ACRA(2015)36

MEETING: Advisory Committee on Resource Allocation
DATE OF MEETING: By correspondence – emailed on 7th March 2016
TITLE OF REPORT/PAPER: Costs of unavoidable smallness due to remoteness
AUTHORS: Allocations Project Team
ACTIONS REQUIRED: This paper is being circulated for information only. Since the previous ACRA paper, it has been updated to reflect the options for the selection criterion for remote sites of a population size of between 200k and 300k. This paper also includes a summary in the expectation that it will be published as part of the Technical guide to allocations.

Cost of unavoidable smallness due to remoteness

Executive Summary

1. This paper provides estimates of the unavoidable costs associated with providing health care services in remote areas.

2. The focus of the analysis is on acute providers. Unavoidable costs associated with community services, elective provision other than at a Type 1 A&E centre, or service provision in highly remote areas are not within the scope of the current analysis. NHS England envisages considering these in the next allocation round.

3. The analysis focuses on nine hospital sites that have been identified as being remote by NHS England. Previous analysis focused on eight sites, however following ACRA's alternative population threshold of 300k (see below), an additional site has been considered. The approach used to identify remote sites was discussed in a separate paper (ACRA (2015)18A).

4. Two hypotheses have been tested: (1) remote sites have a higher unit cost due to operating at lower scale; (2) remote sites have other additional unavoidable costs (unrelated to scale) (for example due to higher agency staff spend).

5. The results presented in this paper support the first hypothesis, showing there is a diseconomy of scale incurred. When testing the second hypothesis, evidence of additional costs is not identified. However, this could be driven by the limitations of the empirical framework. The remoteness adjustments provided in this paper therefore focus on additional sub-scale costs.

6. The estimated impact of operating at sub-scale cost is between £0m and £7.8m depending on the characteristics of the site and the methodological assumptions applied. In particular, a series of sensitivity checks have been performed.

    - **Activity.** Two alternative definitions of activity have been used to compute the expected additional cost of a remote site. The first one is actual activity, which was used in the analysis presented previously. An alternative measure is the activity implied by the population catchment (population-based activity). This was estimated through a model that links activity to population. The advantage of the population-based approach is that it focuses on the key structural issue – the size of population requiring access to services – rather

than level of activity as such. This means that other factors which make activity levels higher or lower (such as patient choice/competition/quality of care) are not compensated for. The drawback of a population based approach is that activity levels have to be estimated, and this is subject to modelling error.

- **Benchmark site.** A benchmark site is needed to form a point of reference from which to measure the impact of sub-scale costs. The benchmark site therefore refers to the size of (an efficient) hospital site. Four alternative benchmarks have been used: (1) A benchmark determined by a site with 200K population catchment; (2) A site with 250K population catchment; (3) a site with average population catchment (i.e. 300K); and (4) a site with average activity. The latter two effectively reflect a site of average size. The 200K benchmark corresponds to previous work that identifies small providers. The 250K provides a benchmark that is between a "small" and average size site. The advantage of the average size approach is that it is consistent with the approach used in setting tariff. It is also consistent with Monitor's benchmark approach used in Local Modifications.

7. ACRA has recommended to use a population catchment threshold, and therefore benchmark, between 200,000 and 300,000 people and suggested that this should be set in a way that ensures the adjustment is applied only to those areas facing the most significant challenge.

1. Introduction

For the General and Acute target revenue allocation formula, the previous allocation round (2014/15 -2015/16) was made of four key building blocks:

1. Relative need reflecting the relative health care need across CCGs primarily on the basis of demographics and health status;
2. Adjustment for health inequalities;
3. Adjustment for variation in input prices using the Market Forces Factor; and
4. An ambulance emergency cost adjustment.

For the 2016/17 allocation round, NHS England ("NHSE") is considering updating the allocations formula to take into account the unavoidable costs associated with providing some health care services in remote areas.[1]

In principle, an adjustment for remoteness is justified if the following conditions hold.

1. **Remote providers face higher unit costs**. Remote providers typically serve smaller population catchments, operate at lower scale and therefore may incur higher unit costs than other providers (this is hypothesis 1 below). Further, remote providers may face higher staff costs if difficulties in recruiting and retaining permanent staff result in a need for greater utilisation of agency staff. Other possible additional costs may be related to longer length of stay and scheduling inefficiencies (e.g. remote providers may not be able to take advantage of shared rotas between sites) (this is hypothesis 2 below).
2. **Additional costs associated with remoteness are unavoidable**. Providing services in remote areas is primarily driven by clinical considerations (especially with regard to emergency services access) and less by commissioner choice. Therefore, for certain services, the additional costs associated with provision of services in remote areas are unavoidable for commissioners.

The objective of this paper is twofold:

- Test the hypothesis that remote hospital sites face higher costs; and
- Quantify the unavoidable cost associated with remoteness.[2]

A critical aspect of the methodology is the type of services considered within scope. One approach would be to consider only emergency services given that these services are more subject to access considerations/clinical safety standards. However, there are two limitations with an emergency services approach. A hospital configuration with only emergency services does not reflect how services are provided in practice. Further, it is analytically difficult to control for economies of scope and therefore accurately estimate the remoteness uplift for emergency services alone (the result of this is that the degree of additional cost will be underestimated). Given these limitations, the remoteness hypothesis is evaluated for total services delivered at sites which provide Type 1 A&E services.

[1] NHSE has previously investigated whether an adjustment for health care provision in rural rather than remote or sparsely populated areas. Analytical work and engagement with stakeholders suggested that it is remoteness rather than rurality that may drive additional costs.
[2] A separate paper has been developed on the identification of remote sites. This paper focuses only on the quantification of costs for those sites that have been identified as being remote.

Compensating commissioners for unavoidable costs is consistent with the Health and Social Care Act (2012), which makes provision for providers to be compensated for unavoidable higher costs via local modifications ("LM"). Once Monitor has approved a LM for a given provider, CCGs must pay the higher prices set by Monitor. However, to date, allocations to CCGs have not been adjusted to account for the additional costs a commissioner might face in this regard.[3]

The remaining of this paper is organised as follows:

- Section 2 describes the methodology;
- Section 3 presents estimates of the unavoidable cost associated with remoteness; and
- The appendix presents technical details and sensitivity checks.

2. Methodology

The premise that remote hospital sites face higher costs than other sites is tested by considering the following two hypotheses:

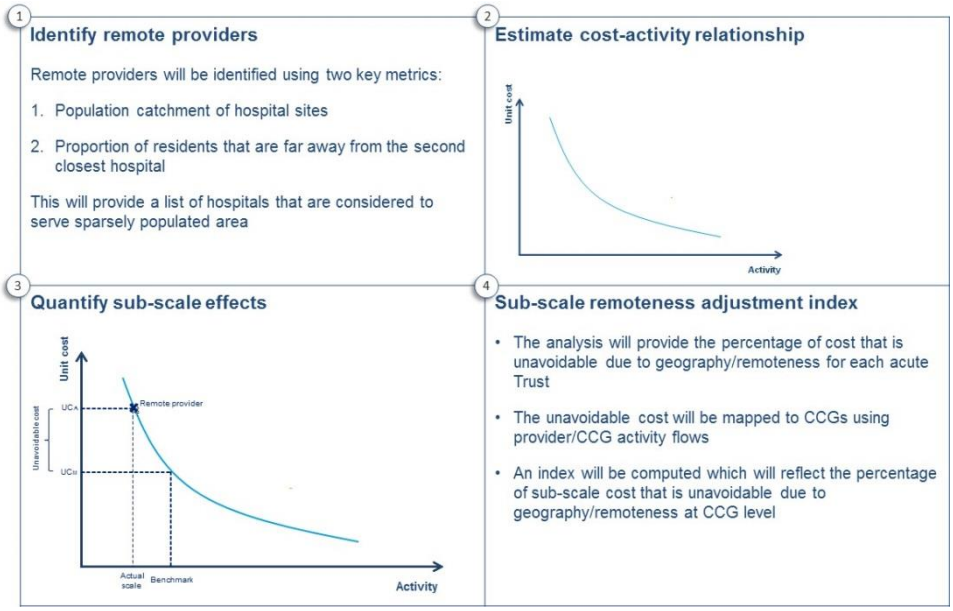
Hypothesis 1: Remote sites have higher unit costs due to operating at lower scale; and

Hypothesis 2: Remote sites face additional unavoidable costs (e.g. costs associated with higher utilisation of agency staff).

Hypothesis 1

The first hypothesis would be true only if there are economies of scale in health care provision, and remote providers operate at a lower scale which means their unit costs are higher than non-remote providers. The approach used is built around these considerations and is set out in Figure 1.

Figure 1: Quantifying sub-scale effects



[3] The first LM was granted by Monitor in 2015 to the University Hospitals Morecambe Bay NHS Foundation Trust. Several other Trusts have applied for LM but their application has been rejected.

Hypothesis 2

The hypothesis that remote providers face higher unit costs (other than sub-scale) is tested by augmenting the econometric model with a "remoteness" variable. Although this is, in principle, a reasonable framework to test this hypothesis, in practice, there are two main limitations. First, specific type of costs (e.g. agency costs) may not be a function of remoteness but rather of attractiveness of the location and career opportunities. Second, the sample of remote sites identified is relatively small therefore the model might not have the power to identify such impacts.

The remaining of this section describes the methodology in detail.

2.1 Step 1: Identification of remote sites

NHS England's work on identifying remote sites uses two criteria:

1. Hospital's maximum population catchment; and
2. The second nearest hospital for a significant portion of the population (e.g. at least 10%) is greater than 60 minutes.

Effectively, remote areas are defined as those that have a relatively small population catchment and for which a significant portion of the population is primarily served by a single site.

This analysis is discussed in a separate paper developed by NHSE; nine remote sites are identified.[4]

2.2 Step 2: Econometric modelling

The relationship between cost and scale is estimated econometrically using the model specification set out in Figure 2.[5] Effectively, provider cost is regressed on the following factors.

- **Activity, case-mix adjusted.** Number of episodes, admissions, attendances, etc. For each provider, the activity variable is the sum of weighted HRG-level activity. The weights reflect the national unit cost of an HRG relative to the national unit cost across all HRGs (see appendix for more details). **The activity variable facilitates testing of Hypothesis 1.**
- **Activity concentration**. This reflects the concentration of activity between sites. Trusts with two equally sized sites might have different unit costs, all other things being equal, relative to trusts that have, for instance, one large and one small site. This variable explores this premise.

[4] ACRA(2015)18A: Unavoidable smallness due to remoteness - identifying remote hospitals.
[5] A similar econometric framework has been used to estimate the relationship between cost and its drivers by other studies (for instance, see Deloitte, "Evidence for the 2015/16 national tariff efficiency factor" commissioned by Monitor and Jacobs et al. (2006), "Measuring efficiency in the health care: Analytic techniques and health policy") although the focus of their analysis was to measure efficiency rather than the unavoidable cost of remoteness.

- **Remoteness**. Two alternative proxies are considered. The first one is a dummy variable that identifies remote sites. The second is a remoteness index which reflects the percentage of population served that is more than 60' away from the second nearest provider.[6] **These proxies, applied separately, facilitate testing of Hypothesis 2**.
- **Case-mix complexity**. While this is primarily captured by case-mix adjusted activity (as above), a number of variables reflecting patient age, gender and ethnicity are included in the model to capture any additional effects.
- **Input prices.** Given the time-series dimension of the sample, variation of prices over time is captured by the cost uplift factor. Variation in input prices between providers is controlled through the Market Forces Factor ("MFF"). It is acknowledged that the MFF may be out-of-date and/or capture variation in input prices insufficiently, which may impact the estimates of the activity and remoteness.
- **Time dummies**. Year dummies are included in the model to control for time-varying factors that are common across providers.

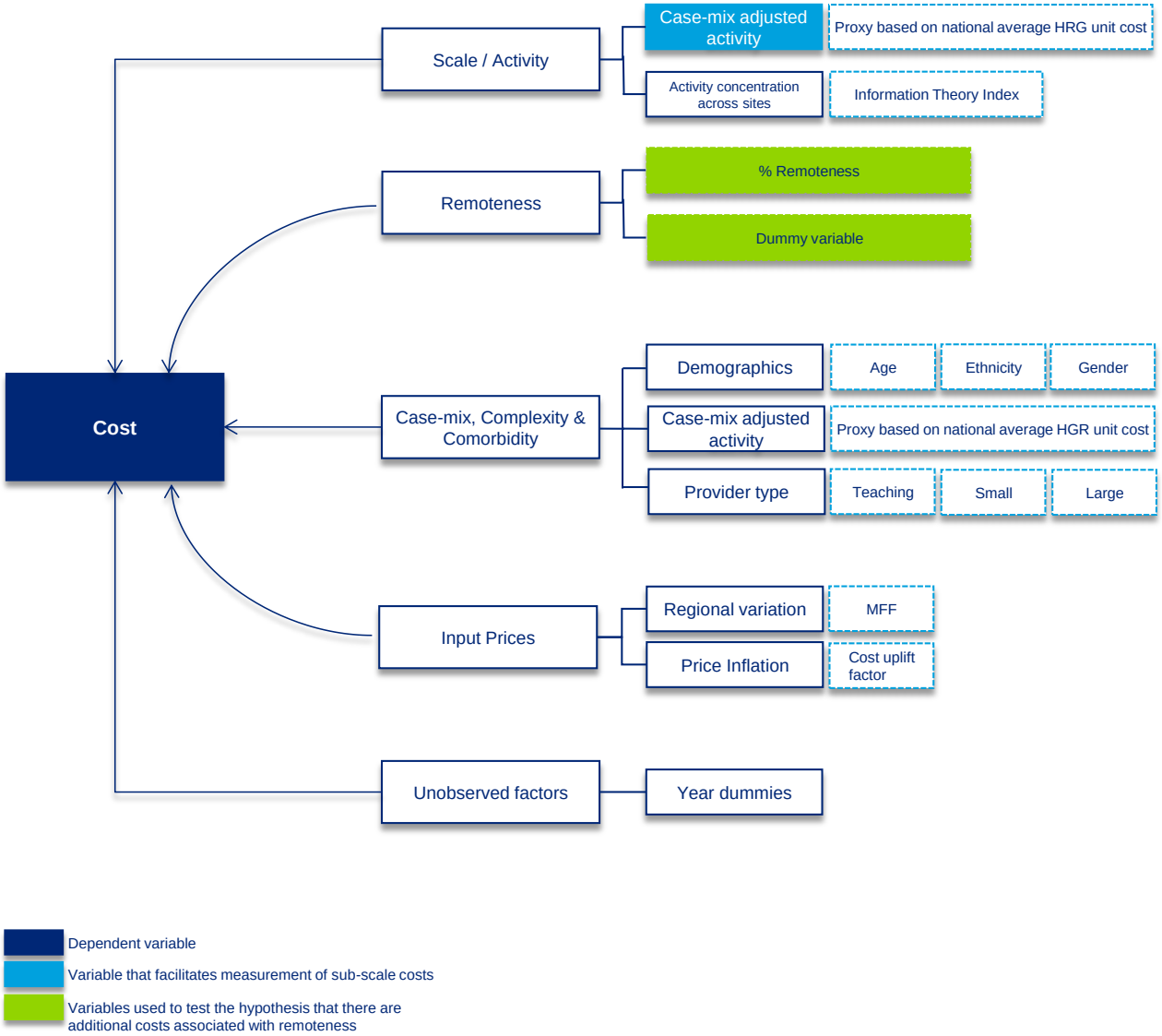
Given that the focus of the analysis is hospital sites as opposed to Trusts, the relationship between cost, activity and remoteness should be estimated at hospital/site level. However, a number of data limitations exist. Firstly, cost data by site are not widely available. Secondly, while site-level activity data are available from HES they don't cover A&E. Given this, the model uses Trust-level data and information on number of sites by trust to compute site-level cost and activity variables. These variables reflect the average cost and activity per site for a given Trust. Differences in activity concentration across sites, which may impact on the model estimates, are controlled for by an activity concentration index as discussed above.

Data from 2009/11 to 2013/14 have been obtained from various sources:

- Cost and activity data, extracted from Reference Costs;
- Remoteness index, from the NHS England allocations team;
- Demographic information, from HES;
- Provider type, from Hospital Estates and Facilities Statistics;
- Number of sites, from Hospital Estates and Facilities Statistics; and
- MFF, extracted from Reference costs.

[6] The data for these variables have been taken from: "ACRA(2015)18A: Unavoidable smallness due to remoteness - identifying remote hospitals".

Figure 2: Econometric model specification



The model is estimated by panel Random Effects using data across acute providers over the period 2009/10 to 2013/14.[7] Cost and activity reflect all services excluding mental health and community.[8] A wide range of sensitivity tests have been conducted (see Appendix) which indicate that the key results of the model are robust to alternate specifications and assumptions.

The results of the estimation are presented in Table 1. The key insights are:

- The elasticity of cost with respect to activity is less than one, suggesting economies of scale and therefore potential sub-scale costs for remote providers operating at lower scale.[9]
- The remoteness variable is statistically insignificant indicating that this model does not identify further additional costs associated with remoteness. However, as it is argued in Section 2, these results should be interpreted with care.

Table 1: Model output[10,11]

Independent Variables	Coefficients
Activity, case-mix adjusted (log)	0.891***
Remoteness dummy	-0.0266
Activity concentration	-0.0796
% patients > 75 years old	-0.00162
% female patients	-0.00580***
% patients BAME	0.00159**
% patients emergency	-0.00270***
Small provider dummy	-0.0230**
Large provider dummy	0.0108
Teaching provider dummy	0.0683***
MFF	0.647***
Time dummy (2010/11)	0.000627
Time dummy (2011/12)	-0.0483***
Time dummy (2012/13)	-0.0455***
Time dummy (2013/14)	-0.0653***
** p-value<.10 ** p<0.05 *** p<0.01*	

Notes: Medium provider dummy is the base against which, Small, Large and Teaching dummies are evaluated. The results remain effectively the same when the provider dummy variables are excluded.

[7] Panel data models exploit the panel structure of the dataset and provide more accurate estimates than conventional regression techniques. The Random Effects estimator has been chosen over Fixed Effects as several independent variables do not vary much over time, which makes Fixed Effects unsuitable.

[8] A model using only emergency services and maternity has also been estimated and presented in the Appendix.

[9] The 95% confidence interval for the activity variable is [.87, .92].

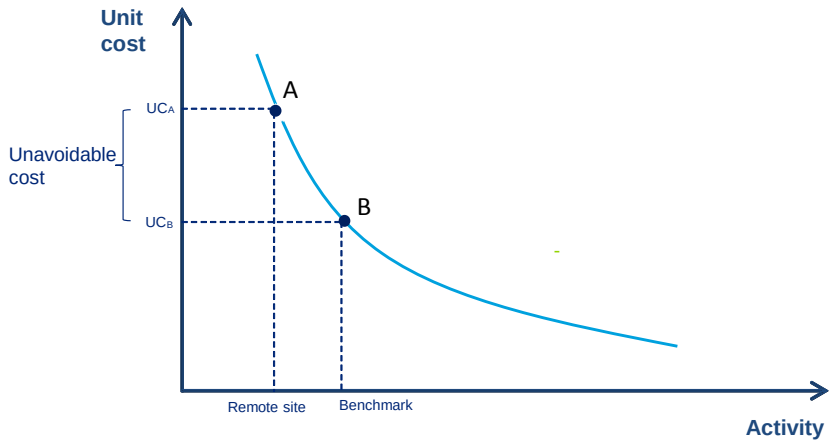
[10] The dependent variable is the logarithm of cost deflated by the cost uplift factor.

[11] The main focus of the analysis is the activity and remoteness variable therefore the remaining results are not discussed in depth. However, it is worth noting that a number of other variables have statistically significant impact on cost. For instance, the higher the proportion of female patients, the lower the cost. The MFF factor although has the correct sign, its magnitude is significantly less than one. This may be the result of MFF being out of date and that it does not accurately reflect input price variation. Finally, most time dummies are negative and statistically significant suggesting that NHS Trusts have reduced cost by an average of 1.6% per year over the period considered, all other things being equal. This is relative consistent with the frontier shift estimates reported in "Evidence for the 2015/16 national tariff efficiency factor" although the two studies use different sample periods. Further, this finding is not the focus of this analysis.

2.3 Step 3: Quantification of sub-scale effects

This step quantifies the sub-scale cost of remote hospital sites relative to a benchmark hospital. In particular, the cost-activity relationship (estimated econometrically) is used to determine the expected cost of the remote site given its scale – see point A in Figure 3. Likewise, the expected cost of the benchmark site is determined using the cost activity curve – see point B. The difference between these two points (that is, the expected cost of the remote site and the benchmark site) reflects the estimated sub-scale cost associated with remoteness.

Figure 3: Quantification of sub-scale effects



Two critical aspects of the methodology relate to the:

- Determination of the benchmark (what should remote sites be compared to); and
- Definition of scale/activity.

Determining the benchmark

In applying this approach to estimating the sub-scale costs of remoteness, clearly the benchmark which is used is central to the degree of additional cost identified. A number of alternative benchmarks have been used.

- **200,000 population catchment.** This reflects the threshold for the "smallest" hospitals. Monitor's work on small hospitals[12] identified the characteristics of such hospitals, with "smallest" being defined as those with a population catchment of less than 195,000.
- **Mean population catchment (300K).** Given that prices are determined using national average cost and therefore computed on the basis of an average provider or site, the sub-scale cost could be benchmarked against an average hospital. Further, this figure is consistent with the definition of remoteness used by NHSE to identify remote hospitals in this work programme.

[12] Monitor (2014), "Facing the future: smaller acute providers".

- **Mean activity.** This reflects the scale of an average site as the above measure but it is based on actual activity rather than population catchment.[13]
- **250,000 population catchment.** This provides an alternative small site benchmark reflecting a scale that is between a "smallest" and average size hospital.

Defining scale/activity for remote sites

In order to determine costs at the remote site its scale needs to be determined (so a reading can be taken from the cost activity curve).There are two ways to do this:

- **Activity-based approach.** This uses the actual activity of the remote site.[14]
- **Population-based approach.** This uses a given hospital's population catchment to project expected activity of the hospital.

A population-based approach focuses on the key structural issue faced by the provider – the size of population requiring access to services – rather than level of activity as such. This means that other factors which make activity levels higher or lower are not compensated for. For example, outturn activity levels might be affected by:

a. The ability of the provider to attract patients relative to other providers in the area; and

b. Discretionary commissioner decisions about how to configure services.

The drawback of a population based approach is that activity levels have to be projected, and this is subject to measurement error. Further, the relationship between activity and population may be different for remote sites. In particular, if average travel times are longer in remote areas then the ratio of activity per population for remote sites might be lower than average. The activity based option also better reflects the challenges faced by the commissioner (supporting a critical provider in its current configuration) rather than the provider (finding a financially sustainable model).

The population-based approach estimates expected activity for a given population catchment through an econometric model that links activity to population and demographics (see Appendix for further information). This is a relatively simple model using aggregate level data.[15] An alternative approach, which is in principle more robust, would be to apply the framework used to estimate relative need in the target allocation formula using patient-level data. However, there has been insufficient time to undertake the latter approach for this exercise.

All alternative approaches described above each have advantages and disadvantages and they may generate significantly different sub-scale cost estimates. Given this, the sensitivity

[13] An alternative benchmark could be determined on the basis of "minimum economic scale", which is defined as the level of activity beyond which additional increases in activity have a small impact on unit cost (e.g. point where cost-activity curve flattens). However, this is difficult to determine given the model specification and inherent challenges around estimating complex non-linear relationships.
[14] The econometric model is specified in terms of case-mix adjusted activity whilst case-mix weights are computed from reference cost data, which are not available at site level. Site-level case-mix adjusted activity is inferred by applying site-level activity shares inferred from HES on trust level case-mix adjusted activity.
[15] Some demographic variables are also counter-intuitive.

of remoteness uplift results is examined across all these alternative approaches. This provides an understanding of the range of possible values for a remoteness uplift.

2.4 Step 4: Construction of the remoteness adjustment index

The unavoidable cost associated with remoteness needs to be expressed as an index at the CCG level in order to adjust the CCG target allocation formula. As the cost uplifts are estimated at the site level, two steps are performed to express the site uplift at the CCG level. These are shown below. The monetary values are calculated as the difference in unit cost between the remote site and the benchmark site, multiplied by the activity of the remote site. This requires two steps:

- **Convert site uplift to provider uplift.** The cost uplift (£) at the provider level is the sum of all the remote site uplifts belonging to a particular Trust.
- **Convert provider uplift to CCG uplift.** The provider uplift is converted to a CCG uplift by utilising the activity flows in the provider-purchaser matrix.[16]

3. Results

Two sets of results are presented in this section. Table 2 sets out the sub-scale cost using the activity-based approach to assess remote sites. Table 3 sets out the sub-scale cost using the population-based approach to assess remote sites. Both tables report results across all four alternative methods used to determine the benchmark site.

The estimated sub-scale cost at a site level using the activity-based approach ranges from £0 to £7.8m depending on the benchmark used.

The zero cost estimated for site 4 in Table 2 is due to the relative large level of activity compared to its population catchment. This larger than anticipated site 4 activity level may be driven by the influx of patients beyond its population catchment (perhaps during peak summer periods).

Table 2: Sub-scale cost uplifts (000s) of remote sites – Activity-based approach

Site (Provider)	Population catchment	Benchmark			
		200K[17] (000s)	250K (000s)	Mean Population (000s)	Mean Activity (000s)
Site 1	111,207	£4,402 (4.8%)	£5,900 (6.5%)	£7,216 (7.9%)	£7,475 (8.1%)
Site 2	130,892	£5,390 (7.3%)	£6,562 (8.9%)	£7,590 (10.3%)	£7,793 (10.6%)
Site 3	138,393	£4,506 (5.0%)	£5,976 (6.7%)	£7,266 (8.1%)	£7,521 (8.3%)
Site 4	169,852	0	£639 (0.4%)	£2,969 (1.9%)	£3,429 (2.2%)

[16] The approach followed is akin to that followed when converting provider level MFF to CCG level MFF.
[17] The 200K population catchment reflects the bottom 22 percentile of the distribution of sites.

Site 5	178,338	£967 (.7%)	£3,162 (2.4%)	£5,090 (3.9%)	£5,471 (4.2%)
Site 6	182,303	£3,842 (3.9%)	£5,480 (5.5%)	£6,919 (7.0%)	£7,202 (7.3%)
Site 7	190,677	0	£1,481 (1.0%)	£3,686 (2.5%)	£4,120 (2.8%)
Site 8	194,103	£3,203 (3.0%)	£4,983 (4.6%)	£6,546 (6.1%)	£6,854 (6.4%)
Site 9	256,816	0	£1,539 (1.1%)	£3,735 (2.6%)	£4,167 (2.9%)
Total		£22,310	£35,722	£50,948	£54,034

Notes: Population-based approach compensates remote sites on the basis of the size of population served rather than actual activity (e.g. activity-based approach). *200K* and 250K reflect the population catchment of the benchmark site against which the uplift is computed; *Mean population* and *Mean activity* evaluate the sub-scale cost against a benchmark that reflects the size of average hospital defined in terms of population catchment and activity, respectively.

The total compensation is c. £22, £36, £50 and £54 million in the 200K, 250K, mean population, and mean activity approaches respectively.

The population-based approach delivers estimates that vary less than those of the activity-based approach and provides percentage uplifts that are a linear function of the size of population (the smaller the catchment, the greater the percentage uplift). The sub-scale costs at a site level are shown in Table 3 and range from c. £300K to £7.4m.

Table 3: Sub-scale cost uplifts (000s) of remote sites – Population-based approach

Site (Provider)	Population catchment	Benchmark			
		200K[18] (000s)	250K (000s)	Mean Population (000s)	Mean Activity (000s)
Site 1	111,207	£4,206 (4.4%)	£5,755 (6.1%)	£7,116 (7.5%)	£7,384 (7.8%)
Site 2	130,892	£3,390 (3.2%)	£5,123 (4.9%)	£6,658 (6.4%)	£6,959 (6.6%)
Site 3	138,393	£3,057 (2.8%)	£4,867 (4.4%)	£6,457 (5.9%)	£6,770 (6.2%)
Site 4	169,852	£1,557 (1.3%)	£3,651 (2.9%)	£5,491 (4.4%)	£5,853 (4.7%)
Site 5	178,338	£1,129 (.9%)	£3,297 (2.6%)	£5,202 (4.1%)	£5,577 (4.4%)
Site 6	182,303	£926 (.7%)	£3,129 (2.4%)	£5,063 (3.9%)	£5,444 (4.2%)
Site 7	190,677	£492 (.4%)	£2,766 (2.1%)	£4,763 (3.6%)	£5,157 (3.9%)
Site 8	194,103	£312 (.2%)	£2,615 (1.9%)	£4,638 (3.4%)	£5,036 (3.7%)
Site 9	256,816	0	0	£2,133 (1.3%)	£2,620 (1.6%)
Total		£11,677	£31,203	£40,863	£43,843

Notes: Population-based approach compensates remote sites on the basis of the size of population served rather than actual activity (e.g. activity-based approach). *200K* and 250K reflect the population catchment of the

[18] The 200K population catchment reflects the bottom 22 percentile of the distribution of sites.

benchmark site against which the uplift is computed; *Mean population* and *Mean activity* evaluate the sub-scale cost against a benchmark that reflects the size of average hospital defined in terms of population catchment and activity, respectively.

3.1 Adjusting for measurement error in the population based approach

Table 4 sets out the model output for the actual and projected activity (e.g. activity implied by the population-based approach). This helps gauge the degree of measurement error associated with this approach. There are sites for which the difference between actual and projected activity is quite large (for instance, actual activity for site 2 is 833,401 whereas projected activity is 1,237,919). It is difficult to explain this with structural factors alone and it is therefore likely to reflect measurement error. To compensate sites adequately, a downward adjustment is applied to projected activity. Further analysis, based on the directional of this bias, could be conducted as part further work on this area.

Table 5 and Table 6 set out the sub-scale cost by revising projected activity by -5%, -10% and -20%. Table 5 presents results using 200K population as the benchmark whereas Table 6 uses mean population.[19]

Table 4: Actual vs. projected activity

Site (Provider)	Activity	Projected Activity (using the population-based approach)
Site 1	1,065,918	1,102,320
Site 2	833,401	1,237,919
Site 3	1,045,798	1,288,011
Site 4	1,888,113	1,490,193
Site 5	1,562,294	1,542,818
Site 6	1,165,408	1,567,157
Site 7	1,786,064	1,618,068
Site 8	1,266,272	1,638,710
Site 9	1,778,875	2,000,114

[19] The downward adjustment increases the cost uplift and therefore errs on the conservative side when making a remoteness adjustment.

4. Conclusion

This paper investigates the hypothesis that hospitals that provide health care services in remote areas face higher costs than other providers. The results of the analysis suggest that:

- Remote hospitals operate at a lower scale and have higher unit costs as they cannot take advantage of economies of scale to the same extent as larger providers. The magnitude of sub-scale cost associated with remoteness depends on the underlying methodological assumptions and ranges between £0 and £7.8 million; and
- The analysis does not find support for the hypothesis that remote hospitals face additional unavoidable costs, other than sub-scale. However, this result should be interpreted with caution as it may be an artefact of the methodology, the available data and the underlying relationships being modelled:
  - Agency costs, for instance, may not be a function of remoteness but rather attractiveness of the location and career opportunities;
  - There are only nine remote sites in the sample and therefore the model may have been unable to isolate causal effects from noise; and
  - This is particularly the case if the relationship between cost and remoteness is not systematic across hospitals but more provider-specific.

In the financial year 2015-16, Monitor granted the first LM to Morecambe Bay ("MB"). The results presented in this paper cannot be contrasted with Monitor's analysis as they are anonymised. However, it is noted that the NHSE's methodology used to identify remote sites and quantify the unavoidable cost is different from Monitor's approach. The key differences are:

- Scope of services. Monitor applies a LM uplift on "six essential services" compared to all services considered by NHSE;
- Granularity. Monitor uses HRG comparisons to estimate the uplift compared to NHSE's analysis where aggregate data can help address confounding factors;
- Number of sites. Monitor considers all three MB sites whereas the NHSE analysis has identified only one site which requires a remoteness uplift; and
- Nature of uplift. NHSE's analysis focuses on sub-scale costs whereas Monitor's analysis may cover broader costs of remoteness.

Table 5: Sub-scale cost uplift adjusted for measurement error in the population-based approach, benchmark site == 200K

Site (Provider)	Cost Uplift ('000 £, %) - Benchmark at 200K			
	Population approach	Population-based approach -5% adjustment	Population-based approach -10% adjustment	Population-based approach -20% adjustment
Site 1	£4,205,593 (4.4%)	£4,498,474 (5.0%)	£4,767,061 (5.5%)	£5,225,156 (6.7%)
Site 2	£3,389,824 (3.2%)	£3,778,292 (3.8%)	£4,139,820 (4.3%)	£4,775,185 (5.6%)
Site 3	£3,056,745 (2.8%)	£3,481,940 (3.4%)	£3,879,228 (3.9%)	£4,582,956 (5.1%)
Site 4	£1,556,823 (1.3%)	£2,137,210 (1.8%)	£2,685,818 (2.4%)	£3,679,582 (3.6%)
Site 5	£1,128,722 (.9%)	£1,751,189 (1.4%)	£2,340,876 (2.0%)	£3,413,551 (3.3%)
Site 6	£925,812 (.7%)	£1,567,960 (1.3%)	£2,176,868 (1.8%)	£3,286,486 (3.1%)
Site 7	£491,663 (.4%)	£1,175,412 (.9%)	£1,824,960 (1.5%)	£3,012,731 (2.8%)
Site 8	£311,965 (.2%)	£1,012,745 (.8%)	£1,678,936 (1.4%)	£2,898,728 (2.6%)
Site 9	0	0	0	0

Notes: Population-based approach compensates remote sites on the basis of the size of population served rather than actual activity (e.g. activity-based approach). 200K reflects the population catchment of the benchmark site against which the uplift is computed; Mean evaluates the sub-scale cost against a benchmark that reflects average hospital size. *Population-approach – x% adjustment* reduces the activity implied by population catchment by *x*%.

Table 6: Sub-scale cost uplift adjusted for measurement error in the population-based approach, benchmark site == population mean

Site (Provider)	Cost Uplift ('000 £, %) - Benchmark at population mean			
	Population approach	Population-based approach -5% adjustment	Population-based approach -10% adjustment	Population-based approach -20% adjustment
Site 1	£7,115,668 (7.5%)	£7,263,045 (8.0%)	£7,386,129 (8.6%)	£7,553,216 (9.7%)
Site 2	£6,657,876 (6.4%)	£6,882,941 (6.9%)	£7,081,066 (7.4%)	£7,389,627 (8.6%)
Site 3	£6,457,037 (5.9%)	£6,712,218 (6.5%)	£6,939,491 (7.0%)	£7,303,190 (8.2%)
Site 4	£5,490,866 (4.4%)	£5,874,552 (5.0%)	£6,226,457 (5.5%)	£6,826,817 (6.7%)
Site 5	£5,201,691 (4.1%)	£5,620,510 (4.6%)	£6,006,549 (5.2%)	£6,671,926 (6.4%)
Site 6	£5,063,038 (3.9%)	£5,498,324 (4.5%)	£5,900,371 (5.0%)	£6,596,266 (6.2%)
Site 7	£4,763,290 (3.6%)	£5,233,457 (4.1%)	£5,669,424 (4.7%)	£6,430,032 (5.9%)
Site 8	£4,638,086 (3.4%)	£5,122,560 (4.0%)	£5,572,445 (4.5%)	£6,359,625 (5.8%)
Site 9	£2,133,322 (1.3%)	£2,882,245 (1.9%)	£3,589,858 (2.5%)	£4,870,617 (3.7%)

Notes: Population-based approach compensates remote sites on the basis of the size of population served rather than actual activity (e.g. activity-based approach). 200K reflects the population catchment of the benchmark site against which the uplift is computed; Mean evaluates the sub-scale cost against a benchmark that reflects average hospital size. *Population-approach – x% adjustment* reduces the activity implied by population catchment by x%.

5. APPENDIX

A.1. Case-mix adjusted activity

The case-mix adjusted activity used in the econometrics model has been computed in three steps.

1. Deflate providers' costs by the market forces factor.

$$C_{MFF_{i,j}} = \frac{C_{i,j}}{MFF_j}$$

2. Compute complexity weights for each HRG.

$$w_i = \frac{\frac{\sum_{j}^{N} C_{MFF_{i,j}}}{\sum_{j}^{N} A_{i,j}}}{\frac{\sum_{i}^{k}\sum_{j}^{N} C_{MFF_{ij}}}{\sum_{i}^{k}\sum_{j}^{N} A_{ij}}} = \frac{National\ average\ \cos t\ \ for\ HRGi}{National\ average\ \cos t}$$

3. Apply weights to the raw activity measure.

$$ACM_j = \sum_{i}^{k} w_i A_{i,j}$$

where

i : HRG i

j : Provider j

ACM : Activity, case - mix adjusted

A : Unadjusted activity

$C_{i,j}$: Total cost incurred by provider j for HRG i

C_{MFF} : Cost deflated by MFF

The weights have been computed separately for each year due to temporal changes in HRG coding.

A.2. Econometric model sensitivity

A series of sensitivity checks have been conducted in terms of:

- Years considered in the analysis;
- Treatment of outliers; and
- Remoteness proxies.

Years considered in the analysis

To test the stability of the results over time, the baseline specification is estimated by consecutively excluding one year from the sample. The results are shown in Table 7.

Table 7: Stability over time

Variable	Baseline	Dropping 2009	Dropping 2010	Dropping 2011	Dropping 2012	Dropping 2013
Activity, case-mix adjusted (log)	0.89***	0.91***	0.90***	0.90***	0.92***	0.89***

Treatment of outliers

Outliers are defined as those observations whose model residuals deviate more than +/- 3.5 times the standard error of the residuals. Table 8 shows the results from the baseline model alongside the baseline model estimated by excluding outliers.

Table 8: Model estimates after controlling for outliers

Independent Variables	Baseline	Baseline excluding outliers
Activity, case-mix adjusted (log)	0.891***	0.914***
Remoteness dummy	-0.0266	0.0140
Activity concentration	-0.0796	-0.0296
% patients > 75 years old	-0.00162	-0.00201
% female patients	-0.00580***	-0.00585***
% patients BAME	0.00159**	0.00136**
% patients emergency	-0.00270***	-0.00207***
Small provider dummy	-0.0230**	-0.0196*
Large provider dummy	0.0108	0.00700
Teaching provider dummy	0.0683***	0.0587***
MFF	0.647***	0.705***
Time indicator (2010/11)	0.000627	-0.00740
Time indicator (2011/12)	-0.0483***	-0.0576***
Time indicator (2012/13)	-0.0455***	-0.0528***
Time indicator (2013/14)	-0.0653***	-0.0760***
** p-value<.10* *** p<0.05* **** p<0.01*		

Remoteness proxies

Ten alternative remoteness proxies are tested to identify the additional effects of remoteness. These proxies are based on the two identification criteria: (1) population catchment and (2) travel time to the next nearest hospital (see Section 2.1 for further details).

The table below describes the remoteness proxies. For instance, *Index 200K 45'* defines remote sites on the basis of their population catchment being less or equal to 200K. The remoteness index reflects the proportion of the population that is less than 45' away from the second nearest hospital and reflects the degree of remoteness.

Table 10 sets out the results of this exercise. In all cases, the remoteness variable is statistically insignificant at 5% level and negative.

Table 9: Remoteness proxies

Variable name	Population catchment	Travel time	Index vs Dummy
Index 200K 30'	200K	30'	Index
Index 200K 45'	200K	45'	Index
Index 200K 60'	200K	60'	Index
Index 250K 60'	250K	60'	Index
Index 300K 60'	300K	60'	Index
Dummy 200K 30'	200K	30'	Dummy
Dummy 200K 45'	200K	45'	Dummy
Dummy 200K 60'	200K	60'	Dummy
Dummy 250K 60'	250K	60'	Dummy
Dummy 300K 60'	300K	60'	Dummy

Table 10: Sensitivity analysis: remoteness proxies

Variable name	Remoteness proxy
Index 200K 30'	-0.036*
Index 200K 45'	-0.026
Index 200K 60'	-0.027
Index 250K 60'	-0.026
Index 300K 60'	-0.019
Dummy 200K 30'	-0.021
Dummy 200K 45'	-0.019
Dummy 200K 60'	-0.013
Dummy 250K 60'	-0.022
Dummy 300K 60'	-0.012
** p-value<.10 ** p<0.05 *** p<0.01*	

Other sensitivity checks

The sensitivity of the econometric model is tested along 2 further dimensions. The organisation type variables (except teaching) are dropped from the baseline model to test if the results remain stable. A new interactive term between the small organisation type and activity is included in the baseline model to see if the impact of activity on cost varies on the basis of the size of the provider.

Table 11 sets out the results of this exercise. In both cases, the results remain stable and do not experience notable changes. The interactive term is not significant at 10%.

Table 11: Other sensitivity checks

Independent Variables	Baseline	Baseline excluding organisation type	Baseline including interaction term
Activity, case-mix adjusted (log)	0.891***	0.895***	0.902***
Interactive term	N/A	N/A	-0.0325
Remoteness dummy	-0.0266	-0.0326	-0.0173
Activity concentration	-0.0796	-0.0593	-0.0693
% patients > 75 years old	-0.00162	-0.00188	-0.00168
% female patients	-0.00580***	-0.00660***	-0.00552***
% patients BAME	0.00159**	0.00172***	0.00155**
% patients emergency	-0.00270***	-0.00281***	-0.00270***
Small provider dummy	-0.0230**	N/A	0.442
Large provider dummy	0.0108	N/A	0.0122
Teaching provider dummy	0.0683***	0.0662***	0.0665***
MFF	0.647***	0.633***	0.654***
Time indicator (2010/11)	0.000627	0.000623	0.000959
Time indicator (2011/12)	-0.0483***	-0.0489***	-0.0478***
Time indicator (2012/13)	-0.0455***	-0.0460***	-0.0451***
Time indicator (2013/14)	-0.0653***	-0.0661***	-0.0651***
** p-value<.10*	*** p<0.05*	**** p<0.01*	

A.3. Emergency cost model

The analysis presented so far uses total services (excluding mental health and community services provided by acute providers) to test the remoteness hypotheses and quantify the sub-scale cost. Table 12 sets out the econometric results using only non-elective and maternity services. The sub-scale cost determined by this model is presented in Table 13 [20]

Table 12: Econometric results - Emergency services model

Independent Variables	Emergency services model
Activity, case-mix adjusted (log)	0.814***
Remoteness proxy	-0.0279
Activity concentration	-0.227
% patients > 75 years old	-0.00354
% female patients	-0.00640**
% patients BAME	0.00195*
Small provider dummy	-0.0295
Large provider dummy	0.00720
Teaching provider dummy	0.0611**
MFF	0.571***
Time indicator (2010/11)	-0.0131
Time indicator (2011/12)	-0.00687
Time indicator (2012/13)	0.0238*
Time indicator (2013/14)	0.0554***
** p-value<.10 ** p<0.05 *** p<0.01*	

Table 13: Sub-scale cost uplift of remote sites – Emergency services model

Site (Provider)	Population	Cost Uplift (£, %[21])
Site 1	111,207	£ 2,798,571 (3.05%)
Site 2	130,892	£ 4,141,573 (5.62%)
Site 3	138,393	£ 4,151,402 (4.60%)
Site 4	169,852	£ 2,880,430 (1.89%)
Site 5	178,338	£ 1,074,458 (.83%)
Site 6	182,303	£ 3,552,233 (3.58%)
Site 7	190,677	0
Site 8	194,103	£ 1,587,142 (1.48%)

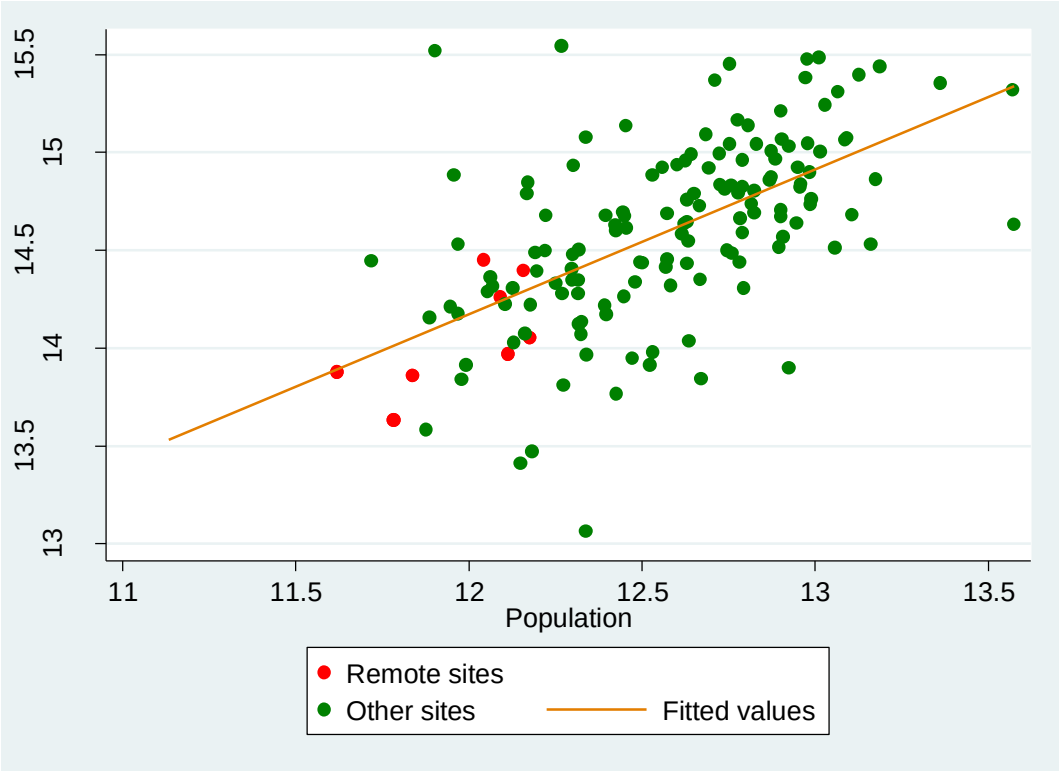
[20] The cost uplift analysis conducted for the specification whereby the benchmark is based on a population catchment of 200K and activity is calculated using the activity-based approach.
[21] The % is the emergency cost uplift of a site relative to the site's total estimated cost.

A.4. Population-based approach

This section describes the model used to project the activity of a remote hospital on the basis of its population catchment. The projected activity is then used to estimate the sub-scale cost associated with providing health services in remote areas. The motivation of this approach is discussed in Section 2.3.

The relationship between activity and population is shown in Figure 4 and as expected there is a clear positive association.

Figure 4: Bivariate relationship between hospital activity and population



The econometric model specification and estimated coefficients are presented in Table 14. This model effectively estimates the expected activity of a hospital given its population catchment after controlling for other factors. The dependent variable is the logarithm of the case-mix adjusted activity. The results of the model are shown in Table 14.

Table 14: Activity-Population model

Independent Variables	Coefficient estimates
Population (in logs)	.712***
% patients < 14 years old	-.006
% female patients	-.039***
% patients emergency	-.017***
% patients BAME	.002
% patients > 75 years old	-.014*
** p-value<.10 ** p<0.05 *** p<0.01*	

A.5. Cost-activity curve

Figure 5 shows the estimated unit cost – activity curve implied by the econometric model, which have been used to calculate the sub-scale cost of remote providers relative to the benchmark.

An alternative way of selecting the benchmark could be based on the minimum economic scale defined as the level of activity after which additional increases in scale have relatively small incremental impact on unit cost. Although the largest incremental reduction in unit cost is observed up to 2m level of activity, there are significant economies of scale at relatively large levels of activity. This is the product of the model specification (i.e. log-log), the inherit challenges with capturing complex non-linear relationships and therefore inferring the minimum economic scale.

Figure 5: Unit cost-activity curve