What Drives Broadband Traffic?

Robert Kenny*, Charles Kenny, and Zack Gehan†‡

June 20, 2023

Abstract

Worldwide there is an ongoing policy and regulatory push to make very high speed broadband available as widely as possible. Underlying the policy interventions to support higher speeds is an implicit assumption that higher speeds will enable different (and socially valuable) use. In this paper we empirically test whether higher speed lines are associated with greater household data usage in the UK. We find that after allowing for demographic factors, higher speed in fact has a very limited relationship to traffic. This suggests that mid-speed broadband is not in fact a constraint on household usage (as measured by traffic), and thus the benefits of policy interventions to support higher speeds remain somewhat speculative.

Keywords: Broadband, internet traffic, demographics

- This paper uses extensive data on usage published by the UK telecoms regulator to explore the linkage between broadband speeds and traffic per line.
- We find that demographic factors are closely linked to traffic volumes:areas with larger households, younger population, and high percentages of those able to speak English are all associated with higher use, for example
- The link between the average broadband connection speed in an area and traffic volumes is considerably weaker, and weakens further once demographic factors are accounted for.

*R. Kenny (corresponding author): Communications Chambers, 3 Oakthorpe Rd, Oxford OX2 7BD, rob@commcham.com

†C. Kenny (ckenny@cgdev.org) and Z. Gehan (zgehan@cgdev.org): Center for Global Development 2055 L St NW Washington DC, 20036.

‡This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1 Introduction

Around the world there is an ongoing policy and regulatory push to make very high speed broadband available as widely as possible.[1] In Europe for example, the EU has set 'Digital Decade' targets[2] that call for gigabit broadband to be available to all households by 2030. In the UK, the government's "Project Gigabit" will deploy 5 billion GBP to support ultrafast rollout.[3] (Gigabit coverage had reached 70% in Europe[4] as of 2021 and the same percentage in the UK[5] as of September 2022).

Underlying the policy interventions to support higher speeds is an implicit assumption that higher speeds will enable different (and socially valuable) use cases. This might be because certain applications are only possible with higher speeds (the metaverse is sometimes offered as an example) or that other applications work much better with higher speeds.

Given the already widespread deployment of gigabit-capable networks, adoption of high speed broadband is already material (9% of lines in England and Wales, for example)[6]. This creates the possibility to empirically test whether higher speed lines are associated with different patterns of household usage.

One perspective on this is usage as measured by traffic—the number of gigabytes of traffic used per month. Broadly speaking, household traffic is driven by the following equation:

$$\begin{aligned}\text{Household traffic} &= \text{Number of people in the household}\\ &\times \text{ Average duration of usage per person of different applications}\\ &\times \text{ Bandwidth used by those applications}\end{aligned}$$

Duration of usage and bandwidth (in Mbps) are in turn functions of 'inherent' demand - what the occupants would like to do - and constraints. A key potential constraint is the bandwidth to the home. If this is insufficient for peak inherent demand, then applications may become frustrating to use, and occupants might simply reduce their time spent with the application. Alternatively, the application may still be usable, but in a degraded manner. For instance, an HD TV stream might 'drop back' to SD (which requires less bandwidth). In either of these scenarios, traffic will drop compared to the traffic generated by inherent demand.

This paper uses extensive data on usage published by Ofcom (the UK telecoms regulator) to explore the linkage between broadband speeds and traffic. We find that after allowing for demographic factors, higher speed has a very limited relationship to traffic. This suggests that mid-speed broadband is not in fact a constraint on household usage (as measured by traffic), and thus the benefits of policy interventions to support higher speeds are somewhat speculative.

[1] Strube Martins S, Wernick C. Regional differences in residential demand for very high bandwidth broadband internet in 2025. Telecomm Policy. 2021 Feb;45(1):102043

[2] EC, Europe's Digital Decade: digital targets for 2030 [accessed 1 February 2023]

[3] https://projectgigabit.campaign.gov.uk/

[4] EC, Digital Scoreboard [accessed 20 December 2022]

[5] https://www.ofcom.org.uk/research-and-data/multi-sector-research/infrastructure-research/connected-nations-2022

[6] Ofcom, Fixed performance census output area data

2 Literature Review

The existing literature on broadband speed focuses on output, employment and productivity. At a macro level, the cross-country economic growth literature lacks a convincing identification strategy, but with that significant caveat, predominantly suggests a positive link between faster broadband and economic performance.[7] However, this literature primarily relates to improvements in speed at the lower end of today's range - that is, improvements within the range of 0-30 Mbps. Such speeds are now very widely available. In the UK, for example, 97% of households are able to receive 30 Mbps or more, and 83% of households are using such connections.[8] Such literature is thus of limited relevance to the potential benefits of the upgrade to gigabit networks (either via fibre-to-the-premise [FTTP], or DOCSIS 3.1 upgrades to cable networks).

The existing literature on the rollout of FTTP has focused on the impacts on business and suggests a mixed picture. There is some evidence from the rollout of ultrafast (over 300 Mbps) in Italy that it may have boosted firm total factor productivity overall, while fostering consolidation at the expense of small firms who became more likely to exit.[9] Analysis of French data suggests broadband of over 30 Mbps is associated with a heterogeneous, but mostly positive effect on firm creation, while the association between faster speed and impact is contested.[10] New Zealand data suggests a link between FTTP adoption by firms and their productivity (but a negative link with employment).[11] Similarly, in Tennessee, availability of 100 Mbps+ is associated with lower unemployment, although faster broadband has no additional association.[12]

---

[7] On the positive side: Rohman, I., & Bohlin, E. (2013). Does Broadband Speed Really Matter for Driving Economic Growth? Investigating OECD Countries. International Journal of Management and Network Economics, 2, issue 4, p. 336-356. Gruber, H., Hätönen, J., & Koutroumpis, P. (2014). "Broadband access in the EU: An assessment of future economic benefits." Telecommunications Policy, 38(11), 1046-1058.(though note they suggest the benefits of faster speeds begin to level off after 1 Mbps, far below levels now very widely available) See also Briglauer, W., Dürr, N., & Gugler, K. (2021). "A retrospective study on the regional benefits and spillover effects of high-speed broadband networks: Evidence from German counties." International Journal of Industrial Organization, 74, 102677. On the less positive side: Mayer, Walter, Gary Madden, and Chen Wu. "Broadband and economic growth: a reassessment." Information Technology for Development 26.1 (2020): 128-145. Again, Briglauer, W., & Gugler, K. (2019). "Go for gigabit? First evidence on economic benefits of high-speed broadband technologies in Europe." JCMS: Journal of Common Market Studies, 57(5), 1071-1090 find a positive impact but suggest it is small enough that universal subsidy of superfast would have negative returns. Koutroumpis, P. (2019). "The economic impact of broadband: Evidence from OECD countries." Technological Forecasting and Social Change, 148, 119719 suggests an association between broadband speed and growth, but one that tops out at 10Mbps . Related, see Lobo, B. J., Alam, M. R., & Whitacre, B. E. (2020). "Broadband speed and unemployment rates: Data and measurement issues." Telecommunications Policy, 44(1), 101829 on employment and Xu, X., Watts, A., & Reed, M. (2019). "Does access to internet promote innovation? A look at the US broadband industry." Growth and Change, 50(4), 1423-1440.on innovation. Ford, G. S. (2018). "Is faster better? Quantifying the relationship between broadband speed and economic growth." Telecommunications Policy, 42(9), 766-777.

[8] Ofcom (2022), Connected Nations 2022

[9] The productivity impact is discussed in Carlo Cambini, Elena Grinza, Lorien Sabatino, "Ultra-fast broadband access and productivity: Evidence from Italian firms" International Journal of Industrial Organization, Volume 86, 2023, The consolidation effect is described in Cambini, C., & Sabatino, L. (2022). Digital highways and firm turnover. Available at SSRN 4119355

[10] Duvivier, C., Cazou, E., Truchet-Aznar, S., Brunelle, C., & Dubé, J. (2021). "When, where, and for what industries does broadband foster establishment births?" Papers in Regional Science, 100(6), 1377-1401. Hasbi, M. (2020). "Impact of very high-speed broadband on company creation and entrepreneurship: Empirical Evidence" Telecommunications Policy, 44(3), 101873

[11] Fabling, R., & Grimes, A. (2021). "Picking up speed: Does ultrafast broadband increase firm productivity?" Information Economics and Policy, 57, 100937.

[12] Lobo, B. J., Alam, M. R., & Whitacre, B. E. (2020). "Broadband speed and unemployment rates: Data and measurement issues" Telecommunications Policy, 44(1), 101829.

In Sweden data on 23 million observations from 2009 to 2017 suggest no simple link between broadband speed and income and unemployment, although some potential impacts on particular subgroups, while a study of 100 Mbps+ broadband in Germany found no link with firm creation.[13] A study of a UK superfast (over 30 Mbps) subsidy program suggests it was associated with small but significant employment and labor productivity gains on firms in the areas where the subsidies were provided. It also found mixed but broadly positive impacts for moves to higher speeds (100 Mbps+).[14] On the other hand, an analysis using UK Ofcom data suggests a "small but significant negative impact of high-speed (30 Mbps+) broadband on new business creation."[15]

The more limited literature on the impact of 100 Mbps+ speeds delivered by FTTP and DOCSIS 3.1 is to be expected, given that it has been deployed more recently. Moreover, any impact depends not just on deployment but uptake, and consumers (business and residential) have generally been slow to uptake ultrafast speeds. As of May 2022, 68% of UK consumers had access to gigabit services, but as of March of that year, just 8% had chosen to take services with speeds of 300 Mbps or higher.[16] In the EU in 2022, gigabit capable networks were available to 70% of households, but uptake of gigabit services was just 8%.[17] (One reason for this is that consumers appear to have limited willingness-to-pay for higher speeds. A recent conjoint-based survey in the US found that the marginal value of an incremental 1 Mbps fell from $6.81 at 5 Mbps to $0.34 at 100 Mbps.[18])

A further limitation of the literature is that much of it focuses on the impact on firms, and/or the effect of usage by firms of higher speeds. However, policy interventions are primarily focused on domestic availability and uptake (albeit widespread coverage will also benefit businesses). The European Commission's 'Digital Decade' target mentioned above is for 100% availability to households, for example. While there have been some targeted interventions (for connectivity to schools and hospitals, for example), these have been much rarer.

Thus much policy is based on a view that households have (or will have) insufficient bandwidth absent gigabit networks. Sometimes such views are based on an assumption that requirements will climb ever upwards as traffic grows, which we discuss further below. Such views may also be based on explicit forecasts of bandwidth demand.

These forecasts are complicated by a number of factors. Households are highly heterogeneous—to take one simple example, an older person living alone is likely to need far less bandwidth than a family of five. Further, these forecasts assess technical demand, rather than demand in the standard economic sense (moderated by willingness-to-pay). In that regard, they have to make assumptions about need for a particular bandwidth (If - say - 100 Mbps rather than 50 Mbps will save ten

[13] Hasbi, Maude, and Erik Bohlin. "Impact of broadband quality on median income and unemployment: Evidence from Sweden." (2021). Sarachuk, K., Missler-Behr, M., & Hellebrand, A. (2021, August). "Ultra High Speed Broadband Internet and Firm Creation in Germany." In Innovations in Digital Economy: Second International Scientific Conference, SPBPU IDE 2020, St. Petersburg, Russia, October 22–23, 2020, Revised Selected Papers (pp. 40-56). Cham: Springer International Publishing.

[14] Ipsos, M. O. R. I. "Superfast Broadband Programme Evaluation. Annex B: Economic Impacts." (2018). Interestingly, a well-identified analysis in Germany found no impact on employment of DSL rollout Czernich, N. (2014). "Does broadband internet reduce the unemployment rate? Evidence for Germany." Information Economics and Policy, 29, 32-45.

[15] Chen, P., Oughton, E. J., Tyler, P., Jia, M., & Zagdanski, J. (2020). "Evaluating the impact of next generation broadband on local business creation." arXiv preprint arXiv:2010.14113.. The paper uses roll-out data for Asymmetric Digital Subscriber Line (ADSL) technology by British Telecom (BT) between 1999-2007 as instrumental variable for broadband rollout 2011-2015 (percentage of premises by postcode with NGA)

[16] Ofcom (2022), Connected Nations 2022. Ofcom (2022), UK Home Broadband Performance

[17] EC, DESI by components

[18] Rabbani, Maysam and Bogulski, Cari and Eswaran, Hari & Hayes, Corey, Willingness to Pay for Internet Services (February 27, 2023)

minutes on a game download, does a household 'need' 100 Mbps?). And since such exercises are generally forward looking, they inevitably depend on subjective assumptions.

These factors contribute to a very wide range of outcomes from forecast exercises. Ofcom, for instance, published a 2018 study by WIK that 40% of UK households would demand 1 Gbps by 2025.[19] By contrast, an earlier study by Kenny & Broughton for BSG anticipated median demand of just 19 Mbps in 2023.[20]

Interlinked with required bandwidth (peak flow rates) is traffic (total data moved over, say, a month). These are sometimes used interchangeably, and often traffic growth is assumed to require ever higher bandwidth. For instance, according to the UK's House of Commons Public Accounts Committee:

> "As of November, 2021, 97% of premises in the UK have access to superfast broadband, which has a download speed of at least 30 Mbps. This speed is adequate for most household use, but global internet traffic is growing by around 40% each year. To ensure the UK has access to the broadband speeds necessary to match future demand, the [government announced] ... "Project Gigabit"."[21] [Note that the claimed 40% growth is far higher than actual UK growth rates].

However, the link between traffic and bandwidth is complex. As UK operator BT has put it:

> "There is no direct relationship between data speeds and total traffic. The latter has been increasing significantly as customers spend more time online. Customers are using fixed-line connections ever more intensively, streaming HD content requires a connection speed of around 5Mbps, and people can double their data consumption by watching twice as much content, but without demanding higher speeds. Demand for higher usage does not therefore directly feed into demand for higher speeds, and the drivers for each are distinct albeit both reflect digitalisation in its many forms."[22]

The simple equation for household traffic we laid out above suggests it is a product of household size, duration of usage per person and bandwidth required by that usage (note that this excludes traffic from the 'internet of things', such as smart meters, but generally speaking such traffic is modest). To BT's point, traffic can be increased by a growth in the duration of usage, without the required bandwidth increasing. That said, increased duration does increase the probability of overlapping usage, and so may cause an increase in peak demand.

There are numerous factors that will drive variation in the inputs to this equation. Household size obviously may vary, but so too will the usage of household members. For example, Ofcom suggests that the average UK Internet user aged 18-24 spent 5:06 hours a day on the Internet compared to 2:58 for those 55 and older (and there will be further variation around these averages).[23] Employment status might be expected to have an impact too—time use surveys suggest unemployed people in the UK spend 222 minutes a day on computing and other mass-media or

[19] WIK (2018), The Benefits of Ultrafast Broadband Deployment. For a critical review of this paper, see Kenny, R (2018), Examining WIK's bandwidth demand forecast

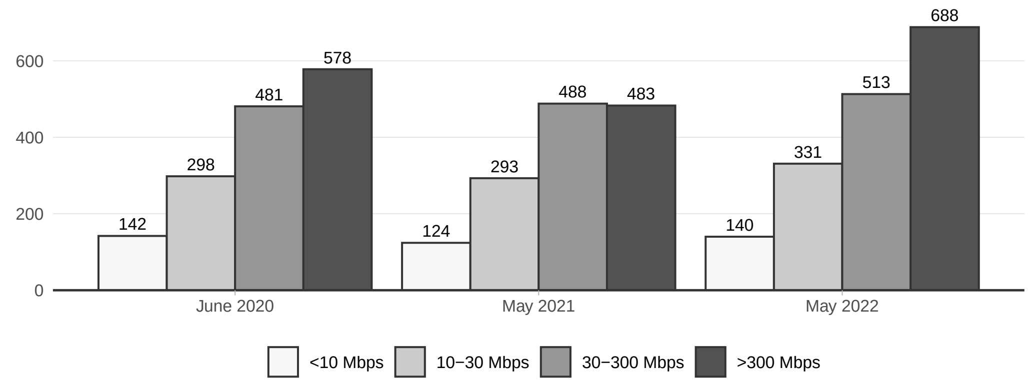
[20] Kenny, R. & Broughton T. (2013), Domestic demand for bandwidth

[21] House of Commons Public Accounts Committee (2022), Delivering gigabit capable broadband

[22] BT (2018), DCMS Future Telecoms Infrastructure Review: Call for Evidence. BT's response (January 2018) - Annex 2 Sources of Demand for ultrafast networks

[23] https://www.ofcom.org.uk/__data/assets/pdf_file/0023/238361/online-nation-2022-report.pdf

Figure 1: Traffic Per Line (in GB), United Kingdom



watching TV/Video, compared to 127 minutes for the employed.[24] Available devices will also have an impact—a person with a 4K TV will use much more bandwidth to watch a given film than someone watching it on a smartphone.

Regardless, if higher speed has benefit, then we might expect higher speed lines to have higher traffic. Lower speed lines—which made certain use cases impossible or difficult—might be expected to reduce usage, either by reducing duration of usage (through frustration or simple impossibility) or by reducing bandwidth (by forcing a downgrade in the quality of the application, from HD to SD, say). Certainly, if (per WIK) a high percentage of households need gigabit speeds, then we would expect the average 112 Mbps line speed in use in the UK to be a major constraint.[25]

One perspective on this issue comes from the relationship between line speed and traffic, for which there is some publicly available data in both in both the US and the UK. Ofcom has published high level data for May 2022 (Figure 1).[26]

These data do suggest a positive correlation between speed and traffic, although not a linear one. (For example, the traffic difference between lines of ¡10 and 10-30 Mbps is similar to that between 10-30 and 30-300 Mbps lines).

Data for the US from OpenVault (a provider of broadband IT systems) found a broadly similar pattern (Figure 2).[27]

However, a challenge with interpreting both the Ofcom and the Openvault data is that even if higher speeds are correlated with higher traffic, it is not clear that those speeds cause (or enable) higher traffic. There are two other possible explanations.

First, households with high usage may select higher speeds, whether or not they actually need them. An assessment of required bandwidth is complex, and likely well beyond most consumers. Given this, consumers are likely to rely either on heuristics (for example, 'I think I'm a heavy user, so I probably need the highest speed offered'), or on guidance given by broadband providers. For

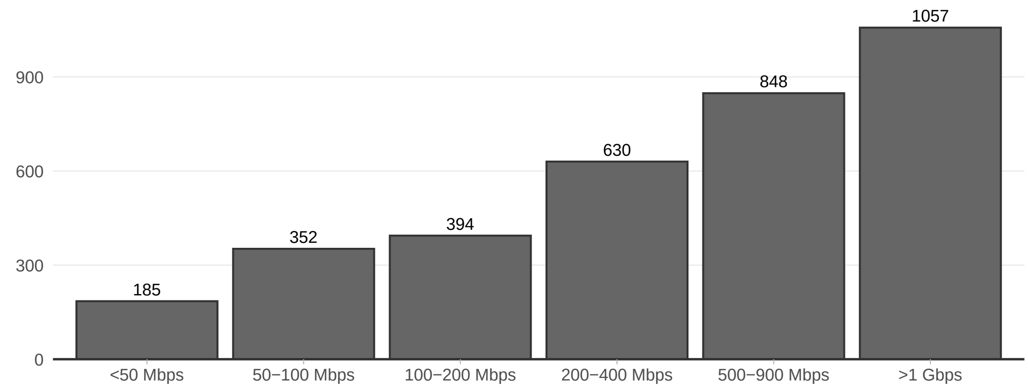
[24] Hoang, Thi Truong An; Knabe, Andreas (2019) : "Time use, unemployment, and well-being: an empirical analysis using British time-use data", CESifo Working Paper, No. 7581, Center for Economic Studies and ifo Institute (CESifo), Munich Krueger, A. B., & Mueller, A. I. (2012). "The lot of the unemployed: a time use perspective." Journal of the European Economic Association, 10(4), 765-794 suggests a similar pattern holds across rich countries.

[25] Ofcom (2022), Connected Nations 2022

[26] Analysis of data from 'Data Downloads' associated with Ofcom's Connected Nations reports

[27] Openvault (2022), Broadband Insights Report (OVBI) 4Q21

Figure 2: Traffic Per Line (in GB), United States



instance, BT's marketing[28] says:

> "Full Fibre: Our next generation of fibre broadband gives you an ultrafast fibre optic broadband connection 25x faster than superfast fibre. Enjoy high quality video calls with friends and family, stream your favourite 4K shows on multiple devices, download HD movies and boxsets faster than ever and upload videos and high-res photos."
>
> "Fibre: Our Fibre broadband gives you a superfast fibre optic broadband connection with average speeds of 67Mbps. Everyone in your household can get online at the same time, you can stream and download in HD on lots of devices at once and play games online."

In practice, high-quality video calls, streaming multiple 4K shows and uploading videos and photos are all possible using basic fibre with average speeds of 67 Mbps. However, BT is steering those using these high traffic applications to full fibre, and thus full fibre might be correlated with higher traffic simply for this reason. (Other broadband providers have similar marketing material).

Second, both higher speeds and higher traffic may be correlated with a third factor. For example, higher income households might be more willing to pay for higher speed connections and might be more likely to have 4K TV sets that trigger higher data consumption. Or these higher income households might on average have more occupants.

Another perspective on traffic and line speeds comes from international comparisons, which suggests at the national level there is essentially no correlation between per-household traffic and the FTTP share of overall broadband (a proxy for higher speeds). Rather, traffic appears to be linked to drivers such as frequency of internet use and video-on-demand penetration.[29]

As far as we have been able to identify, there is no academic literature that addresses the drivers of traffic at the household level. In particular, we believe there is nothing that addresses the impact of speed as a driver, once other factors (such as household size) have been taken into account. That is the gap this paper seeks to address: to find whether speeds currently available may be

[28] BT, Broadband speed: What is it and how does it work? [accessed 13 March 2023]
[29] Kenny, R (2021). Patterns of fixed traffic growth, 2021

constraining usage, and hence whether there are likely to be material benefits from an upgrade in speed.

Primarily we use a standard OLS regression technique to present conditional correlations between demographic factors, connection speeds and data usage. We make the assumption that recent data usage is unlikely to have been a major causal factor in demographic characteristics like household size or income while (demand for) data usage and connection speed may well be co-determined, but connection speed may in some cases be a binding constraint on use. As a robustness exercise, we apply propensity score matching, widely used to separate out the impact of 'treatments' (in this case fast internet connections) on outcomes (in our case data usage) abstracting from causal factors that might explain likelihood of treatment and outcomes (in this case demographic variables).[30]

3 Data Description

The basis for our analysis is disaggregated data published by Ofcom on monthly data usage (traffic) of fixed broadband.[31] The underlying data is sourced from UK broadband providers and relates to May 2022.

For each area treated, this dataset provides (amongst other datapoints):

- The number of lines in use of different speeds, 0-2 Mbps, 2-5 Mbps and so on, up to >300 Mbps
- Overall average and median download and upload speeds for the area
- Overall average and median traffic per month per line
- Average traffic per month for lines of each speed category

Ofcom provides these data at several degrees of geographic granularity, from local/unitary authority (with an average of 66,000 lines) to postcode (15 lines). To strike a balance between sufficient granularity to pick up meaningful variation in drivers (such as area age profiles) on the one hand, and a manageable data set without undue statistical noise in small areas, we make use of Ofcom's data for 'Output Areas' (OAs). There are 239,000 such areas in the UK, and 80% of them have between 110 and 139[32] households. The average number of broadband lines per OA is approximately 100.

To these data we have appended a range of geographic and demographic data sourced from ONS (the Office of National Statistics), discussed further below. These data derive from the 2021 UK census. However, this additional data was only available for England and Wales at the time of our analysis, and thus we have excluded from the Ofcom data set those OAs in Scotland and Northern Ireland. This leaves 188,833 records.[33]

We additionally excluded from the data set firstly those OAs with fewer than 51 or more than 999 lines (these outliers represented 1.6% of the data set). Given that OAs typically have around 125 households, OAs with a number of lines either well below this or far above are likely to be

[30] See Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011 May;46(3):399-424

[31] Ofcom, Connected Nations 2022: data downloads

[32] ONS, Output areas

[33] The indicator 'cannot speak English well' is drawn from from the 2021 England and Wales census questions on language proficiency where 'cannot speak English well' was one of the possible responses.

anomalous. For instance, a high number of lines might perhaps represent heavy business use, and a low number of lines could represent a large number of unoccupied premises (such as holiday homes).

Secondly we excluded OAs with average usage of 0 or over 1000 GB (this excluded a further 0.4% of the data set). Again, such OAs are likely to be anomalous. Zero usage for an entire OA might perhaps be an as-of-yet unoccupied new build, say. Very heavy usage might indicate the presence of (say) intense commercial user(s) that happen to be using a mass market product.

After these exclusions, 185,177 records remained for our analysis. Summary statistics are presented in Table 1.

Table 1: Summary Statistics

	Min	Median	Mean	Max	SD
Percent connections under 30 Mbps	0.0	11.9	16.3	100.0	14.2
Percent connections 30 to 300 Mbps	0.0	78.9	74.9	100.0	15.0
Percent connections over 300 Mbps	0.0	8.0	8.8	100.0	8.7
Average download speed (Mbps)	2.2	112.9	112.2	1019.4	53.0
Median download speed (Mbps)	1.1	71.8	71.1	1000.0	36.5
Average data usage (GB)	1.0	475.0	486.3	999.0	141.4
Median data usage (GB)	0.0	311.0	322.3	1009.0	129.3
Age 0-24 percent	0.0	16.8	17.1	59.3	6.0
Age 25-64 percent	0.9	52.4	52.6	94.9	7.8
Age 65+ percent	0.0	18.0	19.6	99.1	11.3
Log(popn density)	-0.1	8.4	8.0	14.5	1.5
HH size	1.0	2.3	2.4	5.9	0.4
Student percent	0.0	5.6	7.0	97.3	6.6
Unemployed percent	0.0	2.5	2.9	37.2	2.0
Retired percent	0.0	21.0	22.3	96.3	12.2
Economically inactive percent	0.0	10.1	12.2	80.9	7.2
Percent works mainly from home	0.0	13.6	15.1	75.1	8.8
Percent managerial	1.1	32.4	33.1	86.7	13.7
Percent cannot speak English well	0.0	0.6	1.6	38.7	2.7

Figure 3 displays the distribution (kernel density) of average monthly traffic at the OA level across our 185,177 OAs. Note this is a tighter distribution of outcomes than would be seen at the individual household level, since each OA data-point is an average of approximately 100 households.

Figure 4 displays the distribution of average download speeds at the OA level across our 185,177 OAs.

Figure 3: Distribution of Average Data Usage

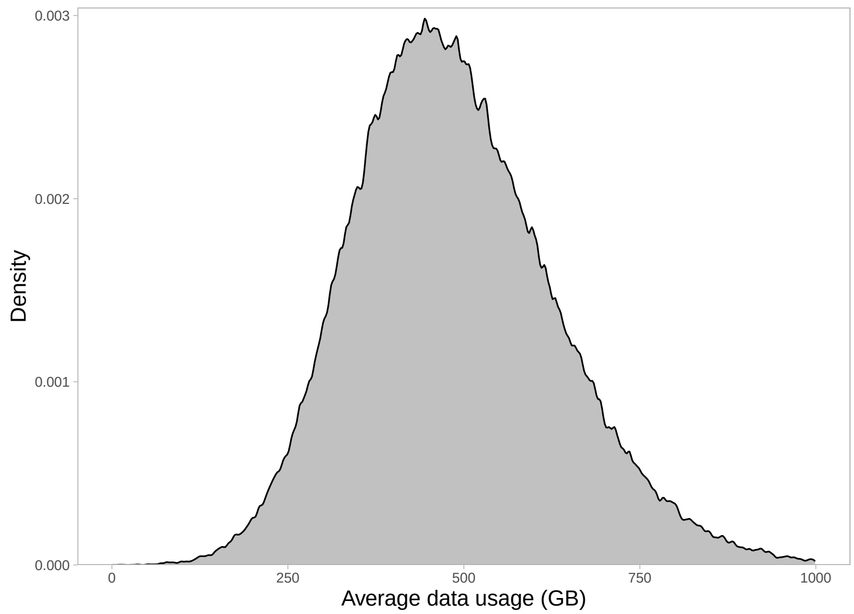
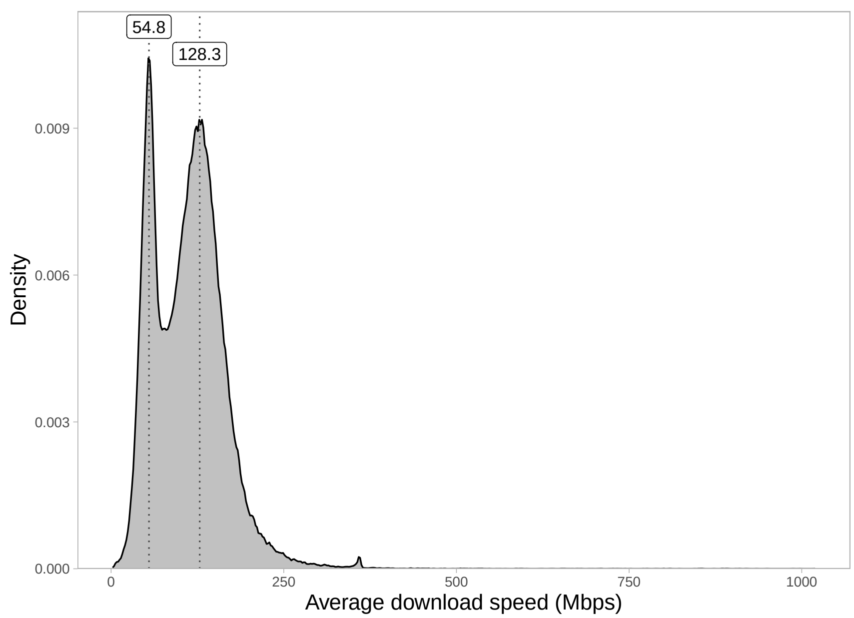


Figure 4: Distribution of Average Download Speed



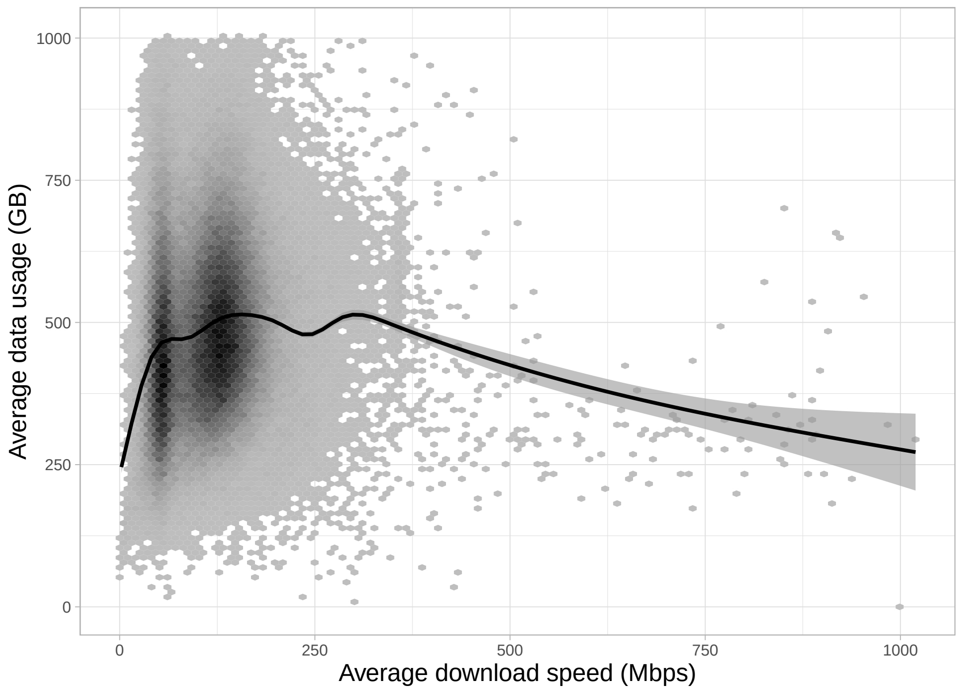


Figure 5: Average Data Usage vs Average Download Speed.

4 Analysis

As a first step in the analysis, we present graphs that display the bivariate relationship between average monthly traffic and characteristics of connection speeds and area populations. The lines represent generalized additive model (GAM) estimations[34] of the relationship of the variables, with the hexagons displaying the density of data points (darker hexagons represent more data points in that region).

Figure 5 shows the simple relationship between average download speed and average monthly traffic. It trends sharply upward at low speeds until around the 30 Mbps mark, where the relationship distinctly flattens. Figure 6 explores the link between data usage and the proportion of connections slower than 30 Mbps and suggests as more connections drop below this threshold there is a consistent association with lower monthly traffic. On the other hand, Figure 7 looks at traffic against the percentage of connections 300 Mbps or faster and finds essentially no relationship.

Figure 8 looks at the relationship between the percentage of population in an area between the ages of 0-24 and monthly traffic: through the considerable majority of datapoints there is a consistent and steep relationship between usage and the share of young people. That more young people in an area drives greater traffic is consistent with the Ofcom data showing that younger people spend significantly more time online. This also points to the potential importance of demographic factors in explaining traffic. We turn to regression analysis to explore the link between line speeds,

[34]GAM lines are similar to the more familiar Locally Estimated Scatterplot Smoothing (LOESS) lines, but are less intensive to calculate with the large number of data points (185k).

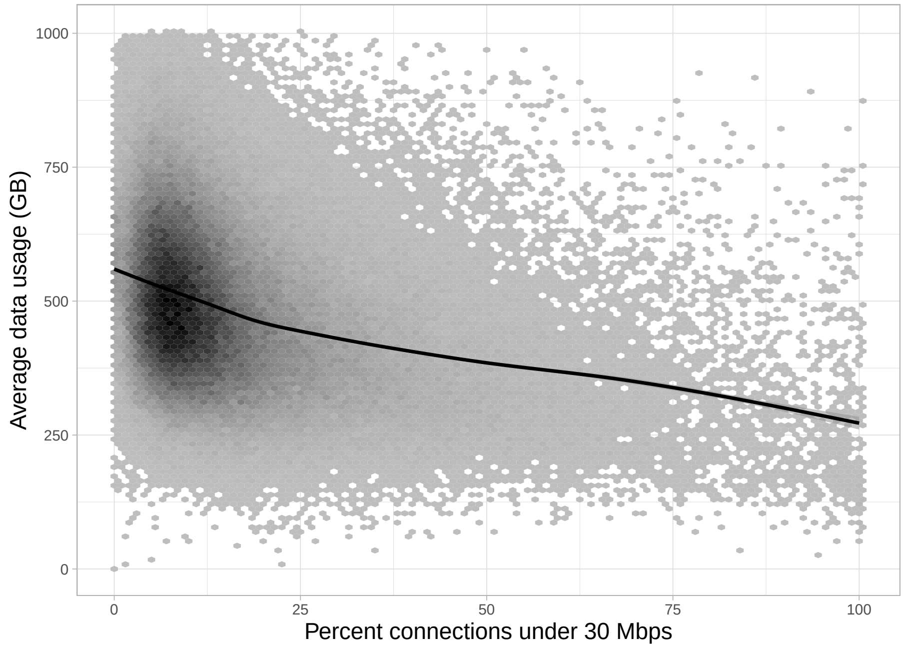


Figure 6: Average Data Usage vs % of connections which are slower than 30 Mbps.

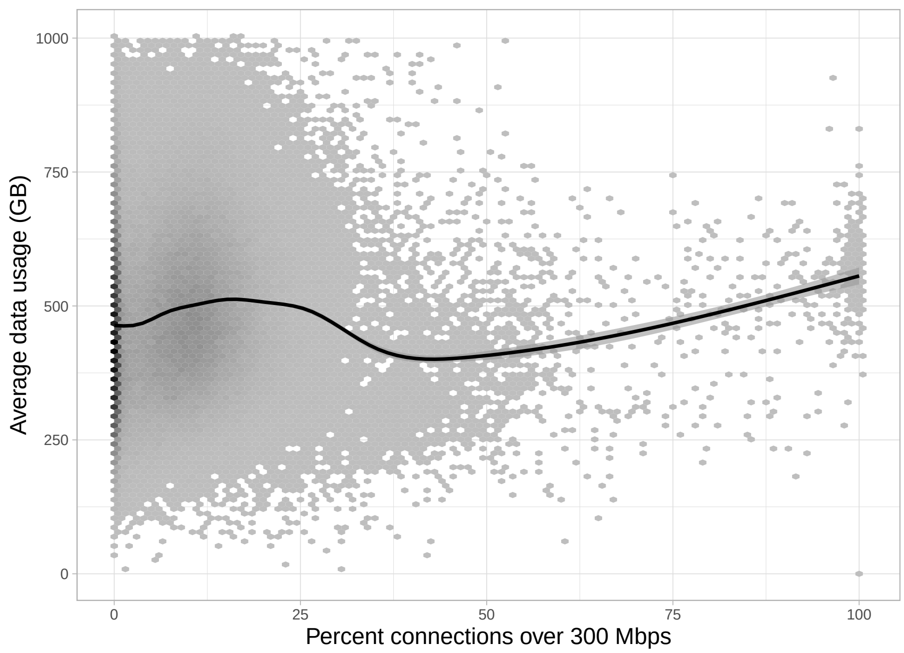


Figure 7: Average Data Usage vs % of connections which are faster than 300 Mbps.

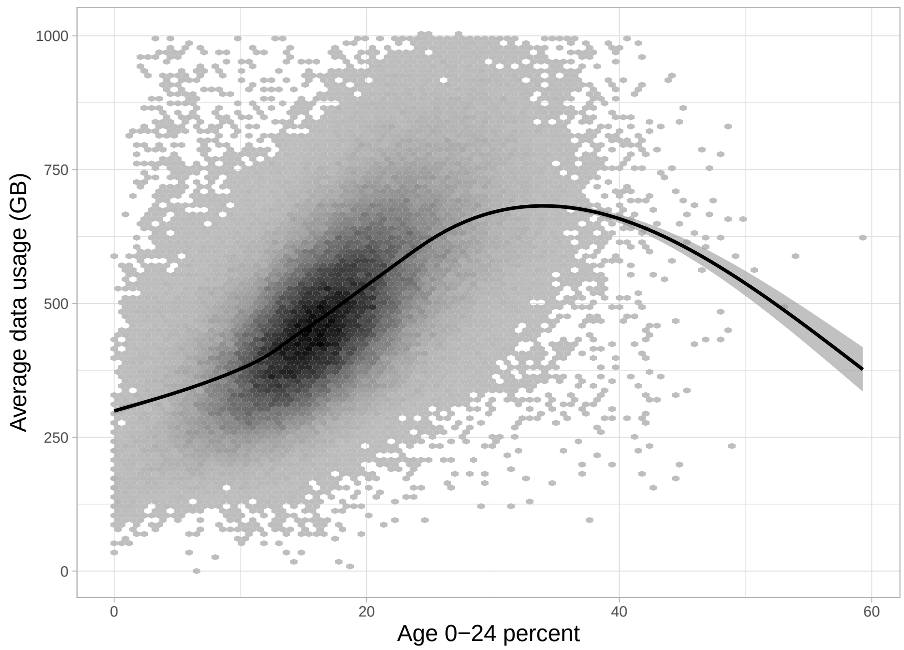


Figure 8: Average Data Usage vs % of population aged 0-24.

demographic, employment and household factors and traffic at greater depth.

The relationship between average download speeds and the percentage of households on slower (less than 30 Mbps) or fast (greater than 300 Mbps) connections is unsurprisingly strong. Those two variables alone are associated with 84 percent of the variation of average download speeds across output areas (Table 2).

In contrast, demographic and density variables explain very little of the variation of download speeds: age, employment and managerial status, average household size, the ability to work from home, language abilities and population density combined are associated with only 16 percent of the variation in average download speeds (and 8 percent of median download speeds, Tables 2 and 3). These same factors are associated with ten percent of the variation across output areas in the percentage of connections that are above 300 Mbps in speed (Table 4). One reason for a relatively weak linkage between demographics and line speed is that higher speeds are not universally available—regardless of its demographics, a household in a rural area will likely be unable to choose a 300 Mbps service, for example.

Turning to monthly traffic, Table 6, the situation is reversed. Fast connections (and average download speed) is very weakly connected with usage, while household, demographic and employment factors are strongly associated with use. The percentage of households on slower (less than 30 Mbps) or fast (greater than 300 Mbps) connections explains only about ten percent of the variation in average traffic, and the coefficient on connections over 300 Mbps suggests a small, insignificant negative relationship with average traffic.

It should be noted that the coefficient on slow connections suggests moving from 100 to zero percent slow connections would be associated with a 325 GB increase in traffic, which is about the

Table 2: Regressions on Average Download Speed in Mbps.

	Average download speed (Mbps)	
	(1)	(2)
% con'ns <30 Mbps	-0.706^{***} (0.004)	
% con'ns $\geq$ 300 Mbps	5.224^{***} (0.006)	
Age 25-64 %		-0.011 (0.035)
Age 65+ %		-0.539^{***} (0.054)
Log(Popn density)		7.477^{***} (0.084)
Economically inactive %		-0.085^{***} (0.025)
Student %		-0.063^{**} (0.027)
Retired %		-0.122^{**} (0.051)
Unemployed %		-0.301^{***} (0.075)
HH size		27.059^{***} (0.418)
% Managerial		-0.081^{***} (0.021)
% Cannot speak English well		-1.523^{***} (0.055)
% Works mainly from home		0.935^{***} (0.032)
Constant	77.829^{***} (0.098)	-4.346 (3.299)
Observations	185,177	185,177
Adjusted R^2	0.842	0.161
Note:		*p<0.1; **p<0.05; ***p<0.01

same as median traffic usage across the UK. Slow connections really may be a significant barrier to traffic consumption. At the same time, moving from 100 percent of the population aged 65 or older to 100 percent under 25 in an output area would be associated with a 1,146 GB increase in average traffic.

Demographic factors are much more closely associated with monthly traffic than connection and download speeds. Household, demographic and employment factors combined are associated with about 62 percent of the variation across output areas in average traffic. Once demographic factors are accounted for, adding connection speeds is only associated with one to three percentage points more of the variation in traffic across output areas (Table 6).

The data suggest that areas with a high share of economically inactive people see a particularly high level of traffic. Population-dense areas and areas with large average household sizes are also associated with higher traffic per connection. A large population of non-English speakers is associated with lower use.

The two variables of aged 25-64 and aged 65 plus alone are associated with 44 percent of the variation across output areas in traffic, compared to two percent for average download speed (Tables 7 and 6). Very approximately, adding a ten percent share of the population aged 25-64 or 65 plus, which is about one standard deviation in these variables, is associated with an approximate 100 GB fall in average data use (a little more than two thirds of the standard deviation of this variable). It is clear that what explains traffic variation across England and Wales in 2022 is not the speed of connections but the demographic factors, and particularly age.

Table 3: Regressions on Median Download Speed in Mbps.

	Median download speed (Mbps)	
	(1)	(2)
% con'ns <30 Mbps	-0.682^{***} (0.004)	
% con'ns $\geq$ 300 Mbps	2.755^{***} (0.007)	
Age 25-64 %		-0.107^{***} (0.025)
Age 65+ %		-0.225^{***} (0.039)
Log(Popn density)		3.853^{***} (0.061)
Economically inactive %		-0.234^{***} (0.018)
Student %		-0.099^{***} (0.020)
Retired %		-0.369^{***} (0.037)
Unemployed %		-0.321^{***} (0.054)
HH size		4.593^{***} (0.302)
% Managerial		0.065^{***} (0.015)
% Cannot speak English well		0.004 (0.040)
% Works mainly from home		0.104^{***} (0.023)
Constant	58.047^{***} (0.110)	48.438^{***} (2.384)
Observations	185,177	185,177
Adjusted R^2	0.576	0.079
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 4: Effect of demographic variables on Percent of Connections faster than 300 Mbps.

	Percent connections over 300 Mbps
Age 25-64 %	-0.045^{***} (0.006)
Age 65+ %	-0.044^{***} (0.009)
Log(Popn density)	0.671^{***} (0.014)
Economically inactive %	-0.027^{***} (0.004)
Student %	-0.034^{***} (0.005)
Retired %	-0.084^{***} (0.009)
Unemployed %	-0.072^{***} (0.013)
HH size	4.057^{***} (0.071)
% Managerial	-0.001 (0.004)
% Cannot speak English well	-0.307^{***} (0.009)
% Works mainly from home	0.109^{***} (0.005)
Constant	-1.418^{**} (0.561)
Observations	185,177
Adjusted R^2	0.100
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5: Effect of demographic variables on Percent of Connections slower than 30 Mbps.

	Percent connections under 30 Mbps
Age 25-64 %	0.046*** (0.009)
Age 65+ %	0.307*** (0.014)
Log(Popn density)	−4.077*** (0.022)
Economically inactive %	0.114*** (0.007)
Student %	0.194*** (0.007)
Retired %	−0.106*** (0.013)
Unemployed %	0.262*** (0.019)
HH size	−3.062*** (0.108)
% Managerial	−0.103*** (0.005)
% Cannot speak English well	0.254*** (0.014)
% Works mainly from home	0.166*** (0.008)
Constant	47.075*** (0.854)
Observations	185,177
Adjusted R^2	0.217
Note:	*p<0.1; **p<0.05; ***p<0.01

In these regressions, we note that that the coefficient for the variable 'percent of connections $\geq$ 300 Mbps' is negative - that is, these higher speeds are associated with lower traffic, albeit weakly. This is counterintuitive. One possible explanation relates to in-home wifi. In practice, this is often the binding constraint on broadband speed, with the wireless capacity or reliability being less than those of the fixed connection to the home.[35] However, this is not well understood by consumers. It is possible that consumers frustrated by poor internet experience have been more likely to upgrade to the fastest speeds, unaware that their frustrations are in fact caused by wifi issues. Such consumers might perhaps have lower traffic than other customers with better wifi. Another reason for this might be that a direct measure of income is omitted from our analysis by data constraints. Higher-income consumers may be more likely to subscribe to the fastest data packages, but less likely to be using data intensive applications like gaming and streaming because of other leisure options.[36]

[35]For a detailed discussion of the impact of wifi on home broadband performance, see Sanhueza, Antonia and Richardson, Marc (2022), Is Your WiFi Limiting Your Home Internet Performance?

[36]The American Time Use Survey suggests that in 2021 high-earning individuals spent thirty minutes less per day than low-earning individuals watching TV and using the computer for leisure. Source: https://www.bls.gov/news.release/atus.t11A.htm

Table 6: Regressions on Average Data Usage in GB.

	Average data usage (GB)				
	(1)	(2)	(3)	(4)	(5)
% con'ns <30 Mbps	-3.258^{***}				-1.647^{***}
	(0.022)				(0.016)
% con'ns $\geq$ 300 Mbps	-0.036				-1.831^{***}
	(0.037)				(0.024)
Average download speed (Mbps)		0.348^{***}		-0.197^{***}	
		(0.006)		(0.004)	
Age 25-64 %			-3.888^{***}	-3.890^{***}	-3.893^{***}
			(0.063)	(0.063)	(0.061)
Age 65+ %			-7.572^{***}	-7.678^{***}	-7.148^{***}
			(0.097)	(0.096)	(0.093)
Log(Popn density)			16.345^{***}	17.820^{***}	10.861^{***}
			(0.152)	(0.154)	(0.160)
Economically inactive %			1.781^{***}	1.765^{***}	1.919^{***}
			(0.046)	(0.045)	(0.044)
Student %			-3.544^{***}	-3.557^{***}	-3.288^{***}
			(0.049)	(0.049)	(0.048)
Retired %			-0.499^{***}	-0.523^{***}	-0.828^{***}
			(0.092)	(0.092)	(0.089)
Unemployed %			-0.851^{***}	-0.910^{***}	-0.550^{***}
			(0.134)	(0.134)	(0.129)
HH size			99.187^{***}	104.524^{***}	101.574^{***}
			(0.753)	(0.757)	(0.731)
% Managerial			-2.039^{***}	-2.055^{***}	-2.212^{***}
			(0.038)	(0.038)	(0.036)
% Cannot speak English well			-11.349^{***}	-11.649^{***}	-11.491^{***}
			(0.099)	(0.099)	(0.096)
% Works mainly from home			-1.931^{***}	-1.747^{***}	-1.458^{***}
			(0.058)	(0.058)	(0.056)
Constant	539.681^{***}	447.198^{***}	605.809^{***}	604.952^{***}	680.726^{***}
	(0.620)	(0.763)	(5.948)	(5.912)	(5.771)
Observations	185,177	185,177	185,177	185,177	185,177
Adjusted R^2	0.107	0.017	0.617	0.622	0.646
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 7: Regressions on Average Data Usage in GB.

	Average data usage (GB)		
	(1)	(2)	(3)
Age 25-64 %	−7.922***	−8.504***	−3.888***
	(0.042)	(0.056)	(0.063)
Age 65+ %	−10.903***	−13.190***	−7.572***
	(0.029)	(0.097)	(0.097)
Economically inactive %		0.598***	1.781***
		(0.048)	(0.046)
Student %		−5.582***	−3.544***
		(0.049)	(0.049)
Retired %		0.674***	−0.499***
		(0.099)	(0.092)
Unemployed %		−2.241***	−0.851***
		(0.141)	(0.134)
% Managerial		−1.547***	−2.039***
		(0.040)	(0.038)
% Works mainly from home		−1.955***	−1.931***
		(0.062)	(0.058)
Log(Popn density)			16.345***
			(0.152)
HH size			99.187***
			(0.753)
% Cannot speak English well			−11.349***
			(0.099)
Constant	1,116.734***	1,295.844***	605.809***
	(2.622)	(4.073)	(5.948)
Observations	185,177	185,177	185,177
Adjusted R^2	0.442	0.558	0.617
Note:		*p<0.1; **p<0.05; ***p<0.01	

5 Robustness

As robustness exercises, we repeated the regression analysis using median data use and median download speed, as well as dropping areas that reported average download speeds of greater than 250 Mbps (see Appendix A). Both approaches had the effect of reversing the coefficient on percentage of connections over 300 Mbps from negative to positive in the regression only accounting for that variable and the percentage of connections under 30 Mbps, but otherwise left the results similar in magnitude, significance and R-squared.

As a further robustness exercise, we employ propensity score matching. In this case, the purpose is to determine the effect of faster connectivity on data usage comparing output areas with very similar demographics. We define four 'treatments': (i) average output area download speed is greater than 100 Mbps, (ii) percent of output area connections above 300 Mbps is above national average (8.78%) (iii) percent of output area connections above 300 Mbps is above 15%; (iv) percent of output area connections above 30 Mbps is above national average. We match 'treatment' output areas to 'control' areas on demographic variables (the full suite used in Table 4 above) using the nearest neighbor approach implemented through Stata's `teffects psmatch` command.[37]

Table 8 reports the results: the impact of our 'treatment' on data use. Areas with average connection speeds faster than 100 Mbps saw 24GB lower data usage than areas with similar demographic conditions but average connection speeds slower than 100 Mbps; (ii) areas with an above average proportion of connections above 300 Mbps saw 19GB lower data usage than areas with similar demographic conditions but a below average proportion of connections above 300 Mbps; (iii) areas with above 15 percent of connections above 300 Mbps saw 20GB lower data usage than areas with similar demographic conditions but below 15 of connections above 300 Mbps; (iv) areas with an above average proportion of connections above 30 Mbps saw 31GB higher data usage than areas with similar demographic conditions but a below average proportion of connections above 30 Mbps.[38]

These results broadly match what we have seen above: there is evidence that slow connection speeds (below 30 Mbps) are a (limited) constraint on data use. There is no evidence that fast data speeds (above 100 Mbps) are a constraint on use—indeed they are associated with slightly lower data usage.

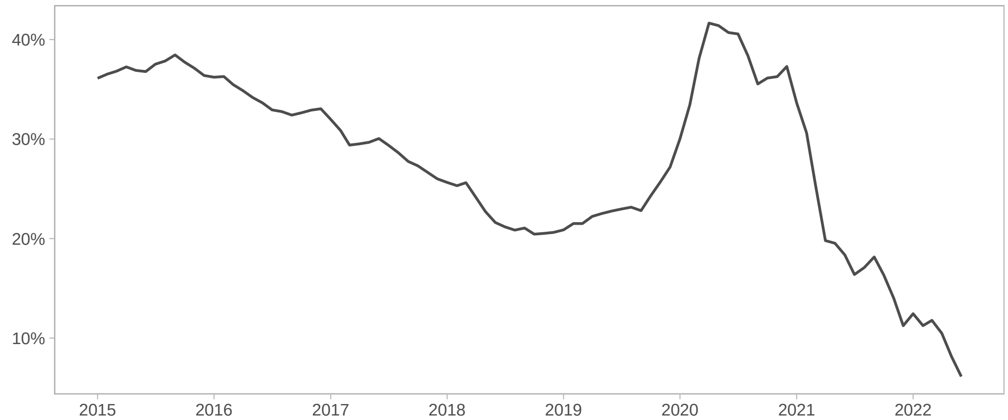
Table 8: Effect sizes from Propensity Score Matching.

	(i)	(ii)	(iii)	(iv)
Effect	−23.59	−19.31	−20.34	31.53
Robust SE	0.66	0.59	1.03	0.66
P-value	0	0	0	0

[37] We were able to match all treatments with a caliper of 0.03 or narrower.

[38] Results regarding average treatment effect on the treated were of the same sign and magnitude.

Figure 9: Year-on-year growth in traffic per fixed broadband line, average across countries.



6 Limitations of our analysis

Our analysis suggests that barring exceptional cases, today's broadband speeds do not appear to materially constrain internet usage, as measured by traffic. However, there are several caveats. First, simply because traffic is unchanged does not mean the user's experience is unchanged. For example, lower bandwidth might cause a game download to take much longer. This might frustrate the user, even though the associated traffic was not reduced. Second, our analysis is based on usage today. There might be a policy argument for investing in higher speeds for the future, even if they were not currently required (though we note (i) there is considerable headroom between the speeds already available and those currently taken up by consumers; and (ii) pandemic aside, traffic growth has been on a downward trend since 2015 - see Figure 9). Third, it is at least possible that there is a lag effect when users upgrade to higher speeds. Habits developed under the constraints of lower speeds may take time to change, and thus for a time usage remains lower even on a higher speed connection. Fourth, our analysis does not take account of any constraints due to in-home wifi issues.

7 Conclusion

'Does high bandwidth increase (or enable) traffic', and 'does low bandwidth decrease it' are two independent questions. Our data would seem to suggest the answers are 'no' and 'perhaps a little' respectively.

If higher bandwidth does not (now) result in higher traffic, this has a number of implications. First, this may suggest that traffic growth is likely to slow. If in the past bandwidth was a constraint, then improvements in broadband will have contributed to traffic growth by easing that constraint. However, if it is no longer a constraint, then this contribution to traffic growth will fall away, even if speeds continue to improve. This would be consistent with the growing evidence that (after a pandemic-induced spike) fixed traffic growth is slowing appreciably around the world.[39]

[39] Kenny, R (2022). Patterns of fixed traffic growth, 2022.

Second, we believe these results have implications for the policy debate around traffic charges (since if network upgrades don't result in more traffic, traffic charges won't increase revenue from upgrades) and for subsidies for fiber to the premises and other improvements to broadband (in that usage does not appear to be currently constrained by speeds achievable with fiber to the cabinet).

Third, broadband providers face a marketing challenge. Customers with 30 Mbps+ broadband are likely not facing material 'pain points' at the moment, and the 'hard' benefits of the higher speed lines they are now seeking to sell may be limited. (This is consistent with evidence that take-up rates for higher speeds are modest unless they are priced at levels very close to existing speeds). Looking ahead, this suggests that broadband providers would benefit from any future highly-bandwidth-intense applications. This may have implications for how broadband providers think about their relationships with CAPs. It also has implications for the skills required by FTTP deployers. To date, this has been seen as primarily an engineering challenge, and the executive teams of many FTTP deployers reflect this. However, if the product is not one that will 'sell itself' through technical superiority, then marketing skills will be equally important.

Fourthly and finally, the strong linkage between demographics and traffic suggests that there is enormous latent potential in the infrastructure we already have. Many households (notably older ones) are underutilising their connection. Thus the best route to enhancing the societal and economic value of broadband likely lies in training not trenching - in helping those making only limited use of the internet to get more out of it, rather than in further network deployment.

A Robustness Checks

Table 9: Regressions on Average Data Usage in GB, where Average Download Speed is < 250.

	Average data usage (GB), where Average Download Speed is < 250 Mbps				
	(1)	(2)	(3)	(4)	(5)
% con'ns <30 Mbps	-3.225^{***}				-1.646^{***}
	(0.023)				(0.016)
% con'ns $\geq$ 300 Mbps	0.195^{***}				-2.381^{***}
	(0.045)				(0.030)
Average download speed (Mbps)		0.473^{***}		-0.238^{***}	
		(0.007)		(0.005)	
Age 25-64 %			-3.973^{***}	-3.932^{***}	-3.871^{***}
			(0.064)	(0.063)	(0.061)
Age 65+ %			-7.540^{***}	-7.649^{***}	-7.134^{***}
			(0.097)	(0.096)	(0.093)
Log(Popn density)			16.276^{***}	18.225^{***}	11.308^{***}
			(0.152)	(0.156)	(0.160)
Economically inactive %			1.751^{***}	1.738^{***}	1.915^{***}
			(0.046)	(0.045)	(0.044)
Student %			-3.623^{***}	-3.622^{***}	-3.292^{***}
			(0.050)	(0.049)	(0.048)
Retired %			-0.622^{***}	-0.593^{***}	-0.797^{***}
			(0.092)	(0.092)	(0.089)
Unemployed %			-0.876^{***}	-0.875^{***}	-0.443^{***}
			(0.135)	(0.134)	(0.129)
HH size			99.086^{***}	106.028^{***}	104.663^{***}
			(0.754)	(0.762)	(0.737)
% Managerial			-1.981^{***}	-2.011^{***}	-2.208^{***}
			(0.038)	(0.038)	(0.037)
% Cannot speak English well			-11.424^{***}	-11.789^{***}	-11.729^{***}
			(0.099)	(0.099)	(0.096)
% Works mainly from home			-2.121^{***}	-1.939^{***}	-1.587^{***}
			(0.059)	(0.059)	(0.057)
Constant	537.512^{***}	434.407^{***}	615.497^{***}	607.991^{***}	673.757^{***}
	(0.661)	(0.837)	(5.974)	(5.936)	(5.788)
Observations	182,808	182,808	182,808	182,808	182,808
Adjusted R^2	0.107	0.024	0.622	0.627	0.651
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 10: Regressions on Average Data Usage in GB, where Average Download Speed is < 250.

	Average data usage (GB), where Average Download Speed is < 250 Mbps		
	(1)	(2)	(3)
Age 25-64 %	-7.934^{***}	-8.603^{***}	-3.973^{***}
	(0.042)	(0.056)	(0.064)
Age 65+ %	-10.952^{***}	-13.135^{***}	-7.540^{***}
	(0.029)	(0.097)	(0.097)
Economically inactive %		0.548^{***}	1.751^{***}
		(0.048)	(0.046)
Student %		-5.672^{***}	-3.623^{***}
		(0.050)	(0.050)
Retired %		0.519^{***}	-0.622^{***}
		(0.099)	(0.092)
Unemployed %		-2.329^{***}	-0.876^{***}
		(0.141)	(0.135)
% Managerial		-1.478^{***}	-1.981^{***}
		(0.041)	(0.038)
% Works mainly from home		-2.165^{***}	-2.121^{***}
		(0.063)	(0.059)
Log(Popn density)			16.276^{***}
			(0.152)
HH size			99.086^{***}
			(0.754)
% Cannot speak English well			-11.424^{***}
			(0.099)
Constant	$1,119.037^{***}$	$1,306.101^{***}$	615.497^{***}
	(2.655)	(4.105)	(5.974)
Observations	182,808	182,808	182,808
Adjusted R^2	0.445	0.563	0.622
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 11: Regressions on Median Data Usage in GB.

	Median data usage (GB)				
	(1)	(2)	(3)	(4)	(5)
% con'ns <30 Mbps	−2.940***				−1.321***
	(0.020)				(0.014)
% con'ns ≥ 300 Mbps	1.378***				−0.683***
	(0.033)				(0.022)
Median download speed (Mbps)		0.594***		−0.005	
		(0.008)		(0.005)	
Age 25-64 %			−2.890***	−2.890***	−2.859***
			(0.057)	(0.057)	(0.055)
Age 65+ %			−5.257***	−5.258***	−4.882***
			(0.087)	(0.087)	(0.085)
Log(Popn density)			16.204***	16.223***	11.275***
			(0.136)	(0.137)	(0.145)
Economically inactive %			1.025***	1.024***	1.157***
			(0.041)	(0.041)	(0.040)
Student %			−2.950***	−2.950***	−2.717***
			(0.044)	(0.044)	(0.043)
Retired %			−2.253***	−2.255***	−2.450***
			(0.082)	(0.082)	(0.081)
Unemployed %			−0.886***	−0.888***	−0.589***
			(0.120)	(0.120)	(0.118)
HH size			106.183***	106.206***	104.908***
			(0.673)	(0.674)	(0.664)
% Managerial			−0.948***	−0.947***	−1.085***
			(0.034)	(0.034)	(0.033)
% Cannot speak English well			−7.923***	−7.923***	−7.796***
			(0.089)	(0.089)	(0.087)
% Works mainly from home			−1.380***	−1.379***	−1.085***
			(0.052)	(0.052)	(0.051)
Constant	358.157***	280.107***	322.768***	323.016***	384.008***
	(0.560)	(0.648)	(5.321)	(5.327)	(5.241)
Observations	185,177	185,177	185,177	185,177	185,177
Adjusted R^2	0.126	0.028	0.633	0.633	0.650
Note:					*p<0.1; **p<0.05; ***p<0.01