ARTICLE TEMPLATE

WFH and broadband speed (title needs rework)

A. N. Author[a], John Smith[b]

[a]Taylor & Francis, 4 Park Square, Milton Park, Abingdon, UK; [b]Institut für Informatik, Albert-Ludwigs-Universität, Freiburg, Germany

ARTICLE HISTORY
Compiled December 18, 2020

ABSTRACT
TBC

KEYWORDS
covid; internet; working from home; broadband speed; time-series clusters

1. Introduction

During the pandemic, working from home using digital technologies, whether partially or exclusively, was transformed from a niche means of accessing work, albeit one that had been on a slow, upward trend, to a widespread way of life in many countries. The ability to work from home or telecommute meant millions retained their jobs and, to a varying extent, maintained productivity during periods of strict lockdowns around the world. However, this ability has not been evenly distributed socially or spatially, creating a new type of digital divide. On one side are those who can work from home, supported by digital technologies, and have thus been able to enjoy both economic resilience and greater personal safety. On the other side, previously employed individuals have been forced to accept furlough or redundancy packages unless they are part of the cadre of essential workers, who are potentially at high risk of infection. Whilst the basis for this new digital divide has been viewed as mainly occupational, here we consider whether the divide is also technological.

Using the UK as a case study, this paper aims to understand how the quality and reliability of internet service, as reflected in *experienced* internet speeds, may reinforce or redress the spatial and social dimensions of the digital division exposed by the pandemic. To do so, we employ volunteered geographic data on individual broadband speed tests and state-of-the-art time-series clustering methods to create clusters of UK local authorities with similar temporal signatures of experienced internet speeds. We then associate these clusters of local authorities with their socioeconomic and geographic characteristics to explore how they overlap with or diverge from the existing economic and digital geography of the UK. Our analysis enables us to better understand how the spatial and social distribution of both occupations and online accessibility intersect to enable or hinder the practice of telecommuting at a time of

CONTACT A. N. Author. Email: latex.helpdesk@tandf.co.uk, John Smith. Email: john.smith@uni-freiburg.de

extreme demand. We will also consider what lessons can be learned from this time for a future where telecommuting is likely to remain a more common means of accessing work, at least in comparison to the pre-Covid era, and broadband services and infrastructure must be fit for purpose. **LET'S LEAVE IT FOR NOW, BUT I THINK WE CAN CRYSTALISE MORE THE RQ**

The capability to work from home has previously been studied from the perspective of whether work tasks in a given occupation both can be and are allowed to be performed using digital technologies independently of location or co-location with colleagues, including supervisors (Allen, Golden, and Shockley 2015; Singh et al. 2013). However, successful telecommuting also requires that the quality and reliability of digital services, particularly home internet connection speeds, enable the completion of work tasks with a minimum of delay or interruption. Prior to the pandemic, the performance of broadband services with respect to telecommuters was never tested at scale, as working from home and connecting to colleagues and workplace resources via the internet was the purview of a small minority of workers. Instead, leisure use in the evening, when video streaming services are at their peak, has been used to benchmark broadband performance and service delivery by different Internet Service Providers (ISPs), at least in the UK (OfCom 2017).

Yet the shift towards telecommuting during various stages of lockdown around the world has been drastic and there are speculations that post-Covid, the tendency to work from home will be much higher than pre-Covid, raising questions around whether internet services can accommodate the increased demand. For example, 47% of people in employment in the UK worked soley from home in April 2020, whilst the same figure only reached 5% the year before (ONS 2020a,b). A back of the envelope calculation suggests that up to 40% of the working force could work from home on an ongoing basis (Batty 2020). Similar figures have been reported for other countries (Felstead and Reuschke 2020). For instance, 37% of the European workforce worked from home in April 2020 with countries like Finland reaching 60% (Eurofound 2020). In the US, almost half of the working population worked from home during the same period because of the pandemic (Brynjolfsson et al. 2020), and a recent estimate indicated that 37% of all jobs in the US can be permanently performed entirely from home (Dingel and Neiman 2020).

None of these changes could have happened in the absence of reliable information and communication technology (ICT) infrastructure – both in terms of software and hardware. But while software innovations are easily diffused across space and society[1], the same does not apply for ICT hardware infrastructure such as internet broadband connectivity. The literature describes first level digital divides in terms of the availability and quality of internet connectivity, such as that manifest in different geographies in the UK (Riddlesden and Singleton 2014; Philip et al. 2017). Second level digital divides consider the presence or lack of the necessary skills to effectively utilise digital technologies and the internet (Blank and Groselj 2014; Van Deursen and Van Dijk 2011). The third level focuses on the heterogenous returns of internet usage among different socioeconomic groups and, consequently, how digital technologies can assist in bridging or further enhancing existing socioeconomic divides. (Stern, Adams, and Elsasser 2009; Van Deursen and Van Dijk 2014; Van Deursen and Helsper 2015). The capability to telecommute is related to all three levels of digital divides, but more importantly leads to differentiated outcomes regarding the economic resilience of people and places to overcome a systemic shock such as the current pandemic. The extreme

[1] See for example the huge success of videoconferencing apps such as Zoom (Marks 2020).

level of demand for telecommuting fundamentally alters the potential returns of internet use for the user and wider community, assumes skills or functions that are present in some occupations but not in others, and relies upon access to high quality internet services.

As the quality of internet infrastructure and services, as well as variation in occupations are spatially dependent and clustered in space, our approach offers a framework for understanding the impact of and interactions between the different levels of digital division in different places with different characteristics. By asking how resilient broadband speeds, and particularly upload speeds are as experienced in different parts of the UK during a time of extreme demand, we interrogate which places benefit from the greater economic resilience digital technologies can offer, not only during the pandemic, but also into the future. The structure of this paper is as follows. First we review the literature on telecommuting and digital divides to better understand the structural and spatial development of these practices pre-pandemic, and thus their importance to the economic resilience of different places. We then describe our data and methodology. Our results section first offers a classification of how internet services vary across the UK local authorities and then assesses whether these clusters replicate or repudiate other socio-economic and geographic patterns of economic resilience.

2. Literature review

2.1. *From telecommuting to #WFH*

In this analysis, the terms 'telecommuting' and 'working from home' are used interchangeably, as most remote labour during the Covid-19 crisis was carried out in the homes of individual employees rather than any other location (Eurofound 2020). However, it should be noted that previous research has explored how telecommuting can occur in other places, including satellite offices or on public transport (Felstead 2012; Siha and Monroe 2006). Previous research has also used a variety of definitions to measure the level of telecommuting within different workforces, distinguishing, for example, between those directly employed, indirectly employed, self-employed, full-time or part-time, and those who use digital technologies to work remotely full-days or part-days (Allen, Golden, and Shockley 2015; Bailey and Kurland 2002; Haddad, Lyons, and Chatterjee 2009). No matter the definition, the option and capability to telecommute or work from home has never been equally distributed spatially or socio-economically any more than different industries and employment opportunities have. For example, studies from the United States, the Netherlands, and the UK found that telecommuters are most likely to hold professional, managerial, and technical occupations where the workforce is better educated and wealthier, and that there is suppressed demand among women and part-time workers (Headicar and Stokes 2016; Peters, Tijdens, and Wetzels 2004; Singh et al. 2013).

Opportunities for working from home during the current pandemic have likewise not been equally spread across the workforce. Dingel and Neiman (2020) indicated that in the US, managers, educators, as well as those working in computer-related occupations, finance, and law can easily work from home, and that occupations with opportunities to telecommute are associated with higher earnings. This is not the case for the workforce occupied in more spatially fixed occupations, from farming, construction and manufacturing to hospitality and care services. In the US, these occupations tend to be lower-income, non-white, without a university degree, live in rental accommodation and lack health insurance (Mongey, Pilossoph, and Weinberg 2020). Similar trends can be observed for other countries. For example, 75% of workers with tertiary education worked from home in Europe during spring 2020, whilst only 34% of workers with secondary education and 14% of those primary education did so (Eurofound 2020).

2.2. *Digital divides and economic resilience*

Our understanding of telecommuting as a product of enabled occupations can be described as a manifestation of the third level digital divide, as those who are able to use digital technologies to work from home benefit from a high rate of return on their use of the internet in terms of autonomy, flexibility, and time saved from commuting (Peters, Tijdens, and Wetzels 2004; Siha and Monroe 2006; Singh et al. 2013). These returns have been even greater during the Covid-19 crisis, when those with the ability to telecommute also have the ability to maintain their employment whilst protecting their health. However, the success of these arrangements has been dependent upon the first level digital divide, which is associated with access and quality of internet connectivity at a time of extremely high demand. Salemink, Strijker, and Bosworth (2017) provides a systematic review of the pre-pandemic, first level digital divide in infrastructure quality between urban and rural areas in various advanced economies. Rural areas, predictably, fare worse. Yet whether this variation in infrastructure quality affects the spatial footprint of telecommuting has not previously been measured, in part because telecommuting has not previously been the cause of greatest demand and pressure on internet services.

There are indications that those who purchase high speed connections consume more data of all sorts and use their connections for a variety of purposes (Hauge, Jamison, and Marcu 2011), and that there is a correlation between access to internet services and a reduction in household transport spend (Bris, Pawlak, and Polak 2017). Whether the implication is more internet use and less travel because of increased telecommuting, these studies suggest that better internet services enable households to make savings and efficiencies, an example of the first level digital divide reinforcing the third level. Yet did the purchase of high speed connections and increased internet access also prepare households for long-term home-working, enforced by government restrictions? The extreme demand during the pandemic provides a new opportunity to understand how infrastructure accessibility, quality, and reliability affects telecommuting, particularly in light of the high volumes of bandwidth-intensive video conferencing required in order to avoid the face-to-face contact that could increase the spread of infection. We seek to answer how internet service resilience contributes to or reduces economic resilience when the latter is dependent upon the capability to work from home. We also aim to improve our understanding of the impact of first level digital division on telecommuting, and whether this results in much more fundamental third level digital division than has previously been perceived.

Furthermore, these multi-layered digital divides intersect with material divides and the economic geography of the UK. Following the regional economic resilience literature, which underlines the differentiated capacity of cities and regions to escape or recover from economic crises (Martin 2012; Kitsos and Bishop 2018), different places have different industrial and occupational profiles, and these affect the aggregated potential capacity of places for telecommuting. Such profiles are associated with long-standing inequalities in the UK and their spatial representation as a North-South divide (Martin 1988). Various studies have illustrated severe inequalities between the north and the south regions of England **replace 'UK'** regarding, for example, skills and human capital, unemployment, productivity and prosperity (Lee 2014; McCann 2020; Dorling 2018). Some scholars have even argued that the UK suffers some of the highest level of interregional inequalities in the global north (Gal and Egeland 2018; McCann 2016). Not only are all three levels of digital divides associated to a certain extent with or shaped by the geography of the UK, but the intersection of the digital and material divides affects the capacity of places to overcome some of the economic effects of the Covid-19 pandemic. Importantly, this is the first time that digital technologies became an essential tool for economic resilience for such a great part of the population.

3. Methods and data

##Time-Series Clustering {#sec:3.1} The starting point of our methodological framework is cluster analysis, which can be defined within the modern machine learning framework as an unsupervised learning task, partitioning unlabelled observations into homogeneous groups known as clusters (Montero, Vilar et al. 2014). The key idea is that observations within clusters tend to be more similar than observations between clusters. Clustering is particularly useful for exploratory studies as it identifies structures within the data (Aghabozorgi, Shirkhorshidi, and Wah 2015). Therefore, cluster analysis is a widely used family of techniques in geography (Gordon 1977; Everitt 1974). For instance, clustering methods are the basis of *geodemographics*, a research domain which aims to create small area indicators or typologies of neighbourhoods based on various and sometimes diverse variables (Singleton and Longley 2009; Harris, Sleight, and Webber 2005). Clustering techniques have also been employed to solve *regionalisation* problems (Niesterowicz, Stepinski, and Jasiewicz 2016).

Common characteristics of these studies are the cross-sectional nature of the data they employ. Indeed, most clustering problems in geography deal with observations that are fixed in time. However, for this paper we are interested in internet speeds, which vary over time and, therefore, create clusters of local authorities in the UK with similar temporal signatures of experienced internet speeds. Hence, we deviate from the established geographical clustering tools and employ time-series clustering methods.

Time-series clustering methods have been developed in order to deal with clustering problems linked to, for instance, stock or other financial data, economic, governmental or medical data as well as machine monitoring (Aggarwal and Reddy 2013; Aggarwal, Hinneburg, and Keim 2001; Hyndman, Wang, and Laptev 2015; Warren Liao 2005). The main challenge – and also the difference with cross-sectional clustering problems – is data dimensionality given the multiplicity of data points for every individual object, local authorities in our case, included in the data set. Time-series are dynamic data as the value of the observations change as a function of time (Aghabozorgi, Shirkhorshidi, and Wah 2015). This high dimensionality leads to (i) computational and algorithmic challenges regarding handling these data and building algorithms to perform clustering over long time-series, and (ii) open questions regarding the choice of similarity measures in order to cluster similar times series objects together considering the whole length of the time-series and overcoming issues around noise, outliers and shifts (Lin et al. 2004; Aghabozorgi, Shirkhorshidi, and Wah 2015).

REMOVE THIS PARAGRAPH - IRRELEVANT DETAIL: Time-series clustering methods utilising the whole length of time-series can be grouped

in three categories. The first – model-based approaches – is based on recovering the underlying model for each time-series and then applying clustering algorithms on the model parameters of each time-series (Aghabozorgi, Shirkhorshidi, and Wah 2015). The main criticism is the cluster accuracy for nearby clusters (Mitsa 2010).The second approach is based the formation of vectors of features derived from the original time-series. These new data of reduced dimensionality are then clustered using conventional clustering algorithms.

For this paper we utilise a category of time-series clustering methods known as shape-based approaches. These methods match two sperate time-series objects based on the similarity of their shapes through the calculation of distances between the shapes, and are thus better equipped to capture similarities between short length time-series (Aghabozorgi, Shirkhorshidi, and Wah 2015). This approach serves best this paper because (i) we identify clusters of UK local authorities with similar temporal signatures – i.e. shapes – of experienced internet speeds and (ii) the length of our time-series is short (see the data discussion in this section).

Another import element of time-series clustering is the actual clustering algorithm. Similar to the clustering of cross-sectional data, we can employ partitioning algorithms, which lead to non-overlapping clusters, hierarchical clustering, which classifies clusters at different levels, and fuzzy algorithms, which create overlapping clusters (Sardá-Espinosa 2019). Because of the simplicity of the implementation and the interpretability of the results, we utilise here partitioning clustering based on the widely used *k-Means* algorithm. This is an iterative algorithm, which begins with defining the desired number of clusters k. Then each observation is randomly assigned to a cluster from the $[1, k]$ space. This initial cluster assignment is followed by iterations in order to minimise the distance between the centroids of the clusters and the observations assigned to these clusters (James et al. 2013).

There are a number of differences between the above described application of *k-Means* for cross-sectional data and its application for times series data. Instead of creating clusters around centroids, a common approach is to create clusters around *medoids*, which are representative time-series objects with a minimal distance to all other cluster objects (Sardá-Espinosa 2019). Also, instead of calculating the Euclidean distance between centroids and data points, more complex distance measures need to be employed in order to capture the similarity between a time-series object and a medoid. A common distance measure for shape-based time-series clustering is Dynamic Time Warping (DTW). Using its underpinning dynamic programming algorithm, DTW compares two time-series objects to find the optimum warping path between them. DTW is widely used in order to overcome limitations linked to the use of Euclidean distance (Sardá-Espinosa 2019; Berndt and Clifford 1994; Ratanamahatana and Keogh 2004). The `R` package `dtwclust` has been used for the time-series clustering (Sarda-Espinosa 2019).

##Experienced Broadband Speeds {#sec:3.2} To assess the quality and reliability of internet across local authorities in the UK during the time when the population were told to work from home if at all possible we utilise unique data comprising individual internet speed tests from Speedchecker Ltd[2]. This is a private company that allows internet users to check their own broadband upload and download speeds, and stores every speed-check with timestamp and geolocation information. These data have been used before to assess digital divides (Riddlesden and Singleton 2014) and the impact

[2] `https://www.broadbandspeedchecker.co.uk/`

of local loop unbundling regulatory processes (Nardotto, Valletti, and Verboven 2015). By using volunteered geographic data, we are able to assess the *experienced* internet speed by users, which may differ from the *advertised* maximum speeds of Internet Service Providers (ISPs).

We are particularly interested in upload speeds and the frequency of speed tests over the period from March to May 2020, as government statements indicate this is when UK workers were told to work from home if at all possible (GOV.UK 2020). Average upload speeds are slower than average download speeds, at 9.3Mb/s mean upload speed for the whole sample during the period of interest, compared to 29.6Mb/s for download speeds, but they are also less associated with internet-based, high-demand, leisure activities such as video streaming. Therefore, upload speeds are more relevant to work-related activities such as uploading documents or two-way audio, video, and text-based communication systems. Meanwhile, the frequency of speed tests was important in identifying the temporal profile which would give us most insight into experienced internet service and resilience, and provide an indication of the volume of experience over particular units of time.

The first step in the workflow after dropping some outliers following Riddlesden and Singleton (2014) was to transform the individual, geolocated and time-stamped tests to more meaningful aggregates both in terms of space and time. The frequency of testing indicates that whilst there is an overall trend of increased testing from March to April and then a slight reduction from April to May, this trend masks substantial variation by not only the day of the week, but also time of day, as can be seen in Figure 1. Thus, a daily aggregation of upload speeds would mask the variation in experienced service over the course of each weekday. Furthermore, the importance of this variation is highlighted by a comparison with the same period in 2019, as in Figure 2, when the volume of testing and thus of experience of internet service quality was much more tied to the evening leisure peak and presumably to download speeds. Since the increase in testing during the working day in 2020 is an indication that users have a greater perception of the variable quality of internet service, particularly during a new morning peak of testing for service reliability, we decided to include a measure of hourly variation in our temporal profiles.

However, although this is a large data set – $241,088$ individual tests performed during weekday hours across the study period – there are not enough observations for each Local Authority District (LAD) and for each working hour of each working day – 631 speed tests per LAD on average – to profile speeds at that level of detail. Therefore, we aggregate all the speed-checks during the 13 weeks of March to May inclusive for weekdays in 2020 by each hour of the day and day of the week. As our research aims to identify the geography of internet service resilience for work purposes, bank holidays and the hours between midnight and 6:00 were excluded, as well as weekend days. The composite week time-series thus comprise 18 hours multiplied by 5 weekdays or 90 time points per series. We also aggregate these data spatially because we could not follow individuals or households and connect data points. The time-series were calculated for each of the 382 LADs in the UK, standardised, and then a k-means partitioning around medoids clustering algorithm was applied using DTW. We initially run the algorithm for $k \in \mathbb{N} \bigcap [5, 15]$ and used cluster validity indices (CVIs) to pick the optimal solution of $k = 13$. Following Sardá-Espinosa (2019) the majority vote for the following CVIs was used: Silhouette (max), Score function (max), Calinski-Harabasz (max), Davies-Bouldin (min), Modified Davies-Bouldin (DB*, min), Dunn (max), COP (min).

In section, we review the temporal profile of upload speed by hour of the day and day of the composite week, as well as the experienced speed characteristics of each

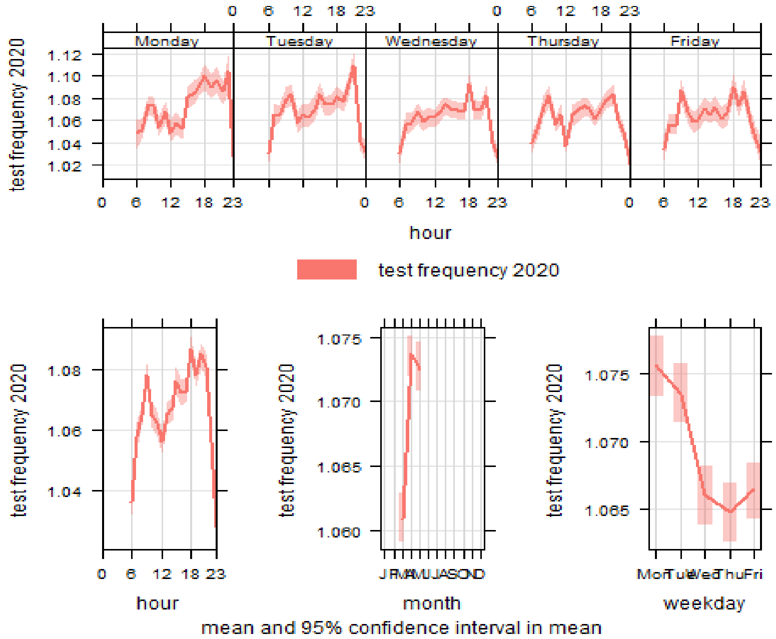

Figure 1. Speed tests over time, 2020

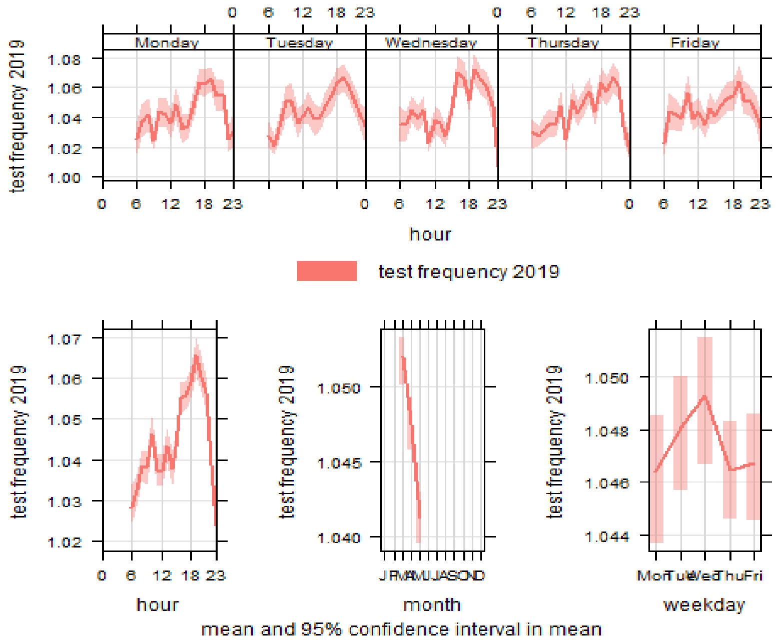

Figure 2. Speed tests over time, 2019

cluster. Since the quality and reliability of internet services vary in time and space due to both supply and demand-side influences, we use a number of different measures to describe experienced upload speeds per cluster. These include: a) mean, experienced connection speed, b) standard deviation or the amount of fluctuation from the mean, and c) the variation in speeds at particular times of day when working from home is more likely to take place. We take account of all three measurements in our descriptives statistics of upload speeds in order to determine how resilient broadband speeds are as experienced in different parts of the UK during a time of extreme demand.

The cause of these different experiences of broadband resilience may be different in different areas, as they may reflect either similarities in patterns of demand or similar quality of infrastructure. Our approach is also limited by potential endogeneity, as for example, better quality connections with high mean speeds may enable more working from home, but greater demand may cause slower speeds, less reliability and greater variability of speed at different times of day or week. Therefore, we avoid attributing any cause to our analysis of the experienced level of quality and reliability of upload speeds. Instead, we run an auxiliary regression in order to understand how the spatial and temporal patterns of internet service relate to the economic geography of the UK. More specifically, we estimate the following multinomial model:

$$Pr(Y_i = j) = \frac{exp^{X_i \beta_j}}{\sum_{i=1}^{j} exp^{X_i \beta_j}} \begin{cases} i = 1, 2, ..., N \\ j = 1, 2, ..., J \end{cases} \tag{1}$$

Based on the outcomes of the time-series clustering, we identify J distinct and crisp clusters. We then regress this cluster membership against a vector X_i of socio-economic and geographic variables, which are discussed in detail in the relevant section, in order to explore how the different patterns might support or undermine efforts to work from home and maintain safe productivity and whether they reinforce existing spatial and social inequalities. This analysis enables us to provide a more nuanced understanding of how telecommuting and technology intersect at a time of extreme demand, and what lessons this time has for a future where telecommuting is likely to remain a common means of accessing work and broadband services, as well as infrastructure, must be fit for purpose.

4. Results

4.1. *Upload Clusters / cluster description*

The temporal profiles of the local authority clusters have been summarised in Figures 3 and 4 and Table 1. The graphs show a composite profile of mean upload speeds per hour per day for each cluster, with the largest, in terms of the LAD membership and population, five clusters in Figure 3, and the next six in Figure 4. These figures and table provide a comprehensive overview of the quality and reliability of experienced broadband in different parts of the UK, the temporal clusters offering a novel approach to understanding spatial disparity.

The second largest cluster, comprising 126 local authorities and over 20 million people, is cluster 13. Cluster 13 has the slowest aggregate mean upload speed of any of the clusters, and the second highest ratio of the standard deviation to the mean. This suggests that those living in local authorities in this cluster experienced some of the

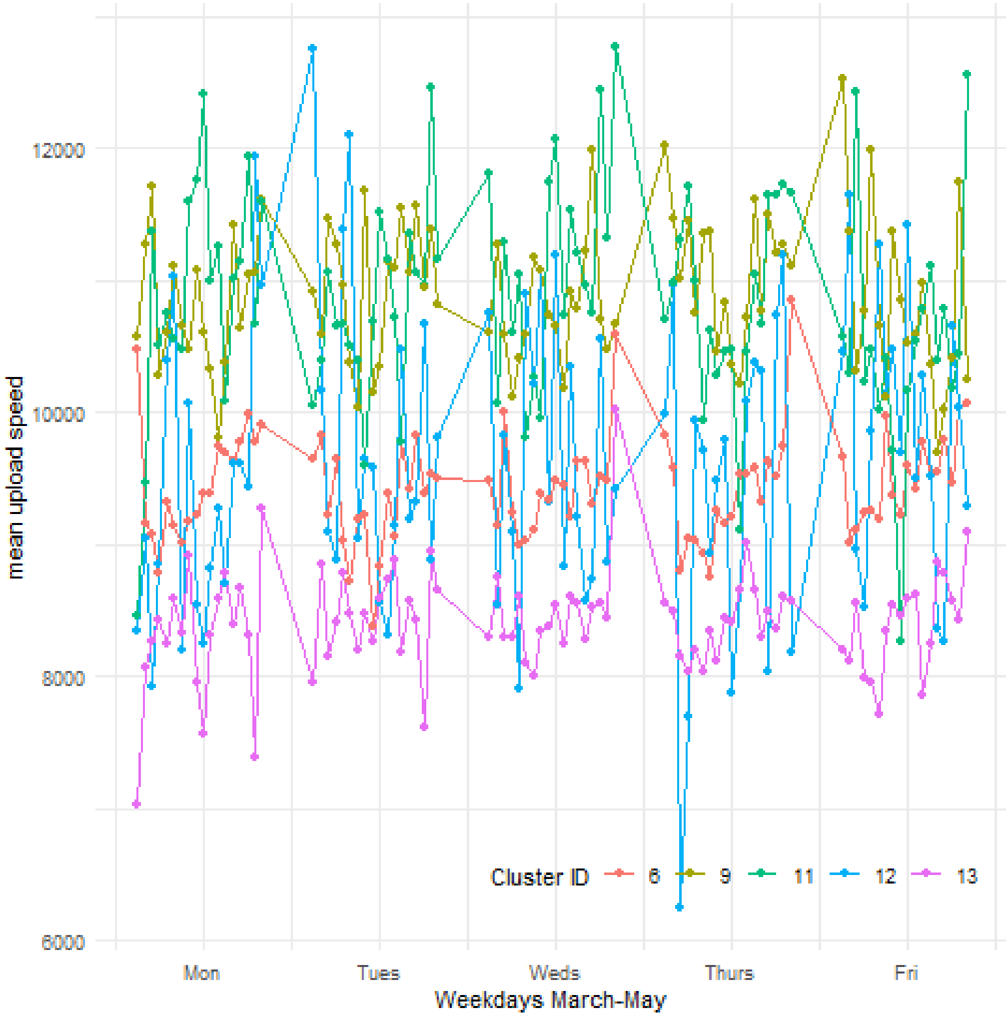

Figure 3. Temporal profilies for upload speed large clusters

Table 1. Upload speed cluster characteristics

Cluster	N. of LADs	LAD population	mean speed	SD speed	mean AM speed	mean PM speed
1	5	343100	8557	6139	7747	9563
2	2	265600	10922	6687	9674	10645
3	4	474700	10201	5658	9470	11236
4	1	91100	9689	6122	7816	9689
5	1	79800	10127	6024	9030	11101
6	155	29535700	9397	5839	9161	9580
7	4	559800	10119	6102	9813	11070
8	5	436300	9429	6254	8682	10434
9	32	6355500	10878	5957	10832	11071
10	4	699600	10795	6005	9258	10697
11	33	5771400	10845	5936	10781	10988
12	10	1544900	9551	6166	9254	9048
13	126	20277700	8392	5849	8299	8522

Note: All speed measures are upload speeds

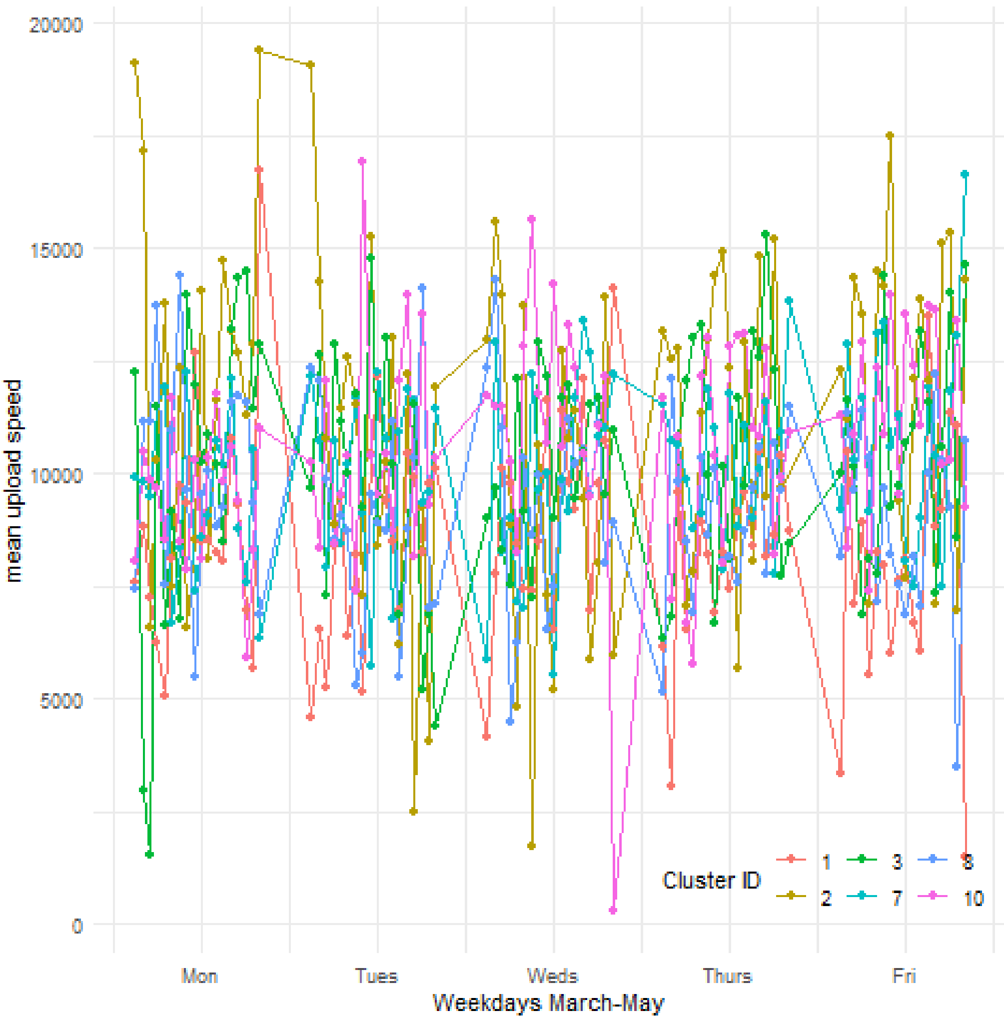

Figure 4. Temporal profilies for upload speed small clusters

lowest quality broadband services in terms of upload speeds in the UK. However, the variation in upload speeds in cluster 13, which can be an indication of its reliability, does not seem to disproportionately affect the morning peak from $9:00$-$10:59$, as upload speeds are, on average, only 2.6% slower than in the evening peak period between $19:00$ and $20:59$, when entertainment purposes are likely to be using the most bandwidth. In comparison, the five LADs that are home to 343 thousand people in cluster 1 not only experience the second slowest mean upload speeds and the highest ratio of standard deviation to the mean, but are also much more affected during the morning peak.

Meanwhile, those living in the largest cluster – 6, with 155 LADs home to 29.5 million people, experience aggregate mean upload speeds of about 1Mb/s faster than those in cluster 13, but still lower than the other three large clusters and most of the smaller clusters, suggesting a middling quality of service. However, the time profile for cluster 6 in Figure 3 shows that upload speeds drop quickly from $6:00$ to $7:00$ on a Monday and peak between $23:00$ and midnight on Wednesday and Thursday, but tend to be lower during the working day. Indeed, experienced mean upload speeds in the morning peak are 4.4% lower than in the evening peak – a greater, more noticeable change than any of the other large clusters experience, but smaller than any of the clusters with temporal profiles shown in Figure 4. Clusters 8 and 12 also have mean upload speeds under 10Mb/s, but higher than clusters 1 and 13, and their standard deviation is not dissimilar. However, this masks great variation in when lower speeds are experienced, with the mean upload speeds much lower between $9:00$-$10:59$ than between $19:00$-$20:59$ in cluster 8, but slightly faster in the morning in cluster 12.

Indeed cluster 12 is the only cluster to experience higher speeds in the evening peak, compared to the morning peak, suggesting that widespread telecommuting has generally changed the temporal profile of internet activity throughout the UK. Yet even if all but one cluster is showing slower speeds in the morning than the evening, the reliability of internet services in different clusters during the working day still varies considerably. Interpreting this variation from the large spikes and dips shown on Figures 3 and 4 is difficult, but the statistics in Table 1 show that clusters 9 and 11 have the most reliable internet services. The ratio of standard deviation to mean in both these clusters is below 55%, and the ratio of upload speeds in the peak periods indicates that speeds are only about 2% slower in the morning. Mean speeds are also higher than in any other cluster, excluding cluster 2, where measures of reliability suggests poorer performance.

Thus, broadband services in clusters 9 and 11, home to over twelve million people performed the best during the study period, in terms of both quality and reliability. In Figure 3, cluster 11 shows more noticeable peaks and troughs, but the lowest points are not at the peak times described in Table 1. Rather, the slowest upload speeds on average occur between 6:00-7:00 on Monday morning, 14:00-15:00 on Friday, and $16:00 - 17:00$ on Thursday. These slowest times are still mostly faster than the average hourly upload speeds in cluster 13. Finally, ignoring the smallest clusters in terms of population, clusters 3, 7 and 10 also have relatively high mean upload speeds. Clusters 3 and 10 pair high mean speeds with low standard deviations relative to the mean speeds, suggesting reliability and resilience, as well as quality broadband services. Cluster 7 has a higher ratio of standard deviation to mean, but there is less difference in average speeds between the morning and evening peaks than in clusters 3 and 10.

In summary, the local authorities in clusters 9 and 11 experienced resilient broadband that could support high levels of telecommuting. Those in clusters 2, 3, 7, and

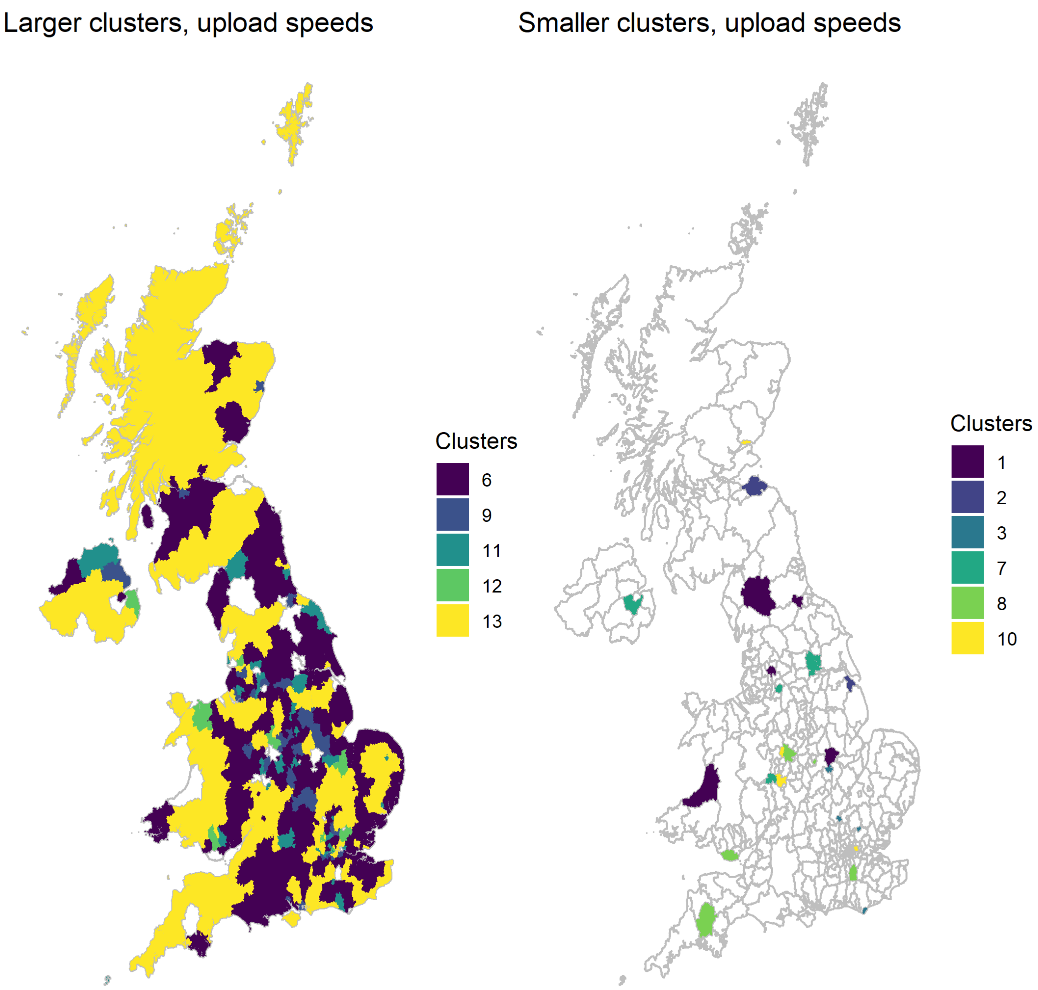

Figure 5. Upload speed clusters for LADs

10 also experience higher than average mean speeds and rank high to middle on measures of service realiability. These LADs are *not* on the wrong side of the first level digital divide, but how likely are they to be able to take advantage of their resilient ICT infrastructure and services? Meanwhile, cluster 6 is not only the largest in terms of number of LADs and population, it has the closest mean upload speed to the pre-clustered average for the whole sample. As well as average quality internet services, those in cluster 6 also experience average reliability for work purposes, ranking fifth behind the four other clusters with populations over one million, but ahead of the smaller clusters. Clusters 8 and 12 are also close to average mean upload speeds, but show very different patterns in terms of reliability, whilst clusters 1 and 13 appear to suffer most from a lack of quality internet services, with slow speeds and high standard deviations. With those in cluster 1 in particular more likely to experience that poor reliability during the morning peak, is this first level digital divide occurring in areas where few are occupationally able to telecommute anyway, and what are the implications for economic resilience?

4.2. *aux regressions*

Using auxiliary regressions, we test whether the clusters that have higher mean speeds and more reliable services are more urban, more wealthy, and / or more likely to benefit from a choice of high quality internet services. We also measure which of our clusters are more likely to have a higher proportion of occupations which enable telecommuting because of the nature of the work. The results of these auxiliary regressions are presented in Table 2. The dependent variable is the LAD cluster membership as described in the Methods and data section and equation 1 and each column represent a different cluster. The reference case is cluster 4, which includes only the local authority of Hambleton in North Yorkshire, a rural area of just over ninety thousand people. Mean, experienced upload speeds in cluster 4 are close to both the average speeds for the 13 clusters (9.9Mb/s) and the pre-clustered average for the whole sample (9.3Mb/s) as shown in 1. However, the standard deviation for cluster 4 and the difference between average speeds in the morning compared to the evening peak periods are indications of worse reliability than many of the other clusters. Hence, the results in Table 2 should be seens as relative rather than absolute probabilities.

First, the number of speed tests run per cluster inhabitant between 9:00-10:59 is an indication of satisfaction or at least a lack of concern over broadband speeds and quality of service. People in clusters 6 and 13 ran the most speed tests (at a per capita basis) at this time of day, whilst people in cluster 11 ran the fewest. This high level of satisfaction matches the high quality and reliability of service we would expect from the analysis in section. Those in clusters 2, 3, 9 and 10 also ran fewer speed tests per capita than those in the other six multi-authority clusters. Cluster 7 was the only cluster with higher mean speeds and middle to high measures of reliability which was more likely to run more speed tests per capita. This might be a reflection of the choice of ISP and connection options in those LADs, as, along with cluster 12, those in cluster 7 were least likely to have Virgin Media connections. Fast Virgin Media cable connections have historically only been available to 45% of premises in the UK (OfCom 2016), where the more lucrative and competative market originally attracted the cable TV provider. Those in clusters 2, 3, 9 and 11 benefit from a higher proportion of Virgin connection, which is an indication that people in these clusters are more likely to live in urban areas, with more choice of broadband services, or in other words, are more

Table 2. Auxiliary multinomial regression of upload speed clusters on socio-economic and geographic LAD variables

	1	2	3	6	7	8	9	10	11	12	13
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
pop, 2018	−0.00004***	0.00002*	0.00001	0.00002***	0.00002***	0.00000	0.00002***	0.00002***	0.00002***	0.00002**	0.00002***
	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
job density, 2018	−0.536***	−1.834***	−0.132***	−0.925***	−1.208***	−0.299***	−1.746***	−1.436***	3.350***	3.400***	0.630***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
distance to nearest met. area	−0.034***	−0.014***	0.002***	−0.020***	−0.074***	−0.044***	−0.013***	−0.036***	−0.031***	−0.036***	−0.024***
	(0.0005)	(0.001)	(0.0002)	(0.002)	(0.0001)	(0.0002)	(0.002)	(0.0005)	(0.0003)	(0.0002)	(0.002)
distance to London	0.007***	0.002	−0.016***	0.001	0.004***	0.004***	−0.002*	0.005**	−0.002	0.003	0.006***
	(0.001)	(0.002)	(0.0004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
South of the UK	−0.410***	−1.451***	−0.039***	−0.111***	−0.048***	−0.841***	−0.798***	1.492***	0.610***	0.798***	2.403***
	(0.00000)	(0.00000)	(0.00000)	(0.00001)	(0.00000)	(0.00000)	(0.00001)	(0.00000)	(0.00001)	(0.00001)	(0.00001)
managerial jobs, 2020	0.939***	0.704***	0.435***	0.704***	0.316***	0.786***	0.576***	0.311***	0.476***	0.594***	0.615***
	(0.00004)	(0.0001)	(0.00004)	(0.00004)	(0.00002)	(0.0001)	(0.00003)	(0.00003)	(0.00002)	(0.00004)	(0.00003)
tech jobs, 2020	0.096***	−0.257***	−0.071***	−0.111***	−0.206***	0.199***	−0.126***	−0.606***	−0.180***	−0.398***	−0.112***
	(0.00004)	(0.00004)	(0.00004)	(0.00003)	(0.00003)	(0.0001)	(0.00003)	(0.00003)	(0.00002)	(0.00004)	(0.00003)
skilled trade jobs, 2020	0.651***	0.160***	−0.191***	0.236***	0.604***	−0.184***	0.205***	0.597***	0.108***	−0.022***	0.295***
	(0.00004)	(0.00004)	(0.00003)	(0.00003)	(0.00003)	(0.00004)	(0.00002)	(0.00003)	(0.00003)	(0.00004)	(0.00003)
professional jobs, 2020	−0.118***	−0.234***	−0.121***	−0.172***	−0.514***	−0.172***	−0.349***	−0.351***	−0.245***	−0.344***	−0.229***
	(0.00005)	(0.0001)	(0.0001)	(0.00005)	(0.00003)	(0.0001)	(0.00004)	(0.0001)	(0.00003)	(0.00005)	(0.0001)
administrative jobs, 2020	0.019***	−0.836***	−0.040***	−0.117***	−0.139***	0.206***	−0.058***	−0.200***	−0.055***	−0.168***	−0.177***
	(0.00003)	(0.00002)	(0.00004)	(0.00001)	(0.00003)	(0.00003)	(0.00001)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
leisure jobs, 2020	−0.198***	−0.180***	−0.225***	−0.476***	−0.654***	−0.820***	−0.537***	−0.935***	−0.353***	−0.625***	−0.491***
	(0.00002)	(0.00004)	(0.00004)	(0.00002)	(0.00002)	(0.00003)	(0.00002)	(0.00001)	(0.00002)	(0.00003)	(0.00002)
machine operation jobs, 2020	−0.336***	0.207***	0.392***	0.010***	−0.433***	0.139***	−0.099***	−0.139***	−0.144***	0.098***	−0.179***
	(0.00002)	(0.00003)	(0.00003)	(0.00002)	(0.00001)	(0.00003)	(0.00002)	(0.00001)	(0.00001)	(0.00002)	(0.00001)
earnings, 2019	−0.003*	0.010***	0.012***	0.020***	0.027***	0.001	0.020***	0.016***	0.015***	0.025***	0.014***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
n. business est. per hab., 2019	0.126***	−0.120***	−0.094***	−0.133***	0.123***	−0.051***	−0.334***	−0.133***	−0.150***	0.289***	0.377***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
NVQ4+	−0.141***	0.064***	−0.091***	−0.070***	−0.010***	0.004***	0.016***	0.170***	−0.110***	−0.038***	−0.035***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
AM tests per hab., 2020	0.0005***	−0.002***	−0.005***	0.010***	0.0004***	−0.001***	−0.002***	−0.005***	−0.013***	−0.001***	0.016***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Virgin Media %, 2020	−0.044***	1.578***	1.210***	0.248***	−1.724***	−0.242***	3.109***	−0.085***	1.214***	−3.889***	−0.745***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Constant	0.321***	−0.436***	0.199***	−2.953***	0.278***	0.002***	−0.866***	0.788***	2.600***	0.017***	0.022***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
McFadden's R squared	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338
N	323	323	323	323	323	323	323	323	323	323	323
Akaike Inf. Crit.	1,148.027	1,148.027	1,148.027	1,148.027	1,148.027	1,148.027	1,148.027	1,148.027	1,148.027	1,148.027	1,148.027

Note: *p<0.1; **p<0.05; ***p<0.01

likely to be on the right side of the first level, infrastructure-based digital divide.

We measure distance from London and from London or another metropolitan areas as two more indicators of rurality and potentially first level digital divides. The broadband speed tests run in the authorities in cluster 3 are more likely to be taking place close to London than those run in any of the other clusters, a result that makes sense considering two of the four authorities in cluster 3 are the London commuter towns of Harlow and Luton. However, even though London was also included in the variable calculating distance from the centre of one of either the ten largest metropolitan areas in England, Glasgow or Cardiff, tests run in cluster 3 are likely to be furthest away. Still, whilst significant, the effect is tiny and the other members of cluster 3 are Corby and Eastbourne, two towns which, whilst not part of metropolitan areas, are also not in remote parts of the country, and are home to an estimated population of 175 thousand people between them. **move maps to here?** Meanwhile, LADs in cluster 7 are most likely to be near the centre of a large metropolitan area, even though the four local authorities of cluster 7 include no central urban boroughs, and only one LAD that is part of a metropolitan area of governance - Tameside in Greater Manchester. This may be why those in cluster 7 are not served by Virgin Media. It also is a demonstration of the complexity of experienced broadband upload speeds as captured by time-series clustering, and their likely interaction with first level digital divides. Cluster 1, which our analysis suggests lacks broadband resilience, contains five, mainly rural authorities. However, they are scattered around the country – Ceredigion in West Wales, Darlington in the Northeast, Eden between the Lake District and the North Penines, Rossendale in the Northwest, and Rutland in the East Midlands – and therefore, the results in Table 2 indicate that these authorities are not as far from the centre of a metropolitan area as authorities in cluster 3, but furthest from London.

Internet resilience as measured here does not seem to follow the North-South economic divide either, as broadly defined in our regression as being part of the Greater London, Southeast, Southwest and East of England regions. The authorities most likely to be in the South are those in cluster 13, which was identified as having the slowest mean upload speeds of any of the clusters, and thus a low level of service. However, cluster 13 does include some rural, remote areas of the country, such as Northwest Scotland, Cornwall and Powys in Wales as shown on Figure 5 It also includes the major metropolitan centres of Bristol, Liverpool, and Newcastle, as well as nine (of 32) London Boroughs and plenty of home county and suburban areas. Although by the standard deviation measure cluster 13 has unreliable internet services, speed variation was less during the morning peak.

5. Conclusions

Upload speeds are not used as benchmarks in the UK, nor have been seen as integral to universal service, considering there has never before been such extreme demand for telecommuting and operations such as video calls.

those in cluster 6, including people living in Birmingham, Leeds and Sheffield, in twelve London Boroughs, and many suburban areas and smaller cities like Cardiff, Oxford and Cambridge This lower level of reliability may be as a result of increased demand, e.g. for telecommuting, especially as there are few truly remote areas in this cluster.

The eight LADs in clusters 3 and 10 are home to over one million people, and are all more urban and suburban than cluster 1, including the London Borough of Newham, the Southeast commuter towns of Harlow and Luton, the Birmingham suburbs of Bromsgrove and Cannock Chase, Dundee City, and the towns of Eastbourne and Corby.

Although not all of the local authorities in these clusters 9 and 11 are urban, among the 65 LADs are Manchester and three of the nine other boroughs of Greater Manchester, Glasgow, Nottingham, both Portsmouth and Southampton, eight London Boroughs, four of the seven constituent authorities of the West Midlands Combined Authority. There are also a number of other tightly bounded urban areas, such as Aberdeen, Blackpool, Ipswich, Norwich, Slough, and Stevenage, and urban areas at the centre of less confined districts like Burnley, Milton Keynes, and Northampton.

Clusters 3, 9, 10 and 11 seem to benefit most from high quality and resilient broadband services. The dips in mean upload speeds in clusters 6, 3 and 10 during the morning peak are suggestive of more use during the working day, and potentially more telecommuting.

The long-term effects of such drastic changes in telecommuting and attitudes towards working from home are difficult to predict. Nevertheless, they span through various aspects of economy and society: from changes to transportation planning due to altered commuting patters, to changes in land use and urban planning to accommodate people who work from home (Budnitz, Tranos, and Chapman 2020)**also 2020 Swedish article from JTG**; and from productivity and innovation changes, to changes in agglomeration externalities and the attraction of large cities (Nathan and Overman 2020) just to name a few. This paper is positioned to support endeavours in understanding the effects of increased telecommuting by exposing the spatial and social dimensions of telecommuting including the resilience of broadband speeds in terms of both quality and reliability of service, and whether this reinforces or redresses prior digital divisions. **took this bit you wrote to put in the discussion at the end?**

Acknowledgement(s)

An unnumbered section, e.g. `\section*{Acknowledgements}`, may be used for thanks, etc. if required and included *in the non-anonymous version* before any Notes or References.

Funding

An unnumbered section, e.g. `\section*{Funding}`, may be used for grant details, etc. if required and included *in the non-anonymous version* before any Notes or References.

References

Aggarwal, CC, and CK Reddy. 2013. "Time-series data clustering." *Data Clustering: Algorithms and Applications* 1.

Aggarwal, Charu C, Alexander Hinneburg, and Daniel A Keim. 2001. "On the surprising behavior of distance metrics in high dimensional space." In *International conference on database theory*, 420–434. Springer.

Aghabozorgi, Saeed, Ali Seyed Shirkhorshidi, and Teh Ying Wah. 2015. "Time-series clustering–a decade review." *Information Systems* 53: 16–38.

Allen, Tammy D, Timothy D Golden, and Kristen M Shockley. 2015. "How effective is telecommuting? Assessing the status of our scientific findings." *Psychological Science in the Public Interest* 16 (2): 40–68.

Bailey, Diane E, and Nancy B Kurland. 2002. "A review of telework research: Findings, new directions, and lessons for the study of modern work." *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior* 23 (4): 383–400.

Batty, Michael. 2020. "The Coronavirus crisis: What will the post-pandemic city look like?" *Environment and Planning B: Urban Analytics and City Science* 47 (4): 547–552. https://doi.org/10.1177/2399808320926912.

Berndt, Donald J, and James Clifford. 1994. "Using dynamic time warping to find patterns in time series." In *KDD workshop*, Vol. 10, 359–370. Seattle, WA, USA:.

Blank, Grant, and Darja Groselj. 2014. "Dimensions of Internet use: amount, variety, and types." *Information, Communication & Society* 17 (4): 417–435.

Bris, Myriam, Jacek Pawlak, and John W Polak. 2017. "How is ICT use linked to household transport expenditure? A cross-national macro analysis of the influence of home broadband access." *Journal of Transport Geography* 60: 231–242.

Brynjolfsson, Erik, John J Horton, Adam Ozimek, Daniel Rock, Garima Sharma, and Hong-Yi TuYe. 2020. *COVID-19 and remote work: an early look at US data*. Technical Report. National Bureau of Economic Research.

Budnitz, Hannah, Emmanouil Tranos, and Lee Chapman. 2020. "Telecommuting and other trips: an English case study." *Journal of Transport Geography* 85: 102713. http://www.sciencedirect.com/science/article/pii/S0966692319305472.

Dingel, Jonathan I, and Brent Neiman. 2020. *How Many Jobs Can be Done at Home?* Working Paper 26948. National Bureau of Economic Research. http://www.nber.org/papers/w26948.

Dorling, Danny. 2018. *Peak Inequality: Britain's ticking time bomb*. Policy Press.

Eurofound. 2020. "Living, working and COVID-19." *COVID-19 series* https://www.eurofound.europa.eu/publications/report/2020/living-working-and-covid-19.

Everitt, B. 1974. "Cluster Analysis Heinemann Educational London." .

Felstead, Alan. 2012. "Rapid change or slow evolution? Changing places of work and their consequences in the UK." *Journal of Transport Geography* 21: 31–38.

Felstead, Alan, and Darja Reuschke. 2020. "Homeworking in the UK: before and during the 2020 lockdown." .

Gal, Peter, and Jagoda Egeland. 2018. "Reducing regional disparities in productivity in the United Kingdom." .

Gordon, AD. 1977. "Classification: methods for the exploratory analysis of multivariate data." Wiley.

GOV.UK. 2020. "News and Communications." https://www.gov.uk/search/news-and-communications?level_one_taxon=5b7b9532-a775-4bd2-a3aa-6ce380184b6c&order=updated-oldest.

Haddad, Hebba, Glenn Lyons, and Kiron Chatterjee. 2009. "An examination of determinants influencing the desire for and frequency of part-day and whole-day homeworking." *Journal of Transport Geography* 17 (2): 124–133.

Harris, Richard, Peter Sleight, and Richard Webber. 2005. *Geodemographics, GIS and neighbourhood targeting*. Vol. 8. John Wiley & Sons.

Hauge, Janice, Mark Jamison, and Mircea Marcu. 2011. "Consumer Usage of Broadband Internet Services: an analysis of the case of Portugal." In *Adoption, Usage, and Global Impact of Broadband Technologies: Diffusion, Practice and Policy*, 198–213. IGI Global.

Headicar, P, and G Stokes. 2016. *On the Move 2: Making sense of travel trends in England 1995–2014*. Technical Report. Technical Report. Independent Transport Commission. http://www. theitc. org . . . .

Hyndman, Rob J, Earo Wang, and Nikolay Laptev. 2015. "Large-scale unusual time series detection." In *2015 IEEE international conference on data mining workshop (ICDMW)*, 1616–1619. IEEE.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An introduction to statistical learning.* Vol. 112. Springer.

Kitsos, Anastasios, and Paul Bishop. 2018. "Economic resilience in Great Britain: the crisis impact and its determining factors for local authority districts." *The Annals of Regional Science* 60 (2): 329–347.

Lee, Neil. 2014. "Grim down South? The determinants of unemployment increases in British cities in the 2008–2009 recession." *Regional Studies* 48 (11): 1761–1778.

Lin, Jessica, Michail Vlachos, Eamonn Keogh, and Dimitrios Gunopulos. 2004. "Iterative incremental clustering of time series." In *International Conference on Extending Database Technology,* 106–122. Springer.

Marks, Paul. 2020. "News: Virtual collaboration in the age of the coronavirus." *Communications of the ACM* 63 (9): 21–23.

Martin, Ron. 1988. "The Political Economy of Britain's North-South Divide." *Transactions of the Institute of British Geographers* 13 (4): 389–418. http://www.jstor.org/stable/622738.

Martin, Ron. 2012. "Regional economic resilience, hysteresis and recessionary shocks." *Journal of economic geography* 12 (1): 1–32.

McCann, Philip. 2016. *The UK regional-national economic problem: Geography, globalisation and governance.* Routledge.

McCann, Philip. 2020. "Perceptions of regional inequality and the geography of discontent: Insights from the UK." *Regional Studies* 54 (2): 256–267.

Mitsa, Theophano. 2010. *Temporal data mining.* Boca Raton, FL: CRC Press.

Mongey, Simon, Laura Pilossoph, and Alex Weinberg. 2020. *Which Workers Bear the Burden of Social Distancing Policies?* Working Paper 27085. National Bureau of Economic Research. http://www.nber.org/papers/w27085.

Montero, Pablo, José A Vilar, et al. 2014. "TSclust: An R package for time series clustering." *Journal of Statistical Software* 62 (1): 1–43.

Nardotto, Mattia, Tommaso Valletti, and Frank Verboven. 2015. "Unbundling the incumbent: Evidence from UK broadband." *Journal of the European Economic Association* 13 (2): 330–362.

Nathan, Max, and Henry Overman. 2020. "Will coronavirus cause a big city exodus?" https://www.coronavirusandtheeconomy.com/question/will-coronavirus-cause-big-city-exodus.

Niesterowicz, J., T.F. Stepinski, and J. Jasiewicz. 2016. "Unsupervised regionalization of the United States into landscape pattern types." *International Journal of Geographical Information Science* 30 (7): 1450–1468. https://doi.org/10.1080/13658816.2015.1134796.

OfCom. 2016. "Connected Nations Report 2016." https://www.ofcom.org.uk/__data/assets/pdf_file/0035/95876/CN-Report-2016.pdf.

OfCom. 2017. "UK Home Broadband Performance: The performance of fixed-line broadband delivered to UK residential consumers. Research Report." https://www.ofcom.org.uk/__data/assets/pdf_file/0015/100761/UK-home-broadband-performance,-November-2016-Technical-report.pdf.

ONS. 2020a. "Coronavirus and homeworking in the UK: April 2020." https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/coronavirusandhomeworkingintheuk/april2020.

ONS. 2020b. "Coronavirus and Homeworking in the UK Labour Market: 2019." https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/coronavirusandhomeworkingintheuklabourmarket/2019.

Peters, Pascale, Kea G Tijdens, and Cécile Wetzels. 2004. "Employees' opportunities, preferences, and practices in telecommuting adoption." *Information & Management* 41 (4): 469–482.

Philip, Lorna, Caitlin Cottrill, John Farrington, Fiona Williams, and Fiona Ashmore. 2017.

“The digital divide: Patterns, policy and scenarios for connecting the ‘final few’in rural communities across Great Britain.” *Journal of Rural Studies* 54: 386–398.

Ratanamahatana, Chotirat Ann, and Eamonn Keogh. 2004. “Everything you know about dynamic time warping is wrong.” In *Third workshop on mining temporal and sequential data*, Vol. 32. Citeseer.

Riddlesden, Dean, and Alex D Singleton. 2014. “Broadband speed equity: A new digital divide?” *Applied Geography* 52: 25–33.

Salemink, Koen, Dirk Strijker, and Gary Bosworth. 2017. “Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas.” *Journal of Rural Studies* 54: 360 – 371. `http://www.sciencedirect.com/science/article/pii/S0743016715300176`.

Sardá-Espinosa, Alexis. 2019. “Time-Series Clustering in R Using the dtwclust Package.” *The R Journal* 11 (1): 22–43. `https://doi.org/10.32614/RJ-2019-023`.

Sarda-Espinosa, Alexis. 2019. *dtwclust: Time Series Clustering Along with Optimizations for the Dynamic Time Warping Distance*. R package version 5.5.6, `https://CRAN.R-project.org/package=dtwclust`.

Siha, Samia M, and Richard W Monroe. 2006. “Telecommuting’s past and future: a literature review and research agenda.” *Business Process Management Journal* .

Singh, Palvinder, Rajesh Paleti, Syndney Jenkins, and Chandra R Bhat. 2013. “On modeling telecommuting behavior: Option, choice, and frequency.” *Transportation* 40 (2): 373–396.

Singleton, Alexander D., and Paul A. Longley. 2009. “Geodemographics, visualisation, and social networks in applied geography.” *Applied Geography* 29 (3): 289 – 298. `http://www.sciencedirect.com/science/article/pii/S0143622808000726`.

Stern, M, Alison Adams, and Shaun Elsasser. 2009. “How levels of Internet proficiency affect usefulness of access across rural, suburban, and urban communities.” *Sociological Inquiry* 79 (4): 391–417.

Van Deursen, Alexander, and Jan Van Dijk. 2011. “Internet skills and the digital divide.” *New media & society* 13 (6): 893–911.

Van Deursen, Alexander JAM, and Ellen J Helsper. 2015. “The third-level digital divide: Who benefits most from being online?” In *Communication and information technologies annual*, edited by Laura Robinson, Shelia R. Cotten, and Jeremy Schulz, 29 – 52. Emerald Group Publishing Limited.

Van Deursen, Alexander JAM, and Jan AGM Van Dijk. 2014. “The digital divide shifts to differences in usage.” *New media & society* 16 (3): 507–526.

Warren Liao, T. 2005. “Clustering of time series data—a survey.” *Pattern Recognition* 38 (11): 1857 – 1874. `http://www.sciencedirect.com/science/article/pii/S0031320305001305`.

6. Appendices

Any appendices should be placed after the list of references, beginning with the command `\appendix` followed by the command `\section` for each appendix title, e.g.

```
\appendix
\section{This is the title of the first appendix}
\section{This is the title of the second appendix}
```

produces:

Appendix A. This is the title of the first appendix

Appendix B. This is the title of the second appendix

Subsections, equations, figures, tables, etc. within appendices will then be automatically numbered as appropriate. Some theorem-like environments may need to have their counters reset manually (e.g. if they are not numbered within sections in the main text). You can achieve this by using `\numberwithin{remark}{section}` (for example) just after the `\appendix` command.

Please note that if the `endfloat` package is used on a document containing appendices, the `\processdelayedfloats` command must be included immediately before the `\appendix` command in order to ensure that the floats in the main body of the text are numbered as such.

Appendix A. Troubleshooting

Authors may occasionally encounter problems with the preparation of a manuscript using LaTeX. The appropriate action to take will depend on the nature of the problem:

(i) If the problem is with LaTeX itself, rather than with the actual macros, please consult an appropriate LaTeX 2_ε manual for initial advice. If the solution cannot be found, or if you suspect that the problem does lie with the macros, then please contact Taylor & Francis for assistance (`latex.helpdesk@tandf.co.uk`).
(ii) Problems with page make-up (e.g. occasional overlong lines of text; figures or tables appearing out of order): please do not try to fix these using 'hard' page make-up commands – the typesetter will deal with such problems. (You may, if you wish, draw attention to particular problems when submitting the final version of your manuscript.)
(iii) If a required font is not available on your system, allow TeX to substitute the font and specify which font is required in a covering letter accompanying your files.

Appendix B. Obtaining the template and class file

B.1. *Via the Taylor & Francis website*

This article template and the `interact` class file may be obtained via the 'Instructions for Authors' pages of selected Taylor & Francis journals.

Please note that the class file calls up the open-source LaTeX packages booktabs.sty, epsfig.sty and rotating.sty, which will, for convenience, unpack with the downloaded template and class file. The template calls for natbib.sty and subfigure.sty, which are also supplied for convenience.

B.2. *Via e-mail*

This article template, the `interact` class file and the associated open-source LaTeX packages are also available via e-mail. Requests should be addressed to `latex.helpdesk@tandf.co.uk`, clearly stating for which journal you require the template and class file.