17 March 2008

The Comparative Efficiency of BT Openreach

A Report for Ofcom

NERA
Economic Consulting

MMC Marsh & McLennan Companies

Project Team

Nigel Attenborough
Gordon Hughes
Tim Miller
Mat Pearson
Sumit Sharma

NERA Economic Consulting
15 Stratford Place
London W1C 1BE
United Kingdom
Tel: +44 20 7659 8500
Fax: +44 20 7659 8501
www.nera.com

Contents

Executive Summary

Specification of the Study

NERA was commissioned by Ofcom to carry out a comparative efficiency assessment of BT Openreach. More specifically, the focus is on those BT Openreach activities which are involved in providing wholesale line rental (WLR) and unbundled local loop (LLU) services.

In order to assess the relative efficiency of Openreach, NERA used a benchmark dataset which comprises data on costs, network size, environmental and quality of service variables for approximately 70 US local exchange companies (LECs) for the years 1999 to 2006. The data and methods that we have used to assess the efficiency of Openreach are consistent with those in previous studies of the efficiency of BT's network operations undertaken by NERA[1].

One important difference between this study and our previous studies arises from the fact that the activities which Openreach performs are a subset of the activities of BT Network – and, indeed, of the activities for which the LECs report their costs. The difference in the scope of this study presents a substantial problem in that it is impossible to compile data for the LECs and the WLR and LLU activities of Openreach on a basis that allows direct comparisons to be made. This is because even the disaggregated cost information for the LECs includes a substantial number of items where the costs of operating the core network and access network are combined and cannot therefore be separately identified. In particular it is impossible either to identify separately the core and access network components of cable, duct and pole costs, or to identify fibre costs separately from copper cable costs.

One major consequence is that the costs of providing leased lines or partial private circuits cannot be stripped out of the data for the LECs. This is likely to be important because the numbers of leased lines provided by the LECs has grown rapidly in recent years, whereas the numbers of switched lines has been falling. BT has not experienced a similar shift from switched to leased lines.

To address this issue, we have examined the efficiency of two Openreach 'pseudo-companies'. The first pseudo-company (referred to as OR excl LL) closely resembles the WLR and LLU activities of Openreach, the only difference being that line cards have been excluded in order to allow a more direct comparison to be made with the LECs. The second pseudo-company (referred to as OR incl LL) adds back the identifiable costs - extracted from BT's regulatory accounts - of providing leased lines (excluding transmission equipment) in order to create an entity that more closely resembles the LECs after removing their switching and transmission equipment costs. As a cross-check on these results we have also examined the efficiency of BT Network compared with the LECs on the same basis as in our previous studies.

Results of the Study

We have considered various methods of assessing the comparative efficiency of the Openreach pseudo-companies and of BT Network. As in the past, we regard the results of the panel Stochastic Frontier Analysis (SFA) as the most reliable. The results of applying

---

1 See for example 'The Comparative efficiency of BT in 2003', NERA (2005)

this method are shown in the table below. The relative efficiency of OR incl LL, OR excl LL and BT Network are calculated with respect to the upper decile of the US LECs. The figures are the percentage by which the cost per switched line for Openreach/BT Network exceeds (for positive numbers) or are less than (for negative numbers) that of the company which falls on the upper decile for the US LEC efficiency rankings.

Summary of SFA Results

	Efficiency relative to upper decile			
Model specification	Log-linear	Log-linear	Log-linear	Linear
Independent variables	Basic	Basic excl leased lines	Basic plus QoS variables	Basic
OR incl LL	-7.2%	-8.3%	-8.3%	-1.3%
OR excl LL	6.8%	-9.2%	-8.2%	-0.9%
BT Network	-3.8%	-5.4%	-4.5%	0.8%

NOTE: BT Network SFA efficiency estimates obtained using efficiency models including costs and cost drivers which are appropriate to standard network activities rather than just the access network.

One further difference from our previous studies is that the decline in the number of switched lines provided by the LECs – for many companies this amounted to a fall of 20-25% over the period from 2000 to 2006 - has left them with substantial investments in what are now stranded assets. We found that the best statistical specification for the cost frontier includes a variable defined as the ratio of the peak number of switched lines to the current number of switched lines, which we refer to as the stranded assets ratio. With this variable included, the model displays constant returns to scale, so that we focus upon average cost per switched line. Our results show that, using a basic log-linear specification, Openreach incl LL is 7.2% **more** efficient than the upper decile, while Openreach excl LL is 6.8% **less** efficient than the upper decile. Using the specification for full network costs we find BT Network is 3.8% more efficient than the upper decile.

We have carried out various sensitivity tests on the model specification to identify what factors drive the difference between the estimated efficiencies for Openreach with and without leased lines. These tests show that the critical factor is the role of the number of leased lines per switched line as a cost driver. The cost frontier that is estimated is a standard log-linear Cobb-Douglas specification, so that it is necessary to take the logarithm of the number of leased lines per switched line. The logarithm is not defined if the number of leased lines is zero, so it is necessary to set the value of leased lines to some arbitrary number – we have used 1 which is the standard assumption in such cases. However, Openreach has nearly 30 million switched lines, so that log (leased lines per switched lines) is -17.2, whereas the minimum value for this parameter for any of the LECs is -3.8. Consequently, Openreach is well outside the span of the LEC dataset. This gives rise to a problem when making a comparative efficiency assessment.

Each of the potential solution to this problem is more or less unsatisfactory. The main alternatives are:

(a) Drop leased lines per switched line as a cost driver. The elasticity is 0.023 which is different from zero at a significance level of 10%. If the cost frontier is estimated without leased lines per switched line, Openreach excl LL is 9.2% **more** efficient than the upper decile and Openreach incl LL is 8.3% **more** efficient than the upper decile.

(b) Replace the value of log (leased lines per switched line) by a different value, say -4.0. This has no effect on Openreach incl LL but raises the estimated efficiency of Openreach excl LL to 7.2% **more** efficient than the upper decile (identical to Openreach incl LL).[2]

It follows that the way in which the technical problem caused by the zero number of leased lines for Openreach excl LL is resolved affects the magnitude of the difference between the two pseudo-companies. We have also examined a linear cost model, which does not involve taking logarithms. In that case, Openreach incl LL (1.3%) & Openreach excl LL (0.9%) are marginally more efficient than the upper decile.

Finally, we have experimented with adding quality of service variables to our efficiency models. We find that the efficiency of Openreach improves relative to the decile for the US LECs. However, the signs of the coefficients on the two QoS variables that are significant in our model, namely % of installation commitments met on time (negative) and the average number of reported faults per switched line (positive), are not consistent with prior expectations. This may indicate that these variables are picking up the impact of some exogenous cost drivers, for example the condition of the network, not otherwise captured by our models.

Conclusion

Overall, our results suggest that Openreach is somewhat more efficient than the upper decile of the US LECs. The only outlier is the Openreach excl LL pseudo-company when leased lines appear as a significant cost driver in the log-linear specification. As explained above, this result appears to be an artificial consequence of technical factors and hence should probably be discounted.[3] This conclusion is reinforced by the results obtained when BT's network costs are compared with the network costs for the LECs, which show that BT's efficiency appears to be consistently on or marginally better than the upper decile.

---

2 This value of -7.2% relative to the decile is produced using the log linear specification of the efficiency model when the number of leased lines for Openreach excl LL is assumed to be four. This compares to the value of +6.8% reported in the Summary of SFA results table (above), which is produced using the same model specification but using the assumption that the number of leased lines for Openreach excl LL is one.

3 When we include QoS variables in the efficiency models Openreach excl LL is no longer an outlier. This is because the coefficient in these models for log(leased lines/switched lines) is significantly lower than in the models which exclude QoS variables. The reduced importance of this coefficient in determining the efficient level of costs means that extreme values for this variable (such as the one for Openreach excl LL) are less of an issue. However, the fact that the quality of service variables have signs that are the opposite of a priori expectations means that this result needs to be treated with some scepticism.

Notwithstanding this, the difficulty of defining a reliable basis for comparing Openreach with the LECs suggests that the results of this study must be regarded with some caution. In future, it may be better to rely on comparisons of BT total network costs with those of the LECs, rather than attempting to make substantial but inevitably uncertain adjustments to cost information for the LECs and Openreach in an attempt to estimate and compare costs for the access network on its own. Alternatively, bearing in mind that the relative efficiency of BT's whole network may not necessarily be a guide to the relative efficiency of BT Openreach's WLR and LLU activities, it would be possible to look at how the measured relative efficiency of OR excl LL and OR incl LL had changed between now and the next point in time when such a comparison is made.

1. Introduction

NERA was commissioned by Ofcom to carry out a comparative efficiency assessment of BT Openreach. More specifically the focus is on those BT Openreach activities which are involved in providing wholesale line rental (WLR) and unbundled local loop (LLU) services.

The reason for concentrating the analysis on these activities is that Ofcom has to determine what future price controls should be applied to Openreach WLR and LLU and this efficiency study is designed to shed light on what future efficiency improvements Openreach might reasonably be expected to achieve in the provision of these services.

In carrying out this assignment, we compared BT Openreach's costs with those of the equivalent activities of 70 US local exchange carriers, using econometric and mathematical methods to take account of differences in the level of outputs and in operating environments. These companies were chosen because, unlike in the case of other European operators, detailed cost and operational data are publicly available and because the best performing US LECs are likely to provide an appropriate benchmark for efficient operation.

1.1. Report Structure

The remainder of this report is structured as follows:

§ Section 2 reviews the theory behind comparative efficiency measurement and the methodology which NERA has used in this study;

§ Section 3 describes the precise way that Openreach has been defined, the manner in which the equivalent costs of the US LECs have been extracted, the data that has been collected and how this has been processed;

§ Section 4 set out the results NERA has obtained, and

§ Section 5 summarises the main conclusions of the comparative efficiency assessment.

2. Comparative Efficiency Measurement

2.1. Introduction

The efficiency of a company can be defined as the extent to which it is able to minimise its costs for producing a given set and volume of outputs, taking into account the environment in which it operates (including demographic and geographical circumstances). A perfectly efficient company is one which has the lowest costs possible given the outputs that it produces and the environment in which it operates.

There are a variety of statistical and mathematical programming techniques that can be used to assess the comparative efficiency of different companies. In considering the most appropriate approach to take, it is important to examine the relative merits and drawbacks of the alternative techniques that could be used. This section looks at the most frequently used techniques and examines their main advantages and disadvantages when used in comparative efficiency assessments.

Statistical techniques use regression analysis to estimate a model, based on past data for different companies, that relates costs to different types of output (such as exchange lines, call minutes, leased lines and so on) and environmental factors (population density and dispersion, network size, relative number of business and residential lines, factor input prices and so on).

2.2. Functional Form of Regression Model

The first step in the estimation of a regression model is to consider the functional form of the equation relating the level of costs to the factors that determine costs. For a study of this type a log-log functional form is commonly used. We considered both the Cobb-Douglas and translog functional forms. The Cobb-Douglas functional form is much less flexible than its translog counterpart, which allows the functional form of the regression equation to be influenced by the data to a far greater extent. However, to achieve this greater flexibility, the translog specification includes many more explanatory variables in the model (made up of the squared and cross-product terms of each explanatory variable included in the model) and therefore requires a far larger number of observations in order to derive a statistically significant relationship.

As part of NERA's initial investigations, the availability of data that would enable the translog functional form to be used was investigated, and it was found that, despite the expanded dataset used in this study, the number of observations was too small to accommodate the large number of variables that would have to be included in the regression. Moreover some of the required data was not available.[4] Consequently, a Cobb-Douglas (log-log) specification was chosen. The equation below is an example of the Cobb-Douglas specification:

---

4 When estimating a translog function it is necessary to estimate simultaneously the demand functions for the inputs to the cost function (labour, capital and other inputs), and the total cost function itself. Investigation into the availability of data for the estimation of the input demand functions indicated that, particularly when looking at other costs, it would not be possible to obtain LEC specific data of sufficient reliability to allow the estimation of these functions.

$$\log(C) = a + b_1 \log(L) + b_2 \log(P) + ...$$

In this regression, *C* is a measure of cost, and *L* and *P* are explanatory variables, such as the number of switched lines, and population density.

Some advantages of using a cost function with this functional form are that:

- Firstly, it allows for non-constant returns to scale;
- Secondly, it limits the impact of heteroscedasticity; and
- Thirdly, a log-log transformation of the Cobb-Douglas functional form results in an equation that is linear in explanatory variables (which is a requirement of regression analysis) and which can be easily interpreted (the coefficient on an explanatory variable indicates the percentage change in total cost that would result from a 1% increase in the explanatory variable, all other variables remaining constant).

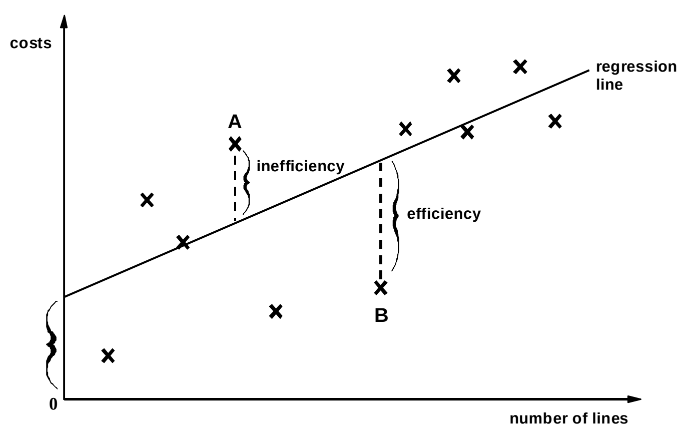
The estimation of the relationship between costs, outputs and environmental factors, based on the use of the Cobb-Douglas functional form, is described in the subsequent paragraphs.

2.3. Ordinary Least Squares (OLS) Regression Analysis

Ordinary Least Squares analysis is one of a variety of techniques which fall under the heading of regression analysis. It involves the identification of the statistical relationship between different variables. In the case of this study, therefore, the objective is to derive the relationship between total cost and a variety of exogenous cost drivers such as the number of lines, the number of call minutes, the dispersion of population etc.

OLS regression analysis can be best understood through the use of a simple example. If the cost of building and operating a network (*C*) depended only on the number of exchange lines provided (*L*), then each operator's level of costs and number of customer lines could be plotted on a graph, as in Figure 2.1 below, where each point represents a different operator.

Figure 2.1
Ordinary Least Squares Regression Analysis



Ordinary least squares regression analysis fits a line of "best fit" to these points, such that the line minimises the sum of the squared vertical distances of the observed company costs (represented by crosses) from the line, hence the technique's formal name, ordinary least squares.

The line of best fit can be written in equation form as:

$$C_i = a + bL_i + u_i$$

where ***i*** represents the observations for the different operators, ***a*** is the fixed cost involved in providing a network regardless of the number of exchange lines, ***b*** is the cost of providing each additional line (the marginal cost), and ***u*** is the regression residual (the difference between actual costs and those "predicted" by the line of best fit).

If there are many companies in the sample, it is very unlikely that they would all lie on the best-fit line, but rather some would be above and others below. The best-fit line therefore represents the costs that a company of 'average' efficiency would be expected to incur at each volume of exchange lines. Those companies with an observation above the line (for example, company A in Figure 2.1) have costs above those of a company of average efficiency with the same number of lines. Such companies are, in this relative sense, inefficient. Conversely, those companies that lie below the regression line (for example, company B) may be viewed as being relatively efficient (above average efficiency).

In practice, rather than plotting all the companies' observations on a graph, a computer program is used to estimate the regression coefficients (***a*** and ***b***) using the data on all the companies in the sample. Individual companies are then judged by substituting their actual output numbers into the equation to give a predicted level of costs, ***Z***, as if the company were of average efficiency. If the company's actual cost level were larger than ***Z***, then it would lie above the regression line and, therefore would be deemed inefficient (compared to "average performance"). Likewise, if its predicted costs were to exceed its actual costs, it would be judged to be efficient compared to "average performance".

The difference between a company's actual costs and its predicted costs is termed the residual. A positive residual therefore indicates inefficiency relative to the sample "average", and a negative residual indicates efficiency relative to the sample "average".

Most cost functions are likely to have more than one cost driver. So, for example, the cost function for a telecommunications operator will in reality have additional cost drivers apart from the number of exchange lines. OLS regression analysis deals with this through the use of multivariate regressions, which take the general form:

$$C_i = a + b_1L_i + b_2P_i + b_3Q_i + ... + u_i$$

As before, ***a*** represents the level of fixed costs, b_1 measures the marginal cost of explanatory factor ***L***, and ***u*** is the regression residual. However, in addition, b_2 and b_3 now measure the marginal cost of the new explanatory factors ***P*** and ***Q*** respectively (assuming in each case that the other two explanatory factors are held constant).

2.3.1. Multi-year least squares regression analysis

The analysis described above uses data for a single year to assess how efficient one firm is compared to others. However, depending upon the number of firms for which data is available, such analysis has limitations with regards to accuracy and robustness. If, for example, a number of firms have low costs for spurious reasons (such as misreporting of accounting data in a particular year) this could skew the model significantly, making other firms look less efficient than they actually are. Also, the number of observations is limited to the number of companies for whom the required data are available.

Where a number of years of data are available, it is possible to create a data panel (or "pool"), which includes data for different companies over a number of years. This helps overcome problems associated with a limited number of observations, and reduces or eliminates the impact of peculiarities in the data, as these tend to "average out". The use of a panel dataset should therefore lead to a more robust and stable model.

However, including more than one year's worth of data from any firm can lead to problems due to the existence of heterogeneity both within observations across time and between the different observations in the panel. This can lead to difficulties in obtaining efficient and unbiased estimates of the regression coefficients. In addition, panel data can also lead to problems of autocorrelation, if the within-observation heterogeneity is low (if the figures for each year for an observation do not differ by a large amount).

Ordinary Least Squares analysis is neither able to control for the heterogeneity both within and between observations, nor for the autocorrelation problems that can arise with panel data, and hence it is not an appropriate technique to use with this type of data. In its place a two-step Generalised Least Squares (GLS) approach can be used, which takes account of the repeat observations for each firm.

The model estimated using data for a number of years is similar to that used in single-year analysis, but has an additional term measuring the time trend. This variable, which effectively allows the constant term to change over time, takes account of technological progress, inflation, or other such items that cause changes in the costs of all companies over time. The regression equation in this case is:

$$C_{i,t} = a + b_1 L_{i,t} + b_2 P_{i,t} + b_3 Q_{i,t} + ... + T + u_{i,t}$$

where T is the time trend, and $L_{i,t}$ is the value of variable L for company i in time period t, and so on. Finally, $u_{i,t}$ is the regression residual which indicates the gap between actual and predicted (average) efficiency for each company in each time period.

It is possible to run panel data analysis with an "unbalanced panel"; that is a dataset that does not contain an observation for each company in every year in the panel. If, for example, the panel covers eight years, it is possible to include firms in the panel, which are missing data for some of those years (for example a firm which has data for only 5 of the 8 years), without the model being adversely affected.

2.4. Stochastic Frontier Analysis (SFA)

A significant drawback of both OLS and GLS regression analysis is that they both implicitly assume that the whole of the residual that is obtained for any company in any period of time can be attributed to relative inefficiency (or efficiency). However, it is possible, if not probable, that the residuals from such an analysis will include unexplained cost differences that are the result of data errors and other factors affecting costs that have not been picked up in the regression equation. Stochastic Frontier Analysis (SFA) builds on the methodologies outlined above and aims to address this shortcoming.

There is an extensive academic literature on efficiency measurement using SFA, and this technique is increasingly being used by utility regulators to measure efficiency. It is based on regression analysis, but has two distinctive features:

- In contrast to OLS and GLS regression analysis, SFA models incorporate the possibility that some of the model residual may result from errors in measurement of costs or the omission of explanatory variables, as opposed to the existence of genuine inefficiencies. This decomposition of residuals between 'error' and 'genuine inefficiency', which is based on assumptions made about the distributions of the 'error' and 'genuine inefficiency' terms, is intended to provide a more accurate reflection of the true level of inefficiency.
- Secondly, the regression for SFA looks not at the average firm, but at the theoretically most efficient one.

In the case of data for just one year SFA estimates the equation:

$$C_i = a + b_1 L_i + ... + v_i + u_i$$

where '…' indicates the other variables included in the model.

The residual in a stochastic frontier model is assumed to have two components: the u_i component, which represents the genuine inefficiency; and the v_i component, which represents the genuine error. In econometrics literature, u_i is often referred to as the inefficiency term and v_i is often referred to as the random error.

In order to be able to decompose the residual into inefficiency and random error it is necessary to make assumptions about the distributions of its two components. For single year SFA models, the inefficiency term is assumed to follow a non-negative distribution (such as the half-normal or truncated normal distributions), whilst the genuine error term is assumed to follow a symmetric distribution. By making these assumptions the technique is able to decompose the residual by fitting the assumed non-negative distribution to the residuals to identify the proportion of the residuals that can be explained by this distribution.

Having to make such assumptions is a key disadvantage of single year SFA, as the appropriateness of these assumptions cannot accurately be measured.

SFA is described in greater detail in Appendix A below.

2.4.1. Multi-year stochastic frontier analysis

SFA can also be applied to panel data. This involves estimating a regression equation of the following form:

$$C_{i,t} = a + b_1 L_{i,t} + b_2 P_{i,t} + b_3 Q_{i,t} + ... + T + v_{i,t} + u_{i,t}$$

where T is a time trend variable that identifies the change over time in the regression constant, i represents an individual company observation and t represents the time period. With this specification, residuals can be different for each firm and for each year. Once again, in a multi-year setting, SFA decomposes the residual between inefficiency and error by making assumptions about the statistical distributions of these two components of the residual.

The advantages of using panel data over simple cross-sectional data (single year data) is that, with cross-sectional data in SFA analysis, strong assumptions are required about the statistical distribution of the inefficiency component of the regression residuals and, in many practical cases when cross-sectional data are used, insufficient data are available to support these assumptions. There is often little evidence to suggest which statistical distribution is appropriate in constructing a model, and in many cases, more than one distribution may be deemed to 'fit' the data. The use of panel data, in contrast, allows for these distributional assumptions to be relaxed. By observing each firm more than once, inefficiency can be estimated more precisely as firm data is embedded in a larger sample of observations. Specifically, with panel data, it is possible to construct estimates of the efficiency level of each firm that are consistent as the number of time-series observations per firm (t) increases.

In early SFA panel data studies, however, the benefits described above came at the expense of another strong assumption, namely that relative firm efficiency does not vary over time (that is, $u_{i,t} = u_i$). This may not be a realistic assumption, especially in long panels. Recent studies on this issue, however, have shown that this assumption of time-invariance can be tested, and can also be relaxed, without losing the other advantages of panel data.

Reflecting these points, NERA has applied two different possible parameterisations of the inefficiency term u to the SFA panel.

§ A time-invariant model where the inefficiency term is assumed to be constant over time within the panel; and

§ A parameterisation of time effects (time-varying decay model) where the inefficiency term is modelled as a random variable multiplied by a specific function of time:

$$u_{i,t} = \varepsilon_{i,t}.e^{\eta(t-T)}$$

where T corresponds to the last time period in each panel and η is the decay parameter to be estimated.

2.5. Assessing the Regression Model

Before drawing conclusions about relative efficiency, it is essential to verify that the regression equation is theoretically and statistically valid and that it represents the best possible model, if there is more than one possibility. The types of questions likely to be raised in this context are:

§ How well does the cost model fit the observations? Is there a large proportion of cost variation that is left unexplained by the variation in the chosen explanatory factors? Under Ordinary Least Squares analysis this is measured by the coefficient of determination R^2 (or a variation on it).

- Are the coefficients sensible? For example, does the model predict that costs will rise (rather than fall) as the number of exchange lines increases, as intuition and experience would suggest? Care must be taken here to consider the possible impact of multicollinearity, which may make some coefficients appear unintuitive when they in fact are closely related to other variables.
- Are the coefficients statistically significant? In other words, can we be confident that the relationship described is a statistically valid one?

Even if the model appears to be satisfactory, there are several potential sources of inaccuracy. These concern:

- Inaccuracies of functional form; it is unlikely that in practice the model's functional form is known exactly in advance. For example, are costs linearly related to the number of lines or is the functional form more complex? Does logarithmic transformation of explanatory factors give a better or worse fit?
- The omission of relevant variables. The accuracy of regression analysis in measuring relative efficiency depends to a large extent on the degree to which all relevant explanatory factors have been included. If, for example, hilly countryside had a significant adverse effect on costs but was ignored in the regression study, then those companies serving hilly terrain might appear to have unduly high costs simply because of their location rather than because of inefficiency; and
- A lack of independence among the cost drivers. For meaningful results, there need to be many more independent observations than the number of cost-driver coefficients being estimated (in econometric terms, there need to be many degrees of freedom).

In some cases these inaccuracies can be tested for, and wherever this is possible, NERA has completed such tests. However, it is not always possible to eliminate all such problems. Consequently the results of analysis using a mathematical programming rather than regression analysis techniques are also often considered. One such mathematical programming technique is data envelopment analysis (DEA).

2.6. Data Envelopment Analysis (DEA)

DEA can be used as an alternative to regression-based techniques. It does not involve statistical estimation, but instead makes use of mathematical programming methods, without the need to rely on a precise parametric cost function. This, in fact, is its main advantage as it allows a complex non-linear (concave or convex) relationship to exist between outputs and costs, whereas regression analysis usually restricts such relationships to be either linear, or to have fairly simple non-linear forms.

DEA operates by searching for a 'least cost peer group' of comparator companies for each individual target company. The 'peer group' is defined such that a linear combination of these companies can be shown to have at least as great an output and no more favourable operating conditions than the target company (with output and environmental variables measured in the same way as in regression analysis). If such a 'peer group' exists, and the linear combination of their costs is lower than that of the target company, this cost difference is assumed to be attributable to inefficiency on the part of the target company.

It should be noted that since it is a mathematical technique, DEA offers no statistical framework for modelling the performance of firms outside the sample, and cannot offer predictions on the effect of changes in any particular firm's costs or outputs. Furthermore, DEA is unable to assess the relevance or significance of variables, and it is therefore necessary to make assumptions based on other analysis over which variables to use.

This analytical technique is discussed in greater detail in Appendix B below (the discussion includes further detail concerning its use and its potential disadvantages).

2.7. Conclusions

Given the shortcomings of ordinary and generalised least squares regressions in predicting efficiency, NERA considers it appropriate to examine in the first instance stochastic frontier analysis (SFA), and more specifically SFA run over a number of years. In addition, data envelopment analysis has the potential to provide a useful estimate from a non-statistical point of view. The results of these different techniques can then be reviewed in the light of their relative strengths and weaknesses in order to provide a more informed view of comparative efficiency. Additionally, if the different techniques provide a common picture as to the relative efficiency of an individual firm, greater weight can be placed upon the overall efficiency result.

This common picture can be either in terms of the actual efficiency results or in the rankings of the firms under the different techniques. It is possible, if not likely, that different techniques will produce different efficiency results, as they are based on different underlying assumptions; however if there are similarities between the rankings of firms under the various techniques, this indicates that one can be confident in the relative position of the firm within the sample. It then remains to decide which efficiency result is most appropriate given the purpose for which it is to be used.

NERA has undertaken a number of efficiency studies for Ofcom. The most recent of these is '*The Comparative efficiency of BT in 2003*'published in 2005. The SFA method that we use in the current study is consistent with that used in the 2005 report.

3. Data Collection and Processing

A requirement of the study is to construct similar entities whose efficiency levels can be compared. Ideally the comparator companies should have the same structure as BT Openreach ("Openreach") and provide the same services. At the same time, given that Ofcom wishes to focus on WLR and LLU services, it would be desirable if the costs and outputs of these activities could be separately identified.

Openreach only provides wholesale access network services. In contrast, the US local exchange carriers ("LECs") provide both retail and network services and within the latter category offer core and access services. It is therefore necessary to attempt to adjust and process the LEC data to derive entities that are as close as possible to Openreach. To the extent that this is not possible to replicate Openreach WLR and LLU, it is necessary to adjust the latter's costs and output to produce an entity that is similar to the closest US look a likes that it is possible to produce.

The process of producing similar entities is described below, followed by a detailed description of the necessary adjustments to, and processing of, the Openreach and LEC data.

3.1. Comparison with Previous Efficiency Studies

Our approach to data collection and processing is broadly consistent with the 2005 BT efficiency study[5]. There are some cases where we take a slightly different approach to data issues and these are summarised below:

- In this study we are comparing the LECs to Openreach, and more specifically its WLR and LLU operations, rather than BT network. We therefore take a different approach to the construction of comparable entities from the LEC and Openreach data. – see sections 3.2 and 3.4 for more details.

- In this study, unlike in 2005, we include some quality of service variables in our efficiency analysis – see sections 3.3.4 and 3.4.3.2 for further information.

- In our 2005 study we included dummy variables in our efficiency models to take into account a structural break in the US LEC data after 1998. For this study because we have data available for an extra 3 years (2004 to 2006) we chose not to include data from 1996 to 1998 in our main efficiency equations – see section 3.4.5.1 for further explanation.

3.2. Construction of Comparable Entities

The LEC cost and operating data in the ARMIS database covers all retail and network activities. In previous comparative efficiency studies we have developed a methodology for separating LEC wholesale and retail costs, including in each case an allocation of overheads, in order to compare the efficiency of the LECs' network activities with those of BT. Now, however, an additional step is required, namely the separate identification of the access network costs of the LECs.

---

5 'The Comparative efficiency of BT in 2003', NERA (2005).

We therefore started with the total LEC cost base and removed:

- retail costs and associated overheads; and then
- the network costs and overheads associated with switching equipment, transmission equipment, radio transmission and payphones.

Ideally we would also have wanted to remove the costs of cable, duct and poles in the core network. However, cable, duct and pole costs can each only be identified at the aggregate level and it is not possible to separate them into their access and core components. In addition, fibre and copper cable costs are inseparably aggregated, as are the costs of line cards and other switch equipment.

It is not therefore possible to convert the LECs into precise Openreach look a likes. Consequently it was necessary to make adjustments to Openreach. In doing so, we constructed two Openreach pseudo-companies. The first, referred to as Openreach excluding leased lines (OR excl LL), consists of Openreach WLR and LLU minus line cards. The exclusion of the latter reflects the fact that we also unavoidably exclude line card costs from the LECs when removing switching equipment from their cost base. The second pseudo-company, referred to as Openreach including leased lines (OR incl LL), comprises Openreach WLR and LLU minus line cards plus leased lines (excluding transmission equipment). The reason for adding leased lines (excluding transmission equipment) is that the LEC costs that we obtained after removing retail costs and the costs of switching, transmission etc include the cable, duct and pole costs associated with leased lines.

The cost bases of OR excl LL and the LECs (after removing retail costs and switching and transmission equipment costs etc) are not the same but this is dealt with via the cost drivers used in the cost model. The cost model for the LECs is estimated using exchange lines, leased and special access lines and environmental variables (e.g. cable route length, proportion of fibre etc) that relate to both the core and access networks. OR excl LL's efficient level of costs is then predicted by the cost model, taking account of the fact that it provides no leased lines and using environmental variables that only relate to the WLR and LLU component of the access network. OR excl LL's actual costs are then compared with its efficient costs.

In contrast, the cost bases of OR incl LL and the LECs are similar (after removing retail costs and switching and transmission equipment costs etc). In this case, OR incl LL's efficient level of costs is predicted by the cost model, taking account of the fact that it provides leased lines and partial private circuits and using environmental variables that relate to the full access network. Its actual costs are then compared with its efficient costs.

These two approaches have their respective advantages and disadvantages. The use of OR excl LL involves comparing a company with no leased lines, which is therefore an outlier, with the extrapolation of a cost function derived from companies which do provide leased lines. It does, however, attempt to assess the comparative efficiency of an entity that is very similar to Openreach WLR and LLU. The use of OR incl LL avoids the outlier problem but assesses the efficiency of an entity that is more extensive than Openreach WLR and LLU. The measured comparative efficiency is that of a combination of Openreach WLR and LLU and BT leased lines (minus transmission equipment). Thus, if the level of efficiency varies between these two activities, the measured comparative efficiency will not be that of Openreach WLR and LLU.

Reflecting these considerations, we have used both approaches when carrying out the comparative efficiency assessment of Openreach WLR and LLU.

3.3. Data Supplied by Openreach

Openreach supplied data on operating costs, depreciation, capital employed, the cost of capital, and volume of outputs (PSTN and ISDN exchange lines and unbundled local loops) for their WLL and LLU services.

3.3.1. Costs

3.3.1.1. OR excl LL

Openreach's cost data was presented on a current cost accounting (CCA) basis. Supporting this was data indicating how Openreach's costs were derived as a subset of BT's total network costs. Schedules were provided showing the different cost categories in BT's network accounts and which of these categories related to Openreach WLR and LLU. Inspection of these schedules suggested that the categories defined as being part of Openreach WLR and LLU were appropriate and that nothing obvious was missing.

Schedules were also provided showing how depreciation, GRC and NRC for Openreach WLR and LLU were derived. These involved starting with the total asset base and applying apportionment factors to derive the share appropriate to Openreach WLR and LLU. Although it was not possible for us to verify the apportionment factors, they appeared plausible. For example, taking the largest items, we found that 70% of BT's total duct (core and access) was attributed to Openreach WLR and LLU, as was 98% of access network copper and 100% of drop wires and related equipment.

In order to derive costs for OR excl LL, it was necessary to remove the costs of line cards. This was straightforward as the assets, depreciation and operating expenses for line cards were all separately identified and could therefore be excluded.

3.3.1.2. OR incl LL

In order to derive the costs of OR incl LL it was also necessary to obtain information on the costs of BT leased lines and partial private circuits (PPCs). Openreach did not provide this data but it was possible to derive it from BT's published current cost financial statements.[6] In particular, it was possible to identify the operating expenses, CCA depreciation and NRC for terminating segments (TISBO and AISBO) of leased lines and PPC

s and for wholesale trunk segments.

It was then necessary to remove the identifiable costs of transmission equipment within these categories, because such costs had been removed in the case of the LECs. This was possible for NRC and HCA depreciation. However it was not possible to identify the relevant costs in

---

6 BT, *Current Cost Financial Statements for 2007 including Openreach Undertakings*

the case of supplementary CCA depreciation and operating expenses. Not excluding the operating costs associated with leased line transmission equipment will lead to an overstatement of the costs of OR incl LL. On the other hand, not excluding the CCA supplementary depreciation associated with such equipment will lead to an understatement of costs because supplementary depreciation is negative (and hence too much of a negative item is included). It is not clear which of these too effects is more important. However, whatever the answer, it is unlikely that any residual impact would be large in relation to the overall OR incl LL cost base.

The latter is obtained by adding leased line and PPC costs (excluding transmission equipment) to the costs of OR excl LL. In both cases we assumed a cost of capital of 11.4%, which we understand to be the rate that Ofcom applies to Openreach for regulatory purposes.

3.3.1.3. Differences between US and UK Accounting Principles

Under FCC guidelines, LECs submit their accounting figures in accordance with US GAAP (generally accepted accounting principles). In contrast BT's figures are prepared IFRS principles. There are differences between the two sets of principles in the treatment of pension costs (including incremental pensions associated with redundancies), interest costs associated with major construction projects, the sale and leaseback of properties and the impairment of assets. These give rise to differences in operating costs and capital employed.

In its Annual Report Form 20F for 2007, BT provides figures for the impact of applying US GAAP rather than IFRS accounting principles for each of the items where the accounting treatment differed. Using this information, we were able to calculate what impact it would have made on Openreach's costs had BT reported its accounts using US GAAP.

The method used to calculate this adjustment was as follows:

§ To calculate the pension cost adjustment relating to Openreach's relevant network activities, the total pension cost adjustment for BT Group was multiplied by Openreach's relevant activities' share of total BT Group employment.

§ Capitalised interest adjustments were apportioned to the relevant Openreach activities using the proportion of BT Group's current cost capital employed accounted for by Openreach's relevant network activities.

§ To restate the sale and leaseback of properties under the requirements of US GAAP it is necessary to reverse the gain on the disposal of the fixed assets and the rental charge BT paid under the leaseback, and to replace the rental charge with a finance lease interest charge and a depreciation charge. These adjustments were allocated to the relevant Openreach activities using the proportion of BT Group's current cost capital employed accounted for by Openreach's relevant network activities.

§ To calculate the share of the impairment of assets adjustment that should be allocated to the Openreach's relevant activities, the share of the relevant Openreach activities in BT Group's current cost capital employed was used.

In total, the conversion of Openreach's relevant network costs to US GAAP results in a less than 0.4% increase in costs. Given the small size of this adjustment and the margin of error associated with its calculation we chose not to apply it to Openreach's costs when carrying out the comparative efficiency assessment.

3.3.1.4. Currency conversion

To achieve comparability between the data for BT and the US LECs it is necessary to express the data in a common currency either by converting the US LECs data into pounds sterling or the BT data into US dollars. As this conversion involves the use of a single exchange rate in each year for all firms in one country, the efficiency comparison is not affected by whether the US LEC data is converted into pounds sterling or BT's data is converted into US dollars. However, as the conversion of BT data into US dollars involves converting the data of only one firm whilst converting the data of the US LECs into pounds sterling involves converting the data for each LEC in every year covered by this study (from 1996 to 2003), the conversion of BT rather than the US LECs reduces processing time and minimises the risk of any inadvertent data processing errors. Therefore the data for BT was converted from pounds sterling into US dollars.

In general, it would not be appropriate to use actual market exchange rates in this conversion, since actual market exchange rates can be subject to considerable volatility and often reflect other influences in addition to differences in price levels between countries (and, therefore, do not reflect the comparative costs of labour and material purchases made by telecommunications operators). Exchange rates based on PPP (purchasing power parities), on the other hand, eliminate the impact of differences in price levels between countries. It might be argued that, for goods that can be purchased in international markets, actual market exchange rates would be appropriate. However, even in this case, when exchange rates are fluctuating the prices of goods rarely fluctuate to the same extent and hence the use of actual rates could give a misleading indication of the actual costs

In our analysis we have used PPP exchange rates for all operating expenses and asset categories. The PPP data used in this study has been obtained from the OECD publication "Purchasing Power Parities and Real Expenditures" and the IFS Yearbook.

The OECD publication provides data only for the PPP of GDP in recent years. A previous publication, last published in 1993, calculated PPPs for specific asset categories (for example, non-residential buildings, electrical equipment, non-electrical equipment, transport equipment and civil engineering works). However, this has not been updated for the relevant asset categories. Also, the PPP for GDP is generally considered to be more statistically reliable as it is based on a significant number of data points, whilst the PPPs for specific asset categories, even when they are available, are based on a relatively small number of data points for each category.

Therefore, NERA has used the GDP PPP exchange rate for each year to convert BT's costs into US$.

3.3.2. Output

Openreach supplied data on:

- the number of PSTN switched lines (both business and residential);

- the number of ISDN 2 and ISDN 30 lines (which we converted into 64Kbps channels);
- the number of unbundled local loops.

When carrying out the comparative efficiency assessment using OR incl LL, it was also necessary to have information on the number of leased and partial private circuit ends. This information is available from BT's current cost financial statements for different capacities of circuit. This enabled us to calculate the number of leased and partial private circuit ends in 64 Kbps channels and hence use comparative data for OR incl LL.

3.3.3. Network and environmental variables

Openreach provided data on the route (sheath) length of aerial copper cable, non-aerial copper cable and fibre, both in the access network as a whole and for Openreach WLR and LLU in isolation. From this we were also able to calculate the share of fibre in total sheath length.

In addition to these variables, a number of environmental variables have been compiled by NERA that are likely to provide a partial explanation of differences in costs. These include the population density of the UK and the percentage of the population living in metropolitan areas (defined as areas with a population above 100,000).

3.3.4. Quality of Service Data

We constructed a number of quality of service variables for Openreach that were comparable with quality of service measures that could be derived for the LECs given the data available in the ARMIS database. The source of the data that we used to construct these variables was published Openreach key performance measures for wholesale line rental in April 2007. We constructed the following variables:

- The percentage of orders met by the commitment date for both new lines and transfers;
- The number of faults per line; and
- The average interval between faults being reported and being repaired.

These values were a weighted average of all of the reported values for business wholesale access, residential wholesale access, ISDN2 and ISDN30.

3.3.5. BT Network

As a reference point, both with respect to OR efficiency and BT network efficiency in previous years (from previous studies), we also carried out a brief comparative assessment of BT Network as a whole. This required the following additional data:

- Costs for all the different components of BT Network (OR plus BT wholesale), which were provided by Openreach and cross-checked against the more aggregated numbers in BT's regulatory accounts. The following costs were then removed (as in previous studies) to ensure that BT Network was comparable to LEC network:
  - Operator assistance and emergency services (which are not included in the LEC costs);

  - Data network costs apart from broadband (which again are not included in LEC costs);
  - Payments to mobile operators for call termination (of which there are none in the US);

- Cable sheath length for the whole of BT's inland network (core and access) which was derived by projecting forward figures from earlier years and sanity checking the result against the current amount of cable sheath in the access network;
- Local and main switch minutes, which were derived from BT's regulatory accounts.

3.4. US LEC Data

3.4.1. Data Sources

We collected data on US LEC asset values, depreciation, operating expenses, outputs and network characteristics for the years 1999 to 2006 from the ARMIS database, which is compiled by the Federal Communications Commission (FCC). This database can be accessed at http://www.fcc.gov/wcb/armis/. It provides extensive information for around 70 LECs and it is these companies that have been used as comparators in this study. A less complete dataset is available for a larger number of companies.

The ARMIS dataset is more comprehensive than the Statistics of Common Communications Carriers (SOCC) tables which were used in earlier NERA studies. The SOCC tables are derived from the ARMIS database, but aggregate a number of state operators to the regional Bell operating company (RBOC) level (most noticeably, Bellsouth and Qwest). In addition, the ARMIS database is more up-to-date, with revisions being entered into this database but not the SOCC tables.

We also collected some data from the US census bureau statistical abstract for the purposes of deriving variables to describe the characteristics of the operating environment of the LECs.

3.4.2. US LEC costs

In order to enable comparisons to be made with Openreach cost data, which was reported on a current cost accounting basis, where appropriate the LEC asset values have been adjusted so that they are also in current cost terms.

Costs for the LECs comprise three elements:

- depreciation;
- the cost of capital; and
- operating expenses.

3.4.2.1. US LEC asset values

As the asset book values reported for the LECs are all on a historical cost accounting (HCA) basis, each operator's reported asset values are influenced by the point in time at which the assets were purchased. Due to changes in asset prices over time, operators with relatively older asset bases will report different gross book values (GBVs) to those with relatively newer asset bases, for comparable assets. This difference impacts on the depreciation and cost of capital of operators. Hence, it was necessary to adjust the reported asset values to express the book values reported by all operators on a comparable basis. To do this, the assets bases of the LECs were converted from historic to current costs. The derived costs from this process were compared to Openreach's costs as derived under current cost accounting (CCA).

The process by which this adjustment was made was as follows:

- Firstly the average age of the assets of each LEC was estimated for each of the following asset categories: poles, cable, duct, switches, transmission equipment, radio transmission, payphones, buildings, accommodation plant, computing equipment, and vehicles.[7] The actual age of assets is not reported by the LECs, so it is necessary to estimate the average age for each asset category using the following formula:

$$Average\ Age = Asset\ Life \times \left(1 - \frac{Net\ Book\ Value}{Gross\ Book\ Value}\right)$$

- Secondly, data on US price indices for each of these asset categories was collected from the US Statistical Abstract and the US Bureau of Labour Statistics. For each year of data these price indices allow the identification of the change in asset prices. For example, the price index for Telephone Switching and Switchboard Equipment was 105.3 in 1992 and 109.6 in 2006. This indicates that the price of this asset increased by 4.0% between 1992 and 2006. Two points to note are that:
  - In 2003 the construction of producer price indices in the US was revised and consequently some of the price indices that we have used in previous efficiency studies were discontinued. Where this has occurred we have created a pseudo index, where the discontinued indices have been updated using the values taken from the closest available index; and
  - The index that we use to identify changes in cable asset prices is constructed from two separate indices, one representing the costs of purchased communications cable and one representing the cost of telecoms construction workers. This is to reflect the fact that the capitalised cable asset values are a combination of the physical cable cost and the cost of putting the cable in place.

- Thirdly, the GBV of the assets of each LEC, for each asset category, was adjusted for the asset price changes that occurred between the average point in time at which the assets were purchased (which is determined by their average age) and the year for which the data was presented (1999, 2000, and so on). For 2006, this adjustment can be represented by the following equation:

$$GRC_{Cable2006} = GBV_{Cable2006} \times \left(\frac{Cable\,Price_{2006}}{Cable\,Price_{2006-AverageAge}}\right)$$

[7] As noted in Section 3.4.2.5, the costs of LEC switches, transmission equipment, radio transmission and payphones are not used in the efficiency comparisons because they do not form part of Openreach's asset base.

Where $cableprice_{2006}$ / $cableprice_{2006\text{-average age}}$ is the estimated LEC CCA/HCA ratio

An alternative approach to the adjustment suggested above would be to use Openreach's CCA/HCA ratio as a proxy for the CCA/HCA ratios for the LECs. In a previous efficiency study undertaken for BT as a whole we took this approach because the information required to calculate LEC specific CCA/HCA ratios was unavailable. Our reasons for not using this alternative approach in this study are that it:

- constrains the average age of assets to be the same across all companies, which is unlikely to be the case in practice; and
- does not allow the average asset age for the LECs to change over time, which is something that we would expect to happen over the period 1999-2006 for which that dataset has been constructed.

3.4.2.2. US LEC depreciation rates

To calculate the depreciation costs of the LECs, it is necessary to use appropriate asset lives. We calculate the implied asset life for each category of asset, for each year and for each LEC. The implied asset life is taken to be the historic cost depreciation charge for the asset category divided by the gross book value of the assets. The effective depreciation rate is then calculated as the inverse of the implied asset life.

An alternative to this is to use Openreach's effective depreciation rates and then apply these to the LECs to calculate the depreciation costs. We used this approach for some previous studies of BT's efficiency when the data that would enable us to calculate LEC specific asset lives was not available. The reasons for not taking this approach in this study are outlined below.

- The appropriate asset life to use may be different for different companies, reflecting the different mix of assets within a certain asset category. This is because the asset mix within asset categories will reflect decisions taken by the companies. For example the asset category cable includes both copper cable and fibre cable which may have different asset lives. The relative proportions of copper cable and fibre cable included in the cable category for each LEC may be a result of efficient (or inefficient) decisions taken by the companies and the calculated depreciation costs should reflect this.
- When using a dataset spanning ten years, the assumption that the effective depreciation rate will remain constant is unlikely to hold true, since the effective depreciation rate will be influenced by the proportion of capital which has been fully written down. Furthermore, this approach allows for the fact that older assets may have been depreciated at a higher or lower rate than currently exists.
- If we were to apply Openreach depreciation rates to the LECs it would be inconsistent with our use of LEC specific average asset ages to adjust LEC asset values from historic cost values to current cost values.
- Ofcom has previously raised concerns over the fact that BT has a procedure whereby assets are removed from their GRC asset base when they are fully depreciated, even if they continue to be used. As far as we know Openreach does the same thing. Compared to the situation where this procedure is not applied, this has the effect of raising Openreach's effective depreciation rates and NRC/GRC ratios. It is not clear that the

LECs adopt the same procedure. If they do not, applying Openreach's effective depreciation rates to LEC GRCs would overstate the LECs' depreciation costs.

For these reasons, given that the relevant data is now available from the ARMIS dataset, NERA has used the actual effective depreciation rates for the LECs when estimating their depreciation costs. Using actual and year-specific depreciation calculations should ensure that the dataset is more robust.

A further alternative would be to use Openreach's stated (as opposed to effective) depreciation rates. However, this will not necessarily lead to a directly comparable depreciation figure for the LECs, since Openreach's actual depreciation itself may not follow these assumptions. This is especially true where there are assets which have been purchased when stated asset lives were different to those used now. It is preferable, therefore, to look at the actual depreciation recorded by each US operator in each year[8].

3.4.2.3. US LECs cost of capital

It is also necessary to calculate the cost of capital for the LECs. To do this, the net replacement cost for each asset type is multiplied by an appropriate weighted average cost of capital (WACC).

There are two possible approaches that can be followed. Firstly, the cost of capital could be calculated using a constant WACC for all firms over all years. The second approach would be to use a cost of capital that was specific to each LEC but constant for all years. A variation on these approaches would be to allow the cost of capital to vary over time.

We have adopted the first approach and applied Openreach's current weighted average cost of capital to each of the LECs in each year. We believe that this is preferable to using a LEC specific cost of capital because Openreach's WACC relates to access network activities and reflects the fundamental risk associated with these activities. In contrast, for each LEC, WACC could only be estimated for the whole business and this would include retail and network activities and, within the latter, core as well as access network activities. Moreover, the calculation of a specific WACC for each LEC would be a significant task in itself.

As regards using the same cost of capital in each year, we do not have estimates of Openreach's WACC for earlier years, not least because Openreach did not exist. The cost of capital does change with time, although it is typically fairly stable over the long run. However, this is not likely to be a major concern for an efficiency study of this kind because the relative WACC of each of the companies will remain fairly constant over time.

The WACC has been applied to the net replacement cost (NRC) to derive the cost of capital for each company. NRC is calculated, as described in section 3.3.1, by multiplying the GRC by the NBV/GBV ratio.

---

8 Although the depreciation written down each year is influenced by accounting practices rather than actual economic usage, it is not possible to measure the latter for either the US LECs or BT. Where accounting lives differ from economic lives, therefore, there may be a slight inaccuracy in our calculation of the actual depreciation of a firm's assets. However, this difference is not likely to be material.

3.4.2.4. US LEC operating expenses

For each year between 1999 and 2006 we have collected operating expense data for the LECs. These costs are allocated to the asset categories within the ARMIS database.

3.4.2.5. Adjustments to costs

A number of adjustments have been made to the US costs:

- The LECs separately report the value of "Telecommunications Plant under Construction" (TPUC), without identifying the component assets. Therefore, it is not possible to identify to which asset categories this construction activity relates. However, it only represents a very small element of the overall asset base of the LECs. In 2006, for example, TPUC accounted for, on average, only around 2% of the total telecommunications asset base. Since there is no information available concerning exactly which types of asset are included in this category, we allocate these assets to the GRC of each asset category based on the assumption that it is exclusively replacement investment (i.e. we allocate plant under construction to the various asset groupings in proportion to the relative level of depreciation for each type of asset).
- The LEC cable asset values were reduced by 10%. This adjustment is made to reflect the fact that historic cable asset values contain some double counting of the costs of installing the cable. It is these asset values that are indexed to derive current costs asset values for the LECs. The size of the adjustment that we have made reflects the estimated size of this so called "double dig" effect for BT. The same adjustment does not need to be made to BT's cable asset values because these have been derived by direct revaluation rather than by indexing historic book values.
- The LECs consider bad debts to be a reduction in revenue. However, Openreach includes bad debts in its operating expenses. Therefore, bad debts have been included in the calculation of the US LECs' operating costs.
- The data for the LECs covers both network (core and access) and retail activities. Openreach's scope of activities incorporates only the access network, in particular wholesale line rentals and local loop unbundling. Therefore, it is necessary to remove those costs of the LECs that are related to core network and retail operations. To do this, all cost categories for the LECs were first categorised as either:

  - directly attributable to the running of the core network;
  - directly attributable to the running of the access network;
  - directly attributable to retail activities; or
  - indirectly attributable to all of the above.

  We have removed all categories of costs from the total LEC costs which can be directly attributed to core network and retail activities. These includes costs related to: switching equipment; transmission; radio transmission; payphones; operator service; customer service and marketing. It was not, however, possible to identify separately and remove the costs of duct and cable in the core network because these are aggregated together with the same types of cost in the access network.

Indirect costs were allocated to the access network activities according to the proportion of total directly attributable costs accounted for by access network activities. The exceptions to this are capital and operating costs that relate to buildings for which the proportion of total employees employed by the access network is more likely to reflect the costs attributable to this activity. To ensure consistency in the allocation across all of the LECs and Openreach we allocated building costs to the LEC access network activities according to the proportion of total BT Group employees who work for Openreach (which is approximately 30%).

3.4.3. US LEC Outputs

As for Openreach, output data for the LECs comprises the number of switched lines, leased lines (including special access lines), switch minutes and a variety of quality of service variables.

3.4.3.1. Line numbers

US LECs operate three types of access line:

- switched lines;
- leased (private) circuits; and
- special access lines.

While the numbers of switched and special access lines are reported in the ARMIS database for every year in our sample, the number of leased circuits has only been reported since 2002. Therefore, we have calculated the number of leased lines using the leased line revenues reported to the ARMIS cost database. These revenues have been divided by the average price of a 64kbps-equivalent leased line, as reported by the OECD[9], to give an estimate for the number of 64kbps-equivalent leased lines operated by each LEC.

For purposes of comparability with BT, it is necessary to aggregate leased circuits and special access lines into one variable. To do this, NERA has adopted the methodology used in previous studies. Rather than assuming an equal weight for both types of line, we consider the components of each type of line. A leased line has two customer ends and generally a main link, while a special access line has only one customer end (special access lines connect with an inter-exchange carrier point of presence) and generally has a main link and an interconnecting circuit at the inter-exchange carrier point of presence. Therefore, we place a weight of 2 on leased lines, and a weight of 1.5 (more than a pure switched line but less than a leased line) on special access lines.

We also noted that the line numbers data for a number of Verizon companies exhibits unusual growth between since 2004.[10] The number of leased lines (in 64 kbps equivalents) reported

---

9 OECD Communications Outlook 2006, OECD

10 The LECs affected by the change in Verizon's reporting of leased lines are: Verizon NE – Maine; Verizon NE – Massachusetts; Verizon NE – New Hampshire; Verizon NE – Rhode Island; Verizon – New York Telephone; Verizon – Delaware; Verizon – Maryland; Verizon – New Jersey; Verizon – Pennsylvania; Verizon – Virginia; Verizon – Washington, DC; and Verizon – West Virginia. The Verizon NE companies plus New York Telephone were part of Nynex, while the other LECs were part of Bell Atlantic.

for certain LECs that are part of the Verizon group increased by more than 100% from 2004 to 2005. These large increases were far greater than might have been expected on the basis of growth in earlier years and were systematically associated with LECs which formed part of the previous Bell Atlantic and Nynex regional Bell companies. Our conclusion was that Verizon changed the way in which it reported leased lines for these companies from 2005 onwards. Unfortunately, it is not clear whether the earlier or current data is more consistent with the practice followed by other LECs. To allow for this change we have included a Verizon dummy variable in our equations. We explain our approach to this "Verizon anomaly" further in section 4.1.1.

There is one further issue regarding the LEC line data that we take into account in our efficiency analysis. We have noted that the numbers of switched lines reported by most LECs peaked at the beginning of the current decade (usually in the year 2000); these numbers have fallen substantially since 2000 – in many cases by 20-25%. Since the number of switched lines is the most important determinant of costs, this pattern gives rise to the possibility that the total cost of the LECs are affected by costs associated with stranded assets installed to meet a higher level of demand for switched lines in the past. We discuss how we deal with this issue further in section 4.1.4.

3.4.3.2. US LEC quality of service

For each year we collected data on the quality of service provided by the LEC's. From this data we constructed three variables that are comparable to quality of service data that is reported for Openreach. These variables are:

- The percentage of orders met by the commitment date;
- The number of faults per line; and
- The average interval between faults being reported and being repaired.

3.4.4. US LEC Network and environmental data

Data for the US LECs has been collected to match that provided by BT. This includes:

- Sheath length in the network (split into aerial and non-aerial);
- Duct route length;
- The proportion of the network which uses optical fibre;
- The proportion of business lines to residential lines; and
- The average wage of employees.

In addition, a number of environmental variables have been derived including the population density and proportion of the population living in metropolitan areas.

3.4.4.1. Regional dummy variables

In some, but not all, previous studies, NERA has attempted to account for the impact of higher wages and other costs in the North East of the United States by including a dummy variable for those companies operating in this area. There is not an overwhelming a priori case for the inclusion of these variables, so the decision whether to include them or not has depended on the statistical performance of the estimated efficiency models when they are included.

This time, we considered an alternative approach, namely using an average staff cost variable for each company in our efficiency modelling. Using the ARMIS data, it was possible to derive an average staff cost variable on a state by state basis. This allowed us to consider the effect of variations in average staff costs, and we have found it not to be a significant cause of cost variations between companies.[11]

Similarly, in previous studies BT argued that a dummy should be included covering the Midwest region of the US. BT claims that companies operating in this region, (particularly those formerly part of the Ameritech group) have reacted in particular ways to the change in the general US regulatory environment (that is, the move away from rate of return regulation) since 1994. More specifically, BT believes that these LECs have:

- reduced their capital bases by lowering annual capital expenditure;
- increased profitability by cutting their number of employees and simultaneously increasing productivity and reducing expenditures;
- increased profitability by stimulating revenue (more features and lines sold) with little increase in their asset base; and
- reduced costs and increased profits by reducing service quality.

Dealing with the first point, it is a well known result that rate of return regulation provides incentives that can lead to overinvestment and an inefficiently high capital to labour ratio[12]. Given a move to price cap regulation, such as has occurred in the US since the 1990s, one might therefore expect there to be a reduction in the rate of investment as firms strive to return to efficient capital to labour ratios. Rather than being an adverse development this would be a manifestation of the efficiency inducing properties of price cap regulation.

As regards the second and third points above, these would appear again to reflect the beneficial effect on efficiency of price cap regulation, to which BT itself is also subject.

Finally, in relation to the fourth point above, while some of these companies have a relatively high number of customer complaints (reflecting the quality of service, as reported by the FCC), the data concerned are far from complete and it is unclear whether there are differences between companies in terms of the way that complaints are reported and defined or in the way that the data are collected.

NERA has carried out an analysis into the relative efficiency of these LECs, to see whether it has in fact significantly improved since 1994. To do this we compared the position of these companies in a study undertaken for Ofcom in 1995 with that in the current study. There are a number of difficulties in comparing the numbers obtained in NERA's 1995 report[13] with the

---

11 The implication is that wage rates affect productivity and vice versa.

12 Averch H and Johnson LL (1962) "Behaviour of the Firm under Regulatory Constraint", American Economic Review 52, pp1052-69

13 "The Efficiency of BT", NERA, September 1995

results outlined in this study. These difficulties include a differing methodology used to calculate efficiency, and a different set of companies. However, we can obtain a rough idea of whether the relative efficiency of these companies has improved by looking at the relative rankings of companies between studies.

For total costs, the companies' rankings are as shown in Table 3.1:

Table 3.1
Rankings of Ameritech companies

Company	Rank in 1995 study (out of 30 firms)	Rank in current study (out of 68 firms)
Illinois Bell	3	22
Indiana Bell	7	5
Michigan Bell	13	7
Ohio Bell	8	20
Wisconsin Bell	4	9

Source: NERA analysis

This table shows that there has not been a uniform improvement in the relative efficiency of these firms compared to the sample as a whole over the period since rate of return regulation was abandoned. The group of five companies is in the top half of the sample in both studies, but the relative positions of the companies vary.

As stated above, BT have also argued previously that companies in the Midwest, particularly those formerly part of Ameritech, have cut back the number of employees in order to increase profitability and reduce costs. Analysis of employee numbers shows no such trend. For each of the five companies, there was only a small decrease in employee numbers up until 1999, after which employee numbers have increased. This is no different to the pattern exhibited by the rest of the LECs.

We have therefore found that there is no evidence that conclusively proves BT's assertions. The relative efficiencies of the companies in question have not all significantly improved over the years that NERA has been carrying out studies of BT's comparative efficiency. Furthermore, there is also no evidence that the number of employees for these firms shows a decline compared to those of other firms.

In this study we have explored the performance of efficiency models which include North East and Mid-West dummy variables compared with models that do not and have found that the models which include the dummies perform better. We have also examined efficiency model which includes quality of service variables.

3.4.5. Other data issues

3.4.5.1. Structural breaks

In previous efficiency studies statistical analysis of the data collected for US companies indicated that there is a structural break in the cost data between the years 1998 and 1999. A structural break occurs when there is a step change in the level of costs or other variables.

There are a number of possible explanations for this break. For example, from analysis of NERA's dataset, there was a sharp decrease in cable and wire prices between 1998 and 1999[14] which led to a parallel fall in the LECs' CCA asset values for cable.

To deal with this structural break, NERA considered two possible options. Firstly, it is possible to run the model only considering years since 1999. However, this reduces the number of observations and, other things being equal, may not produce such a robust model. In particular, for variables which do not experience a structural break, we would be losing explanatory power.

Secondly, we could introduce a dummy variable to measure the effect of the break between 1998 and 1999, as well as a number of interaction terms to allow the impact of output variables to change between the period up to 1998 and the period from 1999 onwards. By including these interaction terms, we allow the model to fit closely to data over only the last eight years, while using the full dataset to produce the cost model.

In previous efficiency studies our preferred option has been to use dummy variables so that we minimise the loss of data to improve the statistical robustness of the efficiency modelling. For the purposes of this study our preferred approach is to drop the LEC data for the years prior to 1999 from the dataset that we use for the efficiency modelling. The reason for this is that by updating our dataset to 2006 we increase the number of available data-points. We consider the period 1999 to 2006 to provide sufficient data to ensure that the model is statistically robust.

3.4.5.2. Excluded companies

A number of US companies which report to the ARMIS database are excluded from our analysis. These are:

- Puerto Rico Telephone Company and Verizon Hawaii, which were excluded because data was not available for these companies for a number of years.
- A number of small Contel companies, which provide cost data separately from a sister company in the same area. For example, Verizon North Contel Indiana provides cost data separately from Verizon North Indiana. However, output and environmental data for these companies is not available on a disaggregated basis. NERA has been advised by the FCC that operations within these Contel companies are mainly governed by the larger companies in each state, and it is therefore reasonable to consider Verizon North's operations in Indiana as a whole. Other companies that fall into this category are: Contel California; Qwest-Idaho North; Verizon SO-Contel-North Carolina; Verizon SO-Contel-South Carolina; Verizon SO-Contel-Virginia; Verizon NO-Contel/Illinois; Verizon NO-Contel/Indiana; Verizon NO-Contel/Pennsylvania; Verizon NO-Contel/Quaker State; andVerizon SW-Contel-Texas.

[14] As taken from the US Statistical Abstract's PPI data.

4. Results

In this section the results of the comparative efficiency assessment are set out and discussed.

Firstly, we look at the final model specification derived for analysis of BT's costs, and examine the definition of costs used. Secondly, we set out our results obtained from using this model.

4.1. Model Specification

As described in Section 3, NERA has collected a number of potential cost driver variables for each operator. Our main analysis is based upon a Cobb-Douglas log-log model, as discussed in Section 2.2, which results in a specification of the following form:

$$\ln(\text{Total Cost}) = a + b_1 \text{ Outputs} + b_2 \text{ Environmental} + b_3 \text{ Time}$$

In this equation, the output variables include switched lines and leased lines. Environmental variables include factors such as the population density, the cable sheath length used in the network, and the ratio of business lines to residential lines.

4.1.1. Differences from previous NERA studies

While the general form of this specification is similar to that used in our previous studies it is necessary to bear in mind that changes in the nature of the data that is available have affected the detailed implementation of our analysis. Thus, the following points should be noted:

- In previous studies we have estimated cost frontiers for operating costs and operating costs plus depreciation as well as for total costs. However, our review of the data for Openreach showed that the latter relies heavily upon using equipment that is leased from BT. The cost of equipment leased is recorded as an operating cost, whereas the LECs normally own the equivalent equipment which means that the associated costs will mainly be recorded under depreciation and cost of capital. Hence, operating costs mean different things for Openreach and the LECs. We have confirmed that Openreach is an extreme outlier with respect to the cost frontier for operating costs. Consequently, the results reported here focus exclusively on total costs – i.e. operating costs plus depreciation plus cost of capital.
- In our previous studies we found that there was strong evidence of a structural break in the data in the late 1990s. This led to the adoption of a specification that included a dummy variable for the years 1996-98 plus separate coefficients for many of the independent variables for the same years. Our preliminary analysis of the LEC data confirmed that the structural break can still be identified in the data. However, we have a larger number of years of data for each LEC – a maximum of 11 observations from 1996 to 2006. Thus, the loss in estimation efficiency from dropping data for the years 1996-98 is considerably lower than in the past. After some investigation we concluded that the option of retaining the dummy variable and separate slope coefficients in the model for the full period 1996-2006 produced less reliable results than could be obtained by analysing the data for 1999-2006 without any need to correct for the structural break. Hence, the results reported here refer to the estimation of a cost frontier using data for 8 years from 1999 to 2006.

- We have included regional dummy variables for (a) LECs in the New York and New England region (NYNE), and (b) LECs in the MidWest region (MidWest) among our environmental variables. This is a change from the specification adopted in our last study, though it is consistent with our practice in earlier studies. The arguments for including regional dummies are finely balanced. In principle, it would be desirable to capture regional effects via some more explicit independent variable or variables such as the level of wages. Unfortunately, our attempts to do this have never produced satisfactory results, possibly because the data on wages or average earnings is highly aggregated. In addition, our results suggest that the regional influences on total costs are more than a matter of differential wage rates. There seem to be a variety of factors which combine to produce substantially higher costs for LECs operating in the NYNE region and lower costs for LECs operating in the MidWest. For this study we found that the NYNE and MidWest dummy variables had consistent and highly significant coefficients, especially once the data for the year 1996-98 was dropped. Hence, the models discussed in this section include these dummy variables.

- There is one other data-related adjustment to our model that requires comment. As explained in section 3.4.3.1, we noted that the number of leased lines (in 64 kbps equivalents) reported for certain LECs that are part of the Verizon group increased by more than 100% from 2004 to 2005. These large increases were far greater than might have been expected on the basis of growth in earlier years and were systematically associated with LECs which formed part of the previous Bell Atlantic and Nynex regional Bell companies. Our conclusion was that Verizon changed the way in which it reported leased lines for these companies from 2005 onwards. Unfortunately, it is not clear whether the earlier or current data is more consistent with the practice followed by other LECs. To allow for this change we have included a Verizon dummy variable in our equations.[15] We experimented with two ways of defining the dummy variable: (a) a dummy variable (VDUMMY) that takes the value 1 for the relevant LECs in the years 2005 and 2006 and the value 0 for all other cases; and (b) a dummy variable that takes the value 1 for the relevant LECs in the years 1999-2004 and the value 0 for all other cases. We found that the second option consistently generates a higher value for the log-likelihood, so we have used this definition in the results reported here. In practice, the coefficient on the Verizon dummy is rarely significant at the 5% level, but we have retained the variable because there is patently an anomaly in the data.

4.1.2. Treatment of Openreach

When running this regression model, we include Openreach in the dataset but give it an extremely low importance weighting in the calculations.[16] It would have been preferable to have excluded Openreach altogether but the manner in which the efficiency calculations are performed by the statistical software (Stata) requires that Openreach should be part of the

[15] The LECs affected by the change in Verizon's reporting of leased lines are: Verizon NE – Maine; Verizon NE – Massachusetts; Verizon NE – New Hampshire; Verizon NE – Rhode Island; Verizon – New York Telephone; Verizon – Delaware; Verizon – Maryland; Verizon – New Jersey; Verizon – Pennsylvania; Verizon – Virginia; Verizon – Washington, DC; and Verizon – West Virginia. The Verizon NE companies plus New York Telephone were part of Nynex, while the other LECs were part of Bell Atlantic.

[16] This means that the values for Openreach make a very small contribution to the log-likelihood function that is maximised by the estimation. The weight given to Openreach is 0.000001 – i.e. one divided by one million – whereas all other observations have a weight of 1.

dataset analysed. It is essential to give BT a low weight because it is much the largest company in the sample, so that giving it the same weight as other LECs would tend to skew the equation by "pulling" the slope of the regression line towards itself. This would have the effect of reducing Openreach's observed efficiency or inefficiency, and worsening the fit of the model. By giving Openreach a very low weight in the estimation, we ensure that the regression is not biased in this way, and we may then use the estimated equation to predict BT's costs outside of the model.

4.1.3. Derivation of model specification

We started by using a general-to-specific methodology to identify the independent variables to be retained in the model. Under this methodology, a general specification is formed using all available variables which may be considered potential drivers of costs. A step-by-step process is then used to identify those variables which are empirically (rather than just theoretically) significant explanatory variables and to reject those that are not. Decisions on the step-by-step exclusion of variables were made generally on the basis of statistical significance[17]. One particular variable that was dropped at this stage was the log of switched minutes. In the past we have found that this variable was a significant cost driver for operating costs but rarely for total costs. Since the current exercise focuses on the costs of running the access network, it is not surprising that switched minutes are not significant in our models.

The basic specification derived from this analysis takes the form:

$$\begin{aligned}\ln(\text{Total Cost}) &= a + b_1 \ln(\text{Switched Lines}) + b_2 \ln(\text{Leased Lines}) + b_3 \ln(\text{Total Sheath}) \\ &+ b_4 \ln(\text{Population Density}) + b_5 \ln(\text{Business Residential Ratio}) + b_6 \ln(\text{Fibre Proportion}) \\ &+ b_7 \,\text{Vdummy} + b_8 \,\text{NYNE} + b_9 \,\text{MidWest} + b_{10} \,\text{Time}\end{aligned}$$

The functional form of the model outlined above allows us to interpret the coefficients on each variable as an elasticity. This means that, if we see a 1% increase in the number of switched lines, we would see a β_1% increase in costs. The Time variable is set at 1 in 1996, and increases by 1 each year, so the coefficient β_{10} gives the percentage increase or decrease in costs over time after allowing for the impact of the various cost drivers and environmental variables.[18]

This model allows the possibility of non-constant returns to scale – i.e. costs which increase more or less than proportionately with the overall scale of the operations of the company. In the past we have found that the cost models tend to display constant returns to scale since the sum of the coefficients β_1 for switched lines, β_2 for leased lines, and β_3 for total sheath is usually very close to 1. We tested whether it would be reasonable to impose an assumption of constant returns to scale by testing the basic specification against the specification for the Constant Returns Model:

---

[17] The significance of variables was tested using both the t statistics for the variable in the estimated regression and by testing that excluding the variable from the regression did not significantly alter the model fit (as measured by the log likelihood function).

[18] All costs are measured in nominal terms, so in order to find the change in real costs over time it is necessary to take account of the rate of inflation. In the US, over the period studied, the average annual inflation rate has been 2.50%. Therefore, the change in real costs over time can be estimated by the change in nominal costs minus 2.50%.

$$\ln(\text{Total Cost / Switched Lines}) = a + b_2 \ln(\text{Leased Lines / Switched Lines})$$
$$+ b_3 \ln(\text{Total Sheath / Switched Lines}) + b_4 \ln(\text{Population Density})$$
$$+ b_5 \ln(\text{Business Residential Ratio}) + b_6 \ln(\text{Fibre Proportion})$$
$$+ b_7 \text{ Vdummy} + b_8 \text{ NYNE} + b_9 \text{ MidWest} + b_{10} \text{ Time}$$

We found that the constant returns to scale restriction cannot be rejected and have used that when reporting our results.

4.1.4. Stranded assets

There is one further modification to the basic specification that we have investigated. We noted that the numbers of switched lines reported by most LECs peaked at the beginning of the current decade (usually in the year 2000); these numbers have fallen substantially since 2000 – in many cases by 20-25%. Since the number of switched lines is the most important determinant of costs, this pattern gives rise to the possibility that the total cost of the LECs are affected by costs associated with stranded assets installed to meet a higher level of demand for switched lines in the past. The number of switched lines provided by BT/Openreach has not followed a similar path, so if stranded assets are an important factor in determining costs then there is a danger that our cost frontier would be mis-specified, leading to an unreliable estimate of Openreach's efficiency.

As a measured of the proportion of stranded assets we have defined the following variable:

Stranded Ratio = Max (Switched Lines) / Switched Lines

where the maximum is taken over the numbers of switched lines for the years 1999 onwards. The stranded ratio is set equal to 1.0 for Openreach. Introducing this variable into the model above yields the following specification for the Stranded Assets Model:

$$\ln(\text{Total Cost / Switched Lines}) = a + b_1 \ln(\text{Stranded Ratio})$$
$$+ b_2 \ln(\text{Leased Lines / Switched Lines}) + b_3 \ln(\text{Total Sheath / Switched Lines})$$
$$+ b_4 \ln(\text{Population Density}) + b_5 \ln(\text{Business Residential Ratio}) + b_6 \ln(\text{Fibre Proportion})$$
$$+ b_7 \text{ Vdummy} + b_8 \text{ NYNE} + b_9 \text{ MidWest} + b_{10} \text{ Time}$$

4.2. Comparisons of the Cost Frontier for Alternative Specifications

4.2.1. Basic Model v Constant Returns Model

Table 4.1 compares the results obtained by estimating the three alternative models discussed above. Comparing the Basic Model using ln(Total Cost) with the Constant Returns Model using ln(Total Cost per Switched Line) shows that the coefficients for the matched variables are very similar. Their log-likelihood values differ by only 0.8, implying a likelihood ratio test statistic of 1.6 with 1 degree of freedom on the restriction of constant returns. This does not reject the restriction, so that we may proceed with the Constant Returns Model.

Table 4.1
Estimates of the cost frontier for alternative models

Dependent Variable	Basic Model Ln(Total Cost)			Constant Returns Model Ln(Total Cost/Switched Line)			Stranded Assets Model Ln(Total Cost/Switched Line)		
	Coeff	**Std Err**	**Z-ratio**	**Coeff**	**Std Err**	**Z-ratio**	**Coeff**	**Std Err**	**Z-ratio**
Independent Variables									
Ln(Switched Lines)	0.689***	0.042	16.47						
Ln(Stranded Ratio)							0.492***	0.066	7.49
Ln(Leased Lines)	0.019	0.014	1.36						
Ln(Leased Lines per Switched Line)				0.022	0.014	1.59	0.023*	0.013	1.75
Ln(Total Sheath)	0.307***	0.038	8.12						
Ln(Total Sheath per Switched Line)				0.300***	0.038	7.98	0.162***	0.033	4.84
Ln(Population Density)	0.037**	0.019	2.00	0.042**	0.020	2.12	0.014	0.014	0.96
Ln(Business-Residential Ratio)	0.145***	0.036	4.04	0.139***	0.036	3.87	0.044	0.036	1.22
Ln(Fibre Proportion)	0.098***	0.037	2.63	0.116***	0.035	3.28	0.071**	0.032	2.25
Vdummy	-0.031	0.022	-1.38	-0.029	0.022	-1.29	-0.029	0.021	-1.38
NYNE	0.162***	0.047	3.44	0.160***	0.050	3.23	0.170***	0.044	3.91
MidWest	-0.196***	0.043	-4.52	-0.192***	0.046	-4.17	-0.146***	0.040	-3.66
Time	0.038***	0.004	8.48	0.036***	0.004	8.40	0.023***	0.004	5.41
Constant	-0.544**	0.258	-2.11	-0.327*	0.173	-1.89	-0.806***	0.151	-5.33
Log-Likelihood	505.4			504.8			531.4		

Source: NERA analysis

*Note: Significance levels for coefficients: * = significant at 10% level, ** = significant at 5% level, *** = significant at 1% level.*

4.2.2. Stranded Assets Model

The second comparison is between the Constant Returns Model and the Stranded Assets Model. In this case, we needed to test whether the inclusion of the log of the Stranded Ratio improves the performance of the Constant Returns Model. The z-ratio for ln(Stranded Ratio) is 7.49, while the likelihood ratio test statistic is 53.2 with 1 degree of freedom. The results imply that the hypothesis of a zero coefficient for this variable is decisively rejected. Thus, we concentrate on the Stranded Assets version of the model as the basic specification for the remainder of our analysis.

The coefficient on the stranded assets variable confirms the suspicion that the gradual decline in the numbers of switched lines for the LECs from the peak levels attained at the beginning of this decade has increased their efficient costs per switched line. A fall of 25% in the number of switched lines – implying a stranded ratio of 1.33 – is associated with an increase of 14% in the efficient cost per switched line. This is a substantial effect. Though it is not relevant to the situation of Openreach in 2006, there is the possibility that Openreach's efficient costs may increase in future if it were to experience a downward trend in the number of switched lines.

Apart from stranded assets, the most important cost drivers for the LECs are the amount of total sheath per switched line and the two regional dummies. The inclusion of stranded assets greatly reduces the coefficients for population density and the business-residential ratio with the consequence that they are no longer significantly different from zero in the Stranded Assets Model. The fibre proportion remains significant but again the estimated coefficient is reduced.

Inclusion of stranded assets also has an important effect on the time trend. The time trend in the Constant Returns Model implies that costs were increasing at an exogenous rate of 3.6% per year. The Stranded Assets Model suggests that only a part of that trend was exogenous – about 2.3% per year – while the remainder was the effect of a steady reduction in the number of switched lines relative to its peak level.

Overall, the Stranded Assets Model seems to be an intuitively plausible specification which yields reasonable results with respect to the estimation of the cost frontier.

4.3. Efficiency Estimates

Table 4.2 shows the efficiency ranking for the 68 LECs included in the analysis plus the two variants of Openreach. It should be recalled that there are considerable difficulties in ensuring that the coverage of the Openreach data matches that for the LECs. To deal with this problem we have created two variants of Openreach for comparison with the LECs. One variant - referred to as "Openreach incl LL" - includes costs associated with leased lines (excluding transmission equipment) and the number of leased line local ends provided by BT. The second variant – referred to as "Openreach excl LL"– excludes all Openreach costs associated with leased lines, while the number of leased line local ends is set equal to 1, so that ln(Leased Lines) = 0 in the model. Other variables such as total sheath length and the fibre proportion are adjusted to reflect the inclusion or exclusion of leased lines.

Table 4.2
Efficiency ranking for the stranded assets model

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
1	Qwest-Idaho South	3.7%	0.0%	-8.6%
2	Openreach incl LL	5.2%	1.5%	-7.2%
3	Qwest-Iowa	5.3%	1.6%	-7.1%
4	Wisconsin Bell	7.4%	3.5%	-5.3%
5	Contel-Nevada	8.5%	4.6%	-4.3%
6	Verizon NE-Rhode Island	10.7%	6.8%	-2.4%
7	Qwest-North Dakota	10.9%	6.9%	-2.2%
8	Qwest-Montana	13.4%	9.4%	0.0%
9	Verizon NO-Illinois	13.8%	9.7%	0.3%
10	BellSouth-Tennessee	16.0%	11.9%	2.3%
11	Verizon SO-Illinois	16.4%	12.3%	2.7%
12	Qwest-Minnesota	16.6%	12.5%	2.8%
13	Verizon-Virginia	17.6%	13.4%	3.7%
14	SBC/SNET-Connecticut	17.8%	13.6%	3.9%
15	Verizon-Delaware	18.8%	14.6%	4.8%
16	Verizon-Maryland	19.3%	15.0%	5.2%
17	Verizon-New Jersey	20.2%	16.0%	6.0%
18	Verizon NE-Maine	20.7%	16.4%	6.5%
19	Openreach excl LL	21.1%	16.8%	6.8%
20	Verizon-Washington D.C.	21.2%	16.9%	6.9%
21	BellSouth-North Carolina	21.3%	17.0%	6.9%
22	Verizon NO-Wisconsin	22.4%	18.0%	7.9%
23	BellSouth-Florida	22.5%	18.1%	8.0%
24	Qwest-South Dakota	22.5%	18.2%	8.0%
25	Indiana Bell	22.8%	18.5%	8.3%
26	Verizon SO-Virginia	23.1%	18.7%	8.6%
27	Verizon-Pennsylvania	23.4%	19.0%	8.8%
28	SBC-Kansas	23.7%	19.3%	9.1%
29	Pacific Bell-California	24.3%	19.9%	9.6%
30	BellSouth-Alabama	24.3%	19.9%	9.6%
31	Verizon NO-Ohio	24.7%	20.3%	10.0%
32	Qwest-Utah	26.4%	21.9%	11.4%
33	Nevada Bell	26.9%	22.4%	12.0%
34	Verizon NE-New Hampshire	27.0%	22.5%	12.0%
35	BellSouth-South Carolina	27.4%	22.9%	12.3%
36	BellSouth-Kentucky	28.0%	23.5%	12.9%
37	Qwest-Washington	28.6%	24.0%	13.4%
38	GTE-California	28.8%	24.2%	13.6%
39	SBC-Oklahoma	29.2%	24.6%	13.9%
40	SBC-Missouri	29.5%	24.9%	14.2%
41	Qwest-New Mexico	29.9%	25.3%	14.5%
42	Verizon NE-Vermont	30.4%	25.8%	15.0%
43	Qwest-Oregon	31.3%	26.7%	15.8%
44	Illinois Bell	31.9%	27.2%	16.3%
45	Verizon NE-Massachusetts	31.9%	27.3%	16.4%
46	Verizon NO-Michigan	32.3%	27.6%	16.6%
47	Verizon NW-West Coast California	32.7%	27.9%	17.0%
48	BellSouth-Louisiana	33.5%	28.8%	17.8%
49	Verizon NW-Oregon	34.2%	29.4%	18.3%
50	Ohio Bell	34.3%	29.5%	18.4%
51	Verizon SO-South Carolina	34.4%	29.6%	18.5%
52	Michigan Bell	34.4%	29.7%	18.6%

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
53	Verizon NO-Indiana	34.8%	30.0%	18.9%
54	Verizon NW-Idaho	35.1%	30.3%	19.1%
55	Qwest-Arizona	36.9%	32.0%	20.7%
56	Verizon NW-Washington	36.9%	32.1%	20.8%
57	BellSouth-Mississippi	37.0%	32.1%	20.8%
58	BellSouth-Georgia	37.4%	32.5%	21.1%
59	SBC-Arkansas	38.6%	33.7%	22.3%
60	Verizon-West Virginia	38.9%	34.0%	22.5%
61	Verizon NO-Pennsylvania	39.0%	34.1%	22.6%
62	Qwest-Nebraska	42.3%	37.3%	25.5%
63	Qwest-Colorado	42.4%	37.3%	25.6%
64	Verizon SW-Texas	43.4%	38.3%	26.5%
65	Verizon-Florida	44.3%	39.2%	27.3%
66	Qwest-Wyoming	45.4%	40.3%	28.2%
67	SBC-Texas	49.7%	44.4%	32.0%
68	Verizon SO-North Carolina	53.7%	48.2%	35.5%
69	Contel-Arizona	58.9%	53.2%	40.1%
70	Verizon-New York Telephone	60.5%	54.8%	41.5%

Source: NERA analysis

The figures show that OR incl LL ranks 2nd in the list in terms of efficiency, while OR excl LL is 19th. As in the past we have computed the percentage inefficiency of each company relative to the estimated cost frontier, the best company and the company which falls on the upper decile - in this case Qwest-Montana.[19] In each case, the percentage inefficiency is the amount by which the standardised cost for the company, after allowing for its operating and environmental characteristics and setting aside random variations, exceeds the reference cost – i.e. the frontier, the best company or the company on the upper decile for each of the inefficiency measures. A negative value for the percentage inefficiency means that the standardised cost is less than the reference cost. This is only possible for the percentage inefficiency relative to the upper decile.

The usual measure of Openreach's efficiency focuses on its performance relative to the decile. For OR incl LL the results suggest that it is 7.2% **more** efficient than the upper decile while for OR excl LL the results suggest that it is 6.8% **less** efficient than the upper decile. It is unfortunate that the results are not closer together, but the difference between the results illustrates the difficulty of defining an appropriate basis for comparing the adjusted data for the LECs and for Openreach. We discuss this difference further in Section 4.4.1 below. Our conclusion is that the result for OR excl LL cannot be relied upon for technical reasons. This leaves the efficiency estimate for OR incl LL as the best estimate that can be derived from our standard model, although this also has limitations in that it does nor measure the comparative efficiency of Openreach's WLR and LLU activities.

As a cross-check we have re-estimated our models using all BT Network and LEC costs and the drivers used in our previous studies but updated to include all years up to 2006. We find that BT Network is ranked 5th with an inefficiency of -3.8% relative to the upper decile.

---

[19] It should be noted that the upper decile is defined by considering the LECs on their own, so that Qwest-Montana is ranked 8th in the ranking of all LECs plus OR incl LL and OR excl LL but it is ranked 7th among the 68 LECs in the analysis.

Based on the comparisons that we have been able to make, Openreach is probably slightly more efficient than the upper decile. However, this result needs to be treated with caution because of the difficulties of obtaining data for the LECs and for Openreach that reflects the same costs and activities.

4.4. Sensitivity Tests and Model Robustness

This section describes a number of sensitivity analyses conducted to test the robustness of our model using the total cost methodology.

4.4.1. Dropping leased lines from the specification

The main reason for the difference between the relative efficiency estimates for OR inc LL and OR excl LL is a technical problem associated with the treatment of leased lines for the OR excl LL pseudo-company. This arises because our model includes ln(leased lines per switched line) as an explanatory variable. However, this cannot be computed when the number of leased lines is zero. To estimate the efficiency of OR excl LL it is necessary to make an arbitrary assumption about the value of this variable. We have adopted the standard approach in such circumstances of setting the number of leased lines to 1, so that ln(leased lines per switched line) is equal to -17.3. However, this is far removed from the lowest value of this variable for the LECs, which is -3.8. The consequence is that the efficiency of OR excl LL is estimated using a value for one variable that is a complete outlier with respect to the range used in the estimation. In these circumstances, minor mis-specification of the equation may lead to very large errors in the estimation of the relative efficiency of OR excl LL.

One possible solution is to adopt an alternative value for ln(leased lines per switched line) that is within or close to the range of observed values. For this exercise we have adopted a value of ln(leased lines per switched line) of -4.0, just below the minimum of the range in the LEC data. On this basis the relative efficiency of OR excl LL becomes -7.2%, i.e. the same as the value for OR incl LL. This illustrates the nature of the uncertainty about the true efficiency for OR excl LL caused by the necessity of adopting what must be an arbitrary assumption about the value of ln(leased lines per switched line).

An alternative approach is to avoid reliance upon the number of leased lines as a cost driver altogether. We have examined what happens if the cost frontier is estimated without including ln(leased lines per switched line) as an independent variable. The consequence is the relative efficiency for OR incl LL is -8.3% when compared with the upper decile, while the relative efficiency for OR excl LL is -9.2% - i.e. it is more efficient than OR incl LL. While the exact figures may not be reliable because of the exclusion of leased lines, nonetheless the result supports the indication that Openreach appears to be more efficient than the upper decile for the LECs.

4.4.2. Additive linear specification

As explained earlier we have used the standard multiplicative log-linear model for our cost frontier. This leads to the interpretation of the estimated coefficients as elasticities – i.e. the coefficient of 0.162 on the length of total sheath per switched line tells us that doubling the length of total sheath per switched line will, other things remaining unchanged, increase the total cost per switched line by just over 16%. One consequence of this specification is that the marginal cost of additional sheath falls as the amount of sheath increases. Similarly, the logarithmic specification exaggerates the marginal effects of changes in variables whose absolute value is much less unity, such as the proportion of fibre in the network or, in some cases, the number of leased lines per switched line.[20] This may or may not be plausible, but it is an assumption that is worth testing, especially because Openreach is an extreme outlier for certain variables – particularly leased lines per switched line for OR excl LL.

Hence, we have estimated the following linear additive version of the cost model:

$$
\begin{aligned}
&(\text{Total Cost/Switched Lines}) = a + b_1(\text{Stranded Ratio}) \\
&+b_2(\text{Leased Lines/Switched Lines})\text{*}(1\text{-Vdummy}) \\
&+b_3(\text{Leased Lines/Switched Lines})\text{*Vdummy} \\
&+b_4(\text{Total Sheath/Switched Lines}) + b_5(\text{Population Density}) \\
&+b_6(\text{Business Residential Ratio}) + b_7(\text{Fibre Proportion}) \\
&+b_8\,\text{NYNE} + b_9\,\text{MidWest} + b_{10}\,\text{Time}
\end{aligned}
$$

Apart from the move from the log-linear to the linear form of the equation the main change concerns the treatment of the leased lines per switched line variable to deal with the Verizon problem. In the linear specification we have two coefficients – β_2 and β_3 – for the numbers of leased lines per switched line when Vdummy=0 (most cases) and Vdummy=1 (for the relevant Verizon companies in the years 1999-2004). This is analogous to our treatment of the Verizon dummy in the log-linear specification.

Table 4.3 gives the results of estimating the linear model with and without the Stranded Ratio. Again, the inclusion of this variable greatly improves the performance of the model, so we concentrate on the efficiency results for the linear specification including the Stranded Ratio.

[20] Suppose that we have a cost driver X whose average value is 20% and we consider the effect of a change of 10% in its absolute value either from (a) 0.05 to 0.15 or (b) from 0.50 to 0.60. In case (a) the change in logs is +1.099, whereas in case (b) the change is 0.182, so the marginal effect of the absolute change of 10% in the value of X on total cost is more than 5 times higher in case (a) than in case (b).

Table 4.4 shows the efficiency ranking of the LECs, the two variants of Openreach and BT Network for this version of the linear specification. There is only a small difference in efficiency between the two variants of Openreach and they are ranked 5th and 6th in the full sample, just below the decile, with inefficiencies relative to the decile of -1.3% and -0.9%. Again, these results would seem to support the conclusion that Openreach's inefficiency is on or slightly better than the upper decile. As before, the caveat needs to be applied that we have not been able to derive Openreach and LEC entities that are completely comparable and hence conclusions regarding relative efficiency need to be interpreted with caution.

Table 4.3
Estimates of the cost frontier for alternative linear specifications

Dependent Variable	**Total Cost/Switched Line**			**Total Cost/Switched Line**		
	Coeff	**Std Err**	**Z-ratio**	**Coeff**	**Std Err**	**Z-ratio**
Independent Variables						
Stranded Ratio				0.174***	0.019	9.34
Leased Lines per Switched Line * (1-Vdummy)	0.012***	0.003	3.47	0.011***	0.003	3.42
Leased Lines per Switched Line * Vdummy	0.001	0.011	0.12	0.002	0.010	0.17
Total Sheath per Switched Lind	1.727***	0.196	8.81	0.828***	0.174	4.75
Population Density	0.000	0.000	-1.14	0.000	0.000	-0.22
Business-Residential Ratio	-0.012	0.030	-0.41	-0.050*	0.027	-1.81
Fibre Proportion	0.583***	0.135	4.31	0.353***	0.125	2.82
NYNE	0.074***	0.019	3.78	0.067***	0.018	3.78
MidWest	-0.071***	0.017	-4.26	-0.056***	0.015	-3.79
Time	0.015***	0.001	10.50	0.009***	0.001	6.54
Constant	-0.070	0.208	-0.33	-0.042	0.032	-1.31
Log-Likelihood	950.6			988.0		

*Source: NERA analysis. Note: Significance levels for coefficients: * = significant at 10% level, ** = significant at 5% level, *** = significant at 1% level*

.

Table 4.4
Efficiency ranking for the stranded assets linear specification

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
1	Qwest-North Dakota	1.4%	0.0%	-2.5%
2	Verizon SO-Illinois	1.8%	0.4%	-2.1%
3	Qwest-Idaho South	1.9%	0.5%	-2.0%
4	Qwest-Iowa	2.5%	1.1%	-1.5%
5	Openreach incl LL	2.7%	1.3%	-1.3%
6	Openreach excl LL	3.1%	1.7%	-0.9%
7	Wisconsin Bell	3.6%	2.2%	-0.4%
8	Verizon NE-Rhode Island	3.8%	2.3%	-0.2%
9	Qwest-Montana	4.0%	2.6%	0.0%
10	Contel-Nevada	4.1%	2.7%	0.1%
11	Verizon-New Jersey	4.1%	2.7%	0.1%
12	Verizon NO-Illinois	5.5%	4.1%	1.4%
13	BellSouth-Tennessee	5.6%	4.2%	1.6%
14	Qwest-Minnesota	5.9%	4.4%	1.8%
15	Verizon-Pennsylvania	5.9%	4.4%	1.8%
16	SBC/SNET-Connecticut	5.9%	4.4%	1.8%
17	BellSouth-Florida	6.2%	4.7%	2.1%
18	Verizon-Virginia	6.2%	4.7%	2.1%
19	Verizon NO-Wisconsin	6.6%	5.1%	2.5%
20	Verizon-Maryland	6.6%	5.2%	2.5%
21	Verizon-Delaware	6.7%	5.2%	2.6%
22	Verizon NE-Maine	6.8%	5.4%	2.7%
23	Qwest-South Dakota	7.4%	5.9%	3.3%
24	BellSouth-North Carolina	7.8%	6.4%	3.7%
25	Qwest-Utah	7.9%	6.5%	3.8%
26	Nevada Bell	8.1%	6.6%	3.9%
27	SBC-Kansas	8.3%	6.8%	4.1%
28	Pacific Bell-California	8.4%	6.9%	4.2%
29	BellSouth-Alabama	8.5%	7.1%	4.4%
30	Verizon SO-Virginia	8.7%	7.2%	4.5%
31	Qwest-Washington	9.1%	7.6%	4.9%
32	GTE-California	9.3%	7.8%	5.1%
33	Verizon NO-Ohio	9.3%	7.8%	5.1%
34	Qwest-New Mexico	9.3%	7.9%	5.1%
35	Qwest-Oregon	9.5%	8.0%	5.3%
36	Verizon NE-New Hampshire	9.6%	8.1%	5.4%
37	Indiana Bell	9.8%	8.3%	5.5%
38	Verizon NO-Michigan	10.1%	8.6%	5.9%
39	BellSouth-South Carolina	10.1%	8.7%	5.9%
40	SBC-Missouri	10.2%	8.7%	5.9%
41	BellSouth-Kentucky	10.3%	8.8%	6.1%
42	SBC-Oklahoma	10.4%	8.9%	6.2%
43	Verizon-Washington D.C.	10.6%	9.1%	6.4%
44	Verizon NE-Massachusetts	11.2%	9.7%	6.9%
45	Qwest-Arizona	11.4%	9.9%	7.1%
46	Verizon NW-West Coast California	11.6%	10.1%	7.3%
47	Verizon NW-Oregon	11.6%	10.1%	7.3%

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
48	Illinois Bell	11.6%	10.1%	7.3%
49	Verizon NE-Vermont	11.7%	10.2%	7.4%
50	Verizon NW-Idaho	11.8%	10.3%	7.5%
51	Verizon NW-Washington	12.1%	10.6%	7.8%
52	Verizon-West Virginia	12.2%	10.6%	7.8%
53	Ohio Bell	12.3%	10.8%	8.0%
54	BellSouth-Louisiana	12.6%	11.0%	8.2%
55	Verizon NO-Indiana	13.0%	11.5%	8.6%
56	BellSouth-Georgia	13.4%	11.9%	9.0%
57	Verizon SO-South Carolina	13.5%	11.9%	9.1%
58	Verizon-Florida	13.7%	12.1%	9.3%
59	Qwest-Colorado	13.8%	12.2%	9.4%
60	Verizon NO-Pennsylvania	13.9%	12.4%	9.6%
61	BellSouth-Mississippi	14.0%	12.4%	9.6%
62	Michigan Bell	14.2%	12.7%	9.8%
63	SBC-Arkansas	14.4%	12.8%	10.0%
64	Qwest-Wyoming	15.9%	14.3%	11.4%
65	Qwest-Nebraska	16.2%	14.6%	11.7%
66	Verizon SW-Texas	16.5%	15.0%	12.1%
67	Contel-Arizona	17.1%	15.5%	12.6%
68	SBC-Texas	17.7%	16.1%	13.2%
69	Verizon-New York Telephone	18.1%	16.5%	13.6%
70	Verizon SO-North Carolina	20.4%	18.8%	15.8%

Source: NERA analysis

We have tested how well the log-linear and linear specifications perform against each other. This can be done using the PE test devised by Mackinnon, White & Davidson.[21] This involves testing the significance of new variables in the log-linear and linear specifications. We denote predicted value of ln(cost per switched line) for the log-linear specification as $\widehat{\log c}$ and the predicted value of cost per switched line for the linear specification as $\hat{c}$. Then, the additional variables are:

- in the log-linear specification: $\hat{c} - \exp(\widehat{\log c})$; and
- in the linear specification: $\log(\hat{c}) - \widehat{\log c}$.

We test whether the coefficient on the additional variable – γ_{log} for the log-linear specification, γ_{lin} for the linear specification - in each specification is equal to zero. If $\gamma_{log} = 0$ is not rejected, the log-linear specification may be preferred, while if $\gamma_{lin} = 0$ is not rejected, the linear specification may be preferred. If both hypotheses are rejected, then it would seem that neither specification is satisfactory.

[21] See M. Verbeek – A Guide to Modern Econometrics (2nd Edition; New York: John Wiley & Sons; 2004), pages 61-62 & 67 or R. Davidson & J. G. Mackinnon – Estimation and Inference in Econometrics (New York: Oxford University Press; 1993), pages 506-507.

The test is not a very powerful one because it is quite possible that neither of the two hypotheses is rejected, in which case one is not able to use the results to decide between the two specifications. This turns out not to be a problem in our case.

- In the log-linear specification the coefficient on the variable $\hat{c} - \exp(\widehat{\log c})$ has a t-ratio of 0.20, so that the hypothesis that the coefficient is 0 is certainly not rejected.
- By contrast in the linear specification the coefficient on the variable $\log(\hat{c}) - \widehat{\log c}$ has a t-ratio of 6.14, so that the hypothesis that the coefficient is 0 is decisively rejected.

Hence, the log-linear specification should clearly be preferred over the linear specification. It is worth noting that after including the additional variable in the linear specification the estimated inefficiency relative to the decile of OR incl LL is -1.6% and of OR excl LL is 3.1%. Thus, even the modification of the linear specification does not alter the support that it offers for the conclusion that Openreach's efficiency is very close to the upper decile.

4.4.3. SFA analysis for 2006

The estimation results described above rely upon the information on each of the LECs over time – i.e. the panel dimension of pooled data for the years 1999-2006 – as well as the cross section of LECs. We have checked whether our conclusions about the relative efficiency of Openreach would be altered by focusing on the cross-section information for the year 2006 alone. This involves the estimation of the stochastic frontier model using the cross-section data for the 69 LECs for which we have data. The disadvantage of this specification is that it sacrifices the information that is available over time for each company. This loss may be offset by more accurate estimates of relative efficiency if the importance of different cost drivers has been changing over time, so that, for example, estimating the equation with data for the year 1999 would generate significantly different coefficients from those obtained when using data for the year 2006. Our tests do not suggest that this is the case, but they are not very powerful.

Estimation of the SFA model using data for 2006 in isolation turns out to be quite difficult or even impossible for some specifications. It is well established that the maximum likelihood estimation of the cross-section stochastic frontier may fail to converge or may generate sets of coefficients that do not satisfy the fundamental concavity property of cost functions – the two are usually closely linked. We have experimented with a range of different assumptions concerning the specification of the error term but we were unable to obtain sensible results for the analysis of full network costs for the LECs and BT Network. Some of the estimations were apparently successful but the results implied that there were almost no differences in efficiency between the companies – all fell within 1% of the upper decile. This pattern is so different from our previous results that it is not credible.

Table 4.5
Stochastic frontier coefficients for Openreach costs in 2006

Dependent Variable	Ln(Total Cost/Switched Line)		
	Coeff	**Std Err**	**Z-ratio**
Independent Variables			
Ln(Stranded Ratio)	0.394***	0.129	3.06
Ln(Leased Lines per Switched Line)	0.122***	0.041	2.98
Ln(Total Sheath per Switched Line)	0.198***	0.045	4.43
Ln(Population Density)	0.017	0.019	0.91
Ln(Business-Residential Ratio)	-0.099*	0.060	-1.64
Ln(Fibre Proportion)	0.051	0.066	0.77
NYNE	0.184***	0.051	3.64
MidWest	-0.031	0.048	-0.64
Constant	-0.453*	0.249	-1.82
Log-Likelihood	52.4		

Source: NERA analysis

*Note: Significance levels for coefficients: * = significant at 10% level, ** = significant at 5% level, *** = significant at 1% level. Note that coefficients for the Time and Vdummy variables cannot be estimated in this model.*

As shown in Table 4.6 we obtained apparently sensible results for the Openreach comparisons for 2006. However, there is one important point about these coefficients that should be noted. The elasticity of costs with respect to the ratio of leased lines to switched lines is much higher (0.122) than in the pooled model (0.023). The coefficient is barely significant at the 10% level in the pooled model whereas it is significant at the 1% level in the cross-section model. However, the standard error is also much higher in the cross-section model, as one would expect as a consequence of the reduction in the sample size, so that the confidence interval associated with the coefficient is rather wide, ranging from 0.04 to 0.20.

Table 4.6
Efficiency rankings for 2006 stochastic frontier coefficients

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
1	Openreach incl LL	0.9%	0.0%	-3.0%
2	Qwest-Iowa	3.1%	2.1%	-0.9%
3	Verizon-Virginia	3.3%	2.3%	-0.7%
4	Qwest-Idaho South	3.5%	2.6%	-0.5%
5	Verizon-New Jersey	3.6%	2.6%	-0.4%
6	Qwest-Minnesota	3.8%	2.8%	-0.2%
7	Verizon NO-Illinois	3.9%	3.0%	-0.1%
8	Verizon-Maryland	4.1%	3.1%	0.1%
9	SBC/SNET-Connecticut	4.5%	3.5%	0.5%
10	BellSouth-Tennessee	4.5%	3.6%	0.5%
11	Qwest-Montana	4.6%	3.7%	0.6%
12	Verizon NE-Maine	4.7%	3.7%	0.7%
13	Verizon NW-Washington	4.8%	3.8%	0.8%
14	Verizon NE-Rhode Island	4.8%	3.8%	0.8%

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
15	Qwest-North Dakota	5.1%	4.1%	1.1%
16	Nevada Bell	5.1%	4.1%	1.1%
17	Indiana Bell	5.2%	4.2%	1.2%
18	Verizon-Washington D.C.	5.5%	4.5%	1.4%
19	Verizon NO-Ohio	5.5%	4.6%	1.5%
20	Verizon-Delaware	5.6%	4.6%	1.5%
21	BellSouth-Alabama	5.7%	4.7%	1.6%
22	Qwest-Washington	5.8%	4.8%	1.7%
23	Wisconsin Bell	5.8%	4.8%	1.7%
24	Qwest-Utah	5.9%	4.9%	1.8%
25	Verizon NE-New Hampshire	6.0%	5.0%	1.9%
26	Verizon NO-Wisconsin	6.2%	5.2%	2.1%
27	SBC-Oklahoma	6.2%	5.2%	2.1%
28	Verizon-Pennsylvania	6.2%	5.3%	2.1%
29	Verizon SO-Illinois	6.3%	5.3%	2.2%
30	Contel-Nevada	6.4%	5.4%	2.3%
31	Qwest-New Mexico	6.4%	5.4%	2.3%
32	Verizon SO-Virginia	6.4%	5.5%	2.3%
33	Verizon NO-Michigan	6.5%	5.5%	2.4%
34	BellSouth-North Carolina	6.5%	5.5%	2.4%
35	Qwest-South Dakota	6.8%	5.8%	2.7%
36	SBC-Kansas	7.0%	6.0%	2.9%
37	Verizon NO-Pennsylvania	7.1%	6.1%	3.0%
38	Qwest-Colorado	7.2%	6.2%	3.1%
39	Qwest-Oregon	7.3%	6.3%	3.1%
40	Verizon NW-Idaho	7.4%	6.4%	3.3%
41	Pacific Bell-California	7.4%	6.4%	3.3%
42	GTE-California	7.6%	6.6%	3.4%
43	BellSouth-Florida	7.7%	6.7%	3.6%
44	Verizon-West Virginia	7.9%	6.9%	3.8%
45	Verizon NW-Oregon	8.0%	6.9%	3.8%
46	Verizon NE-Massachusetts	8.1%	7.0%	3.9%
47	Illinois Bell	8.2%	7.2%	4.0%
48	BellSouth-Kentucky	8.3%	7.3%	4.1%
49	SBC-Arkansas	8.4%	7.4%	4.2%
50	Verizon SO-North Carolina	8.4%	7.4%	4.2%
51	Verizon NE-Vermont	8.4%	7.4%	4.3%
52	Verizon NO-Indiana	8.5%	7.5%	4.3%
53	Qwest-Arizona	8.8%	7.8%	4.6%
54	SBC-Missouri	9.0%	7.9%	4.8%
55	Verizon NW-West Coast California	9.0%	8.0%	4.8%
56	BellSouth-Georgia	9.0%	8.0%	4.8%
57	Qwest-Nebraska	9.3%	8.3%	5.1%
58	BellSouth-South Carolina	9.4%	8.4%	5.2%
59	Ohio Bell	9.5%	8.5%	5.3%
60	Qwest-Wyoming	9.9%	8.9%	5.7%
61	Contel-Arizona	11.1%	10.0%	6.8%
62	Verizon SW-Texas	11.3%	10.2%	7.0%
63	Michigan Bell	11.6%	10.5%	7.3%
64	Verizon-Florida	12.0%	10.9%	7.7%

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
65	Verizon SO-South Carolina	12.4%	11.4%	8.1%
66	BellSouth-Louisiana	13.2%	12.1%	8.8%
67	BellSouth-Mississippi	14.5%	13.4%	10.1%
68	SBC-Texas	17.6%	16.5%	13.0%
69	Verizon-New York Telephone	20.0%	18.9%	15.4%
70	Openreach excl LL	255.1%	251.8%	241.4%

Source: NERA analysis

Table 4.6 above shows the efficiency rankings calculated using the results of the SFA analysis for the 2006 cross-section. This shows OR incl LL at the top with costs that are 3% less than the upper decile. However, the result for OR excl LL is clearly absurd with a cost that is more than 240% above the upper decile. As pointed out earlier, the fundamental difficulty in obtaining sensible results for Openreach lies in the weight that is given to leased lines in the cost frontier. The single year frontier has a much higher elasticity for leased lines than the pooled frontier. In the case of OR excl LL – for which the ratio of leased lines to switched lines is effectively zero (to avoid degeneracy in the calculation of logarithms it is necessary to assume that the number of leased lines is 1) – the procedure generates extreme results. In practice, the efficiency estimate for OR excl LL is simply an artefact of the way in which the data is adjusted so as to avoid the degeneracy caused by attempting to take the logarithm of zero. This is not so important when the coefficient on leased lines is much smaller, but it highlights the fact that there is a fundamental problem with using the OR excl LL data in a logarithmic specification.

4.4.4. Maximum likelihood ‘GLS’ analysis

As described in Section 2.4, NERA believes that a multi-year stochastic frontier analysis is the best measure of efficiency given the size and scope of our dataset. However, we can also run the same specification using an alternative model to test for the robustness of our results.

Generalised least squares (as described in Section 2.3.1) provides one such alternate model. GLS seeks to estimate an ‘average’ cost function rather than an efficiency frontier, and does not decompose the residual into measurement error and inefficiency as stochastic frontier analysis does. This limitation is shared by a random effects error components model, which is estimated using maximum likelihood iterative techniques in the same way as stochastic frontier analysis. Therefore, to examine the effect of moving to an average cost function and of not decomposing the residual, it is best to use this maximum likelihood method.[22] We have adopted a somewhat more general version of the GLS model by allowing for (a) heteroskedastic errors across panels – i.e. the random error is not constrained to have the same variance for each LEC, and (b) autocorrelation of errors within panels – i.e. the error for each company in year t+1 may be correlated with the error of the same company in year t. The autocorrelation parameter is 0.70, so that the errors for each company are quite highly correlated over time.

22 Where residuals are assumed normally distributed, a random effects error component model will give an identical regression equation to a generalised least squares model.

The GLS model generates a wider dispersion of estimates of company inefficiency relative to the decile than does the stochastic frontier analysis. OR incl LL emerges as the most efficient company with an inefficiency relative to the upper decile of -16.6%, while OR excl LL is ranked 16th with an inefficiency relative to the decile of 7.2%, though, as before, this is affected by the assumption made about the value of ln(leased lines per switched line). Again, the qualitative nature of these results appears to be broadly consistent with the results generated by our main model.

4.4.5. Other sensitivity tests

We have carried out a range of other sensitivity tests to determine whether they affect our conclusions about Openreach's inefficiency. We do not report them in detail here because the results are uniformly negative in the sense that they suggest that the broad conclusions are indeed robust. Hence, we will simply list the main tests that have been carried out.

- Alternative specifications for the Verizon leased lines problem as discussed earlier.
- Use of a standard cost of capital of 10% rather than 11.4% for all companies.
- Excluding either or both leased lines and the fibre proportion from the estimated model.
- Use of an alternative cost index for capital equipment.

There is one modification to the model which could be more important but which turns out to generate results that are difficult to interpret. The issue concerns the stability of the efficiency levels for each LEC over time. Our standard application of the stochastic frontier analysis assumes that efficiency is time invariant, i.e. that each company is equally efficient or inefficient in 2006 as in 1999. This validity of this assumption can be examined by estimating a model with a time trend in the relative efficiency of each LEC and testing whether the time trend is significantly different from zero. In the past we have found that the time trend was not significantly different from zero, so we have continued to use the time invariant assumption.

However, this time most of the tests suggest that the time trend in relative levels of efficiency is not zero. In fact, it seems that the relative efficiencies of the LECs are diverging quite rapidly. This is surprising and very difficult to interpret. Further, one consequence of adopting the time varying efficiency assumption is that the time trend in the cost frontier becomes highly negative at about -12% per year. It seems rather implausible that network costs should be falling at such a rate. Thus, we have concluded that the time varying model cannot be relied upon, at least without more evidence.

It should also be noted that using the time varying assumption increases the dispersion of LEC inefficiencies relative to the upper decile in 2006 but it does not substantially alter the ranking of Openreach relative to the LECs. For the log-linear specification OR incl LL is ranked 2nd with an inefficiency relative to the decile of -14.0%, while OR excl LL is ranked 17th with an inefficiency relative to the upper decile of 6.2%. Thus, even if one were to accept the time varying efficiency assumption, it would not alter our broad conclusion.

4.5. Quality of Service

We have examined whether the inclusion of indicators of quality of service has a significant effect on the estimates of frontier costs and on the efficiency rankings. While regulatory bodies collect a considerable amount of information on quality of service for the LECs and Openreach, there are only a small number of variables which are available on a consistent basis for all of the companies. These are:

§ The percentage of orders – both new orders and transfers – completed within the commitment time. The percentages are recorded separately for business and residential orders for the LECs but only the aggregate figure is available for Openreach, so we have used the aggregate figure. The average for Openreach is 90.6%, while the average for the LECs is 98.5% with a range from 90.5% to 99.9%. Thus, Openreach's performance falls right at the bottom end of the range for the LECs.

§ The average number of fault reports per 1000 switched lines. The number of fault reports is recorded separately for business and residential lines by both Openreach and the LECs. We found that we obtained more satisfactory results by using the average number of fault reports taken over all switched lines. For Openreach the number of fault reports per 1000 switched lines was 106 whereas the average for the LECs was 177.5 with a minimum of 65.5 and a maximum of 796. Thus, despite the apparent similarity of the definitions it it may be that reporting conventions differ for the LECs and Openreach.

§ The average number of working hours required to repair faults. These indicators are recorded separately for residential and business customers for the LECs and in an even more disaggregated manner by Openreach, but again we found that there was no advantage to including the disaggregated indicators rather than the average repair time for each company. The average for Openreach is 17.2 hours while the average for the LECs is 19.7 hours with a range from a minimum of 6.7 hours to a maximum of 102.7 hours.

The results of estimating the cost frontier with all three quality of service variables are shown in columns 2 to 4 of Table 4.7. The addition of all three variables improves the value of the log-likelihood from 531.4 to 553.4, so that jointly they are highly significant. Nonetheless, the coefficient on the average time to repair faults has a t-ratio of only 0.52, so that we re-estimated the model after dropping this variable. The estimates of the refined model including quality of service variables are shown in columns 5 to 7 of Table 4.7.

Table 4.7
Estimates of cost frontier with quality of service variables

Dependent Variable	Ln(Total Cost/Switched Line)			Ln(Total Cost/Switched Line)		
	Coeff	**Std Err**	**Z-ratio**	**Coeff**	**Std Err**	**Z-ratio**
Independent Variables						
Ln(Stranded Ratio)	0.479***	0.063	7.57	0.484***	0.063	7.71
Ln(Leased Lines per Switched Line)	0.007	0.013	0.54	0.006	0.013	0.51
Ln(Total Sheath per Switched Line)	0.107***	0.032	3.32	0.108***	0.032	3.37
Ln(Population Density)	-0.013	0.014	-0.93	-0.013	0.014	-0.92
Ln(Business-Residential Ratio)	0.069**	0.034	2.00	0.068**	0.034	1.98
Ln(Fibre Proportion)	0.064**	0.030	2.16	0.064**	0.030	2.16
Vdummy	-0.030	0.020	-1.47	-0.029	0.020	-1.43
NYNE	0.203***	0.040	5.04	0.201***	0.040	5.00
MidWest	-0.126***	0.037	-3.43	-0.125***	0.037	-3.41
Time	0.032***	0.004	7.41	0.031***	0.004	7.41
% of orders completed on time	-0.011***	0.004	-3.11	-0.011***	0.004	-3.08
Average faults per 1000 switched lines	0.001***	0.000	5.48	0.001***	0.000	5.54
Average time to repair faults	0.000	0.001	-0.52			
Constant	0.051	0.368	0.14	0.040	0.367	0.11
Log-Likelihood	553.4			553.2		

Source: NERA analysis

*Note: Significance levels for coefficients: * = significant at 10% level, ** = significant at 5% level, *** = significant at 1% level.*

The signs of the coefficients are not entirely intuitive. It seems that achieving a higher percentage of orders completed on time reduces the efficient level of costs, whereas one might have expected the opposite – i.e. that extra resources are required in order to achieve a high rate of completion on time in face of a random flow of orders. Similarly, a higher average number of fault reports per 1000 switched lines is associated with higher costs. This suggests that fault reports are an exogenous cost driver – perhaps reflecting the state of the network – rather than the outcome of choices about maintenance made by the operator. Thus, it is not clear that the quality of service variables produce results that correspond to prior expectations.

Table 4.8
Efficiency rankings with quality of service variables

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
1	Qwest-Idaho South	4.0%	0.0%	-8.8%
2	Openreach incl LL	4.5%	0.1%	-8.3%
3	Openreach excl LL	4.7%	0.1%	-8.2%
4	Contel-Nevada	8.7%	4.6%	-4.6%
5	Qwest-Iowa	9.1%	4.9%	-4.3%
6	Wisconsin Bell	9.4%	5.2%	-4.0%
7	BellSouth-Tennessee	13.7%	9.3%	-0.3%
8	Verizon NE-Rhode Island	13.8%	9.4%	-0.2%
9	Qwest-Montana	14.0%	9.6%	0.0%
10	Qwest-North Dakota	14.1%	9.7%	0.1%
11	BellSouth-Florida	17.1%	12.6%	2.7%
12	Verizon NO-Illinois	17.4%	12.9%	3.0%
13	SBC/SNET-Connecticut	17.5%	13.0%	3.1%
14	Verizon SO-Illinois	18.8%	14.2%	4.2%
15	BellSouth-North Carolina	18.8%	14.3%	4.3%
16	Verizon-Virginia	19.0%	14.5%	4.4%
17	Verizon-Delaware	19.9%	15.3%	5.2%
18	BellSouth-Alabama	19.9%	15.3%	5.2%
19	Indiana Bell	20.3%	15.7%	5.6%
20	Qwest-Minnesota	20.5%	15.9%	5.7%
21	Verizon NE-Maine	21.4%	16.8%	6.5%
22	Verizon-Maryland	22.6%	17.9%	7.6%
23	Verizon-New Jersey	24.1%	19.3%	8.9%
24	Verizon-Washington D.C.	24.1%	19.4%	8.9%
25	BellSouth-Kentucky	24.7%	19.9%	9.4%
26	BellSouth-South Carolina	24.9%	20.1%	9.5%
27	Qwest-Utah	25.0%	20.2%	9.6%
28	Verizon NO-Ohio	25.4%	20.6%	10.0%
29	Pacific Bell-California	25.5%	20.7%	10.1%
30	SBC-Kansas	25.5%	20.7%	10.1%
31	Qwest-South Dakota	25.9%	21.1%	10.4%
32	Verizon SO-Virginia	26.0%	21.2%	10.5%
33	Verizon NO-Wisconsin	26.4%	21.5%	10.9%
34	Qwest-New Mexico	26.5%	21.7%	11.0%
35	BellSouth-Louisiana	26.9%	22.0%	11.3%
36	Verizon-Pennsylvania	27.0%	22.2%	11.4%
37	SBC-Oklahoma	27.4%	22.5%	11.8%
38	Nevada Bell	27.5%	22.6%	11.9%
39	BellSouth-Mississippi	27.7%	22.8%	12.0%
40	Illinois Bell	28.2%	23.3%	12.5%
41	Verizon NE-New Hampshire	28.3%	23.4%	12.6%
42	Verizon NO-Indiana	29.0%	24.0%	13.2%
43	Verizon SO-South Carolina	29.3%	24.4%	13.5%
44	Verizon NE-Vermont	30.4%	25.4%	14.4%
45	SBC-Missouri	30.4%	25.4%	14.4%
46	GTE-California	30.5%	25.5%	14.5%
47	Qwest-Washington	30.6%	25.6%	14.6%

Rank	Company	Estimated inefficiency (%) relative to:		
		Frontier	Best	Decile
48	Qwest-Oregon	31.4%	26.4%	15.3%
49	Ohio Bell	32.4%	27.3%	16.1%
50	BellSouth-Georgia	32.6%	27.5%	16.3%
51	Michigan Bell	32.8%	27.7%	16.5%
52	Verizon NE-Massachusetts	33.3%	28.2%	16.9%
53	Qwest-Arizona	33.8%	28.7%	17.4%
54	Verizon NO-Michigan	34.5%	29.4%	18.0%
55	Verizon NW-West Coast California	35.4%	30.2%	18.8%
56	Verizon NW-Idaho	35.6%	30.4%	19.0%
57	Verizon NW-Oregon	36.0%	30.8%	19.3%
58	Verizon-West Virginia	37.1%	31.9%	20.3%
59	SBC-Arkansas	37.8%	32.5%	20.9%
60	Verizon NO-Pennsylvania	41.4%	36.0%	24.1%
61	Verizon NW-Washington	41.4%	36.0%	24.1%
62	Verizon-Florida	42.3%	36.9%	24.9%
63	Qwest-Colorado	42.8%	37.3%	25.2%
64	Verizon SW-Texas	42.9%	37.4%	25.4%
65	Qwest-Wyoming	45.4%	39.9%	27.6%
66	SBC-Texas	47.1%	41.5%	29.1%
67	Qwest-Nebraska	48.0%	42.4%	29.9%
68	Verizon SO-North Carolina	52.7%	46.8%	33.9%
69	Verizon-New York Telephone	52.8%	46.9%	34.0%
70	Contel-Arizona	53.1%	47.2%	34.3%

Source: NERA analysis

Table 4.8 shows the results of the efficiency ranking for the refined model including quality of service variables. The two Openreach variants are ranked 2nd and 3rd with inefficiencies of -8.3% and -8.2% relative to the decile.[23] Thus, including the quality of service indicators in the cost frontier appears to improve Openreach's performance relative to that of the LECs. However, the fact that the quality of service variables have signs that are the opposite of a priori expectations means that this result needs to be treated with some scepticism.

4.6. Data Envelopment Analysis

In addition to the stochastic frontier and generalised least squares analysis outlined above, NERA has also attempted to measure the efficiency of both versions of Openreach using a mathematical technique, data envelopment analysis (DEA). DEA uses a mathematical programme, rather than statistical techniques to assess comparative efficiency. We describe this technique further in Appendix B. DEA operates by searching for a 'least cost peer group' of comparator companies for Openreach. The 'peer group' is defined such that a linear combination of these companies can be shown to have at least as great an output and

---

[23] When we include QoS variables in the efficiency models the relative efficiency of Openreach excl LL changes significantly. This is because the coefficient in these models for log(leased lines/switched lines) is significantly lower than in the models which exclude QoS variables. The reduced importance of this coefficient in determining the efficient level of costs means that the extreme value for this variable taken by Openreach excl LL (the model specification means that we are forced to arbitrarily assume value for leased lines which is not zero) has a much reduced impact on relative efficiency.

no more favourable operating conditions than Openreach. If such a 'peer group' exists, and the linear combination of their costs is lower than that of Openreach, this cost difference is assumed to be attributable to inefficiency on the part of Openreach.

It should be noted that, since it is a mathematical technique, DEA offers no statistical framework for modelling the performance of firms outside the sample, nor does it allow us to assess the significance of the variables used in the analysis. Therefore, DEA analysis was completed using the same explanatory variables as the regression analysis, as the statistical validity of including these variables in a total costs model has already been established.

Using these variables, DEA is unable to find peers to compare against the total costs of either version of Openreach in 2006. It is difficult to draw firm conclusions from this because there are two possible explanations for why DEA cannot find peers for a company:

- Firstly, it could be that the company is 100% efficient.
- The second possibility is that DEA is unable find comparable firms within the sample against which to assess the efficiency of the firm.

In a previous efficiency study of BT Network we concluded in the light of the regression results that the latter possibility was the most likely. Despite the fact that the regression results obtained suggest that Openreach and BT Network are relatively efficient, subject to the proviso that it has not proved possible to find true comparators with which to compare Openreach, it would be difficult to argue that the DEA result is strong evidence that the two Openreach pseudo-companies are 100% efficient. What we can say is that the DEA results are consistent with the results that we obtained from the regression analysis.

4.7. Comparisons with Previous Studies

The SFA results suggest that Openreach and the BT Network comparator are slightly more efficient than the decile although the exact margin is uncertain. In the 2005 study BT Network was found to be slightly less efficient than the decile for the year 2003. As we explain below there are no significant methodological changes between this study and the 2005 study that would provide a systematic explanation for the apparent efficiency improvement of BT Network. Therefore, the results support the conclusion that there has been an improvement in the relative efficiency of BT Network operations over the period 2003 to 2006. In this context, we note that:

- The SFA methodology, with time invariant inefficiency, for assessing the relative efficiency of Openreach is the same as we used in 2005. We find that an efficiency model that allows for time varying efficiency is statistically significant but implies that the overall rate of efficiency growth is implausibly high, although this time varying efficiency model produces relative efficiency result that are consistent with those reported above.

- We have made only a small number of changes to how we collect and process the LEC and BT Network data compared with the 2005 study.

Turning to our attempts to measure the comparative efficiency of Openreach, because we cannot identify separately the costs associated with the provision of leased lines, we constructed two Openreach pseudo-companies to compare with the LECs – one including leased line costs and one excluding these costs. Whilst there is a gap between the relative efficiency ranking of these two pseudo-companies, the overall body of evidence – which includes a BT Network relative efficiency ranking – suggests that the Openreach pseudo companies may be slightly more efficient than the upper decile. However, neither of these Openreach pseudo companies is completely representative of Openreach WLR and LLU, the target of the comparative efficiency assessment.

We exclude data from the years 1996-1998. In earlier versions of the efficiency models we included data from these years and find that the general results are consistent with those reported above.

We include quality of service variables in some of our efficiency models. The inclusion of these variables does not change the conclusion that Openreach is below the upper decile. In fact when these variables are included in the efficiency models the relative efficiency of Openreach and the BT Network comparator improve. However, as noted above, the counter intuitive signs on the coefficients of the quality of service variables mean that little weight should be attached to this finding.

The model specification is slightly different to that used in 2005 and we find that this specification gives a better statistical fit to the data. However, results from our initial analysis using a specification very similar to that used in 2005 are consistent with our overall findings.

For the record we should note that we have estimated the relative efficiency of BT Network relative to LEC network costs for various models that were used in testing the sensitivity of our conclusions to alternative specifications. The key results are:

§ For our basic log linear model the relative efficiency of BT Network Compared to the decile is -3.8% (i.e. it is more efficient).

§ Using the linear model the relative inefficiency of BT Network is +0.8%, marginally worse than the upper decile.

§ Finally, for the standard model with Quality of Service variables added the efficiency of BT Network relative to the upper decile is -4.5%.

These results support the general conclusion that, at the level of all network costs, BT is more efficient than the upper decile for the LECs and that there has been an improvement in BT's performance since 2003.

5. Lessons for Future Efficiency Studies

We have found that making a comparison between Openreach and the US LECs to be problematic because of the way in which the LEC data is reported. In particular, it is impossible to determine how costs from a number of the cost categories reported by the LECs should be divided between core network and access network activities. Consequently we are forced to make a comparison between Openreach and the LECs that is less direct than has been the case with previous efficiency studies which have focused on the whole of BT network.

To try and overcome the problem of comparability we assess the efficiency of two Openreach 'pseudo-companies' relative to the LECs. The first pseudo-company closely resembles Openreach (OR excl LL), whilst the second psuedo-company adds in the costs of the provision of leased lines (excluding transmission equipment), which are taken from BT's regulatory accounts, to create an entity that more closely resembles the LECs after switching and transmission activities (OR incl LL). Our SFA analysis finds that there is a substantial gap between the assessed efficiency of the two pseudo-companies relative to the LECs. In the case of OR excl LL there are technical reasons for treating the results with considerable caution. At the same time OR incl LL is not a close proxy to Openreach WLR plus LLU. Consequently, neither of the results for the two pseudo companies is likely to provide a reliable indication of the relative efficiency of Openreach WLR plus LLU.

The results of this study therefore suggest that the comparison of the costs of Openreach to those of the US LECs (after excluding identifiable core network costs) is less reliable than the comparison of BT Network's costs with the total network costs of the LECs. On the other hand, the latter measures the relative efficiency of all of BT's network activities and may not therefore provide a guide to the relative efficiency of Openreach WLR plus LLU.

Future studies of this type are likely to continue to produce uncertain estimates of the relative efficiency of Openreach, although some useful information may be provided by estimated relative efficiency at different points in time. Furthermore, it may be that similar data comparability problems will arise as part of any prospective future study that would attempt to use the US LEC data to assess the relative efficiency of the relevant network activities of BT.

6. Summary of Results and Conclusions

This section outlines our results and conclusions regarding the comparative efficiency of Openreach using the cost equations estimated above. Table 6.1 summarises our SFA efficiency results for both Openreach variants and the BT Network cross-check.

Table 6.1
Summary of SFA Results

	Position Relative to Decile		
	CRS Log linear stranded assets model	**Linear additive specification**	**CRS log linear stranded assets model including QoS variables**
Openreach incl LL	-7.20%	-1.50%	-8.30%
Openreach excl LL	6.80%	-1.3%	-8.20%
BT Network	-3.8%	0.8%	-4.5%

NOTE: BT Network SFA efficiency estimates obtained using efficiency models including costs and cost drivers which are appropriate to standard network activities rather than just the access network.

The constant returns to scale log-linear model suggests that Openreach incl LL is 7.2% **more** efficient than the upper decile of the US LECs, while Openreach excl LL is 6.8% **less** efficient than the upper decile. Ideally the results would be closer together. However, the difficulty of constructing an appropriate basis for comparing the US LEC cost data with that of Openreach means that this is not possible. Our cross-check of the Openreach efficiency results, whereby we assess the comparative efficiency of BT Network as a whole suggests that BT Network is 3.8% more efficient than the upper decile.

Overall we can conclude that the efficiency of Openreach lies somewhere within the range of -7.2% and +6.8% relative to the upper decile, the mid-point of this range being -0.2%. We have no basis for deciding exactly where Oprenreach lies relative to the upper decile. Both comparisons are associated with difficulties and cannot be regarded as completely reliable.

When we add quality of service variables to our efficiency models we find that the efficiency of Openreach improves relative to the US LECs. However, the two QoS variables that are significant in our model have coefficients that are not consistent with prior expectations. This may indicate that these variables are picking up the impact of some exogenous cost drivers, for example the condition of the network, not otherwise captured by our models.

Appendix A. Stochastic Frontier Analysis

There is an extensive academic literature on efficiency measurement using Stochastic Frontier Analysis (SFA), and it is increasingly being used by utility regulators as a robust measure of inefficiency. It is based on regression analysis, but with a key distinction. In the context of efficiency comparisons, OLS regression models implicitly assume that the whole of the residual for a particular observation (company) is the result of genuine inefficiency. In contrast, SFA models incorporate the possibility that some component of the model residual may result from errors in measurement of costs or the omission of explanatory variables, as opposed to the existence of genuine inefficiencies. This decomposition of residuals between ‘error’ and ‘genuine inefficiency’ is intended to provide a more accurate reflection of the true level of inefficiency.

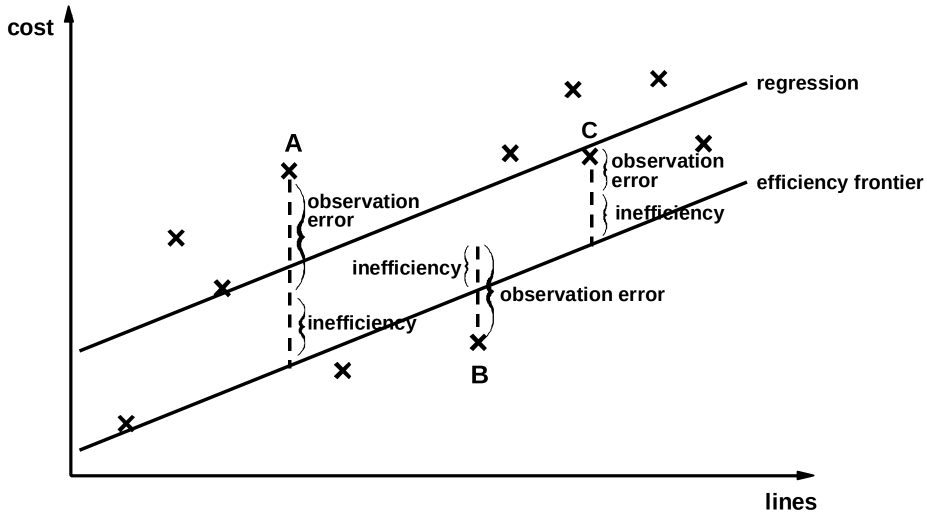
Under the simplest implementation of SFA, an Ordinary Least Squares (OLS) regression is first estimated to capture the sensitivity of costs to differences in operating environments, scale of production or mix of outputs. This model will leave residuals that can be attributed to the two remaining causes of measured cost differences:

- those that are the result of remaining incompatibilities in data that have not been picked up in the preceding adjustments, or even just plain errors in the observed data;
- genuine inefficiencies of the operators compared to best practice, in which we are primarily interested.

It is assumed that remaining data errors and incompatibilities are equally likely to increase or reduce measured costs for any one operator. However, inefficiencies, where they exist, will only increase costs. Therefore, to the extent that there are genuine operator inefficiencies, the distribution of the residuals from the regression model will be skewed in the direction of being positive (actual cost exceeds that estimated by the regression). By measuring the extent of this skew to the distribution (estimated by the third moment of distribution) it is possible to calculate the relative importance of observation errors vs. genuine inefficiencies.

Once the relative importance of the genuine inefficiencies in the residual distribution has been estimated, it is possible to adjust the estimated regression line downward by this amount to form an efficiency frontier. Thus, distances of each observation from the original OLS regression line measure departures of cost from “mean” industry performance, whilst in the new line they measure departures from the efficiency frontier. Figure A.1 provides an illustration of this. We have used the simplified example of just one explanatory variable in the regression model – the number of lines.

Figure A.1
Stochastic Frontier Analysis



Once an efficiency frontier has been estimated, it is still necessary to estimate the efficiency of each company, bearing in mind that the vertical distance of the measured cost of each company from the efficiency frontier will still be composed of two components:

§ observation error and data incompatibilities for that particular company;

§ genuine inefficiency for that particular company.

Indeed, it will still be possible for a particular company to have a cost below that of the efficiency frontier if the downward bias due to observation error exceeds the company's inefficiency (as is the case for observation B in Figure A.1). However, knowledge of the company’s cost relative to the efficiency frontier does provide information on which to base an estimate of the company’s inefficiency. For example, if the company’s cost is high relative to the efficiency frontier (as in the case of observation A), its inefficiency is also likely to be high, whereas if its cost is low relative to the efficiency frontier, its inefficiency is likely to be low (as in the case of observation C).

It is also possible to estimate the regression line taking account of the distribution of the residuals, using a maximum likelihood estimator. This will produce a more statistically efficient estimate than that given by OLS, and is the technique used in the analysis contained in this report. This method involves analysing a number of statistical distributions, to see if they will fit the distribution of the ‘inefficiency’ component of the residual error (which, as we have discussed above, is skewed in the direction of being positive). Three different distributions are commonly considered:

§ the half normal distribution;

§ the negative exponential distribution; and

§ the truncated normal distribution.

Essentially, SFA attempts to estimate a maximum likelihood model using each of these statistical distributions. If a distribution is found not to fit the data, the model is abandoned. However, if a distribution fits the data, the model produces estimates of the position of each observation (company) relative to the theoretical efficiency frontier. The model also yields estimates of the average distance of all observations from the efficiency frontier, as well as the proportion of the total residual error which is characterised as ‘inefficiency’ (as opposed to ‘measurement error’).

The main disadvantage of SFA is that strong assumptions are required in order to decompose the regression residuals, and in many practical cases insufficient data is available to support this decomposition. That is, there is often little evidence to suggest which of the above statistical distributions is most appropriate in constructing a model (note that, in many cases, more than one distribution may be deemed to ‘fit’ the data).

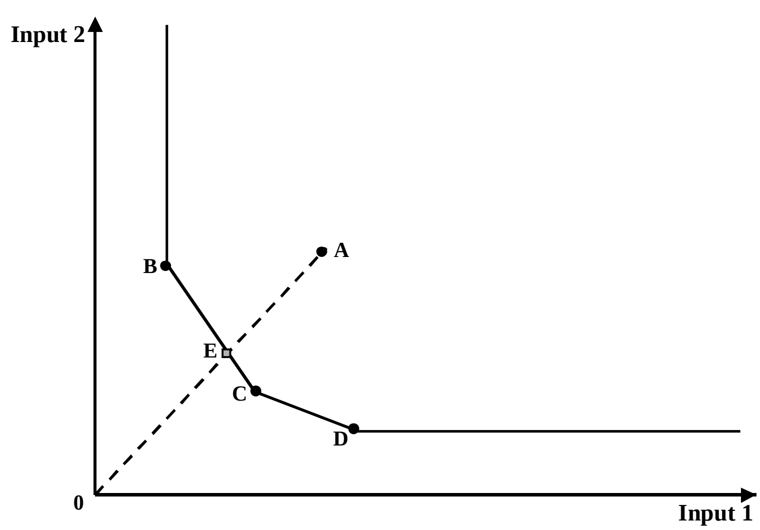
Appendix B. Data Envelopment Analysis

Data Envelopment Analysis (DEA) can be used as an alternative to regression-based techniques. It is not a statistical estimation technique, but rather makes use of mathematical programming methods, without the need to rely on a precise parametric cost function. This, in fact, is its main advantage; it allows a complex non-linear (concave or convex) relationship to exist between outputs and costs, whereas regression usually restricts such relationships to be either linear, or to have fairly simple non-linear forms. One additional important advantage of DEA is that it does not suffer from statistical problems, such as multicollinearity between variables, which may affect OLS, panel data and SFA models.

DEA operates by searching for a ‘least cost peer group’ of comparator companies for each individual target company. The ‘peer group’ is defined such that a linear combination of these companies can be shown to have at least as great an output and no more favourable operating conditions as the target company (with output and environmental variables measured in the same way as in regression analysis). If such a ‘peer group’ exists, and the linear combination of their costs is lower than that of the target company, this cost difference is assumed to be attributable to inefficiency on the part of the target company.

This is explained more clearly with the aid of a diagram. Figure B.1 below shows an example of an input minimisation model. Points A, B, C and D represent different companies, such that company A produces the same amount or less of each (common) output than the other three. Now, DEA assumes that, since B and C are technically feasible positions, point E (a linear combination of the two) is also technically feasible. Point E lies on a radial to A from the origin, and hence is a point using proportionately less of every output than A (a ‘radial contraction’), in order to produce the same (or greater) amount of each output. Therefore, point A is considered to be inefficient. The degree of inefficiency is reflected in the Farrell measure, which is defined as the ratio 0E/0A, and companies B and C form a ‘peer group’ for company A.

Figure B.1
DEA: Input Minimisation



It should be noted that DEA has a number of disadvantages:

- it provides no statistical framework to model the possibility of errors in the observations;
- it has a tendency to give observations the 'benefit of the doubt'. In arriving at an efficiency score for an observation, DEA adopts a weighting structure which emphasises the particular inputs and outputs which show that observation in the best possible light;
- related to the last point, unless restrictive assumptions are imposed (such as constant returns to scale), DEA is unable to assess the efficiency of companies with outputs or environmental variables on the "edges" of the data set. For example, if a company operates in an environment that is more adverse than any other in the sample, DEA will always assume that this company is 100% efficient, regardless of its cost base (since it has no other equivalent comparator). Regression analysis, on the other hand, by assuming a particular functional form for the impact of the environmental variable, is able to estimate the impact that the environmental variable should have on the company's costs, and calculate the company's underlying efficiency on this basis.

NERA
Economic Consulting

NERA Economic Consulting
15 Stratford Place
London W1C 1BE
United Kingdom
Tel: +44 20 7659 8500
Fax: +44 20 7659 8501
www.nera.com

NERA UK Limited, registered in England and Wales, No 3974527
Registered Office: 15 Stratford Place, London W1C 1BE

MMC Marsh & McLennan Companies