PROMOTING DIGITALIZATION THROUGH INFORMATION DISSEMINATION*

Tian Gao† Maria-Teresa Marchica‡ Stefan Petry§

Abstract

We exploit the heterogeneous regional allocation of funding from a U.K. government program that aimed at helping train and educate small and medium-sized enterprises on the use of digital technologies. Using a novel firm-level dataset of web technologies, our matched difference-in-differences estimations, robust to several identification strategies, show that treated SMEs were more likely to enhance their web presence, adopt more advanced web technologies, and expand their operating boundaries through e-commerce following the program. Consequently, firms observed positive effects on performance and labor outcomes. Additionally, the program enabled digitalization in geographically remote locations. *JEL* Codes: G33, J24, O15, O33, O38.

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

*We would like to thank Diana Bonfim, Murillo Campello, Mara Faccio, Thomas Flanagan (discussant), Todd Gormley, Bruce Grundy, Melanie Koch, Roni Michaely, Steven Ongena, Raghu Rau, Anna Toldrà-Simats, David Yermack, Sujiao Zhao, and conference and seminar participants at BAM Conference, EEA, FIRS, FMA Annual Meeting, HK-TLV Conference, Royal Economic Society Annual Conference, World Finance Conference, Alliance Manchester Business School, Banco de Portugal, Bayes Business School, Birmingham Business School, Bristol Business School, CUNEF Madrid, HEC Montreal, HK Polytechnic University, Liverpool Management School, Nottingham Business School, Sun Yat-sen University, University of Coimbra, Universidad de San Andrés, Universidad Torcuato Di Tella, and University of Chile for useful comments and suggestions; and Annalisa Ferrando from the European Central Bank for sharing the SAFE index. An earlier version of this paper circulated under the title "Promoting Digitalization without Subsidies".

†tian.gao-9@postgrad.manchester.ac.uk, Alliance Manchester Business School, University of Manchester

‡maria.marchica@manchester.ac.uk, Alliance Manchester Business School, University of Manchester

§stefan.petry@manchester.ac.uk, Alliance Manchester Business School, University of Manchester

"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, and several studies suggest they can have positive effects on macroeconomic growth and productivity (e.g., Syverson 2011; Brynjolfsson and McAfee 2014; Greenstein 2020). There is little firm-level evidence, however, on the beneficial effects of investing in digital technologies.[1] Furthermore, despite their potential benefits, it is not clear why particularly small- and medium-sized enterprises (SMEs) hesitate to invest in these technologies. Academic research has traditionally focused on the absence of capital as the main friction that can help explain inefficient investment decisions (e.g., Hubbard 1998; Campello, Graham, and Harvey 2010; Denis and Sibilkov 2010; Campello et al. 2011; Li 2011; Cingano, Manaresi, and Sette 2016). However, when investing in digital technologies, one of the main barriers to implementation that has received little attention is the lack of managerial awareness of such technologies, their potential benefits, and, more generally, the lack of digital skills.

Such a lack of awareness and skills is particularly prevalent among managers of SMEs as well as entrepreneurs, who lack not only hard, technical knowledge but also soft skills (e.g., professional attitude). SMEs have been identified as companies digitalizing their businesses slower than large firms in most economies, leading to the so-called *digital divide*. Since SMEs represent the backbone of most economies, such a digital divide can have negative effects on aggregate productivity and income distribution (e.g., Arendt 2008; Acemoğlu et al. 2014; Comin and Mestieri 2018; Berlingieri et al. 2020; Millán et al. 2021; OECD 2021a).[2] Surprisingly, there has not been much empirical research done to

[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 and/or implementation of e-commerce), to strategic planning (e.g., Big Data analysis), general administration and IT systems (e.g., cloud computing services), production, pre-production, and logistics (e.g., supply chain management software).

[2]Across economies, SMEs account for 99% of all businesses, between 50% and 60% of value-added, and employing the majority of the labor force (BEIS 2021; OECD 2021b).

date on the relationship between lack of awareness of new technologies and digital skills, on the one hand, and the degree of digitalization, on the other hand. In this paper, we investigate the impact of improvements in managerial awareness of digital technologies and training in digital skills among SMEs on investments in such technologies and firm performance. We exploit a U.K. government initiative that allows us to inform the debate on the link between information dissemination of digital technologies, digitalization by firms, 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 to support the provision of training and information dissemination of digital technologies to SMEs in England. Instead of subsidizing directly the firms' investments in digital technologies, the funding was distributed 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 online business opportunities.

More in detail, the BIS distributed the funding on a competitive basis following criteria such as value for money, reach, the originality of the methods proposed by LEPs to disseminate information and training on digital knowledge to the local SMEs, economies of scale, and sustainability (BIS 2015). Eventually, 22 out of 39 LEPs in England were funded. From an empirical perspective, the potential concern is that the BIS might have selected areas that already had better technological infrastructures and stronger local economies to ensure the program's success. Alternatively, there is a concern that LEPs might have self-selected into applying for funding if they believed to have a higher chance of being selected, for example, due to the strength of their local economy and/or the LEPs' boards' ability to raise external funds. This would imply that firms in treated areas (i.e., those that received funding) would show better post-program performance measures compared to firms in the control areas (i.e., those that didn't receive funding) even in the absence of treatment. A second concern might be that the composition of firms and industries across treated and control areas might be different. For instance, firms operating in industries more prone to adopt digitalization might be concentrated in areas that eventually received funding. Third and finally, SMEs in the treated areas could choose whether to participate in the events organized by their LEPs.

To ensure that our estimated coefficients are indeed causal, we address these concerns using different samples and estimation strategies. We first show via univariate analysis that treated and control areas are insignificantly different from each other with respect to several regional characteristics such as broadband access and usage, a number of metrics of economic strength, capital availability, and political affiliation in the years before the implementation of the program. This suggests that the allocation of the funding undertaken by the BIS does not seem to have been determined by regional observable characteristics that could have predicted the success of the program.

In our main empirical setting, we exploit the geographical variation of the allocation of funding and compare firms in the treated areas with those in control areas. We implement pairwise-matched difference-in-differences (DID) estimations using the years 2011-2013 (2015-2019) as the pre- (post-) program period. The model specification accounts not only for firm, year, and industry-by-year, but also for region fixed effects to further reduce the concerns that unobservable characteristics might have driven the LEPs' selection. We match each treated firm with one control firm along several dimensions (i.e., size, age, financial constraints, growth opportunities, and industry) to control for the possibility that the firm compositions in treated and control areas are different. We require both treated and control SMEs to be without internet presence before the program to isolate the firms that are more likely to lack digital awareness and skills and are more likely to benefit from the program. Since we do not have information on the actual participation of SMEs in the various activities organized by the LEPs, our tests estimate the intention-to-treat effect (ITT). That means we compare SMEs that are *more likely* to receive information and training from the program with those that are not exposed to the benefits of the program. estimated effect should be viewed as a lower bound to the treatment-on-the-treated (TOT) effect.

A novel contribution of this paper is our measure of corporate digitalization, which is available at the firm level and for the entire spectrum of firm sizes.[3] We obtain a detailed database of U.K. firms' websites and the technologies embedded therein from BuiltWith Pty Ltd (BuiltWith). BuiltWith collects information such as the type of web hosting the website uses, e-commerce functionalities, and analytics software through regular snapshots of every single website. This enables us to construct a time series of website technologies for every firm in our sample from 2011 to 2019. In our empirical analysis, we 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.

[3]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).

Across our matched sample, we count 4,244 unique technologies with a total number of 636,607 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 156,837 firm-year observations generated by 20,860 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.[4] The top three industries are manufacturing, wholesale and retail trade, human health- and social work-related industries, which account 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 3 percentage points (pp) higher compared to that of SMEs in control areas. Given an average estimated engagement rate of 34% in the *Challenge Fund* program, this represents a TOT effect of 9pp percentage points. On the intensive margin, treated SMEs employ 7% more different technologies in their websites than control SMEs, corresponding to a three times larger TOT effect, making their websites more digitally sophisticated. There has also been an increasing trend over time for treated SMEs to have a website.

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 imply that, on the extensive margin, treated firms are 8pp more likely to start an e-commerce platform on their websites than control firms (TOT effect). On the intensive margin, treated firms adopt the equivalent of 6% more e-commerce-related technologies relative to the control firms in the aftermath of the program (TOT effect). The findings show that higher sophistication of website technologies spurred through enhanced awareness and training in digital skills is associated with the adoption of e-commerce among SMEs.

We validate the results from our baseline regression estimations by confirming the existence of parallel trends in the outcome variables pre-treatment, by performing two

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

placebo tests in which we either randomize the treatment firms or randomize the timing of treatment, and by implementing the same baseline model around the borders between treated and control areas to ensure that firms are exposed to similar economic conditions. We further address the potential concern of LEPs' self-selection when applying to participate in the program by using an alternative sample and research design that focuses only on LEPs that received funding. First, we confirm that treated and control LEPs are similar along several dimensions that capture both the strength of the local economies and the ability of LEPs' boards to raise funds. Within each treated LEP, we match SMEs with non-SMEs, both of which were without internet presence before the program, based on age, financial constraints, and industry. Consistent with the main baseline results, SMEs are 6pp more likely to have a website after the program compared to non-SMEs in the same treated LEPs, corresponding to a TOT effect of approximately 18pp. Similarly, when it comes to e-commerce, treated SMEs are, on average, four percentage points more likely to set up an e-commerce business than their counterparts, which represents a TOT effect of about 12pp. The results confirm that, indeed, only firms that were targeted by the program, namely SMEs, benefited and not all firms in the same treated LEP.

Overall, the combination of all our validation tests points in the same direction, namely the implementation of the *Challenge Fund* program enables us to examine the causal link between information dissemination of digital technologies, on the one hand, and actual digitalization by firms, on the other hand.

Does higher digitalization have real effects? We examine the economic outcomes on both the performance (i.e., revenue and return on assets (ROA) growth) and labor (i.e., employment and productivity growth) dimensions and find that in the post-program period, relative to the control firms, treated SMEs witnessed on average 3% higher revenue, which is economically significant when compared with the average sales growth before the program (15%). The increased revenue also improved the growth rate of companies' return on assets, which is around 2% higher than that of the control firms. Evidence on web traffic growth further suggests that these results are likely due to an increased pool of new customers. We also find that treated SMEs experienced 2% higher employment growth rates and 5% higher value-added per employee. These results are validated by establishment-level data showing a significantly positive increase in both IT workers and IT expenses of about 6% and 12%, respectively, for treated firms relative to the control firms.

In the last part of our investigation, we discuss further potential barriers to digitalization, namely financial constraints, director age, and the headquarter location in rural areas. First, we discuss and rule out financial constraints as an alternative channel that may prevent firms from digitalizing (e.g., OECD 2021a). Since the *Challenge Fund* program funded training at the LEP level rather than directly subsidizing SMEs, our empirical setting particularly captures the knowledge and skills channel. In addition, our empirical strategy based on firm-matching compares SMEs with similar levels of financial constraints, among other characteristics. Unsurprisingly, we do not find a statistically significant difference in digitalization outcomes when we compare relatively more financially constrained SMEs with less constrained ones.

Second, 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 fewer skills tailored to technological innovations and are more exposed to organizational inertia (e.g., OECD 2021b). We show that not only treated SMEs with younger directors but also those with mature directors have been affected by the program in the likelihood of setting up a website and adopting more digital technologies. It suggests that the program successfully enhanced digital awareness even beyond certain behavioral traits.[5]

Finally, it has been observed that, mainly due to differences in human capital distribution across different geographical areas, firms in the rural part of the country digitalize at a slower pace than those located in the urban areas. Results of the sub-sample analysis show that, while all treated firms digitalized after the program at both the extensive and intensive margins, those in rural areas did so significantly more than those in urban areas. This suggests that disseminating information about digital technologies and providing training in digital skills can be an effective channel to narrow the *digital divide* at the spatial level.

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 in digital skills eases 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 making digitalization more diffuse within a country by reaching remote

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

rural areas and narrowing the digital divide with companies in urban areas.

In addition, 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, firm size, and entrepreneurial survival (for extensive reviews see Draca, Sadun, and Van Reenen 2009; Syverson 2011; Gal et al. 2019; Alhorr 2024); while others find either no effect on firm productivity or heterogeneity in the performance, firm creation, and geography effects (e.g., Acemoğlu et al. 2014; DeStefano, Kneller, and Timmis 2018; Couture et al. 2021; Hvide and Meling 2023). 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. Coupled with the pairwise-matching DID estimations around the implementation of the *Challenge Fund*, we overcome several identification issues that have affected previous studies and show that the digitalization of SMEs has, ultimately, positive effects on performance, employment and labor productivity. Finally, our paper 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 invest in such technologies and that government intervention seems to be effective in reducing such an information barrier.

A back-of-the-envelope calculation shows that the 3pp points higher revenue of the average treated SME over the average control SME corresponds to a £34,353 increase in sales. Since the overall cost of the *Challenge Fund* program was £2 million, this has 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 research design and the data. Sections IV and V discuss the digitalization and real outcomes results, respectively. Section VI provides robustness tests, and section VII concludes.

II INSTITUTIONAL DETAILS

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 organization ("Go ON UK") specializing in projects that contribute toward digital inclusions.[6] Although available nationwide, neither campaign was 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

[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 organization digitalskills.com provided access to digital skills and a forum for SMEs to submit digital-related inquiries. In addition, it offered an interactive map showing locations of digital resources, such as Wi-Fi spots (Al Harbi 2014)

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 enterprise 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 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). The proposals' summaries reveal that almost all LEPs proposed initiatives to make SMEs aware of the project, such as flyers, local press, and/or local media to promote the projects, and activities to involve SMEs in the project, such as 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]

[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. 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] For instance, Cornwall & Isles of Scilly LEP advertised the Digital Capabilities program on local buses across the major towns in Cornwall, by distributing flyers by hand to major retailers and town centers across Cornwall, and by setting up a Twitter page showcasing the training events. Further, they launched engagement events for SMEs from across Cornwall, which outlined the opportunities for developing their digital capabilities and showcased the impact of going digital on business performance. Finally, they delivered three digital high street skills training modules in each business district across Cornwall to emphasize and help SMEs develop a website design, use of email marketing, social media, Google Analytics, and Google+. Another example is the one provided by Swindon & Wiltshire LEP, which used the challenge fund to run a series of day-long workshops delivered by local ICT specialists to provide hands-on support to help businesses use ICT in innovative ways. They also developed a

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." (BIS 2015, 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 (p.16).

After a competitive bidding process, in October 2014, the BIS selected 20 projects (out of 24) 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).[11] 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 successful LEPs in South West England as well as in the East Midlands. In our methodology, we address this by using propensity score matching to identify control firms that share similar characteristics as treated firms. In the following section, we discuss the characteristics of treated and control areas in more detail and how we build our main identification strategy.

III RESEARCH DESIGN AND DATA

III.A Treatment Assignment

From an empirical perspective, the setting of the *Challenge Fund* program enables separating firms in Great Britain into areas that received funding from the program (*Treated areas*) and areas that did not receive funding (*Control areas*). One possible concern

knowledge-sharing website for additional support to provide information and user guides as well as tips and hints on ICT best practices.

[11] A potential confounding event called Digital Gateway was launched in Scotland in 2016. To avoid a potential bias, we exclude Scottish SMEs in our sample from 2017 onward. We also notice the implementation of the Small Business GREAT Ambition program in the UK in 2013, which distributed broadband vouchers to SMEs in major cities across the UK. We believe that such an initiative did not have any bearing on the effects of the *Challenge Fund* program. The initiative targeted all UK SMEs located in city areas. As such, both treated and control areas would have been affected, and the effect of information dissemination and training promoted by the *Challenge Fund* would be diminished. In the following sections, we show instead that there is a significant difference in digitalization between SMEs located in the treated and control areas. We also show that such difference is particularly relevant in rural areas rather than in urban ones.

could be that certain characteristics of the treated areas, which also determine future SME digitalization, might have made their LEPs' proposals more likely to be selected for the *Challenge Fund*. For example, the BIS may have preferred to select applications from LEPs that have a better technological infrastructure to increase the chances of the program being successful. This would imply that the results of our study were endogenously determined by the characteristics of the treated areas rather than by the increased awareness of digitalization spurred by the *Challenge Fund*.

We mitigate this concern via univariate tests in Table I, which show 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 of 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.[12] It also includes the fraction of internet users and 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), macroeconomic conditions (i.e., total employment, percentage of educated workers, gross domestic products (GDP), growth in GDP, and growth value-added), and capital availability (i.e., number of bank branches) across treated LEPs and all other areas.[13] 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,

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

[13] The percentage of educated workers is the number of workers aged between 16 and 63 with a qualification that is at least in the National Vocational Qualification (NVQ) 4 level or equivalent (e.g., bachelor degree) divided by the total workforce (ONS 2021)). NVQ levels are work-based qualifications that can be achieved through assessment or training.

Labour Party, Liberal Democrats, and all other parties. We define a local authority as one 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 coalition between the Conservative Party and Liberal Democrats 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-Liberal Democratic coalition to the number of local authorities controlled by all other parties in that area. We use the geographical areas based on LEPs for England, while the Nomenclature of Territorial Units for Statistics 2 (NUTS-2) for Wales and Scotland.[14] Differences in the averages between treated and control areas are again statistically insignificant.

(Insert Table I here)

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

III.B Identification Strategy

Given the statistically insignificant difference in characteristics between areas that did and did not receive any funding, we build our identification strategy by exploiting the geographical variation of the allocation of funding by the *Challenge Fund* program across Great Britain and by focusing 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*.

The previously observed geographical clustering of some of the treated LEPs in South West England and in the East Midlands may imply that the composition of firms in the treated and control areas is different despite the geographical areas being statistically similar, as shown in Table I. Indeed, 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

[14] NUTS is a geocode classification used by the ONS for referencing the subdivisions of the U.K. 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.

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 and real outcome 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, and growth opportunities 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.

(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)*, *Ln(B-score)*, *E-commerce*, or *Ln(E-score)*), performance outcomes (*Sales Growth* and *ROA Growth*), and labor market outcomes (Δ*Ln(Employees)* and Δ*Ln(VPE)*), as defined in the following sub-sections. $Post_t$ is a dummy variable equal to one (zero) during the 2015-2019 (2011-2013) period. $Treated_i$ equals one (zero) for SMEs without a website before the program and located in treated (control) areas. Standard errors are clustered at the geographic-year level. $Z_{i,t,j,g}$ 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 codes. The geographic fixed effect is defined at the NUTS-2 level.[15] Controlling for firm, industry-by-year, and region time-invariant unobservable effects further alleviates the concern over selection on unobservables.

The coefficient of our interest is β. While we observe when firms implemented their websites (and/or e-commerce) and changed the amount and types of website technologies, we do not have information on their actual participation in the various activities organized locally by the LEPs. As such, we are not able to explicitly account for their

[15] Using a log-linear transformation of the intensive margin digitalization measures 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.

selection into those events. Therefore, β captures the intention-to-treat (ITT) effect, i.e., we compare firms that could have been exposed to information dissemination and training in digital skills provided locally by the LEPs with firms that didn't have such exposure because they are located in areas without such initiatives. As a result, the ITT effects we report should be viewed as a lower bound to the TOT effects. TOT effects are obtained by scaling up the ITT by the engagement rate, which is approximately 34%. We define the engagement rate as the ratio of the number of SMEs that actually participated in the events (*engaged firms*) over the number of targeted SMEs, as estimated and reported by the funded LEPs in the aftermath of the program (BIS 2015, p.11).

III.C Digital Technology Variables

To implement our research design, we use two main databases: 1) BuiltWith and 2) Financial Analysis Made Easy (FAME). BuiltWith retrieves the web technologies on each firm's website since 2000. On a regular basis, BuiltWith collects data on the web technologies used on firms' websites by using website crawlers to identify and track all websites.[16] It 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 636,607 firm-technologies observations that are based on 4,244 unique technologies and 26 tags.

We exploit the richness of this novel source of digital technology data to construct several measures of digitalization at the firm level. 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 then use BuiltWith's technology tags to construct several measures that capture the intensive margin of digitalization. As explained in Section *IA.I*, we manually classify all tags into *General* and *Business* technologies, where general technologies are essential for either the construction and existence of any website, while business technologies apply only to businesses. 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*).

[16] Please see Section *IA.I* in the Internet Appendix for further details on BuiltWith's data.

Further, within the business tags, several technologies relate to e-commerce (e.g., online shops, payment, shipping providers, 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 is likely to impact the 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. This captures the extensive margin of the enlarged firm's boundary. 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.D Financial Variables

The other main database we use for our analysis is FAME, the U.K.-based product of Bureau Van Dijk, which provides detailed firmographic data, financial statements, and directors' information for a very large sample of U.K. public and private firms. To perform the PSM estimations, we use the following variables: *Ln(Total Assets)* as a proxy for firm size measured as the natural logarithm of one plus 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; *Employment* measured as the natural logarithm of one plus the total number of employees; and *Turnover* measured as the natural logarithm of one plus turnover.

As for the real outcomes tests, we use *Sales Growth* as the growth rate of annual turnover and the growth rate of ROA measured as earnings before interest and tax over total assets (*ROA Growth*). 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 amortization (EBITDA) scaled by the number of employees (Δ*Ln(VPE)*), respectively. All variables are winsorized at the 1% level to remove outliers. Table IA.II reports the definitions of all variables.

Finally, when measuring financial constraints, in addition to using firm size and leverage as proxies, we measure the degree of financial constraints using the European Central Bank (ECB) Survey on Access to Finance of Enterprises (SAFE) Index. Section

IA.II in the Internet Appendix reports details on the construction of such an index.

III.E Sample Construction

Our sample period covers the years 2011-2019, spanning around the launch of the *Challenge Fund* in 2014. We exclude 2014 as the event year. 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.[17] 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 to 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 *Challenge Fund* targeted SMEs with no prior digital knowledge. As such, our sample includes companies without a website before the launch of the *Challenge Fund* as a proxy for the lack of digital knowledge, and we track those companies over the post-program period. 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 a sample.[18] This leaves us with a preliminary sample of 27,696 SMEs *without a website before* the start of the program. After implementing the PSM mentioned above, the final sample includes 10,430 treated and an equal number of matched control firms.

The *Ln(Total Assets)* and *Ln(Turnover)* of the average (median) treated firm is 7.2 (8.1) and 7.3 (7.9), respectively, which corresponds to $\approx$ £12mil (£3mil) in total assets and £6.8mil (£2.9mil) in turnover, respectively. The average (median) treated firm is 16 (11) years old, employing about 47 (22) workers (Panel B in Table II). Treated firms operate in 15 different sectors: 12.17% of the sample is from manufacturing, followed by human health and social work activities (10.51%), education (10.33%), administrative and support activities (10.47%) and wholesale and retail trade (10.18%).

In the aftermath of the program, treated firms show on average a higher digitalization, with

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

[18] 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.

significantly more treated firms setting up a website and e-commerce and adopting a larger number of different types of technologies (Table IA.III). This 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) in the following Section.

IV DIGITALIZATION OUTCOMES

IV.A 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)), the estimated ITT effect of treated SMEs to have a website after the *Challenge Fund* is about 3 pp, which translates into a TOT effect of about 8 pp. Even after controlling for potential time-varying industry shocks or area-specific shocks that may confound with the effect of the *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)* suggest that the website set-up by treated SMEs involves a wider range of technologies than that of all other SMEs. The estimated ITT effect suggests that treated SMEs employ almost 7% more different technologies in their websites relative to the control firms (column (4)). That corresponds to a TOT effect of 21%. We further delve into the types of newly adopted technologies to analyze 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. For instance, treated SMEs seem to adopt about 5% more business-related and general technologies than the control SMEs based on the estimated ITT effects that correspond to an average TOT of 15%.

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 average digitalization outcomes of the control firms in the post-*Challenge Fund* period as a benchmark, as reported in Panel A of Table IA.III. For example, the unconditional probability of the control SMEs to have a website after the program is 25%. Therefore, the estimated ITT effect of approximately 3 pp of column (1) in Table III implies that the average treated SME is approximately 12% more likely to have a web presence relative to the control SMEs (0.03/0.25*100 = 12%). Similarly, for the intensive margin outcomes, the estimated ITT of*Ln(B-score)*, 0.05, implies that treated SMEs employ, on average, 16% more different technologies related to business in their websites compared with their counterparts (0.05/0.32*100 $\approx$ 16%).

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.B 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 a 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 pp more likely to adopt e-commerce technologies on their website than the control SMEs (ITT effect). Such estimated ITT effect corresponds to an 8 pp TOT effect, and, in economic terms, it implies that the average treated SME is approximately 20% more likely to expand its boundaries with e-commerce relative to the control SMEs ($0.03/0.15*100 \approx 20\%$). On the intensive margin, the estimated ITT effect of 0.02 is equivalent to a TOT effect of 6% (column (4)). This suggests that treated SMEs adopt 15% more e-commerce technologies ($0.02/0.13*100 \approx 15\%$) 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, expanding the firm's boundaries.[19]

(Insert Table IV here)

IV.C Parallel Trends Condition, Placebo Tests, and Local Effect

To corroborate the validity of our identification strategy, we perform several different tests. First, we confirm that the parallel trends 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 IA.I in the Internet Appendix shows the estimated coefficients of the PSM-DID yearly interactions for *Web*, *Ln(BG-score)*, *Ln(B-score)*, and *Ln(G-score)*. 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 for both extensive and intensive margins, suggesting that the benefits from increased awareness and knowledge of new digital skills tend to build gradually over time.

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

[19]In a robustness test, we estimate the model on digital outcomes on subsamples of firms based on their target customers. We distinguish Business-to-Business (B2B) firms that sell products and services to other businesses from the Business-to-Customers (B2C) firms that sell products and services directly to consumers. We find that after the program, both types of firms digitalized, although B2C firms did so significantly more than B2B ones (Table IA.VIII in the Internet Appendix).

in FAME. We randomly assign the treatment to half of those, 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 IA.IV in the Internet Appendix reports the results. Second, we construct a new sample with 2011 as the placebo event year, requiring firms not to have a website in the years before the presumed treatment. The years 2012 and 2013 are defined as the post-event period. We then re-run the same baseline specifications using this identification strategy. Panel B shows the results of these tests. As expected, across all tests in both panels, the estimated coefficients are statistically insignificant, confirming the validity of our identification strategy and, therefore, the causal relation between the implementation of the *Challenge Fund* program and the digitalization of SMEs.

Since our identification strategy also relies on geographical boundaries, a concern could be whether local area characteristics have changed differently over time and whether such changes have determined the differential effects on digitalization and real economic outcomes. Further, 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 aware of such initiatives. To address both concerns, we re-estimate our baseline digitalization models using as control firms those that are located relatively close to the borders of areas covered by the treated LEPs. Focusing on areas close to the borders might help keep local characteristics constant over time. Further, if the *Challenge Fund* program also promoted digitalization for the SMEs in the neighborhood areas, we should detect either a small or no significant difference between the treated and the new control SMEs. Table IA.V in the Internet 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 local area characteristics did not affect the treatment assignment and that the *Challenge Fund* program had no significant spillover effects on the neighborhood areas.

Finally, we control whether there is a "London effect" on the results. When we exclude the London area, results remain unchanged (Table IA.VI).

IV.D LEPs' Self-selection into the Challenge Fund program

One further concern with the setting of the program might be that not all LEPs *chose* to submit their bid to BIS. In general, the priorities of most LEPs before the start of the program were mainly related to providing support to businesses in terms of skills and employment (Peck et al. 2013). Nonetheless, it might be that the LEPs that submitted their bids were of a better quality than those that did not. One could argue that LEPs' quality is the reflection not only of the LEPs' boards' ability to raise funds and their political affiliation but also of the local economic strength and growth opportunities. Similar to what we posit above, LEPs that *chose* to submit their bid might have been those that oversaw areas with already better technological infrastructure, stronger local economies, closer local political affiliation with the central government, and/or more skillful boards to increase the chances of success of their bid.

The implication of such a potential self-selection issue might be that the results on digitalization and real outcomes were endogenously determined by the characteristics of those LEPs that submitted and won the bid rather than by the increased awareness of digitalization and digital training promoted by the *Challenge Fund*.

To mitigate such concern, we first compare the LEPs that submitted their bids (and won) with those LEPs that didn't receive any funding.[20] Similar to Table I, we compare characteristics that relate to internet/digital infrastructure (Panel A) and business demographics and economic conditions (Panel B) of the areas under LEPs' jurisdiction. In addition, we look at the characteristics of LEPs' resources and board structures (Panel C). In detail, we compare the amount of public funding that LEPs were able to raise, the size of the boards that take LEPs' decisions, and characteristics of directors seated on LEPs' boards such as their age and government affiliation as possible determinants of LEPs' ability/strength to submit (and win) public bids.[21]

Table IA.VII shows that before the implementation of the program, LEPs that received funding from the program were not statistically different from those that didn't receive any funding, along dimensions that possibly predict the likelihood to submit and win a bid.

To further address the self-selection issue, we implement an alternative identification strategy focusing only on companies located within those areas where LEPs received funding. In particular, we compare SMEs without a website before the implementation of the program (*Treated2*) with all other companies located in the same LEPs and without a

[20] Detailed information about all bidders was not retrievable, possibly due to the fact that BIS undertook a major restructuring by merging with the Department of Energy and Climate Change in 2016 to create the Department for Business, Energy & Industrial Strategy until 2023 when it was split into three separate Departments.

[21] Public funding allocated to LEPs refers to the various funding rounds that the government launched to distribute funds on a competitive basis to support local economic growth (*Regional Growth Funds*).

website before the program, matched along with their age, financial constraints, and industry.[22] Next, using this pairwise matched sample, we implement the DID estimation of digital outcomes as in Equation (1).

Although there is an important drop in the number of observations, Table V shows that across all specifications and outcome variables, results are similar to those reported in Tables III and IV. For instance, treated SMEs are 6 pp and 4 pp more likely to have a website and implement e-commerce, respectively, than their counterparts within the same local area (estimated ITT effects in column (1) of Panels A and B, respectively). Those represent a TOT effect of 18 pp and 12 pp, respectively.[23]

More importantly, it further confirms that the program was indeed effective in disseminating information on digital technologies and training in digital skills among the local SMEs.

(Insert Table V here)

Overall, although none of our tests separately can fully address potential concerns with our research setting, nonetheless, all together point to the same direction. That is, the implementation of the *Challenge Fund* program has created an interesting 'quasi-exogenous' change in the provision of information and training in digital technologies for SMEs that allows us to tease out a causal relationship between information dissemination of digital technologies, on the one hand, and actual digitalization by firms, on the other hand. In the following section, we are going to investigate whether such enhanced digitalization has any real outcome effect.

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

[22] Firm size and turnover cannot be included among the PSM covariates here as they are the criteria used to define a company as SME.

[23] We also validate the parallel trends condition of such identification strategy by estimating a year-by-year treatment effects model around the implementation of the program. Figure IA.II in the Internet Appendix shows the estimated coefficients of the PSM-DID yearly interactions for all digital outcomes. Patterns are similar to those reported in Figure IA.I for the baseline strategy.

on digitalization 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 the growth of revenue (*Sales Growth*) and profitability (*ROA Growth*).

Table VI Panel A shows that in the aftermath of the program, sales of treated SMEs grew by approximately 3 pp more than those of control SMEs (estimated ITT effect in column (1)). This corresponds to a 9 pp TOT effect and, economically, is equivalent to an increase of 21% compared to the average growth rate of sales of control SMEs during the pre-program period, 14%. Further, Panel B shows that treated firms were more able to efficiently convert sales into profits with *ROA* growing on average by 2 pp more than the control group across all specifications.

Panels A and B in Figure IA.III in the Internet Appendix also show the estimated coefficients of the PSM-DID yearly interactions for both the performance outcomes, validating the parallel trends condition for both *Sales Growth* and *ROA Growth* estimates.

(Insert Table VI here)

One way for SMEs to increase their sales is to increase their pool of customers. Indeed, higher digitalization in general, and e-commerce in particular, would enable a company to reach customers away from the local area. This could, in turn, translate into more sales. To investigate this channel, we look at the website traffic. If sales of treated SMEs benefited more from an enlarged pool of customers, their websites should experience higher traffic following the *Challenge Fund* compared to that of control firms.

We retrieve the website traffic data from SEMrush, which collects minute-based click-stream data and Google search results of the 500 million most popular keywords from third-party data providers.[24] 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.

Table IA.IX in the Internet Appendix shows that the growth rate of total web traffic of treated SMEs is higher than that of the control group by 400 pp (column (1)). This difference in website traffic growth is also economically significant when compared with

[24]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 first use 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.

the 258% average website traffic growth rate of our sample firms during the post-program period.[25]

Overall, this result highlights a potential mechanism that explains the evidence reported in Table VI. It also adds to the evidence provided by (Armstrong, Konchitchki, and Zhang 2023), who document the economic implications of firms' digital traffic for financial performance, among others.

V.B Employment and Productivity

There is a long-standing debate about the effects of technology adoption on the labor market. On the one hand, new technologies replace labor through automation, which has 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 VII 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 estimated ITT effect of the rate of employment growth of treated SMEs relative to the control firms is approximately 2 pp and statistically significant across all specifications. This translates into a 6 pp TOT effect. 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 around 1%.

In line with the growth rate of employment, results in Panel B report that employees of treated SMEs, on average, also became more productive than those of the control group after the program. The treated SMEs' value-added per employee rate is at least 5 pp higher than

[25] Although the sales growth rate appears very small compared to the website traffic one, it seems to mirror the average global e-commerce conversion rate (i.e., number of orders divided by number of visits), which in 2022 was approximately 1.8% (https://www.invespcro.com/blog/the-average-website-conversion-rate-by-industry).

that of the control SMEs across all specifications (estimated ITT effect). This translates into a 15 pp TOT effect, on average. 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%.

More importantly, Panels C and D in Figure IA.III in the Internet Appendix show that parallel trends condition also holds for labor outcomes, similarly to the performance outcomes, confirming that there are no pre-trends in either of the real outcomes we investigate.

(Insert Table VII here)

When combined, the above evidence suggests 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 and improved labor productivity (Table VII). 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.[26] 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 on 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

[26]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).

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:

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

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 $Ln(IT\ Expenses)$).[27] *Post* is a dummy equal to 1 (zero) in the 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 VIII shows that, regardless of the specification, there is an increase in both IT workers and IT expenses in the post-*Challenge Fund* period. The estimated ITT effects suggest that relative to the control establishments, the number of IT workers increased, on average, by 6 pp in treated establishments (column (1)), while the expenses for IT items increased on average by about 12 pp more (column (4)).

Overall, these results deepen our understanding of the mechanisms behind the evidence reported in Table VII. They offer further support for the effectiveness of the program in digitalizing SMEs and also provide external validity of our previous results at the firm level.

(Insert Table VIII here)

VI BARRIERS TO DIGITALIZATION

VI.A Financial Constraints

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

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

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. 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 four different subsample pairs of firms to verify whether the information and skills channel is more relevant than the financial capital constraints one for SMEs' digitalization.

We first employ firm size as a commonly used proxy for financial constraints. In detail, 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 the median of total assets distribution) and less financially constrained (above the median of total assets distribution) firms. Second, we use the number of employees as a proxy for firm size, and we split the sample between more financially constrained (below the median of number of employees distribution) and less financially constrained (above the median of number of employees distribution) firms. Third, we use leverage in 2013 as a proxy for access to debt and identify financially constrained (below the median of the leverage distribution) and less financially constrained (above the median of the leverage distribution) firms. Fourth, we use the ECB SAFE Index for the U.K. to identify firms that, before the implementation of the program, are likely to be more financially constrained (above the top 12% of the SAFE Index distribution) from those less constrained (below the top 12% of the SAFE Index distribution).

Table IX reports the results of the baseline models as in Tables III and IV with a full set of fixed effects across the four pair subsamples. Across all Panels, 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.

(Insert Table IX here)

VI.B Directors' Age

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 traits among treated SMEs, those with younger managers should be more likely to benefit from the 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 86,357 directors in our entire sample, with the median firm employing about six directors. We split the sample into firms with younger directors, where at least one director is under the age of 55, and firms with more mature directors, where the age of all directors is above 55.

Results in Table X indicate that the estimated ITT effects on digitalization are 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. Further, tests of differences across the estimated coefficients indicate that such an effect is similar for both firms with young and mature directors.

Overall, these results 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 behavioral traits.

(Insert Table X here)

VI.C Spatial Digital Divide

Over the past decade, it has been observed a digital divide at the spatial level, that is, firms in rural areas lag behind in adopting advanced digital technologies compared to firms in urban areas (e.g., Thonipara et al. 2022). 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 a digital divide has been to particularly focus the allocation of government resources on the provision of training and education (Wielicki and Arendt 2010).

If the channel of enhancing digital awareness and training is the one that indeed spurs digitalization, we should observe that after the *Challenge Fund* program, SMEs in rural areas should be able to digitalized at least equally to or more than SMEs in urban areas and, as such, help narrow the geographical digital divide.

To investigate that, we split the sample along two alternative definitions of spatial divide. First, we sort our SME 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.[28] For the second definition of spatial divide we split the SME 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*.[29] We then test our baseline digitalization models as in Equation

[28]The top ten largest cities are identified using the first part of their postcodes, and they are London, Birmingham, Sheffield, Manchester, Nottingham, Newcastle-upon-Thyme, Cardiff, Leicester, and Bristol. Data retrieved from: (https://www.nomisweb.co.uk/query/select/getdatasetbytheme.asp).

[29]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 university areas are based in London, Oxford, Cranfield, Cambridge, Edinburgh, Glasgow, Manchester, Nottingham, Southampton, and Warwick.

(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 XI reports that in both urban and rural areas, treated SMEs have become significantly more digitalized at the extensive and intensive margins (Panel A). The differences between the estimated ITT effects across the pairs of subsamples are statistically significant from each other, suggesting that after the program, digitalization occurred more in the rural rather than the urban areas. Similarly, in university areas and all other areas, treated SMEs have become more digitalized at the extensive and intensive margins (Panel B). The differences between the estimated ITT effects across the pairs of subsamples here are, nonetheless, statistically insignificant from each other.

Overall, 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. More importantly, the results speak to the relevance of disseminating information about digital technologies and providing training in digital skills as an effective way to spur digitalization.[30]

(Insert Table XI here)

VII CONCLUDING REMARKS

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 indicates 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 granular dataset of web technologies that, to the best of our knowledge, has not been used in the literature, so far. It allows us to assemble several types of technologies at the firm level every year over the 2011-2019 period. We classify those technologies into "general" and "business". Within the business category, we further

[30] Interestingly, the *Challenge Fund* program also had a positive effect on the highly debated issue of the digital divide between large companies and SMEs, which characterizes most of the economies worldwide. Details on the investigation of this matter and its results are included in Section *IA.V* of the Internet Appendix.

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 invest in more sophisticated e-commerce technologies. Importantly, the increased digitalization has significant and positive effects on revenue and profitability, and employment growth and labor productivity. Our results, robust to several falsification tests and different sample constructions, seem to be mainly driven by the information and skills channel, while financial constraints and directors' age 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 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 a barrier.

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/.

Alhorr, Layane. "Virtual Windows Through Glass Walls? Digitalization for Low-Mobility Female Entrepreneurs." Policy Research Working Paper No. 10803, World Bank Group, 2024.

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, 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.

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, Erasmo Giambona, John R Graham, and Campbell R Harvey. “Liquidity management and corporate investment during a financial crisis.” *The review of financial studies* 24, no. 6 (2011): 1944–1979.

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.

Cingano, Federico, Francesco Manaresi, and Enrico Sette. "Does credit crunch investment down? New evidence on the real effects of the bank-lending channel." *The Review of Financial Studies* 29, no. 10 (2016): 2737–2773.

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.

Couture, Victor, Benjamin Faber, Yizhen Gu, and Lizhi Liu. "Connecting the Countryside via E-Commerce: Evidence from China." *American Economic Review: Insights* 3, no. 1 (March 2021): 35–50. https://www.aeaweb.org/articles?id=10.1257/aeri.20190382.

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.

Denis, David J, and Valeriy Sibilkov. "Financial constraints, investment, and the value of cash holdings." *The Review of Financial Studies* 23, no. 1 (2010): 247–269.

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.

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.

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.

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.

Hubbard, R. Glenn. "Capital-Market Imperfections and Investment." *Journal of Economic Literature* 36, no. 1 (1998): 193–225.

Hvide, Hans K., and Tom Meling. "New Technology and Business Dynamics." Fisher College of Business Working Paper No. 2023-03-019, Ohio State, 2023.

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.

Li, Dongmei. "Financial constraints, R&D investment, and stock returns." *The Review of Financial Studies* 24, no. 9 (2011): 2974–3007.

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.

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

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

Peck, Frank, Stephen Connolly, Jonathan Durnin, and Keith Jackson. "Prospects for 'place-based' industrial policy in England: The role of Local Enterprise Partnerships." *Local Economy* 28, nos. 7-8 (2013): 828–841.

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.

TABLES

Table I
Pre-Program Characteristics between Treated and Control Areas

This table compares several macro characteristics across treated (Column (1)) and control areas (Column (2)) in the pre-program period from 2011 to 2013. 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 demographics 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. *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. *Number of Bank Branches* is the number of bank branches per 100 firms in each area. Panel C reports the political affinity in the treated and control areas. *#Conservative-LibDem LAs / #Other Parties LAs* is the ratio of local authorities controlled by the Conservative Party and Liberal Democrats coalition 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.

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	1,059	1,292	-233
			(-0.55)

(continued on next page)

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 demographics and economic conditions			
Business Population	43,691	43,122	569
			(0.0765)
Number of Business Births	5,237	5,704	-467
			(-0.32)
Number of Business Deaths	5,193	5,213	-20
			(-0.02)
%Businesses survived aft 1 yr	0.93	0.93	-0.00
			(-0.38)
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)
Number of Bank Branches	0.546	0.584	-0.038
			(-1.244)
Panel C: Political Affinity			
$\frac{\#\text{Conservative-LibDem LAs}}{\#\text{Other Parties LAs}}$	2.584	1.317	1.267
			(1.58)

Table II
Sample Characteristics Before and After the PSM

This table reports firm-level characteristics of treated (Columns (1 - 3)) and control firms (Columns (4 - 6)) in 2013 before (Panel A) and after (Panel B) propensity score matching (PSM). The last column 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. *Ln(Turnover)* is the natural logarithm of one plus the annual turnover. *Ln(Employment)* is the natural logarithm of one plus aggregate employment. *p-score* is the probability of being treated estimated via a probit model using 2013 values of *Ln(Total Assets)*, *Ln(Age)*, *Leverage*, *Cash*, *Ln(Turnover)*, *Ln(Employment)*, and 1-digit U.K.-SIC dummies as covariates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Before PSM

	Treated Areas (N = 10,909)			Control Areas (N =16,787)			
	Mean	Median	SD	Mean	Median	SD	Diff
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(1)-(4)
Ln(Total Assets)	7.157	8.005	2.472	7.062	7.791	2.673	0.095***
							(2.99)
Ln(Age)	2.410	2.485	0.973	2.405	2.398	0.897	0.005
							(0.47)
Leverage	0.228	0.104	0.275	0.236	0.090	0.294	-0.008**
							(-2.26)
Cash	0.260	0.107	0.316	0.277	0.119	0.326	-0.016***
							(-4.11)
Ln(Turnover)	7.218	7.833	2.319	6.761	7.291	2.659	0.458***
							(14.70)
Ln(Employment)	2.961	3.091	1.478	2.649	2.485	1.476	0.312***
							(17.20)
p-score	0.410	0.418	0.072	0.385	0.394	0.081	0.026***
							(26.23)

Panel B: After PSM

	Treated Areas (N = 10,430)			Control Areas (N =10,430)			
	Mean	Median	SD	Mean	Median	SD	Diff
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(1)-(4)
Ln(Total Assets)	7.205	8.050	2.441	7.219	8.025	2.510	-0.014
							(-0.41)
Ln(Age)	2.415	2.485	0.977	2.434	2.485	0.904	-0.019
							(-1.49)
Leverage	0.227	0.103	0.275	0.231	0.102	0.281	-0.004
							(-1.10)
Cash	0.261	0.107	0.316	0.264	0.118	0.312	-0.003
							(-0.75)
Ln(Turnover)	7.266	7.901	2.294	7.314	8.000	2.401	-0.048
							(-1.47)
Ln(Employment)	2.992	3.135	1.468	2.997	3.135	1.473	-0.005
							(-0.25)
p-score	0.409	0.417	0.071	0.410	0.417	0.071	-0.000
							(-0.10)

Table III
Digitalization Outcomes: Baseline Results

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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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.

	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)
$Post \times Treated$	0.027***	0.026***	0.025***	0.067***	0.065***	0.063***	0.049***	0.048***	0.046***	0.052***	0.051***	0.048***
	(2.21)	(2.54)	(2.44)	(2.32)	(2.66)	(2.58)	(2.39)	(2.73)	(2.67)	(2.31)	(2.64)	(2.56)
Observations	148,645	148,645	147,096	149,293	147,293	147,739	149,293	149,293	147,739	148,293	148,293	147,739
Adjusted R^2	0.51	0.53	0.53	0.52	0.53	0.53	0.50	0.52	0.52	0.51	0.53	0.53
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

Table IV
Digitalization Outcomes: E-commerce

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. $\boldsymbol{Z}$ 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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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.

	Extensive Margin			Intensive Margin		
	E-commerce			*Ln(E-score)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Post × *Treated*	0.026***	0.025***	0.025***	0.022***	0.022***	0.021***
	(2.45)	(2.75)	(2.71)	(2.39)	(2.64)	(2.62)
Observations	149,293	149,293	147,739	149,293	149,293	147,739
adjusted R^2	0.44	0.46	0.46	0.44	0.46	0.46
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

Table V
Digitalization Outcomes within LEPs

This table shows the results from the model estimation:

$$Y_{i,t} = \beta(Post_t \times Treated2_i) + \mathbf{Z_{i,t,j,g}} + \epsilon_{i,t} \ ,$$

where $Y_{i,t}$ includes the digitalization outcomes *Web*, *Ln(BG-score)*, *Ln(B-score)*, *Ln(G-score)* in Panel A, and *E-commerce* and *Ln(E-score)* in Panel B. *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. *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. *Treated2* equals one (zero) for SMEs (all other companies) without a website before the program and located within the LEPs that received funding, i.e., *Treated* LEPs. $\boldsymbol{Z}$ 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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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.

Panel A: Baseline Digitalization Outcomes

	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)
$Post \times Treated2$	0.059***	0.057***	0.057***	0.131***	0.127***	0.127***	0.089***	0.087***	0.087***	0.101***	0.099***	0.099***
	(4.28)	(4.19)	(4.18)	(4.01)	(3.95)	(3.95)	(3.28)	(3.76)	(3.76)	(3.95)	(3.89)	(3.89)
Observations	22,360	22,360	22,360	22,431	22,431	22,431	22,431	22,431	22,431	22,431	22,431	22,431
Adjusted R^2	0.52	0.53	0.53	0.53	0.53	0.53	0.51	0.52	0.52	0.53	0.53	0.53
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

(continued on next page)

Table V
Digitalization Outcomes within LEPs (cont'd)

Panel B: E-commerce Outcomes						
	Extensive Margin			Intensive Margin		
	E-commerce			*Ln(E-score)*		
	(1)	(2)	(3)	(4)	(5)	(6)
$Post \times Treated2$	0.041***	0.040***	0.040***	0.033***	0.032***	0.032***
	(3.59)	(3.55)	(3.55)	(3.18)	(3.15)	(3.15)
Observations	22,431	22,431	22,431	22,431	22,431	22,431
adjusted R^2	0.45	0.46	0.46	0.45	0.46	0.46
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry $\times$ Year FE	No	Yes	Yes	No	Yes	Yes
Geographic FE	No	No	Yes	No	No	Yes

Table VI
Sales and ROA Growth

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 growth* (Panel B). *Sales Growth* is the growth rate of the annual turnover. *ROA growth* is the growth rate of ROA 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. $\boldsymbol{Z}$ 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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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.

	(1)	(2)	(3)
	Panel A: *Sales Growth*		
$Post \times Treated$	0.029**	0.027*	0.029**
	(2.04)	(1.95)	(2.10)
Observations	107,581	107,581	106,463
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
	Panel B: *ROA Growth*		
$Post \times Treated$	0.021**	0.020**	0.021**
	(2.26)	(2.18)	(2.29)
Observations	109,491	109,491	108,347
Adjusted R^2	0.40	0.38	0.39
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes
Geographic FE	No	No	Yes

Table VII
Employment and Labor Productivity Growth

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). 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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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.

	(1)	(2)	(3)
	Panel A: Δ*Ln(Employees)*		
$Post \times Treated$	0.020***	0.020***	0.020***
	(5.02)	(4.98)	(5.01)
Observations	106,468	106,468	105,402
Adjusted R^2	0.04	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
	Panel B: Δ*Ln(VPE)*		
$Post \times Treated$	0.051**	0.055***	0.057***
	(2.43)	(2.62)	(2.71)
Observations	76,025	76,025	75,288
Adjusted R^2	0.06	0.06	0.06
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry × Year FE	No	Yes	Yes
Geographic FE	No	No	Yes

Table VIII
Establishment-level IT Outcomes

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 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 codes. The geographic fixed effect is based on the 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.

	$Ln(IT\ Staff)$			$Ln(IT\ Expenses)$		
	(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

Table IX
Digitalization Outcomes:
Financial Constraints

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 constraints are defined by firm *Ln(Total Assets)* in 2013 where 'Small Firms' ('Medium Firms') are those with total assets below (above) median (Panel A); by firms' number of employees in 2013 where 'Few Employees' ('Many Employees') are firms with total number of employees below (above) median (Panel B); by firm *Leverage* in 2013 where 'Low-Leverage Firms' ('High-Leverage Firms') are those with leverage below (above) median (Panel C); by the SAFE Index where 'More Constrained' ('Less Constrained') are those above (below) the top 12% of the ECB SAFE Index distribution (Panel D). $Y_{i,t}$ 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. $\boldsymbol{Z}$ 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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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 are reported at the bottom of each sub-panel. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*	
	(1)	(2)	(3)	(4)	
	Reported Coefficient: *Post* × *Treated*				Observations
Panel A: Firm Size					
Small Firms	0.027***	0.064***	0.047***	0.048***	70,456
	(2.65)	(2.62)	(2.72)	(2.53)	
Medium Firms	0.022***	0.060***	0.045***	0.047***	76,640
	(2.10)	(2.38)	(2.43)	(2.44)	
Diff. (Small – Medium)	0.005	0.004	0.002	0.001	
F-stat	[0.70]	[0.06]	[0.02]	[0.01]	
Panel B: Number of Employees					
Few Employees	0.022**	0.056***	0.040***	0.045***	71,642
	(2.43)	(2.63)	(2.67)	(2.65)	
Many Employees	0.027**	0.068***	0.052***	0.051**	76,097
	(2.52)	(2.62)	(2.73)	(2.57)	
Diff. (Few – Many)	-0.005	-0.012	-0.012	-0.006	
F-stat	[0.91]	[0.67]	[1.03]	[0.36]	
Firm FE	Yes	Yes	Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	
Industry × Year FE	Yes	Yes	Yes	Yes	
Geographic FE	Yes	Yes	Yes	Yes	

(continued on next page)

Table IX
Digitalization Outcomes:
Financial Constraints (cont'd)

	Web (1)	*Ln(BG-score)* (2)	*Ln(B-score)* (3)	*Ln(G-score)* (4)	
	Reported Coefficient: $Post \times Treated$				Observations
Panel C: Leverage					
Low-Leverage Firms	0.024**	0.057***	0.041***	0.045***	72,706
	(2.58)	(2.65)	(2.64)	(2.66)	
High-Leverage Firms	0.025**	0.065**	0.049**	0.049**	75,033
	(2.17)	(2.36)	(2.50)	(2.31)	
Diff. (Low – High)	-0.001	-0.008	-0.008	-0.004	
F-stat	[0.03]	[0.16]	[0.35]	[0.08]	
Panel D: SAFE Index					
More Constrained	0.036***	0.089***	0.068***	0.067***	17,360
	(3.64)	(3.78)	(4.04)	(3.63)	
Less Constrained	0.023**	0.058**	0.043**	0.045**	130,378
	(2.12)	(2.26)	(2.31)	(2.25)	
Diff. (More – Less)	0.013	0.031	0.025	0.022	
F-stat	[1.78]	[1.59]	[2.21]	[1.33]	
Firm FE	Yes	Yes	Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	
Industry × Year FE	Yes	Yes	Yes	Yes	
Geographic FE	Yes	Yes	Yes	Yes	

Table X
Digitalization Outcomes:
Directors' Age

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 directors' age. Directors' age is based on the age of the 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. $Y_{i,t}$ 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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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 directors' age are reported at the bottom of each sub-panel. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*	
	(1)	(2)	(3)	(4)	
	Reported Coefficient: $Post \times Treated$				Observations
Young Directors	0.028**	0.071**	0.052**	0.054**	95,415
	(2.12)	(2.20)	(2.29)	(2.16)	
Mature Directors	0.014**	0.036***	0.029***	0.027***	49,167
	(2.09)	(2.34)	(2.53)	(2.29)	
Diff. (Young – Mature)	0.014	0.034	0.023	0.027	
F-stat	[1.94]	[1.96]	[1.93]	[1.88]	
Firm FE	Yes	Yes	Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	
Industry × Year FE	Yes	Yes	Yes	Yes	
Geographic FE	Yes	Yes	Yes	Yes	

Table XI
Digitalization across Geographical Areas

This table shows the results from the model estimation:

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

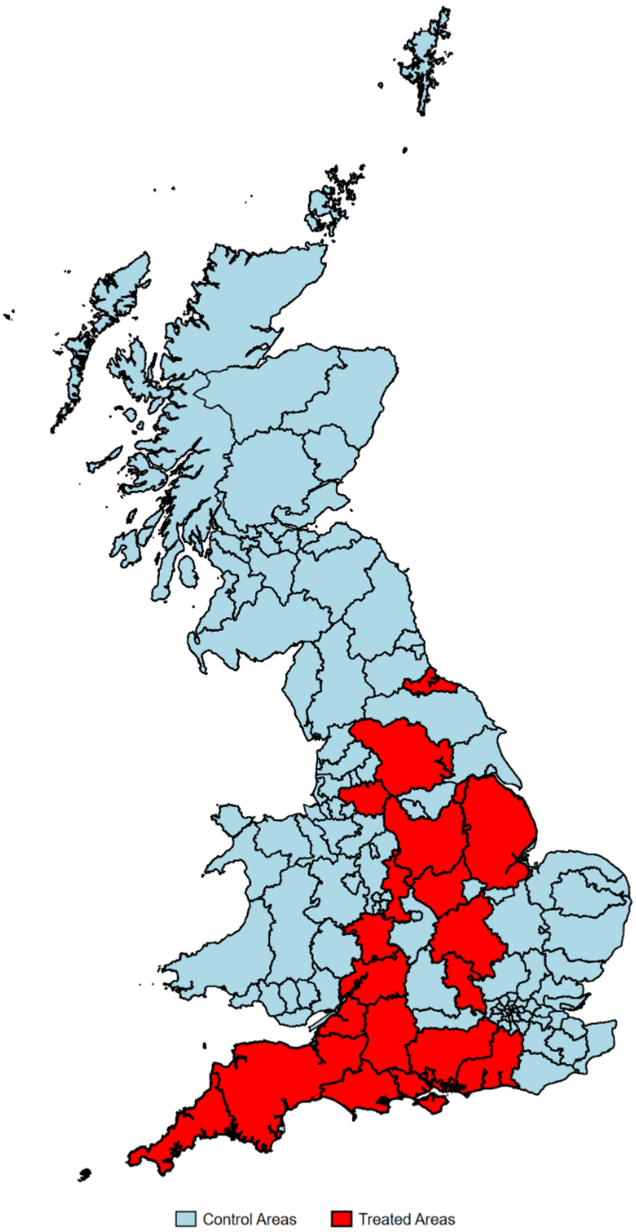
where the model is separately estimated on subsamples of urban and rural firms (Panel A) and universities areas and all other areas (Panel B). The *Urban Areas* subsample contains SMEs in the largest 10 cities by population according to the 2011 census, while the *Rural Areas* subsample contains SMEs 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* subsample contains SMEs in all other locations. $Y_{i,t}$ 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), year (t), industry-by-year (j), and geographic (g) fixed effects. Industry is based on 1-digit U.K.-SIC codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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 at the bottom of each sub-panel. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*	
	(1)	(2)	(3)	(4)	
	Reported Coefficient: $Post \times Treated$				Observations
Panel A: City Population					
Urban Areas	0.023**	0.059**	0.045**	0.045**	96,678
	(2.05)	(2.25)	(2.39)	(2.23)	
Rural Areas	0.045***	0.108***	0.077***	0.084***	41,889
	(3.67)	(3.67)	(3.63)	(3.67)	
Diff. (Urban - Rural)	-0.022***	-0.049***	-0.032**	-0.039***	
F-stat	[7.47]	[7.59]	[6.17]	[7.59]	
Panel B: University Ranking					
Universities Areas	0.029**	0.070**	0.050**	0.054**	29,545
	(2.14)	(2.23)	(2.30)	(2.18)	
All Other Areas	0.023**	0.059***	0.044***	0.046***	118,194
	(2.48)	(2.64)	(2.72)	(2.64)	
Diff. (University - All Others)	0.006	0.011	0.006	0.008	
F-stat	[0.42]	[0.21]	[0.13]	[0.17]	
Firm FE	Yes	Yes	Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	
Industry × Year FE	Yes	Yes	Yes	Yes	
Geographic FE	Yes	Yes	Yes	Yes	

FIGURES

Figure I
Treated LEPs and Control Areas

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



Internet Appendix for

"PROMOTING DIGITALIZATION THROUGH INFORMATION DISSEMINATION"

This Internet Appendix includes descriptions and results from supplementary and robustness tests not reported in the paper.

IA.I BuiltWith Data

We employ a novel source of digital technology data that is available at the firm level and covers the business population across all size groups. The data consist of the web technologies that BuiltWith retrieves from each firm's website. Its 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).

On a regular basis, BuiltWith collects data on the web technologies used on firms' websites by using website crawlers to identify and track all websites.[31] Web technologies captured by BuiltWith might include, among others, the type of web hosting the website uses, its e-commerce functionalities, and its analytics software.[32] Such technologies can change over time depending on how sophisticated the firm's website becomes. BuiltWith tracks when technologies are added or removed from a website through either weekly, bi-weekly, or quarterly updates depending on the activity of the website.[33] This enables

[31]BuiltWith's website crawler works similarly to Google Search by crawling and indexing websites. The crawler robot automatically finds websites from the internet and downloads information from each website ("indexing"). BuiltWith targets to index all country-code top-level domains (ccTLD) as well as any active generic top-level domains (gTLD), such as '.com', '.net' and '.org' in all regions (BuiltWith 2022).

[32]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 is displayed. Customers can browse, select the product from the online shop, pay for it, and, ultimately, arrange the shipping of such product. Web analytic software (e.g., Google Analytics) are tools that can collect and report web traffic data to improve the user's experience and the website's performance. It can help firms better understand their customers and refine their marketing strategies.

[33]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 (BuiltWith 2022,

us to construct a panel dataset of firms' website digital technologies at the end of their fiscal years. In addition, BuiltWith classifies all technologies in *tags* according to their functions.

Overall, 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 636,607 firm-technologies observations that are based on 4,244 unique technologies and 26 tags.

We then 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 apply 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: advertising, analytics and tracking that enable digital marketing, audio and video media, e-commerce shop, email, feeding content, language, mobile commerce technologies, payment, shipping provider, SSL certificate, widgets that include social media sharing. Among them, we further classify e-commerce shop, payment, shipping provider and SSL certificate as e-commerce related technologies. Table IA.I shows all technology tags included in our sample and their classifications; while Tables IA.II includes the definitions of the digitalization variables and IA.III reports their descriptive statistics, respectively.

IA.II SAFE Index

In Section VI, we use several measures of financial constraints to construct sub-sample pairs of (un)constrained firms. In addition to using firm size and leverage as proxies, we measure the degree of financial constraints using the ECB SAFE Index. For the U.K., the ECB defines that index as follows:

BuiltWith 2022).

$$
\begin{aligned}
SAFEIndex_{i,t} = &+0.015FinancialLeverage_{i,t} + 0.0004CoverageRatio_{i,t} \\
&- 0.070Profitability_{i,t} + 0.043Tangibility_{i,t} \\
&- 0.134CashHolding_{i,t} - 0.024TotalAssets_{i,t} + \alpha_j,
\end{aligned} \tag{IA.1}
$$

where $FinancialLeverage$ is measured as the sum of short-term debt and long-term debt divided by total assets. $CoverageRatio$ is interest expenses scaled by earnings before interest, taxes, and depreciation. $Profitability$ is earnings before interest and taxes divided by total assets. $Tangibility$ is tangible fixed assets divided by total assets. $CashHolding$ is cash and cash equivalent divided by total assets. $TotalAssets$ is the log of one plus total assets. All these variables are averaged over the pre-program period. α_j is the industry fixed effects, where the four main sectors used by the ECB are construction, trade, services, and industry. Following ECB standards, we identify the threshold to define firms as financially constrained by using the distribution of *SAFE Index* from the survey data for the U.K (Ferrando et al. 2015).[34] We classify firms as financially constrained if they are in the top 12% of the distribution, where *12%* is the weighted average of constrained firms calculated directly from the SAFE survey for the U.K.

The ECB categorizes firms as more or less financially constrained based on a combination of qualitative information from the SAFE and quantitative data from financial statements available in ORBIS. The SAFE survey is stratified across different sectors, euro area countries, firm age, and firm ownership. Some of the early waves also include a stratified sample from the U.K. used to compute the above estimates. The SAFE index categorizes firms as more financially constrained based on their responses to the following: (a) they report that either loan applications or bank overdraft/credit lines applications or both were rejected; (b) they report that for either loan applications or bank overdraft/credit lines applications or both, only a limited amount of credit was granted; (c) they report that either loan applications or bank overdraft/credit lines applications or both were declined by the firms because the borrowing costs were too high; and (d) they did not apply for either loans or bank overdraft/credit lines or both for fear of rejection (*discouraged borrowers*). Once the firms in the SAFE survey have been matched with firms in ORBIS, a probit model estimates the probability that a firm is financially constrained as a function of its leverage, coverage ratio, profitability, tangibility, cash holding, size, and industry. The coefficients of the probit model are then used

[34] Access to the anonymized SAFE microdata is available at www.ecb.europa.eu/stats/ecb_surveys/safe/html/data.en.html

to calculate the SAFE Index as in Equation IA.1.[35]

IA.III Identification Strategy Tests

This section includes tests to validate our identification strategy. In particular, Table IA.IV reports the results of falsification tests that were constructed using a placebo sample or a placebo event year. Table IA.V reports the results of the local effect test, where we compare the treated SMEs with control SMEs that are located near the geographic boundaries of the treated LEPs. Table IA.VI reports another robustness test where we test the baseline results excluding London from the control group.

Last but not least, Figures IA.I, Figure IA.II, and IA.III show the plots of the dynamic coefficients estimated using year-by-year interaction terms for our baseline digitalization outcomes, digitalization outcomes within treated LEPs, and real economic outcomes, respectively, to support the parallel trends assumption.

IA.IV Supplementary Tests

This section includes supplementary tests. Table IA.VIII reports results of the effects of the program on digitalization across sub-samples of B2B and B2C firms. Table IA.IX reports findings related to the effect on the growth rate of web traffic.

IA.V The Digital Divide

In this section, we present results of tests on the digital divide between large corporations and SMEs.

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).

[35] A potential limitation of this approach is the stability of the coefficients, which means that the estimated coefficients may not apply to firms outside the estimation sample and the estimation period (see, e.g., Hoberg and Maksimovic (2015)). Results based on the SAFE Index are consistent with those based on alternative proxies of constraints.

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 fewer 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 a digital divide has been to particularly focus the allocation of government resources 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:

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

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

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

where $\overline{Y^l_{j,g,t}}$ is the median digitalization outcome across large firms (l) operating in the same industry j and 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*.[36] The coefficient of interest is δ, which 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 reduced the digital gap between them and large companies more in the treated areas than the control ones.

Table IA.X reports the findings of the catching-up analysis. In terms of the number of newly adopted technologies (*BG-score*), results in columns (1) to (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 0.30 (column (1)). Given that the average gap between large firms and control SMEs in the post-program period is 5.48, this corresponds to a reduction of approximately 6% (0.31/5.48*100 = 0.057%).

Similarly, the results show negative and statistically significant treatment effects on the gaps between SMEs and large firms for both business-related (columns (4) to (6)) and general technologies (columns (7) to (9)). The effect is similar across both types of technologies, ranging between -0.12 and -0.16 across all specifications.

Finally, columns (10) to (12) of Table IA.X report the results of catching-up on *E-score*. Compared with the controls, we find that the gap in the number of adopted e-commerce technologies between large firms and treated SMEs is significantly smaller. For instance, the gap reduces by 0.06 more for treated SMEs compared to the control ones (column (10)).

Overall, the evidence suggests that the program was also successful in helping narrow the digital divide between large firms and SMEs.[37]

[36]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.

[37]We also perform a year-by-year PSM-DID estimation of the catching-up analysis presented in Table IA.X. Figure IA.IV in the Internet 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 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 observed in that period are due to the digitalization of large firms in both treated and control areas. This suggests that large firms across all areas have a similar degree of digitalization and, as such, they can be used as a benchmark to construct the *Gap* outcome variables. It also confirms that the parallel trends 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 than the control group. Further, such gaps narrow more and more over time, consistent with the evolution of the digitalization outcomes as in Figure IA.I.

REFERENCES FOR THE INTERNET APPENDIX

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.

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.

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

Ferrando, Annalisa, Matteo Iudice, Carlo Altomonte, Sven Blank, Marie-Hélène Felt, Philipp Meinen, Katja Neugebauer, and Iulia Siedschlag. "Assessing the financial and financing conditions of firms in Europe: the financial module in CompNet." ECB Working Paper No. 1836, ECB, 2015.

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

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.

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

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.

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.

Table IA.I
List of Technology Tags

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.

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*
Verified Links	Widgets
Web Master	
Web Server	

**indicates e-commerce technologies*

Table IA.II
Definitions of Variables

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.

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 Growth	Growth rate of ROA measured as 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).

Table IA.III
Descriptive Statistics Post-Program Period

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 a 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.

	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Diff.
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(4)-(7)
Panel A: Before log-transformation										
	All Firms (N = 20,860)			Treated (N = 10,430)			Control (N = 10,430)			
Web	0.269	0.000	0.443	0.284	0.000	0.451	0.252	0.000	0.434	0.032***
										(12.41)
BG-score	2.181	0.000	4.589	2.324	0.000	4.711	2.027	0.000	4.447	0.297***
										(10.88)
B-score	0.968	0.000	2.150	1.034	0.000	2.209	0.897	0.000	2.083	0.137***
										(10.73)
G-score	1.213	0.000	2.521	1.289	0.000	2.586	1.130	0.000	2.445	0.160***
										(10.65)
E-commerce	0.161	0.000	0.368	0.173	0.000	0.378	0.148	0.000	0.355	0.025***
										(11.22)
E-score	0.225	0.000	0.565	0.243	0.000	0.585	0.207	0.000	0.542	0.036***
										(10.78)
Panel B: After log-transformation										
	All Firms (N = 20,860)			Treated (N = 10,430)			Control (N = 10,430)			
Ln(BG-score)	0.503	0.000	0.973	0.534	0.000	0.996	0.470	0.000	0.947	0.064***
										(11.12)
Ln(B-score)	0.341	0.000	0.696	0.363	0.000	0.714	0.317	0.000	0.675	0.046***
										(11.23)
Ln(G-score)	0.395	0.000	0.770	0.419	0.000	0.788	0.369	0.000	0.749	0.050***
										(10.83)
Ln(E-score)	0.137	0.000	0.324	0.147	0.000	0.334	0.126	0.000	0.313	0.021***
										(11.09)

Table IA.IV
Placebo Tests

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 2012 to 2013 and zero from 2009 to 2011 (*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. $\boldsymbol{Z}$ 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 codes. The geographic fixed effect is based on the 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 (2010) values of age, leverage, cash, turnover, employment, 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.

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
Panel A: Placebo Sample				
Post × *Treated_P*	0.003	0.004	0.002	0.003
	(1.24)	(0.65)	(0.48)	(0.68)
Observations	150,895	150,204	150,204	150,204
Adjusted R^2	0.73	0.76	0.75	0.74
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
Geographic FE	Yes	Yes	Yes	Yes
Panel B: Placebo Event Year				
Post_P × *Treated*	-0.001	-0.003	-0.003	-0.001
	(-0.18)	(-0.24)	(-0.74)	(-0.16)
Observations	120,024	120,022	120,022	120,022
Adjusted R^2	0.25	0.23	0.22	0.24
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
Geographic FE	Yes	Yes	Yes	Yes

Table IA.V
Local Effect

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. $\boldsymbol{Z}$ 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 codes. The geographic fixed effect is based on the 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, turnover, employment, 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 of 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.

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
$Post \times Treated$	0.017*	0.044**	0.037***	0.030***
	(1.87)	(2.44)	(2.61)	(2.61)
Observations	87,317	87,314	87,314	87,314
Adjusted R^2	0.53	0.53	0.53	0.52
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
Geographic FE	Yes	Yes	Yes	Yes

Table IA.VI
Digitalization Outcomes: Baseline Results Excluding London

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. The London area is excluded from the sample. ***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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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.

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*
	(1)	(2)	(3)	(4)
$Post \times Treated$	0.017*	0.046*	0.036**	0.036*
	(1.72)	(1.90)	(2.02)	(1.88)
Observations	132,298	132,907	132,907	132,907
Adjusted R^2	0.53	0.54	0.52	0.53
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry $\times$ Year FE	Yes	Yes	Yes	Yes
Geographic FE	Yes	Yes	Yes	Yes

Table IA.VII
Pre-Program Characteristics within LEPs

This table compares several macro characteristics within LEPs across treated LEPs (Column (1)) and those that didn't receive any funding (Column (2)) in the pre-program period from 2011 to 2013. Treated LEPs include those LEPs that received funding from the *Challenge Fund*. Other LEPs include those LEPs that did not receive any funding from the *Challenge Fund*. Column (3) reports differences in all characteristics between the treated and other LEPs 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 demographics 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. *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. *Number of Bank Branches* is the number of bank branches per 100 firms in each area. Panel C reports characteristics at the LEP level. *RGF Amount* is the total amount of the Regional Growth Funds (RGF) that LEPs were able to raise. *Board Size* is the total number of directors seated on the LEP's board. *Directors' Age* is the average age of directors seated on the LEPs' boards. *Government-affiliated Directors* is the total number of directors seated on the LEP's board that have a political affiliation. The total number of observations used in this analysis is 114: 63 observations for the treated areas and 51 for the control ones. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Variables	Treated LEPs (1)	Other LEPs (2)	Diff. (1)-(2)
Panel A: Internet Infrastructure			
% Areas without good internet connection	0.09	0.10	-0.01
			(-0.28)
% Internet Users	0.82	0.80	0.02
			(1.25)
Total Websites	1,059	1,568	-509
			(-1.03)

(continued on next page)

Table IA.VII
Pre-Program Characteristics within LEPs (cont'd)

Variables	Treated LEPs (1)	Other LEPs (2)	Diff. (1)-(2)
Panel B: Business demographic and economic conditions			
Business Population	43,691	49,311	-5620
			(0.0765)
Number of Business Birth	5,237	6,719	-1,482
			(-0.86)
Number of Business Death	5,193	6,054	-861
			(-0.63)
%Businesses survived aft 1 yr	0.93	0.93	0.00
			(0.37)
Employment	570,543	649,559	-79,016
			(-0.62)
%Educated Workers	0.33	0.32	0.01
			(1.16)
GDP	30,630.98	43,472.24	-12,841.25
			(-0.73)
GDP Growth	0.04	0.03	0.01
			(0.64)
GVA	27,015.98	39,358.49	-12,342.51
			(1.26)
Number of Bank Branches	0.546	0.539	0.007
			(0.23)
Panel C: LEPs characteristics			
RGF Amount (£mil)	3.668	3.782	-0.114
			(-0.079)
Board Size	20.00	21.88	-1.88
			(-0.322)
Directors' Age	65.29	64.59	0.698
			(0.257)
Government-affiliated directors	2.750	1.750	1.000
			(0.864)

Table IA.VIII
Digitalization Outcomes:
B2B versus B2C effect

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 based on their target customers. Business-to-Business firms sell products and services to other businesses ('B2B Firms'), while Business-to-Customers firms are those that sell products and services directly to consumers ('B2C Firms'). $Y_{i,t}$ 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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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 at the bottom of each sub-panel. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Web	*Ln(BG-score)*	*Ln(B-score)*	*Ln(G-score)*	
	(1)	(2)	(3)	(4)	
	Reported Coefficient: $Post \times Treated$				Observations
B2B Firms	0.017	0.043*	0.030*	0.034*	64,286
	(1.51)	(1.69)	(1.74)	(1.71)	
B2C Firms	0.033**	0.077***	0.056**	0.058**	69,066
	(2.31)	(2.19)	(2.22)	(2.14)	
Diff. (B2B – B2C)	-0.016**	-0.034**	-0.026**	0.025**	
F-stat	[5.44]	[4.04]	[4.37]	[3.46]	
Firm FE	Yes	Yes	Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	
Industry × Year FE	Yes	Yes	Yes	Yes	
Geographic FE	Yes	Yes	Yes	Yes	

Table IA.IX
Web Traffic Growth

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. ***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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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.

	Website Traffic Growth		
	(1)	(2)	(3)
Post × *Treated*	0.408**	0.384**	0.373**
	(2.03)	(2.30)	(2.24)
Observations	129,341	129,341	127,988
Adjusted R^2	0.04	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

Table IA.X
Digitalization Catching-up between SMEs and Large firms

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:

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

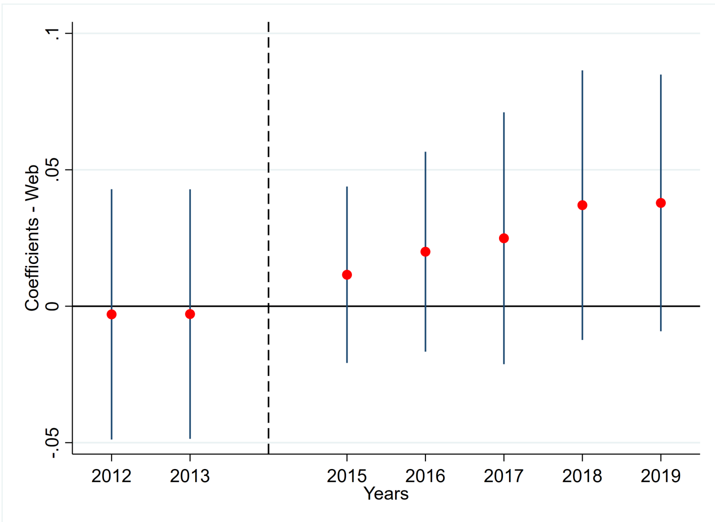
where $\overline{Y^l_{j,g,t}}$ is the median digitalization outcome across large firms operating in the same industry j and located in the same area g of SME i at time t, and $Y_{i,t}$ is the digitalization outcome of SME i at time t. Large firms are defined as companies with more than 250 employees 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. ***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 codes. The geographic fixed effect is based on the NUTS-2 level. All regressions are estimated on a propensity score matched sample using 2013 values of size, age, leverage, cash, turnover, employment, 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.

	Gap(BG-score)			*Gap(B-score)*			*Gap(G-score)*			*Gap(E-score)*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$Post \times Treated$	-0.666***	-0.653***	-0.639**	-0.420***	-0.407***	-0.401***	-0.836**	-0.818**	-0.806**	-0.108***	-0.107***	-0.106***
	(-3.18)	(-3.28)	(-3.22)	(-2.15)	(-2.10)	(-2.07)	(-2.52)	(-2.45)	(-2.42)	(-3.20)	(-3.11)	(-3.08)
Observations	145,954	145,954	144,402	145,954	145,954	144,402	145,954	145,954	144,402	145,954	145,954	144,402
Adjusted R^2	0.46	0.49	0.49	0.63	0.65	0.65	0.72	0.75	0.75	0.49	0.51	0.51
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

Figure IA.I
Dynamic Models for Digitalization Outcomes

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

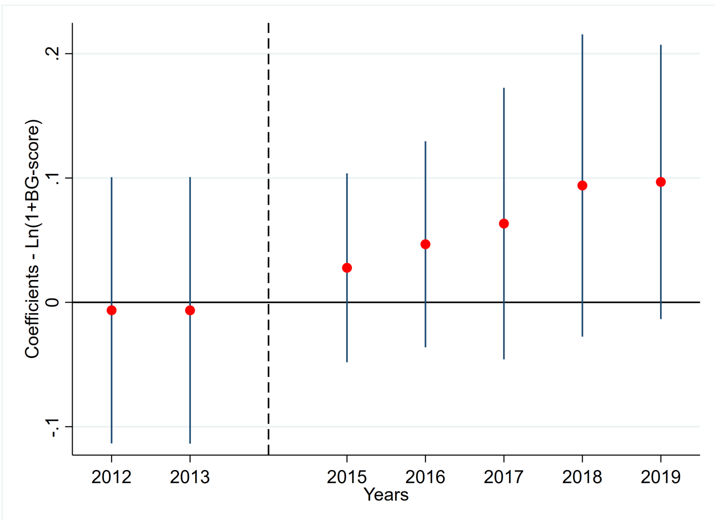
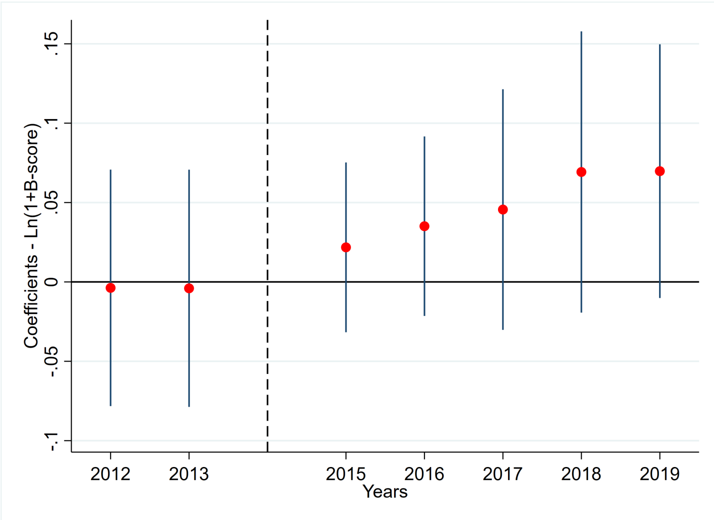


Figure IA.I
Dynamic Models for Digitalization Outcomes (cont'd)

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



Panel D: $Ln(G\text{-}score)$

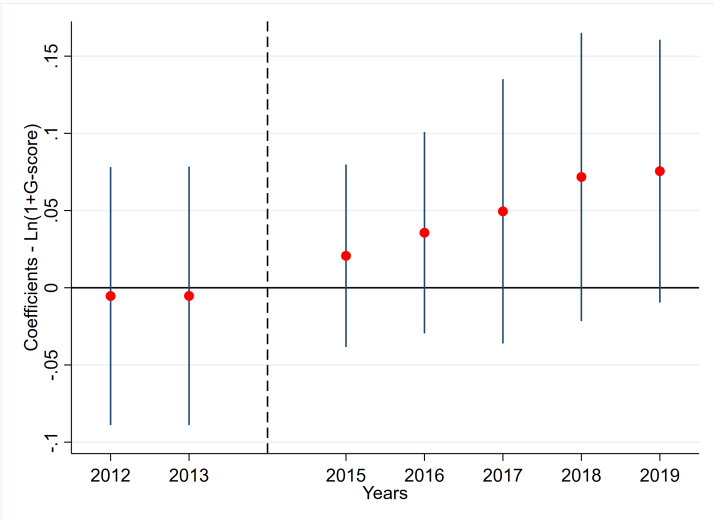
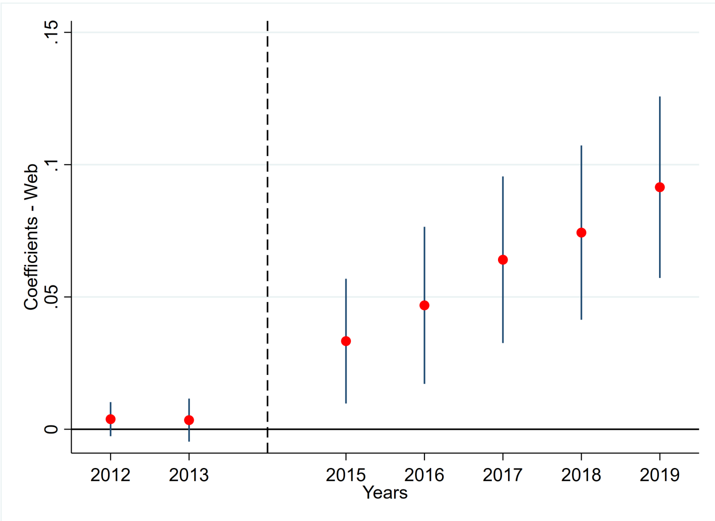


Figure IA.II
Dynamic Models for Digitalization Outcomes: Within Treated LEPs

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

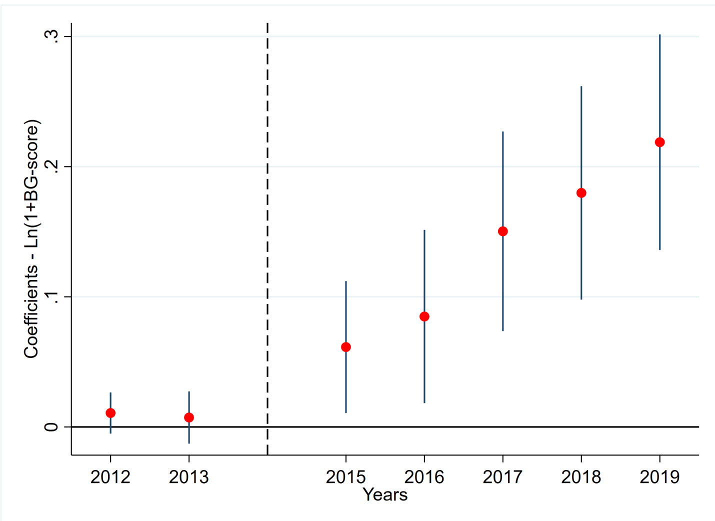
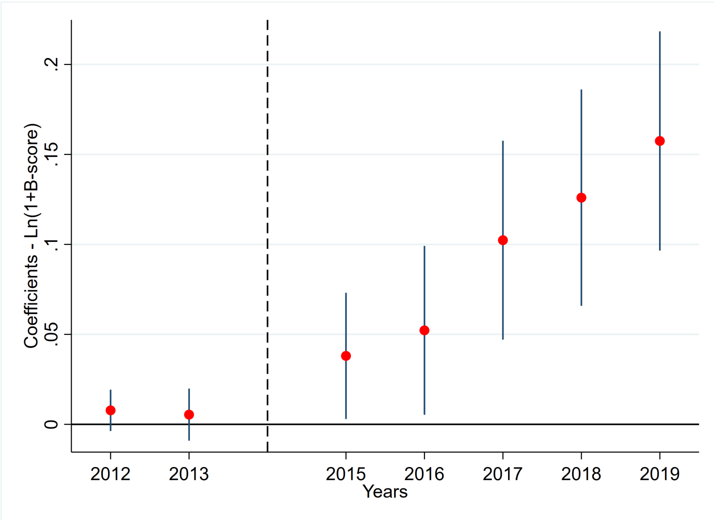


Figure IA.II
Dynamic Models for Digitalization Outcomes (cont'd)

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



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

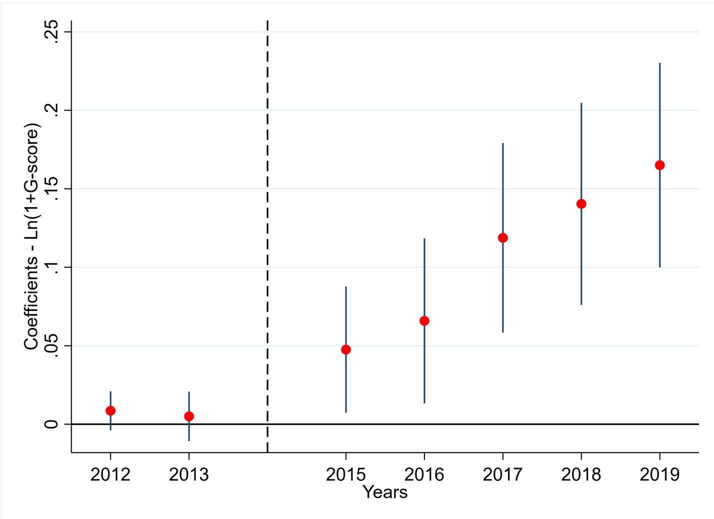
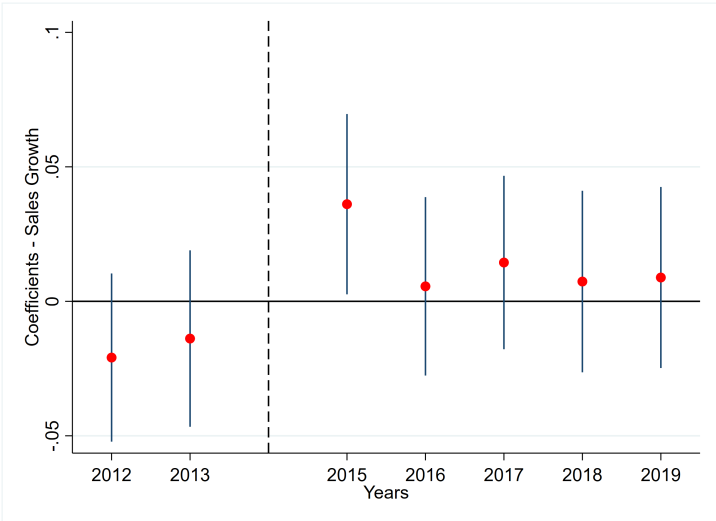


Figure IA.III
Dynamic Models for Real Outcomes

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*Sales Growth*, *ROA Growth*, Δ *Ln(Employment)*, or Δ *Ln(VPE)*, respectively.

Panel A: *Sales Growth*



Panel B: *Growth ROA*

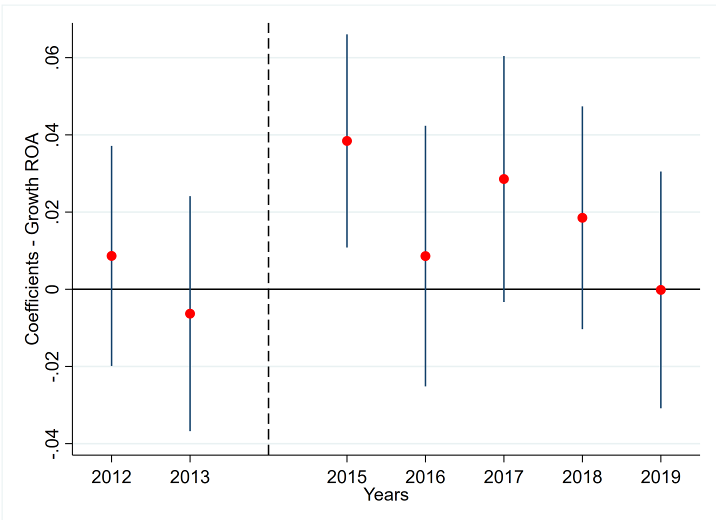
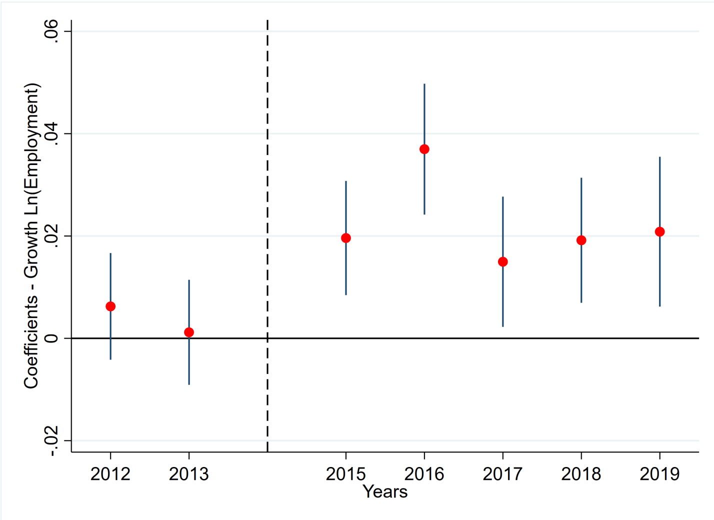


Figure IA.III
Dynamic Models for Digitalization Outcomes (cont'd)

Panel C: Δ *Ln(Employment)*



Panel D: Δ *Ln(VPE)*

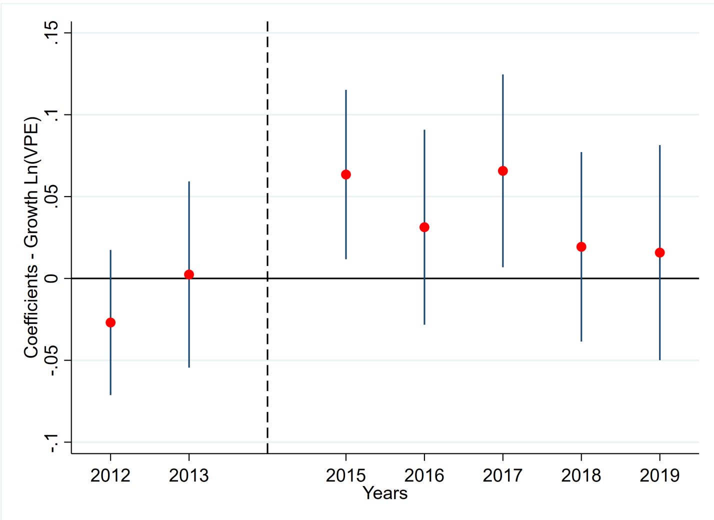
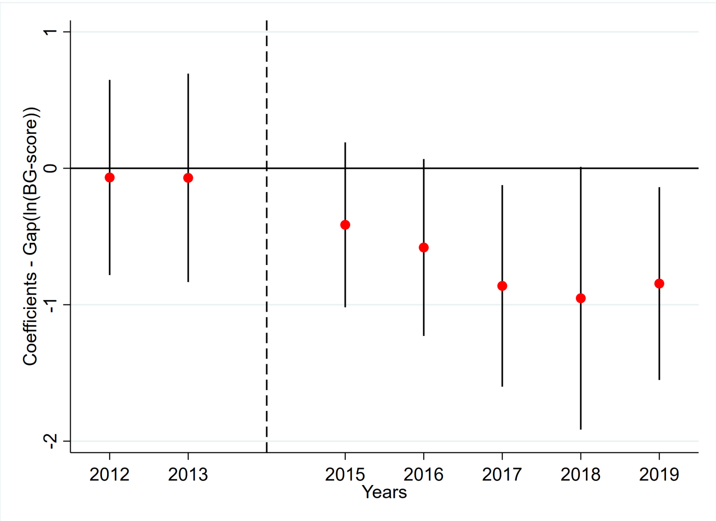


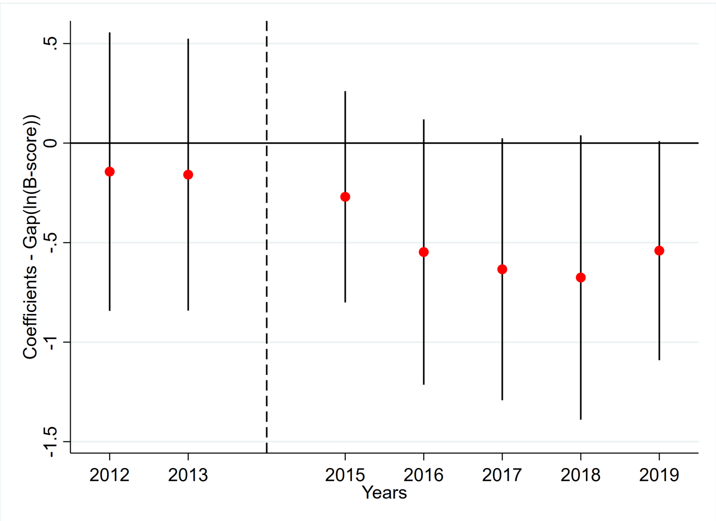
Figure IA.IV
Dynamic Models for Digitalization Catching-up

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.

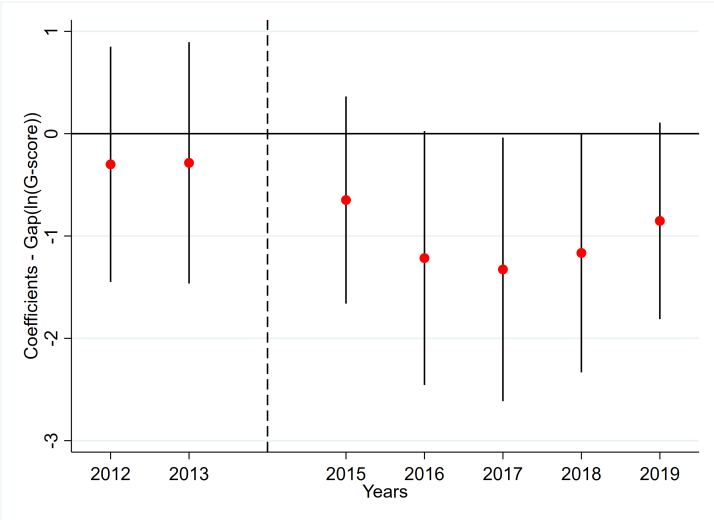
Panel A: *Gap(BG-score)*



Panel B: *Gap(B-score)*



Panel C: *Gap(G-score)*



Panel D: *Gap(E-score)*