Applied Geography 52 (2014) 25–33

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

Broadband speed equity: A new digital divide?

Dean Riddlesden*, Alex D. Singleton

Department of Geography and Planning, School of Environmental Sciences, University of Liverpool, United Kingdom

Keywords:
Broadband
Internet
Digital divide
GIS
Crowdsourced data
Geodemographics

A B S T R A C T

The availability and performance of broadband connectivity is becoming an increasingly important issue across much of the developed world as the prevalence of richer media services and growing populations have generated increasing demands on existing networks. The heterogeneous geography of broadband infrastructure and investments results in variable service provision, and as such, there exist large disparities in access and performance within different spatio-temporal locations. This paper presents analysis of 4.7 million crowdsourced Internet speed test results that were compiled between 2010 and 2013 alongside various indicators of socio-spatial structure to map disparities in English broadband speed between and within urban areas. Although average speeds have improved over time, inequity is shown to emerge between different societal groups and locations. Short-term dynamics also reveal that in areas of different density, speeds can fall dramatically during peak hours, thus influencing the availability of services. The apparent disparities in access and performance represent a major issue as Internet use becomes increasingly ubiquitous in our everyday lives, with inequalities evoking social and economic disadvantage at local and national scales. This work resonates with UK government policy that has stimulated considerable investment in improving infrastructure, and presents analysis of an expansive crowd sourced "big data" resource for the first time.

Broadband and the Internet

Within developed countries, access to the Internet in urban areas is predominantly available in the home (Dutton & Blank, 2013) and enabled through wired broadband infrastructure. This availability reduces traditional constraints on communication such as time and cost, enables the consumption of rich media services (e.g. video chat, streaming movies or music) and most importantly, enhances access to a plethora of online information and services that engender benefits for economic development, education, health and wellbeing (Broadbent & Papadopoulos, 2013; Kraut et al., 2002). Particular emphasis has been placed by governments on the link between the provision of broadband infrastructure and economic growth (Freund & Weinhold, 2004; Picot & Wernick, 2007; Xavier, 2003; Yiu & Fink, 2012), leading to promotion and investment in these technologies as part of national infrastructure plans. In the USA, the National Broadband Plan, developed by the Federal Communications Commission (FCC) aims to promote broadband availability through ensuring robust competition and universal service, as well as maximising the benefits of broadband in government influenced sectors (Federal Communications Commission, 2011). Similarly, UK government response has come in the form of the Digital Britain Report, which outlined a nationwide Universal Service Commitment (USC) to be achieved by 2012 (DCMS, 2009). The importance of such investments have been widely cited, with previous studies estimating that broadband infrastructure accounted for 9.53% of the United Kingdom's GDP growth in the period 2002–2007 (Koutroumpis, 2009). Investment in developing countries has also increased in recent years, with heavily subsidised rollout of high-bandwidth infrastructure across much of the global south (Graham & Mann, 2013), although not necessarily fixed line connectivity. For example, some private sector funded projects aim to provide widespread Internet access through the launch of medium orbit satellites (O3B, 2013). Other projects, such as the One Laptop Per Child program (OLPC), have aimed to promote the distribution of inexpensive computing equipment with wireless connectivity to children (Graham, 2011).

Early research into Internet inequalities was concerned with disparities in connectivity between developed and developing countries, and what social impacts these differences would likely engender. Within this context, the term 'Digital Divide' was introduced in the late 1990s (Norris, 2001) to describe differences between the 'haves' and 'have-nots'. Although a valuable concept at the time when the Internet was first developing, as the access divide narrowed (Kyriakidou, Michalakelis, & Sphicopoulos, 2011;

* Corresponding author. Tel.: +44 (0)1517941355.
E-mail address: d.riddlesden@liv.ac.uk (D. Riddlesden).

http://dx.doi.org/10.1016/j.apgeog.2014.04.008

Peter & Valkenburg, 2006), more recent discussion has diversified this binary concept, adopting instead the term 'Digital Differentiation' (Hargittai, 2002; Longley, Webber, & Li, 2008) to signify those new complexities that have emerged in differentials of Internet access. The digital differentiation approach aims to depict more nuanced differences between societal groups in terms of use and engagement patterns, that are evident especially within those countries with more developed Internet infrastructure (Longley, 2003). Such differentiating factors explored by this research have included age (van Dijk & Hacker, 2003; Lenhart, Madden, & Hitlin, 2005; Livingstone & Helsper, 2007; Loges & Jung, 2001), gender (Bimber, 2000; Cooper, 2006), rurality (Pigg & Crank, 2005; Prieger, 2013) and ethnicity (Fairlie, 2004; Prieger & Hu, 2008). Most recently, research has diversified to utilise crowdsourced information generated through location-based social media and Internet use to better understand the complex emerging geographies linked with the built environment (Arribas-Bel, 2014; Sui & Goodchild, 2011). Manifest from digital differentiation are patterns of digital exclusion where individuals will have varying degrees of engagement with the Internet (Bunyan & Collins, 2013). The complexity of such patterns has been discussed elsewhere (Longley, 2003), and have been shown to emerge between the intersections of traditional concepts of material deprivation (Longley & Singleton, 2009). Thus, you may have urban areas with populations that are typically digitally engaged, yet materially deprived, or inversely, urban areas with low digital engagement, but also low levels of material deprivation. Within a UK context, overall levels of digital exclusion have declined steadily in recent years (Dutton & Blank, 2013; Lane-Fox, 2009), albeit, a significant proportion of the population still remain digitally excluded. In 2013, this equated to around 11.4 m people having never used the Internet (18% of the UK population) (Dutton & Blank, 2013). Although these differentials are part of a wider complex of influences (Parayil, 2005), in part, these inequalities will relate to the provision of infrastructure enabling access to the Internet, and as such, underpin a UK government aim to ensure that a minimum threshold of Internet speed is enabled for all of the populations. The Universal Service Commitment outlined in the government's Digital Britain Report set a nationwide minimum connection speed of 2 Mbps to be delivered by 2012. It was stated that this target "would allow virtually everyone to experience the benefits of broadband, including the increasing delivery of public services online" (DCMS, 2009, p. 27).

Geographic disparities in access and performance exist in part due to the physical structure of broadband networks. In particular, performance is affected by the distance (or line length) between a customer's home and the nearest telephone exchange. Such is the effect, that distance to the nearest exchange is often used as a proxy for deliverable speed (Ofcom, 2012). Within rural areas, distance has a vastly limiting effect on service provision due to sparse distribution of both populations and core network infrastructure such as exchanges and backbone networks; as such, large disparities in broadband performance exist between urban and rural areas. Similar limitations of cellular networks also exist as a result of rurality, with mobile broadband coverage being often poor in isolated areas. However, although relevant to the wider field of research, disparities in mobile broadband access fall outside the scope of this paper. Advances in communications technology, such as the use of fibre optics to supply domestic broadband services, aim to increase broadband speed in locations disparate from telephone exchanges. Fibre to the Cabinet (FTTC) connections link street cabinets that supply small neighbourhood areas to a local telephone exchange with a dedicated fibre link: this allows for much faster transmission of data. FTTC connections (where available) can currently deliver speeds of up to 76 Mbps to homes in the UK (BT PLC, 2013). By contrast, Fibre to the Premises (FTTP) connections (often utilised by businesses) supply a direct fibre link to a site, allowing for large volumes of data to be exchanged rapidly. FTTP connections to domestic properties, offering speeds of up to 300 Mbps are currently only available in a small number of areas and as such are very much in their infancy. Within this context, this paper explores how provision and performance of connections to broadband infrastructure within England are both temporally and socio-spatially differentiated; evoking a complex geography of connectivity both between and within urban areas. For the first time, this study utilises a large dataset of crowdsourced Internet speed estimation tests for England, supplied by the company Speedchecker Limited (broadbandspeedchecker.co.uk), and pertaining to 4.7 million test results, with geographic attribution at the level of the unit postcode (zip code). An extract comprising two time periods was provided, covering 1/1/2010 to 31/1/2011 and 1/4/2012 to 31/5/2013.

Measuring broadband connectivity and access

Data were supplied by Speedchecker Limited, who are a provider of a Web-based application (broadbandspeedchecker.co.uk) that enables users to test their Internet connection speed. When users visit the website, a page is loaded with an embedded testing application that when run provides a small file (of known size) that is automatically downloaded and uploaded, thus enabling speeds to be estimated (i.e. size/time). After running a speed test, users are requested to supply a unit postcode and to confirm details of the connecting Internet company/package. The unit postcode details enable speed test results to be geo-located, and the website displays the test outcome within the context of other results proximal to the supplied postcode. All results derived through the website are stored by Speedchecker Limited as part of their terms of use. The geographic resolution of the geo-located speed tests is therefore high, with postcodes relating to on average around 13 households. It is important to note that these tests differ from "official" speed tests in the UK, which are collated on behalf of the industry regulator Ofcom (http://www.ofcom.org.uk) by SamKnows (www.samknows.com), who are an organisation that provide information about broadband performance, providers and usage. The crucial difference between SamKnows data and the data supplied by Speedchecker Limited is the collection method. Rather than relying on users to run speed tests through a Web application, SamKnows supply hardware in the form of a small testing box that sits between participants' existing routers and the rest of their network. Boxes are supplied to a representative sample of Internet users nationwide. In 2010, Ofcom's UK broadband performance report utilised data collected from 1506 testing boxes. The boxes automatically ran speed tests on a user's connection, but only when there was no other network activity. Conversely, data supplied by Speedchecker Limited is much larger, but there are no restrictions to prevent users from performing tests when there is other network activity ongoing (e.g. multiple users online within a property, or a background update being downloaded). As such, the data presented here could be interpreted as those actual speeds people attain when using a service, taking into consideration local constraints related to router configuration, WiFi coverage or coincident household usage. This said, comparison between the regulator estimates and the derived Speedchecker Limited estimates revealed very similar figures. Our data sample suggested a nationwide average download speed of around 4.8 Mbps in 2010, close to that estimated by Ofcom/SamKnows at 5.1 Mbps (Ofcom, 2010).

The data used for this study can be considered as 'Volunteered Geographic Information' (VGI) (Goodchild, 2007; Haklay, Singleton,

& Parker, 2008), which although a useful way to generate large amounts of geographic information, require interpretive caution related to data accuracy, coverage and bias. As such, both internal and external validations were implemented to explore sources of potential bias or error. Firstly, speed test results were explored to identify anomalous results, or records that may pertain to non-broadband connections. After researching those packages on offer by suppliers during the period the data was collected, it was deemed that the lower boundary for broadband speed would be set at 512 Kbps (half a Megabit per second) and the upper at 102,400 Kbps (100 Megabits per second). Any results that fell below the minimum threshold, or above the maximum were deemed outliers and removed prior to analysis. In applying these thresholds for speed, we expect the results retained to represent a mix of connection types; predominantly traditional ADSL connections, followed by a significant proportion of more recent FTTC and cable connections and a very limited number of FTTP and satellite connections, thus reflecting the distribution of broadband technologies available in England. Further records were also removed where there was either an invalid or blank postcode, thus preventing geocoding. These tests resulted in the removal of around 790,000 (27.5%) records from our 2010/11 dataset, and around 450,000 (24.1%) from our 2012/13 dataset.

Following this analysis, the remaining results were assigned to their nearest telephone exchange using a nearest neighbour algorithm. The location of telephone exchanges was derived from the website SamKnows and geo-coded by unit postcode. The Euclidean distance between each speed test result and the nearest telephone exchange was then calculated. This analysis showed that the vast majority of speed tests were run from postcodes within 2.5 km of their nearest telephone exchange; and the average distance of a speed test location to the nearest exchange was shown to be around 1.6 km in both datasets. These results are in line with what might be expected given what is known about the technical limitations of digital subscriber line technology, particularly that long distances result in high levels of signal attenuation and reductions in speed. As such, it would be unlikely to find broadband connections at distances over 5 km (Ofcom, 2013).

Further validation may have involved locating users by their IP address at the time of completing the test, and then comparing this result to their supplied postcode. However, the data contained no IP address information, and thus, such a test was not possible. In general, analysis of download speeds by distance from the nearest exchange showed that speed tests conducted in close proximity to the nearest telephone exchange returned the fastest download results, particularly those within 1000 m. Speeds start to deteriorate noticeably at distances of over 2000 m. Fig. 1 presents analysis of average download speeds by distance to the nearest telephone exchange.

However, these analyses also highlighted an apparent anomaly. In both datasets, results appeared between 5000 m and 7000 m from the exchange with increases in mean download speeds; and may represent a percentage of speed tests run through fibre or coaxial-based connections. Such infrastructure allows speeds of up to 76 Mbps on fibre to the cabinet (FTTC) lines (Ofcom, 2012). However, within both datasets only a small percentage (0.5% and 0.6% respectively) of the speed test results were recorded at distances over 5000 m from the closest exchange. Table 1 presents the raw number of connections recorded at these distances in our 2012/13 dataset.

To explore the geographic distribution of test results, the number of speed test results per head of population was calculated for each English local authority district; which is a local administrative geography. For this analysis, both datasets were combined in order to assess whether there was any geographic bias within the data over the entire period of data collection. Fig. 2 illustrates the output of this analysis. It is apparent that there are some districts, particularly those that are within predominantly rural areas such as Suffolk, the Lake District and the Cotswolds, which record a higher propensity for speed testing. This may be due to poorer performance within these areas in general due to infrastructure constraints, and a greater tendency for Internet users to seek methods

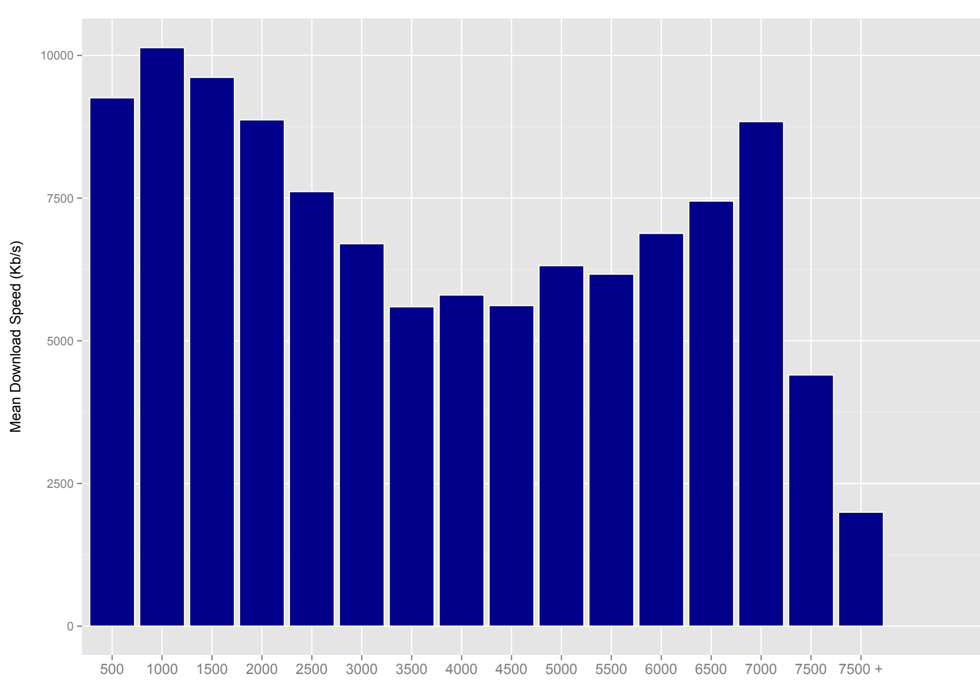

Fig. 1. Mean download speeds by distance to the nearest exchange 2012/13.

Table 1
Number of connections recorded by distance to the nearest exchange 2012/13.

Distance from exchange (m)	Number of speed tests recorded 2012/13
0–500 m	136,697
500–1000 m	279,502
1000–1500 m	292,328
1500–2000 m	247,114
2000–2500 m	181,033
2500–3000 m	119,005
3000–3500 m	73,380
3500–4000 m	41,663
4000–4500 m	22,775
4500–5000 m	11,150
5000–5500 m	4812
5500–6000 m	2458
6000–6500 m	820
6500–7000 m	371
7000–7500 m	150
Over 7500 m	56

of monitoring their connection performance. Conversely, predominantly urban centres such as London, Birmingham, Manchester and Liverpool record fewer speed tests per head of population. This may be due to higher average speeds in these areas as a result of better infrastructure.

To investigate the aggregated socio-spatial structure of test results, the preliminary Office for National Statistics Output Area Classification (OAC) was appended to the test results. This classification was created for England and Wales from 2011 Census Data (www.opendataprofiler.com/). Geodemographic classifications are categorical summary measures of the aggregate social and built environment structure of small areas. These have international use across a wide range of application areas (Singleton & Spielman, 2013) and assist in the analysis and visualization of complex geographies at variable scales (Singleton & Longley, 2009). OAC is appended to Output Area zones, which comprise an area defined by a minimum of 40 households; however, the optimal size is 125 households. OAs contained an average population size of 309 people in 2011. OAC was appended to the records, and the number of speed test results per postcode (the spatial resolution at which our data were geo-tagged) was calculated within each OAC Group (See Fig. 3). For this analysis, we again used the combined datasets, and the number of current postcodes within each OAC group was extracted from the August 2013 Office for National Statistics Postcode Directory.

It is apparent that there are more speed tests performed by residents within certain OAC groups than in others. Group 6b: Established Suburbs records the highest rate at around 4.5 test results per postcode. This is closely followed by groups 6a: Inner Suburbs, 6c: Suburban Aspiration and 4b: Blue Collar Transitions. As we would expect, areas with an over representation of older residents, who are frequently reported to be less engaged with communications technology such as broadband Internet, are under-represented in the dataset. Groups 7c: Elderly in Flats and 8c: Late Retirement record significantly lower rates. The lowest rate is however recorded by group 3A: Urban Deprivation, perhaps where material deprivation has become a more general barrier to Internet use.

Broadband access speeds and equity by indicators of socio-spatial structure

The validated broadband speed test data are utilised in the remaining sections of this paper. Through these analyses we illustrate both the social and spatial patterns of broadband speed inequity, followed by consideration of how these patterns vary over multiple temporal scales. Disparities in broadband speed across England are visualised in Fig. 4 by mapping the average download speeds of our most recent dataset (covering 2012/13) by English local authority district.

There is significant clustering of higher average download speeds around major urban centres such as London, Birmingham, Manchester and Liverpool. As might be expected, predominantly rural regions such as the South West, Norfolk, Yorkshire and Lincolnshire record the majority of the slowest download speeds. On the basis of the averages within the dataset, we estimate that a large proportion of households, particularly those in rural areas, at the time the data were collected, would be falling below the UK government's universal service commitment threshold of 2 Mbps. Of the validated speed test results used for analysis in the 2012/13 dataset, 21.8% fell below the USC threshold.

To examine the relationship between speed and rurality more explicitly, the Office for National Statistics (ONS) Urban/Rural classification was appended to each postcode in the dataset. These definitions were created at a Census Output Area (OA) level of geography for England and Wales, and classified as 'urban' if the majority of the population of an output area lived within settlements with a population of 10,000 or more. In addition, the classification also categorises output areas based on context; such as

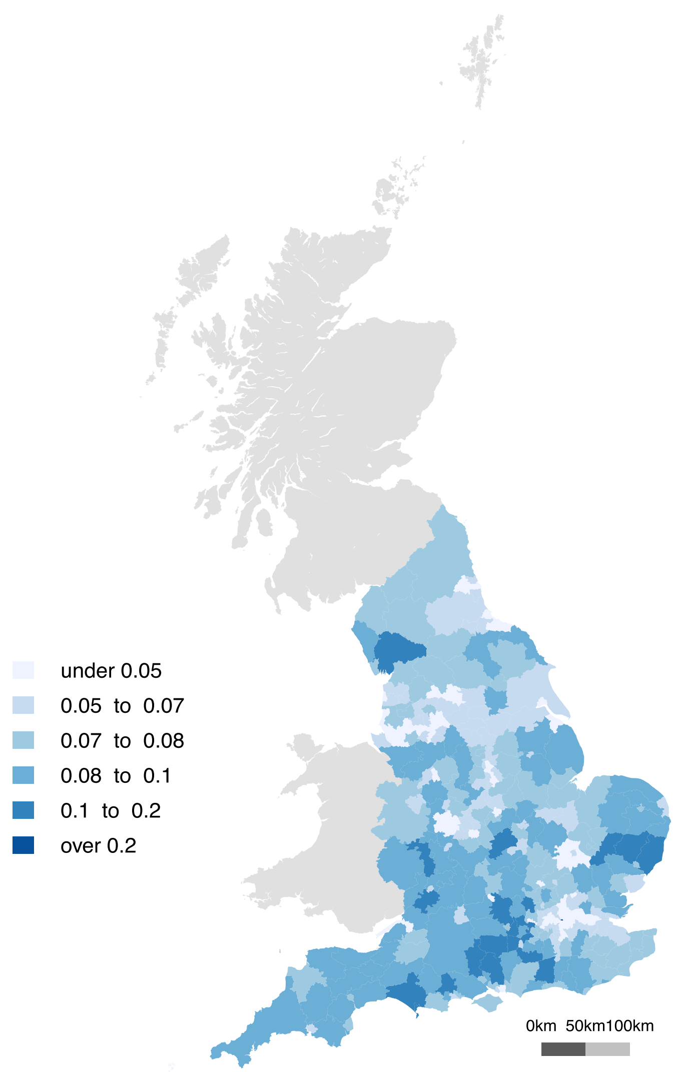

Fig. 2. Speed tests per head of population: English districts.

whether the wider surrounding area of a given output area is sparsely populated or less sparsely populated (Bibby & Shepherd, 2004).

Table 2 shows the average download speed recorded by the Urban/Rural classification alongside average distance to an exchange. The highest average speeds are recorded in those areas classified as 'Urban ≥ 10 K – Less Sparse', essentially areas which are urban and have more densely populated surroundings. Unsurprisingly, the slowest average speeds are recorded in areas classified as 'Hamlets and isolated dwellings' (both 'Sparse' and 'Less Sparse') and villages (again, both 'Sparse' and 'Less sparse'). However, it should be noted that the 'Less Sparse' classes display marginally higher average download speeds than their 'Sparse' counterparts. The results therefore provide further evidence that rurality has a profound effect on the level of service consumers are likely to receive. This is however not a surprising distribution when line length is taken into account, given that longer lines are subject to greater signal attenuation, and as such, likely to deliver slower speeds.

Mean distance to exchange in both the 'Urban ≥ 10 K' and 'Town and Fringe' categories are shorter for the 'Sparse' context than 'Less Sparse'. Initially, this looks anomalous, however, OAs surrounded by less sparsely populated areas would logically be closer to their local exchange, as that exchange would typically be located towards the centre of a populated area in order to supply the best service. In OAs where the surrounding area is classed as 'Less Sparse' (i.e. of higher population density) the logical placement of the exchange may not be as obvious, as the area it must serve is both large and more densely populated throughout. Therefore, it follows that there could be far more OAs at greater distance from their local exchange than those which happen to be closer, thus increasing the average distance.

Although rurality is an important differentiating factor of broadband speed in aggregate, these patterns interact with other social factors. As has been shown elsewhere (Longley & Singleton, 2009), differences in use and engagement with the Internet occur between and at the intersection of patterns of material deprivation. However, to date, equity in broadband speed has not been explored within this context. To examine the relationship between prevailing levels of material deprivation and broadband speeds, the Indices of Multiple Deprivation 2010 (IMD) were converted into ranked deciles and appended to each of the test results. The IMD aims to capture multiple aspects of overall deprivation and is measured at Lower Super Output Area (LSOA) geographic resolution. LSOAs represent small area statistical boundaries that capture between 400 and 1200 households with area populations of between 1000 and 3000 persons. Average speeds were then calculated for each of the IMD decile and are shown in Table 3. These analysis show that average download speeds are higher in those areas ranking as more deprived (particularly so for deciles 1 and 2).

This is due to such areas of high deprivation typically being located within more densely populated urban conurbations, and as such, benefiting from more developed network infrastructure that is necessary to support higher speeds. These results would suggest that prevailing levels of deprivation in England are not necessarily a barrier to broadband access or to higher speed broadband connections. However, one caveat of this analysis is that although speed tests (based on the user supplied postcode) were relatively evenly distributed across all IMD deciles, we possess no data regarding the

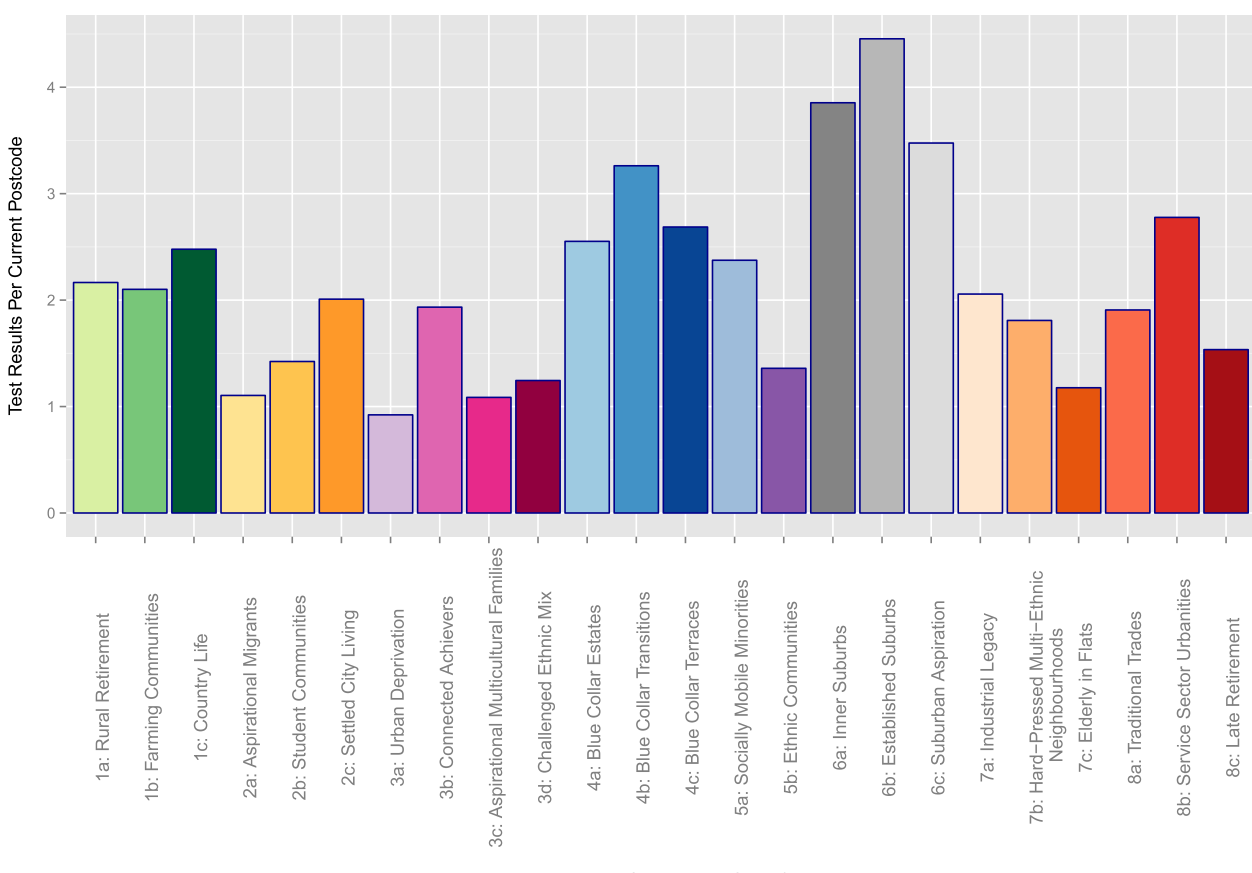

Fig. 3. Test rate (tests per postcode) by 2011 Output Area Classification (OAC).

levels of material deprivation that are experienced by each individual test user, only that at the aggregate level of the LSOA.

The short and long term temporal dynamics of broadband speed

A further attribute for each of the speed test results is a 'time stamp' assigned by the Web server that identifies the time and date when each test was performed. This usefully enables expansion of the socio-spatial analysis across varying temporal scales to provide insight into these changing dynamics. Fig. 5 extends the previously presented urban/rural analysis and illustrates how speeds fluctuate throughout the day within different types of conurbation.

Densely populated urban areas classified as 'Urban ≥ 10 K Less Sparse' record the highest average download speeds throughout the day, with 'Town and Fringe' and sparsely populated urban areas sitting around mid-table. 'Villages' and 'Hamlets and Isolated Dwellings', both sparse and less sparse record the lowest average download speeds throughout the day. In terms of performance variation, there appear to be larger fluctuations in urban areas that are likely due to higher population densities, and more prevalent "bottlenecking" of data traffic when there are a large number of users connected to the Internet. Bottlenecking is a term used to refer to the slowing of a network in times of high use. Within predominantly rural areas, fluctuations in download speed appear to be smaller, although speeds are generally much slower throughout the day. Despite delivering slower speeds, infrastructure in rural areas appears to be less susceptible to bottlenecking, and is most likely due to their composite smaller populations, and as such, lower demand. In all Urban/Rural categories the effects of peak and off peak hours are apparent, with average speeds spiking in the early hours of the morning when there are fewer users online, allowing for faster download speeds. Although we would expect these increases to be measurable, the large spike at 3am in the 'Town and Fringe Sparse' class represents noise in this particular subset of our data. Conversely, the slowest average speeds are recorded between 6pm and 9pm in most areas, which is a peak time for Internet use, and as such, aggregate speeds tend to slow. As well as detailing performance fluctuations, this analysis again highlights the large disparities that exist between urban and rural areas in terms of aggregate broadband speeds.

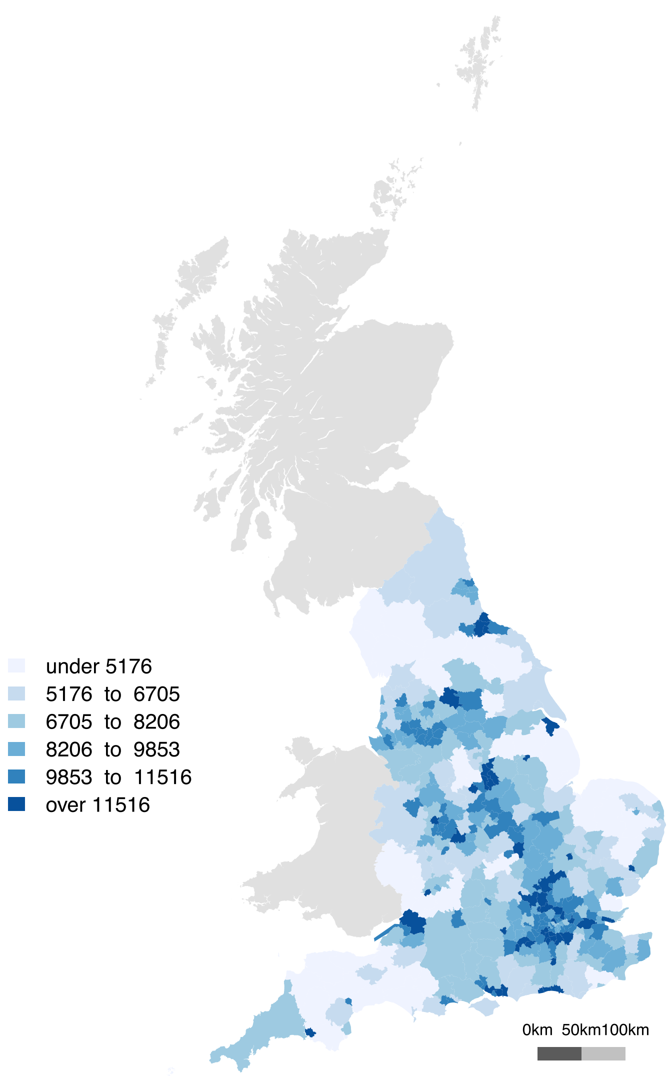

Fig. 4. Mean download speed (Kbps) by English district: 2012/13.

Table 2
Mean download speeds by urban/rural indicator.

Urban/rural indicator	Mean download (Kbps) 2012/13	Mean distance to exchange (m)
Urban ≥ 10 K sparse	6651	971
Town and fringe sparse	6248	736
Village sparse	3380	2272
Hamlet and isolated dwelling sparse	3151	2790
Urban ≥ 10 K less sparse	10,176	1496
Town and fringe less sparse	6952	1575
Village less sparse	4042	2375
Hamlet and isolated dwelling less sparse	3896	2477

However, we can also consider these data at a more granular temporal resolution to highlight how inequalities are changing over time. If the two annual extracts of 2010/11 and 2012/13 are considered in isolation, we find that the national average speed increases from 4.8 Mbps to 8.4 Mbps, with a reduction in the estimated number of people falling bellow the USC 2 Mbps threshold from 30.9% to 21.8%. However, if these results are mapped at the local authority district geography, we see that there is geographic disparity to these improvements (See Fig. 6 and Tables 4 and 5). Unlike the previously presented map of average download speeds (See Fig. 4), change is less obviously clustered around urban conurbations, and instead, high increases in download speed are more widely dispersed. The lowest increases appear predominantly in rural areas, however, some rural areas such as Norfolk and Cornwall have seen relatively high speed increases over the time period. It would be safe to assume that large urban conurbations have not seen large percentage increases in average download speeds as the infrastructure necessary to support high speeds has typically long been in place within these areas. However, the largest increases in average download speeds are predominantly urban, but would not be considered as major urban centres. In most cases, these represent areas that may not previously have benefited from next generation infrastructure that have replaced elements of the cabling carrying broadband with fibre. More densely populated urban centres would logically offer the best return on investment for such enhanced services, and as such, would likely have had infrastructure upgrades prioritised.

As might be expected, the smallest increases in average download speeds were recorded in those districts that are predominantly rural. Districts such as Copeland, Boston and the Isles of Scilly, in particular, are rurally isolated and sparsely populated which constrains the viability of infrastructure upgrades. For example, areas

Table 3
Mean download speeds by IMD decile.

IMD decile	Mean download 2012/13 (Kbps)
Decile 1 (Most deprived)	10,356
Decile 2	9730
Decile 3	8822
Decile 4	8037
Decile 5	7675
Decile 6	7387
Decile 7	8067
Decile 8	8009
Decile 9	8779
Decile 10 (Least deprived)	9326

such as these would probably require government intervention to stimulate the required investment in infrastructure. One such initiative in the UK has been the Rural Community Broadband Fund (RCBF), jointly funded by Defra and Broadband Delivery UK (BDUK), which is aiming to deliver improvements in broadband infrastructure to the most rurally isolated areas.

The relationship between infrastructure investment over these two time periods and broadband speed can be illustrated further by exploring the geography of fibre enabled local exchanges. For a consumer to access a fibre enhanced broadband service, the local telephone exchange serving a property would require this provision to be enabled. In order to investigate the geography of these next generation network upgrades, a comprehensive list of exchanges that are supplying FTTC connections was obtained from BT. The dataset contained the name and a code of each exchange in the UK that was supplying BT's 'Infinity' fibre services. These data were geo-tagged by matching attribute information from our original exchange dataset obtained from SamKnows. The percentage of fibre enabled exchanges within each district was then calculated relative to the total number of exchanges in that district. Fig. 7 shows this analysis on a national scale.

Broadly speaking, the geography of fibre enabled exchanges matches that of broadband speeds, with more FTTC enabled exchanges in predominantly urban areas. London and Manchester in particular have high percentages of FTTC enabled exchanges; however, it is apparent that this is also the case in some rural areas such as Cornwall. Cornwall has been promoting the expansion of fibreoptic broadband through the Superfast Cornwall Program, a £132 m partnership funded by the European Regional Development Fund Convergence Programme, BT and Cornwall Council. This has aimed to upgrade existing broadband infrastructure and enable fibre access to 95% of homes and businesses in Cornwall by 2015. Many other rural areas such as North Yorkshire and Cumbria also have programs in place that make use of BDUK funds to supply superfast broadband to the most hard to reach areas.

Conclusions

This paper has provided an overview of how crowdsourced speed test estimation data can be used to investigate the variable disparities in English broadband infrastructure and access patterns. In doing so, we have presented internal and external validation methods that seek to assess the integrity of the 4.7 million speed test records provided by Speedchecker Limited for the purpose of this study. The validated data were shown to produce broadly similar estimates (here, in terms of nationwide average download speeds) to that of regulator data reported by Ofcom, although, there was some socio-spatial bias within the data which appeared to relate to patterns of overall access participation. In general, analysis of test results by distance to the closest exchange revealed that tests run from close proximity to their nearest exchange returned the highest average download speeds. However, there were distances at which fibre-based connections appeared prevalent, particularly

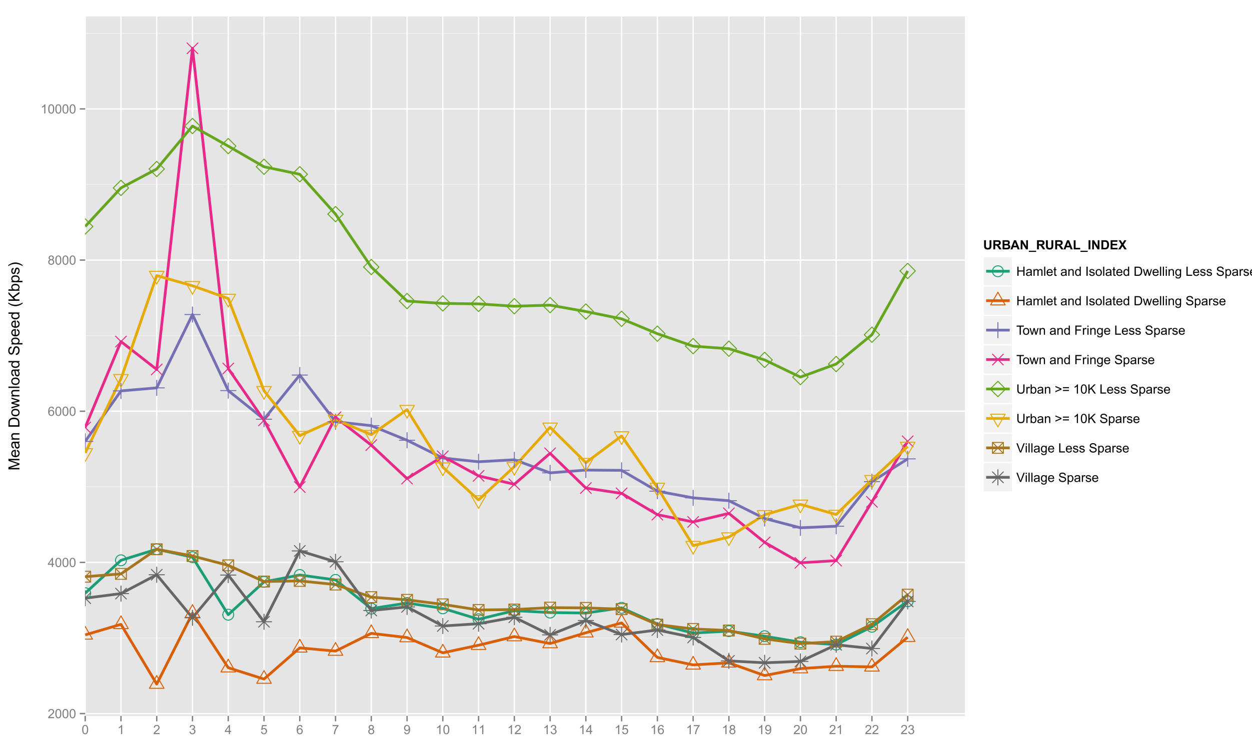

Fig. 5. Hourly fluctuations in mean download speed by urban/rural index.

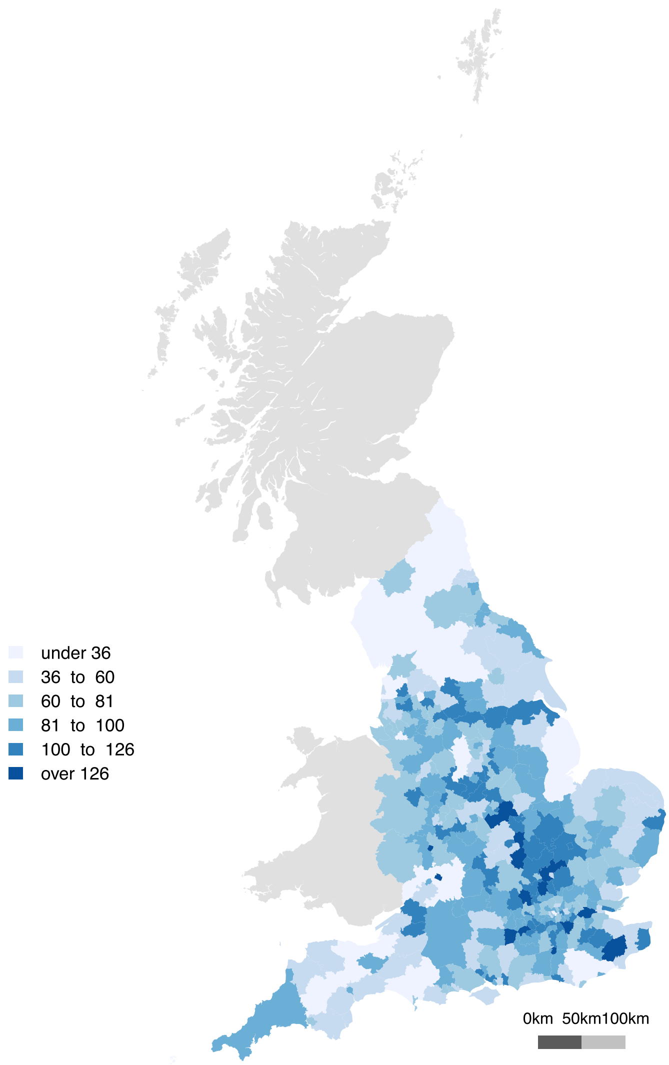

Fig. 6. Percent increase in mean download speed 2010/11 to 2012/13 by English district.

5000–7000 m from a local exchange, and disrupted this trend. More generally, it would appear that predominantly rural areas record a higher propensity for speed testing, possibly due to variable levels of performance as a result of longer lines; motivating a necessity or interest in monitoring speed. Conversely, urban areas tended to record fewer tests per head of population. When profiling by a geodemographic indicator, OAC in this case, further data bias was identified with some groups such as 'Established Suburbs' overrepresented in the data. Equally, some groups such as 'Elderly in Flats' were underrepresented, however, these biases are to be expected and broadly support the notion of digital differentiation between those societal groups likely resident within these areas.

Table 4
Largest ten increases in mean download speeds by English district 2010/11 to 2012/13.

District	Mean download 2010/11 (Kbps)	Mean download 2012/13 (Kbps)	% Increase
Harborough	3763	10,366	175
Ashford	3725	9363	151
Chiltern	4070	10,193	150
Wellingborough	4635	11,601	150
North Hertfordshire	5302	13,112	147
Worcester	3135	7638	144
Surrey Heath	5265	12,600	139
Bromley	5471	12,915	136
Thurrock	5114	12,013	135
Milton Keynes	3474	8128	134

Table 5
Smallest ten increases in mean download speeds by English district 2010/11 to 2012/13.

District	Mean download 2010/11 (Kbps)	Mean download 2012/13 (Kbps)	% Increase
Copeland	3467	3491	0.67
Allerdale	3634	3726	2.53
Craven	4275	4470	4.56
Mid Devon	3248	3667	12.92
Barrow-in-Furness	3702	4226	14.17
West Somerset	3149	3711	17.83
Richmondshire	3332	3932	18.01
Boston	3279	3870	18.02
Cotswold	3599	4279	18.90
Isles of Scilly	2558	3044	18.99

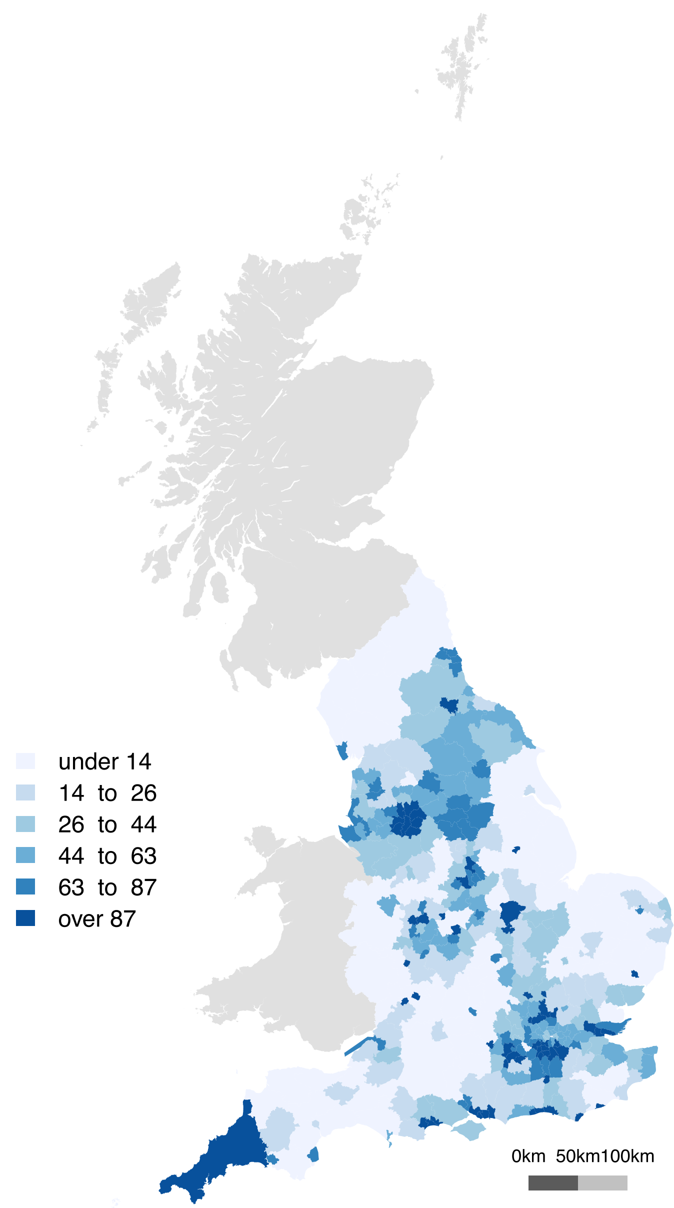

Fig. 7. Percentage of FTTC enabled telephone exchanges by district (October 2013).

Profiling average download speeds at a regional geography revealed that higher speeds are generally clustered around major conurbations. In rural areas, average speeds are shown to be significantly slower. Further investigation revealed that a large proportion of speed test results in our data fell below the USC threshold of 2 Mbps, however the number of these slow connections was shown to be falling across the two annual extracts. Profiling speeds by measures of deprivation showed that average download speeds are higher in those areas ranking as more deprived. It would also appear that prevailing levels of deprivation in England are not necessarily a barrier to broadband access; although, there still remain those constraints of cost and behavioural choice that would result in variable uptake or use.

It is evident that densely populated urban areas record the highest average download speeds throughout the day, but suffer the most in terms of "bottlenecking" of data traffic during peak hours; conversely, small conurbations, and those which are rurally isolated, record far lower speeds, but, are impacted less at peak hours which is likely due to smaller populations. Splitting our data into annual extracts and mapping increases in average download speed at a district level revealed that large increases in speed were not as obviously clustered around urban conurbations, instead, high increases in download speed are more widely dispersed. Some rural areas were shown to be investing heavily in fibre broadband provision to ensure universal access to superfast connections.

The policy implications of this work are diverse, but initially show a requirement for further investment in broadband infrastructure to eliminate poor service below the minimum threshold set by UK Government, and to ensure more equitable delivery of next generation services. Analysis of the dynamics of broadband speeds throughout the day has highlighted implications for use, particularly so for those residents in areas that suffer from large performance fluctuations at peak hours. Upgrading traditional copper networks to carry broadband over fibre will help to eliminate such unstable connections and ensure robust connectivity both geographically and temporally, with longer-term benefits for the UK in an international economic context.

References

Arribas-Bel, D. (2014). Accidental, open and everywhere: emerging data sources for the understanding of cities. *Applied Geography, 49*, 45–53.

Bibby, P., & Shepherd, J. (2004). *Developing a new classification of urban and rural areas for policy purposes – The methodology*. Technical report. Office for National Statistics.

Bimber, B. (2000). Measuring the gender gap on the internet. *Social Science Quarterly, 81*(3), 868–876.

Broadbent, R., & Papadopoulos, T. (2013). Impact and benefits of digital inclusion for social housing residents. *Community Development, 44*(1), 55–67.

BT PLC. (2013). *BT products and services: BT infinity*. http://www.productsandservices.bt.com/products/broadband/infinity Accessed 14.11.13.

Bunyan, S., & Collins, A. (2013). Digital exclusion despite digital accessibility: empirical evidence from an English City. *Tijdschrift voor Economische en Sociale Geografie, 104*(5), 588–603.

Cooper, J. (2006). The digital divide: the special case of gender. *Journal of Computer Assisted Learning, 22*(5), 320–334.

DCMS. (2009). *Digital Britain final report*. http:/interactive.bis.gov.uk/digitalbritain/final-report/ Accessed 20.05.13.

Dutton, W., & Blank, G. (2013). *Cultures of the internet: The internet in Britain*. Technical report. Oxford Internet Institute.

Fairlie, R. (2004). Race and the digital divide. *Contributions to Economic Analysis and Policy, 3*(1), 155–194.

Federal Communications Commission. (2011). *Connecting America: National broadband plan*. http://www.broadband.gov/download-plan/ Accessed 13.05.13.

Freund, C. L., & Weinhold, D. (2004). The effect of the internet on international trade. *Journal of International Economics, 62*(1), 171–189.

Goodchild, M. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal, 69*, 211–221.

Graham, M. (2011). Time machines and virtual portals: the spatialities of the digital divide. *Progress in Development Studies, 11*(3), 211–227.

Graham, M., & Mann, L. (2013). Imagining a silicon savannah? Technological and conceptual connectivity in Kenya's BPO and software development sectors. *The Electronic Journal of Information Systems in Developing Countries, 56*, 1–19.

Haklay, M., Singleton, A. D., & Parker, C. (2008). Web mapping 2.0: the neogeography of the GeoWeb. *Geography Compass, 2*(6), 2011–2039.

Hargittai, E. (2002). Second-level digital divide: differences in people's online skills. *First Monday, 7*(4). http://firstmonday.org/issues/issue7_4/hargittai/index.html Accessed 17.05.13.

Koutroumpis, P. (2009). The economic impact of broadband on growth: a simultaneous approach. *Telecommunications Policy, 33*(9), 471–485.

Kraut, R., Kiesler, S., Boneva, B., Cummings, J., Helgeson, V., & Crawford, A. (2002). Internet paradox revisited. *Journal of Social Issues, 58*(1), 49–74.

Kyriakidou, V., Michalakelis, C., & Sphicopoulos, T. (2011). Digital divide gap convergence in Europe. *Technology in Society, 33*(3–4), 265–270.

Lane-Fox, M. (2009). *Champion for digital inclusion: The economic case for digital inclusion*. Technical report. Pricewaterhousecoopers LLP.

Lenhart, A., Madden, M., & Hitlin, P. (2005). *Teens and technology: Youth are leading the transition to a fully wired and mobile nation*. Technical report. Pew Internet.

Livingstone, S., & Helsper, E. (2007). Gradations in digital inclusion: children, young people and the digital divide. *New Media and Society, 9*(4), 671–696.

Loges, W. E., & Jung, J.-Y. (2001). Exploring the digital divide, internet connectedness and age. *Communication Research, 28*(4), 536–562.

Longley, P. A. (2003). Towards better understanding of digital differentiation, computers. *Environment and Urban Systems, 27*(2), 103–106.

Longley, P. A., & Singleton, A. D. (2009). Linking social deprivation and digital exclusion in England. *Urban Studies, 46*(7), 1275–1298.

Longley, P. A., Webber, R., & Li, C. (2008). The UK Geography of the e-society: a national classification. *Environment and Planning A, 40*(2), 362–382.

Norris, P. (2001). *Digital divide*. Cambridge: Cambridge University Press.

O3B. (2013). *Our technology*. http://www.o3bnetworks.com/o3b-advantage/our-technology Accessed 04.06.13.

Ofcom. (2010). *UK broadband speeds, May 2010: The performance of fixed-line broadband delivered to UK residential consumers* (Technical report).

Ofcom. (2012). *Infrastructure report: 2012 Update*. http://stakeholders.ofcom.org.uk/market-data-research/other/telecoms-research/broadband-speeds/infrastructure-report-2012/ Accessed 08.05.13.

Ofcom. (2013). *UK fixed-line broadband performance, May 2013: The performance of fixed-line broadband provided to UK residential consumers* (Technical report).

Parayil, G. (2005). The digital divide and increasing returns: contradictions of informational capitalism. *Information Society, 21*(1), 41–51.

Peter, J., & Valkenburg, P. M. (2006). Adolescents' internet use: testing the "disappearing digital divide" versus the "emerging digital differentiation" approach. *Poetics, 34*(4–5), 293–305.

Picot, A., & Wernick, C. (2007). The role of government in broadband access. *Telecommunications Policy, 31*(10), 660–674.

Pigg, K. E., & Crank, L. D. (2005). Do information communication technologies promote rural economic development? *Community Development, 36*(1), 65–76.

Prieger, J. E. (2013). The broadband digital divide and the economic benefits of mobile broadband for rural areas. *Telecommunications Policy, 37*(6–7), 483–502.

Prieger, J. E., & Hu, W. M. (2008). The broadband digital divide and the nexus of race, competition, and quality. *Information Economics and Policy, 20*(2), 150–167.

Singleton, A. D., & Longley, P. A. (2009). Geodemographics, visualisation, and social networks in applied geography. *Applied Geography, 29*(3), 289–298.

Singleton, A. D., & Spielman, S. E. (2013). *The past, present and future of geodemographic research in the United States and United Kingdom*. The Professional Geographer.

Sui, D., & Goodchild, M. (2011). The convergence of GIS and social media: challenges for GIScience. *International Journal of Geographical Information Science, 25*(11), 1737–1748.

van Dijk, J., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *The Information Society, 19*(4), 315–326.

Xavier, P. (2003). Should broadband be part of universal service obligations? *Info, 5*(1), 8–25.

Yiu, C., & Fink, S. (2012). *The superfast and the furious: Priorities for the future of UK broad-band policy*. http://www.policyexchange.org.uk/publications/category/item/the-superfast-and-the-furious-priorities-for-the-future-of-uk-broadband-policy Accessed 07.01.13.