PROMOTING DIGITALIZATION THROUGH INFORMATION DISSEMINATION

Abstract

The U.K. Small Business Digital Capability Programme Challenge Fund 2014 aimed at disseminating information and training about the use of digital technologies among SMEs. Using a novel firm-level dataset of web technologies, our matched differences-test strategy shows that, following the implementation of the program, treated SMEs were more likely to increase their web presence, adopted more sophisticated web technologies, and enlarged their operating boundaries through e-commerce. The program reduced the digital divide between SMEs and large companies and enabled digitalization in more geographically remote locations. As a result, firms observed positive effects on performance, labor outcomes, and likelihood of survival. *JEL* Codes: G33, J24, O15, O33, O38.

Keywords: Digital Technology, Human Capital, SMEs, Digital Divide

"Thousands of potential customers are searching online for local small businesses and without an online profile businesses will lose out [...] To make sure consumers get the best deal and small businesses spread their nets far and wide, the government is investing in a range of advice to help them do more online."

Matthew Hancock, U.K. Business and Enterprise Minister, 2014.

I. INTRODUCTION

The past two decades have witnessed a significant development of digital technologies with a number of studies hinting at their possible positive effects on growth and productivity (e.g., Syverson 2011; Brynjolfsson and McAfee 2014; Greenstein 2020).[1] Despite such potential benefits, however, the adoption of these technologies has been very heterogeneous across firms. This has led to a digital divide between the minority of the most productive firms that have been adopting and using these technologies and the majority of firms in the rest of the economy that lag behind. This digital divide has negative implications for aggregate productivity and, ultimately, leads to income differences (e.g., Acemoğlu et al. 2014; Comin and Mestieri 2018; Berlingieri et al. 2020).

In a world without frictions, managers are able to identify and invest in all positive NPV projects to maximize firm value because information and capital are costlessly available. The lack of capital resources is typically examined in academic research as a primary friction that leads to inefficient investment decisions. However, when it comes to investment in digital technologies, the lack of managerial awareness of such technologies, their potential benefits, and, more generally, the lack of digital skills seem to be the main barrier to implement this type of investment, yet the least investigated one.[2]

Such lack of awareness and skills is particularly prevalent among managers of small- and medium-sized enterprises (SMEs) as well as entrepreneurs, who lack not only hard, technical knowledge, but also soft skills (e.g. professional attitude). In fact, SMEs have been identified as the companies digitalizing their businesses at a slower pace than

[1] Digital technologies can be beneficial to businesses on a broad range of dimensions, from marketing, advertising and communication (e.g., use of a website or the implementation of e-commerce), to strategic planning (e.g., Big Data analysis), general administration and IT systems (e.g., cloud computing services), and production, pre-production, and logistics (e.g., supply chain management software).

[2] The misalignment of incentives between employees who are supposed to use the new technology and owners can also act as a barrier to technology adoption (Atkin et al. 2017). However, this potential channel is muted in our setting where mostly owner-managers, not a third party, introduce new technologies.

large firms in most economies worldwide (e.g., Arendt 2008; Millán et al. 2021; OECD 2021a). The implication of a lagged adoption of digitalization could be severe since SMEs represent the backbone of most economies, accounting for 99% of all businesses, between 50% and 60% of value added, and employing the majority of the labor force (BEIS 2021; OECD 2021b). So far, however, there is little empirical evidence on the link between lack of awareness of new technologies and digital skills, on one hand, and the degree of digitalization, on the other hand.

This paper examines whether a U.K. government initiative, aimed at increasing awareness of digital technologies and training in digital skills, helped spur more digitalization among SMEs, with possible effects on their operating boundaries, their digital divide from other firms in the economy, and real economic outcomes. The Small Business Digital Capability Program Challenge Fund (henceforth the *Challenge Fund*), launched in 2014 by the Department of Business, Innovation and Skills (BIS), provided funding with the objective to support the provision of training and information dissemination of digital technologies to SMEs in England. Instead of subsidizing the investments in digital technologies made by individual firms, the funding was provided to so called local enterprise partnerships (LEPs), which cover pre-defined areas in England. Selected LEPs used the funding to organize workshops delivered by ICT experts and networking events with digital suppliers, especially, to inform SMEs of the online business opportunities.

The funding was distributed on a competitive basis by selecting LEP proposals based on their value for money, reach, the originality of their proposed methods to disseminate information and training on digital knowledge to the local SMEs, economies of scale, and sustainability (BIS 2015). Eventually, 22 areas in England were selected for the program, while 24 areas in England and the rest of Great Britain (i.e. Wales and Scotland) did not receive funding. From an empirical perspective, a potential concern with such a government program could be that the BIS might have selected areas that already had better technological infrastructure, stronger local economies, and/or closer local political ties with the central government in order to prove the success of the initiative. Treated areas (i.e. those that received funding) may then show better post-program performance measures compared to the control areas (i.e. those that didn't receive any funding) simply because they had better pre-conditions to begin with. We, thus, start our investigation by comparing treated with control areas in terms of broadband access and usage, several metrics of economic strength, and political affiliation in the years prior to the start of the program. We do not find any statistically significant difference between treated and control areas across all dimensions, suggesting that, even if the allocation of the funding across LEPs was not random, it does not appear to be based on regional characteristics that could predict the success of the program. We conclude that the *Challenge Fund* program represents an interesting quasi-exogenous change in the provision of information and training in digital technologies that can inform the debate on the role of awareness of digital technologies on SMEs actual digitalization.

We exploit the geographical variation of the allocation of funding by implementing pairwise-matched difference-in-differences (DID) estimations in which the years 2011-2014 (2015-2019) represent the pre- (post-) program period. To control for potential confounding effects from firms' characteristics, our sample consists of firms in treated and control areas matched on size, age, financial constraints, growth opportunities, profitability, and industry. In line with the objectives of the program, we build on the prior that SMEs without internet presence before the program were more likely to benefit from the program and implement digitalization in treated areas than similar SMEs located in the control areas.

A novel contribution of our study is the use of a new measure of corporate digitalization, which is available at the firm-level and for the entire spectrum of firm sizes. The large majority of previous studies rely on either industry-/country-level proxies of digitalization or survey evidence for relatively limited samples of large firms (e.g., Fitzgerald et al. 2014; Gal et al. 2019). In contrast, we assemble a detailed database of U.K. firms' websites and the technologies embedded in every website, such as the type of web hosting the website uses, e-commerce functionalities, and analytics software. BuiltWith Pty Ltd collects such information through regular snapshots of every single website. As such, we are able to construct a time series of website technologies for every firm in our sample from 2011 to 2019 that allows us to explore both the extensive margin of the implementation of a website, and the intensive margin of the degree of technological sophistication embedded in each website. Across our matched sample, we count 4,310 unique technologies with a total number of 602,112 firm-technologies observations.

We couple this database with detailed financial data provided by the Financial Analysis Made Easy (FAME) database for a final sample that includes 148,965 firm-year observations generated by 17,598 unique SMEs, where SME is defined following the U.K. government guideline of a company with less than 250 employees and an annual turnover of less than €50 million.[3] The top three industries are manufacturing, whole-

[3]See:https://www.gov.uk/government/publications/fcdo-small-to-medium-sized-enterprise-sme-

sale and retail trade, human health- and social work-related industries, which count for approximately one-third of the sample.

Results from DID regressions show that on the extensive margin, the average web presence of treated SMEs after the program is approximately 13% higher than that of the control group. On the intensive margin, treated SMEs employ 14% more different technologies in their websites compared with the control SMEs, making their websites more digitally sophisticated. There is also an increasing trend over time for treated SMEs to have a website. At the end of our sample period, in 2019, treated SMEs are 17% more likely to have a website compared with their counterparts, which is more than three times the difference estimated in 2015 (7%). This suggests that the program enhanced the SMEs awareness of digital technologies enabling them to, eventually, become more digitally savvy than their counterparts.

Next, we zoom in on a particular type of website technologies, i.e., those related to e-commerce. E-commerce represents an important aspect of companies' digitalization as it allows them to change their boundaries from a simple brick-and-mortar store to a business that is also conducted online (e.g., Kickul and Gundry 2001; Leong et al. 2016). Crucially, e-commerce is enabled by website-based technologies. We use the e-commerce related technologies in the BuiltWith data to investigate whether the improvement in digital skills has also spurred the adoption of e-commerce by SMEs. The DID results show that, on the extensive margin, treated firms are 17% more likely to start an e-commerce platform on their websites than control firms. On the intensive margin, treated firms adopt the equivalent of 17% more e-commerce-related technologies relative to the control firms in the aftermath of the program.[4] The findings further confirm our expectation that enhanced awareness and training in digital skills facilitates the adoption of digitalization among SMEs.

action-plan/small-to-medium-sized-enterprise-sme-action-plan

[4]For example, Russell Building Products Limited is a producer of roofing materials with an average annual turnover and total assets before the program of about £16.67 million and £18.22 million, respectively. Russell Building Products is located in the Greater Birmingham LEP, which received funding from the *Challenge Fund*. Until 2014, the company did not have a website. In 2015, the firm set up a website and adopted analytics and email related technologies. Until 2019, their website has further adopted online payment and mobile supporting technologies. TOA Corporation UK Limited is specialised in the provision of audio equipment to commercial customers across several different industries, with an average annual turnover and total assets before the program of about £4.27 million and £3.36 million, respectively. The company is located in the Coast to Capital LEP, which was supported by the *Challenge Fund*. In 2015, the firm adopted advertising, analytics, content feeding, media, and mobile supporting technologies. Until 2019, their website has further adopted email, online payment, and multi-language supporting technologies.

This evidence prompts us to investigate further whether the increase in digitalization by treated SMEs had any effect on reducing the digital divide between large companies and SMEs, so often debated by academics and policy-makers worldwide. U.K. survey evidence confirms that before the *Challenge Fund*, SMEs adopted proportionally fewer digital technologies and have made less e-commerce sales than large corporations (ONS 2014).[5] Our findings suggest that following the launch of the program, the difference in digitalization between large firms and treated SMEs has reduced by 10% more than the difference observed between large firms and control SMEs. In addition, we examine the geographical dimension of the digital divide. We find that in the aftermath of the program, treated SMEs in rural areas or in areas far from universities were able to start a website, implement sophisticated technologies, and launch e-commerce platforms at the same rate as treated SMEs in urban areas and in areas close to universities. Although this may not be direct evidence that the digital spatial divide has narrowed down, it highlights, nonetheless, that the reach of the program went beyond urban areas by making access to digitalization possible in more remote geographical locations traditionally considered less digitally informed.

In the next part of the paper, we investigate whether digitalization had any impact on real economic outcomes, such as performance, employment and labor productivity growth, and firm survival. We find that in the post-program period, relative to the control firms, treated SMEs had 1% higher revenue that was converted into an improved company's return on assets, almost 1% higher than that of their counterparts. Further, we show that treated SMEs experienced 1% higher employment growth rates and 3% higher value-added per employee. Establishment-level data further validate such findings showing a positive and significant increase in IT workers and IT expenses of about 6% and 12%, respectively, for the treated group compared to the control one. In addition, we find that treated SMEs are 25% less likely to be dissolved than their counterparts. This result holds even when we account for the years affected by the COVID-19 pandemic, suggesting that adoption of digitalization equipped SMEs with a certain degree of resilience against shocks as severe as the pandemic.

In the final step of our investigation we discuss and rule out financial constraints as an alternative channel that may prevent firms to digitalize (e.g., OECD 2021a). First, our setting specifically captures the information and skills channel because of the nature of the *Challenge Fund* program that provided funding for training at the LEP level and

[5]See:https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/articles/monitoringecommerce/2014-08-07#toc.

not direct subsidies to SMEs. Second, our empirical strategy based on firm-matching compares SMEs with similar levels of financial constraints, among other characteristics. As such, we do not find any significant difference in digitalization when we contrast relatively more financially constrained SMEs with less constrained ones. Further, evidence shows that individuals' age shapes the propensity to innovate, making business digital transformation more challenging, when it comes to older managers and employees, who are considered to have less skills tailored to technological innovations and being more exposed to organizational inertia (e.g., OECD 2021b). We show that treated SMEs with mature directors have been affected by the program significantly more than SMEs with younger directors in the likelihood to set up a website and in adopting more digital technologies, suggesting that the program successfully enhanced digital awareness even beyond certain behavioral traits.

Our study contributes to the current debate on corporate digitalization in several ways. We are the first to show that enhancing awareness of digital technologies and training of digital skills eases the access to digitalization for firms traditionally lagging behind, such as SMEs. This has further effects on the firm boundaries through the adoption of e-commerce technologies. On a larger scale, more digitalization for SMEs significantly contributes to narrow the digital divide with large companies and to make digitalization more diffuse within a country by reaching remote rural areas. Unreported results further show that this has, ultimately, benefited those areas in terms of increased economic growth and total factor productivity at the macro level. Further, contrary to previous studies (e.g., Caselli and Coleman 2001; Chinn and Fairlie 2007; Haller and Siedschlag 2011; Gal et al. 2019; Cusolito, Lederman, and Peña 2020; Cong, Yang, and Zhang 2022; Hoffreumon and Labhard 2022)), our novel measures of digitalization at the firm-level allow us to capture both the extensive and intensive margin of digitalization for a large sample of SMEs.

Also, we extend the literature on the real effects of digitalization at the firm level. Several studies show a positive impact of investment in digital technologies on productivity performance and firm size (for extensive reviews see Draca, Sadun, and Van Reenen 2009; Syverson 2011; Gal et al. 2019); while others find either no effect on firm productivity or heterogeneity in the performance and geography effects (e.g., Acemoğlu et al. 2014; DeStefano, Kneller, and Timmis 2018; DeStefano, Kneller, and Timmis 2023). Granular digitalization data at the firm level coupled with the pairwise-matching DID estimations around the implementation of the *Challenge Fund* allow us to overcome several identification issues that have affected previous studies and show that digitalization of SMEs has, ultimately, positive effects on performance, employment and labor productivity, and likelihood of survival.

Our paper further contributes to the literature on the effects of government interventions to support SMEs and their innovation activities. Several studies show that government financial support through loan guarantees and credit quality certifications, as well as tax relief programs, can support SMEs borrowing, investment, and employment decisions (e.g., Columba, Gambacorta, and Mistrulli 2010; Lelarge, Sraer, and Thesmar 2010; Brown and Earle, n.d.; Gonzalez-Uribe and Paravisini 2019; D'Ignazio and Menon 2020; Gonzalez-Uribe and Wang 2020; Bachas, Kim, and Yannelis 2021; Bonfim, Custódio, and Raposo 2023). Other studies have documented the role played by direct government subsidies in promoting R&D spending, patenting, and improving the performance of SMEs (e.g., Lerner 1999; Wallsten 2000; Lach 2002; Almus and Czarnitzki 2003; Bronzini and Iachini 2014; Bronzini and Piselli 2016; Howell 2017). The unique design of the *Challenge Fund* program enables us to explore the causal relation between a government program that enhances the awareness and training of SMEs about digital technologies and firms' adoption of such technologies. Our findings support the notion that the lack of knowledge about digital technologies represents an important barrier for SMEs to become more digitalized and that a government intervention seems to be effective in reducing such information barrier.

This has also relevant policy implications. Over the last decade several countries have promoted programs to support digital adoption, which vary significantly in their approach, focus, and implementation (OECD 2021c). The results from the *Challenge Fund* program show that rather than providing companies with direct monetary subsidies for the implementation of digital technologies, a relatively less expensive alternative that can help diffuse digital adoption with positive effects on real outcomes is a government intervention that aims at training and informing companies.

The remainder of our paper is organized as follows. Section II explains the institutional details surrounding the *Challenge Fund* program. Section III describes our data and the methodology. Sections IV and V discuss the digitalization and real outcomes results, respectively. Section VI provides robustness tests, and section VII concludes.

II. INSTITUTIONAL DETAILS

II.A. *The Challenge Fund and Local Enterprises Partnerships*

In 2014, the Department of Business, Innovation and Skills (BIS) in the U.K. launched the 'Do More Online' campaign. The scope of this campaign was to help SMEs improve their digital skills, including their presence on the internet and their access to e-commerce technology given the increased consumer demand for SMEs' digitalization (BIS 2014).

The 'Do More Online' campaign comprised three initiatives. Two of them consisted of generic online resources made available through a government-backed website that provided general support for businesses, and a website created in partnership with a U.K.-based charity organisation ("Go ON UK") specializing in projects that contribute toward digital inclusions.[6] Although available nationwide, both campaigns were not particularly visible to the interested parties. Between 2015 and 2019, after the launch of the 'Do More Online' campaign, the average traffic of these two online resources was only 61 visits per month, which is approximately 97% lower than the average traffic of other government websites that are related to businesses during the same period (i.e., 2,297 visits per month).[7,8]

In contrast, the third initiative, the Small Business Digital Capability Programme Challenge Fund (*Challenge Fund*), took a more localized approach with a particular focus on SMEs' access to digitalization. The regional focus of the *Challenge Fund* was delivered through the local enterprises public-private partnerships (LEPs) in England. The 39 LEPs that exist in England are business-led partnerships formed between the local government and the private sector.[9] The objective of the *Challenge Fund* was

[6] https://webarchive.nationalarchives.gov.uk/ukgwa/20200102104956/https://www.greatbusiness.gov.uk/domoreonline/ is the government-backed website. The part of the website dedicated to the digitalization initiative provided short and general information on topics such as building a website and setting up social media and online shops. The charity organisation, digitalskills.com, provided access to digital skills and a forum for SMEs to submit digital-related enquiries. In addition, it offered an interactive map showing locations of digital resources, such as Wi-Fi spots (Al Harbi 2014)

[7] We selected websites of government departments and offices that contained either the words "Industry", "Business", "Innovation", or "Technology" in their title. Full list of government website is accessible here: https://www.gov.uk/government/publications/list-of-gov-uk-domain-names.

[8] Monthly website traffic data are retrieved from: https://www.semrush.com/lp/traffic-analytics-7/en/.

[9] LEPs' main tasks include driving local economic growth, boosting employment, improving infrastructure, and raising local workforce skills by integrating resources from the government, private sector, and local educational institutions. Such partnerships have been created in 2010 by BIS and the Department of Communities and Local Government across England. Originally, the number of LEPs was 39.

to support those LEPs' projects that aim at improving the local SMEs' *awareness* of digital technologies and *transfer of digital skills* to SMEs to enable them to trade and grow online. The target group of the LEPs' projects were SMEs with little or no online presence. When the BIS launched the *Challenge Fund* in September 2014, it invited all LEPs to submit project proposals, of which the winning bids would receive a combined £2 million worth of funding.

Although LEPs submitted their own, individually prepared project proposals, they shared some commonalities, as advised by the BIS (BIS 2015). More specifically, they included a description of initiatives to make SMEs aware of the project (e.g., flyers, local press, and/or local media to promote the projects), and activities to actually involve SMEs in the project (e.g. launch events to create an opportunity for networking and/or online or face-to-face advice, and in depth workshops delivered by ICT experts).[10]

The BIS evaluated each proposal using multiple criteria, such as "[..] value for money; reach; innovative or different ways of doing things; economies of scale; and sustainability." (p.16). Further, since all projects had to be delivered by March 2015, a crucial decision criterion was the project's feasibility to achieve its objectives within a tight timeline.[11]

After a competitive bidding process, in October 2014 the BIS selected 20 projects submitted by 22 LEPs (two were joint bids). The funded LEPs spread across England and covered approximately 60% of England's business population at that time (2015).[12] Figure I shows in red (light blue) the geographical dispersion of LEPs that received the funding (areas that did not receive any funding). The figure indicates that the area of LEPs that received funding is approximately similar to the area of LEPs that did not receive funding in England, suggesting that there was no preference to support either larger or smaller LEPs, on average. However, there appears to be a clustering of

In 2016, the Northamptonshire LEP was merged into the South East Midland LEP, reducing the total number of LEPs to 38. The typical legal forms of an LEP are either a company limited by guarantee or an unincorporated voluntary partnership. The chair of each LEP, as well as the majority of the board, has to be from the private sector (Shearer 2021).

[10] One representative example is the project launched jointly by the Buckinghamshire Thames Valley and Oxfordshire LEPs. The initial phase of the project was the promotion of the project through flyers, radio, newspapers, email, and social media. The local Chambers of Commerce also helped with the publicity of the project. A second phase involved the organization of digital knowledge workshops. The workshops were delivered to local and family-run firms through a local "growth hub". They offered basic digital knowledge with an emphasis on the commercial use of digital technologies. They taught SMEs how digital technologies can facilitate their strategies and planning.

[11] "[..] The bidding process opened in late September 2014 [..]This made the timetable for delivery of the projects very tight and significantly shaped the bids that were made [..]" (BIS 2015, p.16).

[12] We acknowledge a potential confounding event called Digital Gateway launched in Scotland in 2016. To avoid a potential bias, we exclude Scottish SMEs in our sample from 2017 onward.

successful LEPs in South West England as well as in the East Midlands. In the next sub-section, we discuss the characteristics of the treated and control areas in more detail, showing that they are statistically similar to each other.

II.B. Treatment Assignment

In our empirical analysis, we build on the prior that SMEs without internet presence before the program were more likely to implement digitalization in those LEPs areas that received funding (treated) than SMEs located in LEPs in England and the rest of Great Britain (i.e. Wales and Scotland) that didn't receive funding from the initiative (control).

One underlying assumption of our identification strategy is that the allocation of funds across LEPs was exogenously determined. One possible concern, however, could be that certain LEP characteristics, which also determine future SME digitalization, might have made their proposals more likely to be selected for the *Challenge Fund*. For example, instead of merely looking at the quality of the proposals, the BIS may have preferred to select those from LEPs that already had a better technological infrastructure, stronger local economies, and/or a closer local political affiliation with the central government to increase the chances of success of the program. This would imply that the results on digitalization and real outcomes were endogenously determined by the characteristics of the treated LEPs rather than by the increased awareness of digitalization spurred by the *Challenge Fund*.

Table I reports results of univariate tests that help mitigate such concern. It shows that prior to the launch of the *Challenge Fund*, treated and control areas were *not* statistically different from each other in regards to proxies that could possibly predict a higher digitalization rate in those areas independent of the program. In particular, we compare proxies for the internet/digital infrastructure (Panel A), business demographics and economic conditions (Panel B), and political affiliation (Panel C) across all areas included in our sample over the three years before the implementation of the program for a total number of 135 observations. Panel A shows evidence on broadband availability provided by the Office of Communication and compares the average percentage of areas without good internet connection (i.e., those that receive less than 2Mb/s) between treated LEPs and all other areas.[13] It also includes the fraction of internet users and

[13]Data source: https://www.ofcom.org.uk/research-and-data/multi-sector-research/infrastructure-research/connected-nations-2020/data-downloads.

the total number of websites as proxies for the adoption of digital technologies in each area. Differences in means across treated and control areas are statistically insignificant for each proxy.

Panel B reports the averages of several measures of business demographics (i.e., business population, number of business births and deaths, and survival rates of new businesses after one and three years, respectively) and macroeconomic conditions (i.e., total employment, percentage of educated workers, gross domestic products (GDP), growth in GDP, and growth value-added) across treated LEPs and all other areas.[14] Differences across all these dimensions are statistically insignificant.

Finally, Panel C presents the results on political affiliation. We manually collect information on the number of seats in each local authority (i.e., unitary authorities and boroughs councils) in England, Wales, and Scotland assigned to the Conservative party, Labour party, and all other parties. We define a local authority controlled by either the Conservative, the Labour or the other political parties if that party controls the majority of seats. Since at the time of the *Challenge Fund* the Conservative Party was the governing party in the U.K., we measure the political affiliation of each area by computing the ratio of the number of local authorities controlled by the Conservative Party to the number of local authorities controlled by all other parties in that area. For England, we use the geographical areas based on LEPs while for Wales and Scotland we use the Nomenclature of Territorial Units for Statistics 2 (NUTS-2).[15] Differences in the averages between treated and control areas are again statistically insignificant.

(Insert Table I here)

Overall, our tests do not provide evidence that treated LEPs were selected because of inherently different characteristics that would predict the success of the program and invalidate our identification strategy.

[14]The percentage of educated workers follows the definition of the Office for National Statistics (ONS 2021) as the number of workers aged between 16 and 63 with a qualification that is at least in National Vocational Qualification (NVQ) 4 level or equivalent divided by the total workforce. NVQ levels are work-based qualifications that can be achieved through assessment or training. An example of qualification in NVQ 4 level is the bachelor degrees.

[15]NUTS is a geocode classification for referencing the subdivisions of the U.K. as well as those of the European Union for the application of regional policies (Eurostat 2022). It is used by the ONS. NUTS-2 level corresponds to 40 basic regions such as Lancashire, Oxfordshire, Kent, and Cornwall, with the same level of granularity as the one based on LEPs.

III. DATA AND METHODOLOGY

III.A. Sample Construction

To conduct our empirical analysis, we use two main databases. One is the Financial Analysis Made Easy (FAME) database, the U.K.-based product of Bureau Van Dijk, that provides detailed firmographic data, financial statements, and directors' information for a very large sample of U.K. public and private firms. The other main database is BuiltWith, which collects data on the web technologies used on firms' websites by using website-crawlers to identify and track all websites since 2000. It provides a comprehensive coverage with a very detailed and unique panel data of websites' digital technologies. Our sample period covers the years 2011-2019 spanning around the launch of the *Challenge Fund* in 2014.

The starting point of our sample construction is FAME. We select firms headquartered in England, Wales, and Scotland with at least one non-missing record of total assets before the launch of the *Challenge Fund*. We further require firms to have all their establishments (trading addresses) either within the treated or the control areas to reduce the noise. We exclude firms from the utility, financial, and public administration industries.[16] This yields a sample of 304,745 unique firms. We further require firms to have non-missing employment and turnover data in 2013, the year prior the launch of the *Challenge Fund* and to have a leverage ratio between 0 and 1. This restricts the sample to 64,451 firms, of which we identify 58,356 SMEs and 6,095 non-SMEs (large firms).

The target of the *Challenge Fund* was SMEs with no prior digital knowledge. We use the non-existence of a website prior to the launch of the *Challenge Fund* as a proxy for the lack of digital knowledge. We rely on both the website address information provided by FAME and the date from BuiltWith when the website (if existent) is detected for the first time to construct such sample. In particular, we include: 1) companies without any record of a website over the entire sample period in both FAME and BuiltWith; and 2) companies with a recorded website at some point after the launch of the *Challenge Fund*.[17] This leaves us with a preliminary sample of 27,696 SMEs without website before

[16] They correspond to U.K. SIC2007 64-66, 35-39, and 84, respectively.

[17] We verify FAME's accuracy in recording missing websites by manually checking the top 100 largest firms by total assets without website information in FAME. While, according to Google Search, only 9 of them seem to have an actual website, none of them has verifiable information, such as a business address or contact details which can link the websites to the firms. This suggests that the information provided by FAME is indeed accurate.

the start of the program. We use this sample to implement a propensity score matching (PSM) of which more details are provided below. After the matching, the final sample includes 8,799 treated and an equal number of matched control firms.

Treated firms operate in 17 different sectors: 12.22% of the sample is from manufacturing, followed by wholesale and retail trade (10.57%), and human health and social work activities (10.20%).

III.B. Digital Technology Variables

We employ a novel source of digital technology data that are available at firm level, and cover the business population across all size groups. The data consist of the web technologies that BuiltWith retrieves from each firm's website. Such technologies include, among others, the type of web hosting the website uses, its e-commerce functionalities, and its analytics softwares.[18] Such technologies can change over time depending on how sophisticated the firm's website becomes. As such, BuiltWith's database is a very detailed and unique panel data of websites' digital technologies.

BuiltWith's website crawler works in a similar way as Google Search by crawling and indexing websites.[19] Their targeted coverage is the entire internet, comprising over 16.4 billion websites in 273 countries globally. For the U.K., BuiltWith provides records for approximately 19 million active and inactive U.K. domains. As a comparison, the number of currently active domains in the U.K. is approximately 11 million (Nominet 2021). BuiltWith updates its records either weekly, bi-weekly, or quarterly depending on the activity of the website.[20]. Every time BuiltWith updates its website records, it captures any new technology that a website may have adopted or dropped, and the date of the record update. We use the dates of these regular within-year updates to create a

[18] When a hosting provider allocates space on a web server for a website to store its files, they are hosting a website. Web hosting makes the files that comprise a website (e.g., code, images) available for viewing online. The amount of space allocated on a server to a website depends on the type of hosting. E-commerce technologies enable a firm to set up an online store where its products or services catalog are displayed. Customers can browse, select the product from the online shop, pay for it, and, ultimately, arrange the shipping of such product. Web analytic softwares (e.g. Google Analytic) are tools that can collect and report web traffic data to improve the users experience and website's performance. It can help firms better understand their customers and refine their marketing strategies.

[19] The crawler robot automatically finds websites from the internet and downloads information from each website ("indexing"). BuiltWith targets to index all domains that end in '.com', '.net' and '.org' and any active domains in ccTLD and gTLD in all regions (BuiltWith 2022).

[20] Records are updated weekly for websites with positive technology spending, which is estimated by BuiltWith, bi-weekly for websites with high traffic (identified using ranks such as 'Google top 10k sites'), and monthly for active websites. All remaining websites are updated quarterly (2022)

snapshot of firms' website technologies at the end of their fiscal year, which enables us to analyse the technology usage over time for each firm that has a website.

BuiltWith provides a detailed description of each tracked technology, a classification (tag) of the technologies according to their functions, and the date when each technology was first captured from a firm's website for the first time. Between 2000 and 2022, BuiltWith has identified a total of 375,183,792 technologies from all U.K. websites, grouped into 33 tags. In our sample, there are a total number of 602,112 firm-technologies observations that are based on 4,310 unique technologies and 27 tags.

We exploit the richness of this database to construct several measures of digitalization. Our first measure is a dummy variable equal to one if a firm has a website in a given year, and zero otherwise (*Web*). This captures the extensive margin of digitalization.

We use BuiltWith's technology tags to construct a number of measures that capture the intensive margin of digitalization. We manually classify all tags into *General* and *Business* technologies. General technologies are essential for either the construction and existence of any website (e.g. content delivery network, web servers, frameworks and registrar), to enable certain website functions (e.g. Content Management System), or to enhance the accessibility of the website without requiring any business-related knowledge. We label them "general" because they are not specific to a business website. In contrast, business technologies are applicable only to businesses. The existence of business website technologies changes how the business operates and requires employees to be trained to understand how to use these technologies. Based on prior studies that document the impact of such technologies on businesses (see Goldfarb and Tucker (2019) for a review), we classify the following as business technologies: online advertising; analytic and tracking that enables digital marketing; audio and video media; e-commerce and mobile commerce technologies; email; feeding content; language; and widgets that include social media sharing.[21] For each firm each year, we sum up all general and business tags to create a *G-score* and *B-score* variable, respectively, and use the sum of both *G-score* and *B-score* to capture all tags detected in each year (*BG-score*).

Further, within the business tags, several technologies are related to e-commerce, such as online shops, payment, shipping providers, and online transaction security. In our empirical analysis, we investigate such e-commerce technologies as a separate specific dimension of digitalization that requires a certain level of digital training, and which is likely to impact the firm's boundary. As a measure of the extensive margin of the

[21] Table A.I shows all technology tags included in our sample and their classifications.

enlarged firm's boundary, we construct a binary indicator, *E-commerce*, equal to one if a firm has adopted any of the e-commerce-related technologies in a given year, and zero otherwise. We also construct a measure of the intensive margin of the e-commerce technologies, *E-score*, which is the sum of all tags associated with e-commerce.

III.C. Firm-level Financial Data

Our main source of financial and firmographic data is FAME. We use the growth rate of annual turnover as a measure of revenue growth (*Sales Growth*) and the ratio of EBIT to total assets as measure of *ROA*. The employment-related outcome variables are the growth of firms' employment and labor productivity, measured as the change between year t and t-1 of the natural logarithm of the number of employees (Δ*Ln(Employees)*) and the change between year t and t-1 of the natural logarithm of one plus the firm's earnings before interest, tax, depreciation and amortisation (EBITDA) scaled by the number of employees (Δ*Ln(VPE)*), respectively.

We further collect information on the legal status of the companies to identify bankrupt companies, which we use in the survival analysis. The firm is defined as bankrupt from year t onward if its legal status at time t is reported as either 'in liquidation', or 'dissolved'.

Finally, we construct the following variables to perform the PSM: *Ln(Total Assets)* as proxy for firm size measured as the natural logarithm of one plus the total assets; *Ln(Age)* measured as the natural logarithm of one plus the number of years from the year of incorporation; *Leverage* measured as the sum of short-term debt and long-term liability divided by total assets; *Cash* measured as cash divided by total assets; and *ROA*. All variables are winsorized at the 1% level to remove outliers. Table A.II reports the definitions of all variables.

III.D. Methodology

Our identification strategy exploits the geographical variation of the allocation of funding by the *Challenge Fund* program across Great Britain, and it focuses on the companies targeted by the program, i.e. SMEs with no prior digital experience. We define a company as treated (control) if it is an SME without a website before the start of the program in 2014 and located in an area that received (did not receive) funding from the *Challenge Fund*.[22]

Given the previously observed geographical clustering of some of the treated LEPs in South West England and in the East Midlands, the composition of firms in the treated and control areas may be different despite the geographical areas being statistically similar, as Table I shows. Previous literature has found evidence suggesting that the industry composition and firm characteristics vary across regions within the U.K., due to either the different scales and types of foreign investments (Dicken and Lloyd 1976) or the uneven distribution of human capital and financial resources (Gal and Egeland 2018).

To avoid that observable firm characteristics may drive the results and bias the effect of the program on firm digitalization, we employ the propensity score matched difference-in-difference (PSM-DID) estimator that enables us to better identify the causal impact of the *Challenge Fund* on firms' digitalization, investment, and performance metrics. We first find the closest matched control SME for every treated SME in the year before the program, 2013, without replacement and using a caliper of 0.001. When estimating propensity scores for treated and control groups, we include firm characteristics that capture the propensity to innovate, potential financial constraints, growth opportunities, and profitability, namely $Ln(Total\ Assets)$, $Ln(Age)$, *Leverage*, *Cash*, *Sales Growth*, and *ROA*, along with 1-digit U.K.-SIC code dummies. Panel A of Table II shows that without matching, treated and control firms are statistically different across all these dimensions. Panel B confirms that after the propensity score matching, treated and control firms are statistically indistinguishable from each other. The final matched sample includes 8,799 treated SMEs.[23]

(Insert Table II here)

Next, using this pairwise matched sample we implement the DID estimation as follows:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t}, \tag{1}$$

where $Y_{i,t}$ represents either the digitalization outcomes (*Web*, *Ln(BG-score)*, *Ln(G-score)*, and *Ln(B-score)*), firms' e-commerce investment proxies (*E-commerce* and *Ln(E-score)*), performance outcomes (*Sales Growth* and *ROA*), and labor market outcomes (Δ*Ln(Employees)* and Δ*Ln(VPE)*) as defined above. $Post_t$ is a dummy variable equal to one (zero) during the 2015-2019 (2011-2014) period. $Treated_i$ equals one (zero) for

[22]The outcome of our analysis captures both the direct effect of the intervention (on companies that attended the workshops) and the indirect one (on those companies that digitalized through e.g. word-of-mouth).

[23]In untabulated regressions, we found that our baseline results hold even without matching.

SMEs without a website before the program and located in treated (control) areas. $Z_{i,t,j,g}$ is the matrix of fixed effects that includes firm (i), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on the 1-digit U.K.-SIC code. The geographic fixed effect is defined at the NUTS-2 level. Standard errors are clustered at the geographic-year level. The coefficient of our interest is β which captures the effect of the *Challenge Fund* program on both SMEs' digitalization and real outcomes.[24]

IV. DIGITALIZATION OUTCOMES

IV.A. Univariate Results

Since our identification strategy requires all sample SMEs not to have a website before the program start, their website-based digital scores are zero in the pre-program period. This leads to satisfying the parallel trend condition as both treated and control firms are indistinguishable in terms of their digitalization level before the program.

Table A.III in the Appendix reports the descriptive statistics of the digitalization measures after the launch of the program in both absolute (Panel A) and log-transformed values (Panel B). It shows that differences between treated and control groups are highly significant. After the program, treated SMEs are, on average, 4 percentage points more likely to set up a website compared with control SMEs. Further, results on *BG-score* suggest that treated SMEs adopt, on average, more than 0.4 different types of technologies, which is almost 19% higher than the average number of technologies of the control group (2.3). Similarly, result on *B-score* and *G-score* imply 19% and 17% more of business and general technologies, respectively. As for the e-commerce technologies, the treated group seems almost 3 percentage points more likely to adopt any technology that enables online selling than the other group. On average, the number of e-commerce technologies adopted by the treated group is 21% higher than the control group. This univariate analysis points to a positive association between the *Challenge Fund* program and SMEs' digitalization, which motivates the further investigation of a potential causal relationship by estimating the PSM-DID models from equation (1).

[24] Using a log-linear transformation can potentially lead to bias (Cohn, Liu, and Wardlaw 2022). Alternatively, we estimate a zero-inflated Poisson model that accounts for the larger number of zero values in our pre-treatment period by dropping firm and year fixed effects. The estimated coefficients have the same sign, and similar magnitude and statistical significance.

IV.B. Digital Adoption

We start our analysis by investigating the effect of the *Challenge Fund* on digitalization outcomes both at the extensive and intensive margins. Table III reports the baseline results from the PSM-DID estimators. Columns (1) to (3) report the digital outcome on the extensive margin, *Web*, with three different fixed effects specifications. Across all models, the estimated coefficients are positive and statistically significant at the 1% level. For example, in the specification with firm and year fixed effects (column 1), treated SMEs are 3 percentage points more likely to have a website after the *Challenge Fund* compared to the matched SMEs in the control areas. Even after controlling for potential time-varying industry shocks or area-specific shocks that may confound with the effect of *Challenge Fund*, the results are still strongly significant.

Columns (4) to (12) report results on the intensive margin. In particular, results on *Ln(BG-score)* in columns (4) to (6) suggest that the website set-up by treated SMEs involves a wider range of technologies than that of all other SMEs. Treated SMEs employ 8% more different technologies in their websites relative to the control firms (column (1)). We further delve into the types of the newly adopted technologies to analyse whether the *Challenge Fund* had any relevance for the business activities. To this end, we distinguish between business-related and general technologies classes. Columns (7) to (9) and columns (10) to (12) report the estimated coefficients for *Ln(B-score)* and *Ln(G-score)*, respectively. The increase in the adoption of web technologies is significant for both classes across all specifications: treated SMEs seem to adopt about 6% more business-related and general technologies than the control SMEs.

The economic magnitude of the treatment effect is not trivial. Since our identification strategy requires SMEs not to have a website in the period before the *Challenge Fund*, i.e., all of our web-related variables are zero for both treated and control firms during that period, we use the descriptive statistics of digitalization outcomes after the *Challenge Fund* (Table A.III). Since the unconditional probability for the control SMEs to have a website after the program is 23% (as reported in Panel A Table A.III), the estimated DID coefficient of 0.03 in Table III implies that the average treated SME is approximately 13% more likely to have a web presence relative to the control SMEs (0.030/0.234*100 = 12.8%). We find similar results with the other digital outcome variables that measure the intensive margin. In particular, the estimated *Ln(B-score)*, 0.057, implies that treated SMEs employ, on average, 16% more different technologies related to business in their websites compared with the controls (0.057/0.358*100 = 15.9%, where

0.358 is the average *Ln(B-score)* of the control sample as reported in Panel B of Table A.III).

Overall, our findings show that the websites of treated SMEs have become more sophisticated than those of the control SMEs after the program launch, suggesting that the *Challenge Fund* had a significant impact on enhancing the digital knowledge of SMEs, including the one more likely to affect their business activities. Further, compared to previous studies (e.g., Bronzini and Iachini 2014; Akerman, Gaarder, and Mogstad 2015; Howell 2017; Bloom, Van Reenen, and Williams 2019), this evidence implies that a government program focused on increasing awareness and training in digital skills, rather than on direct subsidies to companies, can be effective in making SMEs become more digitalized.

(Insert Table III here)

IV.C. E-commerce Investment

One important aspect of digitalization is the adoption of e-commerce technology. Previous studies show that e-commerce adoption has positive impacts on firm's productivity, innovation, connection to the international market (e.g., Bertschek, Fryges, and Kaiser 2006; McElheran 2015; Cassetta et al. 2020), and on SMEs resilience to adverse shocks (e.g., Cong, Yang, and Zhang 2022). Since e-commerce is enabled by website-based technologies and results in the baseline regressions show that business-related technologies have been positively affected by the *Challenge Fund*, we use the e-commerce-related technologies from the BuiltWith database to investigate whether the improvement in digital skills of SMEs has also spurred the adoption of e-commerce.

Table IV reports the results of SMEs' use of e-commerce technologies, estimated on both the extensive (columns (1) to (3)) and intensive margins (columns (4) to (6)). Column (1) shows, for instance, that treated SMEs are, on average, 3 percentage points more likely to adopt e-commerce technologies on their website than the control SMEs. In economic terms, that means the average treated SME is approximately 17% more likely to expand its boundaries with e-commerce relative to the control SMEs (0.028/0.168*100 = 17%). On the intensive margin, the DID coefficient 0.024 suggests that treated SMEs adopt 2% more e-commerce technologies, equivalent to an increase of about 17% (0.024/0.143*100 = 17%) relative to the control firms in the aftermath of the program. This supports our conjecture that SMEs with more digital knowledge are more likely to exploit digital opportunities for their businesses by facilitating the adoption of e-business technologies and, in turn, expand the firm's boundaries.

(Insert Table IV here)

IV.D. The Digital Divide

Over the past decade, there has been an intense debate about the causes of and remedies for the observed high adoption rate of digital technologies among large companies and their relatively low usage among SMEs, often dubbed the "digital divide", which characterizes economies worldwide. A digital divide has also been found at the spatial level, showing that firms in rural areas lag behind in adopting advanced digital technologies compared to firms in urban areas (e.g., Thonipara et al. 2022).

A common view is that the digital divide represents a lost opportunity for SMEs to grow and possibly to become more competitive, which has negative consequences for the number of available jobs in and the wealth of local economies. In the U.K., a survey carried out by the Office for National Statistics (ONS) before the *Challenge Fund* shows that SMEs adopted strikingly fewer digital technologies and have made less e-commerce sales than large corporations (ONS 2014).

Previous evidence across the U.S. and Europe suggests that the major cause of the digital divide between large corporations and SMEs is the lack of digital knowledge and skills (Arendt 2008) that becomes exacerbated in entrepreneurial firms and micro businesses (Millán et al. 2021), and in the presence of adverse shocks such as the recent Covid-19 pandemic (Willcocks 2020). Similarly, one factor that the literature has identified to help explain the spatial digital divide is at the socio-demographic level, namely the human capital differences between rural and urban areas (e.g., Billon, Lera-Lopez, and Marco 2016; Thonipara et al. 2022). A suggested solution to eliminate such digital divide has been to focus the allocation of government resources particularly on the provision of training and education (Wielicki and Arendt 2010).

In this section, we investigate whether the increased digitalization spurred by the *Challenge Fund* program has enabled SMEs to catch-up and narrow the digital divide with large companies as well as the digital divide between rural and urban areas.

We start the analysis by looking at the digital divide between large corporations and SMEs. We implement the following catching-up model:

$$Gap_{i,t} = \delta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t}, \tag{2}$$

where $Gap_{i,t}$ is defined as:

$$Gap_{i,t} = \overline{Y^l_{g,t}} - Y_{i,t}, \tag{3}$$

where $\overline{Y^l_{g,t}}$ is the average digitalization outcome across large firms (l) located in the same area g (LEP or NUTS-2 level areas) of SME i at time t and $Y_{i,t}$ is the digitalization outcome of SME i at time t. We define large firms as companies with more than 250 employees and an annual turnover of more than €50 million. Digitalization outcomes include: *BG-score*, *B-score*, *G-score*, and *E-score*.[25] The coefficient of interest is δ that captures the effect of the program in closing the gap between large companies and SMEs. We expect the coefficient to be negative if the increased digitalization among SMEs has actually reduced the digital gap between them and large companies more in the treated areas than the control ones.

Table V reports the findings of the catching-up analysis. In terms of the number of newly adopted technologies (*BG-score*), the results in columns (1)-(3) show that in the post-program period, the distance between large firms and treated SMEs relative to the distance between large firms and control SMEs is significantly smaller by about 0.40 (column (1)). Given that the average gap between large firms and control SMEs in the post-program period is 4.13, it corresponds to a reduction of approximately 10% (0.40/4.13*100 = 0.097%).

Similarly, the results show negative and statistically significant treatment effects on the gaps between SMEs and large firms for both business-related (columns 4-6) and general technologies (columns 7-9) The effect is similar across both types of technologies ranging between -0.19 to -0.21 across all specifications.

Finally, columns 10-12 of Table V report the results of catching-up on *E-score*. Compared with the controls, we find that the gap in number of adopted e-commerce technologies between large firms and treated SMEs is significantly smaller. For instance, the gap reduces by 0.05 more for treated SMEs compared to the control ones (column 10).

Overall, the evidence suggests that the program was also successful in helping narrow

[25] Due to the construction of the dependent variable, *Gap*, we are able to test the catching-up hypothesis only on the intensive margin dimension of digital outcomes.

the digital divide between large firms and SMEs.

(Insert Table V here)

Next, we investigate whether increasing digital awareness and training through the *Challenge Fund* program also helped narrow the geographical digital divide.

We split the sample along two alternative definitions of spatial divide. First, we sort our SMEs sample into urban and rural subsamples based on the address of their primary trading location. Urban (rural) areas are defined as the areas inside (outside) the largest ten cities in the U.K. by the 2011 Census population.[26] For the second definition of spatial divide we split the SMEs sample into areas where the top 20 universities are located versus all other areas. Previous studies show that in areas with a higher share of college education, skill-intensive technologies are adopted more intensely (Beaudry, Doms, and Lewis 2010). As such, areas without a top university may face greater skills barriers to adopt new technologies and hence may benefit more from the *Challenge Fund*.[27] We then test our baseline digitalization models as in equation (1) across these subsamples to see whether SMEs in rural areas or areas farther away from top universities have benefited from the program as compared to those in urban areas or areas where the top universities are located.

Table VI reports that in both urban (Panel A) and rural (Panel B) areas treated SMEs have become significantly more digitalized at the extensive (2% and 3%, respectively) and intensive margin (3%-4% and 5%-7%, respectively). Similarly, in universities areas (Panel C) and all other areas (Panel D), treated SMEs have become more digitalized at the extensive (3% and 2%, respectively) and intensive margin (4%-6%). The differences between the estimated DID coefficients across the two pairs of subsamples are statistically insignificant from each other. While the findings do not provide direct evidence of a reduction in the digital spatial divide, the results indicate that the reach of the program went beyond the urban areas and those with a higher level of education, making access to digitalization possible in more remote geographical locations as well.[28]

[26]The ten cities, from the largest to the smallest, are: London, Birmingham, Leeds, Sheffield, Bradford, Manchester, Edinburgh, Liverpool, Bristol, and Cardiff (https://www.ons.gov.uk/peoplepopulationandcommunity/ populationandmigration/ internationalmigration/datasets/populationoftheunitedkingdombycountryofbirthandnationality).

[27]To define the top 20 universities, we use the average of the overall FTE results across all units of assessment submitted by each university for the Research Excellence Framework in 2014 (REF2014.). The top 20 universities areas are based in: London, Oxford, Cranfield, Cambridge, Edinburgh, Glasgow, Manchester, Nottingham, Southampton, and Warwick.

[28]We also use an alternative classification of urban/rural areas created by the U.K. government based on the fraction of the population residing in the rural areas of England

(Insert Table VI here)

V. REAL OUTCOMES

V.A. Performance

After having established the causal relation between the implementation of the *Challenge Fund* program and the digitalization of SMEs, in this section we investigate whether the effect of the program also carried through to real outcomes. In particular, since the results on digitalisation indicate that the *Challenge Fund* had a positive effect on e-commerce, here, we examine whether the overall increase in digitalization led to an increase in revenue (*Sales Growth*) and profitability (*ROA*).

Table VII Panel A shows that in the aftermath of the program, sales of treated SMEs grew approximately 1% faster than those of control SMEs. Economically, this is a significant increase compared to the 5.9% average growth rate of sales of all SMEs in our sample during the pre-program period (Panel A of Table A.IV). Further, Panel B shows that treated firms were more able to efficiently convert sales into profits with their average *ROA* 0.9% higher compared to the control group.

One way higher digitalization, and e-commerce in particular, translates into more sales could be an increase in website traffic. As such, we expect that the websites of treated SMEs benefited more from a higher traffic following the *Challenge Fund* compared to those of control firms.

To confirm such channel, we retrieve the website traffic data from SEMrush, which collects minute-based clickstream data and Google search results of the 500 million most popular keywords from third-party data providers. [29] We define the total web traffic as the yearly sum of both organic and paid traffic. We then construct the variable *Website Traffic Growth* as the annual growth rate of the total web traffic.

(https://www.gov.uk/government/collections/rural-urban-classification). We compare urban areas with 'Minor and Major Conurbation' versus all other areas. Although such classification is available only for England, results are similar to what we report here.

[29] SEMrush uses neural network algorithms to estimate both the monthly websites' organic traffic generated from Google searches and the traffic generated from paid Google Ads. Due to data access limitations, we retrieve both organic and paid traffic estimated in all odd months starting with January for all websites of our PSM sample across the sample period. If a month is missing, we use first the information of the month before the missing month, and, if that is also missing, we use the information of the month after the missing one. We believe that this approach should not affect our conclusions.

Table A.V in the Appendix shows that the growth rate of website traffic of treated SMEs is higher than that of the control group by 86%. This difference in website traffic growth is also economically significant when compared with the 514% average website traffic growth rate of our sample firms during the post-program period (see Panel B of Table A.IV).[30] The results presented here add to the evidence provided by (Armstrong, Konchitchki, and Zhang 2023) who document the economic implications of firms' digital traffic for financial performance, among others.

(Insert Table VII here)

V.B. Employment and Productivity

There is a long-standing debate about the effects of technology adoption on the labor market. On one hand, new technologies replace labor through automation with a negative impact on labor demand. On the other hand, the creation of new tasks induced by the adoption of new technologies increases the demand for labor (e.g., Brynjolfsson and McAfee 2014; Acemoğlu and Restrepo 2018) and new technologies can make workers more productive. The empirical evidence on how digital technology adoption affects employment and labor productivity remains limited. Most of the studies rely on either country or industry-level, or survey data with mixed results (e.g., Koellinger 2008; Evangelista, Guerrieri, and Meliciani 2014; Biagi and Falk 2017). Our empirical setting allows us to build on previous findings and test the causal relation between digitalization and labor market outcomes at the firm level.

Table VIII reports the results of the impact of digitalization on firms' growth of employment ($\Delta Ln(Employees)$) in Panel A and growth of labor productivity $\Delta Ln(VPE)$ in Panel B, respectively.

The rate of employment growth of treated SMEs relative to the control firms is approximately 0.9% higher and statistically significant across all specifications. The economic magnitude of this effect is not trivial, considering that the pre-treatment average growth rate of all SMEs' employment in our sample is less than 1% (Panel A of Table A.IV).[31]

[30] The conversion from website traffic to actual sales turnover appears relatively low. Nonetheless, this seems to be in line with the average global e-commerce conversion rate, which in 2022 was approximately 1.8% (https://www.invespcro.com/blog/the-average-website-conversion-rate-by-industry).

[31] In unreported results, we do not find any apparent pre-trends in either employment or productivity growth.

In line with the growth rate of employment, results in Panel B report that employees of treated SMEs, on average, became also more productive than those of the control group after the program. The treated SMEs value-added per employee rate is at least 3% higher than that of the control SMEs across all regression specifications. The economic magnitude is again significant given that the average growth rate of labor productivity of all SMEs during the pre-program period was approximately -1.4% (Panel A of Table A.IV).

Overall, the evidence suggests that adopting digital technologies fosters SMEs' employment growth and, at the same time, increases labor productivity.

(Insert Table VIII here)

V.C. Survival

Finally, we investigate whether the increased digital awareness and training have reduced the probability of bankruptcy. We estimate a Cox-Proportional Hazard model as follows:

$$\lambda_{it} = \phi_t \; exp(\theta \; Treated_i + \gamma(Post_t \times Treated_i)) \tag{4}$$

where λ_{it} indicates the hazard rate, i.e., the probability of firm i being dissolved or in liquidation at time t, conditional on having survived up to that time. ϕ_t represents the baseline hazard rate. We estimate λ_{it} using the Cox-proportional Hazard model over our sample period from 2011 to 2019. In a further test, we estimate the hazard model over an extended sample period that includes also the first two years of the Covid-19 period to explore whether digitalization helps firms become more resilient to severe shocks such as the recent pandemic (Cong, Yang, and Zhang 2022).

Table IX reports the estimated results of the Cox-proportional Hazard model. Column (1) shows a post-treatment coefficient of -0.29, significant at the 10% level. This corresponds to the hazard ratio of 0.75 ($e^{-0.289}$). It implies that treated SMEs were 25% less likely to be dissolved than control ones. We find a similar coefficient in column (2), -0.30, when we include the pandemic years in the sample period. It indicates that treated SMEs seem to be more resilient than their counterparts to a shock as severe as the recent pandemic. This result complements Cong, Yang, and Zhang (2022)'s evidence that shows a positive growth of SMEs online sales thanks to the adoption of e-commerce during COVID, that made them more likely to reopen during and after the lockdown and to hold a more optimistic view of future growth.

(Insert Table IX here)

Overall, the results suggest that SMEs benefited from greater digitalization in terms of higher performance, increased employment and labor productivity, and, ultimately, a higher survival rate with greater resilience to adverse shocks.

VI. ROBUSTNESS TESTS

VI.A. Parallel Trend Condition, Placebo Tests, and Spillover Effect

To corroborate the validity of our identification strategy we perform four different tests. First, we confirm that the parallel trend condition is met. By construction, both treated and control firms in our sample are not digitalized before the program. As such, there is no trend within the two groups before the implementation of the *Challenge Fund* program. Nonetheless, we present a dynamic analysis around the launch of the program by estimating a year-by-year treatment effects model where we augment the baseline specification in equation (1) with yearly dummies using 2011 as the base year. Figure A.I in the Appendix shows the estimated coefficients of the PSM-DID yearly interactions. Panels A to D show results for *Web*, *Ln(BG-score)*, *Ln(B-score)*, and *Ln(G-score)*, respectively. Consistent across all four digitalization outcomes, the treatment effects are nearly close to zero before the program and become positive and significant after the treatment in 2014.

Interestingly, the magnitude of the treatment effect increases over time. At the end of the sample period (2019), the result on the extensive margin of digitalization, *Web*, suggests that treated SMEs are, on average, 17% more likely to have a website than control SMEs; while in the year right after the launch of the program, 2015, this difference is about 7%. Similarly, the difference between the total number of technology tags used by treated and control SMEs (*Ln(BG-score)*) increases from approximately 3% in 2015 to 8% in 2019. The same patterns are also observed for *Ln(B-score)* and *Ln(G-score)*, suggesting that the benefits from increased awareness and knowledge of new digital skills tend to build gradually over time. [32]

[32] We also perform a year-by-year PSM-DID estimation of the catching-up analysis presented in Table V. Figure A.II in the Appendix shows the estimated coefficients of the treatment effects for *Gap(BG-score)*, *Gap(B-score)*, *Gap(G-score)* and *Gap(E-score)*, respectively. Across all measures, the gaps between the SMEs and their benchmarks tend to be around zero and statistically insignificant during the pre-program period. Since both treated and control SMEs have zero digitalization scores before the program, the variations we observe in that period are attributable to the digitalization of large firms

To further validate our identification strategy, we perform two placebo tests. First, we generate a random sample of the same size as our pre-PSM sample from all SMEs in FAME. We assign randomly the treatment to half of those SMEs and we perform a similar matching as in our baseline tests. We re-run the baseline models as in Table III. Panel A of Table A.VI in the Appendix reports the results.

Second, we construct a new sample with 2010 as the placebo event year, requiring firms not to have a website in the years 2009 and 2010 before the presumed treatment. Years 2011 and 2012 are defined as the post-event period. We then re-run the same baseline specifications using this identification strategy. Panel B shows the results from these tests. As expected, across all tests in both panels the estimated coefficients are statistically insignificant confirming the validity of our identification strategy and so the causal relation between the implementation of the *Challenge Fund* program and the digitalization of the SMEs.

Finally, since our identification strategy relies also on geographical boundaries, we want to verify whether the baseline results are affected by a policy spillover effect (e.g., Jardim et al. 2022). As described in Section 2, the way of promoting digitalization by the funded LEPs was highly localized. As such, firms outside the treated areas were unlikely to be even aware of such initiatives. To further corroborate this conjecture, we re-estimate our baseline digitalization models using a sample where the control firms are those located relatively close to the borders of areas covered by the treated LEPs. If the *Challenge Fund* program also promoted digitalization for the SMEs in the neighbourhood areas, we should detect either a small or no significant difference between the treated and the new control SMEs. Table A.VII in the Appendix reports results with control firms located within 5 miles from the borders of the treated LEPs. Across all specifications, the estimations are positive and significant, suggesting that the *Challenge Fund* program had no significant spillover effects on the neighbourhood areas.[33]

in the treated and control areas. It suggests that large firms across all areas have a similar degree of digitalization and, as such, they can be used as benchmark to construct the *Gap* outcome variables. They also confirm that the parallel trend conditions for *Gap* outcomes are satisfied. Negative and significant gaps are detected in the post-program periods only, suggesting that after the program the digital divide narrowed more for the treated group than the controls. Further, such gaps narrow more and more over time, consistent with the evolution of the digitalization outcomes we observe in Figure A.I.

[33] Untabulated tests show that results are qualitatively similar to those reported here when we use the 10 miles cut-off to calculate the distance from the borders.

VI.B. IT Investment

Our previous results show that the heightened awareness of digital opportunities leads to an increased digital presence of SMEs through the creation of (more sophisticated) websites (Table III) coupled with an increased labor demand (Table VIII). Based on findings from studies on the effects of skill-based technological development, we would expect that the increased digitalization of SMEs should particularly affect the demand for *skilled* workers and be associated with increased IT spending (e.g., Autor, Levy, and Murnane 2003).

Since IT-related disclosures are not required for SMEs, we use data provided by Spiceworks Ziff Davis' Aberdeen Technology Data Cloud (ATDC, previously known as Ci Technology Database, CiTDB), which collects establishment-level IT information through monthly surveys and provides estimated IT-related variables.[34] We were able to obtain data for the years 2010, 2011, 2012, 2016, and 2018. ATDC covers firms across all industries and sizes. For our setting, we use information on the number of IT workers (*IT Staff*) and IT budget (*IT Expenses*) that includes hardware-, software-, PCs-, servers and communication-related expenses.

SMEs are defined following the criteria of total number of employees and total revenue across all establishments available in the database, similar to the SME definition we use with the firm-level data. Since we don't have information of the website at the establishment level, we exploit the geographical variation of the areas that received funding from the *Challenge Fund* program to define an alternative identification strategy. In particular, treated (control) establishments are those located within the treated (control) areas. Similar to the baseline models, we employ the propensity score matched difference-in-differences (PSM-DID) estimator where we find the closest matched control establishment for every treated establishment in 2012 based on revenue, employment, and industry. We then estimate our baseline regression model from equation (1) at the establishment-level, as follows:

$$Y_{e,i,t} = \beta(Post_t \times Treated_e) + \mathbf{Z_{i,t,j,g}} + \epsilon_{e,i,t}, \quad (5)$$

[34] These data have been used by several studies both in the U.S. and the U.K. (e.g., Bresnahan, Brynjolfsson, and Hitt 2002; Bloom, Sadun, and Van Reenen 2012; Campello, Gao, and Xu 2023).

where $Y_{e,i,t}$ includes the IT-related outcomes for establishment e of firm i at time t. The IT-related outcomes are the log-transformed variables with one plus the original values ($Ln(IT\ Staff)$ and $Ln(IT\ Expenses)$).[35] *Post* is a dummy equal to 1 (zero) in years 2016 and 2018 (2010, 2011, and 2012). *Treated* equals one (zero) for establishments of SMEs located in the *Treated* (*Control*) areas. $\boldsymbol{Z}$ is the matrix of fixed effects, including firm (i) to exploit variations across establishments within the same firm, year (t), industry-by-year (j), and geographic (g) fixed effects. Standard errors are clustered at the geographic-year level.

Table X shows that, regardless of the specification, there is an increase in both IT workers and IT expenses in the post-*Challenge Fund* period. Relative to the control establishments, IT workers increased, on average, by 6% in treated establishments (column (1)); while the expenses for IT items increased on average by about 12% more (column (4)).

Overall, these results offer further support to the effectiveness of the program in digitalizing SMEs and, also, provide external validity of our previous results at the firm level.

(Insert Table X here)

VI.C. Barriers to Digitalization

Besides the awareness of new technologies, and, more generally, digital skills (information and skills channel), firms' technology spending and, hence, their digitalization, may also be shaped by the firm's ability to access finance (e.g., Campello, Graham, and Harvey 2010) and the propensity of managers to innovate, as indicated by survey evidence (e.g., Arendt 2008; Millán et al. 2021; OECD 2021a).

As for the financial constraints dimension, we argue that our empirical setting is particularly suited to capture mostly the information and skills channel rather than the financial constraints one for two main reasons. The first one pertains to the specific characteristics of the *Challenge Fund*. The program did not provide any direct financial support to SMEs. In other words, treated SMEs were not immediately financially better off just by being exposed to the program. Instead, the program facilitated the creation of digital skills among SMEs by promoting in depth workshops and networking events via the LEPs. The second reason relates to the nature of the companies in our sample.

[35] IT-related variables are winsorized at the top/bottom 1% of the distribution.

Traditionally, SMEs have been considered the most financially constrained firms within the economy (e.g., Fazzari, Hubbard, and Petersen 1987; Beck, Demirgüç-Kunt, and Maksimovic 2005; Hadlock and Pierce 2010). Financial constraints could only explain part of our results if a large fraction of treated SMEs either had more capital to start with or had easier access to additional finance before the program started. Recall, however, that we compare treated SMEs with control ones that are similar along the financial constraints dimension, among several others. The financial constraints channel, therefore, has little bearing, at best, on our empirical setting. Nonetheless, we perform a test across subsamples of firms to verify whether the information and skills channel is more relevant than the financial capital constraints one for SMEs' digitalization.

Given the limited data available for SMEs, we employ firm size as a commonly used proxy for financial constraints. In particular, we use the natural logarithm of SMEs' total assets in 2013, the year before the implementation of the program, to split the sample between more financially constrained (below median of the firm size distribution) and less financially constrained (above median of the firm size distribution) firms. Panel A of Table XI reports the results of the baseline models as in Table III across the two subsamples. The estimations show that the increase in digitalization by treated companies is statistically similar for both more and less financially constrained firms, suggesting that the financial constraints channel did not play a significant role in determining which companies benefited the most from the *Challenge Fund* program.[36]

As for the propensity of managers to innovate, previous studies show that younger people tend to promote more innovation. Within corporations, they show that firms with a younger group of employees or younger managers tend to increase R&D expenditure and generate more radical innovations (Barker and Mueller 2002; Acemoğlu, Akcigit, and Celik 2022) with, ultimately, positive repercussions on both employment and wage growth (Sarada and Tocoian 2019). In fact, MacDonald and Weisbach (2004) argue that with the rapid evolution of technologies, even if experience and learning by doing might help them, the human capital of older workers is eroded by the competition from young workers whose skills are better tailored to new technologies (*"...turns them into has-beens to some degree..."*, MacDonald and Weisbach [2004, p.289]).

As such, we argue that if our baseline results are driven by such individual trait among treated SMEs, those with younger managers should be more likely to benefit from the

[36]We find similar results when we use the number of employees as alternative proxy for firm size. Results are reported in Table A.VIII in the Appendix

program as they are more willing to adopt new (more sophisticated) technologies and exploit the opportunities from digitalization.

From FAME, we collect information on the age of all directors employed by each firm each year. In the U.K., *directors* are the individuals legally responsible for running a firm and who have a duty to promote its success. In 2013, we count 87,771 directors in our entire sample with the median firm employing about 7 directors. We split the sample between firms with younger directors, where at least one director is under the age of 55 and firms with more mature directors, where all directors' age is above 55.

Results in Panel B indicate that the treatment effect on digitalization is significant across both sub-samples suggesting that not only treated SMEs with younger directors have been affected by the program, but also those with mature directors. Nonetheless, tests of differences across the estimated coefficients indicate that such effect is significantly stronger for firms with mature rather than younger directors. For example, the program increased the probability of having a website of treated SMEs with mature directors by an additional 2.6 percentage points compared to that of treated SMEs with younger directors (see column (1)).

Overall, results in Table XI support the idea that the *Challenge Fund* program was more focused on disseminating digital information and enhancing digital skills rather than tackling the potential barriers of financial constraints and behavioral traits.

(Insert Table XI here)

VII. CONCLUSION

Despite making up the majority of the business population, SMEs face many constraints. One of them is their lack of access to digital knowledge and skills, which survey evidence indicate to be a major barrier for SMEs to adopt digitalization. We exploit a U.K. government program that provides training for SMEs and entrepreneurs on the existence and use of digital technologies for business to isolate the role of *awareness* of digital technologies on SMEs actual digitalization and real outcomes.

We make use of a very granular dataset of web technologies that, to best of our knowledge, has not been used in the literature, so far. It allows us to assemble several types of technologies at firm level every year over the 2011-2019 period. We classify those technologies into "general" and "business", and, within the business category, we further define a sub-class of “e-commerce” technologies.

We find that treated SMEs relative to the control group are more likely to set up a business website, adopt significantly more types of digital technologies, and they are more likely to invest in more sophisticated e-commerce technologies. Our results also show that the implementation of the program contributed to narrowing the existing digital divide between large corporations and SMEs. Moreover, we find positive and significant treatment effects in both rural and urban areas, and in areas close and farther away from universities, suggesting that the spatial digital divide has possibly narrowed after digital knowledge has spread across SMEs in rural areas and areas with lower level of higher education. The increased digitalization has significant end positive effects also on revenue, employment growth and labor productivity, and likelihood of survival. Our results, robust to several falsification tests, seem to be mainly driven by the information and skills channel, while financial constraints do not seem to have a significant bearing.

An alternative to training could be the subsidization of digitalization technologies, assuming SMEs do not adopt technologies because of high capital outlay or other expenses. However, as argued in Acemoğlu and Restrepo (2018), subsidizing R&D, or in our case digitalization technologies, of large incumbent firms could possibly lead to the survival of low-type firms. It is not clear in how far their model can be applied to SMEs, and it is worth for future research to investigate whether the subsidization of digital technologies could be a substitute or complement to the human capital training side of government programs.

Our findings suggest that the lack of knowledge about digital technologies represents an important barrier for SMEs to become more digitalized and that a relatively inexpensive government program that aims at the information dissemination and training side can help reduce such barrier.

ALLIANCE MANCHESTER BUSINESS SCHOOL, UNIVERSITY OF MANCHESTER

REFERENCES

Acemoğlu, Daron, Ufuk Akcigit, and Murat Alp Celik. "Radical and Incremental Innovation: The Roles of Firms, Managers, and Innovators." *American Economic Journal: Macroeconomics* 14, no. 3 (2022): 199–249.

Acemoğlu, Daron, David Autor, David Dorn, Gordon Hanson, and Brendan Price. "The rise of China and the future of US manufacturing." *VoxEU. org, September* 28 (2014).

Acemoğlu, Daron, and Pascual Restrepo. "The race between man and machine: Implications of technology for growth, factor shares, and employment." *American Economic Review* 108, no. 6 (2018): 1488–1542.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. "The skill complementarity of broadband internet." *Quarterly Journal of Economics* 130, no. 4 (2015): 1781–1824.

Al Harbi, Khadejah. "Help everyone get connected with digitalskills.com." Accessed: October 31, 2022. 2014. Https://digitalinclusion.blog.gov.uk/2014/10/06/help-everyone-get-connected-with-digitalskills-com/.

Almus, Matthias, and Dirk Czarnitzki. "The Effects of Public R&D Subsidies on Firms' Innovation Activities: The Case of Eastern Germany." *Journal of Business and Economic Statistics* 21, no. 2 (2003): 226–236.

Arendt, Lukasz. "Barriers to ICT adoption in SMEs: How to bridge the digital divide?" *Journal of Systems and Information Technology* 10, no. 2 (2008): 93–108.

Armstrong, Chris, Yaniv Konchitchki, and Biwen Zhang. "Digital Traffic, Financial Performance, and Stock Valuation." Stanford University Graduate School of Business Research Working Papers Series 4416683, Stanford University, 2023.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen. "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan." *Quarterly Journal of Economics* 132, no. 3 (March 2017): 1101–1164.

Autor, David H., Frank Levy, and Richard J. Murnane. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118, no. 4 (2003): 1279–1333.

Bachas, Natalie, Olivia S. Kim, and Constantine Yannelis. "Loan guarantees and credit supply." *Journal of Financial Economics* 139, no. 3 (2021): 872–894.

Barker, Vincent L., and George C. Mueller. "CEO Characteristics and Firm R&D Spending." *Management Science* 48, no. 6 (2002): 782–801.

Beaudry, Paul, Mark Doms, and Ethan Lewis. "Should the Personal Computer Be Considered a Technological Revolution? Evidence from U.S. Metropolitan Areas." *Journal of Political Economy* 118, no. 5 (2010): 988–1036.

Beck, Thorsten, ASLI Demirgüç-Kunt, and Vojislav Maksimovic. "Financial and legal constraints to growth: Does firm size matter?" *Journal of Finance* 60, no. 1 (2005): 137–177.

BEIS. Business population estimates for the UK and regions 2021: statistical release. Https://www.gov.uk/government/statistics/business-population-estimates-2021/business-population-estimates-for-the-uk-and-regions-2021-statistical-release-html.

Berlingieri, Giuseppe, Sara Calligaris, Chiara Criscuolo, and Rudy Verlhac. "Laggard firms, technology diffusion and its structural and policy determinants." OECD Science, Technology and Industry Policy Papers 86, OECD, 2020.

Bertschek, Irene, Helmut Fryges, and Ulrich Kaiser. "B2B or not to be: Does B2B e-commerce increase labour productivity?" *International Journal of the Economics of Business* 13, no. 3 (2006): 387–405.

Biagi, Federico, and Martin Falk. "The impact of ICT and e-commerce on employment in Europe." *Journal of Policy Modeling* 39, no. 1 (2017): 1–18.

Billon, Margarita, Fernando Lera-Lopez, and Rocio Marco. "ICT use by households and firms in the EU: Links and determinants from a multivariate perspective." *Review of World Economics* 152, no. 4 (2016): 629–654.

BIS. "Evaluation of the Small Business Digital Capability Programme Challenge Fund Final Report." Department of Business, Innovation and Skills Research Papers 248, Department of Business, Innovation and Skills, 2015.

———. "Government launches support to help small businesses do more online." Accessed: October 31, 2022. 2014. Https://www.gov.uk/government/news/government-launches-support-to-help-small-businesses-do-more-online.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. "Americans do IT better: US multinationals and the productivity miracle." *American Economic Review* 102, no. 1 (2012): 167–201.

Bloom, Nicholas, John Van Reenen, and Heidi Williams. "A toolkit of policies to promote innovation." *Journal of Economic Perspectives* 33, no. 3 (2019): 163–84.

Bonfim, Diana, Cláudia Custódio, and Clara Raposo. "Supporting small firms through recessions and recoveries." *Journal of Financial Economics, forthcoming*, 2023.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics* 117, no. 1 (2002): 339–376.

Bronzini, Raffaello, and Eleonora Iachini. "Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach." *American Economic Journal: Economic Policy* 6, no. 4 (2014): 100–134.

Bronzini, Raffaello, and Paolo Piselli. "The impact of R&D subsidies on firm innovation." *Research Policy* 45, no. 2 (2016): 442–457.

Brown, J. David, and Johns S. Earle. "Finance and Growth at the Firm Level: Evidence from SBA Loans." *Journal of Finance* 72, no. 3:1039–1080.

Brynjolfsson, Erik, and Andrew McAfee. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies.* WW Norton & Company, 2014.

BuiltWith. "Data Coverage." Accessed: October 31, 2022. 2022. Https://builtwith.com/data-coverage.

Campello, Murillo, Janet Gao, and Qin Xu. "Personal Taxes and Labor Downskilling: Evidence from 27 Million Job Postings." *Management Science, forthcoming* 102, no. 1 (2023): 167–201.

Campello, Murillo, John R Graham, and Campbell R Harvey. "The real effects of financial constraints: Evidence from a financial crisis." *Journal of Financial Economics* 97, no. 3 (2010): 470–487.

Caselli, Francesco, and Wilbur John Coleman. "Cross-country technology diffusion: The case of computers." *American Economic Review* 91, no. 2 (2001): 328–335.

Cassetta, Ernesto, Umberto Monarca, Ivano Dileo, Claudio Di Berardino, and Marco Pini. "The relationship between digital technologies and internationalisation. Evidence from Italian SMEs." *Industry and Innovation* 27, no. 4 (2020): 311–339.

Chinn, Menzie D, and Robert W Fairlie. "The determinants of the global digital divide: A cross-country analysis of computer and internet penetration." *Oxford Economic Papers* 59, no. 1 (2007): 16–44.

Cohn, Jonathan B., Zack Liu, and Malcolm I. Wardlaw. "Count (and count-like) data in finance." *Journal of Financial Economics* 146, no. 2 (2022): 529–551.

Columba, Francesco, Leonardo Gambacorta, and Paolo Emilio Mistrulli. “Mutual guarantee institutions and small business finance.” *Journal of Financial Stability* 6, no. 1 (2010): 45–54.

Comin, Diego, and Martıé Mestieri. “If technology has arrived everywhere, why has income diverged?” *American Economic Journal: Macroeconomics* 10, no. 3 (2018): 137–78.

Cong, Lin William, Xiaohang Yang, and Xiaobo Zhang. “SMEs Amidst the Pandemic and Reopening: Digital Edge and Transformation.” The SC Johnson College of Business Applied Economics and Policy Working Papers Series 2022-01, Cornell University, 2022.

Cusolito, Ana Paula, Daniel Lederman, and Jorge Peña. “The Effects of Digital-Technology Adoption on Productivity and Factor Demand.” World Bank Policy Research Working Papers Series 9333, World Bank, 2020.

D’Ignazio, Alessio, and Carlo Menon. “Causal Effect of Credit Guarantees for Small- and Medium-Sized Enterprises: Evidence from Italy.” *Scandinavian Journal of Economics* 122, no. 1 (2020): 191–218.

DeStefano, Timothy, Richard Kneller, and Jonathan Timmis. “Broadband infrastructure, ICT use and firm performance: Evidence for UK firms.” *Journal of Economic Behavior & Organization* 155 (2018): 110–139.

———. “Cloud computing and firm growth.” *Review of Economics & Statistics, forthcoming*, 2023.

Dicken, Peter, and Peter E Lloyd. “Geographical perspectives on United States investment in the United Kingdom.” *Environment and Planning A* 8, no. 6 (1976): 685–705.

Draca, Mirko, Raffaella Sadun, and John Van Reenen. “Productivity and ICTs: A review of the evidence.” Edited by Chrisanthi Avgerou (ed.) et al. Chap. 5 in *The Oxford Handbook of Information and Communication Technologies*, 100–147. Oxford: Oxford University Press, 2009.

Eurostat. NUTS - Nomenclature of Territorial Units for Statistics. Https://ec.europa.eu/eurostat/web/nuts/background.

Evangelista, Rinaldo, Paolo Guerrieri, and Valentina Meliciani. “The economic impact of digital technologies in Europe.” *Economics of Innovation and New Technology* 23, no. 8 (2014): 802–824.

Fazzari, Steven, R Glenn Hubbard, and Bruce C Petersen. "Financing constraints and corporate investment." NBER Working Papers Series 2387, NBER, 1987.

Fitzgerald, Michael, Nina Kruschwitz, Didier Bonnet, and Michael Welch. "Embracing digital technology: A new strategic imperative." *MIT Sloan Management Review* 55, no. 2 (2014): 1.

Gal, Peter, and Jagoda Egeland. "Reducing regional disparities in productivity in the United Kingdom." OECD Economics Department Working Papers Series 1456, OECD, 2018.

Gal, Peter, Giuseppe Nicoletti, Theodore Renault, Stéphane Sorbe, and Christina Timiliotis. "Digitalisation and productivity: In search of the holy grail–Firm-level empirical evidence from EU countries." OECD Economics Department Working Papers 1533, OECD, 2019.

Goldfarb, Avi, and Catherine Tucker. "Digital economics." *Journal of economic literature* 57, no. 1 (2019): 3–43.

Gonzalez-Uribe, Juanita, and Daniel Paravisini. "How Sensitive is Young Firm Investment to the Cost of Outside Equity? Evidence from a UK Tax Relief." LSE Working Papers, London School of Economics and Political Science, 2019. https://drive.google.com/file/d/1a9ifjqZY9y_gW7TY37N10a1PeluKucLB/view?pli=1.

Gonzalez-Uribe, Juanita, and Su Wang. "The Effects of Small-Firm Loan Guarantees in the UK: Insights for the COVID-19 Pandemic Crisis." LSE Discussion Papers Series 795, London School of Economics and Political Science, 2020.

Greenstein, Shane. "The basic economics of internet infrastructure." *Journal of Economic Perspectives* 34, no. 2 (2020): 192–214.

Hadlock, Charles J, and Joshua R Pierce. "New evidence on measuring financial constraints: Moving beyond the KZ index." *Review of Financial Studies* 23, no. 5 (2010): 1909–1940.

Haller, Stefanie A, and Iulia Siedschlag. "Determinants of ICT adoption: Evidence from firm-level data." *Applied Economics* 43, no. 26 (2011): 3775–3788.

Hoffreumon, Charles, and Vincent Labhard. "Cross-country cross-technology digitalisation: A Bayesian hierarchical model perspective." European Central Bank Working Papers Series 2022/2700, European Central Bank, 2022.

Howell, Sabrina T. "Financing innovation: Evidence from R&D grants." *American Economic Review* 107, no. 4 (2017): 1136–64.

Jardim, Ekaterina S, Mark C Long, Robert Plotnick, Emma van Inwegen, Jacob L Vigdor, and Hilary Wething. "Boundary Discontinuity Methods and Policy Spillovers." NBER Working Papers Series 30075, National Bureau of Economic Research, 2022.

Kickul, Jill, and Lisa K Gundry. "Breaking through boundaries for organizational innovation: New managerial roles and practices in e-commerce firms." *Journal of Management* 27, no. 3 (2001): 347–361.

Koellinger, Philipp. "The relationship between technology, innovation, and firm performance — Empirical evidence from e-business in Europe." *Research Policy* 37, no. 8 (2008): 1317–1328.

Lach, Saul. "Do R&D Subsidies Stimulate or Displace Private R&D? Evidence from Israel." *Journal of Industrial Economics* 50, no. 4 (2002): 369–390.

Lelarge, Claire, David Sraer, and David Thesmar. *Entrepreneurship and Credit Constraints: Evidence from a French Loan Guarantee Program.* In *International Differences in Entrepreneurship*, by Josh Lerner and Antoinette Schoar, 243–273. University of Chicago Press, 2010. http://www.nber.org/chapters/c8218.

Leong, Carmen, Shan L Pan, Sue Newell, and Lili Cui. "The emergence of self-organizing e-commerce ecosystems in remote villages of China." *MIS Quarterly* 40, no. 2 (2016): 475–484.

Lerner, Josh. "The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program." *Journal of Business* 72, no. 3 (1999): 285–318. Accessed February 17, 2023. http://www.jstor.org/stable/10.1086/209616.

MacDonald, Glenn, and Michael S. Weisbach. "The Economics of Has-beens." *Journal of Political Economy* 112, nos. S1 (2004): S289–S310.

McElheran, Kristina. "Do market leaders lead in business process innovation? The case (s) of e-business adoption." *Management Science* 61, no. 6 (2015): 1197–1216.

Millán, José Marıéa, Serhiy Lyalkov, Andrew Burke, Ana Millán, and André van Stel. "'Digital Divide' among European entrepreneurs: Which types benefit most from ICT implementation?" *Journal of Business Research* 125 (2021): 533–547.

Nominet. UK Register Statistics - 2021. Https://www.nominet.uk/news/reports-statistics/uk-register-statistics-2021.

OECD. "Digital Economic Report 2021 - Cross-border data flows and development: For whom the data flow." OECD Publishing, OECD G7 Background Report, 2021.

OECD. "SME digitalisation to "Build Back Better"." OECD Publishing, OECD Studies on SMEs and Entrepreneurship, 2021.

———. "The Digital Transformation of SMEs." OECD Publishing, OECD Studies on SMEs and Entrepreneurship, 2021.

ONS. "Monitoring e-commerce: 2014." Accessed: October 31, 2022. 2014. Https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/articles/monitoringecommerce/2014-08-07#toc.

———. "Qualifications question development for Census 2021." Accessed: October 31, 2022. 2021. Https://www.ons.gov.uk/census/planningforcensus2021/questiondevelopment/qualificationsquestiondevelopmentforcensus2021.

Sarada and Oana Tocoian. "Is It All About Who You Know? Prior Work Connections and Entrepreneurial Success." *ILR Review* 72, no. 5 (2019): 1200–1224.

Shearer, Eleanor. "Local enterprise partnerships." Accessed: October 31, 2022. 2021. Https://www.instituteforgovernment.org.uk/article/explainer/local-enterprise-partnerships.

Syverson, Chad. "What determines productivity?" *Journal of Economic Literature* 49, no. 2 (2011): 326–65.

Thonipara, Anita, Rolf Sternberg, Till Proeger, and Lukas Haefner. "Digital divide, craft firms' websites and urban-rural disparities—Empirical evidence from a web-scraping approach." *Review of Regional Research*, 2022, 1–31.

Wallsten, Scott J. "The Effects of Government-Industry R&D Programs on Private R&D: The Case of the Small Business Innovation Research Program." *RAND Journal of Economics* 31, no. 1 (2000): 82–100. Accessed February 17, 2023. https://EconPapers.repec.org/RePEc:rje:randje:v:31:y:2000:i:spring:p:82-100.

Wielicki, Tom, and Lukasz Arendt. "A knowledge-driven shift in perception of ICT implementation barriers: Comparative study of US and European SMEs." *Journal of Information Science* 36, no. 2 (2010): 162–174.

Willcocks, Leslie P. "COVID-19 may exacerbate the digital divide among businesses." Accessed: October 31, 2022. LSE Business Review, 2020. Https://blogs.lse.ac.uk/businessreview/2020/09/03/covid-19-may-exacerbate-the-digital-divide-among-businesses.

TABLES

Table I
Pre-Program Characteristics between Treated and Control

Variables	Treated Areas (1)	Control Areas (2)	Diff. (1)-(2)
Panel A: Internet Infrastructure			
% Areas without good internet connection	0.09	0.10	-0.01
			(-1.03)
% Internet Users	0.82	0.80	0.02
			(1.65)
Total Websites	1,127	1,395	-268
			(-0.58)

(continued on next page)

Notes. This table compares several macro characteristics across treated (Column 1) and control areas (Column 2) in the pre-program period from 2011 to 2014. Treated areas include LEPs that received funding from the *Challenge Fund*. Control areas include both LEPs that did not receive funding from the *Challenge Fund* and other NUTS-2 areas in Wales and Scotland that were not covered by the *Challenge Fund* in the first place. Column 3 reports differences in all characteristics between the treated and control areas as well as *t*-statistics in parentheses from two-tailed, two-sample t-tests of the difference in means. Panel A reports characteristics related to internet infrastructure. *%Areas without good internet connection* is the percentage of local authority districts that receive less than 2Mb/s; *%Internet Users* is the fraction of the population over the age of 16 that used the internet within the three months when the survey took place. *Total Websites* is the total number of business websites. Panel B reports characteristics related to business demographic and economic conditions. *Business Population* is the total number of businesses in each area. *Number of Business Births* is the number of new firms. *Number of Business Deaths* is the number of dissolved firms. *%Businesses survived aft 1 yr* is the percentage of new firms that survived after one year. *%Businesses survived aft 3 yr* is the percentage of new firms survived after three years. *Employment* is the total number of employed workers aged from 16 to 64. *%Educated Workers* is the percentage of workers between the ages 16 and 64 who hold at least an undergraduate degree (NVQ4) or equivalent. *GDP* is the gross domestic product. *GDP Growth* is the rate of GDP growth in that area. *GVA* is the gross value added in that area. Panel C reports the political affinity in the treated and control areas. *#Conservative LAs / #Other Parties LAs* is the ratio of local authorities controlled by the Conservative party to local authorities controlled by the Labour or other parties. The total number of observations used in this analysis is 135; 63 observations for the treated areas and 72 for the control ones. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table I
Pre-Program Characteristics of Treated and Control Areas (cont'd)

Variables	Treated Areas (1)	Control Areas (2)	Diff. (1)-(2)
Panel B: Business demographic and economic conditions			
Business Population	43,691	43,122	569
			(0.0765)
Number of Business Birth	5,237	5,704	-467
			(-0.32)
Number of Business Death	5,193	5,213	-20
			(-0.02)
%Businesses survived aft 1 yr	0.93	0.93	-0.00
			(-0.38)
%Businesses survived aft 3 yr	0.63	0.62	0.01
			(0.92)
Employment	570,543	580,326	-9,783
			(-0.09)
%Educated Workers	0.33	0.33	-0.00
			(-0.30)
GDP	30,630.98	37,238.68	-6,607.70
			(-0.73)
GDP Growth	0.035	0.034	0.001
			(0.41)
GVA	27,015.98	33,544.22	-6,528.24
			(0.78)
Panel C: Political Affinity			
$\frac{\#\text{Conservative LAs}}{\#\text{Other Parties LAs}}$	2.171	1.145	1.026
			(1.55)

Table II
Sample Characteristics Before and After the PSM

Panel A: Before PSM			
	Treated Areas (N = 10,909)	Control Areas (N =16,787)	Diff.
Variables	(1)	(2)	(1)-(2)
Ln(Total Assets)	7.157	7.062	0.095***
			(2.99)
Ln(Age)	2.410	2.405	0.005
			(0.47)
Leverage	0.228	0.236	-0.008**
			(-2.26)
Cash	0.260	0.277	-0.016***
			(-4.11)
Sales Growth	0.035	0.049	-0.015***
			(-3.12)
ROA	0.164	0.121	0.044***
			(5.26)
p-score	0.397	0.400	0.003***
			(7.96)

Panel B: After PSM			
	Treated Areas (N = 8,799)	Control Areas (N = 8,799)	Diff.
	(1)	(2)	(1)-(2)
Ln(Total Assets)	7.349	7.354	-0.005
			(1.32)
Ln(Age)	2.584	2.588	-0.004
			(-0.36)
Leverage	0.227	0.228	-0.001
			(-0.18)
Cash	0.257	0.257	0.000
			(0.07)
Sales Growth	0.036	0.036	0.000
			(0.06)
ROA	0.096	0.096	0.000
			(0.01)
p-score	0.405	0.405	0.000
			(0.10)

Notes. This table reports firm-level characteristics of treated (Column 1)and control firms (Column 2) in 2013 before (Panel A) and after (Panel B) propensity score matching (PSM). Column 3 reports differences in the characteristics between the treated and control firms as well as t-statistics in parentheses from two-tailed, two-sample t-tests of the difference in means. *Ln(Total Assets)* is the natural logarithm of one plus the firm's total assets; *Ln(Age)* is the natural logarithm of one plus the firm's age; *Leverage* is the sum of short-term debt and long-term liability divided by total assets; *Cash* is the ratio of cash to total assets; *Sales Growth* is the growth rate of the annual turnover; *ROA* is measured as earnings before interest and tax (EBIT) divided by total assets; *p-score* is the probability of being treated estimated via a probit model using 2013 values of *Ln(Total Assets)*, *Ln(Age)*, *Leverage*, *Cash*, *Sales Growth*, *ROA*, and 1-digit U.K.-SIC dummies as covariates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table III
Digitalization Outcomes: Baseline Results

	Extensive Margin			Intensive Margin								
	Web			*Ln(BG-score)*			*Ln(B-score)*			*Ln(G-score)*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post × *Treated*	0.030***	0.029***	0.027***	0.075***	0.073***	0.069***	0.057***	0.055***	0.058***	0.058***	0.056***	0.053***
	(2.90)	(3.44)	(3.27)	(3.08)	(3.61)	(3.48)	(3.23)	(3.75)	(3.65)	(3.05)	(3.59)	(3.45)
Observations	148,965	148,965	147,532	148,955	148,955	147,522	148,955	148,955	147,522	148,955	148,955	147,522
Adjusted R^2	0.47	0.49	0.49	0.47	0.49	0.49	0.46	0.47	0.47	0.47	0.48	0.48
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Geographic FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where $Y_{i,t}$ includes the digitalization outcomes *Web* for the extensive margin, and *Ln(BG-score)*, *Ln(B-score)*, and *Ln(G-score)* for the intensive margin. *Web* is an indicator equal to one when the firm has a website in that year and zero otherwise. *Ln(BG-score)* is the natural logarithm of one plus the number of technology tags detected on the SMEs' website. *Ln(B-score)* is the natural logarithm of one plus the number of "Business" technology tags detected on the SMEs' website. *Ln(G-score)* is the natural logarithm of one plus the number of "General" technology tags detected on the SMEs' website. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. *Treated* equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. $\boldsymbol{Z}$ is the matrix of fixed effects, including firm (i), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on 1-digit U.K.-SIC code. The geographic fixed effect is based on NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table IV
Digitalization Outcomes: E-commerce

	Extensive Margin			Intensive Margin		
	E-commerce			*Ln(E-score)*		
	(1)	(2)	(3)	(4)	(5)	(6)
$Post \times Treated$	0.028***	0.027***	0.026***	0.024***	0.024***	0.023***
	(3.17)	(3.61)	(3.54)	(3.07)	(3.43)	(3.37)
Observations	148,955	148,955	147,522	148,955	148,955	147,522
adjusted R^2	0.40	0.42	0.42	0.40	0.42	0.42
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes	No	Yes	Yes
Geographic FE	No	No	Yes	No	No	Yes

Notes. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t}\ ,$$

where $Y_{i,t}$ includes the digitalization outcomes E-commerce for the extensive margin; and *Ln(E-score)* for the intensive margin. *E-commerce* is an indicator equal to one if the SME adopted any e-commerce technology on its website in that year and zero otherwise. *Ln(E-score)* is the natural logarithm of one plus the sum of all tags associated with e-commerce detected on the SME's website. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. *Treated* equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. ***Z*** is the matrix of fixed effects, including firm (i), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on 1-digit U.K.-SIC code. The geographic fixed effect is based on NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table V

Digitalization Catching-up between SMEs and Large firms

	Gap(BG-score)			*Gap(B-score)*			*Gap(G-score)*			*Gap(E-score)*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$Post \times Treated$	-0.395***	-0.384***	-0.373***	-0.186***	-0.181***	-0.176***	-0.209***	-0.203***	-0.196***	-0.047***	-0.046***	-0.045***
	(-4.31)	(-4.71)	(-4.60)	(-4.35)	(-4.63)	(-4.53)	(-3.81)	(-4.15)	(-4.03)	(-3.36)	(-3.47)	(-3.43)
Observations	148,955	148,955	147,522	148,955	148,955	147,522	148,955	148,955	147,522	148,955	148,955	147,522
Adjusted R^2	0.47	0.48	0.48	0.49	0.51	0.50	0.44	0.46	0.46	0.46	0.48	0.48
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry $\times$ Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Geographic FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes. This table shows the results from the model estimation:

$$Gap_{i,t} = \delta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where $Gap_{i,t}$ is defined as:

$$Gap_{i,t} = \overline{Y^l_{g,t}} - Y_{i,t},$$

where $\overline{Y^l_{g,t}}$ is the average digitalization outcome across large firms that are located in the same area g of SME i at time t, and $Y_{i,t}$ is the digitalization outcome of SME i at time t. Large firms are defined as companies with more than 250 employees and an annual turnover of more than €50 million. The geographical area is based on LEPs (NUTS2-level). Digitalization outcomes include: *BG-score*, *B-score*, *G-score* and *E-score*. *BG-score* is the number of technology tags detected on the SMEs' website. *B-score* is the number of "Business" technology tags detected on the SMEs' website. *G-score* is the number of "General" technology tags detected on the SMEs' website. *E-score* is the sum of all tags associated with e-commerce detected on the SME's website. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. Treated equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. $\boldsymbol{Z}$ is the matrix of fixed effects, including firm (i), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on 1-digit U.K.-SIC code. The geographic fixed effect is based on NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table VI
Digitalization across Geographical Areas

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
Panel A: Urban Areas				
$Post \times Treated$	0.018**	0.040**	0.031**	0.029*
	(2.29)	(2.08)	(2.29)	(1.93)
Observations	39,994	39,991	39,991	39,991
Adjusted R^2	0.46	0.45	0.44	0.45
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Panel B: Rural Areas				
$Post \times Treated$	0.025***	0.068***	0.052***	0.052***
	(2.43)	(2.76)	(2.92)	(2.76)
Observations	108,971	108,964	108,964	108,964
Adjusted R^2	0.48	0.47	0.46	0.47
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Difference (A - B)	-0.007	-0.028	-0.021	-0.023
F-stat	[0.64]	[1.74]	[1.79]	[1.96]

Notes. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where the model is separately estimated on subsamples of urban (Panel A) and rural firms (Panel B), and universities areas (Panel C) and all other areas (Panel D). The urban subsample contains SMEs in the largest 10 cities by population according to the 2011 census; while the rural one contains SMEs in all other locations. The universities areas subsample contains SMEs in cities with the 20 top universities according to the REF2014 ranking; while the all other areas one contains SMEs in all other locations. Y includes the digitalization outcomes *Web*, *Ln(BG-score)*, *Ln(B-score)*, and *Ln(G-score)*. *Web* is an indicator equal to one when the firm has a website in that year and zero otherwise. *Ln(BG-score)* is the natural logarithm of one plus the number of technology tags detected on the SMEs' website. *Ln(B-score)* is the natural logarithm of one plus the number of "Business" technology tags detected on the SMEs' website. *Ln(G-score)* is the natural logarithm of one plus the number of "General" technology tags detected on the SMEs' website. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. *Treated* equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. $\boldsymbol{Z}$ is the matrix of fixed effects, including firm (i) and year (t) fixed effects. This table reports the baseline specification with firm and year fixed effects only. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. Chow test results for the difference between estimated coefficients across urban and rural subsample are reported at the bottom of the table. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table VI
Digitalization across Geographical Areas (cont'd)

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
Panel C: Universities Areas				
$Post \times Treated$	0.026***	0.060**	0.044**	0.045**
	(3.11)	(3.04)	(3.15)	(2.84)
Observations	35,839	35,836	35,836	35,836
Adjusted R^2	0.45	0.45	0.44	0.45
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Panel D: All Other Areas				
$Post \times Treated$	0.020***	0.055***	0.043***	0.042***
	(1.89)	(2.18)	(2.39)	(2.16)
Observations	113,126	113,119	113,119	113,119
Adjusted R^2	0.47	0.47	0.46	0.47
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Difference (C - D)	0.006	0.005	0.001	0.003
F-stat	[0.30]	[0.04]	[0.00]	[0.03]

Table VII
Sales Growth and ROA

	(1)	(2)	(3)
	Panel A: *Sales Growth*		
Post × *Treated*	0.010*	0.010**	0.011**
	(1.96)	(2.03)	(2.16)
Observations	113,940	113,940	112,868
Adjusted R^2	0.01	0.02	0.01
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes
Geographic FE	No	No	Yes
	Panel B: *ROA*		
Post × *Treated*	0.009*	0.009*	0.009*
	(1.75)	(1.65)	(1.73)
Observations	120,750	120,750	120,594
Adjusted R^2	0.40	0.40	0.40
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes
Geographic FE	No	No	Yes

Notes. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where $Y_{i,t}$ includes *Sales Growth* (Panel A) and *ROA* (Panel B), respectively. *Sales Growth* is the growth rate of the annual turnover. *ROA* is measured as earnings before interest and tax (EBIT) divided by total assets. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. *Treated* equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. ***Z*** is the matrix of fixed effects, including firm (i), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on 1-digit U.K.-SIC code. The geographic fixed effect is based on NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table VIII
Employment and Labor Productivity Growth

	(1)	(2)	(3)
Panel A: Δ*Ln(Employees)*			
$Post \times Treated$	0.009**	0.009**	0.009**
	(2.20)	(2.16)	(2.18)
Observations	111,218	111,218	111,213
Adjusted R^2	0.02	0.03	0.03
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes
Geographic FE	No	No	Yes
Panel B: Δ*Ln(VPE)*			
$Post \times Treated$	0.027*	0.029*	0.030*
	(1.67)	(1.77)	(1.83)
Observations	81,109	81,109	81,109
Adjusted R^2	0.07	0.07	0.07
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes
Geographic FE	No	No	Yes

Notes. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where $Y_{i,t}$ includes Δ*Ln(Employees)* (Panel A) and Δ*Ln(VPE)* (Panel B), respectively. Δ*Ln(Employees)* is the change in the natural logarithm of the number of employees between year t and (t1). Δ*Ln(VPE)* is the change in the natural logarithm of the firm's EBITDA scaled by the number of employees. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. Treated equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. ***Z*** is the matrix of fixed effects, including firm (i), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on 1-digit U.K. -SIC code. The geographic fixed effect is based on NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table IX
Survival

	2011 - 2019	2011 - 2021
	(1)	(2)
$Post \times Treated$	-0.289*	-0.303*
	(-1.71)	(-1.81)
$Treated$	0.170	0.173
	(1.03)	(1.05)
Observations	144,207	170,775

Notes. This table shows the results from the Cox-proportional Hazard model estimation:

$$\lambda(it) = \phi_t \; exp(\theta \; Treated_i + \gamma_{it} \; (Post_t \times Treated_i))$$

where $\lambda(it)$ indicates the hazard rate, i.e., the probability of firm i having been dissolved or in liquidation at time t, conditional on surviving at that time. ϕ_t is the baseline hazard rate. We estimate the model for 2011-2019 sample period (column (1)) and for the extended period that includes the first two years of the pandemic, i.e., 2011-2021 (column (2)). *Post* is an indicator equal to one after 2014, and zero otherwise. *Treated* is an indicator equal to one (zero) for SMEs without website before the program located in the treated LEPs (control areas). All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. z-statistics (in parentheses) are calculated from standard errors clustered at the firm-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table X
Establishment-level IT Outcomes

	Ln(IT Staff)			*Ln(IT Expenses)*		
	(1)	(2)	(3)	(4)	(5)	(6)
$Post \times Treated$	0.059**	0.064**	0.066*	0.124**	0.128**	0.123*
	(2.14)	(2.28)	(1.90)	(2.02)	(2.45)	(1.91)
Observations	22,174	22,174	21,662	18,196	18,196	17,769
Adjusted R^2	0.56	0.56	0.57	0.64	0.66	0.66
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes	No	Yes	Yes
Geographic FE	No	No	Yes	No	No	Yes

Notes. This table shows the results from the model estimation:

$$Y_{e,i,t} = \beta(Post_t \times Treated_e) + \mathbf{Z_{i,t,j,g}} + \epsilon_{e,i,t} \ ,$$

where $Y_{e,i,t}$ represents the IT outcomes for establishment e of firm i at time t. The IT-related outcomes are the log-transformed variables with one plus the original values ($Ln(IT\ Staff)$ and $Ln(IT\ Expenses)$). *Post* is a dummy equal to 1 in years 2016 and 2018, and zero in years 2010, 2011 and 2012. *Treated* equals one (zero) for establishments located in the *Treated (Control)* areas. ***Z*** is a matrix including firm, year, industry-by-year, and geography fixed effects. Industry is based on 1-digit U.K.-SIC code. The geographic fixed effect is based on NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2012 values of revenue, employment, and industry code as covariates. Each treated establishment is matched with one unique control establishment. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table XI
Digitalization Outcomes:
Financial Constraints and Directors' Age

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
Panel A: Subsamples by Financial Constraints				
	Small Firms			
Post × *Treated*	0.031***	0.074***	0.055***	0.057***
	(2.86)	(2.83)	(2.90)	(2.78)
Observations	73,277	73,273	73,273	73,273
Adjusted R^2	0.47	0.47	0.46	0.46
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
	Medium Firms			
Post × *Treated*	0.029***	0.077***	0.059***	0.060***
	(2.68)	(3.07)	(3.29)	(3.08)
Observations	75,688	75,682	75,682	75,682
Adjusted R^2	0.47	0.47	0.46	0.47
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Difference				
Small Firms – Medium Firms	0.002	-0.003	-0.004	-0.003
F-stat	[0.07]	[0.04]	[0.14]	[0.06]

(continued on next page)

Notes. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where the model is separately estimated on subsamples of firms characterized by financial constraints (Panel A) and directors' age (Panel B). Financial constraints are defined by firm total assets where 'Small Firms' ('Medium Firms') are those with total assets below (above) median. Directors' age is based on the age of firm's directors where firms with younger directors ('Young Directors') are those with at least one director who is under the age of 55; while firms with mature directors ('Mature Directors') are those with all directors aged above 55. Y includes the digitalization outcomes *Web Ln(BG-score)*, *Ln(G-score)*, and *Ln(B-score)*. *Web* is an indicator equal to one when the firm has a website in that year and zero otherwise. *Ln(BG-score)* is the natural logarithm of one plus the number of technology tags detected on the SMEs' website. *Ln(B-score)* is the natural logarithm of one plus the number of "Business" technology tags detected on the SMEs' website. *Ln(G-score)* is the natural logarithm of one plus the number of "General" technology tags detected on the SMEs' website. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. *Treated* equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. ***Z*** is the matrix of fixed effects, including firm (i), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on 1-digit U.K.-SIC code. The geographic fixed effect is based on NUTS-2 level. This table reports the baseline specification with firm and year fixed effects only. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. Chow test results for the difference between estimated coefficients across the pair subsamples based on financial constraints and directors' age constraints are reported at the bottom of each Panel. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

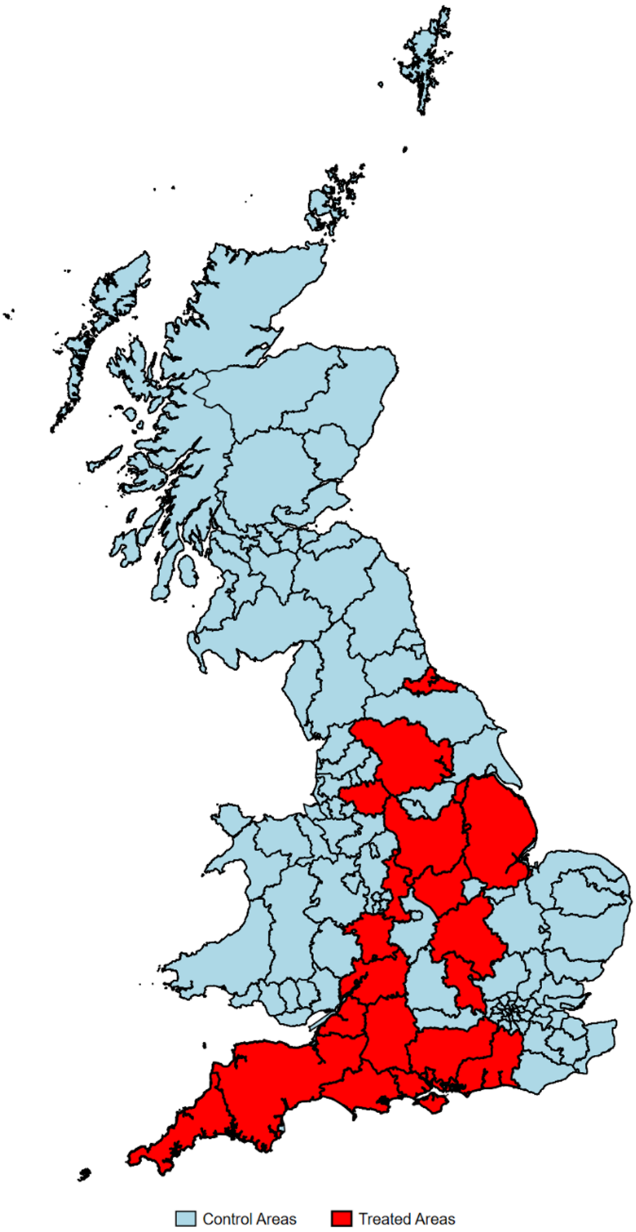
Table XI
Digitalization Outcomes:
Financial Constraints and Directors' Age (cont'd)

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
Panel B: Subsamples by Directors' Age				
	Young Directors			
Post × *Treated*	0.012*	0.032**	0.026**	0.025**
	(1.73)	(2.06)	(2.31)	(2.06)
Observations	44,335	44,335	44,335	44,335
Adjusted R^2	0.47	0.47	0.46	0.47
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
	Mature Directors			
Post × *Treated*	0.038**	0.095***	0.071***	0.073***
	(2.55)	(2.70)	(2.87)	(2.65)
Observations	77,492	77,483	77,483	77,483
Adjusted R^2	0.53	0.53	0.51	0.52
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Difference				
Young Directors – Mature Directors	-0.026**	-0.063**	-0.045**	-0.048**
F-stat	[4.71]	[4.99]	[5.39]	[4.66]

FIGURES

Figure I
Treated LEPs and Control Areas

Notes. The map depicts the treated LEPs (red) in England and the control areas (light blue) in England, Scotland, and Wales.

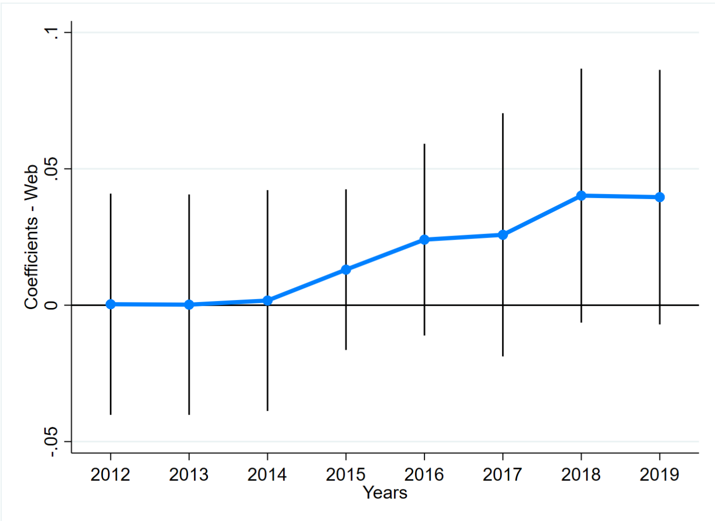


APPENDIX

Figure A.I
Dynamic Models for Digitalization Outcomes

Notes. This figure reports the estimated coefficients obtained from the pairwise-matched difference-in-differences yearly interactions of the augmented baseline model in equation (1) using 2011 as the base year. The depicted results are from the specification with firm and year fixed effects. The vertical axis shows the estimated coefficients. The horizontal axis shows the years. The solid vertical lines show the 95% confidence interval. Panels A to D shows the estimated coefficients where the dependent variable is either *Web*, *Ln(BG-score)*, *Ln(B-score)* or *Ln(G-score)*, respectively.

Panel A: Web



Panel B: $Ln(BG\text{-}score)$

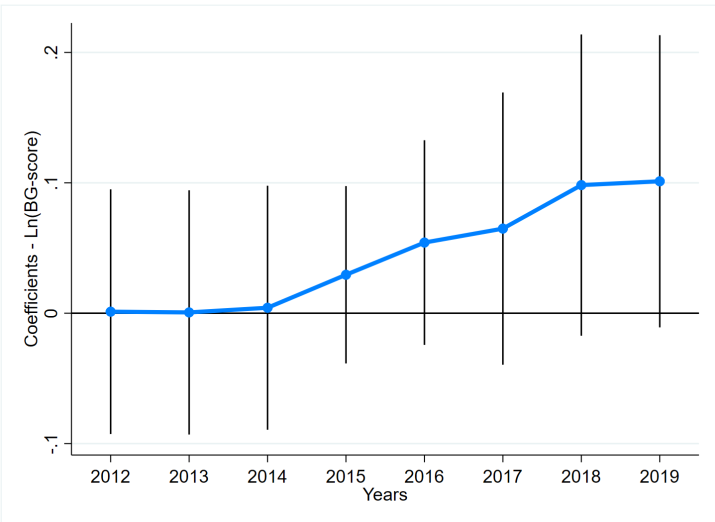
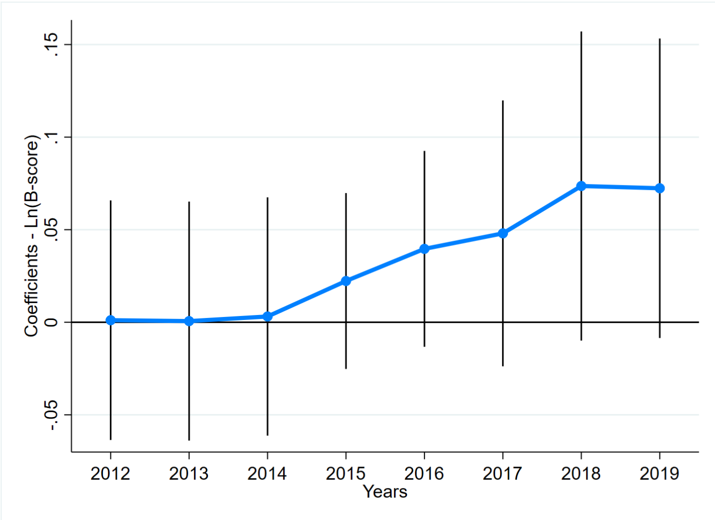


Figure A.I
Dynamic Models for Digitalization Outcomes (cont'd)

Panel C: *Ln*(*B-score*)



Panel D: *Ln*(*G-score*)

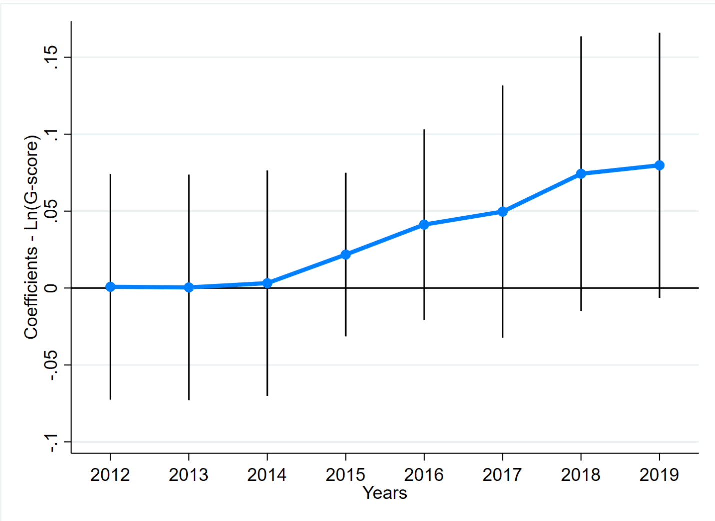
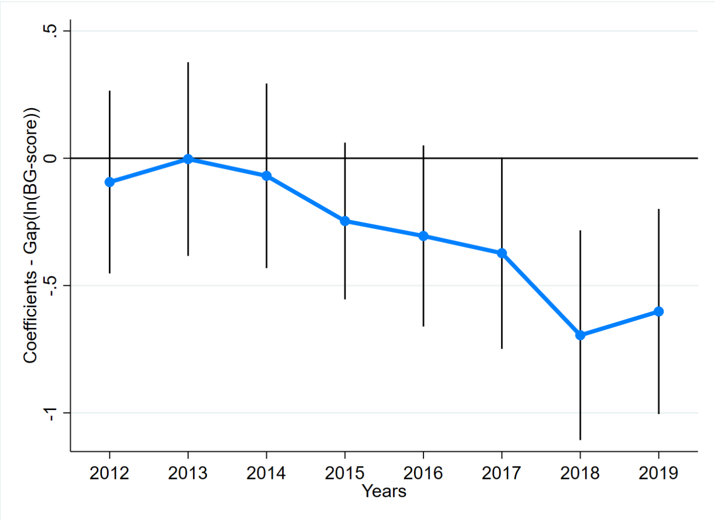


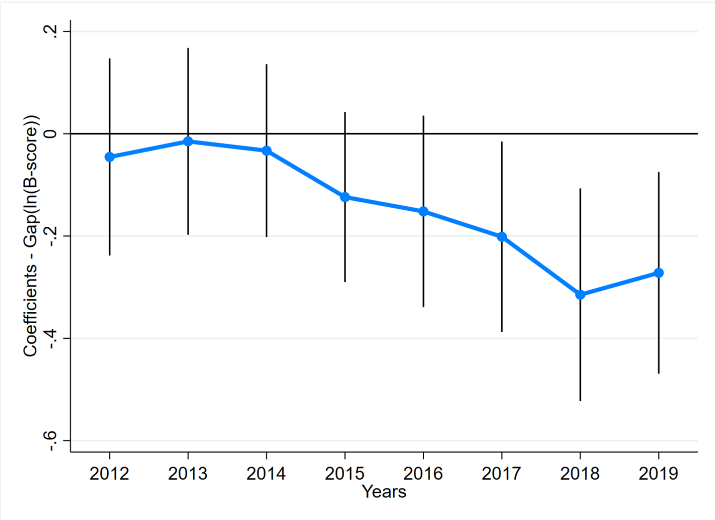
Figure A.II
Dynamic Models for Digitalization Catching-up

Notes. This figure reports the estimated yearly interaction coefficients from the pairwise-matched difference-in-differences regressions of the augmented catching-up model in equation (2) using 2011 as the base year. Results here are from the specification with firm and year fixed effects. The vertical axis shows the estimated coefficients. The horizontal axis shows the years. The solid vertical lines show the 95% confidence interval. Panels A to D show the estimated coefficients where the dependent variable is either *Gap(BG-score)*, *Gap(B-score)*, *Gap(G-score)* or *Gap(E-score)*, respectively.

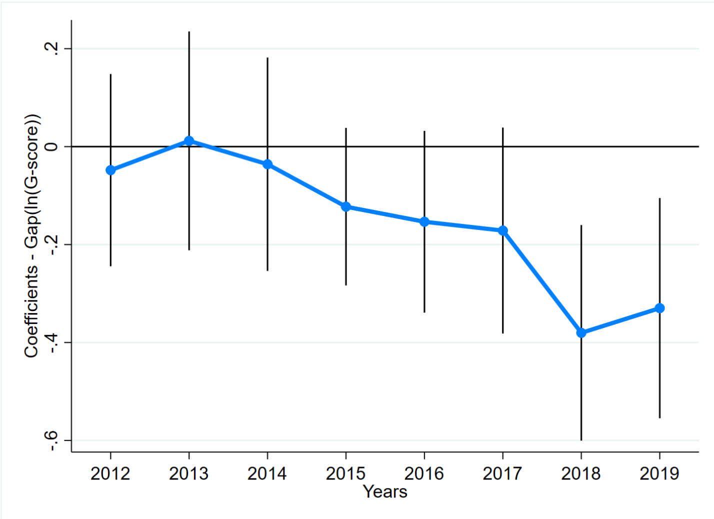
Panel A: *Gap(BG-score)*



Panel B: *Gap(B-score)*



Panel C: *Gap(G-score)*



Panel D: *Gap(E-score)*

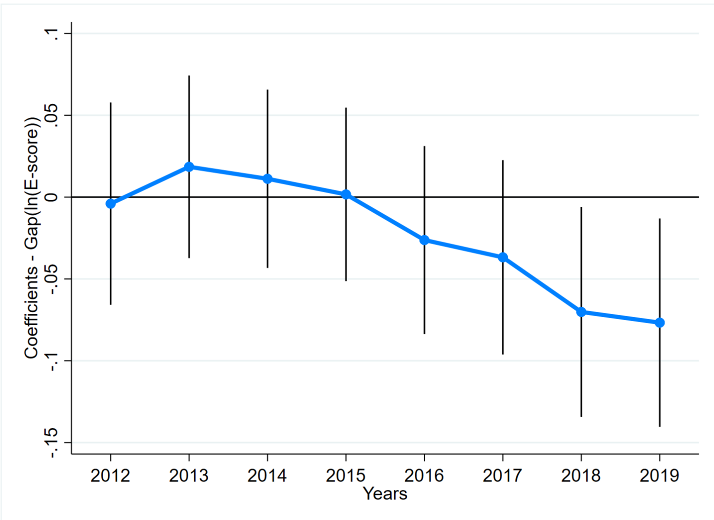


Table A.I
List of Technology Tags

General Web Technologies	**Business Web Technologies**
Content Delivery Network	Advertising
Content Management System	Analytics and Tracking
Copyright	Audio/Video Media
Domain Parking	E-commerce Shop*
Edge Delivery Network	Email
Framework	Feeds
Hosting Providers	Language
Javascript	Mobile
Mapping	Payment*
Name Server	Shipping Provider*
Operating Systems and Servers	SSL certificate*
Robots	Widgets
Verified Links	
Web Master	
Web Server	

**indicates e-commerce technologies*

Notes. This table lists the names of the technology tags obtained from BuiltWith. We manually classify all tags into *General* and *Business* web technologies, and within the business category identify separately e-commerce technologies (indicated with an asterisk). General web technologies are essential for either the construction of any website or to enable certain website functions, and they are not specific to business websites. Business web technologies are applicable only to businesses, and they change how the business operates.

Table A.II
Definitions of Variables

Variables	**Definitions**
Panel A: Variables sourced from BuiltWith	
Web	A dummy variable equal to one if firm i has a website in year t, and zero otherwise.
Ln(BG-score)	The natural logarithm of one plus the sum of all technology tags detected on firm i' s website in year t.
Ln(B-score)	The natural logarithm of one plus the sum of all "Business" technology tags detected on firm i' s website in year t.
Ln(G-score)	The natural logarithm of one plus the sum of "General" technology tags detected on firm i's website in year t.
E-commerce	A dummy variable equal to one if firm i has adopted any of the e-commerce-related technologies in year t, and zero otherwise.
Ln(E-score)	The natural logarithm of one plus the sum of all tags associated with e-commerce detected on firm i' s website in year t.
Panel B: Variables sourced from FAME	
Ln(Total Assets)	Natural logarithm of one plus the firm's total assets
Ln(Age)	Natural logarithm of one plus the firm's age, where age is the number of years from the year of incorporation.
Leverage	The sum of short-term debt and long-term liability divided by total assets.
Cash	Cash divided by total assets.
ROA	Earnings before interest and taxes (EBIT) divided by total assets.
Sales Growth	The growth rate of the annual turnover between year t and (t-1).
Δ*Ln(Employees)*	Change in the natural logarithm of the number of employees between year t and (t-1).
Δ*Ln(VPE)*	Change in the natural logarithm of one plus the firm's EBITDA scaled by the number of employees between year t and (t-1).
$\lambda(t)$	A dummy variable equal to zero if the firm is active and equal to one from the year that the firm has been in administration, in liquidation or dissolved.

Notes. This table reports the definitions of all variables. Panel A reports definitions of the measures of digitalization, the data for which was obtained from BuiltWith. Panel B reports definitions for variables sourced from FAME.

Table A.III
Descriptive Statistics Post-Program Period

Variables	Mean (1)	Median (2)	SD (3)	Mean (4)	Median (5)	SD (6)	Mean (7)	Median (8)	SD (9)	Diff. (4)-(7)
	Panel A: Before log-transformation									
	All Firms (N = 17,598)			Treated (N = 8,799)			Control (N = 8,799)			
Web	0.252	0.000	0.434	0.268	0.000	0.443	0.234	0.000	0.423	0.035***
										(11.41)
BG-score	2.511	0.000	4.852	2.718	0.000	5.020	2.290	0.000	4.656	0.428***
										(12.56)
B-score	1.118	0.000	2.279	1.218	0.000	2.365	1.102	0.000	2.178	0.205***
										(12.82)
G-score	1.393	0.000	2.662	1.501	0.000	2.745	1.278	0.000	2.565	0.223***
										(11.92)
E-commerce	0.185	0.000	0.388	0.201	0.000	0.401	0.168	0.000	0.374	0.033***
										(11.96)
E-score	0.262	0.000	0.605	0.286	0.000	0.630	0.236	0.000	0.576	0.049***
										(11.65)
	Panel B: After log-transformation									
	All Firms (N = 17,598)			Treated (N = 8,799)			Control (N = 8,799)			
Ln(BG-score)	0.575	0.000	1.025	0.617	0.000	1.054	0.529	0.000	0.992	0.088***
										(12.22)
Ln(B-score)	0.392	0.000	0.735	0.424	0.000	0.759	0.358	0.000	0.707	0.066***
										(12.76)
Ln(G-score)	0.451	0.000	0.811	0.484	0.000	0.832	0.416	0.000	0.786	0.068***
										(11.87)
Ln(E-score)	0.158	0.000	0.345	0.172	0.000	0.358	0.143	0.000	0.331	0.029***
										(11.93)

Notes. This table compares digitalization outcomes across matched treated and control firms during the post-program period (2015 - 2019). Panels A and B report digitalization outcomes before and after log-transformation, respectively. Treated (Control) are SMEs without website before the program located in the treated LEPs (control areas). *Web* is an indicator equal to one if the firm has a website in that year, zero otherwise. *BG-score* (*Ln(BG-score)*) is the (natural logarithm of one plus the) number of technology tags detected on the SMEs' website. *B-score* (*Ln(B-score)*) is the (natural logarithm of one plus the) number of "Business" technology tags detected on the SMEs' website. *G-score* (*Ln(G-score)*) is the (natural logarithm of one plus the) number of "General" technology tags detected on the SMEs' website. *E-commerce* is an indicator equal to one if the SME adopted any e-commerce technology on its website in that year and zero otherwise. *Ln(E-score)* is the natural logarithm of one plus the sum of all tags associated with e-commerce detected on the SME's website. *t*-statistics are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A.IV
Descriptive Statistics of Firm Characteristics

Variables	Mean (1)	Median (2)	SD (3)	Mean (4)	Median (5)	SD (6)	Mean (7)	Median (8)	SD (9)	Diff. (4)-(7)
Panel A: Pre-Program										
	All Firms (N = 17,598)			Treated (N = 8,799)			Control (N = 8,799)			
Sales Growth	0.059	0.022	0.337	0.059	0.024	0.330	0.060	0.020	0.343	-0.001
										(-0.19)
ROA	0.107	0.052	0.538	0.111	0.052	0.533	0.104	0.052	0.543	0.007*
										(1.68)
Website Traffic Growth	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
										(0.00)
Δ*Ln(Employees)*	0.006	0.000	0.244	0.005	0.000	0.238	0.007	0.000	0.250	-0.002
										(-1.08)
Δ*Ln(VPE)*	-0.014	0.012	0.917	-0.021	0.012	0.922	-0.007	0.011	0.912	-0.014
										(-1.54)
Panel B: Post-Program										
	All Firms (N = 17,598)			Treated (N = 8,799)			Control (N = 8,799)			
Sales Growth	0.040	0.020	0.304	0.042	0.022	0.293	0.038	0.018	0.316	0.005*
										(1.77)
ROA	0.059	0.042	0.500	0.063	0.042	0.468	0.054	0.042	0.531	0.009**
										(2.39)
Website Traffic Growth	5.143	0.227	22.286	4.894	0.219	21.447	5.490	0.234	23.403	-0.596**
										(2)
Δ*Ln(Employees)*	-0.002	0.000	0.263	0.001	0.000	0.258	-0.005	0.000	0.268	0.006**
										(2.43)
Δ*Ln(VPE)*	-0.014	0.003	0.919	-0.013	0.001	0.909	-0.015	0.004	0.930	0.002
										(0.21)

Notes. This table reports the summary statistics of the real outcomes across matched treated and control firms. Panel A and B report statistics for the pre- (2011 - 2014) and post-program period (2015 - 2019), respectively. Treated (Control) are SMEs without website before the program located in the treated LEPs (control areas). *Sales Growth* is the growth rate of the annual turnover between year t and (t-1). *ROA* is measured as earnings before interest and taxes (EBIT) divided by total assets. *Website Traffic Growth* is the growth rate of the total web traffic. We define the total web traffic as the yearly sum of both organic and paid traffic. Δ*Ln(Employees)* is the change in the natural logarithm of the number of employees between year t and (t1). Δ*Ln(VPE)* is the change in the natural logarithm of the firm's EBITDA scaled by the number of employees. *t*-statistics are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A.V
Web Traffic Growth

	Website Traffic Growth		
	(1)	(2)	(3)
$Post \times Treated$	0.866*	0.852**	0.830**
	(3.14)	(3.87)	(3.88)
Observations	122,096	122,096	120,980
Adjusted R^2	0.05	0.05	0.05
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes
Geographic FE	No	No	Yes

Note. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where $Y_{i,t}$ includes *Website Traffic Growth*, which is the growth rate of the total website traffic. The total website traffic is calculated as the sum of both organic traffic and the traffic that originates from paid Google Ads. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. *Treated* equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. $\boldsymbol{Z}$ is the matrix of fixed effects, including firm (i), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on 1-digit U.K.-SIC code. The geographic fixed effect is based on NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A.VI
Placebo Tests

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
Panel A: Placebo Sample				
$Post \times Treated\_P$	-0.001	-0.008	-0.005	-0.008
	(-0.43)	(-1.32)	(-1.22)	(-1.60)
Observations	148,889	148,222	148,222	148,222
Adjusted R^2	0.73	0.76	0.75	0.74
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Panel B: Placebo Event Year				
$Post\_P \times Treated$	0.002	0.008	0.006	0.004
	(0.44)	(0.70)	(0.79)	(0.47)
Observations	42,735	42,728	42,728	42,728
Adjusted R^2	0.67	0.67	0.65	0.67
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Note. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where $Y_{i,t}$ includes the digitalization outcomes *Web* for the extensive margin; and *Ln(BG-score)*, *Ln(B-score)*, and *Ln(G-score)* for the intensive margin. *Web* is an indicator equal to one when the firm has a website in that year and zero otherwise. *Ln(BG-score)* is the natural logarithm of one plus the number of technology tags detected on the SMEs' website. *Ln(B-score)* is the natural logarithm of one plus the number of "Business" technology tags detected on the SMEs' website. *Ln(G-score)* is the natural logarithm of one plus the number of "General" technology tags detected on the SMEs' website. In Panel A, *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. *Treated* equals one (zero) for a random subsample selected from a placebo sample similar in size to the subsample used for the baseline estimations (*Treated_P*). In Panel B, *Post* equals one from 2011 to 2012 and zero from 2009 to 2010 (*Post_P*). *Treated* equals one (zero) for a subsample of SMEs without a website before the presumed pre-period and located in the Treated (Control) areas. This table reports the baseline specification with firm (i) and year (t) fixed effects only, represented by matrix $\boldsymbol{Z}$. All regressions are estimated on a propensity score matched sample using 2013 (2010) values of age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates for estimations in Panel A (Panel B). Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A.VII
Geographical Spillover

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
$Post \times Treated$	0.074***	0.133***	0.144***	0.184***
	(6.98)	(7.56)	(7.66)	(7.51)
Observations	87,666	87,663	87,663	87,663
Adjusted R^2	0.48	0.46	0.47	0.48
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where $Y_{i,t}$ includes the digitalization outcomes *Web* for the extensive margin; and *Ln(BG-score)*, *Ln(G-score)*, and *Ln(B-score)* for the intensive margin. *Web* is an indicator equal to one when the firm has a website in that year and zero otherwise. *Ln(BG-score)* is the natural logarithm of one plus the number of technology tags detected on the SMEs' website. *Ln(B-score)* is the natural logarithm of one plus the number of "Business" technology tags detected on the SMEs' website. *Ln(G-score)* is the natural logarithm of one plus the number of "General" technology tags detected on the SMEs' website. *Post* is a dummy equal to 1 during the period 2015-2019, zero otherwise. *Treated* equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. This table reports the baseline specification with firm (i) and year (t) fixed effects only, represented by matrix $\boldsymbol{Z}$. The geographic fixed effect is based on NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates for estimations. Each treated firm is matched with one unique control firm that is within 5 miles from the boundaries of the treated LEPs. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A.VIII
Digitalization Outcomes: Financial Constraints by Employment

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
	Panel A: Few Employees			
$Post \times Treated$	0.027***	0.069***	0.048***	0.055***
	(2.66)	(2.74)	(2.68)	(2.80)
Observations	82,732	82,735	82,735	82,735
Adjusted R^2	0.45	0.45	0.44	0.45
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
	Panel B: Many Employees			
$Post \times Treated$	0.021***	0.048***	0.038***	0.035**
	(2.28)	(2.31)	(2.59)	(2.15)
Observations	83,875	83,873	83,873	83,873
Adjusted R^2	0.49	0.48	0.47	0.48
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Difference				
Few – Many	0.006	0.021	0.010	0.020*
F-stat	[1.22]	[1.79]	[0.78]	[2.80]

Notes. This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t}\ ,$$

where the model is separately estimated on subsamples of firms characterized by financial constraints. Financial constraint is defined by the firm's number of employees where 'Few Employees' ('Many Employees') are firms with total number of employees below (above) median. Panel A and B report results estimated for firms with few and many employees, respectively. Y includes the digitalization outcomes *Web*, *Ln(BG-score)*, *Ln(B-score)*, and *Ln(G-score)*. *Web* is an indicator equal to one when the firm has a website in that year, zero otherwise. *Ln(BG-score)* is the natural logarithm of one plus the number of technology tags detected on the SMEs' website. *Ln(B-score)* is the natural logarithm of one plus the number of "Business" technology tags detected on the SMEs' website. *Ln(G-score)* is the natural logarithm of one plus the number of "General" technology tags detected on the SMEs' website. *Post* is a dummy equal to 1 during the period 2015-2019, and zero otherwise. *Treated* equals one (zero) for SMEs without a website before the program and located in the *Treated (Control)* areas. This table reports the baseline specification with firm (i) and year (t) fixed effects only, represented by matrix $\boldsymbol{Z}$. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, sales growth, profitability, and 1-digit U.K.-SIC dummies as covariates. Each treated firm is matched with one unique control firm. t-statistics (in parentheses) are calculated from standard errors clustered at the geographic-year level. Chow test results for the difference between estimated coefficients across the pair subsamples are reported. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.