CLOUD COMPUTING AND FIRM GROWTH

Timothy DeStefano[1]
Georgetown University
Richard Kneller
University of Nottingham
Jonathan Timmis
World Bank

Cloud computing has shifted how firms access IT away from investment in fixed capital to pay-on-demand services that facilitate remote and simultaneous use. Using new firm-level data we examine the impact of cloud adoption on firm performance and organizational geography with an IV approach that exploits cross-section and time-series variation in fiber broadband speeds as instruments. Cloud leads younger firms to increase revenue, employment, and productivity, whereas incumbent firms experience no scale effects and weaker productivity gains. Incumbents however undergo restructuring through establishment deaths and fewer births, while both types of firms experience geographic reorganization as activity shifts farther from the headquarters.

JEL classification: J21; J24; L20; O33

[1] We would like to thank Giuseppe Berlingieri, James Bessen, Nick Bloom, Timothy Bresnahan, Erik Brynjolfsson, Carol Corrado, Chiara Criscuolo, Oliver Falck, Patrick Gaule, Avi Goldfarb, Shane Greenstein, Marco Iansiti, Wang Jin, Nick Johnstone, Bradford Jensen, Karim Lakhani, Kristina McElheran, Frank Nagle, Giuseppe Nicoletti, Taylor Reynolds, Stephen P. Ryan, Raffaella Sadun, Timothy Simcoe, and participants at NBER Productivity, Innovation and Entrepreneurship, NBER Economics of Digitization Spring Meetings, CESIfo Economics of Digitization Conference and the OECD Spatial Workshop on Productivity for helpful comments and suggestions.

Corresponding author: Timothy DeStefano, Georgetown University: timothy.destefano@georgetown.edu
Richard Kneller, University of Nottingham: richard.kneller@nottingham.ac.uk
Jonathan Timmis, World Bank: jtimmis@worldbank.org

Introduction

Over the last decade, a fundamental shift has occurred in the manner by which firms access digital technologies. Traditionally, acquiring information and communication technologies required businesses to make considerable upfront, sunk investments in hardware infrastructure, software and to maintain large IT departments. Now, alternatively, firms acquire their storage, processing and software needs as a service - what is typically referred to as “cloud computing” (Van Ark, 2016; OECD, 2015). Cloud providers offer these services “on demand” via “pay as you go” subscriptions[2]. Purchased in this way, IT shifts from a sunk investment to a largely variable cost, which may lead to changes in firm behavior that go beyond simply acting as a substitute for accessing IT (Iansiti and Richards, 2011; OECD, 2015; OECD, 2014).

In this paper, we use newly available micro data for the UK that measures the adoption of cloud. Detailed measures of cloud adoption at the firm-level have not previously been available to researchers on this topic. These data also allow us to directly explore the extent to which cloud adoption impacts firm performance and organization.

We build on the existing literature to argue that the performance effects of cloud are likely heterogeneous across young and incumbent firms. It has been claimed that the change in the nature of IT costs towards being a largely variable cost has enabled new business models and firm types. Firms can scale operations quickly without the need for acquiring a mass of IT assets, typically referred to as ‘scale without mass’.[3] This is expected to have particularly strong effects on new entrants, since up-front IT investments can be burdensome for young firms

---

[2] Cloud computing is a service delivered by a third party which “enables ubiquitous, convenient on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released” (NIST, 2011).

[3] Uber, NetFlix and Airbnb are often held up as examples of the type of business model made possible from cloud.

given their financial constraints. By avoiding quasi-irreversible investments cloud can also allow for greater flexibility and experimentation, which is key to young firm growth (Decker et al., 2014). In contrast, the transition to the cloud may be more difficult for older firms that may have made large investments in software in the past (Bommadevara et al., 2018).

Within the paper, we also explore a second mechanism, the mobility offered by decentralized data processing and software, and how this affects the spatial distribution of firm activity. Cloud reduces the need for centralized IT departments and allows workers to access IT infrastructure from outside the firm (Iansiti and Richards, 2011). The reduced costs of accessing information across many locations simultaneously would typically facilitate greater geographic dispersion of tasks away from the headquarters (Leamer and Storper, 2001; Duranton and Puga, 2005, Bloom et al., 2014). However, monitoring and problem solving is more likely to be done by senior managers at the headquarters.[4] Which of these geographic effects dominate for cloud adopters and for which type of firm (young or incumbent) is unclear and is assessed in this paper.

We use a number of both traditional and novel measures of firm geographic dispersion. The former includes becoming a multi-establishment firm, establishment birth and death and the number of local regions in which the firm is located. We also introduce two new measures of geographic concentration. First, we measure the unweighted and weighted average distance between establishments and the firm headquarters (weighted by the share of establishment employment in firm employment). Secondly, we construct a distance-employment covariance term to measure how employment is distributed across more proximate or more remote establishments.

[4] Cloud might also encourage hot-desking or flexible work patterns, allowing greater efficiency of space.

A final contribution of the paper is the novel instrumental variable approach. The growth of cloud services is a phenomenon that has gone hand-in-hand with the diffusion of high-speed fiber broadband, driven in part by the removal of the asymmetry between upload and download speeds that was a feature of older broadband technologies, namely ADSL. Symmetric speeds mean that data could be uploaded and shared outside of the firm at much faster rates, enabling real-time collaboration through the cloud.[5]

Our identification strategy relies on time and cross-section (zip-code level) differences in expected fiber broadband speeds that arise from the infrastructure to deliver the technology. We show that firms with access to fiber because the local telephone exchange has been fitted with the requisite technologies and with short cable (local loop) distances to this exchange (which enables faster fiber speeds) are more likely to adopt cloud than those connected to exchanges not yet enabled with fiber, or those attached to an enabled exchange but with a longer cable distance.[6] Importantly, we find that these distance instruments behave in a manner that is closely aligned with the predictions from the telecoms engineering literature.

We take seriously issues surrounding the plausibility of these instruments, which we deal with by a series of sample restrictions, firm fixed effects, as well as providing tests for pre-trends via event studies that account for staggered fiber enablement (following Callaway and Sant'Anna, 2021) and tests of the correlation between the instruments and other technologies.

To preview the main results of the paper. Firstly, in terms of adoption mechanisms we find that cloud does indeed lead to a switch away from investment in tangible IT assets. Secondly,

[5] A stable, high-speed broadband connection is widely seen as a pre-requisite for cloud use (ITU, 2017).

[6] We are only aware of one other paper using the availability of fiber as an instrumental variable. Fabling and Grimes (2016) examines the diffusion of fiber on employment and productivity for New Zealand firms, using proximity to nearby schools as an instrument.

we find strong heterogeneity in the performance and geography effects of cloud. Younger firms that adopt cloud are more likely to grow in employment, sales and labor productivity, whereas for incumbent firms we observe no scale increases with only weak productivity effects. Cloud adoption also impacts the geographic reorganization of firms. Incumbents restructure with cloud, shutting down establishments and reducing the probability of opening new ones. In addition, both young and incumbent firms decentralize activity farther from the headquarters.[7] Taken together cloud appears to have important implications for how firms grow and reorganize.

This paper contributes to a long-established literature on the impact of ICTs on firm performance. A broad literature has shown how digital technology can lead to increases in productivity and scale, particularly for larger or initially productive firms (Bloom et al, 2012; Brynjolfsson, et al 2008, Draca et al, 2006). Recent evidence has linked IT to the slowdown of business dynamism, rising industry concentration and widening disparities between frontier and laggard firms (Calvino, et al., 2016, Crafts and Mills, 2020; Decker et al., 2016). Cloud computing also appears to be relatively more accessible for younger and small entities than older-IT, potentially levelling the playing field between firms (Bloom and Pierri, 2018).

In considering these questions we build on a small literature on the effects of cloud computing on firms. Bloom and Pierri (2018) find for example, that the adoption of cloud is occurring at a faster rate amongst young and small businesses than for previous IT technologies, while Jin and McElheran (2017) find evidence that purchases of IT services, including cloud services, are related to significantly higher survival and growth among young establishments.

[7] When we assess dispersion at the employee level, we find that workers in establishments using cloud technologies are more likely to be relocated compared to establishments that have not adopted the cloud.

This work also contributes to an emerging part of the IT literature that focuses on the impact of the organization and geography of the firm, to which we add new measures of dispersion. Previous work examining the impact of IT on firm organization find that digital technologies lower the cost of communication resulting in more hierarchical firm structures (Bloom et al., 2014). Other research demonstrates that processing and communication IT often push economic activity and decision making in competing directions (Bloom et al., 2014; Garicano and Heaton, 2010). Studies focusing on the geography of the firm have examined the link between the diffusion of broadband on regional concentration of innovation, finding evidence of growth in patenting amongst earlier adopters of the internet (Forman et al., 2015) and that digital technologies can help establish new collaborations across geographic space (Greenstein et al, 2018).

The rest of the paper continues as follows. Section 2 discusses the data used in this paper and presents initial evidence on cloud adoption and IT investment costs, while Section 3 introduces the empirical framework for the analysis. Section 4 presents the main results of the analysis and Section 5 provides some concluding comments.

Data

In this paper we utilize novel micro data from the Office for National Statistics (ONS), which is the UK Census Bureau equivalent. Basic data on firms such as employment, industry and zipcode location of the headquarters and its establishments is sourced from the UK business registry – the Business Structure Database (BSD).

Information on firm cloud adoption is taken from the E-commerce Survey, which includes questions regarding seven different uses of cloud computing.[8] This includes hosting the

[8] The E-commerce survey is a stratified random sample of firms.

business' databases, storage of files, email, office software (such as word-processing and spreadsheets), finance and accounting software, customer reseource management (CRM) software and running the business' own software. These questions are asked in 2013 and 2015 only. Our main measure is a dummy variable of whether the firm uses any form of cloud computing, although we also report results using dummies for these types of cloud technology separately. As we outline in detail below, our estimation strategy relies on the adoption of cloud technologies. Alongside the data for 2013 and 2015, to measure pre-cloud adoption we use data from 2008 – the year the fiber enablement program was first announced in the UK (we discuss fiber rollout later).

Firm outcomes such as employment, sales, labor productivity (sales per worker), and IT investment, are from the Annual Respondent's Database (ARD). Constructed from a mandatory business survey, the ARD is a census of large businesses and a stratified random sample of smaller firms. It covers economic activity in all sectors of the economy asides agriculture and finance from 1997.[9]

As discussed above, there are strong reasons to expect that the impacts of cloud may differ between younger and incumbent firms. Unless otherwise stated, the sample size in our regressions is 17,386 firm-year observations. This is made up of 8,251 firms, of which 25 per cent are defined as young (defined as aged 10 years old or younger in 2008) and the remainder as incumbents (defined as older than 10 years in 2008). We assess further nuances of age heterogeneity in a subsection later in the paper.

Alongside standard firm performance measures such as size and labor productivity, we also use measures of the geographic reorganization of within-firm activity. Our more standard measures are becoming a multi-establishment firm, establishment births and deaths and the

[9] Unfortunately, we do not have comprehensive capital data, which prevents analysis of TFP.

number of different local authorities (equivalent to counties in the US) in which a firm's establishments are located. [10] Our second measure reflects the geographic dispersion of employees from the headquarters – specifically a weighted average distance between establishments and their headquarters (weighted by the share of establishment employment in firm employment). We decompose this weighted average distance into two terms – an unweighted average and a distance-employment covariance term. The covariance term reflects the relationship between establishment distance from the headquarters and establishment employment. A positive (negative) covariance shows that more distant (closer) establishments are relatively larger in terms of employment. A covariance term of this type has been popularized by Olley and Pakes (1996) in productivity decompositions.

Table 1 provides summary statistics for all firms, young and incumbent for all years and then separately for 2013 and 2015 for the cloud variables (in Table A1). [11] Around 22% of firms use some form of cloud within the sample period. The usage of specific cloud services, including relatively popular types such as storage of files (14%), business databases (9%), and email (12%) are lower than overall use, indicating that firms typically adopt some but not all types at once. It is also evident from Table 1 that cloud adoption is lower amongst young (17%) vs incumbent firms (24%). This contrasts with a commonly held view that access to digital technologies via the cloud is particularly attractive for young firms.

[INSERT TABLE 1 HERE]

Mechanism between cloud and traditional IT investment

It has been argued that cloud technologies represent a radical departure in the way firms adopt IT. Non-cloud data transfer systems have though, existed for decades, but require large sunk

[10] Establishment deaths and establishment births are all expressed relative to the total number of establishments.

[11] Table A2 in the Appendix provide summary statistics on all other variables.

costs, in contrast to flexible storage, database hosting and processing offered by cloud technologies. For instance, one common alternative to cloud for data transfer is file transfer protocol (FTP), used to send file transfers to and from in-house servers. The sunk costs associated with FTP include the server hardware to store the data, the investment in centralized software databases to process and update the data and IT departments to manage these systems. Consistent with this, previous research has found strong evidence that earlier digital technologies are large firm biased (Calvino et al, 2016; OECD and World Bank 2015; Brynjolfsson et al., 2008).

We explore initial evidence on how cloud adoption impacts firm IT costs in Table A3 in the Appendix. We consider whether cloud adoption leads to a substitution away from owning IT, measured here as IT investment (per employee). From column 1 we find cloud adoption is correlated with a significant *decline* in IT investments per employee of 42% over the sample period.[12] This effect is stronger for incumbent firms in column 2, with a 48% reduction in IT investment per employee over the sample period, compared to no significant relationship for young firms.[13]

Estimation Strategy

This paper uses instrumental variable estimation to measure how cloud computing affects firm growth and organization. Our second-stage panel fixed effects model is set out in equation (1). The dependent variable y, refers to various firm outcome variables, including employment, sales and sales per worker, but also measures of the concentration of activity, measured by

[12] Since the regression is log-linear, -42% is calculated as exp(-0.550)-1, using the estimated coefficient from column 1 in Table A3 in the Appendix. The same is applied throughout the paper for all log-linear regressions.

[13] These findings continue to hold if we use the instrumental variable approach from Section IV. The results are available upon request.

having multi-establishments, establishment deaths and establishment births (per firm) and the geographic dispersion of the firm.

$$y_{it} = \alpha_i + \alpha_t + \beta cloud_{it} * young + \beta cloud_{it} * incumbent + \gamma X_{it} + \varepsilon_{it} \qquad (1)$$

Our variable of interest, $cloud_{it}$ is a binary variable that measures the firms' use of any of the different forms of cloud computing services at time t. To discern potential heterogeneous effects by firm age we interact the cloud variable with a binary variable which measures whether the firm is young (10 years old or younger in 2008) or an incumbent (older than 10 years in 2008).

We include firm (α_i) and year (α_t) fixed effects in all our estimations (the latter are discussed in the next section). The specification therefore captures changes in firm outcomes driven by cloud adoption, removing the effect of any time invariant firm- industry- or location-specific confounding factors, as well as year specific shocks. X_{it} represents a vector of controls, including firm age, foreign ownership and size measured by the number of establishments.

Our instrumental variable approach combines both time-series variation and cross-sectional variation in broadband speeds that arise from access to fiber technologies in their local exchange and firm distance to this exchange. Accordingly, the first-stage regression in equation (2) relies on two instruments to predict firm cloud adoption: zip-code level access to fiber broadband (lagged one period) signified by a dummy variable $fiber_{it-1}$ and fiber broadband availability interacted with firm distance from the telephone exchange, $fiber_{it-1} *$

$dist$ (see next section for further discussion).[14] [15] [16] The identification is driven by differences between firms close to an exchange with and without fiber, compared with firms far from an exchange with and without fiber. We detail these instruments and their construction next.

$$cloud_{it} = \alpha_i + \alpha_t + \beta fiber_{it-1} + \beta fiber_{it-1} * dist_i + \gamma X_{it} + \varepsilon_{it} \quad (2)$$

Fiber Broadband Instrumental Variables

What is Fiber?

In the UK, fiber is the main source of high-speed broadband. Like its predecessors, dial-up and ADSL, it relies heavily on the structure of the historic telephone exchange network, using pre-existing exchange boxes and street cabinets. We use the mapping of the telephone network employed previously by DeStefano et al. (2018), which includes information on the location of all telephone exchanges in the UK (of which there are over 5,600) and the distances of each zip-code (of which there are over 1.7 million) to the exchange they are connected to. To this we add new information on the date of enablement of the exchange for fiber broadband from OFCOM (the UK telecoms regulator).

We consider the dominant form of fiber in the UK, fiber to the cabinet (FTTC), which uses a fiber optic cable between the exchange box and the cabinet. These fiber cables are more efficient in transmitting data offering faster speeds compared to the pre-existing copper cable

[14] These instruments are calculated using the location of the firm headquarters. Most firms in our sample are multi-establishment. We repeated the baseline regression restricting the sample to observations where the majority of establishments of the headquarters become connected to a fiber enabled exchange at the same time and find results consistent to our baseline.

[15] Instruments are lagged one year to allow for the adjust time between fiber enablement and cloud adoption.

[16] Note we estimate two first stages for each of the endogenous variables $cloud_{it} * young$ and $cloud_{it} * incumbent$ in equation 1, interacting equation 2 with the young or incumbent dummy for each first stage.

used by older vintages of broadband including ADSL. On average in the UK, FTTC averages around 24 mbps, whereas the maximum speeds of the previously dominant ADSL broadband technology are 8.0 mbps (BT Openreach, 2017).[17] Importantly for our study, fiber broadband offers symmetric upload and download speeds; upload speeds for fiber are 5 times faster than the advertised download speeds for ADSL. It is this ability to upload and download data at faster and more reliable speeds which meant that data could be effectively shared outside of the firm increasing the attractiveness of using cloud technologies.

Fiber enablement

Our first instrument, fiber availability in the local exchange, relies on time-series variation in the rollout of fiber broadband in the UK. Our dataset contains enablement information from the start of the rollout program, in 2009, to its completion in 2014 (See Figure 1: Panel A-D). This program enabled around 30% of all exchanges with the necessary technology.[18] The rollout was first announced in October 2008.[19] From this information we know the point in time when each zipcode gains access to fiber.[20]

[INSERT FIGURE 1 HERE]

Our second instrument exploits the fact that fiber speeds decline with distance to the telephone exchange, with longer distances associated with slower internet speeds (Ofcom, 2016). FTTC

[17] A small minority of establishments in the UK have fiber to the premise (FTTP), where the fiber network runs from the exchange to the local cabinet and on to the premises. While we do not have data on locations of FTTP, only 1.5% of households had FTTP the year after our sample (European Commission, 2015).

[18] The total number of telephone exchanges used in the regressions is 2,750, of which 1,627 are fiber enabled by 2014 and 1,123 are enabled after 2014.

[19] We exclude from the sample exchanges in Northern Ireland and Cornwall as these were enabled in a joint venture with BT and there is limited data on exchange enablement dates.

[20] Zipcodes in the UK represent small geographies. On average 12 premises (households/businesses) per zipcode.

speeds deteriorate at faster rates per meter than under earlier ADSL broadband, as shown in Figure 2. For cable distance of 2,000 meters from the cabinet, FTTC connections speeds are less than a quarter of those with a cable distance of 200 meters (80 mbps compared to 17 mbps). Therefore, while fiber provides a substantial improvement over the earlier ADSL technology, this is true only for distances within 1000 meters.

[INSERT FIGURE 2 HERE]

Figure 3 illustrates the differences in the crow-flies distance to the exchange that we use for firms in our sample. When combined with the evidence from Figure 2, these suggest large disparities in fiber speed across firms. The crow-flies distance between the median firm and their exchange is 1.1 kilometers, at the 25th percentile this distance is around 550 meters and the 75th percentile it is roughly 1,900 meters.

[INSERT FIGURE 3 HERE]

Instrument Relevance

In Table 2 we provide evidence that fiber enablement and cable distances within fiber enabled exchange areas predict the adoption of cloud, even when including firm and year fixed effects. We report these regressions using a linear measure of distance (column 1) and a version in which we place firms into separate bins according to their cable distance (column 2).

Across all regressions we find that firms attached to fiber enabled telephone exchanges are significantly more likely to adopt cloud. We also find that this effect declines with the cable distance between the firm and the telephone exchange, within fiber enabled areas. In column 1 the cable distance variable is negative and suggests that for every kilometer increase in distance, the probability of adopting cloud drops by 3%.[21] In column 2 where we place firms in distance bins, we find that firms less than 500 meters from the exchange are significantly more likely to adopt cloud. Between 500 and 1000 meters, the estimated coefficient is positive, but not significant at conventional levels. Beyond this, the effects of fiber distance continue to fall towards zero.[22] In both columns 1 and 2, this is consistent with the engineering prediction on the distance dependency of fiber and that fiber speeds converge with ADSL at roughly 1,000 meters from the exchange (Heath 2013).

[INSERT TABLE 2 HERE]

Instrument Validity

The validity of our instruments requires that fiber enablement and cable distances have no effect on firm performance independent of their relationship with cloud. We discuss this issue below and detail how we deal with potential objections through sample restrictions, including firm fixed effects and testing for pre-treatment trends.

The cable distance instrument depends on the location of the firm and of the telephone exchange. The location of the telephone exchange is based on pre-existing telephone infrastructure dating back in some cases as far as the 19th century. Firms born before the development of fiber broadband are unlikely to choose their location based on a technology that had yet to be invented. Firms also seem unlikely to relocate since moving costs would be

[21] We explore fiber access and the seven different types available as well as the Eurostat classification of low, medium and high-tech cloud (Eurostat, 2018) in Table A4 in the Appendix.

[22] Firms more than 2000 meters from the exchange are the baseline category

large relative to the temporary difference in fiber access.[23] We minimize the above risks by excluding from the sample firms born after 2008 – the year fiber was announced.

A further challenge to the validity of our cable-distance instrument is passive sorting. Telephone exchanges are typically sited near commercial centers and concentrations of residential property and, to aid with the laying of cabling, near major road junctions. Plausibly, firms may also benefit in other ways from locating close to commercial agglomerations and major road junctions. To the extent that these geographic factors or firm characteristics are time invariant over our 8-year time window, such factors will be captured by the firm fixed effects.

Firm fixed effects are not a solution to all possible objections to these instruments, if for example, fiber enablement had been targeted at areas with already fast-growing firms. To consider this, we conduct reduced-form event study analyses to examine the timing of firm performance impacts around the year of fiber enablement.[24] We use the approach of Callaway and Sant'Anna (2021), which accounts for staggered treatment designs and heterogeneous treatment effects, by comparing firm performance effects of fiber enablement for each cohort to those firms that are never fiber enabled during our sample period.[25] We start by examining our first instrument (fiber enablement) with pre and post enablement trends for cohorts of young and incumbent firms (panel a). Next, reflecting on the identifying variation of our second instrument, fiber enablement * distance, we examine the pre- and post- fiber enablement trends for the sub-sample of firms that are near (< 1km) their exchange and then as a separate regression those that are far (>=1km) from the exchange (panel b). We also examine these

[23] In our sample 214 firms report a different address during the sample period. Of the firms that move, 64% move to a new location where the telephone exchange was enabled later than their original exchange. On average these firms move to locations 0.12km farther from an exchange, with slower fiber speeds. We have excluded all movers.

[24] These regressions use annual data from 2008 to 2015.

[25] Similar results are obtained by including not-yet enabled along with the never enabled in the control group.

trends for sub-samples of young firms that are near and far from the exchange and then the same for incumbent firms – which we expect to have different post-treatment effects (panels c and d). We show these for employment in Figure 4 and for sales and labor productivity in Appendix Figures A1 and A2.

Reassuringly, we find no evidence of pre-enablement trends for the timing of enablement for both young and incumbent firms (panel a). Comparing young firms (the blue line) or incumbent firms (the red line) with those connected to never enabled exchanges, the coefficient estimates in the pre-enablement periods are statistically insignificant and, importantly display no obvious trend across the pre-enablement period. Similarly, there are no apparent pre-trends for firms near or far from their exchange (panel b), compared to firms in never enabled areas. Nor do we see evidence of pre-enablement trends for any combination of young, incumbent, near or far firms (panels c and d). After fiber enablement, we find post-treatment effects that mirror our later baseline IV results. Firstly, young firms connected to an enabled exchange experience marked post treatment effects in contrast to incumbent firms (panel a). Secondly, firms closer to their exchange, with fast expected fiber speeds needed for cloud, show marked employment growth in the two years after enablement (panel b). In contrast, firms that are more than 1km from their exchange, where expected fiber speeds are similar to the prior ADSL technology (see panel), show no significant change in employment. We also find stronger post-treatment employment growth for young in panel c, where again these are stronger for firms closer to their exchange (panels c). There are no strong post-treatment effects for incumbents in panel d. In Appendix Figures A1 and A2, we similarly find no evidence of pre-enablement trends for sales or labor productivity, with some evidence of post-treatment growth in sales and labor productivity.

[INSERT FIGURE 4 HERE]

A final validity concern arises from the existing literature, which examines how earlier upgrades to broadband speed, mostly the introduction of ADSL, impacted various important economic outcomes, including skill premiums (Akerman, et al., 2015), firm scale (DeStefano et al., 2018), along with broader economic and social impacts (Falck et al., 2014; Bhuller et al., 2013; Ahlfeldt et al., 2017). Thus, one potential concern is whether the diffusion of fiber impacts firms in ways other than through the use of cloud computing. The event study plots presented in Figure 4 provide additional reassurance on this point. For example, the lack of pre-treatment effects for all firms, as well as post-treatment effects for firms greater than 1km from the exchange in these figures (the point where fiber speeds slow to the predecessor ADSL broadband speeds) suggests any threat to the validity of the instruments would need to come from a confounder that is impacted by high-speed fiber, as opposed to fiber availability per se. Figure 4 panels c and d for the incumbent and young firms also suggest this confounder would need to impact younger firms that are close to the exchange rather than older firms that are similarly close.

In comparison to previous internet technologies fiber is distinct, particularly by offering substantially faster upload speeds (of up to 80MBps versus 5Mbps).[26] Faster fiber upload and download speeds allows real-time data sharing and collaboration through the cloud in a way that was not possible before, owing to the asymmetry between ADSL upload and download speeds. While these speeds could in principle be related to other forms of data sharing, such as sending data to and from in-house servers, these non-cloud data transfer systems have existed for decades and like most non-cloud forms of IT, incur large sunk costs. These costs are in sharp contrast to the decentralized flexible storage, database hosting and processing offered by cloud technologies. If these large sunk-cost, non-cloud data sharing systems were

[26] At the start of our sample period, in 2008, 80% of UK firms had adopted ADSL broadband. The UK-wide rollout of ADSL broadband was completed by 2007, although urban areas (that were the focus of fiber rollout) were enabled much earlier - 60% of UK businesses already had ADSL available by 2002 (DeStefano et al., 2018).

driving our findings, we would anticipate results that mirror those found in studies on earlier digital technologies. Previous research has found strong evidence that earlier digital technologies are large firm biased (Calvino et al, 2016; OECD and World Bank 2015; Brynjolfsson et al., 2008), while we find that performance gains are concentrated amongst young firms, which is more consistent with recent work on cloud and scale without mass (Jin and McElheran, 2017). In fact, we find that the adoption of cloud leads to a reduction in IT investment per employee (see Table A3 in the Appendix). Cloud therefore appears to have effects that are distinct from older IT technologies.

To provide additional support for instrument validity, we first assess whether fiber predicts IT technologies that the literature has shown to be related to predecessor forms of broadband, and secondly, if fiber predicts the use or sharing of data that does not need to be related to the cloud. For the former set of technologies, we use the percentage of employees with PCs, the percentage of employees with mobile phones, whether firms made sales through e-commerce, and the proportion of online sales to total sales. We also assess technology variables which facilitate the use or sharing of data both within and outside of the firm, including measures of production line tracking (radio frequency identification, RFID), using software and databases for marketing and customer planning (Enterprise Resource Planning software, ERPs and Customer Resource Planning software, CRMs), measures of IT skills (use of IT specialists, IT training) and if the firm shares data with suppliers and customers.

We find scant evidence that our instruments are correlated with these other types of digital technology (See Table A5 in the Appendix). This is reflected both in the absence of statistical significance between the instruments and the use of these technologies and demonstrated by the small first-stage F statistics. These results provide reassurance for the absence of potential confounders that might explain our results.

Results

Firm Scale and Performance

Before presenting the instrumental variable results, we begin by examining whether the use of cloud is positively correlated with measures of firm performance using OLS regressions (See Table 3). Note, that the OLS regressions include the same fixed effects and control variables as the instrumental variable results discussed below. As expected, we find cloud adoption is associated with greater employment, sales and labor productivity (columns 1, 2 and 3) for all firms, with stronger correlations found for young firms (columns 4, 5 and 6). [27]

[INSERT TABLE 3 HERE]

Table 4 presents the instrumental variable estimates for the effects of cloud adoption on firm growth. In columns 1-3, we estimate the effects of cloud for all firms where in columns 4-9 we allow for separate effects for young and incumbent firms. The interaction terms are expressed such that they estimate the effect for young and incumbent firms separately, and therefore the estimated coefficient for each type is tested against the null of a zero effect. [28]

In the first stage, reported in Table A6 in the Appendix, we find that being attached to a fiber enabled exchange increases the probability of adopting cloud by 10% for all firms, 17% for

[27] Disaggregated forms of cloud are also positive statistically related to firm performance except when we measure performance by employment and use finance and accounting software and CRM software cloud services. These lie just outside of significance at the 10% level. We choose not to report these regressions for brevity.

[28] In Table A7 and A8 we explore whether they are driven by small rather than young firms. We find for some outcomes, such as employment, sales or labor productivity, that both age and size matter and the results are strongest for small and young.

incumbent firms, and by 29% for young firms. We also find that each kilometer from the exchange reduces the propensity to adopt cloud by 3% for all firms, by just over 2.1% for incumbent firms and by 4.2% for young firms. The first stage F-statistics suggest the instruments are strong predictors of cloud adoption. For the pooled sample the Cragg-Donald F-statistic is 46.63 and 24.62 with the age distinction regressions, exceeding the Stock-Yogo critical values for weak instruments[29], and Kleibergen-Paap F-statistics that account for clustering, exceed 12 in all specifications. The test for overidentification is also comfortably passed for all regressions, with the relevant p-value reported in the table.

In the second stage regressions we find some evidence for performance gains on average for all firms in columns 1-3, where these effects are statistically significant for sales and labor productivity. When exploring heterogeneous gains from cloud by the age of the firm in columns 4-6, we find outcomes that are consistent with this idea of differences across young and incumbent firms. In columns 4 and 5 we find that cloud leads to significant increases in employment and sales for young firms, but not incumbent firms. As our data are measured for the years 2008, 2013 and 2015, this equates to a 14.3% increase in employment each year for young firms over this 7-year period and a 15.9% annual increase in sales.[30][31] These estimated effects compare to the mean annual employment growth rate of 6.4% and sales growth rate of 12.4% for young firms (shown in Table A2). As the mean young firm has fewer than 20 employees, these represent large percentage changes on a small absolute number of employees.

---

[29] Critical value for maximal IV relative bias of 5% is 11.01 and 10% maximal IV size for false positives is 16.87.

[30] 14.3% annual employment growth is calculated as exp(0.938/7)-1, sales growth is calculated similarly.

[31] Following the evidence reported in Table A4 of a stronger effect of the instruments on the use of cloud for data and for storage and high-tech cloud, we report results latter group in Table A9 in the Appendix and the former in the next section. The results are very similar to those in Table 4, suggesting our Local Average Treatment Effect reflects firms that adopt cloud for data and storage or high-tech cloud.

The results in column 6 also suggest evidence of a labor productivity effect for both groups of firms, with larger gains for young firms.

The IV coefficients are somewhat larger than the OLS correlations. If our fiber instrument disproportionately predicts cloud adoption for a subset of firms that have larger treatment effects, then OLS estimates can be downwards biased. We provide evidence in line with the view that the IV results present a local average treatment effect, compared to an average treatment effect for the OLS correlations. We find young firms are 5.5 percentage points more likely to adopt cloud because of fiber enablement than incumbents, with no differential effect of distance.[32]

We subject our results to a barrage of robustness checks and find our baseline firm performance impacts are largely unchanged. In order to ensure that our results are not somehow driven by young firms self-selecting into areas before the rollout was announced, we rerun the results for a sample where all firms were born during or before 2006. These results are also robust to dropping London firms from the sample or adding region-time fixed effects to control for region specific shocks. Additional robustness tests excluded fast growing firms by winsorizing the top and bottom 5% of firms by employment, sales or labor productivity growth and allowing for differential trends by firm size by including employment quartile-year fixed effects.[33]

We can also use the estimates to roughly calculate the aggregate effects of cloud adoption. To do so we repeat our baseline estimation applying sampling weights.[34] From these results,

[32] To assess whether this might be the case, we examine our first stage estimates for young and incumbent firms, by repeating Table 2 column 1, but with a young interaction (see Table A10 in the Appendix).

[33] The results are available upon request.

[34] These aggregate estimates are approximate as they do not account for general equilibrium effects, some sectors and firms below 10 employees are not surveyed in the data we use. The results are available upon request.

we find that the mean young firm was 13 percentage points more likely to use cloud as a result of the fiber rollout, which translates into 32 percentage points more jobs for young firms between 2008 and 2015, or about 4 percentage point more jobs for the UK as a whole.[35] We focus on the jobs created by young firms, since we do not find a statistically significant effect on employment growth of incumbent firms or all firms on average.

[INSERT TABLE 4 HERE]

Heterogeneity

Age heterogeneity

To both assess the heterogeneity of firm age in a more nuanced way and confirm the robustness of our previous results we introduce as alternatives a continuous age interaction term (centered on the mean in 2008 of 16.5 years old) and age quartiles, defined as age groups 0-10, 11-20, 21-30 and older than 30, again based on 2008 values.[36] We note that when using the latter, we estimate four separate first stage regressions for cloud adoption. The effect of this is smaller F statistics, suggesting weaker power from the instruments when using this approach.

Table 5 presents results for both continuous age interaction mean-centered variables (columns 1-3) and age quartiles (columns 4-6) for log employment, log sales and labor productivity. For the mean centered regressions, the (non-interacted) cloud variable provides information on the effects of cloud for the firm of mean-age in 2008. Here we find cloud has an effect on labor productivity of the mean firm (aged 16.5 years), but not employment or sales. The negative and significant coefficients on the age-cloud interaction (in column 1 and 2) implies that the effects of cloud increase in employment or sales as we move below the mean age (and decrease as we move above it), again suggesting stronger effects of cloud on younger firms.

[35] These results are robust to the exclusion of the top 1% of young or incumbent firms based on their employment.

[36] Similar results are obtained when defining the youngest firms as 0-5 years, rather than 0-10, see Table A11.

When looking at heterogeneity in age quartiles we find that the most salient distinction is the bottom quartile of 10 years and younger, albeit with some differences across outcome measures. In terms of employment or sales, firms between the ages of 0-10 obtain the largest scale effects from cloud. The estimated coefficient on cloud is similar to our baseline for the youngest group of firms. This effect declines with age and is insignificantly different from zero for firms aged 11 or older. In column 6 we find labor productivity effect are strongest for firms older than 30 and amongst the youngest cohorts.

[INSERT TABLE 5 HERE]

Cloud use for data and storage

The previous sections examined firm performance effects using an aggregate measure of cloud computing that encompassed cloud functions as varied as email, accessing software, and the storage and processing of data. Our earlier evidence showed that fiber speeds predicted the use of cloud for databases and storage particularly strongly, consistent with the declines in IT investment we also found. In this section we use a cloud measure that reflects either of these two functions.

The results in Table 6 demonstrate consistent results regarding the signs and significance of our outcome variables and similar sized F statistics, to the aggregate cloud variable. However, the sizes of the magnitudes for all outcome variables are larger, suggesting that the outcome from adopting this type of cloud is even more pronounced. For employment, we find that cloud hardware adoption by young firms leads to a 17.4% annual growth rate over the sample period.

[INSERT TABLE 6 HERE]

Industry heterogeneity

While cloud appears to result in distinct performance gains by age, the literature also suggests that there may be further differences according to the industry of the firm. It is well-established that the returns to new digital technologies are greater in industries that are more knowledge intensive (Bloom et al 2012; Bresnahan et al 2002). Early research identified important links between the levels of human capital and knowledge intensity (Acemoglu, 1997). Drawing on detailed data on research and development expenditures from the ARD, we construct knowledge intensity measures to see whether firms residing in these industries obtain greater gains from cloud.[37]

Employment increases for young firms in knowledge intensive industries from cloud by more than young firms in industries that are not knowledge intensive (Table A12), suggesting that the average effect for young firms found in Table 4 may partially be driven by this heterogeneity. The coefficient on employment growth is nearly half the size in less knowledge intensive sectors compared to knowledge intensive sectors (0.636 versus 1.297 = 0.636+0.661). For young firms in less knowledge intensive industries, employment growth because of cloud adoption is 9.5%, compared with 20.3% in young knowledge intensive sectors. In general, however, we find relatively limited heterogeneity across various industries classifications.[38]

Geographic Organization

Cloud is likely to impact how firms organize geographically. The reduced reliance on centralized IT departments combined with the homogenous and flexible information access across the organization may enable greater geographic dispersion of activity within the firm.

[37] The knowledge intensive measure is constructed with R&D expenditures (weighted by employment) at the 5-digit UK SIC level. Knowledge intensive industries are those in the top quartile of the distribution.

[38] We also assessed differences between manufacturing vs service sector firms, Skilled Tradeable Sectors and others (following the industry classification of Eckert et al., 2020), or Knowledge Intensive Activity Sectors compared to others (following Eurostat, 2014). We found limited noticeable differences between firms in these sectors, with results very similar to those in Table 4. Results are available upon request.

Conversely, advances in IT have often gone hand-in-hand with the increased importance of face-to-face communication and the rise of tech clusters (Greenstein et al., 2018).

We introduce different measures of the geographic dispersion of firm activity in Table 7. Our measures reflecting the geographic reorganization of the firm include becoming a multi-establishment and the births and deaths of establishments. In addition, we examine changes in the organization of employees from the headquarters such as the (employment) weighted and unweighted average distance between establishments and their headquarters and the distance-employment covariance term. Finally, we add a measure of the number of local authorities in which the firm has establishments in. Equations detailing the geographic dispersion measures can be found in the Appendix.

For young and incumbent firms, we find that cloud adoption impacts geographic dispersion (see Table 7). Young firms adopting cloud are less likely to become multi-establishments. Taken together with the results found in Table 4 on employment, implies a scale without mass effect from cloud for young firms previously documented anecdotally. For incumbents, the results demonstrate restructuring effects from cloud, signified by the positive and significant coefficient for establishment deaths and the negative and significant coefficient for establishment births.

The results in Table 7 also demonstrate that cloud adoption leads to the average employee working 21.71km and 18.93km farther from their headquarters for young and incumbents, respectively. For both types of firms, we fail to find evidence that they are systematically more likely to close or open farther or more proximate establishments – as reflected in the unweighted distance variable. There is some evidence that firms are redistributing employment towards establishments more distant from the headquarters, there is a positive effect of cloud on the employment-distance covariance variable albeit borderline significant. Cloud therefore appears to enable decentralization of information within the firm, with some nuances for young and incumbent firms. First stage results are reported in Table A13 in the Appendix.

To track whether these are general movements in employment within the firm, as opposed to the shifting of employment towards or away from the headquarters we combine our firm level information with data for a random 1% sample of all workers in the UK. The results are presented in Table B2 and discussed in more detail within the Appendix. These results suggest that cloud adoption affects employee mobility within the firm, but this reorganization of activity is largely across different establishments rather than to and from the headquarters.

Taken together, these results suggest that cloud facilitates decentralizing activity to local establishments away from the headquarters, even relocating workers throughout the firm. Within the firm, cloud enables employees on average to work farther from the headquarters without increasing their geographical footprint.

[INSERT TABLE 7 HERE]

Conclusion

This paper presents new evidence on the mechanisms of cloud adoption and its impact on firm growth and geographic reorganization. We use novel instrumental variables on zip-code level availability and expected speeds (using local loop distances) of fiber broadband to predict firm cloud adoption.

The empirical evidence suggests there are differential impacts of cloud adoption on younger and incumbent firms. Younger firms that adopt cloud are more likely to increase employment, sales, and labor productivity. For incumbent firms, we find no scale and weaker productivity impacts. We find instead they are more likely to reorganize activity by closing establishments. A back of the envelope calculation suggests cloud leads to about 4 percentage points more jobs for the UK as a whole between 2008 and 2015. For both young and old firms, we find they disperse employment farther from the headquarters as a result of cloud. Cloud along with the fiber infrastructure therefore enables young firms to scale, and allow firms more generally to reorganize, increase their productivity and geographically disperse.

Consistent with discussions in the IT literature, our evidence suggests that cloud is distinct from earlier IT technologies, which reinforced the scale advantages of incumbents (see for instance Lashkari et al., 2019). Cloud reduces a firm's fixed costs of IT, which we measure as a decline in firm investments in IT. Cloud also decentralizes data, processing and software availability throughout the firm, going beyond earlier IT that allowed access to information for specific tasks or workers, such as Enterprise Research Planning and CAD/CAM software (Bloom et al., 2014). Consistent with these earlier technologies, the dispersion of economic activity appears to follow the dispersion of information.

References

Acemoglu, Daron, 1997. "Technology, unemployment and efficiency," *European Economic Review*, vol. 41(3-5), pages 525-533,

Ahlfeldt, G. M., Koutroumpis, P. and Valletti, T. (2017). 'Speed 2.0 evaluating access to universal digital highways.' *Journal of the European Economic Association*, Vol. 15(3). 586-625.

Akerman, A., Gaarder, I., and Mogstad, M., (2015). 'The skill complementarity of broadband internet' *Quarterly Journal of Economics*, 130 (4), 1781-1824.

Bhuller, M., Havnes, T., Leuven, E., and Mogstad, M. (2013). 'Broadband Internet: An Information Superhighway to Sex Crime', *The Review of Economic Studies*, *80*(4), 1237–1266.

Bloom, N., Garicano, L., Sadun, R., Van Reenen, J., (2014). 'The distinct effects of information technology and communication technology on firm organization'. *Management Science*, 60, 2859-2885.

Bloom, N., and Pierri, N. (2018). 'Cloud Computing Is Helping Smaller, Newer Firms Compete'. *Harvard Business Review*.

Bloom, N., Sadun, R., and Van Reenen, J., (2012). 'Americans Do IT Better: US Multinationals and the Productivity Miracle', *The American Economic Review*, 102, 167–201.

Bommadevara, N., Del Miglio, A. and Jansen, S. (2018). 'Cloud adoption to accelerate IT modernization', *Digitial McKinsey: Insights,* December, "Creating value with the cloud".

Bresnahan, T., Brynjolfsson, E., and Hitt, L. (2002). 'Information Technology, Workplace Organization, And The Demand For Skilled Labor: Firm-Level Evidence.' *The Quarterly Journal of Economics*, 117 (1): 339–76.

Brynjolfsson, E.et al. (2008). ‘Scale Without Mass: Business Process Replication and Industry Dynamics’, Harvard Business School Technology & Operations Mgt., Unit Research Paper, No. 07-016.

BT Openreach. (2017). ‘Annual Report and From 20-F’, Online at file:///Users/lish/Downloads/annual-report-2017%20(2).pdf

Callaway, B., and Sant’Anna, P.. (2021). ‘Difference-in-Differences with multiple time periods’, *Journal of Econometrics,* Vol 225(2). 200-230,

Calvino, F., Criscuolo, C. and Menon, C. (2016). ‘No Country for Young Firms?: Start-up Dynamics and National Policies’, OECD Science, Technology and Industry Policy Papers, No. 29, OECD Publishing, Paris.

Crafts, N., and Mills, T. (2020). ‘Is the UK Productivity Slowdown Unprecedented?’ *National Institute Economic Review,* 251, R47-R53.

Decker, R., Haltiwanger, J. Jarmin, R. and Miranda, J. (2014). ‘The role of entrepreneurship in U.S. job creation and economic dynamism’, *Journal of Economic Perspectives*, 28, 3-24.

Decker, R., Haltiwanger, J. Jarmin, R. and Miranda, J. (2016). ‘Where Has All the Skewness Gone? The Decline in High-Growth (Young) Firms in the U.S.?’, *European Economic Review*, 86, 4-23.

DeStefano, T., Kneller, R. and Timmis, J. (2018). ‘Broadband Infrastructure, ICT use and Firm Performance: Evidence for UK Firms’, *Journal of Economic Behavior and Organization,* 155, 110-139.

Draca, M., Raffaella S., and Van Reenen, J. (2006). ‘Productivity and IT: a Review of the Evidence’, CEP Discussion Paper, 749.

Duranton, G. and Puga, D. (2005). ‘From sectoral to functional urban specialization’, *Journal of Urban Economics*, 57(2): 343–370.

Eckert, F., and Ganapati, S., and Walsh, C. (2020). 'Skilled Scalable Services: The New Urban Bias in Economic Growth', Available at http://dx.doi.org/10.2139/ssrn.3439118

European Commission. (2015). Broadband Coverage in Europe 2015. Online at https://digital-strategy.ec.europa.eu/en/library/broadband-coverage-europe-2015

Eurostat. (2014). 'High-Tech Industry and Knowledge-Intensive Services (htec).' http://ec.europa.eu/eurostat/cache/metadata/EN/htec_esms.htm.

Eurostat. (2018). 'Cloud Computing - Statistics on the Use by Enterprises'. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Cloud_computing_-_statistics_on_the_use_by_enterprises#Enterprises.E2.80.99_dependence_on_cloud_computing

Fabling, R. and Grimes A. (2016). 'Picking up Speed: Does Ultrafast Broadband Increase Firm Productivity?' Motu Working Paper 16-22.

Falck, O., Gold, R. and Heblich, S. (2014). 'E-lections: Voting behavior and the internet.' *American Economic Review*, 1047, 2238-2265.

Forman, C., Goldfarb, A. and Greenstein, S. (2015). 'Information Technology and the Distribution of Inventive Activity'. In The Changing Frontier: Rethinking Science and Innovation Policy, eds. Adam Jaffe and Ben Jones, University of Chicago Press.

Garicano, L. and Heaton, P., (2010). 'Information Technology, Organization, and Productivity in the Public Sector: Evidence from Police Departments', *Journal of Labor Economics*, 28, 167–201.

Greenstein, S., Goldfarb, A. and Forman, C. (2018). 'How Geography Shapes—and Is Shaped by—the Internet', In *The New Oxford Handbook of Economic Geography*, edited by Gordon Clark, Maryann Feldman, Meric Gertler, and Dariusz Wojcik, 269–285.OUP

Heath, M. (2013). 'Chart of BT Fibre Broadband FTTC (VDSL2) Speed Versus Distance From the Cabinet

Iansiti, M., and Richards, G. (2011) ‘Economic Impact of Cloud Computing’ White Paper,

ITU (2017). The State of Broadband: Catalyzing Sustainable Development”, Online at: https://www.itu.int/dms_pub/itu-s/opb/pol/S-POL-BROADBAND.18-2017-PDF-E.pdf

Jin, W. and McElheran, K. (2017). ‘Economies Before Scale: Survival and Performance of Young Plants in the Age of Cloud Computing’. *Rotman School of Management Working Paper No.* 3112901.

Lashkari, D., Bauer, A. and Boussard, J. (2019). ‘Information Technology and Returns to Scale’, *Mimeo.*

Leamer, E.E. and Storper, M. (2001). ‘The economic geography of the Internet age’, *Journal of International Business Studies*, 32(4): 641–665.

National Institute of Standards and Technology, (2011). ‘The NIST Definition of Cloud Computing’, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

OECD (2015). ‘OECD Digital Economy Outlook 2015’, OECD Publishing, Paris.

OECD (2014). ‘Cloud Computing: The concept, impacts and the role of government policy’,

OECD and World Bank (2015), “Inclusive global value chains: Policy options in trade and complementary areas for GVC integration by small and medium enterprises and low-income developing countries”, OECD and World Bank Group Publishing.

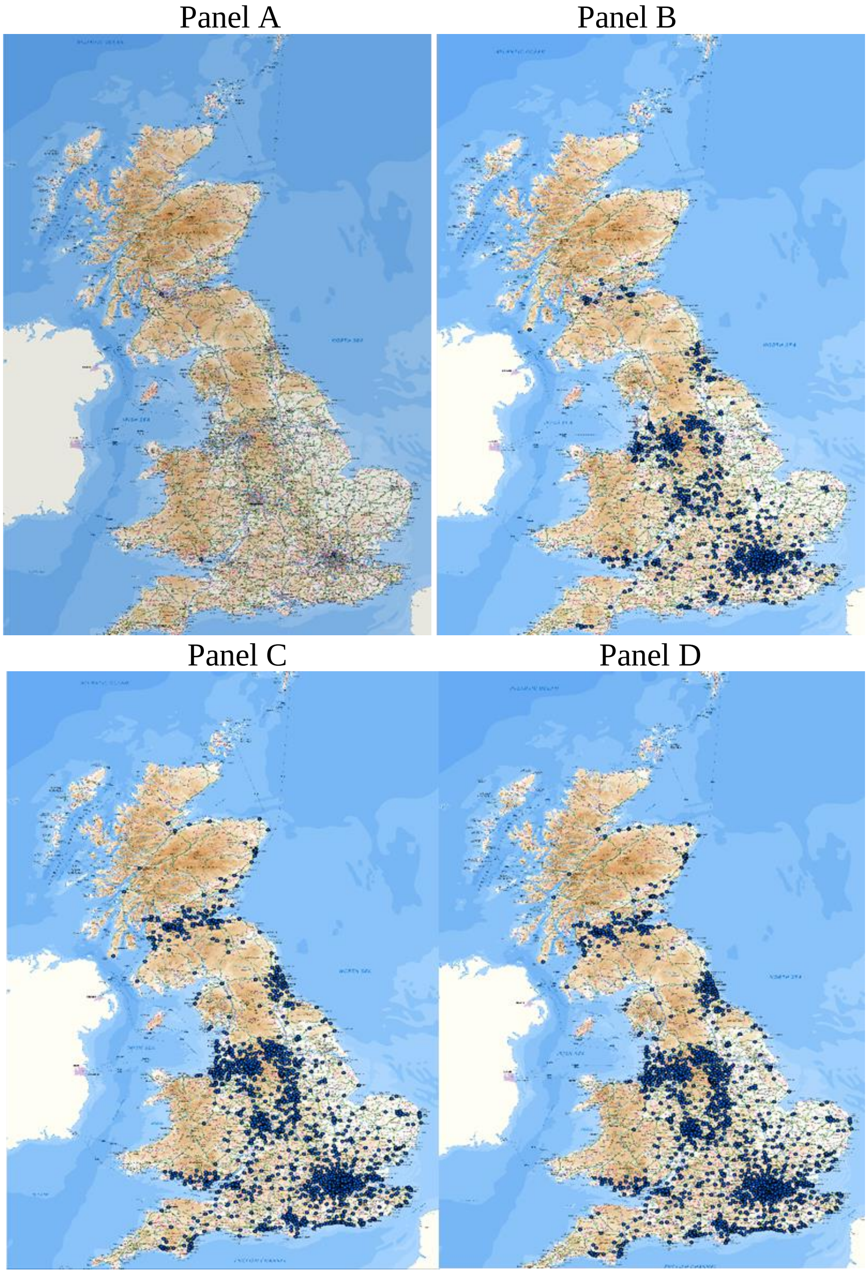
Ofcom (2016). ‘UK Home broadband performance: A consumer summary of fixed-line broadband performance provided to residential consumers’

Olley, S. and Pakes. A. (1996). ‘The Dynamics of Productivity in the Telecommunications Industry’, *Econometrica*, 64(6), 1263-1298.

Van Ark, B., (2016). ‘The Productivity Paradox of the New Digital Economy’, *International Productivity Monitor*, 31, 3-18.

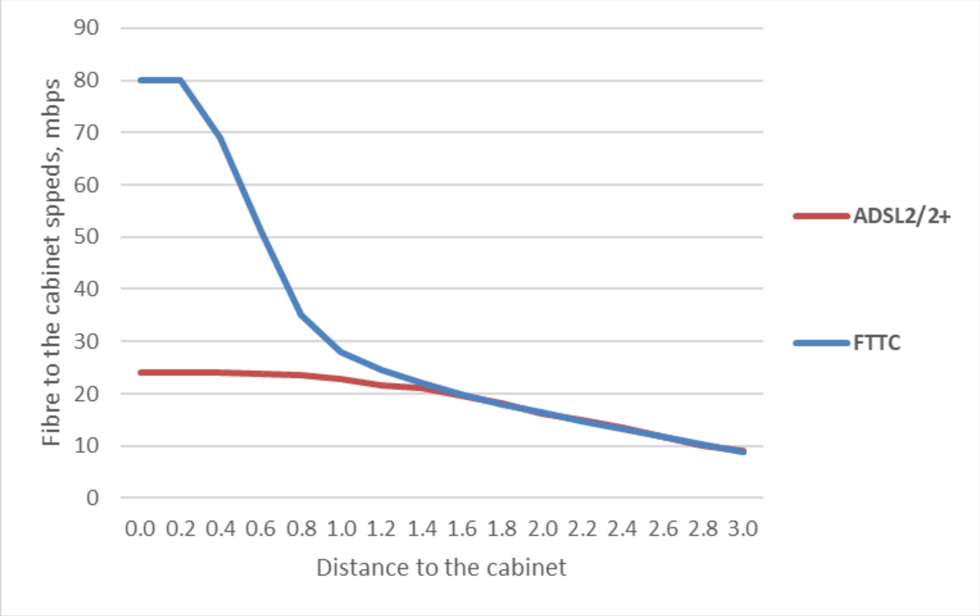
FIGURES AND TABLES

Figure 1: Location of Fiber Enabled Exchanges by 2009, 2011, 2013, 2014



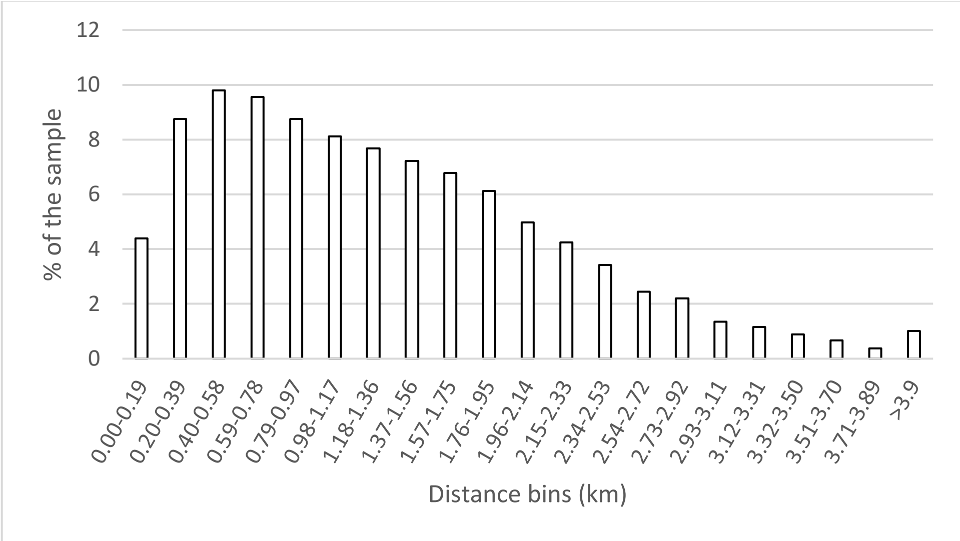
Notes. Points represent the location of fiber enabled exchanges in each year.

Figure 2: Fiber to the cabinet connection (FTTC) speeds and distance to the cabinet



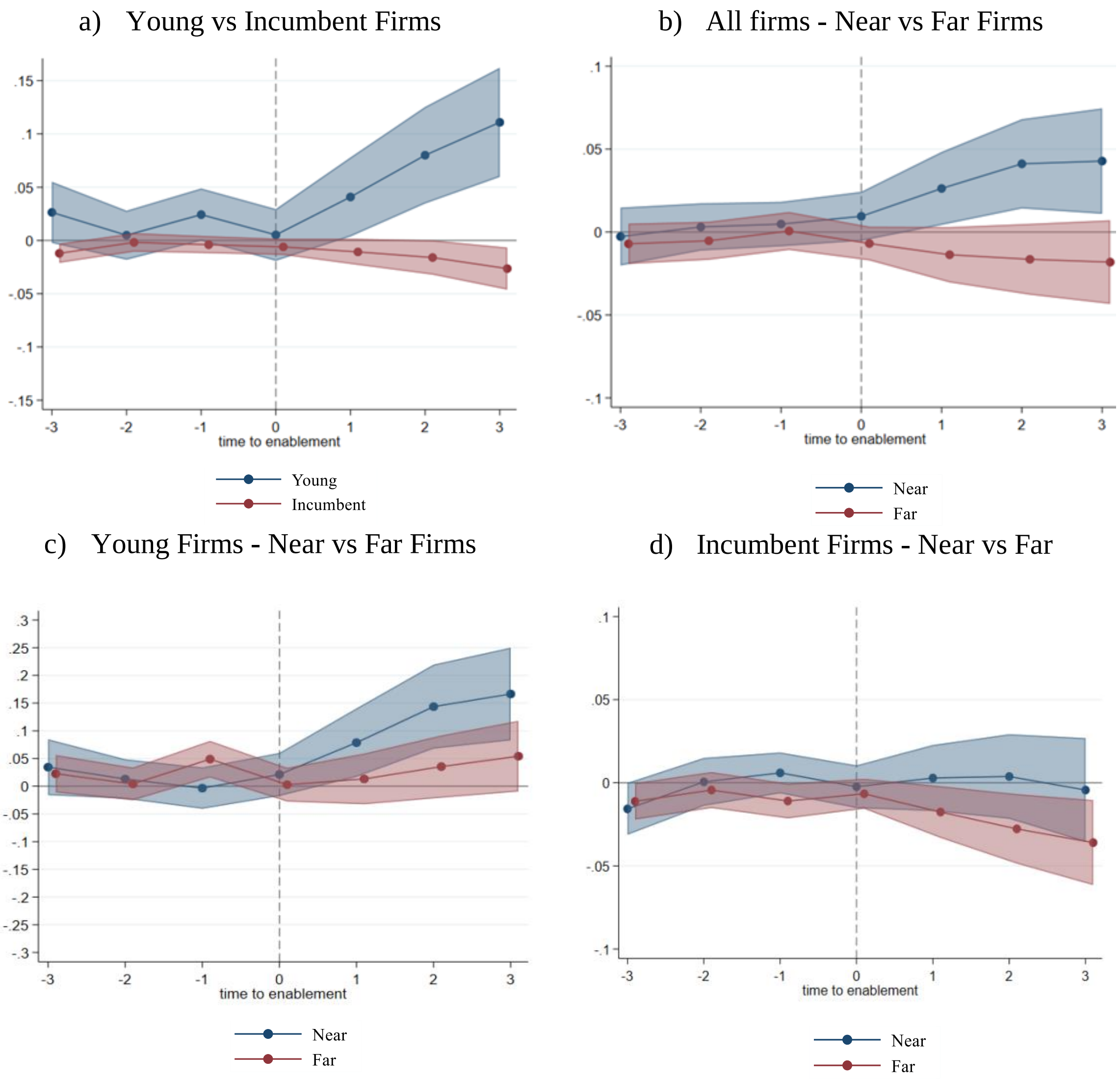
Notes. The figure illustrates expected fiber to the cabinet (FTTC) and ADSL (the prior technology) broadband speeds by distance (km) from the cabinet and telephone exchange respectively.

Figure 3: Firm crow-flies distance to the local telephone exchange



Notes. The figure shows a histogram of the distribution of firms based on their distance from their local telephone exchange.

Figure 4: Employment pre-trends with timing of fiber enablement



Notes: The above figures present the coefficients and 95 per cent confidence intervals of event study regressions – reflecting firm employment in the periods before and after fiber enablement (our instrumental variable). Estimation follows the method of Callaway and Sant'Anna (2021), which estimates enablement-cohort by cohort compared to never fiber enabled firms. Panel a) presents event study plots for young and incumbent firms respectively. Panel b) plots firms near and far from their exchange (less than and more than 1km respectively), and panel c) and d) present similar plots for the subsamples of young and incumbent firms, respectively. Note that the regressions that underlie the lines on each of the graphs are estimated relative to never enabled areas and should be interpreted as such. For example, in panel c) the near-young firms are compared to the near-young firms connected to never enabled exchanges. Note the differing scale across the four panels. Observations during the firm's year of fiber enablement are given by event time equal to 0, the preceding year is event time of – 1, the year after enablement is + 1, and so on. For parsimony, a subset of event-time dummies (three years before and after) are shown in the figures.

Table 1: Summary Statistics of Cloud Adoption

	All firms		Young firms		Incumbent firms	
Variable	**mean**	**st.dev.**	**mean**	**st.dev.**	**mean**	**st.dev.**
Cloud	0.219	0.413	0.168	0.374	0.235	0.424
Cloud Databases	0.090	0.286	0.077	0.267	0.094	0.292
Cloud Storage of files	0.139	0.346	0.112	0.316	0.148	0.355
Cloud Email	0.115	0.319	0.097	0.296	0.121	0.326
Cloud Office Software	0.077	0.267	0.067	0.251	0.081	0.272
Cloud Finance Software	0.049	0.215	0.051	0.221	0.048	0.213
Cloud CRM	0.064	0.244	0.051	0.220	0.068	0.252
Cloud Processing Own Software	0.054	0.226	0.046	0.209	0.057	0.231

Notes. These statistics are from a balanced panel of observations for comparison of adoption across time for the same set of firms – a subset of our estimation sample of firms over years 2008, 2013 and 2015. Young are defined as being aged 10 years old or younger in 2008 and incumbent are defined as being older than 10 years old in 2008. There is zero adoption of cloud in 2008, hence lower values are reported for all years (2008, 2013 and 2015). Use of the cloud technologies by year and grouped into low, medium and high technology uses following Eurostat (2018) are in Appendix Table A1.

Table 2: First stage: fiber enablement and distance on cloud adoption

	(1)	(2)
Dependent variable: Cloud adoption		
Fiber Enablement	0.103***	0.056***
	(0.015)	(0.017)
Fiber *Distance	-0.026***	
	(0.008)	
Fiber, Dist. < 500 meters		0.052**
		(0.021)
Fiber, Dist. 500-1000 meters		0.031
		(0.020)
Fiber, Dist. 1000-1500 meters		-0.022
		(0.020)
Fiber, Dist. 1500-2000 meters		0.005
		(0.021)

*Notes: The estimation in column 1 corresponds to equation 2, with column 2 defined similarly but with distance categories replacing the continuous distance variable. All regressions include year and firm fixed effects and firm controls of a multi-establishment dummy, foreign owned dummy and log age, which are not reported for brevity. Robust standard errors clustered at the firm-level are presented in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Regressions reflect years 2008, 2013 and 2015.*

Table 3: Cloud and firm performance OLS regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**	**Log Employment**	**Log Sales**	**Labor productivity**
Cloud	0.080***	0.148***	0.132***			
	(0.015)	(0.019)	(0.019)			
Cloud – incumbent				0.012	0.078***	0.112***
				(0.016)	(0.018)	(0.019)
Cloud - young				0.353***	0.427***	0.211***
				(0.045)	(0.057)	(0.054)

*Notes: Young denotes firms aged 10 years old or less in 2008. Incumbent denotes firms aged more than 10 years old in 2008. All regressions include year and firm fixed effects and firm controls of a multi-establishment dummy, foreign owned dummy and log age, which are not reported for brevity. Robust standard errors clustered at the firm-level are presented in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Regressions reflect years 2008, 2013 and 2015.*

Table 4: IV regressions: Impact of cloud on firm growth: young vs incumbents

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**	**Log Employment**	**Log Sales**	**Labor productivity**
Cloud	0.325	0.443*	0.648**			
	(0.218)	(0.261)	(0.266)			
Cloud - incumbent				0.133	0.345	0.638**
				(0.213)	(0.249)	(0.253)
Cloud -young				0.938***	1.031***	0.831**
				(0.308)	(0.369)	(0.371)
Cragg-Donald F	46.63	46.63	46.63	24.62	24.62	24.62
Kleibergen-Paap F	25.17	25.17	25.17	12.69	12.69	12.69
J-stat(p-value)	0.17	0.64	0. 32	0.53	0.33	0.31

*Notes: The table presents 2SLS estimation, where the second stage is given by equation 1 and the first stage is denoted by equation 2. We estimate two first stages for each of the endogenous variables Cloud -incumbent and Cloud-young, where we interact equation 2 with the young or incumbent dummy for each first stage. The first stage results are presented in the Appendix Table A6. Young are defined as being aged 10 years old or younger in 2008 and incumbent are defined as being older than 10 years old in 2008. All regressions include year and firm fixed effects and firm controls of a multi-establishment dummy, foreign owned dummy and log age, which are not reported for brevity. Robust standard errors clustered at the firm-level are presented in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Regressions reflect years 2008, 2013 and 2015.*

Table 5: IV regressions: Impact of cloud on firm growth: different age groups

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**	**Log Employment**	**Log Sales**	**Labor productivity**
Cloud	0.324	0.492	0.648**			
	(0.254)	(0.287)	(0.280)			
Cloud-Age (Mean Centered)	-0.698***	-0.544**	-0.149			
	(0.190)	(0.213)	(0.203)			
Cloud- Age Quartiles						
*Cloud * Age= <10 years old*				0.993***	0.988**	0.681*
				(0.309)	(0.367)	(0.366)
*Cloud * Age 11-20 years old*				0.314	0.317	0.422
				(0.230)	(0.267)	(0.269)
*Cloud * Age 21-30 years old*				0.141	0.325	0.561**
				(0.217)	(0.254)	(0.257)
*Cloud * Age >30 years old*				-0.015	0.281	0.626***
				(0.18)	(0.212)	(0.216)
Cragg-Donald F	24.95	24.95	24.95	12.57	12.57	12.57
Kleibergen-Paap F	11.56	11.56	11.56	6.60	6.60	6.60
J-stat (p-value)	0.44	0.60	0.50	0.23	0.45	0.03

Notes: Regressions mirror those in Table 4 with the exception of the treatment variable. In columns 1-3 the age interaction is mean centered (mean=16.5) and in columns 4-6 we include age quartile interactions. Note this implies two first-stage equations in 1-3 and four first-stage equations for each of the endogenous variables in 4-6, here we interact equation 2 with the age variable for each first stage. Otherwise see notes to Table 4.

Table 6: IV Regressions: Impact of cloud on firm growth: young vs incumbents. Cloud data and storage

	(1)	(2)	(3)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**
Cloud Data and Storage	0.098	0.360	0.720**
- incumbent	(0.242)	(0.282)	(0.285)
Cloud Data and Storage	1.125***	1.225***	0.932**
- young	(0.358)	(0.429)	(0.426)
Cragg-Donald F	22.22	22.22	22.22
Kleibergen-Paap F	11.61	11.61	11.61
J-stat (p-value)	0.50	0.38	0.40

Notes: Regressions mirror those in Table 4 with the exception of the treatment variable. Here the measure of cloud is the use of cloud either for data or storage or both, pooling two of the disaggregated cloud technologies. Otherwise see notes to Table 4.

Table 7: IV regressions: Impact of cloud on firm geographic reorganization: young vs incumbents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	**Multi-establishment**	**Establishment Deaths**	**Establishment Births**	**Avg distance (weighted)**	**Avg distance (unweighted)**	**Covariance**	**No. Local authorities**
Cloud -incumbent	-0.005	0.165***	-0.120***	18.929**	9.599	9.33	-0.008
	(0.09)	(0.051)	(0.046)	(8.34)	(9.77)	(7.54)	(0.09)
Cloud-young	-0.279**	0.069	-0.059	21.710*	19.687	2.023	0.065
	(0.13)	(0.068)	(0.062)	(10.80)	(12.50)	(9.19)	(0.11)
Cragg-Donald F	24.62	24.41	24.41	24.62	24.62	24.62	24.62
Kleibergen-Paap F	12.70	12.61	12.61	12.69	12.69	12.69	12.69
J-stat (p-value)	0.50	0.49	0.82	0.79	0.85	0.78	0.47

Notes: The regressions mirror those in Table 4 columns 4-6 with the exception of outcome variables. Multi-establishment status us an indicator variable equal to one if a firm becomes multi-establishment during the sample period and zero otherwise. Establishment births and deaths are calculated as over the period of the last 2 years and expressed as a share of the total number of establishments. Weighted and unweighted average distance refers to the average distance of establishments from their headquarters, where the weights are the share of establishment employment in firm employment. The covariance term measures the correlation between establishment employment and distance from the headquarters, i.e. whether farther establishments are larger (a positive covariance), or closer establishments are larger (a negative covariance) in terms of employment. Number of local authorities reflects the log of the number of different local authorities in which the firm has establishments located. The first stage results are presented in Table A13 in the Appendix. See additional notes in Table 4.

ONLINE APPENDIX

Appendix Section A: Firm level Analysis

Weighted average distance of establishments from the headquarters

Intuition: distance of the mean employee from their headquarters.

It is a firm-level measure and is calculated $\overline{wdist_f}$:

$$\overline{wdist_f} = \sum_{p \in f} s_p \cdot dist_p$$

where $dist_p$ is the distance (in km) of establishments from their headquarters, and $s_p = \frac{emp_p}{emp_f}$ is the share of establishment employment in total firm employment.

Decomposition

Following Olley and Pakes (1996) we can decompose the weighted average as:

$$\overline{wdist_f} = \overline{dist_f} + Cov(dist_p, emp_p)$$

Unweighted average distance of establishments from the headquarters

Intuition: distance of the mean establishment from their headquarters.

It is a firm-level measure and is calculated $\overline{dist_f}$:

$$\overline{dist_f} = \sum_{p \in f} \frac{1}{N_f} \cdot dist_p$$

where $dist_p$ is the distance (in km) of establishments from their headquarters, and N_f is the number of establishments of the firm.

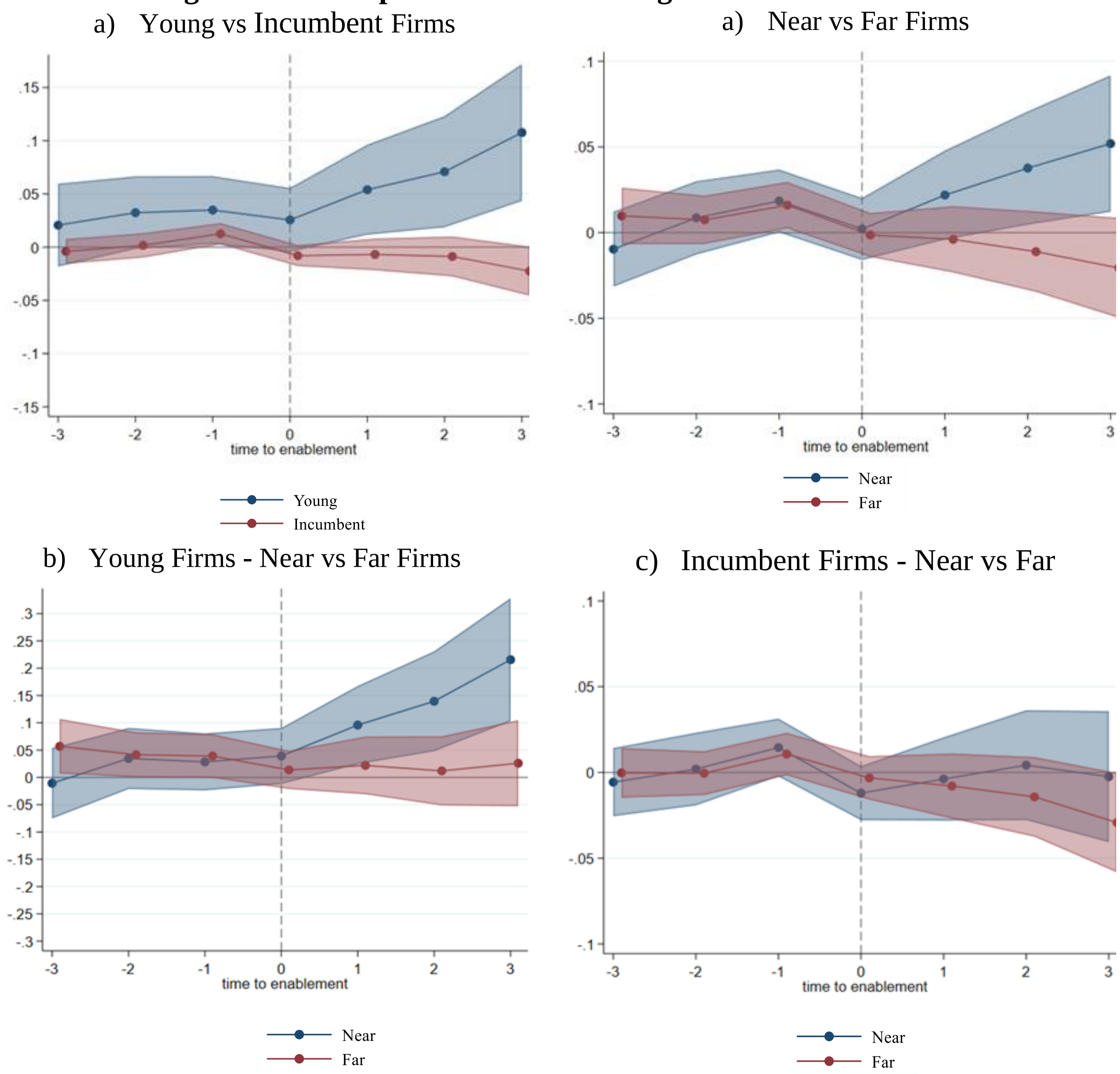
Covariance between establishment employment and establishment distance from the headquarters

Intuition: measures how employment is distributed across establishments by their proximity - are farther establishments larger (+ve covariance) or closer establishments larger (-ve covariance).

$$Cov(dist_p, emp_p) = \sum_{p \in f} (s_p - \bar{s}_f) \cdot (dist_p - \overline{dist_f})$$

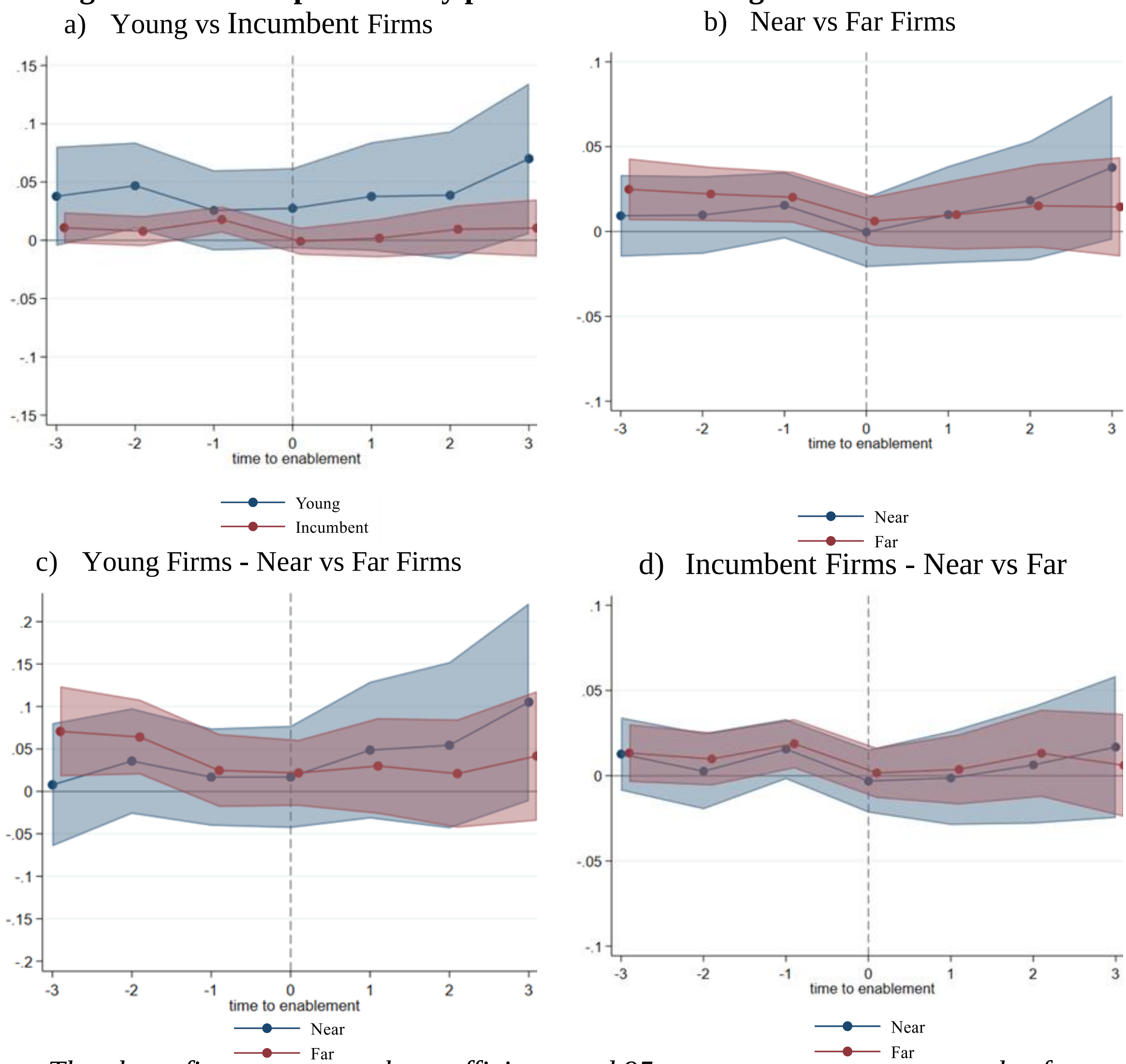
where $\bar{s}_f$ is the unweighted mean share of establishment employment. Other terms are defined as above.

Figure A1: Sales pre-trends with timing of fiber enablement



Notes: The above figures present the coefficients and 95 per cent confidence intervals of event study regressions – reflecting firm sales in the periods before and after fiber enablement (our instrumental variable). Estimation follows the method of Callaway and Sant'Anna (2021), which estimates enablement-cohort by cohort compared to never fiber enabled firms. Panel a) presents event study plots for young and incumbent firms respectively. Panel b) plots firms near and far from their exchange (less than and more than 1km respectively), and panel c) and d) present similar plots for the subsamples of young and incumbent firms, respectively. Note that the regressions that underlie the lines on each of the graphs are estimated relative to never enabled areas and should be interpreted as such. For example, in panel c) the near-young firms are compared to the near-young firms connected to never enabled exchanges. Note the differing scale across the four panels. Observations during the firm's year of fiber enablement are given by event time equal to 0, the preceding year is event time of – 1, the year after enablement is + 1, and so on. For parsimony, a subset of event-time dummies (three years before and after) are shown in the figures.

Figure A2: Labor productivity pre-trends with timing of fiber enablement



Notes: The above figures present the coefficients and 95 per cent confidence intervals of event study regressions – reflecting firm labor productivity in the periods before and after fiber enablement (our instrumental variable). Estimation follows the method of Callaway and Sant'Anna (2021), which estimates enablement-cohort by cohort compared to never fiber enabled firms. Panel a) presents event study plots for young and incumbent firms respectively. Panel b) plots firms near and far from their exchange (less than and more than 1km respectively), and panel c) and d) present similar plots for the subsamples of young and incumbent firms, respectively. Note that the regressions that underlie the lines on each of the graphs are estimated relative to never enabled areas and should be interpreted as such. For example, in panel c) the near-young firms are compared to the near-young firms connected to never enabled exchanges. Note the differing scale across the four panels. Observations during the firm's year of fiber enablement are given by event time equal to 0, the preceding year is event time of – 1, the year after enablement is + 1, and so on. For parsimony, a subset of event-time dummies (three years before and after) are shown in the figures.

Table A1: Summary Statistics of Cloud Measures by Year

	2013		2015	
Variable	**mean**	**st.dev.**	**mean**	**st.dev.**
All firms				
Cloud	0.385	0.487	0.441	0.497
Cloud Databases	0.158	0.365	0.181	0.385
Cloud Storage of files	0.229	0.420	0.293	0.455
Cloud Email	0.181	0.385	0.250	0.433
Cloud Office Software	0.096	0.294	0.190	0.392
Cloud Finance Software	0.082	0.274	0.101	0.301
Cloud CRM	0.114	0.318	0.127	0.333
Cloud Own Software	0.094	0.291	0.110	0.313
Cloud Low Tech	0.118	0.322	0.142	0.350
Cloud Medium Tech	0.158	0.365	0.181	0.385
Cloud High Tech	0.190	0.392	0.223	0.416
Young firms				
Cloud	0.317	0.466	0.349	0.477
Cloud Databases	0.156	0.363	0.152	0.359
Cloud Storage of files	0.208	0.406	0.235	0.424
Cloud Email	0.169	0.375	0.213	0.409
Cloud Office Software	0.102	0.303	0.158	0.365
Cloud Finance Software	0.090	0.286	0.112	0.316
Cloud CRM	0.099	0.299	0.103	0.304
Cloud Own Software	0.100	0.300	0.085	0.279
Cloud Low Tech	0.085	0.280	0.107	0.310
Cloud Medium Tech	0.156	0.363	0.152	0.359
Cloud High Tech	0.172	0.377	0.191	0.393
Incumbent firms				
Cloud	0.405	0.491	0.472	0.499
Cloud Databases	0.158	0.365	0.191	0.393
Cloud Storage of files	0.235	0.424	0.312	0.463
Cloud Email	0.185	0.388	0.262	0.440
Cloud Office Software	0.094	0.292	0.201	0.401
Cloud Finance Software	0.080	0.271	0.097	0.296
Cloud CRM	0.119	0.324	0.135	0.342
Cloud Own Software	0.092	0.289	0.118	0.323
Cloud Low Tech	0.127	0.333	0.154	0.361
Cloud Medium Tech	0.158	0.365	0.191	0.393
Cloud High Tech	0.195	0.397	0.234	0.423

Notes. These statistics are from a balanced panel of observations for comparison of adoption across time for the same set of firms – a subset of our estimation sample of firms over years 2008, 2013 and 2015. Young are defined as being aged 10 years old or younger in 2008 and incumbent are defined as being older than 10 years old in 2008. There is zero adoption of cloud in 2008, hence lower values are reported for all years (2008, 2013 and 2015). Cloud low, medium and high tech are defined following Eurostat (2018). For full details on definitions of each cloud type see the Data References in the Appendix.

Table A2: Summary Statistics of Other Variables

	All firms		Young firms <=10 years old (in 2008)		Incumbent firms >10 years old (in 2008)	
Variable	**mean**	**Sd**	**mean**	**Sd**	**mean**	**Sd**
(Log) Employment	4.335	2.266	2.856	1.846	4.810	2.182
(Log) Sales	8.938	2.628	7.184	2.179	9.514	2.505
(Log) Labor Productivity	4.567	1.193	4.155	1.369	4.702	1.096
Multi-establishment dummy	0.479	0.500	0.314	0.464	0.534	0.499
Establishment deaths/total establishments	0.056	0.141	0.053	0.150	0.058	0.137
Establishment birth/ total establishments	0.062	0.142	0.078	0.182	0.057	0.126
Weighted average distance establishment to headquarter (km)	34.687	67.527	10.936	39.802	42.497	72.746
Unweighted average distance establishment to headquarter (km)	47.199	79.426	15.696	49.450	57.557	84.530
Covariance establishment distance-establishment employment	-12.512	36.309	-4.760	23.421	-15.061	39.310
Fiber enabled	0.358	0.479	0.337	0.473	0.365	0.481
Exchange distance (km)	1.332	0.893	1.341	0.913	1.329	0.887
Number of local authorities	11.396	40.687	2.495	10.745	14.328	46.131
Foreign owned	0.175	0.380	0.080	0.272	0.206	0.404
Log age	3.023	0.719	2.043	0.688	3.345	0.333
Services	0.599	0.490	0.671	0.470	0.575	0.494
Urban region	0.764	0.425	0.739	0.439	0.772	0.420
Knowledge intensive industries	0.163	0.369	0.138	0.345	0.171	0.376
Δ log Employment	0.038	0.304	0.064	0.408	0.030	0.263
Δ log Sales	0.056	0.468	0.124	0.653	0.035	0.390
Δ log Labor Productivity	0.019	0.521	0.062	0.724	0.006	0.438

Notes: Labor productivity reflects log sales per worker, changes in (log) employment, sales and labor productivity reflect growth relative to the previous year. Establishment birth / death ratios are calculated as the number of establishment births / deaths (within a firm) over the period of the last 2 years and expressed as a share of the total number of establishments. Average distances between establishments and the headquarters and the covariance term are defined in the first section of the Online Appendix. Number of local authorities reflects number of different local authorities a firm has establishments in. Knowledge intensity is constructed using R&D expenditures (weighted by employment) at the 5-digit UK SIC level. Knowledge intensive industries are those in the top quartile of the distribution.

Table A3: OLS Regressions: Impact of cloud on IT Investment: young vs incumbent

	(1)	(2)
Dependent variable:	**Log IT investment per employee**	**Log IT investment per employee**
Cloud	-0.550***	
	(0.051)	
Cloud - incumbent		-0.660***
		(0.059)
Cloud- young		-0.117
		(0.082)
Observations	17,544	17,544

*Notes: All regressions include year and firm fixed effects and firm controls of a multi-establishment dummy, foreign owned dummy and log age, which are not reported for brevity. Young are defined as being aged 10 years old or younger in 2008 and incumbent are defined as being older than 10 years old in 2008. Robust standard errors clustered at the firm-level are presented in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. All regressions reflect the time periods 2008, 2013 and 2015.*

Table A4: First-Stage Regressions: Relationship between instruments and adoption by cloud type

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	**Cloud databases**	**Cloud storage of files**	**Cloud email**	**Cloud office software**	**Cloud finance software**	**Cloud CRM**	**Cloud own software**	**Cloud Low-Tech**	**Cloud Med-Tech**	**Cloud High-Tech**
Fiber	0.077***	0.071***	0.067***	0.040***	0.036***	0.043***	0.033***	0.016	0.077***	0.069***
	(0.011)	(0.013)	(0.012)	(0.011)	(0.009)	(0.010)	(0.009)	(0.010)	(0.011)	(0.012)
Fiber*distance	-0.020***	-0.015**	-0.012*	-0.006	-0.011**	-0.009	-0.006	-0.001	-0.020***	-0.017**
	(0.006)	(0.007)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
Observations	17,870	17,870	17,870	17,870	17,870	17,870	17,870	17,870	17,870	17,870
Kleibergen-Paap F	24.40	15.14	16.92	8.38	8.50	10.60	7.08	1.62	24.40	16.72

Notes: The table repeats our baseline first stage results in Table 4 column 1, but with differing measures of cloud. Cloud low, medium and high tech are defined following Eurostat (2018). According to this definition, basic cloud technologies include email, office software, or file storage via cloud. Medium tech cloud use means employing at least one of the basic cloud services along with cloud for hosting the enterprise's database(s). High tech cloud use means employing of at least one of the basic cloud services as well as at least one of the more advanced cloud services including, hosting the enterprise's database(s), Finance Software, CRM and processing services. Other notes follow those described in Table 4.

Table A5: First-Stage Regressions: Relationship between instruments and other digital technologies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Dependent variable:	**% Employees w/ PCs**	**% Employees w/ Cellphones**	**Online sales**	**% Online sales in total sales**	**RFID identification**	**RFID production**	**ERPs**	**CRMs Marketing**	**CRMs customer**	**IT specialists**	**IT training**	**E-data sharing suppliers**	**E-data sharing customers**
Fiber	-1.082	-0.006	-0.001	0.412	-0.044*	-0.017	0.006	-0.006	0.000	0.002	0.007	-0.001	-0.030
	(0.74)	(0.01)	(0.01)	(0.33)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Fiber*distance	0.779*	0.000	0.002	-0.138	0.022	0.018	0.002	0.018	0.003	0.001	-0.002	-0.002	0.019
	(0.44)	(0.00)	(0.01)	(0.21)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	30,727	26,452	30,552	30,552	4,239	4,239	22,025	16,040	16,040	20,107	20,107	11,869	11,869
Cragg-Donald F	5.2	1.81	0.23	2.02	4.92	2.14	0.79	3.96	0.18	0.16	0.15	0.06	2.36
Kleibergen-Paap F	1.54	0.93	0.07	0.90	1.97	0.84	0.42	1.89	0.09	0.08	0.07	0.03	1.20
J-stat (p-value)	0.46	0.59	0.97	0.21	0.72	0.66	0.87	0.93	0.94	0.98	0.51	0.88	0.72

Notes: The table repeats our baseline first stage results in Table 4 column 1 but replaces cloud with alternative measures of digital technologies. Online sales is a binary variable reflecting positive e-commerce sales. RFID identification is a dummy variable reflecting use of RFID for product identification and RFID production reflects RFID for monitoring and control of industrial production. RFID information is only available for a subset of our sample, manufacturing firms. ERP is Enterprise Resource Planning software, CRM is Customer Relationship Management Software. IT specialists and IT training represents a dummy if specialists are employed or training is undertaken respectively. The last two columns are dummy variables reflecting the sharing of electronic data with suppliers and customers respectively. The technologies used in Columns 1,2, 3, 4, 10, 11, 12 and 13 are available for the years 2008 to 2015 inclusively. The technologies used in Columns 5 and 6 are available for the years 2008, 2010, and 2013. The technologies used in Column 7, 8, 9 are available for the years 2008, 2009, 2011, 2013 and 2014. Other notes follow those described in Table 4.

Table A6: First-Stage Regressions: Corresponding to Table 4 in main body

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**	**Log Employment**	**Log Sales**	**Labor productivity**
First stage Cloud						
Fiber	0.103***	0.103***	0.103***			
	(0.015)	(0.015)	(0.015)			
*Fiber *distance*	-0.026***	-0.026***	-0.026***			
	(0.008)	(0.008)	(0.008)			
First stage Cloud-Incumbent						
Fiber -incumbent				0.172***	0.172***	0.172***
				(0.016)	(0.016)	(0.016)
Fiber -young				-0.235***	-0.235***	-0.235***
				(0.010)	(0.010)	(0.010)
*Fiber *distance-incumbent*				-0.021**	-0.021**	-0.021**
				(0.009)	(0.009)	(0.009)
*Fiber *distance-young*				-0.002	-0.002	-0.002
				(0.002)	(0.002)	(0.002)
First stage Cloud-Young						
Fiber -incumbent				-0.057***	-0.057***	-0.057***
				(0.004)	(0.004)	(0.004)
Fiber -young				0.298***	0.298***	0.298***
				(0.026)	(0.026)	(0.026)
*Fiber *distance-incumbent*				-0.000	-0.000	-0.000
				0.000	0.000	0.000
*Fiber *distance-young*				-0.042**	-0.042**	-0.042**
				(0.016)	(0.016)	(0.016)
Cragg-Donald F	46.63	46.63	46.63	24.62	24.62	24.62
Kleibergen-Paap F	25.17	25.17	25.17	12.69	12.69	12.69
J-stat(p-value)	0.17	0.64	0. 32	0.53	0.33	0.31

Notes: The table presents 2SLS estimation, where the second stage is given by equation 1 and the first stage is denoted by equation 2. The second stage estimates are presented in Table 4 in the paper. We estimate two first stages for each of the endogenous variables Cloud -incumbent and Cloud-young, where we interact equation 2 with the young or incumbent dummy for each first stage. Other notes follow those described in Table 4.

Table A7: IV Regressions: Impact of cloud on firm growth: by mean centered size

	(1)	(2)	(3)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**
	1.212***	0.677	0.318
Cloud	(0.425)	(0.477)	(0.474)
	-0.249***	-0.065	0.085
Cloud-Size (mean centered)	(0.049)	(0.054)	(0.054)
Observations	17,516	17,402	17,402
Cragg-Donald F	14.10	13.46	13.46
Kleibergen-Paap F	6.68	6.30	6.30
J-stat(p-value)	0.10	0.70	0.39

Notes: The table presents 2SLS estimation, where the second stage is given by equation 1 and the first stage is denoted by equation 2. We estimate two first stages for each of the endogenous variables Cloud and Cloud-Size, where Size reflects an interaction with (mean-centered) firm employment in 2008. First stage regressions are omitted for parsimony. Other notes follow those described in Table 4.

Table A8: IV Regressions: Impact of cloud on firm growth: Size and Age groups

	(1)	(2)	(3)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**
Cloud * Small and Young	1.611***	0.865*	0.218
	(0.380)	(0.456)	(0.456)
Cloud * Small and Incumbent	1.001***	0.363	0.153
	(0.371)	(0.451)	(0.455)
Cloud * Large and Young	0.420	1.232***	1.286***
	(0.258)	(0.375)	(0.385)
Cloud * Large and Incumbent	0.143	0.278	0.525**
	(0.179)	(0.222)	(0.226)
Observations	17,516	17,402	17,402
Cragg-Donald F	9.86	9.79	9.79
Kleibergen-Paap F	5.85	5.83	5.83
J-stat(p-value)	0.71	0.53	0.59

Notes: The table presents 2SLS estimation, where the second stage is given by equation 1 and the first stage is denoted by equation 2. We estimate four first stages for each of the endogenous variables, cloud interacted with size and age dummies. Young are defined as being aged 10 years old or younger in 2008 and incumbent are defined as being older than 10 years old in 2008. Small defined as firms with 50 employees or less in 2008 and large firms as more than 50 employees in 2008. First stage regressions are omitted for parsimony. Other notes follow those described in Table 4.

Table A9: IV Regressions: Impact of cloud on firm growth: young vs incumbents: Cloud High Tech

	(1)	(2)	(3)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor Productivity**
Cloud - incumbent	0.242	0.502	0.945**
	(0.342)	(0.400)	(0.406)
Cloud- young	1.758***	1.766***	1.197*
	(0.530)	(0.637)	(0.626)
Observations	17,544	17,430	17,430
Cragg-Donald F	15.02	14.62	14.62
Kleibergen-Paap F	7.95	7.70	7.70
J-stat(p-value)	0.80	0.32	0.30

Notes: The table repeats our baseline results in Table 4 columns 4 to 6, but using high-tech cloud. Cloud low, medium and high tech are defined following Eurostat (2018). According to this definition, basic cloud technologies include email, office software, or file storage via cloud. High tech cloud use means employing of at least one of the basic cloud services as well as at least one of the more advanced cloud services including, hosting the enterprise's database(s), Finance Software, CRM and processing services. First stage regressions are omitted for parsimony. Other notes follow those described in Table 4.

Table A10: First-Stage Regressions: Relationship between instruments and: Young vs Incumbent Firms

Regression	**(1)**
Dependent variable: Cloud adoption	**All Firms**
Fiber Enablement	0.115***
	(0.017)
Fiber * Distance	-0.029***
	(0.009)
Fiber Enablement * Young	0.055*
	(0.029)
Fiber * Distance * Young	-0.022
	(0.019)
Observations	17,544
Cragg-Donald F	33.73
Kleibergen-Paap F	18.22
J-stat(p-value)	0.80

Notes: The regression mirrors Table 2 column 1 (or equivalently the first-stage of Table 4 column 1), however we interact our instruments with a young dummy to contrast the power of our instruments across young and incumbent firms. Please see Table 2 or 4 for additional notes.

Table A11: IV Regressions: Impact of cloud on firm growth: Alternative Age Groups

	(1)	(2)	(3)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**
Cloud Age =<5 years old	1.133***	1.276***	0.949**
	(0.386)	(0.460)	(0.461)
Cloud Age 6-10 years old	0.973***	0.813**	0.529
	(0.302)	(0.357)	(0.353)
Cloud Age 11-20 years old	0.371	0.309	0.416
	(0.238)	(0.278)	(0.278)
Cloud Age >20 years old	0.085	0.262	0.574**
	(0.207)	(0.240)	(0.241)
Observations	17,544	17,430	17,430
Cragg-Donald F	12.80	12.81	12.81
Kleibergen-Paap F	6.54	6.53	6.53
J-stat(p-value)	0.37	0.67	0.30

Notes: The regressions mirror Table 5 columns 4-6, but with differing age categories. First stage regressions are omitted for parsimony. Other notes follow those described in Table 5.

Table A12: IV Regressions: Impact of cloud on young firm growth: Knowledge Intensive Sectors

	(1)	(2)	(3)
Dependent variable:	**Log Employment**	**Log Sales**	**Labor productivity**
Cloud - Incumbent	0.031	0.280	0.625**
	(0.218)	(0.252)	(0.256)
Cloud – Young	0.636**	0.859**	0.791**
	(0.319)	(0.382)	(0.383)
Cloud-Young-Knowledge	0.661***	0.368	0.063
	(0.226)	(0.233)	(0.224)
Observations	17,544	17,430	17,430
Cragg-Donald F	16.53	16.54	16.54
Kleibergen-Paap F	8.51	8.48	8.48
J-stat(p-value)	0.19	0.57	0.35

Notes: The table presents 2SLS estimation, where the second stage is given by equation 1 and the first stage is denoted by equation 2. We estimate three first stages, one for each of the endogenous variables Cloud -incumbent and Cloud-young and Cloud-Young-Knowledge, where we interact equation 2 with the young or incumbent or young- Knowledge dummy for each first stage. Knowledge intensity is constructed using R&D expenditures (weighted by employment) at the 5-digit UK SIC level. Knowledge intensive industries are those in the top quartile of the distribution. First stage regressions are omitted for parsimony. Other notes follow those described in Table 4.

Table A13: First-Stage Regressions: Corresponding to Table 7 in main body

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	**Multi-establishment**	**Establishment Deaths**	**Establishment Births**	**Avg distance (weighted)**	**Avg distance (unweighted)**	**Covariance**	**No. Local authorities**
First stage Cloud-Incumbent							
Fiber -incumbent	0.173***	0.173***	0.173***	0.172***	0.172***	0.172***	0.172***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Fiber -young	-0.239***	-0.237***	-0.237***	-0.235***	-0.235***	-0.235***	-0.235***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
*Fiber *distance-incumbent*	-0.021**	-0.021**	-0.021**	-0.021**	-0.021**	-0.021**	-0.021**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
*Fiber *distance-young*	-0.001	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
First stage Cloud-Young							
Fiber -incumbent	-0.058***	-0.058***	-0.058***	-0.057***	-0.057***	-0.057***	-0.057***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Fiber -young	0.300***	0.297***	0.297***	0.298***	0.298***	0.298***	0.298***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
*Fiber *distance-incumbent*	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	0.000	0.000	0.000	0.000	0.000	0.000	0.000
*Fiber *distance-young*	-0.043***	-0.041**	-0.041**	-0.042**	-0.042**	-0.042**	-0.042**
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Observations	17,386	17,386	17,386	17,386	17,386	17,386	17,386
Cragg-Donald F	24.62	24.41	24.41	24.62	24.62	24.62	24.62
Kleibergen-Paap F	12.70	12.61	12.61	12.69	12.69	12.69	12.69
J-stat (p-value)	0.50	0.49	0.82	0.79	0.85	0.78	0.47

Notes: The table presents 2SLS estimation, where the second stage is given by equation 1 and the first stage is denoted by equation 2. The second stage estimates are presented in Table 7 in the paper. We estimate two first stages for each of the endogenous variables Cloud -incumbent and Cloud-young, where we interact equation 2 with the young or incumbent dummy for each first stage. See additional notes to Table 7.

Appendix Section B: Employer-Employee Analysis

In order to further decompose firm geographic dispersion, we use employer-employee data from Annual Survey of Hours and Earnings (ASHE).[39] ASHE is a 1% panel of all the workers in the UK derived from HM Revenue and Customs' Pay As You Earn records. Workers can be matched to establishments and firms in the UK and tracked over their employment lifetime, allowing movements within or across firms over time to be measured.[40] The summary statistics of the worker data, and their firms is given in Tables B1 below.

In the geographic organization section, we found that cloud leads to incumbent firms dispersing employment further from the headquarters. Part of the reallocation story may be driven by the types of establishments that the firms close due to cloud. Alternatively, it may be the fact that the technology enables higher mobility of its workforce and thus it may be driven by existing employees moving throughout the firm across establishment over the sample period. Taking the analysis down to the level of the employee, we examine where employees are being moved to, and in particular whether cloud computing is a key determinant of the mobility of workers across establishments within the firm. Such movements may occur because of changes to the spatial organization of the production (Leamer and Storper, 2001; Duranton and Puga, 2005), because of the ability of management to share information and deal with problems (Bloom et al., 2014) or because of face-to-face interactions (Gaspar and Glaeser, 1998). This may lead to employees being shifted away from the headquarters to other

[39] This data is used by various papers including Bell and Van Reenen (2013) and Aghion et al (2017).

[40] Due to data limitations (we have only a 1% sample of employee jobs) we assess employee movement within/across firms but are not able to examine changes in the distribution of wages or skill compositions.

establishments. Alternatively, cloud may simply induce greater movement across any all of the establishments the firm operates.

This analysis is at the employee-establishment-year level we can also assess the extent to which movement of workers is influenced by whether the HQ has cloud and/or whether the establishment has cloud. [41] The inclusion of establishment and worker fixed effects means we consider movement between existing establishments, neglecting opening and closure. [42]

The results from columns 1 and 2 in Table B2 are consistent with those for the covariance term of employment and distance in Table 7 and confirm a reshuffling of employment within the firm. In particular, the result suggests that workers in establishments using cloud technologies are significantly more likely to move compared to establishments that have not yet adopted the technology.[43] This holds when we include establishment fixed effects, but also worker fixed effects to control for unobservable time invariant characteristics of the individual. We find no evidence that this probability is affected by headquarters cloud use however.

In columns three and four we extend this to explore whether this reshuffling of employment is primarily associated with activity moving to or away from the headquarters. Irrespective of whether the HQ or its establishments adopt cloud we find no systematic movement of employees towards or away from the headquarters. In columns 5 and 6 we consider this in a different way and use a measure of the distance of the worker from the HQ. Again, we find no

[41] The first stages for each of the endogenous variables are not reported for brevity.

[42] Since the data on cloud adoption is at the level of the firm, we construct establishment cloud use based on the typical diffusion of cloud throughout firms (e.g. most firm subscriptions of cloud provide licensing to all establishments of the firm) and the technological prerequisites for adoption (access to high speed internet is essential). As such, establishment cloud use is set to one if the firm has adoption cloud and if the establishment has access to fiber broadband, and zero otherwise.

[43] We present evidence for worker movement between t and t+1, similar results are obtained for 2 periods ahead.

systematic evidence that these distances are affected by cloud adoption. These results therefore suggest that cloud adoption, in particular by establishments, is a determinant for employment mobility within the firm, but this reorganization of activity is across establishments rather than to and from the headquarters.[44]

Table B1: Summary Statistics of Employer-Employee Data

Variable	**Obs.**	**mean**	**std. dev**
1 period Probability of switching establishments (within the firm)	34,108	0.04	0.19
1 period Probability of switching to / from HQ	33,484	0.02	0.14
1 period Change in (Log km) Workplace Distance from HQ (of switchers within firm)	30,467	0.08	1.30
HQ cloud	34,108	0.30	0.46
Establishment *cloud*	34,108	0.29	0.45
HQ Fiber	34,108	0.44	0.50
Establishment *Fiber*HQ Fiber*	34,108	0.42	0.49
Multi establishment	34,108	0.81	0.40
Foreign	34,108	0.38	0.48
Firm Age	34,108	28.73	9.63
Exchange Distance (km)	34,108	1.27	0.82
Worker Age	34,108	40.01	11.77
Tenure	34,108	7.51	8.30
Skilled Worker (Soc 2010 classification)	34,108	0.45	0.50
Female	34,108	0.35	0.48

[44] In unreported results we also find no evidence that firms relocate workers towards regions that were less costly, measured either in terms of the rental cost of commercial floor space or the wage rate of workers.

Table B2: IV Regressions: Worker Movement Regressions using matched employer-employee data

1 Period Ahead	Probability of switching establishments (within the firm)		Probability of switching to / from HQ		Change in Workplace Distance from HQ (of switchers within firm)	
Second Stage IV estimates:	(1)	(2)	(3)	(4)	(5)	(6)
Establishment Cloud	0.100**	0.100**	0.060	0.060	-0.448	-0.443
	(0.047)	(0.046)	(0.039)	(0.039)	(0.309)	(0.308)
HQ Cloud	-0.036	-0.038	-0.045	-0.045	0.122	0.122
	(0.041)	(0.041)	(0.031)	(0.031)	(0.289)	(0.289)
Establishment Fixed Effects	Y	Y	Y	Y	Y	Y
Worker Fixed Effects		Y		Y		Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	34,108	34,066	33,370	33,331	30,339	30,304
Cragg-Donald F	161.72	160.75	159.71	158.26	152.19	152.56
Kleibergen-Paap F	14.35	14.07	14.20	13.89	13.62	13.51
J-stat (p-value)	0.92	0.92	0.69	0.68	0.51	0.51

*Notes: All regressions include controls for multi-establishment, ownership and firm age, as well as worker controls for age, tenure, tenure squared, skilled occupation dummy, sex and their interactions with sex. Robust standard errors clustered at the establishment-level are presented in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Instruments are fiber enablement and an interaction with log employment in 2008, and similarly at establishment level. Establishment (HQ) cloud reflects firm cloud adoption * establishment (HQ) fiber availability. Regressions reflect years 2008, 2013 and 2015. Probability of switching is measured one period ahead.*

References for the Appendix

Aghion, P., Bergeaud, A., Blundell, R., and Griffith, R.. (2017). ‘Innovation, Firms and Wage Inequality’. mimeo.

Bell, B. D., and Van Reenen, J. (2013). ‘Extreme Wage Inequality: Pay at the Very Top.’ *American Economic Review*, 103 (3): 153-57.

Gaspar, J. and E.L. Glaeser, (1998), “Information technologies and the future of cities”, *Journal of Urban Economics*, 43(1): 136–156.

Glaeser, E. L., and Resseger, M. G. (2010). ‘The complementarity between cities and skills’. *Journal of Regional Science*, 50(1), 221-244.

Data References for the Appendix

Types of cloud services in E-commerce survey

Does this business buy any of the following cloud computing services used over the internet?

- *Email*: Email, as a cloud computing service
- *Software*: Office software for example word-processing or spreadsheets, as a cloud computing service
- *Databases*: Hosting the business’ database(s), as a cloud computing service
- *Storage of files*: Storage of files, as a cloud computing service
- *Finance Software*: Finance or accounting software applications, as a cloud computing service
- *CRM*: Customer relations management software, as a cloud computing service
- *Processing Own Software*: Computing capacity to the business’ own software, as a cloud computing service

Office for National Statistics (2019). Business Structure Database, 1997-2018: Secure Access [data collection] 10th Edition. UK Data Service. SN:6697, http://doi.org/10.5255/UKDA-SN-6697-10

Office for National Statistics, Virtual Microdata Laboratory (VML), University of West of England, Bristol (2017) Annual Respondents Database X, 1998-2014. [data collection] 4th Edition Office for National Statistics, [original data producer(s)]. Office for National Statistics. SN:6989, http://doi.org/10.5255/UKDA-SN-7989-4

Office for National Statistics (2017). E-commerce Survey, 2001-2015: Secure Access [data collection] 7th Edition. UK Data Service. SN:6700, http://doi.org/10.5255/UKDA-SN-6700-7

Office for National Statistics. (2019). Annual Survey of Hours and Earnings, 1997-2018: Secure Access. [data collection] 14th Edition. UK Data Service. SN 6689, http://doi.org/10.5255/UKDA-SN-6689-13