National Audit Office

Department of Health

Management of NHS hospital productivity

Regression analysis methodology

DECEMBER 2010

Contents

Design and Production by NAO Communications
DP Ref: 009317-003

Data analysis methodology

Data preparation

Unit of analysis

The analysis was based on 'hospital trusts' that were in existence in 2008-09.

Time period

2006-07, 2007-08 and 2008-09

Software

SPSS v18 and Excel 2002

Explanatory variables

It should be noted that in the case of some datasets there are known to be potential data quality issues, and it was for this reason that the variables 'Number of A&E attendances' and 'I.T. capital and revenue spend' were excluded (see Exclusions).

The size of a hospital was captured through the small/medium/large hospital size variables (based on income as at 2008-09) and so all other variables expressed in absolute terms (no. of admissions, floor space, etc.) were expressed as proportions and averages (percentage of admissions which are day cases, percentage of occupied floor space, etc.), recoding where necessary.

All financial figures were adjusted for **inflation** to be at '08-09 prices', using the Hospital and Community Health Services Pay & Prices Index.

This left the following 46 candidate explanatory variables:

Quality of care

1. CQC Quality Fair
2. CQC Quality Good
3. CQC Quality Excellent
4. Patient Safety Score (Dr Foster) [2008-09 only]

Financial management

5 CQC Financial Fair

6 CQC Financial Good

7 CQC Financial Excellent

8 Operating surplus/deficit (% of all income)

9 Income from activities/Bed

Status/Type

10 Foundation Trust Status

11 Acute – Large

12 Acute – Medium

13 Acute Multi-service

14 Acute Teaching Trust

15 Acute Specialist Trust

Estates

16 PFI Contracted [2008-09 only, as based on an NAO survey administered in that year]

17 Cost of any PFI deal/Total cost of estates [2008-09 only, as based on an NAO survey administered in that year]

18 Premises Expenditure (percentage of all income)

19 Occupied Floor Area (percentage)

20 Functional suitability of estate (percentage of occupied floor area below an acceptable standard)

21 Total Bed Occupancy (percentage)

Patient type

22 Ratio of Emergency to Non-Emergency Admissions

23 Emergency Inpatient Admissions (proportion of total admissions)

24 Mean waiting times in days (Inpatients)

25 Mean age of inpatients

26 Mean Length of Stay for Inpatients

27 Outpatient: Inpatient (percentage)

28 Day cases: Non-emergency admissions

Expenditure type

29 Consultancy Service Costs (percentage of all income)

30 Supplies and Services Costs (percentage of all income)

31 Combined Transport (percentage of all income)

32 Combined Other Costs (percentage of all income)

33 Clinical Negligence Costs (percentage of all income)

34 Total Staff Costs (percentage of all income) [not included in 2007-08]

Staffing

35 Doctor turnover (Staff Stability Index)

36 Consultant turnover (Staff Stability Index)

37 Number of Consultants (percentage of total staff)

38 Number of Registrars (percentage of total staff)

39 Number of Junior Doctors (percentage of total staff)

40 Number of hospital Practitioner/ Clinical Assistant (percentage of total staff)

41 Number of other clinical non-medical staff (percentage of total staff)

42 Number of support to clinical staff (percentage of total staff)

43 Number of NHS infrastructure support (percentage of total staff)

44 Total staff/Bed

45 Total Staff Costs/Bed

46 Sickness Absence (percentage) [2008-09 only]

Exclusions from explanatory variables

- Individual components of RCI and also the BCBV indicators, as these are all **measures of the efficiency** of different components of a hospital, and thus not really explanatory variables.

- For technical reasons, one of the **categorical variables** hospital type (acute-small) and CQC Financial and Quality (Weak) was excluded to prevent perfect multi-collinearity between variables which would have invalidated the model.
- The '**A&E attendances**' variable, following concerns raised about the data quality from the team involved in collating the data.
- **Admin & Clerical Agency Costs as** percentage of Total Expenditure and **Medical Staff Costs** as percentage of Total Expenditure, as these were not available for Foundation Trusts.
- **I.T. capital and revenue spend** as there were obvious data quality issues.

Dependent variable

Reference Cost Index (**RCI**) score.

This is calculated by dividing the amount which a hospital has spent by the amount which the average NHS hospital would have spent to treat the same case-mix, adjusted for 'Market Forces Factor' to account for the different costs of purchasing inputs in different parts of the country and rebased so that the average hospital has a score of 100. Scores under 100 suggest better than average efficiency, scores over 100 suggest inefficiency.

In 2009-10, there was a revision in the calculation of **Market Forces Factor** (MFF). Given that, apart from this revision, MFF had remained quite stable from year to year, the RCIs of previous years were adjusted to account for the 2009-10 MFF. After adjusting for the 2009-10 MFF, all hospitals were re-based so that the mean score remained 100. This method was verified with the Department of Health.

Stepwise approach – Individual years

For each of the years (2006-07, 2007-08, 2008-09), a multivariate regression analysis was performed, using the stepwise method with an F-probability of 0.1 for inclusion and an F-probability of 0.2 for exclusion. Where data was missing, SPSS was instructed to replace with the mean value.

This procedure produced the following three regression equations:

2006-07 (R^2 = 0.493):

RCI = 159.364 + 22.452 (Acute Specialist Trust) – 0.477 (Occupied Floor Area) + 1.014 (Mean Length of Stay) – 0.861 (Operating Surplus / Deficit as a percentage of Income) + 3.995 (Acute – Large) – 0.004 (Income from Activities / Bed) – 0.294 (percentage of Total Bed Occupancy) – 3.454 (Clinical Negligence Costs percentage) + 1.718 (percentage Junior Doctors)

2007-08 (R^2 = 0.485):

RCI = 146.938 + 11.611 (Acute Specialist Trust) – 0.541 (percentage of Bed Occupancy) – 7.241 (CQC Financial Excellent) – 3.352 (percentage Clinical Negligence Costs) + 4.005 (Acute Teaching Trust) – 0.567 (Operating Surplus as percentage of Income) + 0.908 (Total Staff/Bed)

2008-09 (R^2 = 0.559):

RCI = 160.818 – 3.327 (percentage of Junior Doctors) – 0.762 (percentage Support to Clinical Staff) – 0.375 (percentage Occupied Floor Area) + 1.497 (Mean length of stay) – 0.030 (Outpatients expressed as a percentage of inpatients) – 2.661 (Acute – Medium) – 0.430 (percentage Combined Other Costs) – 0.356(Operating Surplus/ Deficit as percentage of Total Income) + 2.562 (percentage of Consultants) – 0.152 (Day Cases Expressed as percentage of Non-Emergency Admissions) – 0.825 (Registrars as percentage of Total Number of Staff)

Excluding specialist trusts

The strongest individual effect from the initial stepwise analysis was from 'acute specialist trust status', which were associated with higher RCIs (lower efficiency). Specialists trusts, by their nature, will tend to take on the severest cases and, therefore, these hospitals are expected to have higher RCIs; in 06-07, the average RCI of acute specialist trusts was 117, against an average for the population as a whole of 100. Because of this skew, acute specialist trusts were removed from the analysis.

Once I had done this, I reran the regressions, using the same process as above. I now got the following results:

2006-07 (R^2 = 0.284):

RCI = 178.236 – 0.513 (percentage of Occupied Floor Area) + 1.931 (Total Staff/Bed) – 0.377 (percentage of Total Bed Occupancy) – 0.713 (Operating surplus or deficit as a percentage of income) – 0.165 (Day cases as a proportion of non-emergency admissions) + 9.024 (Acute multi-service) + 2.345 (Acute – Large)

2007-08 (R^2 = 0.366):

RCI = 139.629 + 1.794 (Acute – Large) + 8.281 (Acute Multi-service) – 0.437 (Operating Surplus/Deficit as a percentage of Income) – 0.293(percentage of Bed Occupancy) + 4.460 (Acute Teaching Trust) – 2.532 (CQC Financial Excellent) – 0.045 (Mean waiting times in days – inpatients) – 0.217 (percentage of Occupied Floor Area) + 0.073 (Emergency Admissions as a percentage of Non-Emergency)

2008-09 (R^2 = 0.361):

RCI = 128.117 – 0.502 (percentage of Support to Clinical Staff) – 3.425 (percentage of Junior Doctors) – 0.249 (percentage of Occupied Floor Area) – 2.737 (Acute – Medium) + 10.301 (Emergency Admissions:Non-Emergency) – 0.441 (Operating Surplus/Deficit as percentage of Total Income) – 0.412 (percentage of Combined Other Costs) + 0.310 (Number of other clinical non-medical staff per total number of staff)

Issues with the Stepwise method

These equations are obviously quite different. However, as there has been no fundamental structural change in the NHS over the past few years, it seems unlikely that the relationship between explanatory variables and RCI has in fact changed from year-to-year.

It is much more likely that some of the variation between equations is an artefact of the way stepwise models are constructed: it picks out the explanatory variable which explains the most variation in the dependent variable and place it in the model, followed by putting in additional variables on the basis of how much more of the dependent variable they explain. Variables with very weak additional effects are excluded. This means that when two variables are quite strongly correlated, it is likely that only one of them will appear in a stepwise model, even if both are quite strongly correlated with the dependent variable.

Whilst there are variations in the equations above, these three models do have certain common features. The following variables are present in at least two models:

- Percentage of Occupied Floor Area (all three models)
- Operating Surplus/Deficit as percentage of Income (all three models)
- Percentage of Bed Occupancy (two models)
- Acute – Large (two models)
- Proportion of emergency to non-emergency admissions (two models)

We also have the following variables which appear in just one model:

- Acute Teaching
- Acute Multi-Service
- CQC Financial Score Excellent
- Acute Medium
- Total Staff/Bed
- Mean waiting times
- Day cases as a proportion of non-emergency admissions
- Percentage of Support to Clinical Staff
- Percentage of Junior Doctors
- Percentage of Combined Other Costs
- Percentage of Clinical Non-Medical Staff

Whilst a variable may not appear multiple times, or indeed at all, this doesn't necessarily mean it doesn't have a strong association with outcome, due to the issue with stepwise models noted above. However, the fact that a variable does appear multiple times certainly suggests it does have a strong association with outcome.

Hierarchical approach – Individual years

An alternative approach to the stepwise method is to input variables into the model 'hierarchically', on the basis of theory about what might be key drivers of productivity. This avoids the potential problem with stepwise of a driver being discarded in favour of a proxy indicator which is strongly correlated with it. This was the approach attempted next; care was taken to ensure that variables which came out strongly in stepwise analysis were strongly considered for inclusion.

This method was carried out using the same exclusions as outlined above.

Explanatory variables

In this case, following variables were hypothesised as being drivers of productivity:

Variable	Rationale
Financial Management	As shown by CQC score, since sound financial management should involve an understanding of efficiency
Percentage of Occupied Space	This variable came out strongly in the stepwise analysis above and, theoretically, unoccupied space is intrinsically inefficient
Hospital Size and Type	Larger hospitals might have greater bureaucracy or perhaps economies of scale
Percentage of Bed Occupancy	Very low bed occupancy suggests there could be consolidation of wards, which could decrease costs. The 'acute – large' variable also came out strongly in the stepwise analysis above
Emergency admissions as a proportion of total	A significant factor in the 08-09 stepwise analysis, emergency admissions are, by their nature, less predictable, which means hospitals might be forced to have excess capacity, to accommodate possible peaks.
Total Staff/Bed	Once a certain level is passed, I would hypothesise that there will be diminishing productivity returns for each extra member of staff per bed
Doctor Turnover	High levels of turnover will mean lots of doctor time is spent learning the procedures of their new hospital, etc., which we would expect to cut productivity

Inputting the variables

These variables were inputted into the model hierarchically, using the 'Enter' method, which forcibly includes all variables, in the following order:

- Financial Management
- Percentage of Occupied Space
- Percentage of Emergency Admissions
- Total Staff/Bed
- Doctor Turnover
- Percentage of Bed Occupancy
- Hospital size/type

Hospital size/type was included last, on the basis that this variable is just a proxy for things which may or may not already be included in our dataset, e.g. it could be that large hospitals are less productive because they tend to make worse use of space. This approach meant that effects of these other variables, which might provide a stronger insight, would not be drowned out by putting hospital size/type in first.

Similarly, I have put percentage bed occupancy near the end because I suspect that low bed occupancy might be caused in part by having a large proportion of emergency admissions, and thus large peaks/troughs in admissions, and I want to separate any effect on productivity which is down to this and any which is down to low bed occupancy for other reasons.

This procedure produced the following three new regression equations.

2006-07 (R^2 = 0.297):

RCI = 158.065 – 1.982 (CQC Financial Fair) – 4.660 (CQC Financial Good) – 4.828 (CQC Financial Excellent) – 0.453 (Percentage of Occupied Floor Area) + 0.222 (Percentage of Emergency Admissions) + 3.360 (Total Staff/Bed) + 0.380 (Doctor Turnover, Staff Stability) – 0.473 (Percentage of Total Bed Occupancy) + 1.068 (Acute – Large) – 0.029 (Acute – Medium) + 9.886 (Acute Multi-Service) + 2.531 (Acute Teaching Trust)

2007-08 (R^2 = 0.316):

RCI = 131.630 – 5.145 (CQC Financial Fair) – 5.150 (CQC Financial Good) – 7.505 (CQC Financial Excellent) – 0.164 (Percentage of Occupied Floor Area) + 0.224 (Percentage of Emergency Inpatient Admissions) + 0.726 (Total Staff/Bed) + 3.846 (Doctor Turnover, Staff Stability) – 0.338 (Percentage of Bed Occupancy) + 1.341 (Acute – Large) – 1.081 (Acute – Medium) + 5.513 (Acute Multi-service) + 5.161 (Acute Teaching Trust)

2008-09 (R^2 = 0.332):

RCI = 121.879 – 7.485 (CQC Financial Excellent) – 5.801 (CQC FinancialGood) – 4.958 (CQC Financial Fair) – 0.027 (Percentage of Occupied Floor Space) + 0.109 (Percentage of Emergency Admissions) + 2.388 (Total Staff/Bed) – 6.871 (Doctor Turnover) – 0.256 (Bed Occupancy) + 1.777 (Acute Large) – 0.859 (Acute Medium) – 1.233 (Acute Multiservice) + 4.612 (Acute Teaching Trust)

The next step was to, for each model, exclude those variables which were not significant at the 5% level. This left the variables listed in the table below:

	2006-07	2007-08	2008-09
CQC Financial Fair	–	✔	✔
CQC Financial Good	✔	✔	✔
CQC Financial Excellent	✔	✔	✔
Occupied Space (%)	✔	–	–
Emergency Admissions (%)	✔	✔	–
Total Staff/Bed	✔	–	✔
Total bed occupancy (%)	✔	✔	✔
Acute Multi-service	✔	–	–
Acute Teaching Trust	–	✔	✔
Acute Large	–	–	–
Acute Medium	–	–	–
Doctor Turnover	–	–	–

This produced the following results:

2006-07 (R^2 =0.274):

RCI = 159.583 – 3.608 (CQC Financial Good) – 3.804 (CQC Financial Excellent) – 0.477 (Percentage of Occupied Floor Area) + 0.229 (Percentage of Emergency Admissions) + 3.870 (Total Staff/Bed) – 0.490 (Percentage of Bed Occupancy) + 9.838 (Acute Multi-Service)

2007-08 (R^2 =0.257):

RCI = 129.598 – 6.282 (CQC Financial Fair) – 6.124 (CQC Financial Good) – 9.098 (CQC Financial Excellent) + 0.194 (Percentage of Emergency Admissions) – 0.391 (Percentage of Bed Occupancy) + 7.325 (Acute Teaching Trust)

2008-09 (R^2 = 0.289):

RCI = 117.173 – 7.535 (CQC Financial Excellent) – 5.697 (CQC Financial Good) – 4.480 (CQC Financial Fair) + 2.272 (Total Staff/Bed) – 0.245 (Percentage of Bed Occupancy) + 5.981 (Acute Teaching Trust)

Pooling data

The next step would have been to pool our three years' worth of data into one dataset, allowing us to make comparisons across time.

The key issue with this method is that, for a hospital, RCIs are not directly comparable across years. This is a result of RCIs being a measure of how efficient a hospital has done compared with the NHS average for their case-mix for that year so, for example, a hospital's RCI might constant from year to year even though its efficiency had improved (if average NHS productivity has improved as well).

To mitigate this, we proposed an alternative measure of productivity, broadly based on the RCI but comparable from year to year. Details are available in Appendix A. However, a combination of lack of computational power and limitations of time for fieldwork meant that this method was not pursued.

Instead, I wanted to perform a multivariate regression analysis on the combined datasets (although this wouldn't be able to investigate trends in individual hospitals from year-to-year), with the resulting equation showing which factors appear to influence a hospital in having below- or above-average performance.

However, this proved impossible due to the fact that one of the requirements of a simple linear regression model (discussed in more detail below) is that all of the data points need to be independent, i.e. generated by different sources. As each hospital would be providing three data points to the pooled dataset, this would not be the case, and therefore any simple linear model generated using this data would be invalid.

Diagnostics of individual models

A variety of diagnostic procedures was undertaken on each model produced during the analyses. There were two main reasons for this:

1 to check whether any of the models were heavily influenced by outlying values, or those with a high influence level.

2 to check the generalisability of the models.

The literature lists a series of assumptions which have to be met if we want to be able to conclude that a model generated from a sample of data is generalisable to the whole population (where in this case the 'whole population' is all hospitals at any time point).

The approach taken to validate these models broadly follows that given in Field 'Discovering Statistics using SPSS' (p162). In particular, the following diagnostic tests were performed:

Checks for undue influence by a few data points:

- Check whether more than 5% of cases has a standardised residual greater than 1.96, whether more than 1% of cases has a standardised residual greater than 2.58, and whether there are any cases with a standardised residual greater than 3.29. This will indicate whether there are any extreme outliers skewing the model
- Examine the Cook's distance of all data points. Field suggests careful examination of any data points with a Cook's distance greater than 1, as this indicates that that individual data point has a very high level of influence on the model.

Checks that the assumptions for generalisability hold:

- (No autocorrelation of errors) Examine the results of the Durbin-Watson test. A score of less than 1 or greater than 3 is taken to indicate problematic autocorrelation of errors.
- (Normally distributed residuals) Examine the residuals histogram to check whether the residuals roughly follow a normal distribution.
- (Homoscedasticity and linearity) Check that the plot of standardised predicted scores against standardised residuals (ZPRED and ZPRESID) is randomly scattered – a discernible curve in the scatterplot indicates that the relationship might be non-linear, whilst a 'funnel' shape indicates possible heteroscedasticity.
- (Multicollinearity) Check whether any variable has a variance inflation factor (VIF) of >10. This suggests that it has a strong linear relationship with the other variables

There are some other assumptions for generalisability specified in Field. These are:

- All predictor variables must be quantitative or categorical, and the outcome variable must be quantitative, continuous and unbounded.
- The predictors should have some variation in value.
- Predictor variables are uncorrelated with external variables (i.e. variables not included in the model).
- Values of the outcome variable are independent (i.e. come from separate organisations).

Due to the way in which the data have been prepared, I can confirm that the first two conditions hold. Whilst the third assumption does not hold, we have minimised any issue of making findings about the significance of one variable in the model when, in fact, the real driver is an external, correlated variable through the use of the hierarchical model.

The results of these validation diagnostics are laid out in the table:

	2006-07 Equation	2007-08 Equation	2008-09 Equation	2008-09 Equation with high Cook's distance point removed
No extreme outliers (whether more than 5% of cases has a standardised residual greater than 1.96, whether more than 1% of cases has a standardised residual greater than 2.58, and whether there are any cases with a standardised residual greater than 3.29)	**No** – there are 7 data points (4.8%) with a standardised residual greater than 2, but 3 (2%) with a standardised residual greater than 2.58 and 2 (3.503 and 4.427) with a standardised residual of greater than 3.29	**Yes** – there are only 5 data points (3.4%) with a standardised residual higher than 2, and none with a standardised residual higher than 2.58	**No** – there were 10 data points (6.8%) with a standardised residual higher than 2, and three (2%) with a standardised residual higher than 2.58, though none higher than 3.29	**No** – there were 9 data points (6%) with a standardised residual higher than 2, and two (2%) with a standardised residual higher than 2.58, though none higher than 3.29
No high-influence data points (Cook's distance >1)	**Yes** – highest Cook's distance is 0.146	**Yes** – highest Cook's distance is 0.083	**No** – highest Cook's distance is 1.064	**Yes** – highest Cook's distance is 0.107
No predictor variable with a VIF >10)	**Yes**	**Yes**	**Yes**	**Yes**
Durbin-Watson score between 1 and 3	**Yes** – 1.994	**Yes** – 2.044	**Yes** – 2.121	**Yes** – 2.131

	2006-07 Equation	2007-08 Equation	2008-09 Equation	2008-09 Equation with high Cook's distance point removed
Approximately normal distribution of residuals	Yes	Yes	Yes	Yes
Plot of standardised prediction against standardised residual shows no 'funnel'	Yes	Yes	Yes	Yes
Plot of standardised prediction against standardised residual shows no 'curve'	Yes	Yes	Yes	Yes

Two items of concern arose from these diagnostic tests. One was the presence of a few extreme outliers, as noted in the first row of the above table. I reran our multivariate analysis excluding the outliers, and it made minimal difference to our regression equations, so I am satisfied this is not an issue. However, the second issue was an item (organisation code RCB) with a Cook's distance of 1.064. When it was excluded from the analysis, it made quite a big difference to the regression equation. What I have read suggests that there is some debate among statisticians as to the most appropriate treatment for high-influence data points. I have chosen to use the model which excludes it, on the basis that a better fit for this one data point comes at the expense of a much poorer fit for the other 146 data points.

This being the case, our final three models are:

2006-07 (R^2 =0.274):

RCI = 159.583 – 3.608 (CQC Financial Good) – 3.804 (CQC Financial Excellent) – 0.477 (Percentage of Occupied Floor Area) + 0.229 (Percentage of Emergency Admissions) + 3.870(Total Staff/Bed) – 0.490(Percentage of Bed Occupancy) + 9.838(Acute Multi-Service)

2007-08 (R^2 =0.257):

RCI = 129.598 – 6.282 (CQC Financial Fair) – 6.124 (CQC Financial Good) – 9.098 (CQC Financial Excellent) + 0.194 (Percentage of Emergency Admissions) – 0.391 (Percentage of Bed Occupancy) + 7.325 (Acute Teaching Trust)

2008-09 (R^2 =0.322):

RCI = 113.137 – 8.682 (CQC Financial Excellent) – 6.908 (CQC Financial Good) – 5.892 (CQC Financial Fair) + 3.808 (Total Staff/Bed) – 0.265 (Percentage of Bed Occupancy) + 4.449 (Acute Teaching Trust)

Using regression equations to estimate potential cost savings

Based on the regression equations, the next step was to estimate the possible NHS cost savings which could result from improvements in the variables our model suggested were associated with improved efficiency. Two different approaches were used.

The first was to attempt to estimate the amount which could be saved if, for a given variable, the bottom 75% moved their performance to that of the 25th percentile.

For the population covered by our regression equation (i.e. all non-acute hospitals), let:

N = no. of hospitals included in our dataset

T = total spend of all hospitals included in our analysis, adjusted to '08-09 prices' using the Hospital and Community Health Services Pay & Prices Index

S = sum of RCIs of all of the N hospitals in our population (where an RCI score is not available, it is assumed to be replaced with the mean)

$\mathbf{R}_h$ = the RCI of a given hospital h

We then have the following identity concerning RCI:

$R_h = a + bx_{1,h} + cx_{2,h} + dx_{3,h} + \ldots\ldots\ldots\ldots + \varepsilon_h$

where a, b, c, . . . are the coefficients of our multivariate regression model, $\mathbf{x}_{i,h}$ is the value of predictor variable x_i in **hospital h**, and ε_h is the error term of our regression equation.

If we then sum across all hospitals in our analysis:

$\Sigma_h R_h = \Sigma_h(a + bx_{1,h} + cx_{2,h} + dx_{3,h} + \ldots\ldots\ldots\ldots + \varepsilon_h)$

$\leftrightarrow$ $\Sigma_h R_h = \Sigma_h a + \Sigma_h bx_{1,h} + \Sigma_h cx_{2,h} + \Sigma_h dx_{3,h} + \ldots\ldots\ldots\ldots + \Sigma_h \varepsilon_h$ (by the properties of summation)

$\leftrightarrow$ $\Sigma_h R_h = aN + b\Sigma_h x_{1,h} + c\Sigma_h x_{2,h} + d\Sigma_h x_{3,h} + \ldots\ldots\ldots\ldots + \Sigma_h \varepsilon_h$ (by the properties of summation)

$\leftrightarrow$ $S = aN + b\Sigma_h x_{1,h} + c\Sigma_h x_{2,h} + d\Sigma_h x_{3,h} + \ldots\ldots\ldots\ldots + \Sigma_h \varepsilon_h$ (by our definition of S, above)

Now, for some given predictor variable x_i, what would be the effect on the right-hand side (RHS) of the equation of altering things so that the bottom 75% get a score as high as the 25th percentile ? (in this case 'top' means most efficient – if an increase in a variable is associated with increased efficiency, the 'top' will be the highest 25%, whereas if an increase in that variable is associated with decreased efficiency the 'top' will be the lowest 25%).

Well, let $\overline{x} = (\Sigma_h\ x_{i,h})/N$, the mean score for variable x_i.

$\leftrightarrow$ $(\Sigma_h\ x_{i,h}) = N\overline{x}$

Then let x_{25} = the score for the 25th percentile of variable x_i.

Then we can take our equation:

$S = aN + b\Sigma_h x_{1,h} + c\Sigma_h x_{2,h} + d\Sigma_h x_{3,h} + \ldots\ldots + q\Sigma_h\ x_{i,h} \ldots\ldots + \Sigma_h \varepsilon_h$

Substitute in the identity above:

$S = aN + b\Sigma_h x_{1,h} + c\Sigma_h x_{2,h} + d\Sigma_h x_{3,h} + \ldots\ldots + q \bullet N \bullet \overline{x} \ldots\ldots + \Sigma_h \varepsilon_h$

Now what happens if the bottom 75% of hospitals improve their score such that they become as good as the hospital at the 25th percentile ? Then the mean for that variable becomes $(\overline{x}_{25} + 3x_{25})/4$, where $\overline{x}_{25}$ is the mean of the top 25% of hospitals.

Then our $q \bullet N \bullet \overline{x}$ becomes $q \bullet N \bullet (\overline{x}_{25} + 3x_{25})/4$, which means adding $q \bullet N \bullet ((\overline{x}_{25} + 3x_{25})/4 - \overline{x})$ to the RHS.

Maintaining the identity by adding this to the LHS also, we end up with this:

$S + (q \bullet N \bullet ((\overline{x}_{25} + 3x_{25})/4 - \overline{x})) = a\Sigma_h 1 + b\Sigma_h x_{1,h} + c\Sigma_h x_{2,h} + d\Sigma_h x_{3,h} + \ldots\ldots + q \bullet N \bullet \overline{x}_{25} \ldots\ldots + \Sigma_h \varepsilon_h$

Which is equivalent to saying that if we changed the scores of variable xi so that the bottom 75% of hospitals moved their performance up to that of the 25th percentile, the total RCI of all of these hospitals would be $S + q \bullet N \bullet ((\overline{x}_{25} + 3x_{25})/4 - \overline{x})$.

This means that their total RCI is $(S + q \bullet N \bullet ((\overline{x}_{25} + 3x_{25})/4 - \overline{x}))/S$ of the old score.

This means that to achieve the same outcome, they are spending only $(S + q \bullet N \bullet ((\overline{x}_{25} + 3x_{25})/4 - \overline{x}))/S$ of their old spend, therefore:

Total saving achieved by the bottom 75% of hospitals improving their scores to the score of the 25th percentile = $T \bullet (1-(S + q \bullet N \bullet ((\overline{x}_{25} + 3x_{25})/4 - \overline{x})/S))$

By similar arithmetic,

Total saving achieved by improving the scores of all hospitals in variable x_i by one percentage point = $T \bullet (1-((S - \mathrm{Abs}(q) \bullet N)/S))$

Deriving from this formula, for an 'average' hospital with RCI score S_a (average RCI score for hospitals included in analysis) and expenditure T_a (average expenditure for hospitals included in our analysis) we have N = 1 and arrive at potential saving in one average hospital = $T_a \bullet (1-((S_a - \mathrm{Abs}(q))/S_a))$

When interpreting these potential savings, it should also be borne in mind that sometimes a hospital's levers to improve even to the conservative level assumed can be limited. For example, hospitals might be able to influence emergency admissions a little via things like better outpatient management of long-term conditions, but to a certain extent the level of emergency admissions will be outside their control. It should also be borne in mind that improvements in one variable could well affect other variables, so an estimated total saving for a year arrived at simply by adding up the estimated total savings for each individual variable can only be treated as a broad estimate.

When presenting these results in the NAO report, it was decided that the best approach was to state estimated potential savings to the NHS as the average potential saving across the three years.

2006-07

Variable	National saving from bottom 75% of hospitals moving to having a mean equal to the current 25th percentile			Saving in one hospital from moving variable 1%/1 staff member		
	95% Lower Bound on Coefficient (£)	Estimated Coefficient (£)	95% Upper Bound on Coefficient (£)	95% Lower Bound on Coefficient (£)	Estimated Coefficient (£)	95% Upper Bound on Coefficient (£)
Emergency Admissions (%)	55,971,955	427,252,591	796,667,495	64,844	494,974	922,942
Occupied Floor Space (%)	133,220,685	525,175,758	916,029,834	261,536	1,032,528	1,798,333
Bed Occupancy (%)	356,736,919	760,004,741	1,163,272,562	497,135	1,059,114	1,621,093
Total Staff/Bed	463,400,556	832,572,030	1,201,743,503	4,655,780	8,364,841	12,073,902

2007-08

Variable	National saving from bottom 75% of hospitals moving to having a mean equal to the current 25th percentile			Saving in one hospital from moving variable 1%		
	95% Lower Bound on Coefficient (£)	Estimated Coefficient (£)	95% Upper Bound on Coefficient (£)	95% Lower Bound on Coefficient (£)	Estimated Coefficient (£)	95% Upper Bound on Coefficient (£)
Emergency Admissions (%)	34,314,148	350,365,509	668,222,789	43,761	446,823	852,189
Bed Occupancy (%)	415,185,671	751,562,950	1,087,940,230	497,494	900,556	1,303,618

2008-09

Variable	National saving from bottom 75% of hospitals moving to having a mean equal to the current 25th percentile			Saving in one hospital from moving variable 1%		
	95% Lower Bound on Coefficient (£)	Estimated Coefficient (£)	95% Upper Bound on Coefficient (£)	95% Lower Bound on Coefficient (£)	Estimated Coefficient (£)	95% Upper Bound on Coefficient (£)
Bed Occupancy (%)	102,098,405	386,515,390	669,473,826	973,435	562,005	148,454
Total Staff/Bed	409,981,296	696,657,196	983,333,096	4,752,652	8,075,903	11,399,154

In the main VfM report, we include the average figures for coefficients and savings. Where a variable was does not appear in the model in a particular year, the average was calculated using a value of 0 for that year:

Variable	Average estimated coefficient	'25th percentile' national saving (£)	'1%/1 staff' hospital saving (£)
Emergency Admissions (%)	0.14	259,206,033	313,932
Occupied Floor Space (%)	-0.16	175,058,586	344,176
Bed Occupancy (%)	-0.38	632,694,360	840,558
Total Staff/Bed	2.56	509,743,075	5,480,248

Appendix A

The problem with comparing RCI scores across years is that each year total cost is divided by average NHS cost for that case-mix **that year**. This allows for comparison of hospitals with other hospitals in the same cohort, but does not allow for comparison between years. This is because dividing by average NHS costs removes the effects of two different things:

- general price inflation, which we would want to remove in a comparison between years; and
- national trend of a change in NHS efficiency, which we would not want to remove in a comparison between years – e.g. as things stand, if a hospital becomes 10 per cent more efficient, but all other hospitals do as well, that hospital's RCI score will stand still, even though its performance has improved).

The situation is further complicated by the effects of rebasing to get an average of 100 each year.

To mitigate these issues, an alternative index of efficiency following two recalculations:

1 Based on data for each hospital for each year on their case-mix, the 08-09 reference costs are used to work out what the NHS average cost of this case-mix would have been in 08-09.

2 Based on data for each hospital for each year on total costs, the Hospital and Community Health Services Pay & Prices Index can be used to express costs in 08-09 pounds (using this index as a measure of general price inflation for the NHS).

Using these two adjusted datasets, a new index can be derived by dividing how much procedures cost in 08-09 pounds by the NHS average for that case-mix in 08-09 reference costs. As figures are still based on case-mix they are still comparable between hospitals, but as they are based on the same reference costs for each year, allowing for inflation, and are not rebased to 100, they are also comparable between years.