process digitalization (even though the average firm in our sample does not). At the same time, product digitalization appears to largely benefit high-productivity firms, whereas low-productivity firms do not gain from introducing digital products.

Figure 1: Representation of the effects of process and product digitalization across the productivity distribution (point estimators and 95% confidence intervals)



4.3 Robustness checks

To further probe the robustness of our results, we performed several robustness checks. One particularly important concern is that digitalised and non-digitalised companies may differ substantially a priori. This may lead to estimation biases if the effects of digitalization do not extend across the equations. The results of the entropy balancing SUR model are presented in Table 6. We see that the coefficients differ only mildly, with significances being largely unchanged.

Table 6: Digitalization, patenting intensity and productivity (SUR with entropy balancing)

	(1) Ln patent intensity	(2) Ln productivity 2020
Ln patent intensity		0.33187**
		(3.22)
Neighbours with patent	0.00060	0.03682**
	(0.57)	(2.64)
Product digitalization	0.00259**	0.03693**
	(3.00)	(3.26)
Process digitalization	−0.00040	0.00471
	(−0.48)	(0.43)
Servitization	−0.00324***	0.06833***
	(−3.47)	(5.59)
Ln employees	−0.00055	0.03650***
	(−1.87)	(9.53)
Ln firm age	−0.00042	0.04704***
	(−0.70)	(6.04)
Multiple establishments	−0.00038	0.10340***
	(−0.39)	(8.02)
Constant	−0.00147	3.51902***
	(-0.18)	(32.66)
Sector dummies	Yes	Yes
Country dummies	Yes	Yes
Latent factor	Yes	Yes
N	16162.00000	16162.00000
r2	0.00490	0.30040
p	0.00000	0.00000

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Beyond the entropy balancing approach, we conducted several further robustness checks. The first consists of an approach to test for dependence on the choice of covariates. For this, we implement a Bayesian model averaging WALS estimator in which we treat all variables—except for the digitalization measures, the latent factor and the patent intensity in Eq. 2—as potentially dispensable. As we can see in Table 7, the results do not vary in any relevant respect, which implies that, overall, the coefficients do not appear to depend much on the choice of the set of control variables.

Table 7: Digitalization, patenting intensity and productivity (WALS model averaging)

	(1) Ln patent intensity	(2) Ln productivity 2020
Ln patent intensity		0.35651***
		(3.76)
Product digitalization	0.00254**	0.03780***
	(2.70)	(3.32)
Process digitalization	−0.00030	0.00602
	(−0.33)	(0.55)
Neighbours with patent	0.00040	0.03986**
	(0.41)	(2.80)
Servitization	−0.00251**	0.07072***
	(−2.62)	(5.61)
Ln employees	−0.00037	0.03550***
	(−1.30)	(9.00)
Ln firm age	−0.00025	0.04527***
	(−0.42)	(5.66)
Multiple establishments	−0.00037	0.10510***
	(−0.39)	(8.22)
Constant	0.00192	3.54135***
	(0.28)	(38.96)
Sector dummies	Yes	Yes
Country dummies	Yes	Yes
Latent factor	Yes	Yes
N	16162.00000	16162.00000

In our third additional robustness check, we included all firm observations from countries with less than 100 observations, which we excluded in the main regressions. Including these observations did not change the results. Fourth, in the main regressions, we restricted the sample to the firms which have information on all variables in all regressions. Using the maximum number of observations per individual equation does not change the results. Fifth, instead of using the patenting intensity variable in Eqs. (1) and (2), we defined a patent activity dummy variable, which is one for firms with a positive patenting intensity. The results remained qualitatively unchanged; however, we observed that the product digitalization measure was significant in patent activity regression only at the 10% level. This may, however, be because the patenting dummy has substantially lower variation. Overall, we conclude that our results are unlikely to be affected strongly by heterogeneous samples, specific choices of the patenting indicator or imposed sample restrictions.

Overall, results confirmed that digitalization has robust and positive effects on firm-level innovation and productivity. In specific, the results on productivity mirror and corroborate the statistically stable relationship between firm level innovation and productivity, which has been documented in a well-established and extensive literature in innovation economics. Our finding that the effects of product and process digitalization differ are also in line with previous findings from that literature, which showed that the productivity effects of product innovations are more stable than those of process innovations. On a conceptual level, our results thus suggest that the nascent literature on the productivity effects of digitalization may benefit from integrating key insights from the innovation literature. This should however clearly not hide away the fact that digital and non-digital innovations may also have important differences, which limits the ability to transfer directly all insights.

5 Discussion and Conclusions

Our analysis showed several important findings. There is a complementary relationship between patenting and digitalization. However, this only holds for product digitalization, while process digitalization does not affect patenting intensity. Also in other respects product and process digitalization differ. While we did not find effects of process digitalization on the expected value of firm-level productivity, we did find positive effects for low productivity firms. For high productivity firms, we did not find any effects for process digitalization. For product digitalization instead, the effects of productivity were increasing across the productivity distribution. The combined consequences of this, and the resultant implications of this, can now be made.

Relying on a large firm-level dataset combining structured information from ORBIS with web-scraped data from company websites, we analysed the role of digitalization in firm innovation processes and its productivity effects by differentiating between process digitalization and product digitalization. The fact that product and process digitalization appear to have differential effects on productivity which differ also across the productivity distribution suggests, that firms need to consider the alignment of their digital transformation investments with the development of digital products according to their productivity. This distinction, we argue, has deep theoretical implications and mirrors, to some degree, the differentiation between product and process innovations. It thus confirms the dual nature of digitalization by showing that the mechanisms governing the productivity effects of digitalization can relate either to digitally upgrading internal processes or to digitalising product offerings.

From a managerial perspective, the findings are important because they convey how and when firms can draw value from digitalization. At least in high-tech manufacturing, the strongest effects appear to be gained from enriching products with digital elements. Instead, the average firm does not benefit much from digitalising its processes.

A further aspect concerns the sectoral aspect. While high-tech manufacturing firms used to derive their productivity gains from technological advances in tangible goods, nowadays, intangible and digital components appear to have become an important driver too. This suggests that to avoid forgoing the potential to increase productivity, firms need to consider where the integration of digital components can provide them with a competitive edge. As a side note, the role of digital intangibles does not appear to be the only driving force. Although not at the centre of the analysis, our results also show that the inclusion of services plays a central role.

A specific focus area that would be of high relevance in future research is the integration of digital services. While we did not analyse this nexus in this paper, we highlight that the underlying data source exploiting a rich type of previously untapped information pool from company websites principally allows for such an analysis.

6 References

Aghion, P., & Howitt, P. (1990). A model of growth through creative destruction.

Agrawal, A. K., Gans, J. S., & Goldfarb, A. (2021). AI Adoption and System-Wide Change (No. w28811). National Bureau of Economic Research.

Ahmad, N., & Schreyer, P. (2016). Are GDP and productivity measures up to the challenges of the digital economy?. International Productivity Monitor, (30), 4.

Annarelli, A., Battistella, C., Nonino, F., Parida, V., & Pessot, E. (2021). Literature review on digitalization capabilities: Co-citation analysis of antecedents, conceptualization and consequences. Technological Forecasting and Social Change, 166, 120635.

Ashouri, S., Suominen, A., Hajikhani, A., Pukelis, L., Schubert, T., Türkeli, S., van Beers, C., & Cunninghum, S. (2021). Indicators on firm level innovation activities from web scraped data. Available at SSRN 3938767.

Battistella, C., De Toni, A. F., De Zan, G., & Pessot, E. (2017). Cultivating business model agility through focused capabilities: A multiple case study. Journal of Business Research, 73, 65-82.

Baum, C. F., Lööf, H., Nabavi, P., & Stephan, A. (2017). A new approach to estimation of the R&D–innovation–productivity relationship. Economics of Innovation and New Technology, 26(1-2), 121-133.

Bäck, A., Hajikhani, A., Jäger, A., Schubert, T., & Suominen, A. (2022). Return of the Solow-paradox in AI? AI-adoption and firm productivity (No. 2022/1). Lund University, CIRCLE-Centre for Innovation Research.

Björkdahl, J. (2020). Strategies for Digitalization in Manufacturing Firms. California Management Review, 62(4), 17–36. https://doi.org/10.1177/0008125620920349

Bloom, N., & Van Reenen, J. (2002). Patents, real options and firm performance. The Economic Journal, 112(478), C97-C116.

Brynjolfsson, E., Rock, D., & Syverson, C. (2019). 1. Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics (pp. 23-60). University of Chicago Press.

Castellacci, F. (2011). How does competition affect the relationship between innovation and productivity? Estimation of a CDM model for Norway. Economics of Innovation and New Technology, 20(7), 637-658.

Crepon, B., Duguet, E., & Mairesse, J. (1998). Research, innovation, and productivity.

Del Giudice, M., Scuotto, V., Papa, A., Tarba, S. Y., Bresciani, S., & Warkentin, M. (2021). A self-tuning model for smart manufacturing SMEs: Effects on digital innovation. Journal of Product Innovation Management, 38(1), 68-89.

Drnevich, P. L., & Croson, D. C. (2013). Information technology and business-level strategy: Toward an integrated theoretical perspective. MIS Quarterly, 483–509.

Dosi, G. (1988). Sources, procedures, and microeconomic effects of innovation. Journal of economic literature, 1120-1171.

Edeh, J. N., & Acedo, F. J. (2021). External supports, innovation efforts and productivity: Estimation of a CDM model for small firms in developing countries. Technological Forecasting and Social Change, 173, 121189.

Eliasson, G. (1991). Modeling the experimentally organized economy: Complex dynamics in an empirical micro-macro model of endogenous economic growth. Journal of Economic Behavior & Organization, 16(1-2), 153-182.

Eller, R., Alford, P., Kallmünzer, A., & Peters, M. (2020). Antecedents, consequences, and challenges of small and medium-sized enterprise digitalization. Journal of Business Research, 112, 119-127.

Fredrich, V., & Bouncken, R. B. (2021). The double-edged sword of digitalization for firm performance. In ISPIM Conference Proceedings (pp. 1-16). The International Society for Professional Innovation Management (ISPIM).

Fudenberg, D., Gilbert, R., Stiglitz, J., & Tirole, J. (1983). Preemption, leapfrogging and competition in patent races. European Economic Review, 22(1), 3-31.

Glaeser, E.L. (2014). Secular joblessness. In: Teulings, C., & Baldwin, R. (2014). Secular stagnation: facts, causes and cures. CEPR Print.

Gordon, R. J. (2014). The turtle's progress: Secular stagnation meets the headwinds. Secular stagnation: facts, causes and cures, 47-59.

Gordon, R.J. (2015). Secular stagnation: A supply-side view. American Economic Review, 105(5), 54-59.

Grillitsch, M., Schubert, T., & Srholec, M. (2019). Knowledge base combinations and firm growth. Research Policy, 48(1), 234-247.

Grömling, Michael (2016) "The Digital Revolution - New Challenges for National Accounting?" World Economics, January-March, 1-13.

Hagedoorn, J., & Cloodt, M. (2003). Measuring innovative performance: is there an advantage in using multiple indicators? Research Policy, 32(8), 1365–1379. https://doi.org/https://doi.org/10.1016/S0048-7333(02)00137-3

Hajikhani, A., Pukelis, L., Suominen, A., Ashouri, S., Schubert, T., Notten, A., & Cunningham, S. W. (2022). Connecting firm's web scraped textual content to body of science: Utilizing microsoft academic graph hierarchical topic modeling. MethodsX, 9, 101650.

Han, J., Kamber, M., Pei, J. (2012) Getting to Know Your Data, i Data Min., 39–82. https://doi.org/10.1016/b978-0-12-381479-1.00002-2.

Hatzius, J., Pandl, Z., Phillips, A., Mericle, D., Pashtan, E., Struyven, D., & Thakkar, A. (2016). Productivity paradox v2. 0 revisited. US Economics Analyst.

Hall, B. H. (2011). Innovation and productivity (No. w17178). National bureau of economic research.

Hall, B. H., & Sena, V. (2017). Appropriability mechanisms, innovation, and productivity: evidence from the UK. Economics of Innovation and New Technology, 26(1-2), 42-62.

Hess, T., Matt, C., Benlian, A., & Wiesböck, F. (2016). Options for formulating a digital transformation strategy. MIS Quarterly Executive, 15(2).

Horvat, D., Kroll, H., & Jäger, A. (2019). Researching the effects of automation and digitalization on manufacturing companies' productivity in the early stage of industry 4.0. *Procedia Manufacturing*, *39*, 886-893.

Janz, N., Lööf, H., & Peters, B. (2003). Firm level innovation and productivity-is there a common story across countries?.

Jaumandreu, J., & Mairesse, J. (2017). Disentangling the effects of process and product innovation on cost and demand. Economics of Innovation and New Technology, 26(1-2), 150-167.

Kerr, W. R., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as experimentation. Journal of Economic Perspectives, 28(3), 25-48.

Kinne, J., & Axenbeck, J. (2020). Web mining for innovation ecosystem mapping: a framework and a large-scale pilot study. *Scientometrics*, *125*(3), 2011–2041. https://doi.org/10.1007/s11192-020-03726-9

Kohtamäki, M., Parida, V., Patel, P. C., & Gebauer, H. (2020). The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. Technological Forecasting and Social Change, 151, 119804. https://doi.org/https://doi.org/10.1016/j.techfore.2019.119804

Kollmann, T., Stöckmann, C., Niemand, T., Hensellek, S., & de Cruppe, K. (2021). A configurational approach to entrepreneurial orientation and cooperation explaining product/service innovation in digital vs. non-digital startups. Journal of Business Research, 125, 508–519. https://doi.org/https://doi.org/10.1016/j.jbusres.2019.09.041

Lanzolla, G., Pesce, D., & Tucci, C. L. (2021). The digital transformation of search and recombination in the innovation function: Tensions and an integrative framework. Journal of Product Innovation Management, 38(1), 90-113.

Lindholm-Dahlstrand, Å., Andersson, M., & Carlsson, B. (2019). Entrepreneurial experimentation: a key function in systems of innovation. Small Business Economics, 53(3), 591-610.

Lööf, H., & Heshmati, A. (2004). Sources of finance, R&D investment and productivity: Correlation or causality?.

Lööf, H., Mairesse, J., & Mohnen, P. (2017). CDM 20 years after. Economics of Innovation and New Technology, 26(1-2), 1-5.

Matt, C., Hess, T., & Benlian, A. (2015). Digital transformation strategies. Business & Information Systems Engineering, 57(5), 339–343.

Mishra, A. N., Konana, P., & Barua, A. (2007). Antecedents and consequences of internet use in procurement: an empirical investigation of US manufacturing firms. Information Systems Research, 18(1), 103–120.

Mohnen, P., & Hall, B. H. (2013). Innovation and productivity: An update. Eurasian Business Review, 3(1), 47-65.

Mokyr, J. (2014). Secular stagnation? Not in your life. In: In: Teulings, C., & Baldwin, R. (2014). Secular stagnation: facts, causes and cures, 83. CEPR Print.

OECD (2018). Oslo manual 2018: guidelines for collecting, reporting and using data on innovation. OECD publishing.

Nathan, M., & Rosso, A. (2022). Innovative events: product launches, innovation and firm performance. Research Policy, 51(1), 104373. https://doi.org/https://doi.org/10.1016/j.respol.2021.104373

Nylén, D., & Holmström, J. (2015). Digital innovation strategy: A framework for diagnosing and improving digital product and service innovation. Business Horizons, 58(1), 57–67.

Parida, V., Sjödin, D., & Reim, W. (2019). Reviewing literature on digitalization, business model innovation, and sustainable industry: Past achievements and future promises. Sustainability, 11(2), 391.

Romer, P. M. (1990). Endogenous technological change. Journal of political Economy, 98(5, Part 2), S71-S102.

Selander, L., Henfridsson, O., & Svahn, F. (2013). Capability search and redeem across digital ecosystems. Journal of Information Technology, 28(3), 183–197.

Solow, R. M. (1957). Technical change and the aggregate production function. The review of Economics and Statistics, 312-320.

Summers, L.H. (2014). Reflections on the ‘new secular stagnation hypothesis’. In: Teulings, C., & Baldwin, R. (2014). Secular stagnation: facts, causes and cures. CEPR Print, 27-38.

Summers, L.H. (2015). Demand side secular stagnation. American Economic Review, 105(5), 60-65.

Sund, K. J., Bogers, M. L., & Sahramaa, M. (2021). Managing business model exploration in incumbent firms: A case study of innovation labs in European banks. Journal of Business Research, 128, 11-19.

Syverson, C. (2017). Challenges to mismeasurement explanations for the US productivity slowdown. Journal of Economic Perspectives, 31(2), 165-86.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. Research policy, 15(6), 285-305.

Urbinati, A., Manelli, L., Frattini, F., & Bogers, M. L. (2021). The digital transformation of the innovation process: orchestration mechanisms and future research directions. Innovation, 1-21.

Utterback, J. M., & Suárez, F. F. (1993). Innovation, competition, and industry structure. Research policy, 22(1), 1-21.

Van Ark, B. (2016). The productivity paradox of the new digital economy. International Productivity Monitor, (31), 3.

Vial, G. (2021). Understanding digital transformation: A review and a research agenda. Managing Digital Transformation, 13–66.

Wang, G. (2021). Digital reframing: The design thinking of redesigning traditional products into innovative digital products. Journal of Product Innovation Management.

Winter, S. G., & Nelson, R. R. (1982). An evolutionary theory of economic change. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.

Appendix 1

Example of two companies with low and high product digitalization score. The following tables show how the product description in the company website can be transformed to FOS ID to reduce the sparsity of textual data.

CIRCONTROL (https://circontrol.com)	Product digitalization Score: 0.33	
Selected products scraped from website	**Product Keywords**	**FOS IDs and Similarity Scores**
eVolve Smart: The most suitable charger for urban environments **Application** Designed to be installed in both public access environments (urban spaces, shopping centres, car parks, airports, petrol stations…) and private areas (companies, community car park sites…) where its intelligent capabilities offer a range of possibilities which improve the user and/or operator experience. **Concept Design** Nowadays, the concept of smart cities demands an innovative design for its urban equipment, especially for EVSE (EV Supply Equipment) due to its innovative nature. With its stylised shape and modern lines, the eVolve series meets this demand. In addition, not only has the exterior design been taken into account, but also the daily conditions (operational and environmental) that EVSE has to withstand.	**eVolve Smart** - charge station, urban environment, energy market, initial capital, strength design, one pole, capital investment, service station, fast charging station	'{"Primary station": 36, "Charging station": 35, "Transfer station": 32, "Transmitter station": 29, "Signal station": 29, "Station P": 29, "Smart market": 28, "Traffic station": 28, "Total station": 26, "Cleaning station": 26, "Harmony (ISS module)": 26, "Smart environment": 25, "Smart system": 25, "Station model": 24, "Tube station": 24, "Fixed capital": 23, "Urban acupuncture": 23, "Capital (economics)": 23, "Smart growth": 23, "Smart device": 23, "Urban design": 23, "Space Station Freedom": 23, "Return on capital": 22, "Capital employed": 22, "Glass Poling": 22, "Agricultural experiment station": 22, "Risk-adjusted return on capital": 22, "Cost of capital": 22, "Capital formation": 22, "Urban density": 22, "Smart products": 22, "Circulating capital": 21, "Urban morphology": 21, "Urban metabolism": 21, "Financial capital": 21, "Poling": 21, "Smart module": 21, "Physical capital": 21, "Urban anthropology": 21, "Capital Consumption Allowance": 21, "Economic capital": 21, "Capital intensity": 21, "Television station": 21, "Marginal cost of capital schedule": 20, "Urban climatology": 20, "Constant capital": 20, "Energy market": 20, "Urban climate": 20, "Smart glass": 20, "Urban structure": 20, "Return on capital employed": 20, "Smart material": 20, "Process capital": 20, "Virtual Reference Station": 20, "Smart objects": 20, "Smart key": 20, "Capital control": 19, "Utility pole": 19, "Weighted average cost of capital": 19, "Minimum capital requirement": 19, "Urban theory": 19, "Capital services": 19, "Marginal product of capital": 19, "Minimum capital": 19, "Capital deepening": 19, "Pole splitting": 19, "Project SMART": 19, "Urban studies": 19, "Capital loss": 19, "Capital adequacy ratio": 19, "Market engineering": 19, "Factor market": 19, "Equity capital markets": 18, "Pole shift hypothesis": 18, "Capital good": 18, "Capital requirement": 18, "Compressor station": 18, "Nursing Stations": 18, "Working capital": 18, "Urban computing": 18, "Market environment": 18, "Urban economics": 18, "Nonmarket forces": 18, "Market microstructure": 18, "Capital structure": 18, "Market saturation": 18, "Urban planning": 18, "Legal capital": 18, "Market transformation": 18, "Smart lock": 18, "Urban spatial structure": 18, "Capital cost": 18, "SMART criteria": 18, "Urban culture": 17, "Market rate": 17, "Urban history": 17, "Market analysis": 17, "Market impact": 17, "Real-time charging": 17, "Capital gain": 17}

eVolve Rapid The most compact and affordable DC charging solution **Application** The eVolve Rapid series is designed to bring fast charging to small private sites that do not have access to large high-power electricity infrastructure (small EV fleets, car dealerships, carsharing companies, small private car parks, etc.) to improve the charging speed of electric vehicles at these locations without large investments. **Concept Design** Thanks to Circontrol's years of experience in the field of slow and/or semi-fast public charging, as well as its fast-charging Raption series, the eVolve Rapid series has been launched as a fast-charging solution that is perfect for small private locations and features two models for wall or floor installation (wallbox and post). Designed to reduce charging times for electric vehicles with larger batteries, which will go from having a range of just over 40 or 60 km (depending on the model) if charged on AC for one hour to being able to travel nearly 150 km with the same charging time and the same enclosure.	**eVolve Rapid charging points** - charge station, load management, fast charging, range anxiety, gas emissions, distribution networks, ultra fast, detect and avoid	'{"Charging station": 45, "Real-time charging": 41, "Charging order": 38, "Online charging system": 38, "Range anxiety": 38, "Standard Charge": 37, "Charge (physics)": 36, "Charge conservation": 36, "Charge amounts": 36, "Charge control": 35, "Charge type": 34, "Charge density": 34, "Method of image charges": 33, "Charge number": 32, "Elementary charge": 31, "Partial charge": 31, "Electric charge": 31, "Effective nuclear charge": 31, "Charge controller": 31, "Space charge": 30, "Floating charge": 30, "Inductive charging": 29, "Surface charge": 29, "Trickle charging": 29, "Moderate anxiety": 28, "Mild anxiety": 28, "Neutral Charge": 28, "Anxiety reduction": 28, "Point particle": 27, "Charge cycle": 27, "Reducing anxiety": 27, "Anxiety test": 26, "Test Anxiety Scale": 26, "Charge sharing": 26, "Anxiety reaction": 25, "Charged current": 25, "Anxiety states": 25, "Test anxiety": 25, "Shaped charge": 25, "Charge carrier": 25, "Charged particle": 25, "Mass-to-charge ratio": 24, "CHELPG": 24, "Anxiety score": 24, "Ionic potential": 24, "Charge ordering": 24, "Acute anxiety": 23, "Anxiety sensitivity": 23, "Convergent charging": 23, "Spacecraft charging": 23, "Chronic anxiety": 23, "Primary station": 22, "FCAPS": 22, "Taylor Manifest Anxiety Scale": 22, "Manifest Anxiety Scale": 22, "Electric distribution network": 22, "Network management application": 22, "Static Charges": 22, "Distribution management system": 22, "Electrostatics": 22, "Element management system": 22, "Central charge": 22, "Foreign language anxiety": 21, "Electrostatic induction": 21, "CHARGE syndrome": 21, "Network management": 21, "Charged particle beam": 21, "Core charge": 21, "Alleviating anxiety": 21, "Compressor station": 20, "Spin\\u2013charge separation": 20, "State of charge": 20, "Charge radius": 20, "Network management station": 20, "Somatic anxiety": 20, "Load management": 20, "Sedimentation potential": 20, "State-Trait Anxiety Inventory": 20, "Formal charge": 20, "Common Management Information Protocol": 20, "Active Network Management": 20, "Network information system": 20, "Highly charged ion": 20, "Load profile": 20, "Ejection charge": 20, "Transfer station": 19, "Policy and charging rules function": 19, "Poisson\\u2013Boltzmann equation": 19, "Load balancing (electrical power)": 19, "Load shifting": 19, "Signal station": 19, "Total station": 19, "Emission intensity": 19, "Emission inventory": 19, "Electric-field screening": 19, "Overcharge": 19, "Electrostatic voltmeter": 19, "Network monitoring": 19, "Anxiety": 19, "Structure of Management Information": 19}
Powerful system able to provide car-finding solutions based on License Plate Recognition within lanes or in each parking space, offering users the location and route to their own car via the user application.	**automated vehicle license recognition** - license plate, parking space, different level, **vehicle detection**, user satisfaction, **smart parking**	'{"License number": 48, "IVMS": 42, "Remotely operated vehicle": 40, "Vehicle tracking system": 39, "Vehicle-to-vehicle": 39, "Remote control vehicle": 38, "License": 38, "License control": 38, "Vehicle search": 37, "Automatic vehicle location": 35, "Vehicle category": 35, "License Status": 34, "MIT License": 31, "Vehicle (Transportation)": 31, "Automobile handling": 31, "Vehicular communication systems": 30, "Police vehicle": 30, "Articulated vehicle": 29, "Monroney sticker": 29, "Vehicle frame": 29, "Vehicle inspection": 29, "Low emission vehicle": 29, "Vehicle dynamics": 29, "Treatment Vehicle": 28, "Sport utility vehicle": 28, "Automated guided vehicle": 28, "Emergency vehicle": 28, "Green vehicle": 28, "Vehicle engineering": 28, "Vehicle infrastructure

		integration": 27, "Fleet telematics system": 27, "Heavy goods vehicle": 27, "CarSim": 27, "Water vehicle": 27, "Cruise control": 26, "Space vehicle": 26, "Vehicle bus": 26, "Miles per gallon gasoline equivalent": 26, "Curb weight": 26, "Rescue vehicle": 25, "Vehicle accident": 25, "Computer user satisfaction": 25, "Dual-mode vehicle": 25, "Active safety": 25, "Yaw": 25, "Software license": 24, "Dedicated short-range communications": 24, "Pharmaceutical Vehicles": 24, "Vehicle emissions control": 24, "Yaw-rate sensor": 24, "Alternative fuel vehicle": 23, "Hybrid vehicle": 23, "All terrain vehicles": 23, "Solar vehicle": 22, "Vehicle fire": 22, "Cooperative Adaptive Cruise Control": 22, "Vehicle Information and Communication System": 21, "Overtaking": 21, "Electronic toll collection": 21, "On-board diagnostics": 21, "Government incentives for plug-in electric vehicles": 20, "Rear-end collision": 20, "Wheelbase": 20, "Treatment satisfaction": 20, "Torque vectoring": 20, "Electronic stability control": 20, "Vehicle routing problem": 20, "Advanced driver assistance systems": 20, "Tailgating": 20, "Platoon": 20, "Toll": 19, "Intelligent transportation system": 19, "Parking sensors": 19, "Chassis dynamometer": 19, "Event data recorder": 19, "Natural gas vehicle": 19, "Vehicle-specific power": 19, "Directional stability": 18, "Telematics": 18, "Patient satisfaction score": 18, "Road train": 18, "Trip computer": 18, "Battery electric vehicle": 18, "IEEE 802.11p": 17, "User journey": 17, "Drug vehicle": 17, "Parking lot": 17, "New Car Assessment Program": 17, "Rollover": 17, "Unique user": 17, "Air Transport Network": 17, "User modeling": 17, "Satisfaction work": 17, "Lane departure warning system": 17, "Collision avoidance system": 17, "Medical license": 17, "Patient satisfaction": 17, "Hydraulic hybrid vehicle": 17, "Traffic violation": 17, "Sprung mass": 17}

SPHERON-VR (www.spheron.com)	Product digitalization Score: 1	
Selected products scraped from website	**Product Keywords**	**FOS IDs and Similarity Scores**
SCENECAM 2.0 Spheron-VR AG, known globally as pioneers in HDR (High Dynamic Range) camera technology and experts in visual content management software, today launch their SceneCam in version 2.0.The SceneCam is able to capture 360º x 180º spherical images automatically in a single scan and produces 26 f-stops of dynamic range (32-bit per channel x RGB = 96-bit image data per pixel) and a clarity of more than 100 mega-pixels* in	**SCENECAM™ -** led array, digital documentation, spherical image, vertical axis, crime scene**, high-dynamic-range imaging, computer-generated imagery, high dynamic range, fill factor**	'{"Dynamic imaging": 33, "Dynamic range": 32, "High dynamic range": 31, "Crime scene": 29, "Wide dynamic range": 29, "Imagery analysis": 28, "High-dynamic-range imaging": 28, "Digital imaging": 28, "Common Source Data Base": 26, "Imaging science": 25, "Spherical image": 25, "Computer-generated imagery": 25, "Digital array": 25, "Dynamic array": 24, "Internal documentation": 23, "Imaging technology": 22, "Imaging problem": 21, "Guided imagery": 21, "Digital image": 21, "Dynamic method": 21, "Imaging chain": 21, "Software documentation": 21, "Computational photography": 20, "Dynamic factor": 20, "Written Documentation": 20, "Tone mapping": 20, "Sparse array": 20, "Spherical mean": 20, "Digital image analysis": 20, "Array data type": 19, "Square array": 19, "Spherical model": 19, "Imaging order": 19, "Image analysis": 19, "Biological imaging": 19, "Dynamic equilibrium": 19, "Spherical sector": 19, "Computed tomography laser mammography": 19, "Spherical segme

resolution (* when compared to a DSLR sensor).		nt": 19, "Optical imaging": 19, "Array factor": 19, "Auditory imagery": 19, "Dynamic Extension": 19, "Hashed array tree": 19, "NIIRS": 19, "Structural dynamics": 19, "Parallel array": 19, "Array data structure": 19, "Hybrid computer": 19, "Liquid dynamics": 19, "Dynamic structure factor": 19, "Spherical design": 18, "Optoacoustic imaging": 18, "Hybrid array": 18, "Complex dynamics": 18, "Array access analysis": 18, "Imaging Tool": 18, "Digital geometry": 18, "Technical documentation": 18, "Mental image": 18, "Additional documentation": 18, "Dynamic problem": 18, "Documentation": 18, "Planar array": 18, "Integral imaging": 18, "Digital image processing": 18, "Molecular imaging": 18, "Spherical cap": 18, "Spherical geometry": 17, "Dynamic range compression": 17, "Array gain": 17, "Digital radiography": 17, "Spherical shaped": 17, "Joint imaging": 17, "Digital photography": 17, "System dynamics": 17, "Microscope image processing": 17, "Filling factor": 17, "Dynamic data": 17, "Preclinical imaging": 17, "Phased array ultrasonics": 17, "Ranging": 17, "Chemical imaging": 17, "Digital microscopy": 17, "Ground sample distance": 17, "Dynamic Scan": 17, "Medical imaging": 17, "Spectral imaging": 17, "Image resolution": 17, "Array processing": 17, "Classes of computers": 17, "Spherical joint": 16, "Imaging Procedures": 16, "Dynamic loading": 16, "Cellular imaging": 16, "Analog image processing": 16, "Phased array": 16, "Imaging Signal": 16, "L\\u00e9n\\u00e1rt sphere": 16, "Image quality": 16}
SCENECENTER Spheron-VR AG, known globally as pioneers in HDR (High Dynamic Range) camera technology and experts in visual content management software, announce its SceneWorks division is today launching the latest version of their SceneCenter Forensic and SceneCenter Framework visual content management software. The technology offers the production for virtual onsite scene documentation. SceneWorks have tailored solutions for areas such as Police Crime Scene forensic documentation, Security, Critical Infrastructure and also Industrial Industries such as Rail, Nuclear, Oil and Gas, Utilities, Construction and other Visual Asset Management applications.	**SCENECENTER™** crime scene, **high dynamic range, software technology,** critical infrastructure protection, workflow	'{"Critical infrastructure protection": 49, "Windows Workflow Foundation": 40, "Workflow technology": 40, "Workflow engine": 39, "Workflow Management Coalition": 39, "XPDL": 39, "Workflow management system": 38, "Crime scene": 36, "Workflow application": 36, "Workflow": 36, "Workflow patterns": 33, "Critical infrastructure": 32, "Event-driven process chain": 31, "Backporting": 24, "Software sizing": 24, "Software review": 23, "Software verification and validation": 23, "Long-term support": 23, "Package development process": 23, "Software construction": 23, "Software distribution": 23, "Software quality analyst": 23, "Software peer review": 23, "Software asset management": 23, "Software design description": 23, "Software quality management": 23, "Social software engineering": 22, "Software technical review": 22, "Goal-Driven Software Development Process": 22, "Avionics software": 21, "Personal software process": 21, "Software quality control": 21, "Software rot": 21, "Software measurement": 21, "Software walkthrough": 21, "Resource-oriented architecture": 21, "Software development": 21, "Software quality": 21, "Study software": 21, "Software crisis": 21, "Software Engineering Process Group": 21, "Software system": 20, "Software factory": 20, "Software reliability testing": 20, "Software business": 20, "Software metric": 20, "Software requirements": 20, "Software quality assurance": 20, "Software framework": 20, "COSMIC software sizing": 19, "Protection mechanism": 19, "Scientific workflow system": 19, "Software development process": 19, "Software": 19, "\'Active\' protection": 19, "Software Problem": 19, "Software feature": 19, "Team software process": 19, "Software release life cycle": 18, "Software analytics": 18, "Medical software": 18, "Software analysis pattern": 18, "Software Design and Development": 18, "Software evolution": 18, "Software archaeology": 18, "Software Evaluation": 18, "Real-time C

		ontrol System Software": 18, "Commercial software": 18, "Component-based software engineering": 18, "Integrated software": 18, "System software": 18, "Software Process simulation": 18, "Software engineering": 18, "Software standard": 17, "Custom software": 17, "Software design": 17, "Software maintenance": 17, "Dynamic range": 17, "Connascence": 17, "Protection procedure": 17, "Software assurance": 17, "Programming complexity": 17, "Stress testing (software)": 17, "Software bus": 17, "Pair testing": 17, "Monolithic application": 17, "Technology dynamics": 17, "Software repository": 17, "IEC 62304": 16, "Software configuration management": 16, "Bloodstain pattern analysis": 16, "Use Case Points": 16, "Embedded software": 16, "Lehman\'s laws of software evolution": 16, "Computer-aided technologies": 16, "Protection ring": 16, "Wide dynamic range": 16, "National data protection authority": 16, "Educational software": 16, "DO-178B": 16}
Todays Crime Scene Investigation (CSI) is executed by dedicated experts from multiple disciplines. One of their goals is to present a concise analysis of their findings. With our business unit SceneWorks-forensics, we offer an integrated workflow to increase the objectivity and the efficiency of modern crime scene investigation.	**SCENEWORKS™** crime scene, **high dynamic range**, forensic science, documentation system;	'{"Crime scene": 45, "Common Source Data Base": 38, "Internal documentation": 33, "Questioned document examination": 31, "Software documentation": 30, "Written Documentation": 28, "Technical documentation": 26, "Additional documentation": 25, "Documentation": 24, "Bloodstain pattern analysis": 22, "Regulatory documentation": 20, "Dynamic range": 19, "Nursing documentation": 18, "Wide dynamic range": 17, "Language documentation": 17, "CSI effect": 17, "User analysis": 17, "Trace evidence": 17, "System dynamics": 17, "Forensic biology": 16, "High dynamic range": 16, "Dynamic method": 16, "Javadoc": 15, "Complex dynamics": 15, "Dynamic equilibrium": 15, "Dynamic Extension": 15, "Forensic science": 15, "Liquid dynamics": 14, "Dynamic problem": 14, "Structural dynamics": 13, "Dynamic simulation": 13, "Offender profiling": 13, "Documentation science": 13, "Dynamic data": 13, "Documentalist": 13, "Dynamic simulation model": 13, "Dynamic loading": 13, "High-dynamic-range imaging": 12, "Forensic profiling": 12, "Ranging": 12, "Dynamic scaling": 12, "Dynamic testing": 12, "Quantum dynamics": 12, "Chemical Dynamics": 12, "Dynamical system": 12, "Dynamic structure factor": 11, "Shape dynamics": 11, "Forensic chemistry": 11, "Forensic geology": 11, "Forensic photography": 11, "Inverse dynamics": 11, "Dynamic positioning": 11, "Force dynamics": 11, "Dynamic range compression": 11, "Dynamic pressure": 11, "Dynamic logic (digital electronics)": 10, "Dynamic factor": 10, "Serial crime": 10, "Progress note": 10, "Dynamic load testing": 10, "Dynamic decision-making": 10, "Symbolic dynamics": 10, "Dynamic Tension": 10, "Dynamic balance": 10, "Newtonian dynamics": 10, "Contact dynamics": 10, "Human dynamics": 10, "Passive dynamics": 10, "Dynamic covalent chemistry": 10, "Working range": 10, "Gas dynamics": 10, "Dynamic efficiency": 10, "Dynamic relaxation": 10, "Differential dynamic microscopy": 10, "Range safety": 10, "Rigid body dynamics": 10, "Forensic archaeology": 10, "Time-variant system": 9, "Dynamic modulus": 9, "Vehicle dynamics": 9, "Dynamic similarity": 9, "Dynamic compliance": 9, "Dynamic aperture": 9, "Historical dynamics": 9, "High-dynamic-range video": 9, "Dynamic Scan": 9, "Dynamic frequency scaling": 9, "Dynamic assessment": 9, "System of systems": 9, "Protein dynamics": 9, "System equivalence": 9, "Time-invariant system": 9, "Multibody system": 9, "Reaction dynamics": 9, "Dynamic demand": 9, "Newmark-beta method": 9, "System requirements": 9, "So

		cial dynamics": 9, "Mid-range": 9, "Range sear ching": 9}

For more information, please contact

Dr. Arho Suominen (Consortium leader)
Tel. +358 50 5050 354
arho.suominen@vtt.fi

About BIGPROD

BIGPROD is a research project focusing on Big Data based analysis of productivity using webscraped data. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 870822.

The project partners in the project are Quantitative Science and Technology Studies team, Foresight-driven Business Strategies, 1) VTT Technical Research Centre of Finland, Competence Center Innovation and Knowledge Economy (Coordinator), 2) Fraunhofer ISI, Economics of Knowledge and Innovation team, 3) UNU-MERIT, Maastricht University, 4) Public Policy and Management Institute, 5) Economics of Technology and Innovations, Faculty of Technology, Policy and Management, 6) Delft University of Technology, Economics of Technology and Innovations, 7) Faculty of Technology, Policy and Management, Delft University of Technology

www.bigprod.eu