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 November 27, 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 lockdown 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. **need to add urban-rural to analysis** 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

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

of telecommuting at a time of extreme demand. We will also consider what lessons can be learned from this time for a future where telecommuting is likely to remain a common means of accessing work, 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. We identify this as a new digital divide, one that fundamentally alters the potential returns of

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

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. **WE NEED TO CONNECT OUR FINDINGS WITH THIS AND JUSTIFY THAT THIS IS INDEED A NEW DIVIDE** 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 which types of places, are more likely to land on the right side of this new digital divide. By asking how resilient broadband 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 origins of the new digital divide revealed by the pandemic and its impact on the economic resilience of different places. We then describe our data and methodology. Our results section first offers 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 indicate 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. For example, the systematic review from Salemink, Strijker, and Bosworth (2017) highlights the infrastructure quality differences between urban and rural areas in various advanced economies. Whether this variation in infrastructure quality affects the spatial footprint of telecommuting has not previously been measured, although 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). Another study identifies a correlation between access to internet services and a reduction in household transport spend across 33 countries (Bris, Pawlak, and Polak 2017). Whether the implication is less travel because of increased telecommuting, or whether internet access enables more efficient travel, this finding is an indication of the potential household savings better internet services offer. Thus, the extreme demand during the pandemic provides a new opportunity to understand how infrastructure accessibility, quality, and reliability affects telecommuting, particularly as working from home during the pandemic required high volumes of bandwidth-intensive video conferencing in order to avoid the face-to-face contact that could increase the spread of infection. By first answering questions about internet service resilience, we can also refine our understanding of how this has contributed to or reduced the new digital divide, where economic resilience has been dependent upon the capability to work from home.

The multi-layered digital divides intersect with materials 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 longstanding 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 part of the 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 all three levels of digital divides are, to a certain extend, associated 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, at least partially, 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.

2.3. *other subsection?*

I THINK WE HAVE ENOUGH OF A LIT REV - circumstances where work can be carried out more flexibly in space and time - how relate to digital divides - capability theory of mobility? - broadband tech stuff? broadband studies / divides / resilience

Some new papers google recommended to me:

- `https://urbanstudies.uva.nl/binaries/content/assets/subsites/centre-for-urban-studies/working-paper-series/wps_43.pdf`
- `https://link.springer.com/article/10.1007/s11116-020-10136-6`
- `https://www.sciencedirect.com/science/article/pii/S0966692319311305`
- check who cites the above and what they cite

MORE SOURCES:

- `https://www.coronavirusandtheeconomy.com/question/why-has-coronavirus-affected-cities-more-rural-areas`
- EPB commentaries
- `https://www.coronavirusandtheeconomy.com/question/what-has-coronavirus-taught-us-about-working-home`
- `https://www.coronavirusandtheeconomy.com/question/who-can-work-home-and-how-does-it-affect-their-productivity`
- `https://www.coronavirusandtheeconomy.com/question/how-will-economic-effects-coronavirus-vary-across-areas-uk`
- `https://www.coronavirusandtheeconomy.com/which-parts-uk-have-been-hit-hardest-covi`
- `https://www.coronavirusandtheeconomy.com/question/why-has-coronavirus-affected-cities-more-rural-areas`

3. Methods and data

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, which involves 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 is the basis of *geodemographics*, which aims to create small area indicators or typologies of neighbourhoods based on various and some times 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 characteristic of these studies is 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 creating clusters of local authorities in the UK with similar temporal signatures of experienced internet speeds over time. 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 authority 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).

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 a based the formation of vectors of features based on the original time-series. These new data of reduced dimensionality is then clustered using conventional clustering algorithms.

For this paper we utilise the third 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 among these two shapes and are better equipped to capture similarities between short length time-series (Aghabozorgi, Shirkhorshidi, and Wah 2015). This approach serves best this paper because (i) we aim to 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 clustering of cross-sectional data, we can employ partitioning algorithms, which lead to non-overlapping clusters, hierarchical clustering, which result to a hierarchy of clusters 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 is the initial cluster assignment, which 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 in the 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).

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 regarding individual internet speed tests from Speedchecker Ltd[2]. This 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 of local loop unbundling regulatory processes (Nardotto, Valletti, and Verboven 2015). These volunteered geographic data enable to assess the *experienced* internet speed by users, which may differ from the *advertised* maximum speeds provided by Internet Service Providers. We are particularly interested in upload speeds and the frequency of speed tests. While the former is less associated with internet-based leisure activities such as video streaming, which involve downloading large amounts of data, the latter can be linked more with work-related activities such as uploading documents or two-way communication and interaction systems. **ADD MORE. HOW ABOUT SPEED TESTS?**

The first step in the workflow after dropping some outliers following Riddlesden and Singleton (2014) is to transform and the individual, geolocated and time-stamped tests to more meaningful aggregates both in terms of space and time. Regarding the temporal dimension, 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 6am were also excluded. The composite week time-series thus comprise 18 hours multiplied by 5 weekdays or 90 time points per series. We aggregate these data because, we could not follow individuals or households and connect data points. Also, although this is a large data set ($241,088$ individual tests during the working hours of study period), there are not enough observations for each local and for each working hour (631 speed test per LAD in average). These time-series were calculated for each of the 382 Local Authority Districts (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 $k = 5, 10, 15 and 20$ and used cluster validity indices (CVIs) to pick the optimal solution of $k = 10$. Following Sardá-Espinosa (2019) the following CVIs were employed: Silhouette (max), Score function (max), Calinski-Harabasz (max), Davies-Bouldin (min), Modified Davies-Bouldin (DB*, min), Dunn (max), COP (min).

We then match the cluster membership to the LAD-level data to identify the characteristics of each cluster, including number of LADs, the descriptive statistics of upload speeds in that cluster, and the temporal profile by hour of the day and day of the

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

composite week. Thus we aim to answer the first part of our research question: How resilient are broadband speeds as experienced in different parts of the UK during a time of extreme demand?

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 of experienced upload speeds. 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 order to describe upload speeds as fully as possible. The following **graphs / tables** show these statistics for the sample as a whole, as well as similar statistics related to frequency of testing.

Data details and some figures, descriptive stats - include whole sample time profile for 2019 and 2020 frequency of tests run as part of why we chose to create the time profiles by hour of the day and day of the week rather than daily over the whole period.

Just looking back, it was partly the frequency of tests, but that was also by geography, so do we need maps? And I feel there was something else I'm forgetting...

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. We discuss 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. More specifically, we estimate the following multinomial model:

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

The identification strategy is as follows. 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 we are discussed in details in **section XX**. 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 used to cluster the local authorities have been summarised in graph [.], which shows a composite profile of mean upload speeds per hour per day for each cluster. For upload speeds, 345 of 382 local authorities, or over 62 million people, fall into cluster 6 or cluster 9. Graph [.] suggests that these two clusters have relatively similar temporal profiles, which are flatter than the other, smaller clusters, suggesting better reliability of service. However, the upload speeds at all times for cluster 9 are substantially higher than for cluster 6, which is an indication of better quality of service. This difference is reflected not only in the mean speeds for these clusters for the whole sample, but also the mean upload speeds in the morning peak from 9:00-10:59, as well as the evening peak period from 19:00-20:59. In comparison, the time profile in graph [.] shows upload speeds in cluster 1 are on average lower at certain times of day during the study period than any other cluster, whilst the profile for cluster 3 appears to show speeds fluctuating as much as cluster 1, but at levels usually higher even than cluster 9.

The variability of the smaller clusters may be related to the fewer speed tests from fewer local authorities that have been averaged, whilst averaging greater numbers of speed tests could artificially flatten the profile. This appears not to be the case for cluster 9, as there is negligible difference between morning and evening upload speeds, at 1% slower in the morning, confirming a high level of reliability of service. In comparison, upload speeds in cluster 6 are 4% slower in the morning than in the evening and experienced the joint highest ratio of standard deviation to mean across the time period under assessment. This suggests that although the time profile is relatively flat in graph [.], the experience is one of speeds that fluctuate from a lower mean, and therefore might more often impact on online activities. Still, in most of the smaller clusters the reliability of service is worse than in cluster 9 in terms of the ratio of standard deviation to mean. The smaller clusters are mainly worse than clusters 9 and 6 in terms of the ratio of upload speed in the morning peak compared to the evening peak. In five of the smaller clusters, which are home to almost 3 million people, speeds are 12-19% slower between 9:00-10:59 compared to 19:00-20:59.

The quality and reliability of broadband service is much better in the 115 Local Authority areas in cluster 9, which are mostly in urban or suburban areas, compared to cluster 6. These include 13 London Boroughs (of 32), 8 of the 10 local authorities of Greater Manchester, 5 of the 7 constituent authorities of the West Midlands Combined Authority, as well as cities like Glasgow, Leicester, Nottingham, Sheffield, and the Portsmouth and Southampton conurbation. There are also some notable medium-sized cities, including Aberdeen, Cardiff, Oxford, Milton Keynes, and York, and many suburban districts from the South East of England to South Tyneside. Meanwhile, the 230 local authorities in cluster 6, which have lower speeds on average and more variation in service still include major urban areas, such as Bristol, Liverpool and Leeds, and many suburban areas, but also include some of the most rural areas in the country. Meanwhile, Cluster 1, with 10 local authorities that are home to over 1 million people has the second slowest speeds in the morning compared to the evening 'peaks' and the second highest ratio of standard deviation to mean. This cluster's most populous area is Westminster in central London.

further description of temporal profiles of other clusters - include all in graph? Indeed, the only cluster where upload speeds were slower in the evening than in the morning was cluster 5, made up of 2 local authorities with less than 200,000 people: Three Rivers, a suburban district north of London, and Fylde, a seaside suburb of Blackpool. However, these are likely to be outliers and may not have many tests from which the clusters are calcualted. **Take out smaller clusters?** The exceptions can be found in seven of the eight much smaller clusters including 35 local authority districts, where AM peak upload speeds are between 6% and 18% slower than PM peak upload speeds, although the mean speeds for each cluster are higher than cluster 6. Indeed, in 25 local authorities with a combined population of almost 3 million, speeds are 13% or more slower in the morning than in the evening. Included in this latter group are central London borough of Westminster and the London Borough of Newham, rural authorities like Eden and West Devon, and small cities like Dundee and Carlisle.

4.2. *aux regressions*

Auxiliary regressions indicate that the speed tests in cluster 9 authorities are also more likely to have been run on services provided by Virgin Media, suggesting they are in the half of the country with the most lucrative ICT market, which originally attracted the cable TV provider (OfCom...). For example, although auxiliary regressions show that Cluster 6 local authorities are more likely to be in the South of the UK than Cluster 9, the cluster notably includes Southern rural districts from Cornwall to North Norfolk. Slower speeds could reflect the lower quality of service in rural areas compared to urban and suburban areas. There is frustration, however, as auxiliary regressions show that those living in cluster 6 had the highest probability of testing their broadband between 9:00 and 11:00 of any of the 10 upload speed clusters.

the finding in the auxiliary regressions that those in cluster 9 ran the fewest speed tests per person during the morning period of any cluster. Now this may be an indication of fewer people working from home, less contention, and less resultant frustration. Cluster 9 comprises many central urban areas and has the lowest number of established businesses per inhabitant, which could be interpreted as a dominance of large employers. However, the job density is lower than cluster 6, meaning there are not as many jobs per resident in these areas. As the cluster includes many suburban areas too, which may be largely residential, could the quality and reliability of internet service be reinforcing patterns of telecommuting by those in wealthier suburbs who can work from home? Earnings in cluster 9 are second highest of all the clusters, with only Cluster 2 (comprising just North East Lincolnshire and East Lothian, population 265k) earning more per person.

The auxiliary regressions show that compared to the two authorities in cluster 5, all the other clusters had a lower percentage of working people in managerial, professional and administrative jobs.??? Yet the auxiliary regression suggests that there are not many tests being run during the am peak in cluster 1. This may be because there are fewer people working at home checking their broadband than in most other clusters. Indeed, the auxiliary regressions indicate that cluster 1 has the highest job density or proportion of jobs to working-age population, which is likely to due to the presence of Westminster, central London, cluster 1's most populous local authority. Westminster not only has more workplaces than residents, but it is reasonable to presume that many who would normally work in Westminster, but be able to work from home during lockdown are likely to live outside central London and not be subject to the fluctuating speeds there. Workplaces, meanwhile, some of which would still have been open, could be running programmes that cause the slowdown and variation, but would be more likely to have their own in-house diagnostics, rather than using a service like Speedchecker Ltd.

5. Conclusions

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.

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. 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.