Economic Geography 2018

An analysis of the determinants of 3G and 4G coverage in the UK

Publication Date: 18 December 2018

About this document

Since 2013, alongside the Connected Nations Report, we have published a research paper entitled "Economic Geography: An analysis of the determinants of 3G and 4G coverage in the UK".[1] This includes an econometric analysis of factors that affect the level of 3G and 4G mobile coverage available to consumers in different areas. The aim of which is to highlight whether the variation in 3G and 4G coverage between different parts of the UK can be explained by differences in the likely demand for, and cost of providing, mobile services.

This report sets out to update that analysis using 3G and 4G coverage data for 2018.

[1] Ofcom (2017) Economic Geography: An analysis of the determinants of 3G and 4G coverage in the UK

Contents

Section

Annex

1. Introduction

What determines mobile coverage in the UK?

A variety of factors influence a firm's decision to offer coverage in a particular area. The decision to offer mobile coverage is essentially a commercial judgement by the mobile network operators (MNOs). MNOs will typically assess the profitability of providing coverage in an area by considering the likely demand for mobile network services as well as the costs of providing those services.

It is to be expected that the main drivers of local availability are the makeup and density of the local population (which will impact on local demand for mobile services) and the topography of the local area (which will cause local variations in the cost of providing mobile services). We see this pattern in the data: areas with good 3G and 4G coverage tend to be in lower lying and more densely populated areas while areas with poor 3G and 4G mobile coverage tend to be in areas that are higher above sea level and have a more sparsely distributed population.

In 2017 we published a research paper in which we used statistical techniques to investigate the factors that affect MNOs local entry decisions.[2] This allowed us to look behind raw averages and gain a better understanding of what drives variations in mobile coverage.

In the past, this has proven to be a useful exercise. In 2013, for instance, we found that factors different to those examined in our analysis appeared to restrict 3G mobile coverage in Northern Ireland. Having identified this, we undertook further research and identified relatively strict planning laws, community opposition and changing network plans by MNOs as the underlying cause of this result.

Given the usefulness of this analysis in uncovering factors that may influence low coverage, in this report we provide an updated analysis using 3G and 4G coverage data for 2018. Additionally, we re-run our analysis with re-stated June 2017 data[3] to compare this year's findings against last year's results.

A summary of our main findings

The main findings of our 2018 analysis are summarised below. In our findings we mainly focus on areas of the UK that are populated and refer to an area as having 'full coverage' if that particular area receives good coverage from all four MNOs (see Section 2 for a full explanation).

Overall coverage levels

- Both 3G and 4G mobile coverage across the UK is generally good with a high probability that there will be good coverage by all four MNOs, especially in urban areas.

[2] For the previous edition of this report, see Ofcom (2017) Economic Geography: An analysis of the determinants of 3G and 4G coverage in the UK

[3] As we detail in the Connected Nations October 2018 Update, Ofcom's drive testing measurements had identified potential errors in coverage data supplied by two MNOs. Ofcom has re-stated historic mobile coverage levels using operators' resubmitted data; we use re-stated 2017 data in the present analysis.

- A comparison of our results from this year and last year suggests that regions are on average slightly more likely to have full 3G coverage in 2018 than 2017. Further, there has been a significant increase in the likelihood of receiving full 4G coverage between 2017 and 2018.

The effect of cost and demand factors on coverage

- A more densely distributed local population increases the probability of having full 3G and 4G coverage. In 2018, a coverage area at the 75th percentile for population density (which corresponds to 710 people per km^2) has a 7.4 and 7.8 percentage points higher probability of receiving full 3G and 4G coverage respectively than a coverage area at the 25th percentile for population density (which corresponds to 54 people per km^2).
- The composition of the local population affects mobile coverage: a larger working age population and an affluent population increase the probability of good coverage. In 2018, a coverage area at the 75th percentile for affluence (where 66% of the local population are classified as affluent) has a 1.7 and 1.6 percentage points higher probability of receiving full 3G and 4G coverage respectively than a coverage area at the 25th percentile for affluence (where 45% of the local population are classified as affluent).
- Local topography can also affect the probability of good coverage. All else being equal, the probability of receiving full 3G and 4G coverage between a low-lying area and an area that is high above sea level (as defined by the inter-quartile range between 35 and 119 metres above sea level) is around 1.2 percentage points higher for 3G and 1.6 percentage points higher for 4G in 2018.
- Additionally, we find that the probability of receiving full coverage between an area with flat terrain and more uneven terrain (as defined by the inter-quartile range in our dataset) is around 2.0 and 2.5 percentage points higher for 3G and 4G respectively.
- Distance to the backhaul network can further influence the probability of having full mobile coverage. The probability of full coverage between an area that is closer to its backhaul network and an area that is further away (as defined by the interquartile range between 1300m and 4000m) is approximately 4.7 percentage points higher for 3G and 4.3 percentage points higher for 4G.

Variation in regional coverage

- Once the local demand and cost factors are taken into account, the probability of good 3G and 4G coverage is relatively similar between different regions of the UK.
- Compared to 2017 the amount of unexplained regional variation in mobile coverage has decreased significantly, especially for 4G. This indicates that much of the variation in coverage between regions can be explained by our specified demand and cost factors. This also means that factors outside of our model have become less relevant over time.
- However, there still appear to be certain regions where coverage is lower than we would expect given the underlying level of demand and cost factors. Specifically, there appear to be unexplained factors that are negatively affecting 3G coverage in the South East and East of England. Whereas 4G coverage appears to be lower than expected based on local cost and demand factors in several regions, especially the South West.

The structure of this report

The remainder of this paper is structured as follows:

- Section 2 introduces the data and determinants of mobile coverage that we use in our analysis. We provide a brief discussion of the relationship between these determinants and the number of 3G and 4G operators with good coverage.
- Section 3 sets out the approach that we use to assess whether and to what extent the determinants specified in the previous section drive variation in mobile coverage.
- Finally, in Section 4 we present the results of our analysis.

2. Data

In this section, we introduce the data and variables that we use in this study. We start by setting out what we mean by mobile coverage and what constitutes 'good coverage', and then discuss the coverage data we use in this report.

We outline the set of demand and cost factors that we use in this analysis. These variables should capture the considerations an MNO makes when considering whether to provide mobile coverage in an area. Having summarised the demand and cost factors relevant to our study, we discuss the relationship between these factors and the number of 3G and 4G operators with good coverage.

The measure of coverage

To measure coverage, we divided the entire UK land mass into 200m by 200m areas. We refer to each of these areas as a 'coverage pixel'. For each coverage pixel, we collected detailed data from MNOs on predicted outdoor 3G and 4G coverage for May 2018. Using these data, we constructed a dataset of mobile phone signal strength across the UK.

We assess the level of coverage in each pixel by counting the number of operators with good coverage. An operator has good coverage if its reported mobile phone signal strength is equal to or exceeds a defined signal strength threshold (set at -100dBm for 3G and -105dBm for 4G).[4]

As we detail in the Connected Nations October 2018 Update[5], we noted Ofcom's drive testing measurements had identified potential errors in coverage data supplied by two MNOs. Ofcom has re-stated historic mobile coverage levels using operators' resubmitted data. We re-ran the analysis for June 2017 in order to ensure comparability between this year's results and last year's results.

The determinants of coverage in our analysis

As mentioned above, MNOs will typically deploy cell sites based on an assessment of local cost and demand factors in an area. This area is typically larger than the coverage pixel (henceforth referred to as the 'coverage area').

We define this coverage area as an area that is covered by a circle of radius 1km around the coverage pixel. We use a 1km radius because this interval roughly corresponds to the range of a typical cell site across a variety of frequencies and environments.[6,7]

For each coverage area, we obtained data on a variety of variables that influence mobile coverage. These can be grouped into the following broad categories:

---

[4] These signal thresholds are consistent with those used in the Connected Nations Report, see Ofcom (2018) Connected Nations Methodology
[5] Ofcom (2018) Connected Nations Update October 2018
[6] See for instance Table 8 in Recommendation ITU-R M.1768-1 (04/2013)
[7] Our results do not materially change when using different radii ranging from 0.5km to 2.5km.

a) Demand factors. Variables that capture the size of the local population, as well as its composition in terms of age and affluence;[8]

b) Cost factors. Variables such as the density of the local population, the characteristics of the local terrain (height above sea level and variability of height) and the proximity of cell sites to mobile infrastructure are likely to influence the costs of providing 3G and 4G mobile services;[9] and

c) Location-specific indicators. These variables are used to capture unobserved variations in demand and costs specific to a particular region or urban location.[10] For example, the indicator for Scotland will allow us to capture effects on coverage not picked up by our other variables, such as population or topography, which are specific to Scotland.

Our analysis focuses on populated areas since the demand factors we use place a strong emphasis on demographic variables.[11] To illustrate our analysis, consider Figure 1 below which shows a coverage pixel (a 200m by 200m area) and the associated coverage area (the full shaded area) in a medium sized town:

- Our dependent variable is the number of operators with good 3G/4G coverage in the 200m by 200m pixel at the centre of the image;
- The local population is obtained by summing the population numbers in all of the postcodes located within the coverage area;
- Information on age and affluence of the local population is obtained by summing over the number of people belonging to each age and affluence group in all of the output areas within the coverage area;
- Summary statistics for topography of the coverage area (median height and standard deviation of height) are obtained from the elevation database; and
- The distance metric is derived by calculating the distance between the location of the pixel and its closest mobile backhaul location.

[8] To obtain data of this nature, we used UK census data, mapping the data to coverage areas. UK census data is available at what is termed an output area centroid. The "output area" is the most detailed level at which (bulk) census data are released to preserve anonymity.
[9] Population density information was obtained from UK census data at a postcode level. Topographic information was obtained from the Consortium for Spatial Information's (CGIAR-CSI) SRTM 90m Digital Elevation Database v4.1.
[10] We obtained data on the urbanicity of the location using the urban locale classification from Bluewave Geographics. Bluewave Geographics is a provider of digital mapping, geographic analysis and sampling services to market research and fieldwork sectors.
[11] More specifically, our analysis relies on demographic variables (such as age, affluence etc), so we would not be able to analyse unpopulated areas with the current model. We have considered how we could extend the analysis, but this would require new variables (e.g. distance to major road, level of car and pedestrian traffic etc.). Currently, adequate data relating to these variables are not readily available.

Figure 1: Coverage area with an illustrative example of a backhaul location

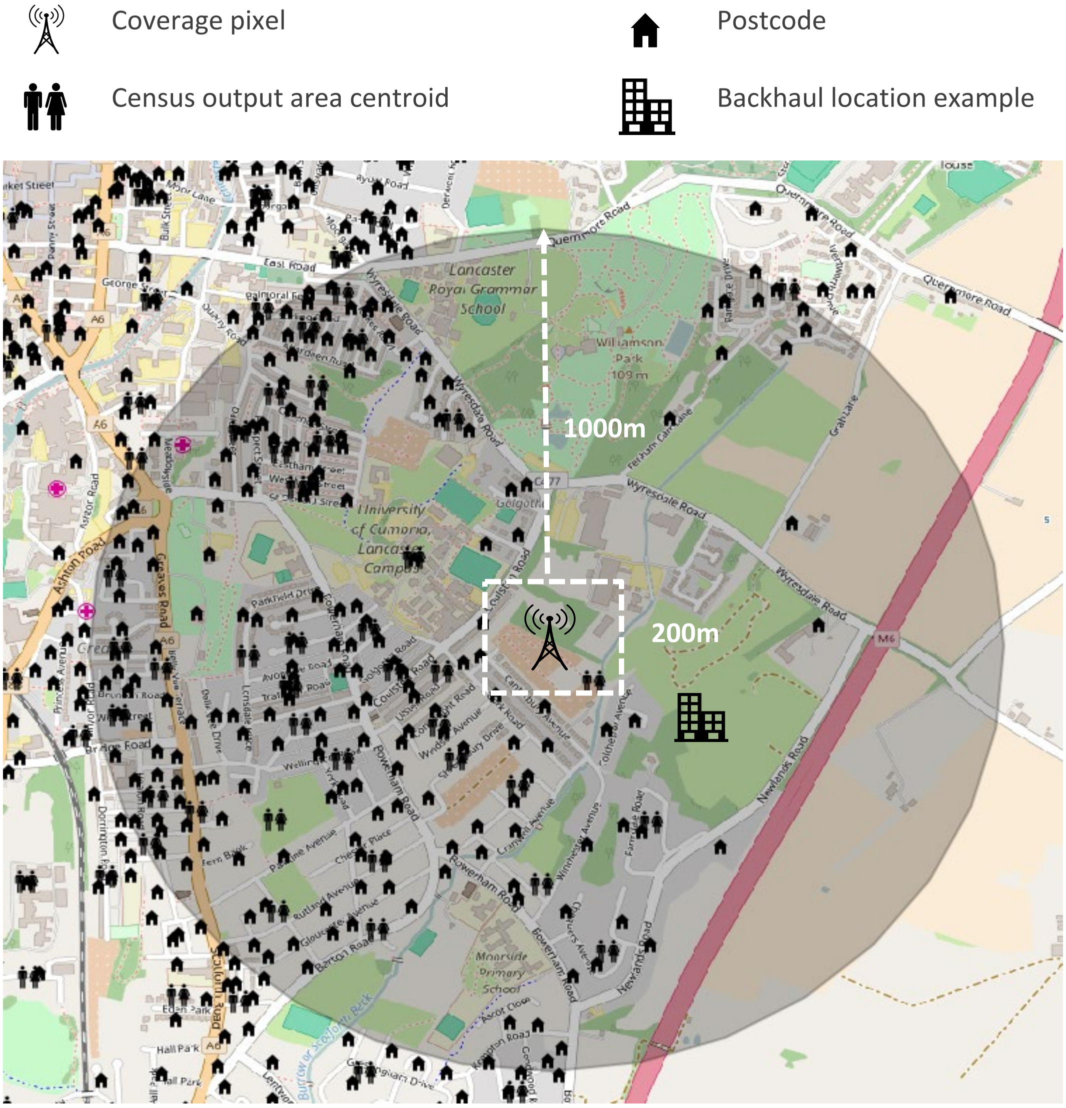


Table 1 below provides a list and description of the variables for which we have collected data and that we use in our analysis.

Table 1: Descriptions of key variables

Variable name	Remarks
num_3G_ops/num_4G_ops	The number of 3G/4G operators with good coverage in a 200m by 200m pixel (operators with signal strength equal to or above the threshold of -100 dBm for 3G and -105 dBm for 4G). This variable can take the value of 0, 1, 2, 3 or 4.
ln_pop	The population within the postcode sector. We measure population in natural logarithms.
pct_abc1	The approximate percentage of the population within the coverage area that reside within a household classified within socioeconomic groups A, B or C1.[12,13]
pct_under25	The approximate percentage of the population within the coverage area that are aged under 25.
pct_over60	The approximate percentage of the population within the coverage area that are aged 60 or over.
height_median	The median height above sea level, in metres, of the coverage area.
height_stdev	The standard deviation of the heights of the coverage area.
urban_code	This is a categorical (dummy) variable that identifies the coverage area as either urban or rural.
dist_backhaul	The distance from the pixel to its nearest backhaul location (see Annex 1 for more information).
region	This is a categorical (dummy) variable that identifies the part of the UK that the coverage area is situated within. This is either the nation in the case of Scotland, Wales and Northern Ireland or the region for England.

[12] These socioeconomic ratings are obtained from the National Readership Survey's (NRS) social classification system. A, B and C1 are typically allocated to occupations that can be described as more affluent. For example, C1 are occupations that can be described as "Supervisory, clerical and junior managerial, administrative and professional".

[13] We appreciate that these ratings are perhaps not the most relevant measures of affluence, so we have explored other measures such as using local unemployment rates, house prices and Index of Multiple Deprivation (IMD). We, unfortunately, found that these new measures were unsuitable. For unemployment rates, we have done some trial testing and we found that it made the overall results look unintuitive (it is likely that this is caused by the variable being highly collinear with other variables and we would need to do more work in the future to use this measure appropriately). For house prices, we found that relevant data (location, price etc.) were only easily available for England and Wales but these were either not collected or were not easy to obtain for the other nations. For IMD, we found that making comparisons using IMD is difficult since each nation uses its own criterion to calculate IMD (for example England places less weight on Income Deprivation than Scotland); IMD is also a relative deprivation measure within each nation and, therefore, comparisons of IMD score of areas between different nations would be erroneous.

Summary statistics

We present summary statistics below for the dependent variables (the number of 3G operators and the number of 4G operators) as well as the continuous explanatory variables in our analysis. Table 2 provides summary statistics for 2018, and Table 3 provides the same statistics for 2017 data.

The tables show that the number of good 3G and 4G operators has been rising over time. This upwards trend has been especially strong for 4G with the average number of operators with good coverage rising from roughly 3.5 to 3.8.

Table 2: Summary statistics, 2018

Variable name	Minimum	25th percentile	Median	Mean	75th percentile	Maximum
num_3g_ops	0	4	4	3.797	4	4
num_4g_ops	0	4	4	3.780	4	4
population	1	170	498	2,270	2,224	63,382
pct_under25	0	24%	27%	27%	30%	95%
pct_over60	0	22%	27%	27%	32%	98%
pct_abc1	0	45%	56%	55%	66%	100%
heightmedian	-3	35	72	85	119	614
heightstdev	0	7	12	16	20	261
dist_backhaul	0	1300	2400	3100	4000	67600

Table 3: Summary statistics, 2017

Variable name	Minimum	25th percentile	Median	Mean	75th percentile	Maximum
num_3g_ops	0	4	4	3.694	4	4
num_4g_ops	0	3	4	3.482	4	4
population	1	170	498	2,270	2,224	63,382
pct_under25	0	24%	27%	27%	30%	95%
pct_over60	0	22%	27%	27%	32%	98%
pct_abc1	0	45%	56%	55%	66%	100%
heightmedian	-3	35	72	85	119	614
heightstdev	0	7	12	16	20	261
dist_backhaul	0	1300	2400	3100	4000	67600

Simple graphical analysis between our cost and demand data and the number of 3G and 4G operators with good coverage

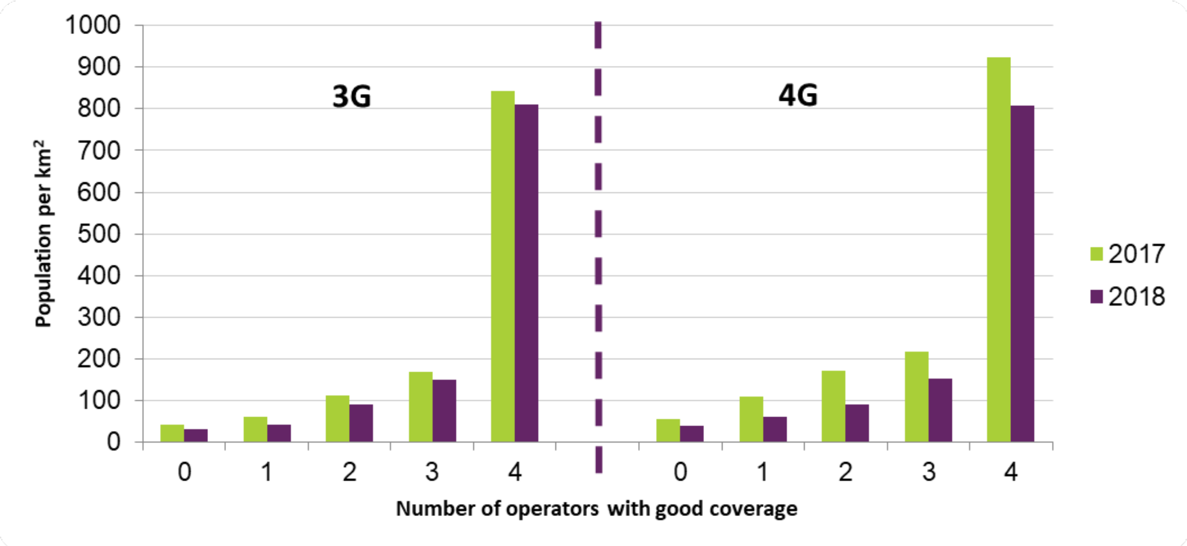
In this section using simple graphical analysis, we illustrate the relationship between the number of 3G and 4G operators with good coverage and various cost and demand factors. Then we consider how good coverage varies with rurality and region.

Population density

Figure 2 shows the relationship between average population density and the number of 3G and 4G operators with good coverage. We can see that areas with good coverage are typically located within densely populated areas in both 2017 and 2018.

We observe that the relationship between population density and the average number of 3G and 4G operators has been falling over time. We would largely expect to see this trend. This is because operators will initially roll out mobile coverage in areas with high demand (areas with higher population density) before rolling out to locations with lower expected demand (more sparsely populated areas).

Figure 2: Comparison of average population density by number of 3G/4G operators, 2017 and 2018



The trend of falling average population over time is stronger for 4G than 3G. This is consistent with the findings in Tables 2 and 3 which show that there has been a considerable increase in 4G coverage between 2017 and 2018. This is further supported by findings from the Connected Nations Report[14], which show that outdoor geographic 4G coverage across the UK has increased significantly between 2017 and 2018.

Height of coverage area

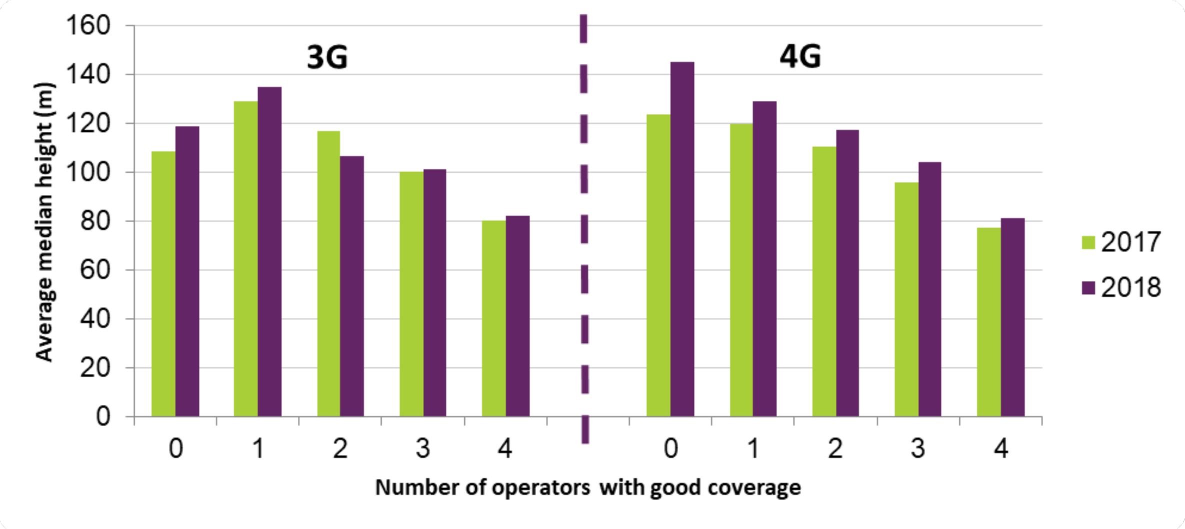
Figure 3 plots the number of operators with good coverage against the average height above sea level in 2017 and 2018. For topography, we would anticipate two patterns:

---

[14] Ofcom (2018) Connected Nations

- We would expect good coverage to be less likely in more elevated areas of the UK. This means that we should observe that a lower average median height is associated with a higher number of operators with good coverage in Figure 3; and
- We would expect MNOs to roll out to harder to reach and costlier places over time. This means that we should observe that average median height would increase over time for a given number of operators with good coverage in Figure 3.

Figure 3: Average median height of coverage areas by the number of 3G/4G operators, 2017 and 2018



From Figure 3, we can see that both patterns hold strongly for 4G. However, we do not observe these trends for 3G. Firstly, on average, areas with zero MNOs offering good 3G coverage are lower than areas with one operator offering good 3G coverage in both years. Secondly, the average median height for two operators offering good 3G coverage decreases between 2017 and 2018.

In relation to the first point, we consider that our median height variable will be an appropriate proxy for local cost conditions in most cases. However, median height will become a poor proxy in locations with extreme terrain characteristics. We found that a high proportion (around 80% in 2018) of the areas with zero 3G MNOs offering good coverage are located in Scotland and Wales. This may imply that our median height variable will be a poor cost indicator in this unique case.

In terms of the second point, we note from above that there are likely to be factors other than the height of the coverage area that will determine the number of MNOs providing good coverage between 2017 and 2018. For example, median height may negatively affect an MNO's entry decisions, but they may still enter if other demand conditions are sufficiently favourable.

Distance to backhaul

Figure 4 plots the relationship between distance of the coverage pixel to its backhaul network and the number of good 3G and 4G operators. We can see that areas with a higher number of operators with good coverage tend to be closely located to backhaul networks.[15]

Figure 4: Average distance of pixels to its backhaul network by the number of 3G/4G operators, 2017 and 2018

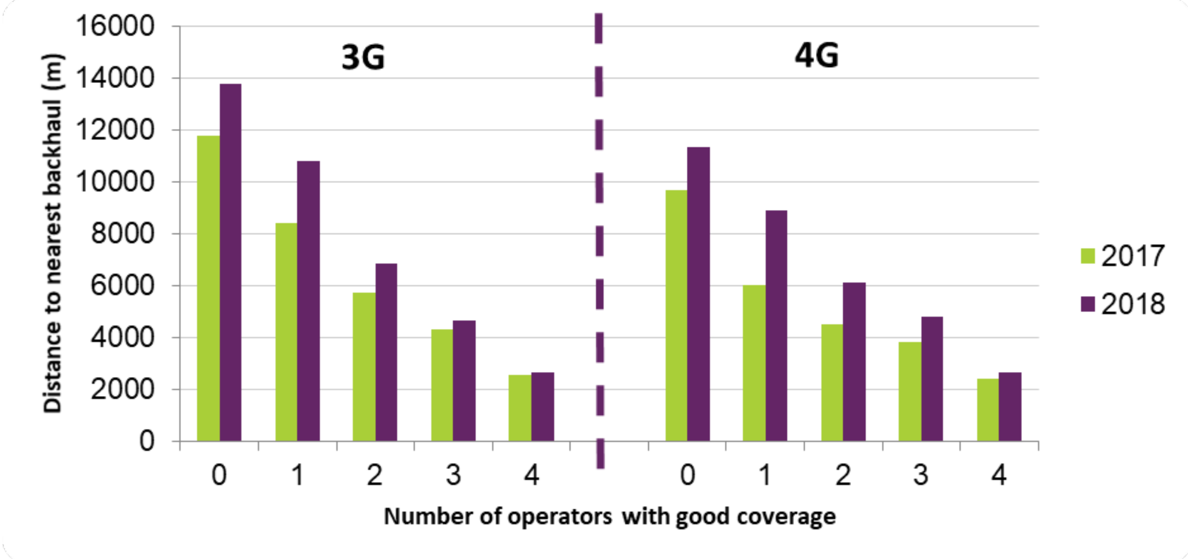


Figure 4 shows that average distance is increasing over time for both 3G and 4G. This can be explained by MNOs choosing to first develop their mobile infrastructure in areas closer to their core network, which are cheaper to deploy in, and then deploying infrastructure in more expensive and distant places.

Rurality

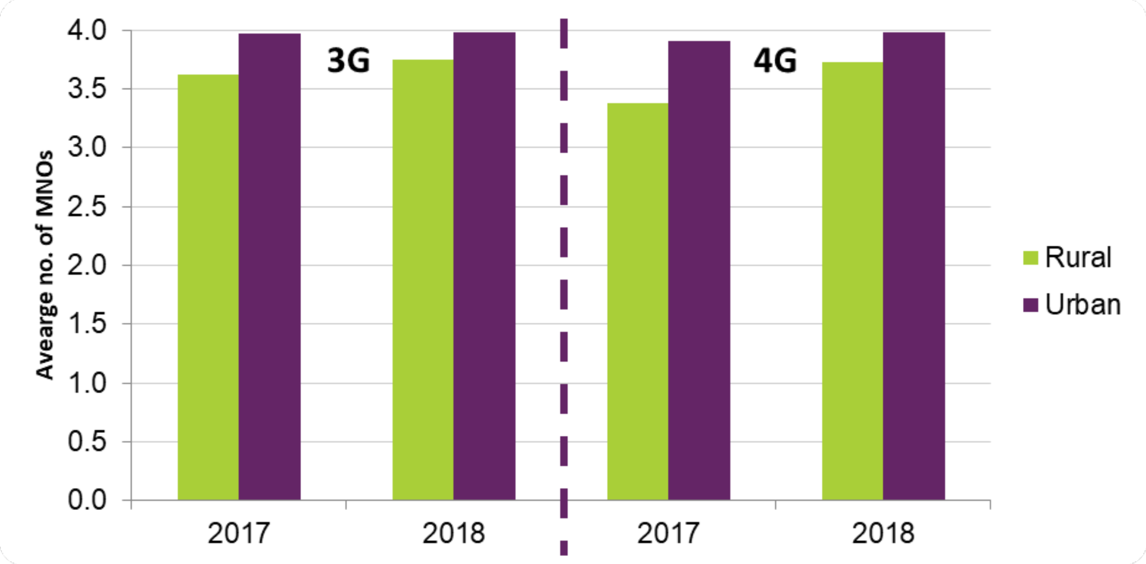
From Figure 5, we can observe that the average number of MNOs with good coverage has been increasing over time in both rural and urban areas for both 3G and 4G. The increase in 4G coverage between 2017 and 2018 is particularly large in rural areas. While there has been significant improvement in rural coverage, we do note that our analysis focuses on populated areas and not relatively unpopulated areas. Moreover, as we document in the Connected Nations Report[16], there remains the need for further improvement in rural coverage.

---

[15] As discussed in Annex 1, there could be a confounding effect between the cost factor of being more expensive for MNOs to deploy mobile cell sites to areas further away to their core network, and a demand factor related to it being less profitable to deploy cell sites to areas further away from major population centres.

[16] Ofcom (2018) Connected Nations

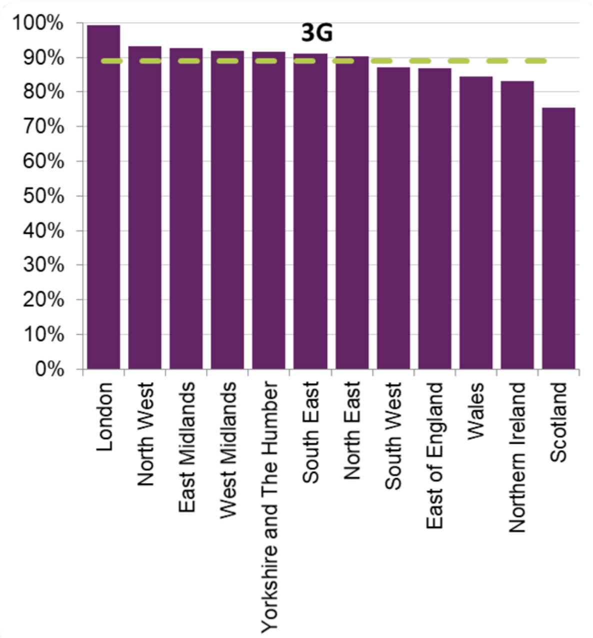
Figure 5: Average number of MNOs with good coverage in urban and rural areas, 2017 and 2018

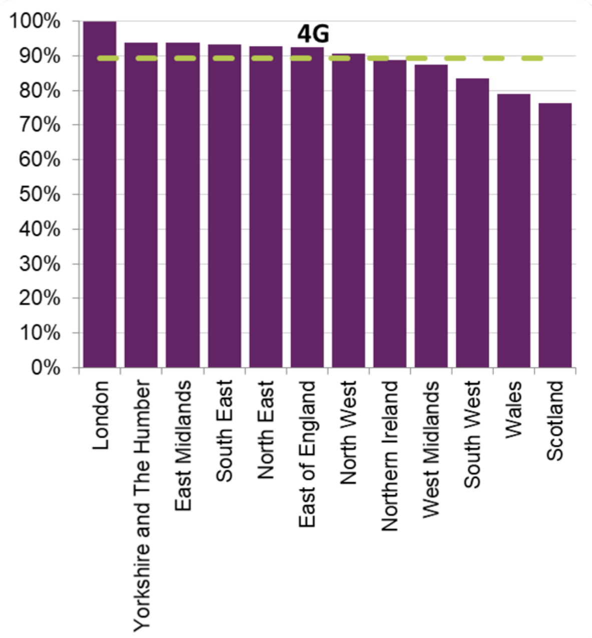


Nations and regions

Finally, in Figure 6 we show the actual level of full 3G and 4G coverage in 2018 for the different nations and regions of the UK (the dashed lines in each graph represent the UK average). We can observe that there are generally more areas in London that have full 3G than the UK average while the opposite is true for areas within Scotland, Northern Ireland and Wales. A similar pattern arises for 4G where more areas in London have full 4G than the UK average while the opposite is true for areas within Scotland, Wales and the South West.

Figure 6: Actual coverage by all four MNOs, 2018





Based on this simple graphical analysis, it may appear that certain areas of the UK are ‘under-served’ in terms of mobile coverage. However, a region may have below average coverage in part because it is less densely populated or has more challenging terrain than other regions. Regression analysis can allow us to examine how much of the regional variation in 3G and 4G coverage can be explained by regional differences in the demand and cost factors.

In the next section, we set out briefly the regression methodology that we will use to consider how a given demand or cost factor affects the likelihood of 3G or 4G coverage in an area whilst holding all other factors constant.

3. Methodology

In Section 2 it was noted that mobile coverage varies across the UK. Certain nations and regions have levels of 3G and 4G coverage below the UK average, and in this sense, it might appear that the data indicate these parts of the UK are 'under-served'.

While the average coverage levels presented in Figure 6 are informative, they do not tell us what factors are causing these differences in coverage between regions. For instance, a certain region may have below average coverage because it is less densely populated than other regions and thus less profitable for MNOs. Alternatively, the true driver of below average coverage for a region may be more challenging terrain, which makes it costlier for MNOs to deploy mobile infrastructure. This lack of clarity about the underlying drivers of coverage levels makes it difficult to draw firm conclusions.

Regression analysis

To advance our understanding of the factors driving the level of mobile coverage across the UK, we undertook multiple regression analysis. Regression analysis has the advantage of allowing us to analyse the impact of each of our local cost and demand factors on coverage while holding all other factors constant.

For our analysis we adopt a relatively simple entry model where we treat each coverage pixel as an area that a potential MNO would be interested to enter and invest in. We use a regression technique[17] that links the number of operators with good coverage in a pixel to a range of factors that determine whether an operator would be interested to enter the area.

The factors that we consider relevant to determine entry are likely to be the demand, cost and location specific factors (population, affluence, age, topography, urban location, distance to core network and region) that we discussed in Section 2.

We conduct our regression analysis with two central aims:

- To assess whether and to what extent our potential explanatory factors influence the likelihood of 3G and 4G mobile coverage; and
- To assess the extent to which regional differences in 3G and 4G mobile coverage can be explained by these factors.

In the next section we present and discuss results from this regression analysis.

[17] Specifically, an ordered probit model. This technique is designed to be used when the dependent variable takes only a limited number of values (such as 0, 1, 2, 3 or 4) and the order of these numbers matters (a larger number of operators implies better coverage).

4. Results

This section presents the key findings from our regression analysis. First, we outline the effect of the local cost and demand factors on the likelihood of 3G and 4G mobile coverage. Second, we summarise the extent to which regional variations in coverage can be explained by these factors.

A brief discussion of the econometric results

Applying the regression techniques outlined in Section 3, we obtain a set of regression results – a summary of the results can be seen in Annex 2. The results of models of this type require careful interpretation.[18] The coefficients returned by our regression technique show only the direction of the effect of each explanatory variable on the likelihood of good 3G and 4G coverage.

Ultimately, we are interested in the magnitude of the impact that each variable has on coverage. Therefore, we have calculated the marginal effect[19] of each explanatory variable. The marginal effect gives us the magnitude of the impact a variable has on the predicted probability that a coverage area will receive good 3G or 4G coverage from all four operators (full coverage).

For the continuous variables, we calculate marginal effects by evaluating the average probability that a coverage area will have full 3G and 4G coverage when that variable is at either the 'low' or 'high' quantile.[20] The difference between the probabilities associated with these 'low' and 'high' quantiles for each explanatory variable allows us to compare their relative impacts.

From this analysis we find the following:

- Coverage across the UK is generally good with a high probability that there will be good coverage by all four MNOs, especially in urban areas.
- A more densely distributed local population increases the probability of having full 3G and 4G coverage. In 2018, a coverage area at the 75th percentile for population density (which corresponds to 710 people per km^2) has a 7.4 and 7.8 percentage points higher probability of receiving full 3G and 4G coverage respectively than a coverage area at the 25th percentile for population density (which corresponds to 54 people per km^2).
- The composition of the local population affects mobile coverage: a larger working age population and an affluent population increase the probability of good coverage. In 2018, a coverage area at the 75th percentile for affluence (where 66% of the local population are classified as affluent) has a 1.7 and 1.6 percentage points higher probability of receiving full 3G and 4G coverage respectively than a coverage area at the 25th percentile for affluence (where 45% of the local population are classified as affluent).

---

[18] Our regression technique uses the data to fit a model that predicts a 'latent variable'. This is a variable that can be thought of, broadly, as a score giving the favourableness of the coverage area depending on its characteristics. The higher this value, the more likely it is that the coverage area has a higher number of operators with good coverage. The model then fits a series of 'cut-offs' that determine how this score translates to the probability of each category (0, 1, 2, 3 or 4 operators with good 3G/4G coverage).

[19] Specifically, we have used the average marginal effects for each variable. The average marginal effect is calculated by estimating the predicted probability of full coverage for each pixel and then taking the average across all pixels.

[20] Corresponding to the 25th and 75th percentiles of the relevant variable respectively.

- Local topography can also affect the probability of good coverage. All else being equal, the probability of receiving full 3G and 4G coverage between a low-lying area and an area that is high above sea level (as defined by the inter-quartile range between 35 and 119 metres above sea level) is around 1.2 percentage points higher for 3G and 1.6 percentage points higher for 4G in 2018.
- Additionally, we find that the probability of receiving full coverage between an area with flat terrain and more uneven terrain (as defined by the inter-quartile range in our dataset) is around 2.0 and 2.5 percentage points higher for 3G and 4G respectively.
- Distance to the backhaul network can further influence the probability of having full mobile coverage. The probability of full coverage between an area that is closer to its backhaul network and an area that is further away (as defined by the interquartile range between 1300m and 4000m) is approximately 4.7 percentage points higher for 3G and 4.3 percentage points higher for 4G.[21]

The results tables, accompanied by a more detailed discussion of these marginal effects, can be found in Annex 3 for 3G and Annex 4 for 4G.

In addition, we have performed various checks on the model to test its robustness.[22] We note that there are limitations to our model. As mentioned previously, the variables used are not perfect proxies for the underlying demand or cost factor. For example, our topographic variables (median and standard deviation of height within the coverage area) may not be granular enough to adequately capture costs in areas that have more extreme terrain characteristics such as valleys.[23]

Implications of our econometric results for regional coverage differences

From the data presented in Section 2, there may be a perception that certain parts of the UK are ‘under-served’ in terms of their level of 3G or 4G coverage. However, as we set out above, a region may have below-average coverage for many different underlying reasons. Regression analysis allows us to say how much of the regional variation in 3G and 4G coverage can be explained by differences in the demand and cost factors used in our model.

3G coverage by all four MNOs in 2018

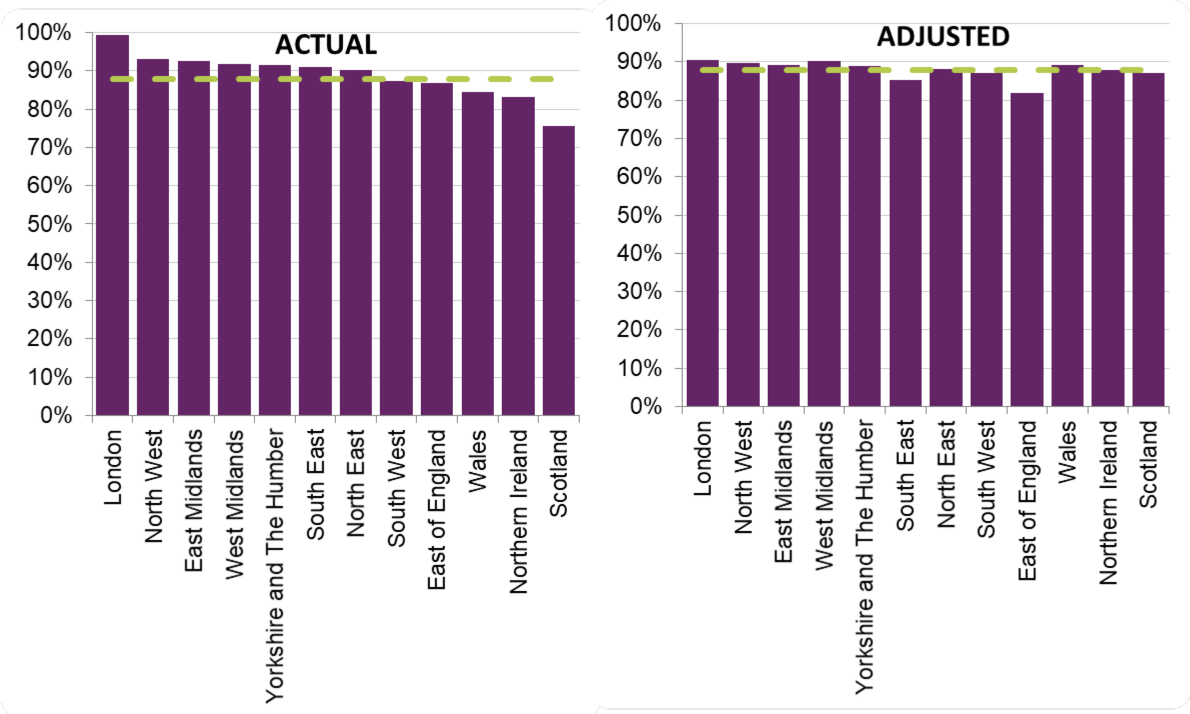
[21] We would expect the difference in the probability of full coverage between the 25th and 75th percentile to be larger in magnitude for 4G than for 3G for each factor. This is because 3G is a more mature technology and, therefore, more likely to serve the 25th and 75th percentiles evenly. However, we do not observe this outcome for affluence or distance to backhaul in 2018. A possible reason is that MNOs may be choosing to upgrade masts to 4G in areas less well served by 3G to improve general access to mobile data. MNOs may also be rolling out additional 4G masts to new locations which may mean that 4G coverage is better at serving the 25th and 75th percentiles than 3G. We will explore the reasons for this effect in the future.

[22] For instance, some care must be taken when assessing the statistical significance of our results as errors are likely to be correlated within clusters of coverage pixels that are in close proximity to each other. As such, we have used clustered standard errors where clusters are defined by coverage pixels that are within the same 10kmx10km grid square. We have conducted robustness checks with respect to different cluster sizes as well as different coverage area radius parameters.

[23] One solution would be to use a more flexible specification with respect to these topographic variables (this could also apply to the population variable as well). For example, it could be argued that it would be more appropriate to use higher order terms (e.g. squared terms) and interaction variables for topography and population based on the diagrams in Section 2. However, we have done some trial testing using higher order terms and we have found the new specification either makes minimal difference or makes the overall results look unintuitive. We will look at this in more detail in the future.

In Figure 7 below, we show both 'actual' and 'adjusted' geographic 3G coverage for the different nations and regions of the UK in 2018 (the dashed lines in each graph represent the UK average). The 'actual' graph is based on the data in Section 2. In contrast, the 'adjusted' coverage graph illustrates a prediction of the level of 3G/4G coverage by all four MNOs if a region were just as densely populated, affluent, urban and hilly etc. as the UK as a whole.

Figure 7: Actual vs adjusted geographic 3G coverage by all four MNOs, 2018



As can be observed from Figure 7, after adjusting for these factors, the apparent differences in 3G coverage have reduced between regions. This is reflected by the smaller variation about the UK average for adjusted coverage relative to actual coverage.

For instance, the level of actual 3G coverage by all four MNOs in Scotland appears to be below the UK average. However, after taking into account local cost and demand conditions, our model predicts that the level of 3G coverage available to consumers in Scotland would not differ too greatly from the UK average. That is, most of the deviation of actual 3G coverage in Scotland from the UK average appears to be explained by the demand and cost factors that we have considered.

In principle, if it were possible to perfectly capture and model all factors that affected the MNOs' coverage decisions, our estimates of adjusted coverage would be exactly the same across all regions, i.e. there would be no deviation about the UK average as all variation would be explained.

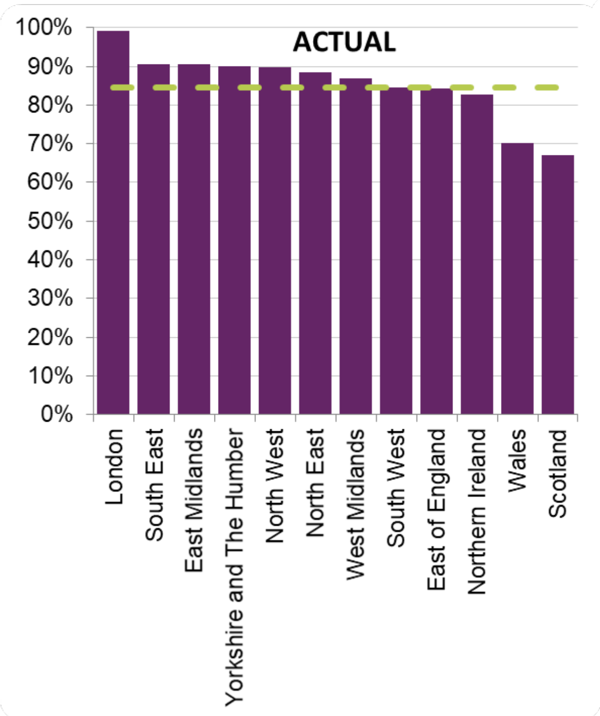
In practice, MNOs' coverage decisions are influenced by factors that we do not control for or only measure imperfectly, which is reflected by certain regions having an adjusted coverage level that differs from the UK average. Figure 7 shows that there are unobserved demand, cost or other factors that are negatively affecting 3G coverage in the South East and East of England. This suggests that there may be other factors in these two regions that influence MNOs' coverage decisions. For instance, there may be regional policies or MNO coverage plans for 3G specific to the South East and

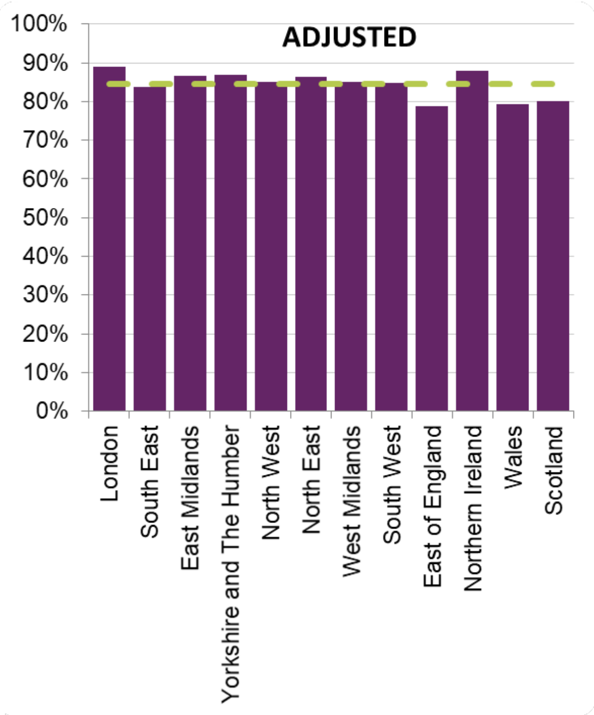
East of England which mean that these regions receive coverage below the average adjusted UK level. Further work will be needed to understand this.

3G coverage by all four MNOs in 2017

Figure 8 shows the same actual and adjusted geographic 3G coverage within nations and regions of the UK in 2017. A slightly weaker pattern than 2018 can be seen, nevertheless, the adjusted coverage figures are able to reconcile differences between regions and nations.

Figure 8: Actual vs adjusted geographic 3G coverage by all four MNOs, 2017



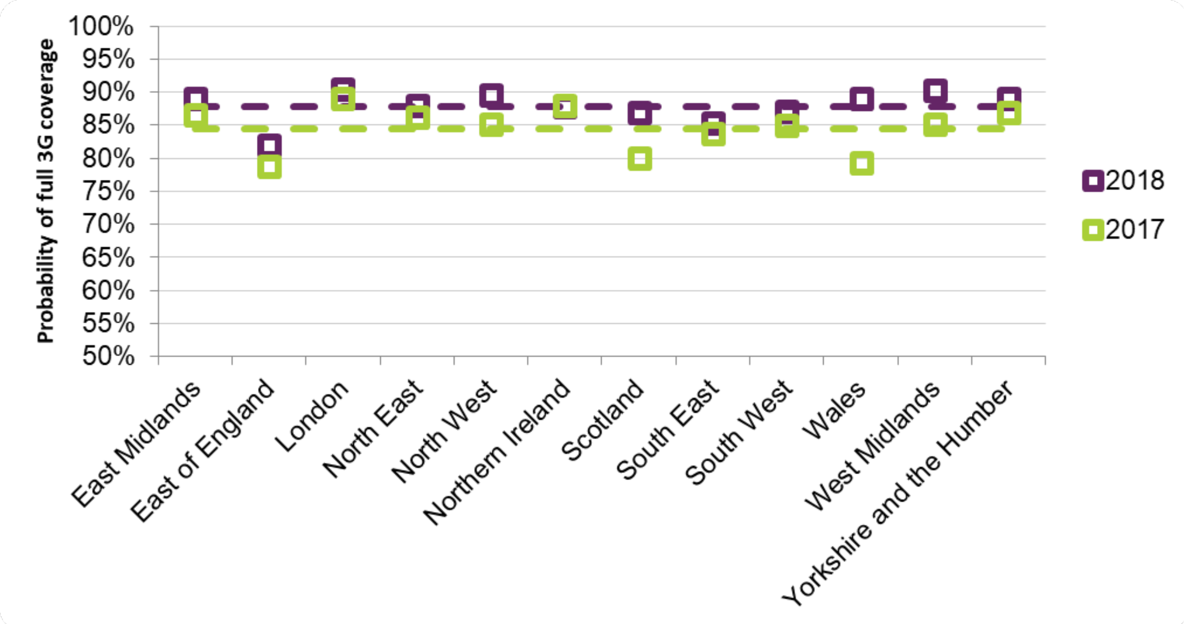


3G coverage results across time

Figure 9 plots the adjusted 3G coverage by region and the UK average for each year (represented by the dashed line). As can be observed, regions are on average more likely to have full 3G coverage in 2018 than in 2017, which suggests that operators are deploying or improving their 3G mobile infrastructure over time.

We note that both Scotland and Wales had adjusted coverage levels below the UK average in 2017, but as Figure 9 shows in both nations the deviation between adjusted coverage levels and the UK average has narrowed in 2018. This means that in both nations, the level of 3G mobile coverage is now closer to what we would expect given underlying cost and demand factors, which implies that factors external to those we model are becoming less relevant over time.

Figure 9: Adjusted geographic 3G coverage by all four MNOs, 2017 and 2018



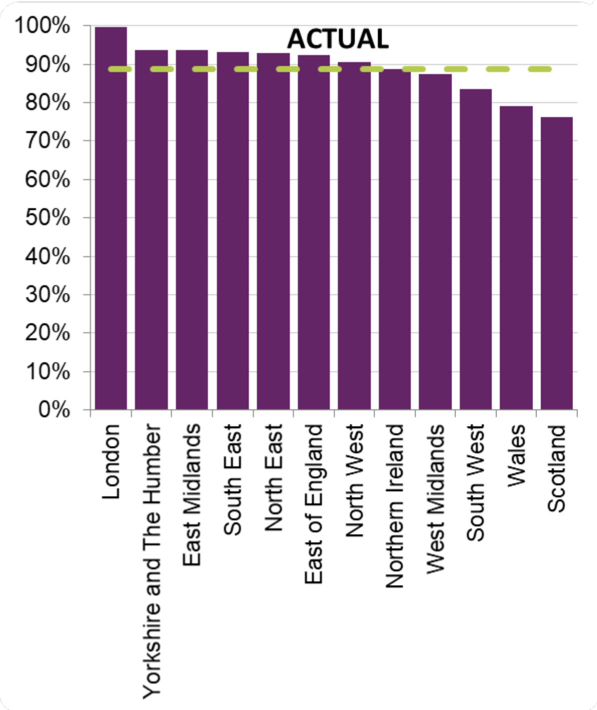
The results above also show that adjusted 3G coverage for the East of England is below average across both years. As set out previously, this difference could potentially be due to our chosen cost and demand factors being insufficient to explain regional differences. Again, this may also be due to regional policies or MNO deployment plans specific to the East of England which lead to below average coverage. Further work is required for us to understand the exact reasons for this result.

4G coverage by all four MNOs in 2018 and 2017

Figure 10 and Figure 11 display the actual and adjusted 4G coverage by regions in 2018 and 2017, respectively. As for 3G, controlling for differences in demand and cost factors appears to explain the regional disparities in coverage, especially in 2018.

For instance, London currently has the highest level of full 4G coverage (99%) in 2018. Yet the difference between the actual and adjusted probabilities (roughly 6 percentage points in 2018) suggests that a great proportion of the disparity in 4G coverage between London and other parts of the UK can be explained by its relatively favourable demand and cost conditions, i.e. it is very densely populated and relatively affluent.

Figure 10: Actual vs adjusted geographic 4G coverage by all four MNOs, 2018



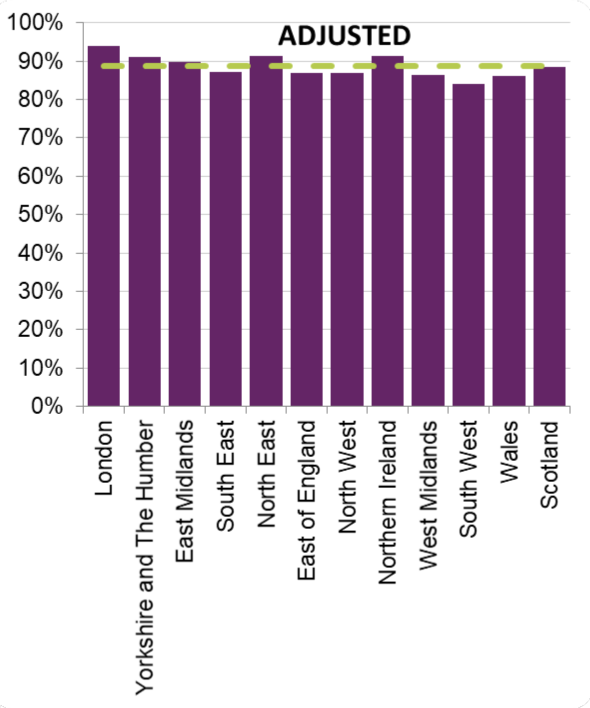
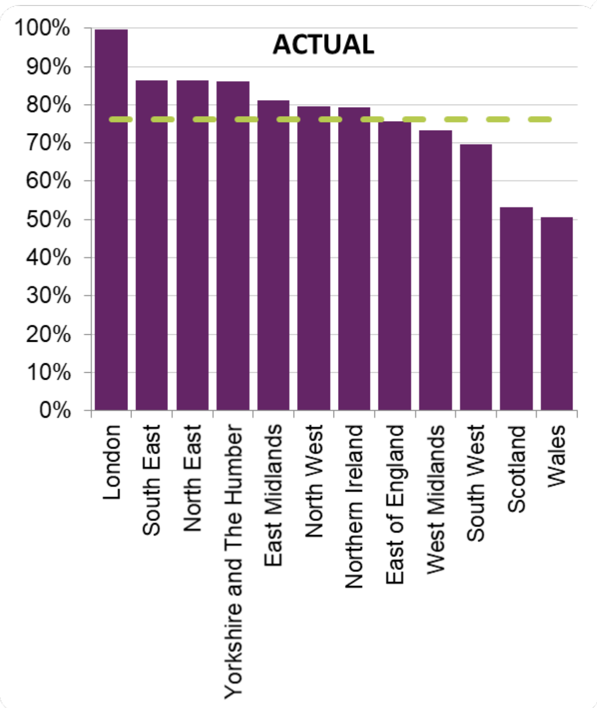
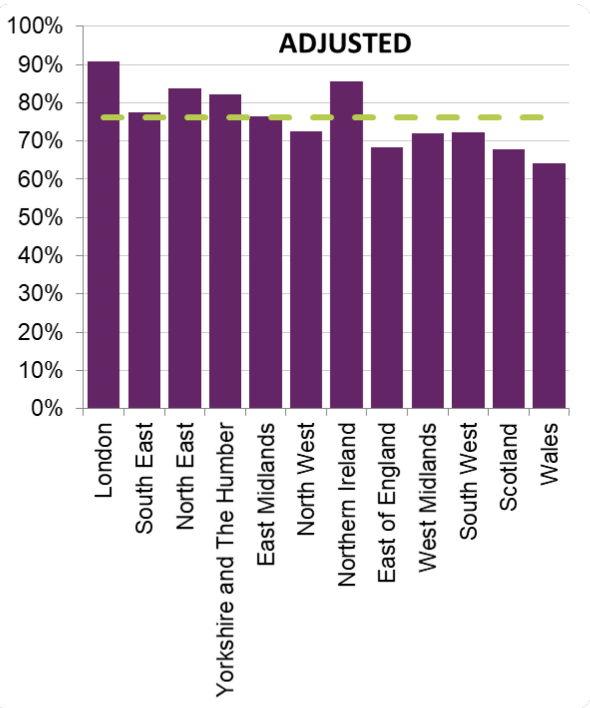


Figure 11: Actual vs adjusted geographic 4G coverage by all four MNOs, 2017

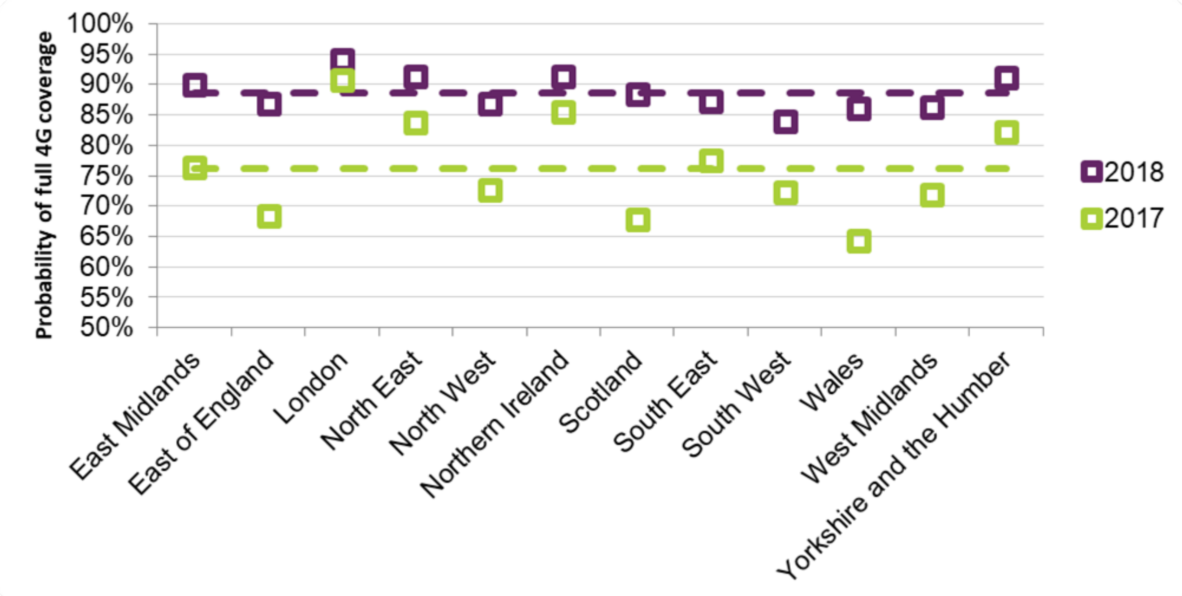




4G coverage results across time

Figure 12 below shows the adjusted 4G coverage across regions of the UK in 2017 and 2018, as well as the UK average for each year (represented by the dashed line). Adjusted full 4G coverage has increased across all regions between 2017 and 2018, with the UK average rising significantly from 76% in 2017 to 89% in 2018. As we explain in the Connected Nations Report[24], the rise in 4G coverage between 2017 and 2018 can be attributed to MNOs meeting two 4G coverage obligations at the end of 2017.

Figure 12: Adjusted geographic 4G coverage by all four MNOs, 2017 and 2018



We observe that the amount of unexplained regional variation has fallen dramatically between 2017 and 2018, which is evidenced by the reduction in the deviation of each region's adjusted 4G coverage from the average level.

This can further be evidenced by looking at Northern Ireland's adjusted coverage. In 2017 Northern Ireland's adjusted 4G coverage was significantly above the UK average. In our previous report[25], we argued that this result could be due to the beneficial effects of Northern Ireland having relatively more conducive land ownership characteristics, uniform population dispersion, and pure engineering-based buildout in relation to other nations and regions. However, we can see that Northern Ireland's adjusted 4G coverage in 2018 is no longer significantly higher than the national average and this is due to 4G mobile infrastructure improving across all regions.

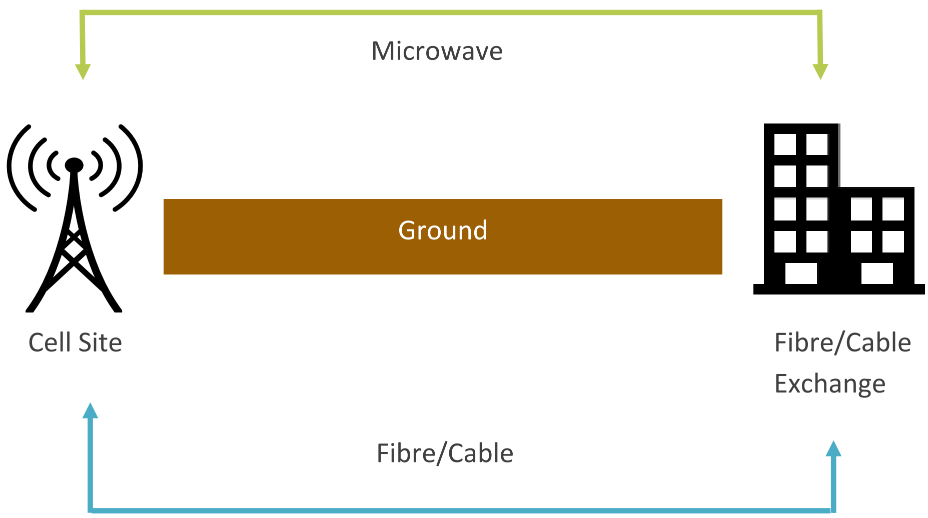
Overall, we have observed significant increases in 4G coverage. In addition, the reduction in regional variation of 4G coverage around the average is roughly in line with what we have seen for 3G technology in previous years. As 4G network rollout continues to develop, we expect to see less unexplained variation in coverage between regions.

[24] Ofcom (2018) Connected Nations
[25] Ofcom (2017) Economic Geography: An analysis of the determinants of 3G and 4G coverage in the UK

A1. Mobile Backhaul Variable

Cell sites deliver mobile services (voice, data and call) to the end mobile user but this requires the cell sites to be connected to its core network. Figure A1.1 below shows a simple example of how a cell site could be connected to a core network. This connection can be via fibre/cable, which directly links an MNO's cell site to its backhaul network at the relevant fibre or cable exchange. This connection may also be made wirelessly using fixed wireless links. Fixed links transmit data wirelessly from one fixed point to another using high frequency spectrum. One end of the link is installed on the mobile cell site and the other end of the link can be installed at an exchange, as in Figure A1.1

Figure A1.1: Example of fibre/cable (blue) and microwave (green) connection of a cell site to its core network[26]



We would expect that MNOs will face a higher cost to roll out mobile infrastructure to areas further away from its core network. As Figure A1.1 shows, this would involve the costly exercise of obtaining a point of presence from a fibre/cable operator at an exchange, extending the fibre connection from the exchange to the cell site or planning and installing microwave fixed links. This implies that areas that are further away from its core networks should also have worse coverage.

Therefore, we have included a relatively simple variable in the model that measures the distance of the pixel to its nearest mobile backhaul network location. These locations are either fibre/cable exchanges or microwave link points which are closest to the core network (belonging to MNOs).

[26] Note Figure A1.1 shows a very simple example of how a cell site is connected to its core network. Cell sites can be connected to the core network in many ways. For example, multiple cell sites could be connected to a central cell site, which is connected to the exchange building (similar to a hub and spoke model). There are likely to be different cost implications from using different types of connections to the core network. We have not modelled connections to the core network in this level of granularity as we believe that our general hypothesis holds on average. Further work will be needed to understand the impact of more complex connection types.

We have tested the addition of this new variable and we found that they do improve the predictive power of the model. However, we are aware that this variable has many limitations, which include the following:

- Firstly, using a variable that chooses the closest mobile backhaul location will mean that we are assuming that MNOs have an equal preference towards microwave or fibre connections to their core network. This might not be the case as some MNOs could prefer fibre or microwave due to historic reasons or commercial agreements. However, our general hypothesis will still hold on average and our variable, although not a perfect proxy, still adds to the predictive power of the model;
- Secondly, having a single variable relating to microwave backhaul ignores the fact that the effect of distance to backhaul could be different for each MNO. We chose one variable because it achieves a good balance between contributing to explaining mobile coverage while also creating the least amount of econometric problems (e.g. multicollinearity). We have further tested the usage of a distance metric for each MNO and the resulting effects are similar to just using one variable;[27]
- Thirdly, this variable also ignores microwave fixed link locations that belong to third party fixed links licensees. We have excluded these licensees because we could not verify if their microwave links were being used for mobile backhaul or other purposes. Additionally, we have tested the impact of adding these extra locations into the regression and the changes (e.g. in regional marginal effects) were minimal; and
- Fourthly, this variable could be partially accounting for other effects rather than just the cost factor mentioned above. For example, fibre and cable exchanges are typically situated in very densely populated areas, which implies that our distance metric could also be accounting for the lack of demand for mobile services in areas further away from heavily populated areas.[28] This means that the result of this variable would be partially confounded by other variables (e.g. demographic variables) but this variable still contributes to explaining the regional effects.

[27] Note that another complication arising from MNOs using their own microwave links is that there are additional complex backhaul sharing agreements between MNOs. We have not included this level of granularity in the model and further work will be needed to incorporate this into the model.

[28] Disentangling these two effects (one relating to cost while other relates to lack of demand) is difficult as we would need to change the model significantly. We will look at developing this in the future.

A2. Regression Results

Table A2.1: Regression table for 2017 and 2018 (1km coverage area radius)

	Dependent variable			
	2018 num_3G_ops	2017 num_3G_ops	2018 num_4G_ops	2017 num_4G_ops
Population				
ln_pop	0.182***	0.183***	0.206***	0.191***
Demography				
pct_abc1	0.00481***	0.00842***	0.00498***	0.00863***
pct_under25	-0.00131	-0.00369*	-5.55E-07	-0.00626***
pct_over60	-0.00877***	-0.0139***	-0.00873***	-0.0171***
Topography				
heightmedian	-0.000913***	-0.000950***	-0.00117***	-0.00130***
heightstdev	-0.00887***	-0.0102***	-0.0122***	-0.0109***
Backhaul network				
dist_backhaul	-0.000111***	-0.000126***	-0.000106***	-0.000145***
Urban_code				
urban	0.384***	0.338***	0.245***	0.156***
Region				
East of England	-0.402***	-0.392***	-0.196***	-0.306***
London	0.0982	0.153	0.366**	0.808***
North East	-0.0721	-0.016	0.119	0.344***
North West	0.0329	-0.0827	-0.196***	-0.150**
Northern Ireland	-0.0835	0.0897	0.122*	0.450***
Scotland	-0.131**	-0.337***	-0.105*	-0.329***
South East	-0.229***	-0.153***	-0.176***	0.0518
South West	-0.129***	-0.0947*	-0.361***	-0.161***
Wales	-0.00272	-0.371***	-0.252***	-0.448***
West Midlands	0.0788	-0.0844	-0.234***	-0.174**
Yorkshire & The Humber	-0.0105	0.0198	0.0946	0.272***
Observations	2,077,771	2,077,771	2,077,771	2,077,771

*** p<0.01, ** p<0.05, * p<0.1

The results in the table above require careful interpretation. Because the ordered probit model is a nonlinear model, the coefficients only give a partial picture of the impact of the variables on the probabilities of good coverage. The coefficients provide an indication of the direction and scale of effects as well as the relative importance of different variables. To compute the precise impacts on the probabilities of having good mobile coverage requires additional calculations. We discuss these computations in Annexes 3 and 4.

The directional effects of each explanatory variable in Table A2.1 can be summarised as follows:

- Population density: the variable ln_pop describes the population within the coverage area. As the size of the coverage area is fixed, ln_pop also describes population density. The positive and significant coefficient on ln_pop in both the 3G and 4G models across both years indicates that increased local population density has a positive effect on the likelihood of good 3G and 4G mobile coverage.
- Affluence: the variable pct_abc1 describes the affluence of the local population. A more affluent local population appears to have a positive influence on both 3G and 4G coverage across both years.
- Age structure: the variables pct_under25 and pct_over60 describe the age profile of the local population. For 3G and 4G, the coefficients on pct_over60 are negative and significant at the 1% level for both years. However, the coefficient for pct_under25 is only significant for 2017 and not 2018. This provides some evidence that a larger "working age" population has a positive impact on the likelihood of good mobile coverage.[29]
- Topography: the variables height_median and height_stdev measure the height of the coverage area above sea level and the variability in height within the coverage area respectively. For both 3G and 4G in both years, increasing both the height and variability of height appears to have a negative influence on coverage.
- Backhaul Network: the variable dist_backhaul measures the distance from the pixel to the closest location that allows access to the core backhaul network. A negative coefficient at the 1% significance level for both technologies and years implies that areas further away from the backhaul network will have a higher likelihood of receiving worse mobile coverage.
- Urbanicity: the variable urban_code identifies whether the coverage area is situated in a rural or urban area. The coefficients show the effect of being in the urban category relative to the comparator category (rural). The results indicate that there is a higher likelihood of good 3G and 4G mobile coverage across both years in urban areas than compared to rural areas.
- Region: this is a categorical (dummy) variable that identifies the region of the UK in which a coverage area is located. The coefficients on these variables show the effect of being within the particular region of interest relative to the comparator region (in this case the East Midlands). We examine how the region in which a given coverage area is located affects the likelihood of 3G/4G coverage in detail in Section 4.

[29] The insignificance of pct_under25 could be due to a mixture of different effects. Firstly, it could be a one-off effect for May 2018. Secondly, it could be that demand for mobile services are rising over time for those that are under 25.

A3. Marginal Effects – 3G

Table A3.1 below shows the effects of the continuous explanatory variables on the likelihood of full 3G coverage for 2018, calculated as described in Section 4. Table A3.2 shows the corresponding effects for 2017.

Table A3.1: Impact of continuous explanatory variables on average predicted probability of full 3G coverage, 2018

Variable	Lower	Upper	Difference
population	85.5%	92.9%	7.4%
pct_abc1	86.5%	88.2%	1.7%
pct_under25	87.3%	87.2%	-0.1%
pct_over60	88.3%	86.7%	-1.6%
height_median	88.2%	87.0%	-1.2%
height_stdev	89.1%	87.1%	-2.0%
dist_backhaul	92.1%	87.4%	-4.7%

Table A3.2: Impact of continuous explanatory variables on average predicted probability of full 3G coverage, 2017

Variable	Lower	Upper	Difference
population	80.9%	89.7%	8.8%
pct_abc1	81.6%	85.0%	3.4%
pct_under25	83.4%	83.0%	-0.4%
pct_over60	85.0%	82.2%	-2.8%
height_median	84.4%	82.9%	-1.5%
height_stdev	85.6%	82.9%	-2.7%
dist_backhaul	89.3%	82.9%	-6.4%

As both tables show, population at higher levels has the greatest effect of the explanatory variables on the likelihood of a coverage area having four 3G operators with good coverage. On average in 2018, a coverage area at the 75th percentile for population has a 7.4 percentage points higher probability of having full 3G coverage than a coverage area at the 25th percentile for population.

Figure A3.1 shows the average predicted probability of full coverage as the population of a coverage area varies continuously from the 25th to the 75th percentile for both years. We observe that population has a diminishing marginal impact. In other words, an extra 1000 people residing in a sparsely populated area has a greater impact on coverage than an extra 1000 people residing in a more densely populated area.

We do note that this diminishing marginal impact is getting slightly weaker over time in Figure A3.1 and we would largely expect this to happen. MNOs are increasingly deploying 3G mobile infrastructure to serve areas with lower population density (see Figures 2 and 3). This implies that mobile coverage is improving in low population density areas and, therefore, additional people residing in those areas will produce a smaller marginal effect over time.

Figure A3.1 Impact of population on average predicted probability of full 3G coverage, 2017 and 2018

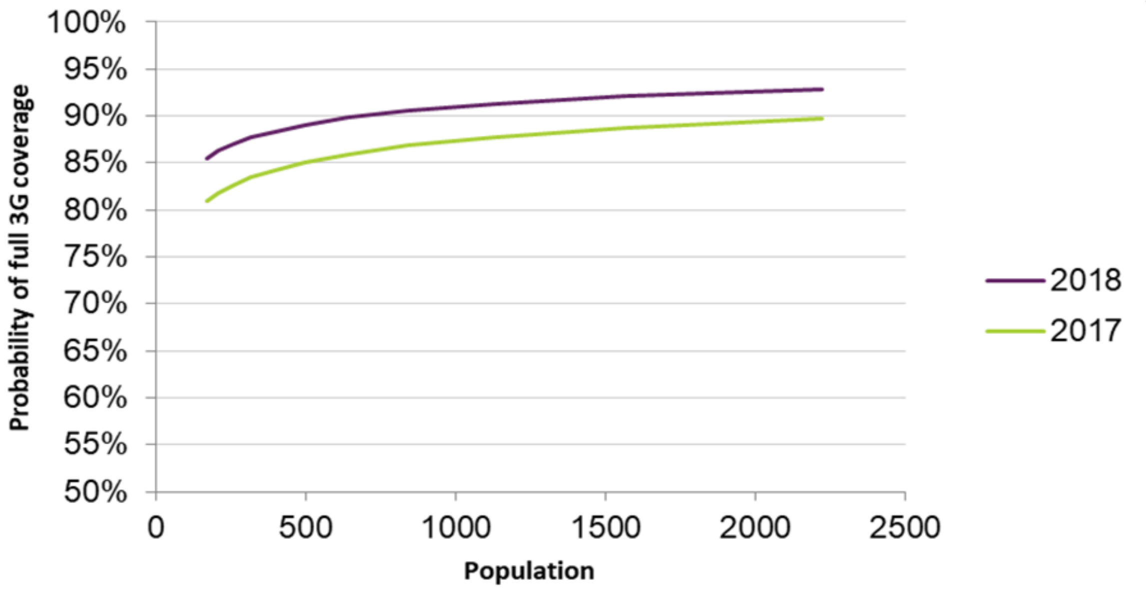


Table A3.3 shows the effect of the variable urban_code on the average predicted probability of full 3G coverage across both years. An area classified as urban has, on average, a higher probability of full coverage than a rural area for both 2017 and 2018.

Table A3.3: Average predicted probability of full 3G coverage by urban location, 2017 and 2018

Urbanity	2018	2017
Rural	87%	83%
Urban	92%	89%

A4. Marginal Effects – 4G

Table A4.1 shows the effects of the continuous explanatory variables on the likelihood of full 4G coverage in 2018. Table A4.2 shows the same for 4G coverage in 2017. Both tables were constructed in the same way as for 3G in Annex 3.

Table A4.1: Impact of continuous explanatory variables on average predicted probability of full 4G coverage, 2018

Variable	Lower	Upper	Difference
population	85.7%	93.5%	7.8%
pct_abc1	86.9%	88.5%	1.6%
pct_under25	87.7%	87.7%	-0.0%
pct_over60	88.6%	87.2%	-1.4%
height_median	88.9%	87.3%	-1.6%
height_stdev	90.1%	87.6%	-2.5%
dist_backhaul	92.1%	87.8%	-4.3%

Table A4.2: Impact of continuous explanatory variables on average predicted probability of full 4G coverage, 2017

Variable	Lower	Upper	Difference
population	68.8%	81.5%	12.7%
pct_abc1	71.0%	75.5%	4.5%
pct_under25	73.8%	72.7%	-1.1%
pct_over60	76.1%	71.6%	-4.5%
height_median	75.2%	72.5%	-2.7%
height_stdev	76.2%	72.5%	-3.7%
dist_backhaul	81.1%	71.1%	-10.0%

As Table A4.1 shows, out of all the explanatory variables, population at higher levels has the greatest effect on the likelihood of a coverage area having full 4G coverage in 2018. On average, a coverage area at the 75th percentile for population has a 7.8 percentage points higher probability of good 4G coverage than a coverage area at the 25th percentile. A similar outcome can be observed for population in Table A4.2 for full 4G coverage in 2017.

The 2017 marginal effects are quite different to the 2018 results. In particular, the level of the lower and upper marginal effects has increased significantly between 2017 and 2018. This can be explained by the rapid deployment of 4G infrastructure between 2017 and 2018 as reflected by Table 2 and Table 3 in Section 2.

Figure A4.1 Impact of population on average predicted probability of full 4G coverage, 2017 and 2018

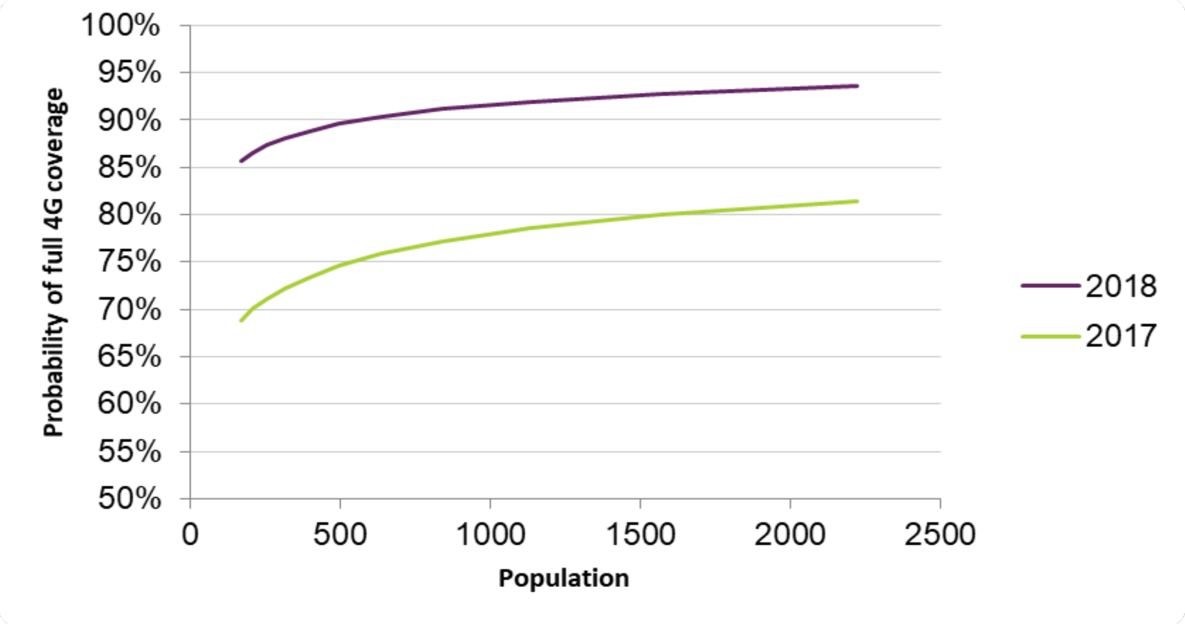


Figure A4.1 shows the average predicted probability for both 2017 and 2018 of full coverage as the population of a coverage area varies continuously from the 25th to the 75th percentile. The jump in predicted probability of full 4G coverage can be attributed to the same reasons outlined above.

Table A4.3 shows the effect of the variable urban_code on the average predicted probability of full 4G coverage in both 2017 and 2018. An area classified as urban has, on average, a higher probability of good coverage than a rural area.

Table A4.3: Average predicted probability of full 4G coverage by urban location, 2017 and 2018

Urbanity	2018	2017
Rural	88%	73%
Urban	91%	77%