Digital economy in the UK

Tranos, Emmanouil; Kitsos, Tasos; Ortega Argiles, Raquel

DOI:
10.1080/00343404.2020.1826420

Digital economy in the UK: Regional productivity effects of early adoption

To appear in *Regional Studies*

Emmanouil Tranos[1], Tasos Kitsos[2], Ortega-Argilés, Raquel[3]

[1]University of Bristol, School of Geographical Sciences, The Alan Turing Institute, e.tranos@bristol.ac.uk, Corresponding author.

[2]University of Birmingham, Birmingham Business School, City-Region Economic Development Institute, A.Kitsos@bham.ac.uk.

[3]University of Birmingham, Birmingham Business School, City-Region Economic Development Institute, R.OrtegaArgiles@bham.ac.uk.

Abstract

Whilst the importance of internet-related technologies and digitisation practices to economic performance is well documented, little is known about the long-term effects of the early adoption of such technologies. We use novel, geolocated data about the volume of online content from the Internet Archive to approximate the active engagement with digital economic activities. Using panel data methods, we find significant positive and long-lasting effects of online content creation in 2000 on subsequent regional productivity levels up to 16 years later. Our findings highlight the sizeable effects of the digital economy that policymakers should consider in developing future rollout strategies.

Introduction

The digital economy has grown in importance during the last decades. Its positive contribution to economic development is highlighted in key policy documents illustrating the role of digital sectors and occupations as key drivers for growth. For instance, the UK, one of the most advanced countries in term of digital capabilities (CHAKRAVORTI and CHATURVEDI, 2017; MILOŠEVIĆ et al., 2018), recently announced its high-level objective to support every business to adopt digital practices (HM GOVERNMENT, 2017). Research suggests that by doing so, the country can add up to 2.5 percentage points to its Gross Domestic Product (GDP) (OXFORD ECONOMICS, 2015). Similarly, the importance of engaging with the digital economy is emphasised at supra-national levels, as, for instance, this was one of EU's ten priorities for the 2015-2019 period[1]. Measuring the growth of the digital economy revealed significant national and sub-national differences in both the availability and the penetration of internet technologies (BLANK et al., 2018; RIDDLESDEN and SINGLETON, 2014). These variations have led to research about the implications of the engagement with the digital economy on interregional inequalities (JONES and HENDERSON, 2019; REUSCHKE and MASON, 2020). This paper contributes to these debates by analysing the long-term productivity effect of the early adoption of internet-related technologies and digitisation practices.

Research so far focused on whether and how the digitisation of market and non-market activities support the economy (KOLKO, 2012; KOUTROUMPIS, 2009). We understand digitisation as the transformation that socio-economic systems undergo because of the wide adoption of digital technologies (KATZ et al., 2014). Likewise, the digital economy refers to the pervasive use of digital technologies in all aspects of the economy (ATKINSON and MCKAY, 2007).

Despite the heterogeneity of focus, scale and methods, there is broad consensus among studies that digitisation is correlated with, or even leads to, positive macro-economic outcomes (e.g. economic growth or productivity) at different scales (from national, to

[1] https://ec.europa.eu/commission/priorities/digital-single-market_en

regional and the individual firm). These studies predominantly (i) investigate the *contemporaneous* relationship between digitisation and economic outputs and (ii) approximate the level of digitisation with data about the quality or the capacity of *digital infrastructure*.

However, the nuances of these relationships are not yet entirely clear. Individual firm characteristics such as agency and sectoral belonging (HENDERSON and ROCHE, 2020; JONES and HENDERSON, 2019) have been found to influence the relationship between digitisation and growth, whilst characteristics (i) and (ii) above, do not fully capture the extent of engaging in digital economic activities and, hence, underestimate their importance for local economies. We, therefore, propose a research framework to address these issues with the use of previously unexplored big data from the Internet Archive[2] (IA).

Contemporaneous approaches examine how digital conditions in time t or t-1 can affect the economy at time t. However, such approaches understate the very nature of the internet as an innovation, which created radical changes to industrial dynamics and affected upstream and downstream sectors. Following evolutionary perspectives (NEFFKE et al., 2011; SIMMIE et al., 2014) we argue that the study of the initial conditions, which led to the evolution of path-dependent technological development trajectories is of equal, if not greater, importance.

Thus, if we want to understand how digitisation and digital economic activities lead to broader macro-economic outcomes, we need to consider the initial stages of digitisation – that is the *early adoption* of digital technologies – and how they vary over space and time. Consequently, we need to explore the effect of digital conditions in time t-n, where n>>1 in order to capture

[2] https://archive.org/

these initial conditions. This is particularly relevant in studies involving information and telecommunication technologies (ICT) since a time-lag between investment and outcomes is very common, persistent in time and is embodied in the Solow paradox (ACEMOGLU et al., 2014; BRYNJOLFSSON et al., 2017; MCCANN, 2018; WOLF, 2018). This approach is also in line with Roger's 'Diffusion of Innovation' framework (1995) and the critical role that early adopters play in how innovations diffuse over time.

Moreover, most digital economy studies approximate the level of digitisation in a country or region using data on internet/broadband infrastructure. WHEELER and O'KELLY (1999) analysed the topology of the internet's hardware and derived city connectivities and, HALLER and LYONS (2015) used broadband speed to examine its effect on business performance. TRANOS (2013) examined the geography and spatial economic effects of the internet's main infrastructural networks in Europe and, RIDDLESDEN and SINGLETON (2014) highlighted the broadband divides in the UK. Research also explored the spatial distribution of internet users in the UK (BLANK et al., 2018; SINGLETON et al., 2015) and provided geodemographic classifications of how individuals engage with the internet (LONGLEY and SINGLETON, 2009; RIDDLESDEN, 2014). Similar digital proxies have been used in examinations of the effect of digitisation on economic performance, which are discussed in the next section. These data sources capture either the supply (provision of internet infrastructure) or the demand for individual connectivity over space. However, they do not offer insights into the actual usage of the infrastructure, or, in other words, the active engagement with the digital economy and its relationship to economic activities that need to be considered to fully understand the economic effects of digitisation.

To move beyond the above approximations of digitisation and to overcome the lagged nature of its potential economic effects, we employ a novel source of historical web data. These data have never been used before in such a context and allow us to capture the geography of the early digital economic activities and test to what extent they contribute to long-term positive productivity effects. The data include all the archived webpages under the .uk country code Top Level Domain (ccTLD) by the IA (TRANOS and STICH, 2020) and is the world's most comprehensive archive of webpages (AINSWORTH et al., 2011b; HOLZMANN et al., 2016a). The .uk was one of the first ccTLD created back in 1985 (HOPE, 2017) and the second most popular one in 1999 (ZOOK, 2001).

We geolocate these webpages by using mentions to UK postcodes in the web text and build annual measures of the volume of web content for UK NUTS3 regions. We consider the creation and maintenance of webpages as a proxy for the active engagement with the digital economy. Hence, we are able to identify the long-run effects of the early engagement with the digital economy on regional productivity. To do this, we use panel data techniques, including fixed-effects regressions that account for unobserved, time-invariant heterogeneity between regions, but also a Hausman-Taylor estimator that allows us to include important time-invariant productivity determinants. By going back to 2000, a time when the evolution of the digital economy was still at its early stages, these data provide a unique opportunity to test the long-term effect of the early adoption of digital technologies.

The results suggest that the volume of online content in 2000 is positively associated with regional productivity in subsequent years until the last year of our panel (2016). These effects are also increasing in time, providing evidence of a long-run impact between early engagement and productivity in later years. The results hold against several robustness tests.

Equally important is the lagged structure of the above effects. With the use of interaction terms, we identify statistically significant differences between the effects of our primary independent variable in 2000 and subsequent years. Our findings quantify the productivity effects of the early adoption of digital technologies and contribute to the debates on the importance of digitisation in explaining interregional inequalities. If digitisation is to be considered as a tool for levelling-up regional disparities in the UK, policymakers should be aware of its significant and positive long-lasting effects, especially concerning rollout strategies.

The paper is structured as follows. The next section reviews empirical studies, which explore the economic effects of the engagement with the digital economy. We then discuss the novel data we use and the processes of cleaning and preparing them. Next, we present and discuss our empirical approach, results and robustness tests. The final section concludes.

The economic effects of digitisation

The importance of technological innovations and ICT on productivity growth has long been recognised

(GRIFFITH et al., 2004; GRILICHES, 1979; JORGENSON, 2009) with several potential ways in which technology contributes to productivity growth. Whether studying the growth of knowledge-intensive industries (HOUSEMAN et al., 2014) or the efficiency gains from the adoption of digital tools (JORGENSON, 2009), a growing number of studies point to a positive relationship between ICT and productivity growth.

More specifically, firm-level studies identified positive causal effects of digitisation on firm performance and productivity (BERTSCHEK et al., 2013; BLOOM et al., 2012). Whilst these studies predominantly verify earlier research (BRYNJOLFSSON and HITT, 1996; LEHR and LICHTENBERG, 1999), they consider a time period characterised by low maturity levels of digital technologies. GAL et al. (2019) proposed that the adoption of digital technologies is linked to firm-level productivity gains in the service sector whilst, BAILIN RIVARES et al. (2019) found that expansion of online platforms is associated with higher productivity in service industries. On the contrary, some studies indicated no effect from access to (or speed of) broadband on firm productivity in several countries (BARTELSMAN et al., 2017; DESTEFANO et al., 2018; HALLER and LYONS, 2015), whilst others found that technology adoption, and its benefits, are not uniform across space and firm types (JONES and HENDERSON, 2019; REUSCHKE and MASON, 2020).

Self-selection in technology adoption and complementarities between digital technologies and firm-specific organisational capital might be significant factors in explaining these contrasting results (GAL et al., 2019). The evidence is stronger, however, in suggesting that larger firms are more likely to adopt digital technologies such as cloud computing (OECD, 2017). If these larger and more productive firms tend to concentrate in larger cities (BEHRENS et al., 2014; COMBES et al., 2012), they may generate spatially-constrained externalities. These externalities can, in turn, benefit firms and industries that are not directly involved in the development and adoption of specific technological advances by just being located nearby (COHEN and LEVINTHAL, 1990).

At the country-level, KOUTROUMPIS (2009) identified positive causal effects of broadband internet investments on economic growth (10% of annual GDP growth) for 22 OECD countries for the period 2002-2007. These results agree with studies on (i) 66 high-income countries for 1980-2002 (QIANG et al., 2009); (ii) 25 OECD countries for 1992-2002 (BELORGEY et al., 2006); (iii) 25 OECD countries for 1996-2007 (CZERNICH et al., 2011; GRUBER et al., 2014); all OECD countries for 2008-2010 (ROHMAN and BOHLIN, 2012). ARVIN and PRADHAN (2014) focused on G-20 countries for the 1998-2011 period, and their results indicated short-run and bidirectional causal effects between broadband penetration and economic growth among the more developed countries. However, for developing countries within G-20, this relationship was unidirectional from economic growth to penetration. NAJARZADEH et al. (2014) studied internet usage in 108 countries during 1995-2010 and found positive effects on labour productivity. CECCOBELLI et al. (2012) examined the relationship of ICT and productivity in 14 OECD countries during 1995-2005 and find that the diffusion of general-purpose ICT needs to be accompanied by broader organisational and business changes in order to generate productivity effects.

From an urban perspective, AHLFELDT et al. (2017) addressed the spatial heterogeneity of these effects at a granular level by estimating the effect of broadband speed on property values. Their results indicated that an upgrade to a first-generation broadband connection led to property price increases of up to 2.8% on average. At a different scale, TRANOS (2012) illustrated the economic effect of direct connectivity to internet backbone networks for city-regions. These effects are not consistent in space and depend upon the absorptive capacity of city-regions. KOLKO (2012) found a positive causal effect between the expansion of broadband provision and local economic growth in the US. Research on US counties also revealed a bidirectional and spatially heterogeneous relationship between the change of broadband providers and the change of knowledge-intensive business service firms (TRANOS and MACK, 2016).

From the above, we observe that although several studies focused on the effects of digitisation – as reflected in digital infrastructure provision and quality – on labour markets, economic growth or entrepreneurship, only a handful of them addressed regional productivity. MACK and FAGGIAN (2013) studied US counties during 2000-2007 and showed that broadband provision has a positive productivity effect only in counties with a high concentration of human capital and highly-skilled occupations. JUNG and LÓPEZ-BAZO (2019) found positive but spatially heterogeneous effects of broadband provision on regional productivity in Brazil with higher benefits for the less developed regions, suggesting a regional convergence role for digitisation.

Consequently, the need to understand the potential productivity effects of digitisation becomes even more apparent when we consider the relationship between space and the use of digital technologies. Heated debates can be found in economic geography and urban economics literature regarding whether digitisation processes can affect urban structure by intervening with agglomeration related benefits (for a review see DADASHPOOR and YOUSEFI, 2018; TRANOS, 2020; TRANOS and IOANNIDES, 2020). Simply put, the question is whether digital technologies can act as a substitute or a supplement of agglomeration forces due to their capacity to decrease distance-related transactions costs as they are communication technologies in their core (TRANOS and NIJKAMP, 2013). Moreover, the extensive adoption of such technologies has led to digitisation of processes and products and, consequently, to an abundance of online content, which can challenge gravitational forces and related externalities. Even static online content, which is what this paper focuses upon, can act as a source of information and codified knowledge that is relevant at the local context. This, may lead to knowledge spillovers and decreased transaction costs and, therefore to productivity gains in places outside the traditional core (BATHELT and COHENDET, 2014; FARAJ et al., 2016).

To summarise, firm productivity, business establishments and labour markets appear to be positively affected by increases in broadband provision. However, such effects are spatially heterogeneous, might also be related to other complementary investments and are higher for technologically sophisticated sectors and for urban areas (THE WHAT WORKS CENTRE FOR LOCAL ECONOMIC GROWTH, 2015). Using the above as a point of departure, we contribute to this literature by (i) moving beyond infrastructural and user-based approximations of the level of digitisation of the economy, and (ii) by exploring the role of early engagement with the digital economy in productivity as well as the longevity of these impacts due to digitisation.

Archiving the web to learn about the digital economy

Capturing the level of digitisation is not trivial. While most common measures illustrate the supply of digital infrastructure, they provide a little insight into the use of this infrastructure. Similarly, measures of individual internet usage patterns offer limited information on the digital economic activities that individuals engage with – except for e-retail (SINGLETON et al., 2016). Consequently, we have a narrow understanding of whether digital engagement is active (i.e. setting up a website) or passive (i.e. following the news) and whether this engagement generates value-added.

To contribute to the measurement of engagement with the digital economy, we employ a novel data set of archived webpages. These public domain data originate from the IA, which started crawling and archiving the web in 1996. We utilise the JISC UK Web Domain Dataset, which is curated by the British Library[3] and contains all the archived webpages from the UK ccTLD. It constitutes a list of 2.5 billion of .uk webpages, which have been archived during 1996-2013, including the archiving timestamp (JISC AND THE INTERNET ARCHIVE, 2013). The contents of these archived webpages have been scanned by the British Library to identify webpages that include a string of text resembling UK postcodes (e.g. B15 2TT)[4]. This subset of the initial pool of all the .uk archived webpages (called Geoindex) includes 0.5 billion archived webpages and is the basis of our analysis (JACKSON, 2013).

Because we can geolocate these webpages, we create measures of the volume of online content anchored to UK regions and test their effect on regional productivity. Such measures serve well as proxies for the level of digitisation. Instead of measuring how often people connect to the internet or how many internet users can simultaneously download a large video file (i.e. broadband speed) the web data capture the outcome of the engagement with the digital economy.

We can distinguish different types of websites based on Second Level Domains (SLD). For instance, we can approximate digital economic activities in a region by measuring the volume of commercial webpages by considering only .co.uk websites, which are dedicated to commercial activities (THELWALL, 2000). Such commercial websites are used to exchange information, support online transactions and share opinions (BLAZQUEZ and DOMENECH, 2018). Although nothing prevents a UK-based company from adopting a generic TLD such as .com and, indeed, such cases escape our data, we do not expect that such omissions could affect our results given the popularity of the .uk ccTLD: UK consumers prefer to visit a .uk website when they are searching for products or services (HOPE, 2017); and, anecdotal

[3] http://data.webarchive.org.uk/opendata/ukwa.ds.2/

[4] False positives (e.g. XX1 1XX, which is not a UK postcode) were excluded from the data.

evidence indicates that during the first half of 2000 three .co.uk domains were registered every minute (OECD, 2001). Also, as Table 1 demonstrates, .co.uk was the most popular SLD under the .uk ccTLD in 2000. Hence, we use the total volume of all the archived webpages in a region as a proxy for the level of digitisation. We start the analysis by focusing on commercial webpages (.co.uk).

Insert Table 1 around here

Essential for the analysis is also the temporal dimension of these data as we use the archival timestamp to create annual aggregates and match them with productivity measures. We use such data from as early as 2000 to measure the early adoption of digitisation practices. This is a period before online social media and smartphones, when Web 1.0 type of applications (e.g. static webpages) were dominating the web and only 25 per cent of UK households had access to the internet (compared to 90 per cent in 2018, ONS, 2018c). The volume of online content during that period reflects the early adoption of internet-related technologies and digitisation practices within the UK regions. The temporal depth of these data is a unique attribute compared to infrastructural and survey-based measures of digital activities.

The IA discovers webpages by following the hyperlinks of every webpage its archives. This almost 25 year-long snowball-like data collection process captured a significant part of the web – more than 371 billion webpages (INTERNET ARCHIVE, 2019). However, these data are not bias-free. Popular webpages, which have a lot of backward links (other webpages with hyperlinks towards them) tend to be archived more often and have a higher likelihood of being archived (HALE et al., 2017b). The frequency bias does not impose issues for this study since we avoid double-counting by considering individual webpages only once per year.

Concerning the extent of the archive, this is not trivial to assess given that the actual size of the web is unknown. Yet, digital humanities studies aimed to assess the coverage of such archives, agree that the IA is the most extensive and complete archive in the world (AINSWORTH et al., 2011a; HOLZMANN et al., 2016b). THELWALL and VAUGHAN (2004) utilised a sample of the IA similar to ours – commercial websites – and concluded that 92 per cent of all the US commercial websites had been archived.

Regarding the archival depth – how many webpages from a specific website have been archived – HALE et al. (2017a) attempted to evaluate it by juxtaposing live and archived webpages from the London website of Trip Advisor. Their analysis indicated that only 24 per cent of these webpages were archived, with webpage popularity being the main driver of the bias. We partially address this issue by using different subsets of the data based on the number of unique postcodes included in each website. Although the initial level of observation is single webpages archived on a given point in time, we are able to reconstruct archived websites and create counts of the number of unique postcodes included in each reconstructed website. Based on the below example, the reconstructed website http://www.examplewebsite.co.uk includes two unique postcodes mentioned in the below three archived webpages, which are linked to this website.

http://www. examplewebsite.co.uk/webpage1	B15 2TT
http://www. examplewebsite.co.uk/webpage2	BS8 1TH
http://www. examplewebsite.co.uk/webpage3	B15 2TT

As Table 2 illustrates, our data include a wide range of reconstructed commercial (.co.uk) websites with regards to the number of postcodes they contain. At one end, we have websites anchored to a unique location (78 per cent of all the reconstructed websites). Such websites may represent a small company with a single plant. At the other end, we have websites with thousands of different postcodes. These can be directory-type websites, advertising professionals around the country (see Appendix A Figure A1 for examples).

We start our analysis by counting all the geolocated webpages in each UK region in order to capture the total volume of digitisation including, for example, the trivial effort of individuals or business to participate to such directories. If a webpage includes multiple postcodes, it is then counted multiple times. To address potential archival depth bias, we replicate our analysis by counting only the volume of archived webpages from websites with a unique postcode. In other words, we only consider websites linked to a unique location, which are less likely to be affected by the archival depth bias.

The geolocation process based on postcode mentions in the web text does not suffer by IP geolocation issues (ZOOK, 2000) or by the 'here and now' complexities of user-generated online social media data (CRAMPTON et al., 2013). Given that these are public-facing websites, we expect that the postcodes refer to trading vis-à-vis registration addresses when we focus on commercial websites (.co.uk) or other anchor points. Similarly, we do not expect that these postcodes will refer to the web designer location, as this is not a common practice.

Insert Table 2 around here

Another potential source of bias is that the IA, like other digital archives, can only archive publicly accessible webpages without robot exclusions[5]. For instance, password-protected Facebook pages cannot be archived. This is not a concern for our study since we are primarily interested in websites in 2000 and social media were not present then. Similarly, we do not expect abandoned websites to bias our data, since there would only be very few abandoned websites in 2000 and such websites would lose popularity and backward links and, eventually, not be captured by the crawler.

Apart from the work of MEIJERS and PERIS (2019), who employed data from a digital archive to illustrate the network embeddedness of Dutch cities, such data have not been used before in regional studies. This is not the case though for digital humanities (BRÜGGER and MILLIGAN, 2018) and social sciences. For example, archived web data have been used by PAPAGIANNIDIS et al. (2015) to analyse the diffusion of specific web technologies (e.g. JavaScript) and by PAPAGIANNIDIS et al. (2018) to build industrial classifications. BLAZQUEZ and DOMENECH (2018) employed archived versions of corporate websites to assess the export orientation of Spanish companies whilst archived data have also been used to examine innovation processes. ARORA et al. (2013) and SHAPIRA et al. (2016) studied the early commercialisation strategies of novel graphene technologies; GÖK et al. (2015) analysed R&D activities, and LI et al. (2016) built Triple Helix indices of green goods for small and mid-size businesses. The same dataset used here was also employed by MUSSO and MERLETTI (2016) to reconstruct the UK business web space during 1996-2001. Most of the above studies were limited in their scope (i.e. focused on small samples of archived web data) and overlooked

[5] These are standard website exclusions practices regarding how accessible they are by other websites and web crawlers and are included in a file called robots.txt

the spatial reflections of the data, despite estimates that around 70% of all websites contain place reference (HILL, 2009).

Methodological approach: The effect of digitisation on regional productivity

Most studies on regional productivity employ the traditional neoclassical growth model (SOLOW, 1956), where regions have identical long-run growth rates determined by exogenous technological progress and steady-state growth paths that are parallel. This assumption excluded heterogeneous regional contexts and led to the development of endogenous growth models during the 1980s enabling the comparative analysis of regional long-run behaviour (ROBERTS and SETTERFIELD, 2010).

The first contributions to the endogenous growth literature (LUCAS, 1988; ROMER, 1986) did not explicitly distinguish between capital accumulation and technological progress, physical or human capital. Further developments (innovation-based growth theory) differentiated physical and human capital, accumulated through saving and schooling, to intellectual capital, accumulated through innovation. ROMER (1990) proposed a version of innovation-based theory where productivity was assumed to be a function of the degree of product variety. Hence, innovation causes productivity growth by creating new, but not necessarily improved, varieties of products. An alternative is the one-country Schumpeterian model developed by AGHION and HOWITT (1992, 1998) which focused on quality-improving innovations that render old products obsolete, through creative destruction.

Many models of endogenous growth (ARROW, 1962; LUCAS, 1988; ROMER, 1986) can be written as follows:

$$Y(t) = A(t)F(K,L) \qquad (1)$$

Capital input *K* and labour input *L* are slow to change, so it is possible to achieve faster output growth only by improving resource utilisation (constant *A*).

Empirical contributions incorporated additional factors such as path dependency, knowledge spillovers and industrial structure characteristics. Expertise in higher technological advances tends to be concentrated in space with clusters expanding around the existing ICT hubs and central research departments (GOLDFARB and TREFLER, 2018; KLINGER et al., 2018). Proximity to universities supports innovation by small and medium enterprises (SMEs) through knowledge exchange since SMEs often lack the capacity to perform R&D (ACS et al., 2009). Overall, results suggest that investments in intangibles such as R&D or ICT, lead to *learning-by-doing* effects and greater benefits when they are complemented with investments in human capital (JONES, 2002; ORTEGA-ARGILÉS, 2013).

Drawing upon this literature, our analysis focuses on the regional productivity gains of digitisation by considering the active engagement with the digital economy measured by the volume of online content anchored to specific regions. Our model tests whether the early adoption of digital economy practices has led to positive and long-term productivity effects. We use a 17-year panel (2000-2016) and methods that allow us to control for time-invariant, unobserved heterogeneity between regions. Equation 2 shows the form of our preferred specification, estimated using fixed-effects[6] and a Hausman-Taylor estimator with heteroscedasticity-robust standard errors, clustered at the NUTS3 level.

$$\log\left(\frac{GVA}{E}\right)_{it} = \beta_0 + \sum_{j=1}^{16}\beta_j \log(DE2000_i) * YEAR_t + \gamma_1 logDE2000_i + Z_1 X_{it} + a_i + \delta_t + u_{it} \quad (2)$$

[6] The Hausman test indicated that the FE specification is preferred to a random-effects (RE) one (BALTAGI et al., 2003).

We measure regional productivity as the gross value added (*GVA*) per employee (*E*) in region *i* and time *t* (CARDONA et al., 2013). This measure offers an estimate of the value creation per employee in each region, approximating average labour productivity for each year. The advantage of using a labour-related measure for productivity is the ease of calculation, its prevalence and applicability across regions, industries and similar studies. GVA, which subtracts intermediate inputs from the gross output, is considered a more accurate measure of the actual surplus created by the regional economy (CARDONA et al., 2013; SENA, 2020).

Our main variable of interest is the interaction between the volume of online content in 2000 in region *i*, which reflects the early adoption of internet-related technologies and digitisation practices (*DE* – digital economy), and yearly dummies (*YEAR)*. This interaction term captures the regional productivity effect of the early adoption of digitisation practices in subsequent years (2001-2016). Aiming at considering the production of different types of online content, several versions of this variable have been analysed (e.g. commercial vs. non-commercial websites and local vs. websites with national reach).

Figure 1 plots two different variables showing the level of digitisation in years 2000 and 2010 standardised by regional population. The top part of the figure plots the volume of commercial archived content, whilst the bottom part shows the volume of non-commercial content. We are mostly interested in year 2000 to capture the early adoption effect. The distribution of online content is far from even across UK regions, and the high concentration in and around London diffuse to the rest of the country over time. Because of this highly skewed distribution, we transform these variables using the natural logarithm for our estimations. The same applies to some of the control variables.

Besides the region (a_i) and time (δ_t) fixed-effects, X_{it} is a vector of regional control variables – and Z_1 is a vector of their estimated coefficients – which includes:

a) Human capital measured as the share of population with National Vocational Qualification level 4 (NVQ4) and above. Expected to have a positive effect on productivity via technological change and innovation.
b) Gross Fixed Capital Formation to capture R&D activity that is expected to positively affect productivity (PRENZEL et al., 2018). Because of lack of NUTS3 data, we use corresponding NUTS2 data.
c) Two variables to capture the existence of Jacobian or Marshallian externalities. We use population density to test for urbanisation economies (JACOBS, 1970) and the inverse of a Herfindahl-Hirschman Index (HHI) to account for the effects of diversification (higher values of the index) or specialisation (lower values of the index) on productivity.
d) Employment density to account for the concentration of economic activity. Due to the construction of this variable (employment/population), it is expected to be negatively related to productivity (GVA/employment).
e) The share of manufacturing, which captures local industrial structures with a large share of manufacturing. To the extent that this is based on high-tech manufacturing, we would expect a positive sign whilst if the focus is in low-tech industries, the relationship to productivity could be negative (PRENZEL et al., 2018).
f) The population of each NUTS3 region, expecting that larger regions offer greater opportunities for agglomeration externalities and local firms are more productive.
g) The number of universities to capture local innovation spillovers.

h) Distance to London to account for the proximity effects to the largest, most productive conurbation in the UK.

Insert Figure 1 around here

Figure 1: Volume of online content in NUTS3

The data for the above control variables come from established sources such as the UK Office for National Statistics (ONS, 2018d, e), the Annual Population Survey (ONS, 2018a) and the Business Register and Employment Survey (ONS, 2018b). To maintain consistency with NUTS3 revisions, we build these datasets on the overlap between Local Authority Districts (LADs) and NUTS3 regions. In doing so, we have three LADs being matched in more than one NUTS3 region in Scotland, which have been, therefore, excluded from the analysis. In total, we create a 2000-2016 panel dataset for 163 NUTS3 regions in Great Britain (see Appendix A Table A1).

Results

The main variable of interest in the first column (Table 3) is the logarithm of the volume of commercial websites during 2000-2010. This regression tests the contemporaneous effect of digitisation on regional productivity in the UK. Interestingly, the estimation did not yield a statistically significant coefficient, a result that is consistent with the Solow paradox on the lack of observable productivity benefits from DE in the short-run.

However, the contemporaneous setting ignores the potentially important role of the early adoption of such technologies. To test the latter, we adopt a different identification strategy in Column 2. Instead of using the volume of commercial online content during 2000-2010, we only consider the 2000 level and introduce an interaction term between the volume of commercial online content in 2000 and yearly dummy variables for all but one years during 2000-2016 (base year is 2000). We interpret the coefficients of this interaction as the effect of the volume of commercial online content in 2000 on regional productivity in subsequent years. Hence, we observe the long-term effect of the early adoption of internet-related technologies and digitisation practices on regional productivity. The relevant coefficients are in Table 3 and illustrated in Figure 2 below.

Overall, the model shows positive and statistically significant coefficients for the interaction term in almost all years in our dataset. Importantly, the magnitude of these coefficients follows a positive trend, indicating the longevity and increasing importance of the effect. The coefficients are not significant for years 2004 and 2005 and then decrease during 2008-2011. The lack of significance for 2004 and 2005 corresponds with the dot-com crash. Empirical evidence suggests that while the productivity of hardware manufacturing decreased post-2000, the productivity linked to ICT usage increased (JORGENSON et al., 2011). Hence, the non-significance coefficients for 2004 and 2005, which are consistent across different specifications discussed below, might be explained by disturbances related to the dot-com crash. Also, we attribute the decreasing magnitude of the coefficients during 2008-2011 to the 2008 financial crisis and its impact on productivity (GRIFFITH and MILLER, 2010). Thus, the FE estimation captures the long-term and upward positive effect of the volume of online commercial content in 2000.

To further validate the above results, Column (3) employs a Hausman-Taylor estimator (HT) to consider the effect of time-invariant variables. We are interested in such variables for three reasons. Firstly, a coefficient for the effect of the initial volume of commercial online content in 2000 accounts for the initial conditions (*(log).co.uk 2000*). Secondly, we directly control for the initial productivity levels (*Productivity 2000*) which may affect subsequent productivity levels (NEFFKE et al., 2011; SIMMIE et al., 2014). Thirdly, we control for other time-invariant characteristics, which may affect regional productivity (*Nr of Universities*; *Distance to London*).

HT allows for some variables to be correlated with the regional effects contrary to the RE model which assumes exogeneity of all regressors and the FE model which permits the correlation of all the explanatory variables and the regional effects (HAUSMAN and TAYLOR, 1981). To achieve this, the HT estimator utilises both the within and between variation of the exogenous variables as instruments for the endogenous variables (HAILEMARIAM and DZHUMASHEV, 2019; PALACÍN-SÁNCHEZ and DI PIETRO, 2016; RODRÍGUEZ-POSE and KETTERER, 2012) and addresses the three challenges outlined above: (i) we obtain a coefficient for the volume of commercial online content in 2000; (ii) we control for the initial level of productivity; and (iii) we control for other time-invariant variables that may affect productivity.

Column (3) in Table 3 presents the HT estimates. The volume of online commercial content in 2000 has a significant positive effect on average productivity during 2001-2016. In addition, the coefficients for the interaction terms are almost identical with the FE estimation (see also Figure 2). This further supports our argument about the long-term positive productivity effects of the early adoption of digital technologies. Importantly, we control for a number of other factors influencing regional productivity. Human capital, agglomeration and physical capital have a significant and positive effect, which is consistent with the existing literature. Manufacturing also has a significant positive effect, but it has been masked out by the interaction terms. Concurrently, population and employment density have – in line with our expectations - negative and consistent effects. The number of regional universities is positively associated with productivity, and there is a negative relationship between distance to London and productivity, but it is marginally significant. Additionally, the HT estimation controls for the initial level of productivity, which is positive and significant as expected. Diversification does not have a systematically significant effect on productivity, a finding which concurs with Caragliu et al. (2016) and de Groot et al. (2016).

Insert Table 3 around here

Insert Figure 2 around here

Figure 2: Magnitude of statistically significant coefficients (p<0.10) for the interaction term $log(DE_{2000}) * YEAR$ in the FE (2) and HT (3) specifications in Table 3

Robustness checks

To further test the robustness of the above findings we create different subsets of the archived dataset and we alternate the main variable of interest among two dimensions: the archival year (see below) and the type of the online content (see discussion in Appendix B and Table B3 due to space constraints). Both sets of tests confirm our results and highlight the positive impacts of early engagement with the digital economy on regional productivity differentials.

The estimations so far used online content from 2000 to capture the effect of the early adoption of web technologies on regional productivity. We now repeat the regression from Column (3) in Table 3 using the archived online commercial content in 2002, 2003, 2004, 2006 and 2008[7] as the main variable of interest. The HT regressions are in Appendix B (Table B1),

[7] These years were selected in order to show the evolution of coefficients for DE from different years whilst still maintaining a sizeable sample. Results for 2005 and 2007 reflect what is already observed and are available upon request from the authors.

and the coefficients for the interaction effect are presented in Figure 3. The use of online content from later years leads to shorter panel data sets and, therefore, to significant coefficients for fewer years. The trend of the coefficients mirrors the main findings, but importantly the magnitude of the coefficients decreases with time.

Insert Figure 3 around here

Figure 3: Magnitude of statistically significant coefficients (p<0.10) for the interaction terms of different yearly DE variables (Table B1 in Appendix B)

In order to compare the coefficient estimates of year 2000 to those presented in Figure 4, we use the coefficients of their *interaction* terms (WOOLDRIDGE, 2010)[8]. In essence, we compare pairs of models based on two different panels: one including the $DE_{2000} * YEAR$ and one including the $DE_{2000+x} * YEAR$ interaction terms, where *x* is equal to 2, 3, 4, 6, 8. We nest these together by appending the two panels of interest and generate a dummy variable d which is equal to 1 for each of the panels containing the $DE_{2000+x} * YEAR$ term and 0 for the original (baseline) panel including the $DE_{2000} * YEAR$ term. This allows us to separate the results of the two interaction terms and run the following model:

$$\log\left(\frac{GVA}{E}\right)_{it} = \beta_0 + \sum_{j=1}^{16} \beta_j \log(DE2000_i) * YEAR_t + \sum_{z=1}^{16-x} \lambda_z \log(DE2000 + x)_i * YEAR_t * d + \gamma_1 logDE2000_i + Z_1 X_{it} + \eta_1 * d + a_i + \delta_t + u_{it} \quad (3)$$

where β_t are the coefficients for the interaction of the original *DE* variable with the year dummies ($\log(DE2000_i) * YEAR$) whilst the λ_z coefficients give us the difference in the coefficients between $\log(DE2000_i) * YEAR$ and each of the $\log(DE2000 + x)_i * YEAR$ for each year and *x*.

[8] For a simple implementation of the test see the following webpage: https://www.stata.com/support/faqs/statistics/test-equality-of-coefficients/.

The results of this test are in Appendix B (Table B2) where it can be seen that the coefficients of the new interaction terms are all significant, suggesting that the estimates are significantly different. In particular, the negative sign of the λ_z coefficients observed, show that our measures of *DE* for the years 2002, 2003, 2004, 2006 and 2008 have a statistically significant smaller effect on productivity than *DE* for 2000. As a result, the information contained in earlier years has larger effects than following years in explaining productivity during the period we examine.

Conclusions

This paper contributes to the literature examining the regional economic effects of digitisation in two ways. Firstly, we focus on the early adoption of digital technologies and on its long-term productivity effects instead of exploring their contemporaneous relationship. Secondly, we use a novel measure to capture the outcomes of the active engagement with the digital economy, namely the creation of online content via data of archived webpages. This complements the latest contributions to the relevant literature, which approximate the level of digitisation using mostly infrastructural measures.

By measuring the volume of static online content, we are able to test whether the related knowledge spillovers and decreased transaction costs can lead to regional productivity gains. Linking to recent debates in the economic geographic literature regarding the potential of digital technologies to reshape agglomeration-related benefits, our research strategy enables us to consider spatial heterogeneity and how the observed and unobserved regional characteristics play a role in the relationship between digitisation and regional productivity.

Our results highlight the long-term effects of the early adoption of digital technologies and digitisation practices in explaining regional productivity in the UK. Our estimates show a clear upward and longstanding positive effect of the volume of online commercial content in early years on regional productivity after controlling for several related variables such as human capital, agglomeration economies and physical capital formation. This productivity gain is associated with the early adoption of digital technologies rather than contemporaneous effects which are absent. We also establish that adoption in earlier years has higher effects on subsequent productivity compared to engagement with the digital economy in later years. In this view, digitisation has the characteristics of radical innovation with early year investments having long-term productivity benefits and higher returns than adoption in later stages. A potential explanation of these diminishing returns in time is offered by the increasing depreciation of ICT capital and first-mover advantages.

The findings are robust to several sensitivity checks. The archived web data allow for the creation of different variables based on the type and location of anchoring of the online content. All these different variables confirm our initial results indicating the role of the early adoption of digital practices in generating long-term productivity effects. Our results are aligned with previous studies, which focus on the supply of digital infrastructure and internet usage. Importantly, our approach allows us to better capture the nuances of engaging in digital economic activities as reflected in the different SLD.

To support the design and implementation of policies and their impact on interregional inequalities, our paper suggests that the decision to roll out digital technologies may have significant and spatially heterogeneous effects on local economic growth. Early engagement with these digital technologies is associated with higher productivity levels in subsequent years. Policymakers should be aware of these spatially varying and long-lasting effects of digitisation in the UK for the implementation of place-based policies with a particular emphasis on evolutionary and path dependence aspects. Place-based strategies should be centred on engaging with digital practises with special emphasis on geographies and firms that are less likely to adopt given their historical records.

Furthermore, the long-term effects on productivity have consequences on the evaluation of relevant policies and initiatives. Most evaluations tend to consider 1, 2 and 5-year effects post-policy implementation. Our findings of positive productivity effects as far as 15 years after engagement with the digital economy in 2000 suggest that evaluation programmes are likely to underestimate the effects of such policies. More long-term evaluation schemes would be better equipped to capture the lagged structure of the aforementioned impacts.

Acknowledgments

The authors would like to acknowledge the useful comments from the participants of the 2019 Winter Seminar of the German speaking section of RSAI and the 2018 ERSA Congress.

References

ACEMOGLU D., DORN D., HANSON G. H. and PRICE B. (2014) Return of the Solow paradox? IT, productivity, and employment in US manufacturing, *American Economic Review* **104**, 394-99.

ACS Z. J., BRAUNERHJELM P., AUDRETSCH D. B. and CARLSSON B. (2009) The knowledge spillover theory of entrepreneurship, *Small Business Economics* **32**, 15-30.

AGHION P. and HOWITT P. (1992) A Model of Growth Through Creative Destruction, *Econometrica* **60**, 323-51.

AGHION P. and HOWITT P. (1998) *Endogenous Growth Theory MIT Press*. MIT Press, Cambridge, MA.

AHLFELDT G. M., KOUTROUMPIS P. and VALLETTI T. M. (2017) Speed 2.0-Evaluating Access to Universal Digital Highways, *Journal of the European Economic Association* **15**, 586–625.

AINSWORTH S. G., ALSUM A., SALAHELDEEN H., WEIGLE M. and NELSON M. (2011a) How much of the web is archived?, *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries*, pp. 133-6. ACM.

AINSWORTH S. G., ALSUM A., SALAHELDEEN H., WEIGLE M. C. and NELSON M. L. (2011b) How much of the web is archived?, *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries*, pp. 133-6. ACM.

ARORA S. K., YOUTIE J., SHAPIRA P., GAO L. and MA T. (2013) Entry strategies in an emerging technology: a pilot web-based study of graphene firms, *Scientometrics* **95**, 1189-207.

ARROW K. J. (1962) The Economic Implications of Learning by Doing, *The Review of Economic Studies* **29**, 155-73.

ARVIN B. M. and PRADHAN R. P. (2014) Broadband penetration and economic growth nexus: evidence from cross-country panel data, *Applied Economics* **46**, 4360-9.

ATKINSON R. D. and MCKAY A. (2007) Digital Prosperity. ITIF, Washington, DC.

BAILIN RIVARES A., GAL P., MILLOT V. and SORBE S. (2019) Like it or not? The impact of online platforms on the productivity of incumbent service providers.

BALTAGI B. H., BRESSON G. and PIROTTE A. (2003) Fixed effects, random effects or Hausman–Taylor?: A pretest estimator, *Economics Letters* **79**, 361-9.

BARTELSMAN E., VAN LEEUWEN G. and POLDER M. (2017) CDM using a cross-country micro moments database, *Economics of Innovation and New Technology* **26**, 168-82.

BATHELT H. and COHENDET P. (2014) The creation of knowledge: local building, global accessing and economic development—toward an agenda, *Journal of Economic Geography* **14**, 869-82.

BEHRENS K., DURANTON G. and ROBERT-NICOUD F. (2014) Productive Cities: Sorting, Selection, and Agglomeration, *Journal of Political Economy* **122**, 507-53.

BELORGEY N., LECAT R. and MAURY T.-P. (2006) Determinants of Productivity per Employee: An Empirical Estimation Using Panel Data, *Economics Letters* **91**, 153-7.

BERTSCHEK I., CERQUERA D. and KLEIN G. J. (2013) More bits–more bucks? Measuring the impact of broadband internet on firm performance, *Information Economics and Policy* **25**, 190-203.

BLANK G., GRAHAM M. and CALVINO C. (2018) Local geographies of digital inequality, *Social Science Computer Review* **36**, 82-102.

BLAZQUEZ D. and DOMENECH J. (2018) Web data mining for monitoring business export orientation, *Technological and Economic Development of Economy* **24**, 406-28.

BLOOM N., SADUN R. and REENEN J. V. (2012) Americans Do IT Better: US Multinationals and the Productivity Miracle., *American Economic Review* **102**, 167-201.

BRÜGGER N. and MILLIGAN I. (2018) *The SAGE Handbook of Web History*. SAGE Publications Limited.

BRYNJOLFSSON E. and HITT L. (1996) Paradox lost? Firm-level evidence on the returns to information systems spending, *Management science* **42**, 541-58.

BRYNJOLFSSON E., ROCK D. and SYVERSON C. (2017) Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. National Bureau of Economic Research.

CARAGLIU A., DE DOMINICIS L. and DE GROOT H. L. F. (2016) Both Marshall and Jacobs were Right!, *Economic Geography* **92**, 87-111.
CARDONA M., KRETSCHMER T. and STROBEL T. (2013) ICT and productivity: conclusions from the empirical literature, *Information Economics and Policy* **25**, 109-25.
CECCOBELLI M., GITTO S. and MANCUSO P. (2012) ICT capital and labour productivity growth: A non-parametric analysis of 14 OECD countries, *Telecommunications Policy* **36**, 282-92.
CHAKRAVORTI B. and CHATURVEDI R. S. (2017) Digital planet 2017: How competitiveness and trust in digital economies vary across the world, *The Fletcher School, Tufts University* **70**, 70.
COHEN W. M. and LEVINTHAL D. A. (1990) Absorptive Capacity: A New Perspective on Learning and Innovation, *Administrative Science Quarterly* **35**, 128-52.
COMBES P.-P., DURANTON G., GOBILLON L., PUGA D. and ROUX S. (2012) The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection **80**, 2543-94.
CRAMPTON J. W., GRAHAM M., POORTHUIS A., SHELTON T., STEPHENS M., WILSON M. W. and ZOOK M. (2013) Beyond the geotag: situating 'big data'and leveraging the potential of the geoweb, *Cartography and geographic information science* **40**, 130-9.
CZERNICH N., FALCK O., KRETSCHMER T. and WOESSMANN L. (2011) Broadband infrastructure and economic growth*, *The Economic Journal* **121**, 505-32.
DADASHPOOR H. and YOUSEFI Z. (2018) Centralization or decentralization? A review on the effects of information and communication technology on urban spatial structure, *Cities* **78**, 194-205.
DE GROOT H. L. F., POOT J. and SMIT M. J. (2016) WHICH AGGLOMERATION EXTERNALITIES MATTER MOST AND WHY?, *Journal of Economic Surveys* **30**, 756-82.
DESTEFANO T., KNELLER R. and TIMMIS J. (2018) Broadband infrastructure, ICT use and firm performance: Evidence for UK firms, *Journal of Economic Behavior & Organization* **155**, 110-39.
FARAJ S., VON KROGH G., MONTEIRO E. and LAKHANI K. R. (2016) Special Section Introduction—Online Community as Space for Knowledge Flows, *Information systems research* **27**, 668-84.
GAL P., NICOLETTI G., RENAULT T., SORBE S. and TIMILIOTIS C. (2019) Digitalisation and productivity: In search of the holy grail – Firm-level empirical evidence from EU countries.
GÖK A., WATERWORTH A. and SHAPIRA P. (2015) Use of web mining in studying innovation, *Scientometrics* **102**, 653-71.
GOLDFARB A. and TREFLER D. (2018) AI and international trade, *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.
GRIFFITH R. and MILLER H. (2010) UK productivity in the recession. Institute for Fiscal Studies.
GRIFFITH R., REDDING S. and REENEN J. V. (2004) Mapping the two faces of R&D: Productivity growth in a panel of OECD industries, *Review of economics and statistics* **86**, 883-95.
GRILICHES Z. (1979) Issues in assessing the contribution of research and development to productivity growth, *Bell Journal of economics* **10**, 92-116.
GRUBER H., HÄTÖNEN J. and KOUTROUMPIS P. (2014) Broadband access in the EU: An assessment of future economic benefits, *Telecommunications Policy* **38**, 1046-58.
HAILEMARIAM A. and DZHUMASHEV R. (2019) Fiscal equalization and composition of subnational government spending: implications for regional convergence, *Regional Studies* **53**, 587-601.
HALE S. A., BLANK G. and ALEXANDER V. D. (2017a) Live versus archive: Comparing a web archive to a population of web pages, in BRÜGGER N. and SCHROEDER R. (Eds) *Web as History: Using Web Archives to Understand the Past and the Present,*. UCL Press, London.
HALE S. A., BLANK G. and ALEXANDER V. D. (2017b) Live versus archive: Comparing a web archive to a population of web pages, in BRÜGGER N. and SCHROEDER R. (Eds) *Web as History: Using Web Archives to Understand the Past and the Present*, pp. 45-61. UCL Press, London.
HALLER S. A. and LYONS S. (2015) Broadband adoption and firm productivity: Evidence from Irish manufacturing firms, *Telecommunications Policy* **39**, 1-13.
HAUSMAN J. A. and TAYLOR W. E. (1981) Panel data and unobservable individual effects, *Econometrica* **49**, 1377-98.

HENDERSON D. and ROCHE N. (2020) Examining the policy mix for broadband deployment in Wales: The role of informal coordination in the last mile, *Local Economy* **35**, 48-67.
HILL L. L. (2009) *Georeferencing: The geographic associations of information*. MIT Press, Cambridge, MA.
HM GOVERNMENT (2017) UK Digital Strategy. Department for Digital, Culture, Media & Sport, London.
HOLZMANN H., NEJDL W. and ANAND A. (2016a) The Dawn of today's popular domains: A study of the archived German Web over 18 years, *Digital Libraries (JCDL), 2016 IEEE/ACM Joint Conference*, pp. 73-82. IEEE.
HOLZMANN H., NEJDL W. and ANAND A. (2016b) The Dawn of today's popular domains: A study of the archived German Web over 18 years., *Digital Libraries (JCDL), 2016 IEEE/ACM Joint Conference*, pp. 73-82. IEEE.
HOPE O. (2017) The changing face of the online world. Nominet, Oxford, UK.
HOUSEMAN S. N., BARTIK T. J. and TIMOTHY J. (2014) Sturgeon. Measuring Manufacturing: How the Computer and Semiconductor Industries Affect the Numbers and Perceptions, in HOUSEMAN S. N. and MANDEL M. (Eds) *Measuring Globalization: Better Trade Statistics for Better Policy Volume 1. Biases to Price, Output, and Productivity Statistics from Trade*, pp. 151-94. W.E Upjohn Institute, Kalamazoo, MI.
INTERNET ARCHIVE (2019) Internet Archive home page.
JACKSON A. N. (2013) *JISC UK Web Domain Dataset (1996-2010) Geoindex*, https://doi.org/10.5259/ukwa.ds.2/geo/1.

JACOBS J. (1970) *The economy of cities*. Random House.
JISC AND THE INTERNET ARCHIVE (2013) *JISC UK Web Domain Dataset (1996-2013)*, https://doi.org/10.5259/ukwa.ds.2/1.

JONES C. and HENDERSON D. (2019) Broadband and uneven spatial development: The case of Cardiff City-Region, *Local Economy* **34**, 228-47.
JONES C. I. (2002) Sources of US economic growth in a world of ideas, *American Economic Review* **92**, 220-39.
JORGENSON D. W. (2009) *The Economics of Productivity*. Edward Elgar Publishing, Cheltenham, UK.
JORGENSON D. W., HO M. S. and SAMUELS J. D. (2011) Information technology and US productivity growth: evidence from a prototype industry production account, *Journal of Productivity Analysis* **36**, 159-75.
JUNG J. and LÓPEZ-BAZO E. (2019) On the regional impact of broadband on productivity: the case of Brazil, *Telecommunications Policy*, 101826.
KATZ R., KOUTROUMPIS P. and MARTIN CALLORDA F. (2014) Using a digitization index to measure the economic and social impact of digital agendas, *info* **16**, 32-44.
KLINGER J., MATEOS-GARCIA J. C. and STATHOULOPOULOS K. (2018) Deep learning, deep change? Mapping the development of the Artificial Intelligence General Purpose Technology, *SSRN (August, 2018)*.
KOLKO J. (2012) Broadband and local growth, *Journal of Urban Economics* **71**, 100-13.
KOUTROUMPIS P. (2009) The economic impact of broadband on growth: A simultaneous approach, *Telecommunications Policy* **33**, 471-85.
LEHR B. and LICHTENBERG F. (1999) Information technology and its impact on productivity: Firm-level evidence from government and private data sources, 1977-1993, *Canadian Journal of Economics*, 335-62.
LI Y., ARORA S., YOUTIE J. and SHAPIRA P. (2016) Using web mining to explore Triple Helix influences on growth in small and mid-size firms, *Technovation* **In Press**.
LONGLEY P. A. and SINGLETON A. D. (2009) Classification through consultation: Public views of the geography of the e-society, *International Journal of Geographical Information Science* **23**, 737-63.

LUCAS R. E. (1988) On the mechanics of economic development, *Journal of Monetary Economics* **22**, 3-42.
MACK E. and FAGGIAN A. (2013) Productivity and broadband: The human factor, *International Regional Science Review* **36**, 392-423.
MCCANN P. (2018) Productivity Perspectives Synthesis, *ESRC Productivity Insights Network*.
MEIJERS E. and PERIS A. (2019) Using toponym co-occurrences to measure relationships between places: review, application and evaluation, *International Journal of Urban Sciences* **23**, 246-68.
MILOŠEVIĆ N., DOBROTA M. and BARJAKTAROVIĆ RAKOČEVIĆ S. (2018) Digital economy in Europe: Evaluation of countries' performances, *Zbornik radova Ekonomskog fakulteta u Rijeci: časopis za ekonomsku teoriju i praksu* **36**, 861-80.
MUSSO M. and MERLETTI F. (2016) This is the future: A reconstruction of the UK business web space (1996–2001), *new media & society* **18**, 1120-42.
NAJARZADEH R., RAHIMZADEH F. and REED M. (2014) Does the Internet increase labor productivity? Evidence from a cross-country dynamic panel, *Journal of Policy Modeling* **36**, 986-93.
NEFFKE F., HENNING M. and BOSCHMA R. (2011) How do regions diversify over time? Industry relatedness and the development of new growth paths in regions, *Economic Geography* **87**, 237-65.
OECD (2001) OECD Communications Outlook 2001, Paris.
ONS (2018a) *Annual Population Survey*. Office for National Statistics (NOMIS)
ONS (2018b) *Business Register and Employment Survey*. Office for National Statistics (NOMIS)
ONS (2018c) Internet access – households and individuals, Great Britain: 2018, *Statistical bulletin*. ONS.
ONS (2018d) *Regional Gross Fixed Capital Formation, NUTS1 and NUTS2, 2000 to 2016. Experimental estimates*. Office for National Statistics
ONS (2018e) *Regional Gross Value Added (income approach) reference tables*. Office for National Statistics
ORTEGA-ARGILÉS R. (2013) R&D, knowledge, economic growth and the transatlantic productivity gap, *Handbook of Industry Studies and Economic Geography*, 271.
OXFORD ECONOMICS (2015) The UK's £92 billion digital opportunity. Virgin Media Business.
PALACÍN-SÁNCHEZ M.-J. and DI PIETRO F. (2016) The role of the regional financial sector in the capital structure of small and medium-sized enterprises (SMEs), *Regional Studies* **50**, 1232-47.
PAPAGIANNIDIS S., GEBKA B., GERTNER D. and STAHL F. (2015) Diffusion of web technologies and practices: A longitudinal study, *Technological Forecasting and Social Change* **96**, 308-21.
PAPAGIANNIDIS S., SEE-TO E. W. K., ASSIMAKOPOULOS D. G. and YANG Y. (2018) Identifying industrial clusters with a novel big-data methodology: Are SIC codes (not) fit for purpose in the Internet age?, *Computers & Operations Research* **98**, 355-66.
PRENZEL P., ORTEGA-ARGILÉS R., COZZA C. and PIVA M. (2018) Interplay between regional and industrial aspects in the R&D–productivity link: evidence from Europe, *Regional Studies* **52**, 659-72.
QIANG C. Z.-W., ROSSOTTO C. M. and KIMURA K. (2009) Economic Impacts of Broadband.‖ In Information and Communications for Development 2009: Extending Reach and Increasing Impact, 35–50, *Washington, DC. World Bank*.
REUSCHKE D. and MASON C. (2020) The engagement of home-based businesses in the digital economy, *Futures*, 102542.
RIDDLESDEN D. (2014) Internet User Classication (IUC) User Guide, in (CDRC) U. O. L. C. D. R. C. (Ed).
RIDDLESDEN D. and SINGLETON A. D. (2014) Broadband speed equity: A new digital divide?, *Applied Geography* **52**, 25-33.
ROBERTS M. and SETTERFIELD M. J. H. O. A. T. O. E. G. (2010) Endogenous regional growth: a critical survey, 431-50.
RODRÍGUEZ-POSE A. and KETTERER T. D. (2012) Do local amenities affect the appeal of regions in Europe for migrants?, *Journal of Regional Science* **52**, 535-61.
ROGERS E. M. (1995) *Diffusion of innovations*. Free Press, New York.

ROHMAN I. K. and BOHLIN E. (2012) Does broadband speed really matter as a driver of economic growth? Investigating OECD countries, *International Journal of Management and Network Economics 5* **2**, 336-56.
ROMER P. M. (1986) Increasing Returns and Long-Run Growth, *Journal of Political Economy* **94**, 1002-37.
ROMER P. M. (1990) Endogenous Technological Change, *Journal of Political Economy* **98**, S71-S102.
SENA V. (2020) Measuring Productivity, in MCCANN P. and VORLEY T. (Eds) *Productivity Perspectives*, pp. 48-77. Edward Elgar, Cheltenham, UK.
SHAPIRA P., GÖK A. and SALEHI F. (2016) Graphene enterprise: mapping innovation and business development in a strategic emerging technology, *Journal of Nanoparticle Research* **18**, 269.
SIMMIE J., STERNBERG R. and CARPENTER J. (2014) New technological path creation: evidence from the British and German wind energy industries, *Journal of Evolutionary Economics* **24**, 875-904.
SINGLETON A., DOLEGA L., RIDDLESDEN D. and LONGLEY P. (2016) Measuring the spatial vulnerability of retail centres to online consumption through a framework of e-resilience, *Geoforum* **69**, 5-18.
SINGLETON A. D., RIDDLESDEN D., BLANK G. and GRAHAM M. (2015) Internet User Map Book, in (CDRC) U. O. L. C. D. R. C. (Ed).
SOLOW R. M. (1956) A Contribution to the Theory of Economic Growth, *The Quarterly Journal of Economics* **70**, 65-94.
THE WHAT WORKS CENTRE FOR LOCAL ECONOMIC GROWTH (2015) Evidence Review: Broadband. The What Works Centre for Local Economic Growth, London.
THELWALL M. (2000) Who is using the .co.uk domain? Professional and media adoption of the Web, *International Journal of Information Management* **20**, 441-53.
THELWALL M. and VAUGHAN L. (2004) A fair history of the Web? Examining country balance in the Internet Archive, *Library & information science research* **26**, 162-76.
TRANOS E. (2012) The causal effect of the Internet infrastructure on the economic development of the European city-regions, *Spatial Economic Analysis* **7**, 319-37.
TRANOS E. (2013) *The Geography of the Internet: Cities, Regions and the Internet Infrastructure in Europe*. Edward Elgar, Cheltenham, UK.
TRANOS E. (2020) Social Network Sites and Knowledge Transfer: An Urban Perspective, *Journal of Planning Literature* **35**, 408-22.
TRANOS E. and IOANNIDES Y. M. (2020) ICT and cities revisited, *Telematics and Informatics*, 101439.
TRANOS E. and MACK E. (2016) Broadband provision and knowledge intensive firms: a causal relationship?, *Regional Studies* **50**, 1113-26.
TRANOS E. and NIJKAMP P. (2013) The death of distance revisited: cyber-place, physical and relational proximities, *Journal of Regional Science* **53**, 855-73.
TRANOS E. and STICH C. (2020) Individual internet usage and the availability of online content of local interest: a multilevel approach, *Computers, Environment and Urban Systems* **20**, in press.
WHEELER D. C. and O'KELLY M. E. (1999) Network topology and city accessibility of the commercial Internet, *Professional Geographer* **51**, 327-39.
WOLF M. (2018) The long wait for a productivity resurgence, *Financial Times* **12**.
WOOLDRIDGE J. M. (2010) *Econometric analysis of cross section and panel data*. MIT press, London, England.
ZOOK M. A. (2000) The Web of Production: The Economic Geography of Commercial Internet Content Production in the United States., *Environment and Planning A* **32**, 411-26.
ZOOK M. A. (2001) Old Hierarchies or New Networks of Centrality? – The Global Geography of the Internet Content Market, *American Behavioral Scientist* **44**, 1679-96.

Tables and Figures

Table 1: Second level domain names, 2000

SLD	Frequency	Percentage	Cumulative Frequency	Cumulative Percentage
co.uk	1,035,955	0.668	1,035,955	0.668
ac.uk	335,682	0.217	1,371,637	0.885
gov.uk	92,827	0.06	1,464,464	0.945
org.uk	67,823	0.044	1,532,287	0.988
net.uk	6,965	0.004	1,539,252	0.993
uk	3,467	0.002	1,542,719	0.995
nhs.uk	2,577	0.002	1,545,296	0.997
ltd.uk	1,393	0.001	1,546,689	0.998

Table 2: Number of unique postcodes per .co.uk website, 2000

Level	Frequency	Percentage	Cumulative Frequency	Cumulative Percentage
[1,2]	48,047	0.83	48,047	0.83
(2,10]	6,163	0.11	54,210	0.94
(10,100]	2,975	0.05	57,185	0.99
(100,1000]	646	0.01	57,831	0.99
(1000,10000]	62	0.00	57,893	1
(10000,100000]	4	0	57,897	1

Table 3: Main estimations

	(1) Productivity	(2) Productivity	(3) Productivity
(log).co.uk	-0.0147 (0.0105)		
(log).co.uk2000*2001		0.00943*** (0.00357)	0.00935*** (0.00358)
(log).co.uk2000*2002		0.0111** (0.00437)	0.0111** (0.00434)
(log).co.uk2000*2003		0.0148** (0.00569)	0.0147*** (0.00564)
(log).co.uk2000*2004		0.00897 (0.00712)	0.00895 (0.00699)
(log).co.uk2000*2005		0.0106 (0.00797)	0.0105 (0.00800)
(log).co.uk2000*2006		0.0136* (0.00820)	0.0135* (0.00817)
(log).co.uk2000*2007		0.0252*** (0.00910)	0.0251*** (0.00911)
(log).co.uk2000*2008		0.0261*** (0.00751)	0.0258*** (0.00747)
(log).co.uk2000*2009		0.0221** (0.00970)	0.0218** (0.00975)
(log).co.uk2000*2010		0.0234** (0.00929)	0.0230** (0.00928)
(log).co.uk2000*2011		0.0168* (0.00895)	0.0164* (0.00892)
(log).co.uk2000*2012		0.0213** (0.00854)	0.0209** (0.00851)
(log).co.uk2000*2013		0.0255*** (0.00866)	0.0251*** (0.00872)
(log).co.uk2000*2014		0.0247*** (0.00925)	0.0242*** (0.00934)
(log).co.uk2000*2015		0.0214** (0.00873)	0.0208** (0.00883)
(log).co.uk2000*2016		0.0237** (0.00948)	0.0229** (0.00944)

(log).co.uk 2000			0.144***
			(0.0492)
Productivity 2000			0.676***
			(0.158)
Nr of Universities			0.0275***
			(0.00882)
Distance to London			-0.0207*
			(0.0109)
Manufacturing	0.00250**	0.000644	0.000704
	(0.00122)	(0.00122)	(0.00118)
Diversification	-0.000747	0.00127	0.00130
	(0.00314)	(0.00320)	(0.00320)
Human Capital	0.00188***	0.00112**	0.00108**
	(0.000595)	(0.000554)	(0.000531)
Agglomeration	1.402*	-0.0947	0.0273***
	(0.816)	(0.675)	(0.00753)
Population	-1.465**	-0.0901	-0.191***
	(0.738)	(0.651)	(0.0571)
Employment Density	-0.578***	-0.476***	-0.456***
	(0.0359)	(0.0421)	(0.0402)
Capital	0.0520***	0.0672***	0.0653***
	(0.0150)	(0.0142)	(0.0136)
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Constant	19.35***	11.68***	4.004**
	(4.185)	(3.918)	(1.832)
Observations	1784	2762	2762
N_g	163	163	163
sigma_e	0.0315	0.0346	0.0343
sigma_u	2.020	0.323	0.186
rho	1.000	0.989	0.967

* p<0.10, ** p<0.05, *** p<0.01

Note: Estimation results of the effect of DE (logarithm of commercial content) on regional productivity (logarithm of GVA/Employee, eq. 2). (1) FE specification where DE is the logarithm of commercial websites during 2000-2010. (2) FE specification with interaction term $log(DE_{2000}) * YEAR$ to assess impact of early adoption on regional productivity subject to controls. (3) HT estimator with interaction term $log(DE_{2000}) * YEAR$ and time-invariant controls. Robust standard errors in parentheses.

Figure 1

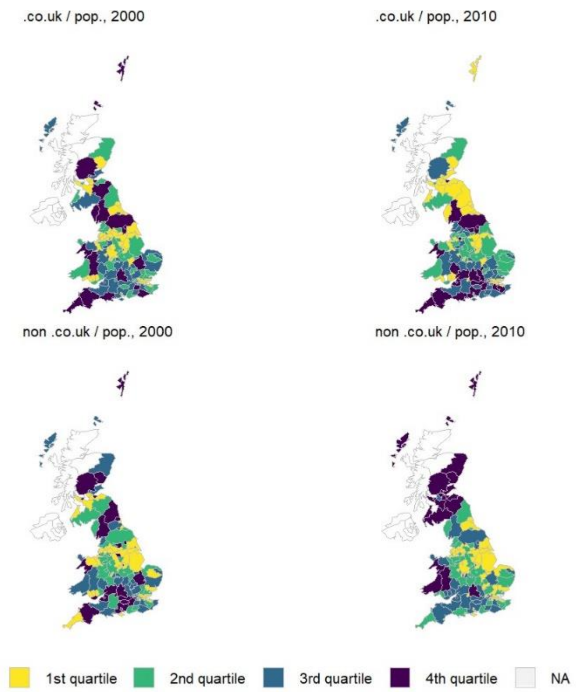
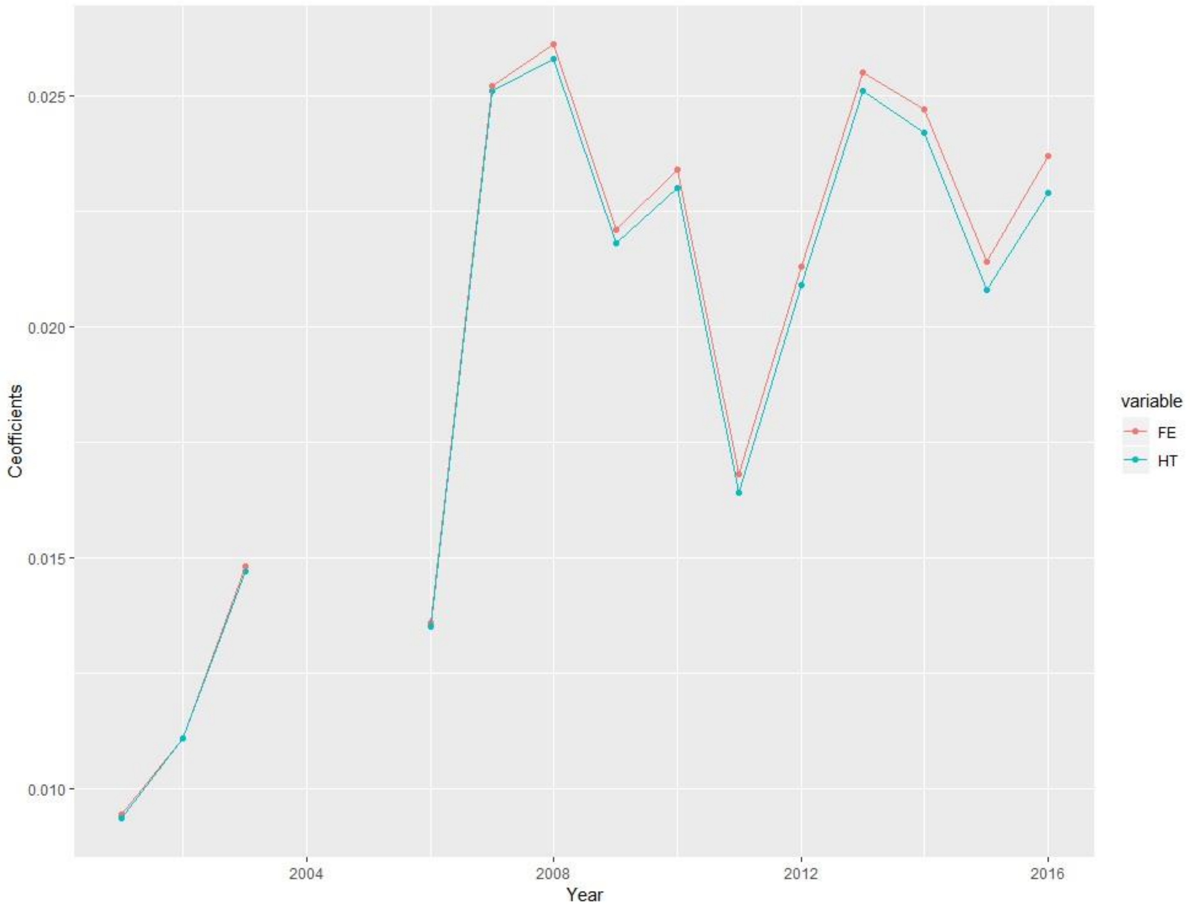
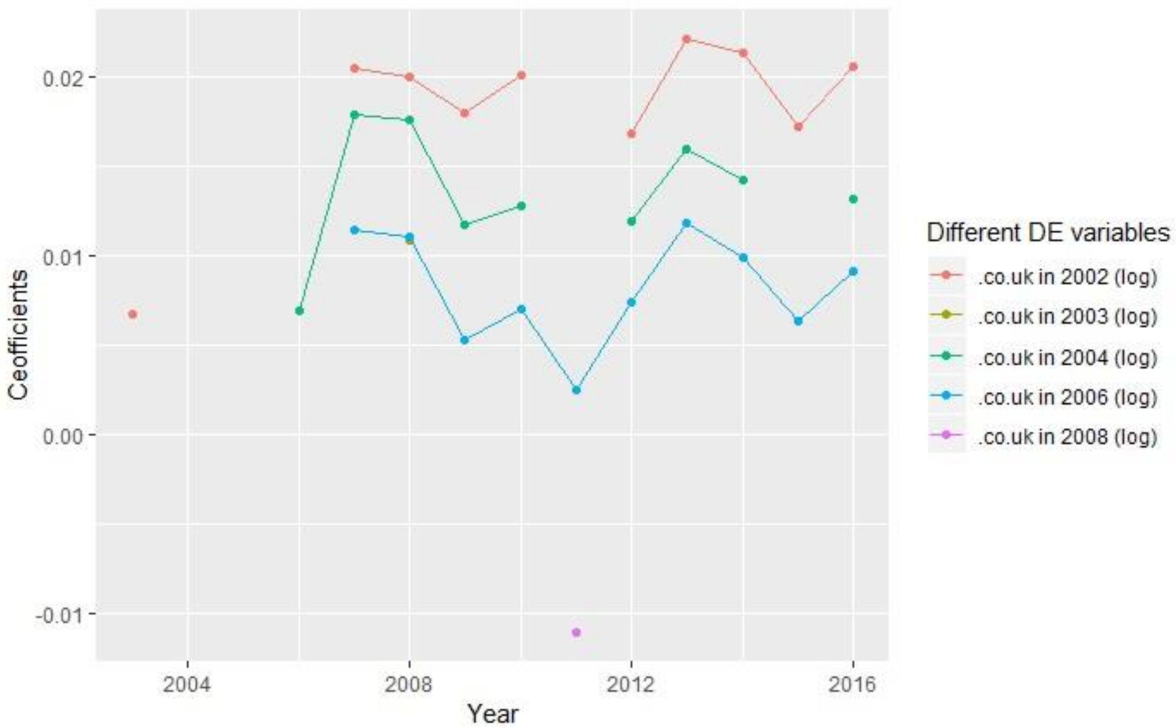


Figure 2



0.025
0.020
0.015
0.010
Ceofficients
2004
2008
2012
2016
Year
variable
FE
HT

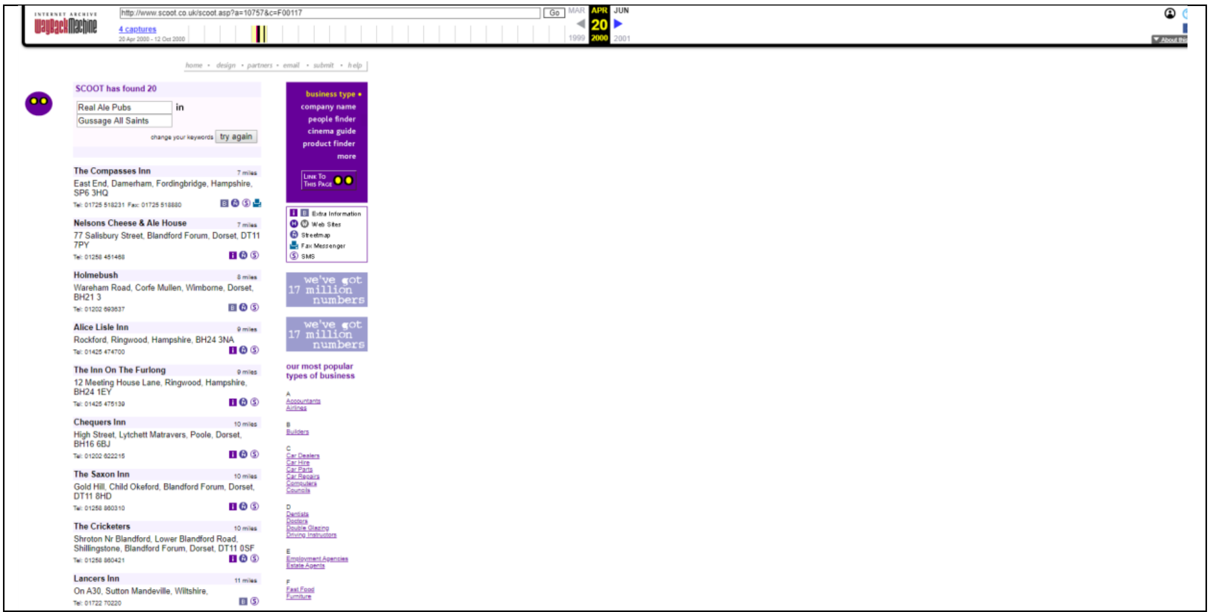
Figure 3



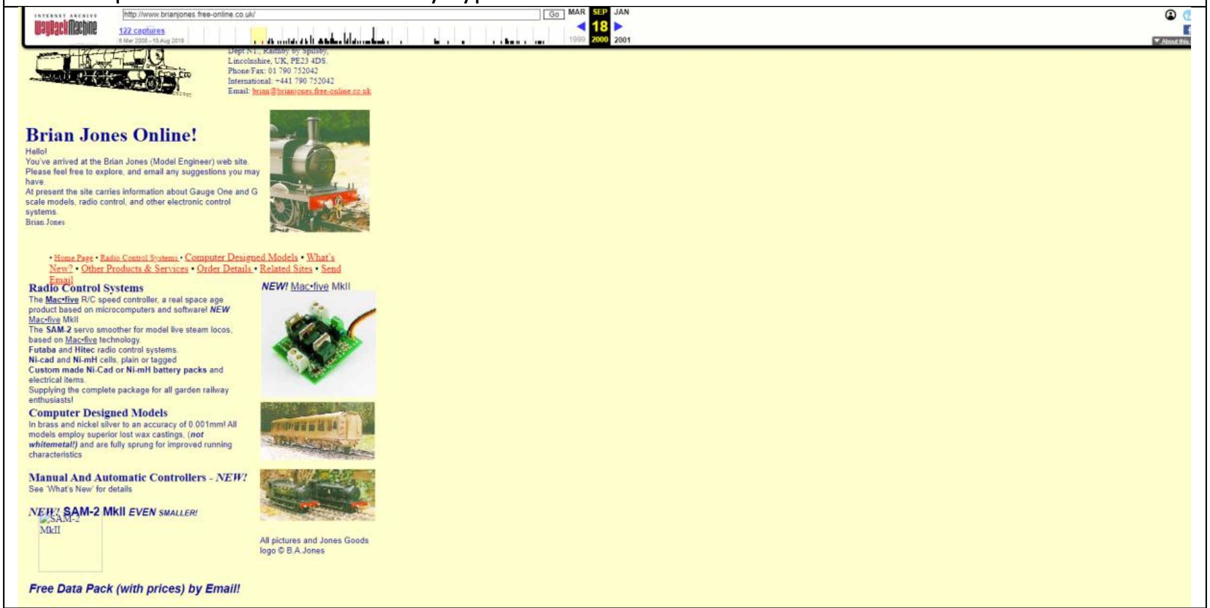
0.02
0.01
0.00
-0.01
Ceofficients
2004
2008
2012
2016
Year
Different DE variables
.co.uk in 2002 (log)
.co.uk in 2003 (log)
.co.uk in 2004 (log)
.co.uk in 2006 (log)
.co.uk in 2008 (log)

Appendix A

Figure A1: Examples of archived webpages



A: Example of an archived directory-type of website



B: Example of a website denoting an activity attached to a unique postcode

Table A1: Descriptive statistics

Label	Variable	Years	Obs.	Mean	Sd	Min	Max	Source	Expected sign
Nr of Universities	N. of universities in NUTS3 area	2019	3024	0.96	1.6	0	15	Own calc.	+
Distance to London	Natural logarithm of distance to London	2019	3024	4.86	1.6	-9.21	6.9	Own calc.	-
Employment Density	Natural logarithm of employed to local population ratio	2000-2016	2771	-0.84	0.29	-1.48	1.2	BRES and ONS population statistics	-
Manufacturing	Employment share of Manufacturing	2000-2016	2771	11.14	5.67	0.49	32.63	BRES	+/-
HHI	Inverse of Herfindahl-Hirschman Index (higher values for diversification)	2000-2016	2771	9.02	1.23	4.64	11.95	BRES	+
Human Capital	Share of population with NVQ4 and above qualification	2000-2016	2925	29.73	9.29	11.54	71.3	APS	+
Population Density	Natural logarithm of population per square kilometre	2000-2016	2856	6.48	1.55	2.15	9.63	ONS	+
Population	Natural logarithm of NUTS3 population	2000-2016	2771	12.66	0.61	9.86	13.98	ONS	-
Capital	Natural logarithm of Gross Fixed Capital Formation in £bn at NUTS2 level	2000-2016	2856	1.8	0.48	0.25	3.02	ONS Experimental statistics	+
.co.uk	Natural logarithm of number of	2000-2010	1848	10.13	1.13	5.82	12.58	IA	+

	.co.uk urls per NUTS3								
.co.uk 2000	Natural logarithm of number of .co.uk urls per NUTS3 in 2000	2000	3024	8.04	0.72	5.82	10.05	IA	+
Productivity	Natural Logarithm of Gross Value Added per employee	2000-2016	2771	10.72	0.21	10.2	11.56	ONS and BRES	
Productivity 2000	Natural Logarithm of Gross Value Added per employee in 2000	2000	2934	10.49	0.15	10.2	11.23	ONS and BRES	

Appendix B

Table B1: Robustness checks – different yearly DE variables

	(1) x=2	(2) x=3	(3) x=4	(4) x=6	(5) x=8
Manufacturing	-0.000412	-0.000102	0.0000790	0.0000480	0.000562
	(0.00140)	(0.00146)	(0.00153)	(0.00148)	(0.00142)
Diversification	0.00202	0.00359	0.00432	0.00503*	0.00699***
	(0.00295)	(0.00292)	(0.00284)	(0.00261)	(0.00270)
Human Capital	0.000568	0.000380	0.000316	-0.000161	-0.000184
	(0.000551)	(0.000552)	(0.000532)	(0.000544)	(0.000533)
Agglomeration	0.0347***	0.0273***	0.0288***	0.0280***	0.0189***
	(0.00885)	(0.00885)	(0.00808)	(0.00828)	(0.00695)
Population	-0.213***	-0.208***	-0.249***	-0.320***	-0.315***
	(0.0538)	(0.0495)	(0.0560)	(0.0715)	(0.0745)
Capital	0.0664***	0.0611***	0.0554***	0.0442***	0.0336**
	(0.0119)	(0.0112)	(0.0103)	(0.0101)	(0.0135)
Employment Density	-0.437***	-0.409***	-0.413***	-0.412***	-0.431***
	(0.0471)	(0.0469)	(0.0499)	(0.0512)	(0.0475)
(log) .co.uk 2000+x*2003	0.00670*				
	(0.00388)				
(log) .co.uk 2000+x *2004	0.00263	-0.00665			
	(0.00584)	(0.00413)			
(log) .co.uk 2000+x *2005	0.00391	-0.00494	0.00201		
	(0.00697)	(0.00656)	(0.00416)		
(log) .co.uk 2000+x *2006	0.0111	-0.000305	0.00694*		
	(0.00718)	(0.00642)	(0.00417)		
(log) .co.uk 2000+x *2007	0.0205***	0.0111	0.0179***	0.0114***	
	(0.00786)	(0.00705)	(0.00514)	(0.00334)	
(log) .co.uk 2000+x *2008	0.0200***	0.0109*	0.0176***	0.0111**	
	(0.00668)	(0.00616)	(0.00526)	(0.00536)	
(log) .co.uk 2000+x *2009	0.0180**	0.00439	0.0117*	0.00528	-0.00719
	(0.00907)	(0.00849)	(0.00642)	(0.00497)	(0.00470)
(log) .co.uk 2000+x *2010	0.0201**	0.00555	0.0128**	0.00703	-0.00655
	(0.00811)	(0.00712)	(0.00580)	(0.00571)	(0.00498)
(log) .co.uk 2000+x *2011	0.0131	0.000706	0.00748	0.00248	-0.0110**
	(0.00796)	(0.00688)	(0.00581)	(0.00677)	(0.00512)
(log) .co.uk 2000+x *2012	0.0168**	0.00488	0.0119*	0.00741	-0.00632
	(0.00763)	(0.00717)	(0.00649)	(0.00741)	(0.00540)
(log) .co.uk 2000+x *2013	0.0221***	0.00921	0.0160**	0.0118	-0.00289
	(0.00809)	(0.00728)	(0.00708)	(0.00821)	(0.00652)
(log) .co.uk 2000+x *2014	0.0213**	0.00746	0.0142**	0.00991	-0.00494
	(0.00860)	(0.00783)	(0.00704)	(0.00774)	(0.00678)
(log) .co.uk 2000+x *2015	0.0172**	0.00341	0.0108	0.00632	-0.00837
	(0.00827)	(0.00737)	(0.00709)	(0.00828)	(0.00663)
(log) .co.uk 2000+x *2016	0.0206**	0.00596	0.0132*	0.00911	-0.00542
	(0.00911)	(0.00790)	(0.00770)	(0.00896)	(0.00749)
Nr of Universities	0.0180*	0.0298***	0.0264***	0.0251***	0.0314***
	(0.01000)	(0.00924)	(0.00846)	(0.00795)	(0.00851)
Distance to London	-0.0287***	-0.00469	-0.00202	0.00622	0.0117

	(0.00995)	(0.0118)	(0.0114)	(0.0106)	(0.0128)
(log) .co.uk 2000+x	0.192***	0.167***	0.197***	0.262***	0.273***
	(0.0518)	(0.0458)	(0.0490)	(0.0601)	(0.0672)
Productivity 2000+x	0.572***	0.768***	0.771***	0.851***	0.926***
	(0.140)	(0.136)	(0.133)	(0.128)	(0.133)
Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Constant	4.762***	2.743*	2.799*	2.168	1.084
	(1.582)	(1.510)	(1.501)	(1.461)	(1.480)
Observations	2442	2282	2119	1793	1467
N_g	163	163	163	163	163
sigma_e	0.0329	0.0325	0.0316	0.0293	0.0274
sigma_u	0.183	0.141	0.167	0.262	0.197
rho	0.969	0.949	0.965	0.988	0.981

Robust standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

HT estimation results of the effect of DE variables for different years (logarithms of commercial content 2002, 2003, 2004, 2006, 2008) on regional productivity (logarithm of GVA/Employee, eq. 2).

Table B2: Robustness checks – different sub-samples

	Productivity difference 2000 – 2000+x				
	x=2	x=3	x=4	x=6	x=8
Manufacturing	-0.00125	-0.00394***	-0.00749***	-0.0144***	-0.0184***
	(0.00118)	(0.00113)	(0.00109)	(0.00112)	(0.00128)
Diversification	0.000931	0.000892	-0.000348	-0.00411	-0.00728**
	(0.00296)	(0.00287)	(0.00280)	(0.00301)	(0.00353)
Human Capital	0.000942*	0.00147***	0.00216***	0.00330***	0.00445***
	(0.000535)	(0.000533)	(0.000534)	(0.000599)	(0.000676)
Agglomeration	0.00902	0.00549	0.0336	0.0287	0.00782
	(0.00777)	(0.00967)	(0.0311)	(0.0321)	(0.0241)
Population	-0.127***	-0.151**	-0.178**	-0.0668	0.0789
	(0.0453)	(0.0613)	(0.0759)	(0.0988)	(0.115)
Capital	0.0718***	0.0837***	0.103***	0.169***	0.232***
	(0.0126)	(0.0122)	(0.0113)	(0.0140)	(0.0201)
Employment	-0.450***	-0.461***	-0.455***	-0.462***	-0.499***
Density	(0.0425)	(0.0439)	(0.0438)	(0.0459)	(0.0519)
(log) .co.uk 2000	-0.0126**	-0.0187***	-0.0234***	-0.0281**	-0.0256**
	(0.00636)	(0.00656)	(0.00753)	(0.0116)	(0.0121)
(log) .co.uk	0.00827**				
2000*2003	(0.00392)				
(log) .co.uk	0.00288	0.00203			
2000*2004	(0.00579)	(0.00488)			
(log) .co.uk	0.00460	0.00491	0.00901		
2000*2005	(0.00705)	(0.00661)	(0.00554)		
(log) .co.uk	0.00970	0.00885	0.0129**		
2000*2006	(0.00711)	(0.00669)	(0.00578)		
(log) .co.uk	0.0210***	0.0211***	0.0249***	0.0251***	
2000*2007	(0.00794)	(0.00750)	(0.00678)	(0.00570)	
(log) .co.uk	0.0209***	0.0207***	0.0232***	0.0199***	
2000*2008	(0.00646)	(0.00604)	(0.00569)	(0.00550)	
(log) .co.uk	0.0181**	0.0165**	0.0206***	0.0208***	0.0172**
2000*2009	(0.00882)	(0.00838)	(0.00763)	(0.00688)	(0.00688)
(log) .co.uk	0.0198**	0.0180**	0.0221***	0.0226***	0.0182**
2000*2010	(0.00804)	(0.00749)	(0.00690)	(0.00702)	(0.00735)
(log) .co.uk	0.0121	0.00978	0.0117*	0.00783	-0.000556
2000*2011	(0.00776)	(0.00719)	(0.00657)	(0.00677)	(0.00704)
(log) .co.uk	0.0164**	0.0145**	0.0163**	0.0115	0.00225
2000*2012	(0.00747)	(0.00725)	(0.00692)	(0.00734)	(0.00766)
(log) .co.uk	0.0212***	0.0190**	0.0210***	0.0168**	0.00758
2000*2013	(0.00784)	(0.00752)	(0.00738)	(0.00793)	(0.00825)
(log) .co.uk	0.0200**	0.0173**	0.0187**	0.0133*	0.00311
2000*2014	(0.00843)	(0.00813)	(0.00775)	(0.00795)	(0.00831)
(log) .co.uk	0.0161**	0.0140*	0.0161**	0.0113	0.00182
2000*2015	(0.00794)	(0.00756)	(0.00727)	(0.00762)	(0.00789)
(log) .co.uk	0.0188**	0.0167**	0.0191**	0.0146*	0.00475
2000*2016	(0.00871)	(0.00821)	(0.00791)	(0.00818)	(0.00823)
d	0.0519***	0.0916***	0.119***	0.138***	0.131***
	(0.00924)	(0.0131)	(0.0200)	(0.0302)	(0.0365)
2003 Difference	-0.00474***				

2000 2000+x	(0.000561)				
2004 Difference	-0.00389***	-0.00582***			
2000 2000+x	(0.000854)	(0.000917)			
2005 Difference	-0.00416***	-0.00638***	-0.00765***		
2000 2000+x	(0.00107)	(0.00130)	(0.00136)		
2006 Difference	-0.00493***	-0.00715***	-0.00859***		
2000 2000+x	(0.00107)	(0.00129)	(0.00139)		
2007 Difference	-0.00672***	-0.00955***	-0.0115***	-0.0123***	
2000 2000+x	(0.00120)	(0.00145)	(0.00166)	(0.00140)	
2008 Difference	-0.00669***	-0.00947***	-0.0111***	-0.0111***	
2000 2000+x	(0.000970)	(0.00118)	(0.00141)	(0.00134)	
2009 Difference	-0.00624***	-0.00865***	-0.0105***	-0.0113***	-0.00960***
2000 2000+x	(0.00133)	(0.00162)	(0.00185)	(0.00169)	(0.00183)
2010 Difference	-0.00651***	-0.00894***	-0.0108***	-0.0117***	-0.00988***
2000 2000+x	(0.00121)	(0.00145)	(0.00168)	(0.00172)	(0.00195)
2011 Difference	-0.00531***	-0.00733***	-0.00830***	-0.00827***	-0.00490***
2000 2000+x	(0.00117)	(0.00140)	(0.00161)	(0.00167)	(0.00186)
2012 Difference	-0.00600***	-0.00825***	-0.00943***	-0.00912***	-0.00564***
2000 2000+x	(0.00115)	(0.00145)	(0.00175)	(0.00183)	(0.00206)
2013 Difference	-0.00673***	-0.00915***	-0.0106***	-0.0104***	-0.00706***
2000 2000+x	(0.00121)	(0.00151)	(0.00187)	(0.00197)	(0.00223)
2014 Difference	-0.00655***	-0.00880***	-0.00999***	-0.00955***	-0.00589***
2000 2000+x	(0.00129)	(0.00160)	(0.00193)	(0.00197)	(0.00224)
2015 Difference	-0.00594***	-0.00816***	-0.00938***	-0.00910***	-0.00555***
2000 2000+x	(0.00121)	(0.00149)	(0.00182)	(0.00188)	(0.00212)
2016 Difference	-0.00636***	-0.00870***	-0.0101***	-0.00986***	-0.00633***
2000 2000+x	(0.00133)	(0.00161)	(0.00196)	(0.00201)	(0.00221)
Nr of Universities	0.0606***	0.0632***	0.0613***	0.0317**	0.00128
	(0.0105)	(0.0124)	(0.0134)	(0.0157)	(0.0178)
Distance to	-0.00770	-0.00660	0.0101	-0.0256	-0.0774***
London	(0.0146)	(0.0166)	(0.0285)	(0.0308)	(0.0261)
Productivity 2000	1.085***	1.076***	1.036***	0.350	-0.408
	(0.177)	(0.210)	(0.263)	(0.303)	(0.330)
Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Constant	0.254	0.741	1.316	7.374***	13.81***
	(1.718)	(1.951)	(2.505)	(2.663)	(2.831)
Observations	5204	5044	4881	4555	4229
N_g	163	163	163	163	163
sigma_e	0.0345	0.0368	0.0393	0.0450	0.0494
sigma_u	0.208	0.405	1.148	1.286	1.083
rho	0.973	0.992	0.999	0.999	0.998

Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

HT estimation results of the effect of early adoption (logarithms of content in 2000) compared to later years (2002, 2003, 2004, 2006 and 2008) on regional productivity (logarithm of GVA/Employee, eq. 2).

With regards to the type of online content, we repeat the analysis using the volume of all – both commercial and non-commercial – online content to capture the broader engagement with the digital economy. To do this we use all the .uk archived webpages. We then focus on the volume of the *local* commercial online content by only considering archived webpages from websites with a unique postcode. By doing this we measure commercial online activities, anchored to specific places. Finally, we consider the non-commercial webpages that is all the SLD names apart from the .co.uk one in order to capture whether the non-commercial online content in 2000 is still associated with long-term productivity effects. The regression tables, replicate the HT estimation from Column (3) in Table 3 can be found in Table B1 and Figure B1 below. Interestingly, these estimates are very similar to the previous ones. All three different measures of archived online content follow an analogous pattern, which mirrors the main results. We can see both the post dot-com dip and the 2008 crisis effect as well as the overall upward trajectory of the coefficients.

Table B3: Robustness checks – different types of DE variables

	(1) Productivity	(2) Productivity	(3) Productivity
Manufacturing	0.000575	0.000903	0.000703
	(0.00118)	(0.00123)	(0.00117)
Diversification	0.000860	0.00105	0.000233
	(0.00322)	(0.00331)	(0.00303)
Human Capital	0.00106**	0.00117**	0.000917*
	(0.000526)	(0.000542)	(0.000517)
Agglomeration	0.0242***	0.0244***	0.00814
	(0.00666)	(0.00695)	(0.00583)
Population	-0.194***	-0.205***	-0.123***
	(0.0508)	(0.0655)	(0.0286)
Capital	0.0638***	0.0670***	0.0579***
	(0.0132)	(0.0138)	(0.0129)
Employment Density	-0.450***	-0.462***	-0.419***
	(0.0397)	(0.0425)	(0.0386)
(log).uk2000*2001	0.0105***		
	(0.00351)		
(log).uk2000*2002	0.0128***		
	(0.00424)		
(log).uk2000*2003	0.0162***		

	(0.00552)	
(log).uk2000*2004	0.0114*	
	(0.00691)	
(log).uk2000*2005	0.0130*	
	(0.00789)	
(log).uk2000*2006	0.0162**	
	(0.00815)	
(log).uk2000*2007	0.0284***	
	(0.00893)	
(log).uk2000*2008	0.0284***	
	(0.00741)	
(log).uk2000*2009	0.0266***	
	(0.00969)	
(log).uk2000*2010	0.0273***	
	(0.00919)	
(log).uk2000*2011	0.0212**	
	(0.00881)	
(log).uk2000*2012	0.0247***	
	(0.00836)	
(log).uk2000*2013	0.0282***	
	(0.00856)	
(log).uk2000*2014	0.0275***	
	(0.00927)	
(log).uk2000*2015	0.0239***	
	(0.00874)	
(log).uk2000*2016	0.0261***	
	(0.00933)	
(log)Local		0.01000***
.co.uk2000*2001		(0.00314)
(log)Local		0.0103***
.co.uk2000*2002		(0.00367)
(log)Local		0.0136***
.co.uk2000*2003		(0.00472)
(log)Local		0.00758
.co.uk 2000*2004		(0.00609)
(log)Local		0.00855
.co.uk 2000*2005		(0.00714)
(log)Local		0.0103
.co.uk 2000*2006		(0.00713)
(log)Local		0.0213***
.co.uk 2000*2007		(0.00789)
(log)Local		0.0212***
.co.uk 2000*2008		(0.00653)
(log)Local		0.0178**
.co.uk 2000*2009		(0.00865)
(log)Local		0.0187**
.co.uk 2000*2010		(0.00828)
(log)Local		0.0138*
.co.uk 2000*2011		(0.00782)
(log)Local		0.0179**
.co.uk 2000*2012		(0.00757)

(log)Local .co.uk 2000*2013		0.0207***	
		(0.00794)	
(log)Local .co.uk 2000*2014		0.0202**	
		(0.00848)	
(log) Local .co.uk 2000*2015		0.0170**	
		(0.00795)	
(log) Local .co.uk 2000*2016		0.0181**	
		(0.00857)	
(log) non .co.uk 2000*2001			0.00835***
			(0.00232)
(log) non .co.uk 2000*2002			0.00769**
			(0.00299)
(log) non .co.uk 2000*2003			0.0109***
			(0.00384)
(log) non .co.uk 2000*2004			0.00791
			(0.00510)
(log) non .co.uk 2000*2005			0.00955*
			(0.00564)
(log) non .co.uk 2000*2006			0.0107*
			(0.00576)
(log) non .co.uk 2000*2007			0.0205***
			(0.00637)
(log) non .co.uk 2000*2008			0.0204***
			(0.00548)
(log) non .co.uk 2000*2009			0.0220***
			(0.00715)
(log) non .co.uk 2000*2010			0.0205***
			(0.00685)
(log) non .co.uk 2000*2011			0.0167**
			(0.00672)
(log) non .co.uk 2000*2012			0.0184***
			(0.00634)
(log) non .co.uk 2000*2013			0.0192***
			(0.00653)
(log) non .co.uk 2000*2014			0.0199***
			(0.00724)
(log) non .co.uk 2000*2015			0.0170**
			(0.00680)
(log) non .co.uk 2000*2016			0.0179**
			(0.00724)
Nr of Universities	0.0211**	0.0323***	0.0254***
	(0.00896)	(0.00767)	(0.00823)
Distance to London	-0.0240**	-0.0179*	-0.0287**
	(0.0110)	(0.00992)	(0.0119)
(log) .uk 2000	0.149***		
	(0.0450)		
Productivity 2000	0.654***	0.582***	0.737***
	(0.154)	(0.153)	(0.141)
(log) Local .co.uk 2000		0.142***	
		(0.0476)	
(log) non .co.uk 2000			0.0603***

			(0.0178)
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Constant	4.241**	5.617***	3.699**
	(1.763)	(2.014)	(1.596)
Observations	2762	2762	2762
N_g	163	163	163
sigma_e	0.0342	0.0344	0.0341
sigma_u	0.155	0.257	0.119
Rho	0.954	0.982	0.924

Robust standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

HT estimation results of the effect of different DE variables (logarithms of all content, local commercial content, non-commercial content) on regional productivity (logarithm of GVA/Employee, eq. 2).

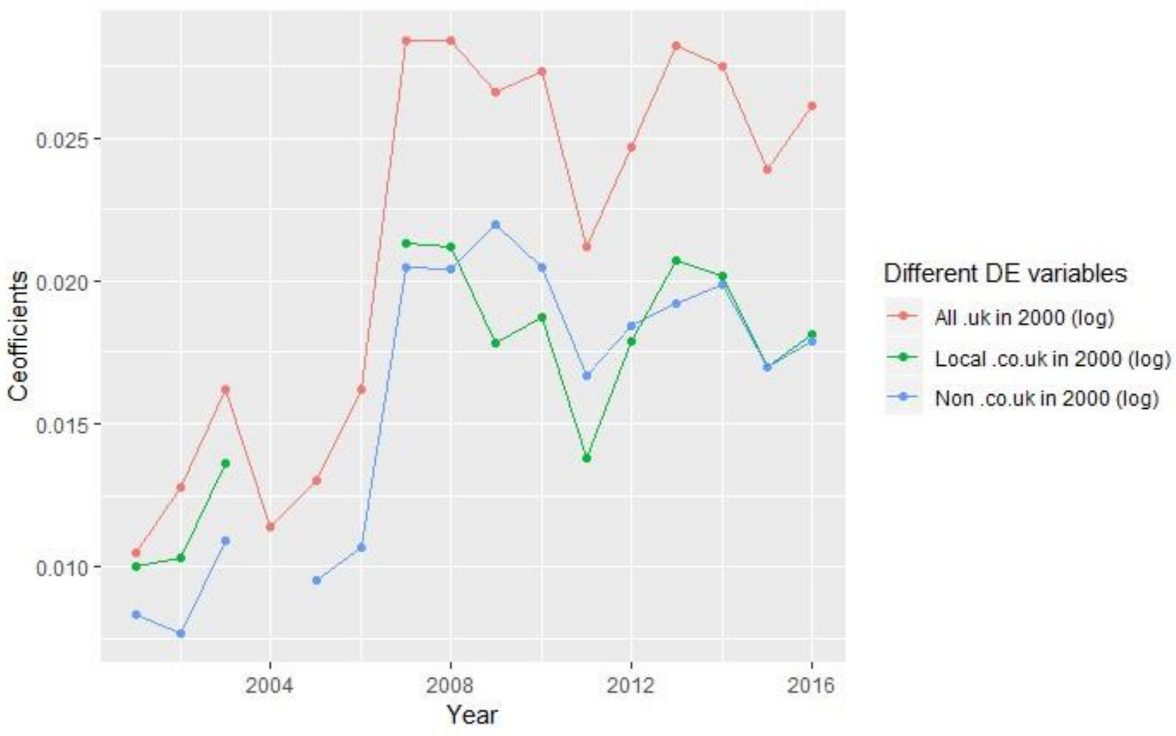


Figure 3: Size and evolution of statistically significant coefficients (p<0.10) for the interaction terms of different DE variables.