BT EFFICIENCY: NETWORK STUDY 2003

A Report for Ofcom

Prepared by NERA

May 2004
London

Project Team:
Nigel Attenborough
Gordon Hughes
Becky Stuttard
Lisa Correa
Tim Miller

n/e/r/a
National Economic Research Associates
Economic Consultants

15 Stratford Place
London W1C 1BE
Tel: (+44) 20 7659 8500
Fax: (+44) 20 7659 8501
Web: http://www.nera.com

An MMC Company

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

This report compares the cost efficiency of BT's network with that of the US LECs. The US LECs were chosen as comparators because a significant amount of detailed cost data is available for these operators, and because the better-performing LECs are generally regarded as providing the international benchmark for efficiency.

The study uses data for the US LECs for the years 1996 to 2001[1] to model the determinants of total network costs. Based on this model, the study then makes use of accounting and other data produced by BT, to assess BT's comparative efficiency in 2000/01 to 2001/02.

There are a variety of statistical and mathematical programming methods that can be used to assess comparative efficiency, each of which has strengths and weaknesses. As a result, NERA considers it appropriate not to rely on any one technique when making such comparisons. If more than one technique is used, the different results can then be reviewed in the light of the relative strengths and weaknesses of the different methods in order to provide a more informed view of comparative efficiency.

The techniques used in this study are:

- Multi-year SFA (estimated using LEC data for 1996-2001);
- Multi-year least squares (estimated using LEC data 1996-2001);
- Single-year OLS (estimated using LEC data for 2001);
- Single-year SFA (estimated using LEC data for 2001); and
- Single-year DEA (estimated using LEC and BT data for 2001[2]).

This report expands and develops the comparative efficiency analysis which NERA has previously carried out for Oftel. The previous comparative efficiency study of BT, completed in 2000, estimated that BT's network activities in 1999/00 were in the region of 3% less efficient than those of the best performing US operators.

The results of the comparative efficiency analysis completed during this study are presented in the table below. These results suggest that BT has become more inefficient since 1999/00.

---

1 Extensive data on the US LECs is published by the Federal Communications Commission (FCC).

2 BT data is not included in the regression analysis (SFA, Multi-year least squares and OLS) because of the possibility of this introducing bias into the results. BT's efficiency result is then estimated using the regression coefficients produced by the LEC data. Excluding BT from the DEA analysis is not possible, nor does the same problem of potential bias exist for this mathematical programming technique.

Table 1
Summary of BT's Comparative Efficiency

Point of Comparison	Full Sample: Best Performer	Full Sample: Upper Decile	Large LECs only: Best Performer	Large LECs only: Upper Decile
Multi-year SFA	n.a.	n.a.	16%	9%
Multi-year least squares	n.a.	n.a.	16%	10%
Single-year OLS	20%	9%	7%	3%
Single-year SFA - Exponential	1%	1%	n.a.	n.a.
Single-year DEA (compared against peer group)	17%	n.a.	13%	n.a.
Range of All Results	1%-20%	1%-9%	7%-16%	3%-10%
Range Excluding Single-year Regressions	**17%**	**n.a.**	**13%-16%**	**9%-10%**

Source: NERA Analysis

Analysis was completed, where possible, using both the full sample of US LECs and using the large LECs only (those with over 1 million exchange lines). It was not possible to estimate the multi-year regressions with the full sample of US LECs as tests identified that, over the period 1996 to 2001 as a whole, there were structural differences between the total cost functions of the large and small LECs. However, comparable tests using data just for 2001 indicated that these structural breaks were not present in that year. Therefore it was possible to estimate the single-year regressions using both the full and large LEC only samples.

There are no results for the single-year SFA (large LECs only) model as there was insufficient data for this technique to estimate inefficiency scores for this regression. Indeed, there were significant problems in obtaining inefficiency scores for the single-year SFA (full LEC sample) regression. Of the three distributional assumptions that can be used to estimate single-year SFA regressions, it was only the exponential distribution which was able to identify inefficiency scores.

The results of all the single-year techniques should be treated with some caution as these regressions were estimated using relatively small samples. The single-year OLS (large LECs only) is based on a sample size of 29 and the single-year SFA and single-year OLS (full LEC sample) are based on a sample size of 46. In contrast, the multi-year regressions are both based on a sample of 174. Hence, the results of the single-year regressions should be considered to be less reliable than those estimated using the multi-year techniques.

Given the greater robustness of the results of the multi-year analysis, and the broad consistency between the results of the multi-year analysis and the DEA analysis using the large LECs only sample, it was concluded that BT's inefficiency is in the region of 9% to 10%

compared to the upper decile point of the large LECs and of 13% to 16% compared to the best performing large LECs[3].

The results discussed above are those estimated using the NERA model specification. BT suggested an alternative model specification, which yields a much higher value for the log-likelihood function, implying that it fits the data better. Additionally, the BT model produces significantly different efficiency results. Using the multi-year SFA technique BT's inefficiency compared to the upper decile is 1% under the BT model compared to 9% under the NERA model.

NERA assessed the suitability of the two alternative model specifications using a number of criteria (see section 5.4 for a discussion of this analysis). The aim of this assessment was to identify the model which is most appropriate for assessing BT's comparative efficiency for the purposes of Ofcom's price cap modelling. The assessment indicated that the NERA model is more appropriate for assessing BT's comparative efficiency for the purposes of deriving an efficiency factor for inclusion in a price cap model. The main reasons for this are that the NERA model is more consistent with BT's regulatory cost base, the results of the NERA model are consistent with analysis of the relative movements over time in the costs and outputs of BT and the US LECs, and the NERA model allows for a comparison of how BT's efficiency has changed since the previous comparative efficiency study completed in 2000.

The results of the NERA model imply a significant deterioration in the performance of BT since 1999/00. To assess the validity of this conclusion we investigated how the change in BT's costs relative to output compared with that for the US LECs. The table below shows how BT's network costs, total lines and call minutes changed between 1999/00 and 2001/02 and compares these changes to the average change for the US LECs.

Table 2
Analysis of Cost and Output Changes 1999/00 to 2001/02

	Percentage change in network costs	Percentage change in total line numbers	Percentage change in call minutes
BT	19%	2%	40%
Average of US LECs	0%	11%[4]	-1%

Source: NERA analysis

It can be seen that, while BT's network costs increased by 19% over this period, the LECs' costs, on average, remained more or less constant. Whilst BT's call minutes increased

[3] The results of the single-year OLS are not included in the range as the very limited sample size used to estimate this regression may have limited the robustness of these results.

[4] A significant driver of the increase in the number of total lines for the US LECs is the growth in the number of leased lines.

significantly faster than the LECs' call minutes between 1999/00 and 2001/02, its line numbers increased more slowly than those of the LECs. If we assume that lines account for 70% of total network costs and calls for 30% (which was the case for BT in 2001/2), then, between 1999/00 and 2001/02, BT experienced a 19% increase in costs in the face of a 13.4% increase in output while the LECs experienced no increase in costs while their output grew by 7.4%. In other words, BT costs grew faster than its output while the opposite was true for the LECs.

These figures support the view that BT has become less efficient since the previous comparative efficiency assessment carried out in 2000 (using data for the financial year 1999/00). In fact they suggest that BT's inefficiency could have increased by something in the region of 13 percentage points since 1999/00. The difference between BT's inefficiency in the previous study (approximately 3%) and the range for BT's inefficiency identified by the NERA model in this study, of between 9% and 10% relative to the upper deciles, is around 7 percentage points. The conclusion therefore is that the efficiency decline implied by the NERA model is broadly consistent with the pattern of changes in costs and output volumes over time.

To assess further the validity of this finding, the percentage change in BT's output(s) over the period, along with the coefficients on these variables in the final NERA regression, were used to estimate more formally the change in network costs that would be expected as a result of the change in outputs.

Given the functional form of the regression model estimated, each of the regression coefficients can be interpreted as the expected percentage change in costs that would occur if the variable concerned were to change by 1%. For example, the coefficient on total lines, which was estimated as 0.81 in the multi-year SFA regression, indicates that if total lines were to increase by 1%, total costs would be expected to increase by 0.81%. The coefficients on switch minutes, of 0.15, and all other variables can be treated in a similar way.

This analysis indicated that the changes in BT's outputs over the period 1999/00 to 2001/02 would suggest a 9% increase in BT's total network costs. As shown above, BT's actual network costs increased by 19% between 1999/00 and 2001/02. Therefore, the results of this more formal estimation of the expected changed in BT's network costs supports the conclusion that BT's efficiency has deteriorated over the period 1999/00 to 2001/02.

The conclusions reached above suggest that BT's inefficiency could have increased by something in the region of 10 to 13 percentage points since 1999/00. This magnitude of efficiency change is consistent with BT's inefficiency score increasing from 3% in 1999/00 to between 9% and 10%, when compared to the upper decile point of the large LECs, and between 13% and 16% when compared to the best performing large LECs.

Reflecting this, it is reasonable to conclude that, for the purposes of Ofcom's price cap modelling, BTs inefficiency score in 2001 is:

- between 9% and 10% when compared with the upper decile point of the large US LECs, and
- between 13% and 16% based on the more demanding comparison against the best performance (i.e. the most efficient large US LEC).

1. INTRODUCTION

1.1. Background

This report, prepared for Ofcom by NERA, examines the comparative efficiency of BT. This report expands upon the comparative efficiency analysis which has previously been undertaken by NERA for Oftel. This report includes an expanded data set for the US Local Exchange Carriers (LECs) (covering the period 1996 to 2001) and extends the analysis of BT for the period 2000/01 to 2001/02.

The report compares the cost efficiency of BT's network with that of the US LECs. The US LECs were chosen as comparators because a significant amount of detailed cost data is available for these operators and because the better-performing LECs are generally regarded as providing the international benchmark for efficiency.

The study uses data for the US LECs for the years 1996 to 2001[5] to model total network costs, and then makes use of accounting and other data produced by BT, to assess BT's comparative efficiency under this model in the period 2000/01 to 2001/02.

Our analysis excludes mobile networks, cable TV networks and value added service activities of BT and the comparator companies. The focus of this study is purely on the core PSTN, access and leased line operations of the companies.

1.2. Report Structure

This report is structured as follows:

- Section 2 discusses the different techniques that can be used to measure comparative efficiency;
- Section 3 provides a detailed description of the methodology used for deriving the comparative cost data, whilst Section 4 describes the methodology for measuring the factors that drive the cost of comparator companies;
- Section 5 presents the results from applying the different estimation techniques discussed in Section 2, using the data derived in Sections 3 and 4. Also included is an analysis of the sensitivity of the central results to changes in the more significant assumptions that have been made; and
- Section 6 summarises the main conclusions of the study.

5 Extensive data on the US LECs is published by the Federal Communications Commission (FCC).

2. COMPARATIVE EFFICIENCY MEASUREMENT

2.1. Introduction

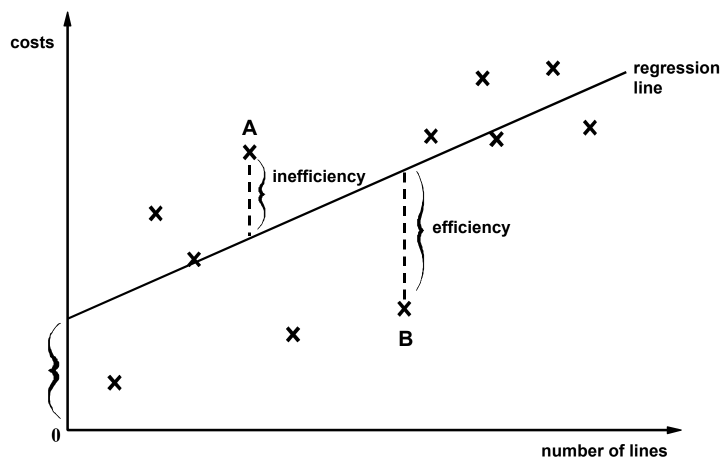
There are a variety of statistical and mathematical programming techniques that can be used to assess comparative efficiency. Therefore, in considering the most appropriate approach to take to comparative efficiency measurement, it is important to examine the relative merits and drawbacks of the alternative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

2.2. Ordinary Least Squares (OLS) Regression Analysis

Ordinary Least Squares analysis is one of a variety of different 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. Suppose that the cost of building and operating a network (C) depends 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 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) would 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 would be used to estimate the regression coefficients (a and b) using the data on all the companies in the sample. Individual companies could then be judged by substituting their actual number of exchange lines 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 as well as the number of exchange lines. OLS regression analysis deals with this through the use of multivariate regressions, which take the general form:

$$C_i = a + bL_i + cP_i + dQ_i + u_i$$

As before, *a* represents the level of fixed costs, *b* measures the marginal cost of explanatory factor *L* and *u* is the regression residual. But in addition, *c* and *d* 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).

2.2.1. 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.

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 (i.e. if the figures for each year for an observation do not differ by a large amount).

Ordinary Least Squares analysis is unable 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 therefore in this case is:

$$C_{it} = a + bL_{it} + cP_{it} + dQ_{it} + T + u_i$$

where T is the time trend, and L_{it} is the value of L for company *i* in time period *t* and so on.

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 6 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 3 of the 6 years), without the model being adversely affected.

2.3. 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 observation 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) 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 inefficiency. 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.

SFA estimates the equation:

$$C_i = a + bL_i + \ldots + v_i + u_i$$

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

Note that 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 the 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. 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.

It is within these assumptions that the key disadvantage of SFA lies, as the accuracy of these assumptions cannot accurately be measured.

SFA is described in greater detail in Appendix A below.

2.3.1. Multi-year stochastic frontier analysis

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

$$C_{it} = a + bL_{it} + \ldots + T + v_i + u_i$$

where T is a time trend variable that identifies the change over time in the regression constant, i represents an individual observation and t represents the time period. Note that residuals can be different for each firm and for each year[6]. 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 cross-sectional data (i.e. 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. Furthermore, 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 firm efficiency does not vary over time, which 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 also be relaxed, without losing the other advantages of panel data.

Reflecting the considerations above, 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.

---

[6] It is possible to estimate SFA panel data models in which the residual is both time variant and time invariant.

$$u_{it} = u_{it} \times e^{\eta(t-T)}$$

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

2.4. 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?
- 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? Would logarithmic transformation of explanatory factors give a better 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.[7] For meaningful results, there need to be many more independent[8] observations than the number of cost-driver

7 This is an example of what is known as multicollinearity in regression.

8 The need for independent observations means that it may not be sufficient simply to pool observations for the same companies over different time periods.

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 discussed below.

2.5. Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) can be used as an alternative to regression-based techniques. It is not a statistical estimation technique, 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.

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

2.6. Summary

In this section we have discussed some of the different techniques that can be used to measure comparative efficiency and have highlighted some of the main strengths and weaknesses of these techniques. Given these strengths and weaknesses, NERA considers it appropriate to make use of all of the techniques discussed above when analysing comparative efficiency. The results of these different techniques can then be reviewed in the light of the relative strengths and weaknesses in order to provide a more informed view of comparative efficiency. Additionally, if the different techniques all 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 feasible, if not likely, that different techniques will produce different efficiency results as different techniques are based on different underlying assumptions, but if there are similarities between the rankings of firms under the different techniques, this indicates that one can be confident in the relative position of the firm within the sample. Then only the question remains, of which efficiency result is most appropriate given the purpose for which it is to be used.

This question will be discussed further in section 5 below, when guidance is provided on the interpretation of BT's comparative efficiency results.

3. COST DATA METHODOLOGY

3.1. Introduction

The last section described the different techniques that can be used to compare the efficiency of different firms. The next two sections of this report discuss firstly the cost data and secondly the output and environmental data required to estimate the comparative efficiency of BT and the US LECs using these techniques.

3.2. Cost Comparisons

Before being able to draw any conclusions concerning relative efficiency, it is important to consider which costs are being compared. The regulatory objective of providing companies with incentives to minimise the overall cost to the customer (that is, the price that the customer pays for the service) means that the appropriate cost measure should, in theory, be total cost rather than other partial measures of cost such as operating expenditure. Total cost is defined as direct operating expenditure plus the cost of capital maintenance plus the required return on assets.

3.3. Structure of this Section

For both BT and the US LECs, this Section provides details of:

- the data sources;
- the methodology used in the estimation of total network costs (operating costs, depreciation and cost of capital);
- the currency conversion used for BT to allow comparison with US data;
- the conversion from historic cost to current cost accounting for US data; and
- the cost of capital for US companies.

3.4. BT Costs

3.4.1. Data sources

All cost and output data for the period 2000/01 to 2001/02 was supplied by BT. Cost data was provided on a CCA basis, and was broken down into the costs of the different network products used in BT's accounting system (for example, International calls (Business), Kilostream connections, and Public Payphone Calls). BT also supplied fixed asset values (Gross Replacement Cost and Net Replacement Cost) for each network product.[9]

3.4.2. Adjustment of costs to remove international network and retail costs

In the previous BT comparative efficiency study completed in 2000, the direct and indirect costs associated with BT's international network (switches and transmission) and the direct and indirect costs associated with the non-UK element of international calls were removed from BT's cost and asset base. This was done to make BT comparable to the US LECs. Whilst BT has both international and long distance operations, the US LECs only cover their local access transport area (or LATA), and hence their calls will either be local, intra-LATA (i.e. long distance but still within the local trading area), or the local "ends" of inter-LATA calls. The same approach was also taken in this study.

The accounting system in use by BT reports the costs of BT's UK network separately from those of its international network for the years covered by this study, as BT's international activities were undertaken by Concert during this period. Therefore the cost data provided by BT for the period 2000/01 to 2001/02 does not include the direct or indirect costs associated with BT's international network or the non-UK element of international calls. Additionally the asset data provided by BT for this period only includes assets in BT's UK network.

3.4.3. Adjustment of US GAAP

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

In its annual reports, BT Plc gives figures for the impact of applying US GAAP rather than UK GAAP for the issues listed above. Using this information we have adjusted the costs of BT's Network to reflect the differences between UK and US GAAP. The methods used to complete these adjustments are set out below:

- To calculate the adjusted pension cost figure a proportion of the total adjustment was allocated to BT Network based on the proportion of staff and other costs in the BT Network cost base as compared to the proportion of these costs in the cost base of BT Plc.

---

[9] To verify the robustness of the data, NERA reconciled the data provided by BT to that provided in BT's Current Cost Financial Statement.

- None of the capitalised interest costs adjustment was allocated to BT Network, because of difficulties in identifying whether any of this adjustment was relevant to BT Network. It is indicated in the financial statements that this adjustment includes the capitalisation of interest incurred on the funding of 3G licences[10]. The latter is not relevant to BT Network. The rest of the adjustment would not be expected to impact significantly on the costs of BT Network for the purpose of this study, as the majority of it is accounted for by the reversal of the interest charge, which is not part of the relevant cost base. This is because the total cost figure used in this study includes an allocation of the cost of capital (see section 3.2 above for further details) and therefore finance costs, such as interest charges, are not also included in the cost base.

- The adjustment for the sale and leaseback of properties needed to be analysed to identify the element of this adjustment which needed to be allocated between BT Network and the rest of BT Plc. To restate this transaction 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 leased interest charge and a depreciation charge. As the gain on the disposal of the fixed assets is reported in the financial statements of BT's Residual Business in its Current Cost Financial Statements, it is only necessary to allocate to BT Network a proportion of the difference between the rental charge used under UK GAAP and the finance lease plus depreciation charge used under US GAAP. This difference was identified, and was allocated to BT Network based on the proportion of total land and buildings assets that had been allocated to BT Network.

- To calculate the proportion of the impairment of assets adjustment that should be allocated to BT Network the proportion of the total asset base of BT Plc which is accounted for by BT Network was used.

- The adjustment for employee share plans was allocated to BT Network based on the proportion of staff and other costs in the BT Network cost base as compared to the proportion of these costs in the cost base of BT Plc.

Conversion to US GAAP increased the total cost of BT's network by 1.8% in 2000/01 and reduced the total cost by 0.5% in 2001/02.

3.4.4. Other costs removed from total costs

Other costs which were removed from BT's cost base for the purposes of this study are set out below:

---

[10] In the first year post the de-merger of BT and mm02 (2002/03) there has been a significant fall in the amount of this adjustment, this supports the assumption that a significant proportion of the adjustment in prior years was due to the capitalisation of interest incurred in relation to the financing of 3G licences.

- Payments by BT to other mobile operators (POMOs) were removed to account for the fact that the called party pays for mobile calls in the US and hence no payment is made by the fixed operator. Also any POLOs relating to international calls were removed from BT's cost base as the LECs do not incur such costs.

- Emergency or "911" costs were removed from the BT cost base because in the US, while the routing of 911 calls to the public safety answering point (PSAP) is under the control of the local telephone company, the processing of the call at the PSAP is controlled by a myriad of different entities, none of which have a regulatory tie to the FCC.

- Property taxes have been excluded from BT's cost base as such costs are not included in the LECs expenses, as they appear under the category other operating taxes (7240), which was not included in the LECs' expenses. In addition, because such costs are outside of the control of the operator (i.e. are mandated by government), it would be inappropriate to include them in the cost base for a comparative efficiency assessment.

- Satellite Service Business and Data Network Costs[11] were excluded from BT's cost base, as NERA has been unable to establish that such costs are included in the cost base of the US LECs. The FCC has indicated that, whilst it does not expect such costs to be included in the data reported for the US LECs, it cannot confirm that they are not included.

- Transit and Admin lease costs were excluded from BT's cost base as, in the previous BT comparative efficiency study completed in 2000, these costs were identified as being relevant to the provision of international network services. The costs included in the Network costs under these headings are very small in both 2000/01 and 2001/02.

Overall, 30% of BT's total costs (inclusive of cost of capital) were eliminated in 2000/01 and 29% in 2001/02 through the removal of the costs mentioned above.

The majority of this adjustment is due to the removal of payments to mobile and international operators. However, the removal of data network costs also had a noticeable impact upon BT's total costs. Given that the FCC have indicated that the US LECs are supposed to exclude data network costs when submitting their returns, it was thought prudent to remove these from BT's cost base. However, a sensitivity to the exclusion of these costs was carried out. The results of this sensitivity are presented in section 5.4.5 below.

[11] Data network costs excluded are the costs of VPN, VPS, ATM, SMDS, SHDS and Framestream.

3.4.5. 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 (UK£) or the BT data into US dollars (US$). As this conversion involves the use of common exchange rate factors for the firms to be converted (either using the rate for UK£ to US$, or the inverse of this rate for the conversion of US$ to UK£), the conversion of either BT into US$ or the US LECs into UK£ will have limited (if any) impact upon the results of the study[12]. However, as the conversion of the BT into US$ involves converting the data of only one firm whilst converting the data of the US LECs into UK£ involves converting the data of up-to 50 firms in every year of data (from 1996 to 2001 in this study), the conversion of BT rather than the US LECs will minimise the risk of any inadvertent data processing errors. Therefore the data for BT was converted from UK£ into US$.

It would not be appropriate to use actual market exchange rates in this conversion, since, with the exception of goods that can be purchased on international markets, actual market exchange rates typically 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 differences in price levels between countries, and so have been used in our analysis. The PPP data used in this study has been obtained from the OECD publication "Purchasing Power Parities and Real Expenditures" and the IFS Yearbook.[13]

The OECD publication provides data for the PPP of GDP and PPP's for specific asset categories (for example, non-residential buildings, electrical equipment, non-electrical equipment, transport equipment and civil engineering works). The PPP for GDP is generally considered to be more statistically reliable as it is based on a significant number of different data points, whilst the PPP's for specific asset categories can be based on a relatively small number of data points for each category. In addition, the GDP PPP is available for all years whilst the breakdown of PPPs by asset categories are only available for a sub-set of years, the latest of which is 1993. Therefore, it is necessary to forecast forward these PPP rates using the trend in the GDP PPP. Given the period for which this forecast forward is required (1993 to 2001), this further calls into question the reliability of the PPPs by asset categories.

---

12 Analysis was undertaken to test this assertion, this involved testing the equivalence of the regression coefficients estimated first with the US LECs in US$ and secondly with the US LECs in UK£s. These tests accepted the hypothesis that the coefficients under both of these regressions were equal. Note, these test excluded the constant term which, as this represents the fixed cost of the operators, will be different when the cost data is expressed in a different currency.

13 The OECD publication gives a variety of PPP exchange rates for different sectors of industry. However, since the most recent figures for this breakdown are for 1993, we have uplifted them each year in proportion to the growth of the overall GDP PPP rate, which is published by the IFS.

Therefore, the approach used to convert BT's data from UK£ to US$ was to use the PPP for GDP for all cost and for all asset categories with the exception of the following[14]:

- Switching equipment, transmission equipment, other network equipment, computers, office machines and payphones: The choice of the most appropriate currency conversion for telecommunications and electronic capital equipment costs raises a number of issues. Telecommunications equipment (and other electronic equipment) can be purchased on international markets and so it can reasonably be assumed that BT and the US LECs effectively face the same prices at market exchange rates.[15] This would mean that, by converting prices using an GDP PPP rate, we would under-state the prices that BT needs to pay now for equipment by up to 5% (the GDP PPP compared to market exchange rate). Therefore, the approach we have taken is to assume that the use of the market exchange rate is the appropriate currency conversion for telecommunications and electronic equipment (averaged over the last 10 years to reflect the purchase schedule of the assets)[16].

A potential alternative approach to converting BT's data from UK£ to US$ would be to use smoothed exchange rates rather than the market exchange rates. The World Bank uses a method called "Atlas" to derive smoothed exchange rates. The purpose of this method is to control for the impact of exchange rate fluctuations. We investigated the impact that using smoothed exchange rates, rather than the market exchange rates, would have upon BT's total cost in US$. We found that using smoothed exchange rates would make BT's total cost in US$ less than 1% higher in 2000 and 1% lower in 2001. A sensitivity test was completed to assess the impact this cost change has upon BT's efficiency score. The results of this sensitivity are presented in section 5.4.5 below.

3.5. US LEC Costs

3.5.1. Data source

The raw data used for calculating the total costs of the US LECs is drawn from the Federal Communications Commission (FCC) publication "Statistics of Communications Common

---

14 BT's costs and assets were converting using both this approach, and the approach used in the previous comparative efficiency study, when the asset PPPs were used (where relevant), to assess the impact of this change in approach. The total cost of BT using the GDP PPP approach were 0.01% lower in 2001/02 and 1.14% lower in 2000/01 than the total costs using the asset PPPs approach.

15 If prices in dollars (for example) differed between countries, this would set up an arbitrage opportunity. Such a differential would not therefore be sustainable for a significant period.

16 A similar assumption was also made during the previous BT efficiency study completed in 2000.

Carriers" for the years 1996 to 2001 (inclusive).[17] Tables 2-10 and 2-11 of this publication provide a balance sheet and an income statement for each of the reporting US LECs.

3.5.2. Data adjustments

Total costs are comprised of the following three elements:

- depreciation;
- cost of capital; and
- operating costs.

The first two are derived from the equipment costs in the balance sheet section of FCC Tables 2-10 and 2-11. Operating costs are obtained from the income statement section of the same table. However, before these costs can be compared to those of BT, a number of adjustments must be made in order to ensure comparability of data.

3.5.3. Streamlining of reporting for medium-sized firms

The FCC reduced the 1999 reporting requirements for local exchange carriers with annual operating revenues less of than $7 billion (such firms are termed 'mid-sized').[18] While the total amounts in each category are still reported, the division into sub-categories is not always reported. In previous years we have used different PPP rates for some of the sub-categories.

We found that the ratios of the cost of sub-categories within cost categories were largely constant between the companies in the sample. It was therefore felt that, from 1999 onwards, it was appropriate to apportion the costs for the mid-sized US LECs between sub-categories using the ratios for the aggregate of all the large US LECs.

3.5.4. Pension adjustment

On examination of the data provided by the US LECs, it was found that several of the entries under item 6728, 'Other General and Administrative Expenses', in 1999, were significantly negative, against intuition and experience from all other years in the sample. This was because of reduced pension commitments following mergers involving the former Ameritech and GTE companies. In order to provide a true measure of underlying costs it was necessary to correct for these effects.

---

17 This data can be found on the FCC website at http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCC-State_Link/socc.html.

18 Through the FCC Accounting Review Order, 14 FCC Rcd 11396 (1999).

For the eleven companies, the 1999 figures for item 6728 have been replaced with the 1998 figures[19]. While this is not ideal, we believe that it removes the distortions in the 1999 figures sufficiently to ensure meaningful results.

3.5.5. Software cost capitalisation

A change in organisational structure which affected several of the US LECs (specifically, those in the Bell Atlantic and GTE groups) in 1999 resulted in a one-off amendment to the treatment of software expenses by these LECs. This resulted in a capitalisation of software expenses in 1999. As this was a one-off amendment, this did not re-occur in subsequent years. This resulted in a fall in expenses, and a corresponding rise in assets (and thus the cost of capital) in this year. Since this amendment would otherwise have led to inconsistent reporting across the LECs in 1999, these costs have been put back into the relevant expense categories for the companies concerned, and the assets and cost of capital reduced accordingly.

With the exception of the GTE companies, this amendment was reversed using the data reported in the relevant FCC accounts. In the case of GTE companies, the increase in 'plant addition intangibles' as per the FCC data did not equal the amount given in the group annual report, so the latter amount was shared between GTE companies according to the ratio of the former.

3.5.6. Conversion from historic cost to current cost accounting

As the asset book values reported for the US 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 US LECs were converted from historic to current costs. The derived costs from which were compared to BT's cost as derived under Current Cost Accounting (CCA).

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

- Firstly the average age[20] of the assets of each US LEC was estimated for each of the following asset categories: cable, duct, switches, transmission equipment, radio transmission, payphones, buildings, accommodation plant, computing equipment,

---

[19] Looking at costs under item 6728 for years prior to 1999, there is a flat or downward trend.

[20] The average age rather than actual age of the assets was used as data on the actual age of assets is not available for the US LECs.

and vehicles. The average age for each asset category for each US LEC was calculated using the following formula:

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

- Secondly, data on US price indices for each of these asset categories was collected from the US Statistical Abstract. These price indices allowed the identification of the change in asset prices between the point-in-time (on average) that each US LEC purchased its assets and the present day. For example, the price index for Telephone and Telegraph Cable and Wire was 154 in 1992 and 161 in 2001. This indicated that the price of this asset had increased by 4.5% between 1992 and 2001.
- Thirdly, the GBV of the assets of each US 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 (i.e. 1996, 1997, etc…). This adjustment can be represented by the following equation:

$$GRC_{Cable2001} = GBV_{Cable2001} * \left(\frac{Cable\,Price_{2001}}{Cable\,Price_{Average\,Age}}\right)$$

An alternative approach to the adjustment suggested above would be to use BT's CCA/HCA ratio as a proxy for the CCA/HCA ratios for the US LECs. Our reason for not using this alternative approach in this study is that the average age of the assets belonging to BT could well differ (possibly even to a significant extent) from the average age of the assets belonging to the different US LECs, and so constraining the LECs to have the same HCA/CCA ratio as BT may introduce errors into the analysis.[21]

The US LECs separately report Telecommunications Plant Under Construction (TPUC), without identifying the component assets. Therefore, it is not possible to identify which asset category this construction related to. However, this is only a very small element in the overall asset base of the US LECs and in 2001, for example, TPUC accounted for, on average, only 2% of the total Telecommunications Plant asset base. Therefore, since there is no information available concerning exactly what assets are included in this account, we allocate these assets to the GRC of each asset category based on the assumption that it is

[21] In previous studies NERA completed for Oftel assessing the efficiency of BT, the alternative approach of applying BT's CCA/HCA ratio to the US LECs' asset bases was used. Tests of the impact of changing from this approach to that used in this study (i.e. using the average age of the US LECs' assets and US asset price changes) have shown the impact of this change to be minimal.

exclusively replacement investment (we allocate plant under construction to the various asset groupings in proportion to the relative level of depreciation in each).[22]

3.5.7. Re-estimation of depreciation

Depreciation policies may differ between operators. Operators with faster depreciation schedules may record greater depreciation in any one year on the same asset base. If the faster depreciation is a reflection of actual shorter asset lives, then the different depreciation charges will nevertheless be comparable (simply reflecting a higher rate of capital usage for the operator with shorter asset lives). Even if the actual lives are the same, different depreciation schedules may still not be a problem in a steady state where assets are depreciated and added at the same rate every year. However, in cases where the capital base is growing, deviations between depreciation schedules (compared to actual asset lives) will be problematic. Therefore, for this analysis we calculate depreciation charges for each operator, based on BT's asset life and depreciation methodology. We obtain estimated depreciation for each asset group and each company in the US by dividing the GRC by the total asset life.

3.5.8. Estimation of Cost of Capital

The cost of capital may also be affected by depreciation policies. This is because different depreciation policies will affect the extent to which the gross replacement cost (GRC) of different operators has been written-down to the net replacement cost (NRC). Furthermore, comparisons based on crude NRCs will be biased against operators with faster growing asset bases, since their assets will tend to be newer and, therefore, their NRCs will be closer to their GRCs than for those operators with slower growing asset bases.

For the purposes of efficiency comparison we do not necessarily wish the capital cost to be affected by differences in depreciation policies or in the rate of asset growth. However, this problem can be avoided by imposing a common NRC/GRC ratio on all operators. Therefore, when deriving capital costs, we have applied the NRC/GRC ratio of BT to derive the NRC of the LECs' assets.[23]

Once the total NRC for all assets had been obtained for each US LEC, we estimated the cost of capital by multiplying the NRC by the required rate of return. To avoid estimated

---

22 Note that, when a construction project is completed, its cost is credited to account 2003 and charged to the account of the relevant asset.

23 Ofcom has informed NERA that the NRC/GRC ratio of BT may be understated due to the presence of fully-depreciated assets within BT's GRC figure, which have been excluded from BT's NRC figure. To assess the potential impact of this understatement NERA carried out a sensitivity test on changing the NRC/GRC ratio used to adjust the US LECs' assets. The results of this sensitivity test are shown in section 5.4.5 below.

efficiency being affected by differences in the cost of capital (rates of return), the required rate of return is set at 13.0% for all companies.[24]

3.5.9. Operating expenses

Total operating expenses are entered in FCC Tables 2-10 and 2-11 under account 720. For our analysis, we have removed two elements of cost:

- station apparatus expenses, and large private branch exchange expenses (which are not included in BT's cost base); and
- total depreciation and amortisation expenses (since these are covered by the above depreciation calculation).

In addition, we have added the following cost item into the operating expenses of the US LECs:

- Bad debts appear as a deduction from revenue for the LECs whereas they appear as an expense for BT. As a consequence, bad debts relating to telecommunications services for the LECs have been included in the operating expenses of the US LECs.

3.5.10. Network, retail and indirect costs

Finally, we also isolated the costs associated with the network business for the US LECs. To begin with, all cost categories for the US LECs were categorised as either:

- directly attributable to the running of the network;
- directly attributable to retail activities; or
- indirectly attributable to both.

Setting aside indirect costs, the proportions of each company's total direct costs (retail and network) that were attributable to each of 'network' and 'retail' were estimated. Following this, indirect costs were allocated between 'network' and 'retail' according to these proportions, for each company.

Having described the methodology used to derive the costs of all the companies in the sample, the next step involves setting out the methodology used to derive the output data used in our analysis.

---

[24] This is BT's cost of capital for its network business in 2001/02: see "Current Cost Financial Statements for the Businesses and Activities 2002 and Restated 2001 Current Cost Financial Statements", BT, September 2002, p14.

4. OUTPUT AND ENVIRONMENTAL DATA METHODOLOGY

4.1. Introduction

The output variables we considered were those most likely to be drivers of total costs in a telecommunications network, such as the number of access lines (including analogue and digital leased lines and special access lines/partial private circuits) and the number of minutes of network usage. These are discussed in more detail below.

A number of operating environment variables were also considered as possible cost drivers. These included:

- population density and the number of local switches (including remotes) as measures reflecting customer dispersion;
- aerial and non-aerial cable sheath length per line, which reflect differences between operators in terms of the average length of their transmission lines and access lines (which is inversely related to line density);
- the respective proportions of business and residential lines and the proportions of leased lines and switched lines, both of which are determined by the composition of the customer base that the operator serves; and
- the proportion of calls that is national (or intra-LATA), which is determined by the calling habits of the population the operator serves and is therefore largely outside the control of the operator.

These are also discussed in more detail below.

In principle, cost drivers should be chosen that are exogenous, that is outside the control of management; otherwise there is no way of distinguishing between a company that has a high cost for a particular cost category (for example, overhead lines) because of its circumstances and one that has a high cost because of inefficient operating practices. In practice, however, identifying the exogenous causes of costs is not always possible and so the use of proxies, like line length, may be unavoidable.

4.2. Number of Lines

Previous studies of costs in the telecommunications industry indicate that the number of access lines is an important cost-driver, since this is a major influence on the overall size of the network, as well as the size of the customer support systems required. Access lines can provide either switched access (normal exchange lines) or non-switched access (e.g. leased lines). Furthermore, leased lines can be of various capacities.

For BT, there are three main types of lines:

- switched exchange lines (including payphone lines);
- private circuits (or leased lines); and
- partial private circuits (equivalent to special access lines for the US LECs).

For the US LECs, there are also three main types of line:

- switched access lines (including payphone lines and Centrex lines);
- leased lines; and
- special access lines (these are dedicated lines from the customer to the inter-exchange carrier point of presence).

Ideally we would wish to use pure volume based measures of all three types of lines. However, whilst FCC Table 2-6 does present data on the total number of switched access and special access lines, it does not include data on the number of leased lines provided by each LEC. FCC Tables 2-10 and 2-11 do however include figures for each LEC for total annual revenue from leased lines. Our analysis uses these revenue figures to estimate the number of leased lines provided by each LEC, based on US leased line prices by capacity published by the OECD.[25]

In order to avoid problems of multicollinearity between the three line variables (which are highly correlated with one another), we have adopted an aggregate measure for lines which sums the three together (note that a 64kbit-equivalent leased line corresponds as closely as possible to a single switched access line). Therefore, our final measure of the number of lines which each LEC provides is:

$$\text{Lines} = \text{Switched Access Lines} + A \times \text{64kbit-Equivalent Special Access Lines and Partial Private Circuits} + B \times \text{64kbit-Equivalent Leased Lines}$$

Where:

- A and B are the weights on special access lines and leased lines respectively; and
- A will lie between 1 and B.

Since the weights that should be applied to special access lines and leased lines when they are compared with each other, and with switched access lines, are not known precisely, sensitivities were carried out using different weights.

25 Communications Outlook 2001, OECD, 2001.

In the previous BT efficiency study completed in 2000, each of the different line types was given equal weight in the aggregate lines measure. Further investigation, however, has suggested that it may be more appropriate to give leased lines and special access lines a greater weighting than switched exchange lines to reflect the fact that a leased line has two customer ends and generally a main link and that, while a special access line has only one customer end (special access lines connect with an inter-exchange carrier point of presence), it generally has a main link and an interconnecting circuit at the inter-exchange carrier point of presence. Therefore, regressions have been run using:

- equal weights on all line types (as in the previous BT study in 2000); and
- using weights of 1 for switched access lines, 2 for leased lines and 1.5 for special access lines.[26]

It should be noted that, for the LECs, the number of digital special access lines decreased sharply between 1998 and 1999, but increased in 2000 to a level that would be expected from the previous (pre-1999) trend. The FCC believes that there may be an error in the reported data in 1999 owing to an ambiguity in the data request sent out that year, and therefore NERA has used estimated figures for 1999 based on the trend in line growth over the years 1997 to 2001.

4.3. Traffic Volumes: NERA Methodology

In addition to the number of access lines, the costs of a telecommunications network will depend on traffic levels. Ideally, traffic should be measured in terms of:

- busy hour Erlangs (driving the cost of transmission and switch port capacity); and
- busy hour call attempts (driving the cost of switch processing capacity).

Unfortunately, there is no publicly available data for these quantities for US LECs. However, call volume data is published in FCC Table 2-6. Provided that the "load profile" of traffic (the percentage of traffic in the busy hour) is similar across all operators, then call minutes can be taken to proxy busy hour Erlangs. Although we have no quantitative data to verify whether this is the case, there are two plausible reasons why US load curves will (if anything) be more peaky than those of BT, and one reason to suppose they may be less peaky:

- there is a smaller overlap of working hours between cities in different US time zones. Therefore, traffic between New York and Los Angeles (for example) must be concentrated into fewer hours;

[26] Intuitively special access lines, as they connect with an inter-exchange carrier point of presence rather than to a customer premises (as is the case with leased lines) should receive a weighting of between 1 and 2.

- free local calls mean that customers have no incentive to keep their calls (including social calls) out of the peak tariff period (which usually contains the busy hour). This is emerging as a particular problem in the US with internet calls using up capacity at peak times on the local network; and
- offsetting these effects is the fact that the different time bands may mean that originating traffic is more spread out during the day according to its destination. For example, business traffic originating in New York and destined for Los Angeles occurs in the early evening when east coast-destined traffic is at a lower level.

Therefore, by using overall call minutes rather than busy hour Erlangs, we anticipate that, if anything, we are underestimating LEC busy hour traffic relative to that of BT.

Apart from busy hour Erlangs, the other main driver of switching costs is busy hour call attempts. These can again be proxied by call minutes provided the load profile and average call lengths are the same across operators. (Unfortunately, there is no data available for the US LECs to test whether this is the case). Again, if traffic is more peaky in the US, the use of minutes of traffic will tend to overstate BT's relative level of busy hour call traffic.

Call minutes are usually stated in terms of the number of conversation minutes and subdivided into local, national and international calls. However, since local calling areas differ in size, and national calls in different countries require different numbers of switching and transmission stages, these measures are not suitable for international comparisons. These problems are avoided if traffic volumes are expressed in terms of the number of switch minutes. If one country has larger local calling areas where a greater proportion of local calls pass through more than one switch, then switch minutes per local call will be correspondingly higher.

4.3.1. Local switch minutes

Local switch minutes are defined as the total number of call minutes (local, national, international, or other) that pass through local switches. For BT, local switch minutes (per annum) are calculated by using the formula:

Local Calls PSTN × Local Switch Routing Factor for Local Calls PSTN +
Local Calls ISDN × Local Switch Routing Factor for Local Calls ISDN +
National Calls PSTN × Local Switch Routing Factor for National Calls PSTN +
National Calls ISDN × Local Switch Routing Factor for National Calls ISDN +
International Calls PSTN incoming × Local Switch Routing Factor for International Calls PSTN incoming +
Public Payphones calls × Local Switch Routing Factor for Public Payphones calls +
International Calls PSTN outgoing × Local Switch Routing Factor for International Calls PSTN outgoing +
International Calls ISDN all × Local Switch Routing Factor for International Calls ISDN all +

International Calls Chargecard × Local Switch Routing Factor for International Calls Chargecard +
International Calls Operator Assistance × Local Switch Routing Factor for International Calls Operator Assistance +
Calls to Mobile × Local Switch Routing Factor for Calls to Mobile +
Local Switch Routing Factor for Linkline/Freefone calls +
Premium Rate Services calls × Local Switch Routing Factor for Premium Rate Services calls +
FeatureNet calls × Local Switch Routing Factor for FeatureNet calls +
Operator Assistance calls × Local Switch Routing Factor for Operator Assistance calls +
Other calls × Local Switch Routing Factor for Other calls +
Chargecard calls × Local Switch Routing Factor for Chargecard calls +
Intermediate calls × Local Switch Routing Factor for Intermediate calls +
AS Network Products - Non Transit calls × Local Switch Routing Factor for AS Network Products – Non Transit calls +
AS Network Products – Transit calls × Local Switch Routing Factor for AS Network Products – Transit calls
= Total local switch minutes

In each case the routing factor is the average number of local switches through which the call passes.

Total local switch minutes for US LECs (per annum) are calculated by using the formula:

Local Switch Routing factor for local calls × local call minutes
+ Local Switch Routing factor for intra-LATA × intra-LATA call minutes
+ Local Switch Routing factor for inter-LATA calls × inter-LATA call minutes
= Total local switch minutes

Routing factors for BT have been supplied to us. For US LECs, we estimate that 46% of local calls are "own exchange"[27], whilst the remainder use two switching stages. This would imply a local call switching factor of 1.54.[28]

Intra-LATA calls pass through two local switches (one at each end of the call), whilst, for inter-LATA calls (where we consider incoming and outgoing calls separately), only one local switch is involved per end. Therefore, the routing factors used are as follows:

27 Estimate based on an input to the Hatfield cost model.

28 Sensitivity tests were completed to estimate the impact of this assumption on the efficiency results. The results of these tests can be found in section 5.4.5 below.

Table 4.1
Local Switch Routing Factors

US LECs	
Type of Call	**Routing Factor**
Local calls	1.5
Intra-LATA calls	2.0
Inter-LATA call ends	1.0

Source: Hatfield model and NERA analysis

4.3.2. Main switch minutes

Total main switch minutes for BT (per annum) are calculated using the formula:

Local Calls PSTN × Main Switch Routing Factor for Local Calls PSTN +
Local Calls ISDN × Main Switch Routing Factor for Local Calls ISDN +
National Calls PSTN × Main Switch Routing Factor for National Calls PSTN +
National Calls ISDN × Main Switch Routing Factor for National Calls ISDN +
International Calls PSTN incoming × Main Switch Routing Factor for International Calls PSTN incoming +
Public Payphones calls × Main Switch Routing Factor for Public Payphones calls +
International Calls PSTN outgoing × Main Switch Routing Factor for International Calls PSTN outgoing +
International Calls ISDN all × Main Switch Routing Factor for International Calls ISDN all +
International Calls Chargecard × Main Switch Routing Factor for International Calls Chargecard +
International Calls Operator Assistance × Main Switch Routing Factor for International Calls Operator Assistance +
Calls to Mobile × Main Switch Routing Factor for Calls to Mobile +
Main Switch Routing Factor for Linkline/Freefone calls +
Premium Rate Services calls × Main Switch Routing Factor for Premium Rate Services calls +
FeatureNet calls × Main Switch Routing Factor for FeatureNet calls +
Operator Assistance calls × Main Switch Routing Factor for Operator Assistance calls +
Other calls × Main Switch Routing Factor for Other calls +
Chargecard calls × Main Switch Routing Factor for Chargecard calls +
Intermediate calls × Main Switch Routing Factor for Intermediate calls +
AS Network Products - Non Transit calls × Main Switch Routing Factor for AS Network Products – Non Transit calls +
AS Network Products – Transit calls × Main Switch Routing Factor for AS Network Products – Transit calls
= Total Main Switch minutes

The relevant routing factors for BT were supplied by BT.

Total main switch minutes for US LECs (per annum) are calculated by using the formula:

Routing factor for local calls × local call minutes
+ Routing factor for intra-LATA calls × intra-LATA call minutes
+ Routing factor for inter-LATA calls × inter-LATA call minutes
= Total main switch minutes

For the US LECs, 98% of local inter-office traffic is directly routed between originating and terminating end offices as opposed to being routed via a tandem switch.[29] This yields a routing factor for local calls passing through tandem switches of 0.02. For long distance calls, however, the situation in the US is more complex. It can be assumed that all long distance calls using LEC switches pass through two local switches - one at each end of the call. However, beyond this, routing is different for intra-LATA and inter-LATA calls.

In the case of intra-LATA toll calls, around 80% are assumed to be routed directly to the terminating local switch[30], whilst the remaining 20% are assumed to transit through a tandem switch – thus giving a routing factor of 0.2. Inter-LATA calls, on the other hand, are assumed to be routed from the LEC local switch directly to a long distance carrier Point of Presence (POP) in 70% of cases[31], whilst in the remainder of cases they will transit through an additional tandem switch on route. At the far end of the long distance operator's network the call will exit at another POP (again usually at a long distance operator switch but sometimes at another location), and be delivered through the destination LEC's network in roughly the same manner as that in which it was originated. This would suggest that each end of an inter-LATA call passes through an average of 0.3 tandem switches within the LEC network.

Therefore, the routing factors used are as follows[32]:

Table 4.2
Main Switch Routing Factors

US LECs

Type of call	**Routing factor**
Local calls	0.02
Intra-LATA calls	0.2
Inter-LATA call ends	0.3

Source: Hatfield model, Bellcore and NERA analysis

---

29 Input to the Hatfield cost model.

30 Based on inputs to the Hatfield cost model.

31 Based on information obtained from Bellcore.

32 Sensitivity tests were completed to assess the impact of changing the main switch routing factors for Intra- and Inter-LATA on the efficiency results. The results of these tests can be found in section 5.4.5 below.

4.3.3. Total switch minutes

Multicollinearity presents a particular problem in the case of local and main switch minutes, where the correlation between the variables is so strong that it is impossible to identify the separate cost impact of an additional local switch minute as opposed to a main switch minute. However, it is possible to construct a composite switch minute variable using external data (particularly network costing studies). Relevant data is shown in the table below.

Table 4.3
Relative Cost of Local and Main Switching

	Source	**Local switching**	**Main switching**	**main / local**
US (LRIC)	*FCC*	0.2-0.4 cents	0.15 cents	0.38-0.75
BT (LRIC)	*BT*[33]	0.135-0.140 pence	0.045 -0.060 pence	0.39-0.57
"average of mid-points"				0.5

Source: as shown

This data indicates some variation in the relative costs of local and main switches in the networks of different operators. However, for the purposes of this analysis it is necessary to determine the appropriate weight for the "output" of a main switch minute relative to a local switch minute, rather than the relative cost. This is because the fact that the cost of main switches in the BT network may differ from those in other networks is part of the cost efficiency we are trying to measure. If we constructed our output measure using different cost factors for each company, we would remove any cost inefficiency difference.

This suggests that we should use the same relative weight for the output of local and main switch minutes for all operators. The relative weight we have chosen is based on the midpoint cost ratio for local and main switch minutes for all operators. The data in Table 4.3 suggests a figure of around 0.5 for this ratio. Therefore, a new variable has been calculated to reflect traffic volumes:

$$\textit{standard switch minutes} = (\textit{ local switch minutes }) + 0.5 \times (\textit{ main switch minutes })$$

Sensitivity tests were completed to assess the impact that the assumed weight on main switch minutes has upon the efficiency results. The results of these tests are presented in section 5.4.5 below.

33 NERA analysis of data in "Current Cost Financial Statements for the Businesses and Activities 2002 and Restated 2001 Current Cost Financial Statements", BT, 2002 and "Current Cost Financial Statements for the Businesses and Activities 2001 and Restated 2000 Current Cost Financial Statements", BT, 2001.

4.3.4. Local call volumes

In December 1999, the FCC revised its definition of what should be included by the LECs in their submissions concerning the number of local calls made using their networks. This revision introduced an explicit instruction to LECs to include the number of unanswered calls in their measures of the total number of local calls made, whereas previously there had been no such explicit request for data on unanswered calls.

This raises the question as to what the local call data used for years prior to 1999 represents. Comments from the FCC indicated that some LECs already included unanswered calls when submitting local call numbers whereas others did not. However, they had no indication as to what proportion of companies included unanswered calls.

If the LECs had not included unanswered calls in the local call numbers prior to 1999 we would expect to see a sharp increase in the number of local calls in 1999 compared to earlier years. However this does not appear to be the case for the vast majority of LECs in 1999. Therefore simple inspection of local call numbers suggests that most were already including unanswered calls in their local call numbers. To allow for this we have reduced all local call minutes by 30%[34] for all years in order to provide estimates of the number of completed local calls for each of the LECs.

4.4. Adjusting for Difference in Calling Rates Between US and UK Market

In modelling the efficiency of BT, we need to consider the differences in calling rates per line between BT and the US LECs. The output data collected for this study suggest that calling rates per line (minutes per line) in the US are approximately 1.5 times greater than in the UK[35].

We have tried to capture this difference implicitly though the basic formulation of the regression model.

$$ln(cost) = \alpha + \beta_1 ln(lines) + \beta_2 ln(minutes) + \ldots + u$$

This could be made more explicit by using the compound variable "minutes/lines", instead of simply "minutes", as an explanatory variable in the model. However, since the model is constructed using a double-log specification, the only impact this will have is to impose some constraints on the model coefficients, as follows:

---

34 This is based on the typical proportion of unanswered calls drawn from the Hatfield model in the US, and is consistent with NERA's experience of the proportion of calls that are unanswered calls in a variety of European countries.

35 In 2001 the average number of call minutes per exchange line over all of the US LECs was 1.6 times the number of call minutes per exchange line of BT. In 2000 this ratio was 1.4.

$$ln(cost) = \alpha + (\beta_1 + \beta_2)\, ln(lines) + \beta_2 ln(minutes/line) + \ldots + u$$

The value of *u* the residual error, would not change. Since it is this residual which reflects the comparative efficiency of the various companies, forcing the model to explicitly consider differences in calling rates would have no effect on the results, and would only serve to impose constraints on the coefficients in the model.

4.5. Environmental Factors

A number of variables exist which could potentially be used to account for environmental differences between operators. Such differences could include both the dispersion of customers, and the make-up of customers (e.g. businesses and residential customers).

Ideally, we would wish to use variables that are totally exogenous to the telephone operators themselves, but which nonetheless capture environmental differences. By exogenous we mean that the environmental variable, and hence differences in that variable, should be outside the control of the operators. Wherever possible, NERA has attempted to use fully exogenous environmental variables (e.g. population density). However, in practice, there is often a "trade-off" between finding genuinely exogenous environmental variables, and variables that can actually be observed for each operator. Therefore, in cases were fully exogenous variables are not observable, NERA has investigated observable variables to identify those which, whilst only partially exogenous[36] (e.g. cable sheath length per line), are still indicators of environmental differences.

4.5.1. Measures of Population Dispersion

One possible candidate for measuring population dispersion is average population density. However, it has serious limitations. By averaging across a whole area, average population density cannot take account of the population distribution. For example, it is possible for two telephone companies of equal size to have the same average population density even though one company serves a population that is concentrated in urban areas (surrounded by uninhabited areas), whilst in the other serves a population that is evenly spread over large rural areas. One would expect companies with significantly different population distributions to have significantly different costs (with the costs of the company with the wider population distribution being higher). Therefore a measure of average population density alone is insufficient to control for differences in population dispersion.

In addition, it is not possible to find population density figures for each of the US LECs. Population density figures are produced on a state-by-state basis but the areas served by each US LEC do not follow state boundaries. For example, some states are served by more

36 Cable sheath length per line is only partially exogenous as, whilst it is largely determined by the dispersion of an operator's customers, it is also dependent upon how efficient the operator is in identifying an optimal network geography and/or structure.

than one US LEC (with the state being split between the different companies) and some US LECs serve part, or all, of more than one state. Therefore, the state-by-state population density figures could provide a highly misleading picture of the average population density of the customers served by each of the US LECs.

An approximate measure of population distribution was also considered. This variable, measured as the percentage of the population living in a metropolitan area of more than 100,000 people, could theoretically work with the population density to create a better proxy of the population related environmental factors faced by a company. However, this data was also only available on a state-by-state basis and hence could provide a highly misleading picture of the population distribution of the customers served by each of the US LECs.

Given the problems mentioned above in accurately mapping state-specific data on to the different US LECs, which could result in a biased picture of the population dispersion of the US LECs customer base, it was decided not to include either the population density or population distribution variables in the regression analysis.

A third possible measure (although not one that is entirely exogenous to the telephone operators) is the geographical extent of the network required by each operator to provide service coverage within its operating area. In the case of the US LECs, data is available for cable sheath length (both aerial sheath length and non aerial sheath length); this is published for each company in FCC Table 2-6.[37] Corresponding data has also been provided by BT. Using this data, it is possible to derive a figure for cable sheath length per (access) line for each operator.[38]

On average, the LECs have a considerably higher percentage of cable above ground than BT. In an attempt to capture this difference in operating environments, sheath length per line was further broken up into aerial and non-aerial sheath length per switch line.

A fourth possible measure is the average local loop length per line, as this should be correlated with the dispersion of an operator's customers (however, as with cable sheath length, this will not be an entirely exogenous variable).

Despite the issues with the partial endogeneity of the cable sheath length and loop length variables, the fact that the data for these variables is produced for each LEC individually, makes them superior to the direct population density and dispersion variables discussed

37 For the purposes of our analysis, "sheath length" includes sheath km of metallic cable and fibre aerial cable, underground cable, buried cable and intrabuilding network cable. However, we have excluded the sheath length of submarine cable and deep-sea cable as it is not part of an inland network such as that operated by the LECs. (All data is from FCC Table 2-6).

38 There was no data available for aerial route length, so it was necessary to use cable sheath length to derive a network dispersion variable.

above, which could produce highly-biased results given the significant problems in accurately mapping the state data on the to different US LECs.

Other environmental variables that were considered were the number of concentrator units and central office switches. In the case of the US companies, data was only available for the total number of switches and the number of remote switches. Therefore, we used the number of remote switches to proxy for the number of concentrator units. However, there are some significant discontinuities in the number of remotes reported by the US LECs over time, which call into question the accuracy of this data. In addition, since the number of remotes is clearly not an accurate proxy of network dispersion, and can in any case be considered to be endogenous (i.e. under the control of the operator) to a significant degree, the number of concentrator units (remotes) and central office switches was rejected as an explanatory variable.

4.5.2. Make-up of Customer Base

The make-up of an operator's customer base can have a significant impact upon costs. Variables that seek to capture some of these differences were considered in this study. These variables are:

- The relative numbers of business and residential lines (expressed as the number of business switched lines per residential switched line);
- The relative number of switched and leased lines (expressed as the number of leased lines - both full leased lines and special access lines or PPCs - in 64kb equivalents per switched line); and
- The proportion of national (or intra-LATA) calls (expressed as the number of national calls per other call).

These variables capture differences in the make-up of the customer base of the operators which can impact upon costs and which are largely outside the control of the operators.

4.5.3. Input Prices

In addition to output levels and geographical factors, cost differences between telecommunication operators may reflect differences in the prices they face when purchasing the necessary inputs. Differences in price levels between the US and the UK have already been taken into account through the use of PPPs in converting BT's cost data into US dollars. However, we need to consider further the special cases of:

- capital markets; and
- labour markets.

4.5.3.1. *Cost of capital*

In order to avoid differences in the cost of capital (required rate of return) directly affecting estimated costs and efficiency levels, the same cost of capital has been used for all companies (see section 3.5.8). This common cost of capital is applied to the asset base of each company, and the cost of capital is included in the total cost figure used for each company.

It is theoretically possible that differences in the cost of capital across companies could result in differences in the mix of capital and labour inputs employed which may in turn affect costs. However, in practical terms the opportunities for telecommunications operators to substitute capital for labour (and vice versa) are limited. This is supported by the fact that when the required rate of return for the US LECs was included in the total cost equation it was shown to be insignificant in explaining differences in total costs[39].[40]

4.5.3.2. *Labour costs*

A similar situation exists with labour costs. Their direct impact on costs is already taken into account in the estimate of total costs. However, if the labour cost per employee in common PPP currency differs widely between companies, it is possible that this could cause variations in the chosen mix of capital and labour inputs and hence indirectly affect costs. However as mentioned above telecommunications operators may face limited opportunities for substituting capital and labour.

Therefore, regression analysis was completed to assess whether staff costs per employee have a significant impact upon total costs. The results of this analysis showed staff cost per employee to be insignificant in explaining total costs.

4.6. US Exclusions

US companies were excluded from the analysis, for the following reasons:

- Puerto Rico Telephone Company and GTE Hawaiian Telephone Company Inc both operate networks on islands, and as such would be likely to face substantially

---

[39] Rate of return information was only available for the US LECs for the years 1997, 1998 and 1999. In addition, data was unavailable for a large number of companies, particularly in 1999. Therefore, it is not possible to include the costs of capital for US LECs as a variable in analysis of the post 1999 period.

[40] Additionally, in a study of comparative efficiency it may be inappropriate to include an operator's actual cost of capital as an explanatory variable, as this cost of capital will be, in part, determined by the efficiency with which the operator arranges its financial structure (i.e. its relative level of debt and equity). Therefore, it would be necessary to include an indicative "efficient" cost of capital figure. As the companies in this study all perform comparable activities, one could assume that the cost of capital is constant across the companies. The cost equation is then estimated conditional on this price and, therefore, the cost of capital does not need to be included in the regression equation. An example of a specification of a cost equation which takes this form can be found in Cubbin and Tzanidakis, "Regression versus data envelopment analysis for efficiency measurement: an application to the England and Wales regulated water industry", Utilities Policy, 1998, pp. 75-85.

different cost structures from the remainder of the sample. Initial attempts to take account of these differences through the use of a dummy variable proved to be insufficient, and so both companies were removed from the sample:

- Alliant was excluded from the 1996 dataset, since its costs for that year were not in line with those reported for all other years, and because it did not provide information on employee numbers or compensation for employees.
- Central of Texas was excluded from all years, as it does not provide information on the staff cost per employee.
- Alltel Carolina was excluded from the 1998 and 1999 datasets due to concerns over the validity of its data. Citizens of New York and the Commonwealth Telephone Company were excluded from 2000 and 2001 datasets for the same reason.
- Frontier of Rochester was excluded from all datasets before 2000 due to accounting errors in their reported data.
- Verizon Washington DC was excluded from all datasets because, as in the previous comparative efficiency study completed in 2000, this company was identified as a potential outlier since its cable sheath length per line variable was significantly lower than those of the other companies in the dataset.

5. RESULTS

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

First of all NERA's approach to specifying the total cost model is explained (section 5.1). Then the results of using this model to estimate BT's comparative efficiency are set out (section 5.2).

In response to NERA's model specification, BT developed an alternative cost model. This model is presented in section 5.3 below.

Section 5.4 then compares the two alternative model specifications and sets-out criteria for identifying the most appropriate model for the assessment of BT's comparative efficiency for the purposes of Ofcom's price cap modelling.

5.1. NERA Model Specification

The purpose of the model specification is to identify a total cost model which is not only a good fit with the data but which is also reasonable from the perspective of both economic theory and *a priori* expectations of the key drivers of the costs of telecommunications networks, and the importance of these drivers in explaining costs (i.e. their magnitude).

Multi-year SFA is used to test the model specification as the use of a multi-year technique increases the sample size used in the modelling and therefore produces more statistically robust results, and secondly because the use of a regression analysis technique, instead of a mathematical programming technique such as DEA, allows for statistical testing of the fit of the model to be undertaken.

The three stages involved in the model specification are listed below. Each of these stages is discussed in detail in the next three sub-sections of this report, namely:

- decide on the appropriate functional form to be estimated;
- identify the data set upon which the model is to be specified; and
- test the different specifications to identify the "best" model.

5.1.1. Functional form

In relation to the functional form of the model, both the Cobb-Douglas and Translog functional forms were considered. The Cobb-Douglas functional form is less flexible than its Translog counterpart, which is advocated by many econometricians as it allows the functional form 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 has potentially onerous data requirements.

The data availability for the Translog functional form was investigated, and it was found that insufficient data was available to justify the use of this technique[41]. Consequently, a Cobb-Douglas (log-log) specification[42] was chosen. The equation below is an example of the Cobb-Douglas specification:

$$log(C) = \beta_0 + \beta_1 log(L) + \beta_2 log(M) + \ldots$$

Some of the advantages of using a cost function with this functional form are as follows:

- 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 a equation that is linear in explanatory variables (which is a requirement of regression analysis) and which can be easily interpreted in terms of percentage changes in the variables (i.e. 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).

5.1.2. Data set

At the outset of this stage of the study any data issues, which could impact (or bias) the results of the econometric analysis, were investigated. Issues that are of particular importance are the inclusion, or exclusion, of BT from the sample used to estimate the regression model, and, given the size of BT, the inclusion of the smaller US LECs in the sample.

It was decided not to include BT in the sample used to estimate the regression model, given the possibility that the inclusion of BT could bias the regression results. This is because BT is towards the top end of the sample in terms of operator size (measured in terms of numbers of lines) and therefore its inclusion in the sample could result in the regression line being spuriously bent towards BT.

---

41 When estimating a Translog function it is necessary to simultaneously estimate the demand functions for the inputs to the cost function (e.g. labour and capital), and the total cost function itself. Investigation into the availability of data for the estimation of the input demand functions indicated that it would not be possible to obtain LEC specific data of sufficient reliability to allow the estimation of these functions to take place.

42 In a Cobb-Douglas (log-log) specification, a 1% increase in a cost driver (e.g. an output) raises costs by a fixed percentage, given by the associated model coefficient. For example, if the coefficient on the number of access lines is 0.5, then the model predicts that a 1% increase in the number of access lines will raise costs by 0.5%.

In relation to the size of the US LECs, in 2001 the distribution of the LECs in terms of size (as measured by the number of lines) was such that approximately 40% of the LECs have less than 1 million lines, 40% have between 1 million and 10 million lines and approximately 20% have more than 10 million lines. Given that relatively few of the LECs are at the top end of the distribution in terms of size, this may bias the regression results as, depending upon the relative performance of these companies, it may be possible for some of them to pull the regression line towards themselves.

To reduce this possibility it was decided to exclude the smaller LECs (i.e. those with less than 1 million lines) from the sample when specifying the model.[43] Given the size of the sample used in the model specification, approximately 50 LECs for a six year period, it was possible to exclude the smaller US LECs without significantly affecting the robustness of the results.

This exclusion of the smaller US LECs during the model specification stage was supported by the results from testing for structural stability. The structural stability test assessed whether the regression coefficients are stable in the sense that there is no statistically significant change in the coefficients when the model is estimated using just the sample of smaller US LECs rather than the full sample. The test rejected the hypothesis of the stability of the regression coefficients, thereby indicating that the smaller US LECs should not be included in the sample when multi-year analysis is being completed. Therefore all multi-year analysis was carried out with the smaller US LECs excluded.

Whilst it is possible to exclude the smaller LECs from the multi-year sample without significantly affecting the robustness of the results, given the available sample size, this is not the case with the single-year analysis. Removing the smaller US LECs from the multi-year sample still leaves 174 observations, while removing the smaller US LECs from the single-year sample leaves only 29 observations.

Therefore tests were also completed to assess whether the smaller US LECs could be included in the single-year sample for the year 2001/02. The test completed is the same as that used to assess structural stability in the multi-year sample discussed above. The test identified that, for the year 2001/02 there is structural stability and that therefore the inclusion of the smaller US LECs in the single-year analysis is justified. However, for comparison purposes, the single-year analysis with the smaller LECs excluded was also completed. The results of these estimations are discussed in section 5.2 below.

5.1.3. Testing of different model specifications

A "general to specific" approach was used to establish the final model specification. This involves including a (relatively) large number of potentially significant explanatory

[43] The alternative of excluding the large LECs would not be appropriate given that BT itself falls into this category.

variables in the regression model at the first stage[44]. These variables are then reduced using a step-by-step process to identify those variables which are empirically (rather than just theoretically) significant explanatory variables. Decisions on the step-by-step exclusion of variables were made based were made on the basis of statistical significance[45].

After excluding all insignificant variables, the following explanatory variables were identified for inclusion in the model:

- Lines;
- Switch minutes;
- Average sheath length per switch line (either total sheath length or sheath length split between aerial and non-aerial cable sheath)[46];
- A dummy variable for those LECs located in the North East of the US[47];
- The ratio of the number of leased lines to the number of switch lines; and
- The ratio of the number of national calls to the number of all other calls.

The next step involved in the model specification was to test for the structural stability of the coefficients on these variables throughout the sample period (1996-2001). To test this the equivalence of the regression coefficients for the periods 1996-1998 and 1999-2001 was tested. These tests highlighted that, whilst many of the coefficients are stable over the period, the sheath length per switch line variables (both total sheath and aerial and non-aerial sheath per switch line), the leased line per switch line variable and the constant term were not stable. Therefore, to capture these differences, interaction terms between these variables and the dummy for the period 1996-1998 (and the dummy for 1996-1998 itself) were included in the regression. The interaction terms allow the coefficients on these

44 Potentially significant explanatory variables are identified by considering those variables which economic theory would predict are significant in determining total costs.

45 The insignificance 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).

46 The sheath length variable is expressed on a per switch line basis rather than a per total line basis (and as an average figure) to control for the trend in the growth in special access lines towards the end of the sample period (note this trend will be captured by the total lines and leased lines per switch line variables). However, as this might conceivably over estimate the efficiency of those operators with relatively more special access and leased lines, we have undertaken sensitivities to the specification of this variable. In section 5.4.5 below we show the impact on BT's inefficiency score of using average sheath length expressed on a per total lines basis and of using total sheath length alone, rather than average total sheath length per switch line. These sensitivity tests show that the specification of the sheath length variable does not have a significant impact upon BT's inefficiency score.

47 Companies located in the North East of the US face different price levels, particularly for labour, to those in the rest of the US. Therefore a dummy was included to capture the impact of this difference on the total cost function.

variables for the period 1996-1998 to differ from those in 1999-2001, thereby allowing all six years of data to be included in the model[48].

The final stage in the model specification involved choosing between the different possible sheath per line variables, namely total sheath per line, aerial sheath per line and non-aerial sheath per line.

Ideally the aerial and non-aerial sheath per line variables would be used, as this would capture not only differences in population dispersion but also environmental differences which are reflected in an operator's choice of aerial or non-aerial cable. However, the estimations showed that, whilst the non-aerial sheath length per switch line variable was always significant, the aerial sheath length per switch line variable was not. Hence the use of total sheath length per line was considered.

Regressions were estimated using firstly a total sheath length per line variable alone, and secondly using this variable alongside the ratio of aerial to non-aerial sheath length. Comparing the log likelihoods of these regressions and the significance of all of the explanatory variables included in them, it was found that the regression with the total sheath length variable alone represented a better fit than either the regression with aerial and non-aerial sheath length or the regression with total sheath length and the ratio of aerial to non-aerial sheath. Therefore, the following model specification was chosen:

$$log(C) = \beta_0 + \beta_1 log(L) + \beta_2 log(M) + \beta_3 log(TS/SL) + \beta_4 NED + \beta_5 T + \beta_6 log(LL/SL) + \beta_7 log(NM/OM) + \beta_8 D9698(TS/SL) + \beta_9 D9698(LL/SL) + \beta_{10} D9698$$

Where:

- *C* = total cost
- *L* = total lines
- *M* = switch minutes
- *TS/SL* = average total sheath length per switch line
- *NED* = North East US dummy
- *T* = Time
- *LL/SL* = leased lines / switch lines
- *NM/OM* = national minutes / other minutes

48 Without these interaction terms it would have only been possible to estimate the model using three years of data, even though there is stability over the full period in terms of some of the models most influential variables in terms of the magnitude of the coefficients (for example lines and switch minutes). This halving of the available sample size would be expected to significantly impact upon the robustness of the results.

- *D9698*log(TS/SL)* = interaction dummy for average total sheath length per switch line in 1996-1998
- *D9698*log(LL/SL)* = interaction dummy for leased lines/switch lines in 1996-1998
- *D9698* = shift dummy for the period 1996-1998.

Table 5.1
Final Model Specification

	Coefficient	Standard Error	T Statistic	Significance
Log of lines	0.82	0.04	20.24	0.00
Log of switch minutes	0.15	0.03	4.83	0.00
Log of average total sheath length per switch line	0.13	0.05	2.62	0.01
North East US dummy	0.21	0.06	3.76	0.00
Time	0.01	0.01	0.95	0.34
Log of leased lines per switch line	-0.23	0.02	-10.67	0.00
Log of national calls per other calls	0.04	0.01	3.55	0.00
Interaction dummy for log of average total sheath per switch line in the period 1996-1998	0.09	0.02	5.22	0.00
Interaction dummy for log of leased lines per switch line in the period 1996-1998	0.09	0.02	5.72	0.00
Shift dummy for the period 1996-1998	0.39	0.06	6.77	0.00
Constant term	5.08	0.46	11.13	0.00
Log Likelihood	**238.91**			

Source: NERA analysis

In this model all of the explanatory variables, with the exception of the time variable, are significant at the 5% level, and all the explanatory variables have the expected sign.

The insignificance of the time variable and the magnitude of the coefficient on this variable, suggest that there are no systematic changes in costs over time, which are not picked up by the other explanatory variables in the regression. It does not suggest in itself that there is no movement in costs over time. This is a useful indicator of the robustness of the model.

The shift dummy for the period 1996-1998 shows that the constant term for this period is 0.39 higher than the constant term for the period 1999-2001 of 5.08. This suggests that the fixed costs of the LECs are slightly higher in the 1996-1998 period than in the period 1999-2001, however, the magnitude of the difference is not great. Possible explanations for this difference could be relatively small changes in the reporting requirements of the US LECs. However, by including the dummy for the 1996-1998 period we are controlling for this difference.

5.2. Comparative Efficiency Assessment Using the NERA Model

After specifying the final model, it was possible to carry out the comparative assessment of BT's efficiency. To do so, the final model specification was estimated using a variety of different techniques, namely:

- Multi-year SFA (1996-2001);
- Multi-year least squares (1996-2001);
- Single-year OLS (2001);
- Single-year SFA (2001); and
- Single-year DEA (2001).

The results of each of techniques are presented in the following five sub-sections of this report.

5.2.1. Multi-year SFA

The multi-year SFA was completed using data for only the larger US LECs. As explained in section 5.1.2 above, tests were completed for the inclusion of the smaller US LECs in the multi-year sample and these tests rejected the structural stability of the regression coefficients between the smaller and larger US LECs. Therefore, it is not valid to include the smaller US LECs in the sample when completing multi-year analysis.

The multi-year SFA was estimated assuming a truncated-normal distribution for the inefficiency term and a normal distribution for the idiosyncratic error term. The only distributional assumption available in the statistical package for running the multi-year SFA is the truncated-normal assumption. However, as the multi-year SFA is better able than the single-year SFA to decompose the residual term given the presence of repeat observations for each operator, this limitation on the number of distributional assumption of the error term that can be tested is of limited importance. For the single-year SFA analysis, we estimated regressions using the half-normal, exponential and truncated normal distributions (see section 5.2.4 for a discussion of these estimations).

The results of the multi-year SFA regression have already been presented in Table 5.1.

The efficiency results estimated using this regression for the US LECs and for BT are shown in the table below.

Table 5.2
Multi-year SFA Results

Rank	Company	Residual	% Inefficiency from Frontier	% Inefficiency from Best Performer	% Inefficiency from Decile	% Inefficiency from Median
1	WISCONSIN BELL	0.03	3%	0%	-6%	-20%
2	S NEW ENGLAND TEL CO	0.04	4%	1%	-5%	-19%
3	INDIANA BELL	0.07	7%	4%	-2%	-17%
4	BA DELAWARE	0.09	9%	6%	0%	-15%
5	MICHIGAN BELL	0.13	14%	11%	5%	-11%
6	VERIZON NORTH	0.15	16%	13%	6%	-10%
7	CENTRAL TEL CO	0.15	16%	13%	6%	-10%
8	OHIO BELL	0.17	19%	15%	8%	-8%
9	BT	0.17	19%	16%	9%	-8%
10	PACIFIC BELL	0.22	24%	21%	13%	-4%
11	CAROLINA TEL CO	0.22	24%	21%	14%	-4%
12	VERIZON NEW ENGLAND	0.22	25%	22%	14%	-3%
13	ILLINOIS BELL	0.24	27%	23%	16%	-2%
14	SPRINT FLORIDA	0.24	27%	23%	16%	-2%
15	UNITED OF OHIO	0.25	29%	25%	18%	0%
16	VERIZON PENNSYLVANIA	0.26	30%	27%	19%	1%
17	BA MARYLAND	0.27	31%	28%	20%	2%
18	VERIZON NEW JERSEY	0.28	32%	28%	21%	2%
19	CINCINNATI BELL	0.28	33%	29%	21%	3%
20	QWEST	0.30	35%	31%	23%	5%
21	BELLSOUTH	0.30	35%	32%	24%	5%
22	VERIZON NORTHWEST	0.31	36%	32%	24%	5%
23	VERIZON SOUTH	0.31	36%	32%	24%	5%
24	VERIZON WEST VIRGINIA	0.32	37%	33%	25%	6%
25	VERIZON VIRGINIA	0.33	39%	35%	27%	8%
26	VERIZON CALIFORNIA	0.34	41%	37%	29%	9%
27	SW BELL	0.35	42%	38%	30%	10%
28	VERIZON FLORIDA	0.36	43%	39%	30%	11%
29	VERIZON SOUTHWEST	0.38	46%	42%	33%	13%
30	BA NEW YORK	0.48	62%	57%	48%	26%

Source: NERA analysis

Note a negative result indicates efficiency compared to the reference point and a positive result indicates inefficiency. Therefore, Wisconsin Bell is 3% inefficient when compared to the frontier and 6% efficient when compared to the upper decile point (i.e. is more efficient than the upper decile point).

Results are shown both for performance relative to the frontier and for performance relative to the upper decile point in the sample. The comparison to the upper decile is shown to ensure that the comparison is against an achievable level of efficiency. This is because the SFA technique, by attempting to fit an envelope around the most efficient observations, could produce a frontier that is biased outwards if, due to any data or estimation problems (for example unobservable heterogeneity), the efficient observations appear spuriously efficient. As the SFA regression estimated here follows the time invariant specification[49] (see section 2.3.1 above for a description of time invariant and time variant specifications), any spurious observations for an operator in a sub-set of the years in the sample will be controlled for in the time invariant result. Therefore, by comparing efficiency relative to the upper decile of the sample, we can be certain that we are comparing against an achievable level of efficiency.

As the regression was estimated using data for the US LECs only (see section 5.1.2 above) it was necessary to compute the decomposition of the regression residual for BT between inefficiency and idiosyncratic error. This computation was completed using the formulation set out in "Stochastic Frontier Analysis" by S. C. Kumbhakar and C. A. Knox Lovell[50].

5.2.2. Multi-year least squares

The multi-year least squares analysis was completed using data for only the larger US LECs. As explained in section 5.1.2 above, tests were completed for the inclusion of the smaller US LECs in the multi-year sample and these tests rejected the structural stability of the regression coefficients between the smaller and larger US LECs. Therefore, it is not valid to include the smaller US LECs in the sample when completing multi-year analysis.

In the table below the estimated multi-year least squares regression is documented:

Table 5.3
Multi-year Least Squares

	Coefficient	Standard Error	T Statistic	Significance
Log of lines	0.82	0.04	20.50	0.00
Log of switch minutes	0.15	0.03	4.81	0.00
Log of average total sheath length per switch line	0.13	0.05	2.67	0.01
North East US dummy	0.21	0.07	3.06	0.00
Time	0.01	0.01	0.94	0.35
Log of leased lines per switch line	-0.23	0.02	-10.74	0.00
Log of national calls per other calls	0.04	0.01	3.56	0.00
Interaction dummy for log of average total sheath per switch line in the period 1996-1998	0.09	0.02	5.22	0.00
Interaction dummy for log of leased lines per switch line in the period 1996-1998	0.09	0.02	5.73	0.00
Shift dummy for the period 1996-1998	0.39	0.06	6.77	0.00
Constant term	5.30	0.45	11.85	0.00
Log Likelihood	**238.62**			

Source: NERA analysis

The efficiency results estimated using this regression for the US LECs and for BT are shown in the table below.

49 Both time variant and time invariant specifications were estimated and tested. These tests showed the time invariant specification to be the appropriate specification for this study.

50 S. C. Kumbhakar and C. A. Knox Lovell, 2000; "Stochastic Frontier Analysis; Cambridge University Press.

Table 5.4
Multi-year Least Squares Results

Rank	Company	Residual	% Inefficiency from Best Performer	% Inefficiency form Decile	% Inefficiency from Average	% Inefficiency from Median
1	WISCONSIN BELL	-0.22	0%	-5%	-20%	-21%
2	S NEW ENGLAND TEL CO	-0.21	1%	-4%	-19%	-20%
3	INDIANA BELL	-0.18	4%	-1%	-17%	-18%
4	BA DELAWARE	-0.16	7%	1%	-15%	-16%
5	MICHIGAN BELL	-0.11	12%	6%	-11%	-12%
6	CENTRAL TEL CO	-0.10	13%	7%	-9%	-11%
7	VERIZON NORTH	-0.10	14%	7%	-9%	-11%
8	OHIO BELL	-0.08	16%	10%	-7%	-9%
9	BT	-0.08	16%	10%	-7%	-9%
10	VERIZON NEW ENGLAND	-0.03	21%	15%	-3%	-5%
11	CAROLINA TEL CO	-0.03	22%	15%	-3%	-4%
12	PACIFIC BELL	-0.02	22%	16%	-2%	-4%
13	ILLINOIS BELL	-0.01	24%	18%	-1%	-2%
14	SPRINT FLORIDA	-0.01	24%	18%	-1%	-2%
15	UNITED OF OHIO	0.01	27%	20%	1%	0%
16	VERIZON PENNSYLVANIA	0.02	28%	21%	2%	0%
17	BA MARYLAND	0.03	29%	22%	3%	1%
18	VERIZON NEW JERSEY	0.03	29%	22%	3%	1%
19	CINCINNATI BELL	0.04	30%	23%	4%	2%
20	VERIZON NORTHWEST	0.06	33%	26%	6%	4%
21	QWEST	0.06	33%	26%	6%	5%
22	BELLSOUTH	0.06	33%	26%	7%	5%
23	VERIZON SOUTH	0.06	33%	26%	7%	5%
24	VERIZON WEST VIRGINIA	0.08	35%	28%	8%	6%
25	VERIZON VIRGINIA	0.09	36%	29%	9%	7%
26	VERIZON CALIFORNIA	0.10	38%	31%	11%	9%
27	VERIZON FLORIDA	0.11	40%	32%	12%	10%
28	SW BELL	0.11	40%	33%	12%	10%
29	VERIZON SOUTHWEST	0.14	44%	36%	15%	13%
30	BA NEW YORK	0.24	59%	51%	27%	25%

Source: NERA analysis

As above, a negative result indicates efficiency compared to the reference point and a positive result indicates inefficiency. Therefore, in the above table, Central Tel Co is 9% efficient when compared to the average and is 7% inefficient when compared to the upper decile point.

5.2.3. Single-year OLS

The single-year OLS analysis was completed for the year 2001/02. Testing for structural stability of the regression coefficients between the smaller and larger US LECs in this one year showed that there was structural stability and therefore that it is valid to include the smaller LECs in the sample for single-year analysis.

Presented below are the results of the OLS regression for 2001/02 using the full sample. For comparison purposes we also present the results of the OLS regression estimated with the smaller LECs excluded.

Table 5.5
Single-year OLS Full Sample

	Coefficient	Standard Error	T Statistic	Significance
Log of lines	0.91	0.09	9.73	0.00
Log of switch minutes	0.08	0.08	1.08	0.29
Log of average total sheath length per switch line	0.12	0.05	2.34	0.02
North East US dummy	0.22	0.08	2.85	0.01
Log of leased lines per switch line	-0.28	0.05	-5.88	0.00
Log of national calls per other calls	-0.01	0.02	-0.49	0.63
Constant term	5.23	0.56	9.33	0.00
R^2	**0.993**			

Source: NERA analysis

Table 5.6
Single-year OLS Large LECs Only

	Coefficient	Standard Error	T Statistic	Significance
Log of lines	0.89	0.10	8.49	0.00
Log of switch minutes	0.14	0.09	1.63	0.12
Log of average total sheath length per switch line	0.08	0.05	1.53	0.14
North East US dummy	0.21	0.07	2.95	0.01
Log of leased lines per switch line	-0.32	0.07	-4.83	0.00
Log of national calls per other calls	-0.02	0.02	-0.99	0.33
Constant term	3.99	0.76	5.25	0.00
R^2	**0.993**			

Source: NERA analysis

The results presented above show that there is some broad consistency between the regression coefficients in the two samples.

The efficiency results estimated using these regressions for the US LECs and for BT are shown in the tables bellow, firstly for the full sample and secondly for the large LECs only sample.

Table 5.7
Single-year OLS Full Sample Results

Rank	Company	Residual	% Inefficiency from Best Performer	% Inefficiency form Decile	% Inefficiency from Average	% Inefficiency from Median
1	CENTRAL TEL CO	-0.23	0%	-9%	-20%	-20%
2	INDIANA BELL	-0.18	5%	-4%	-16%	-16%
3	ALLTEL GEORGIA	-0.17	6%	-4%	-16%	-15%
4	WISCONSIN BELL	-0.16	7%	-2%	-15%	-14%
5	SPRINT FLORIDA	-0.14	9%	0%	-13%	-12%
6	VERIZON MIDWEST	-0.13	10%	1%	-12%	-12%
7	UNITED OF NEW JERSEY	-0.12	11%	2%	-11%	-11%
8	S NEW ENGLAND TEL CO	-0.12	12%	2%	-11%	-11%
9	VERIZON NORTH	-0.11	12%	2%	-11%	-10%
10	CONTEL OF THE SOUTH	-0.11	12%	2%	-11%	-10%
11	CAROLINA TEL CO	-0.10	13%	3%	-10%	-9%
12	UNITED SOUTHEAST	-0.10	14%	4%	-9%	-9%
13	FRONTIER OF ROCHESTER	-0.09	14%	4%	-9%	-8%
14	OHIO BELL	-0.06	18%	8%	-6%	-5%
15	BA DELAWARE	-0.06	18%	8%	-6%	-5%
16	UNITED OF PENNSYLVANIA	-0.06	19%	8%	-5%	-5%
17	UNITED OF OHIO	-0.05	19%	9%	-5%	-4%
18	BT	-0.05	20%	9%	-5%	-4%
19	ALLTEL PENNSYLVANIA	-0.01	24%	13%	-1%	-1%
20	UNITED OF INDIANA	-0.01	24%	14%	-1%	0%
21	VERIZON NORTHWEST	-0.01	25%	14%	-1%	0%
22	SPRINT MISSOURI	-0.01	25%	14%	-1%	0%
23	BA MARYLAND	-0.01	25%	14%	-1%	0%
24	VERIZON SOUTH	-0.01	25%	14%	-1%	0%
25	ALLTEL CAROLINA	0.00	25%	14%	0%	0%
26	ILLINOIS BELL	0.00	25%	15%	0%	1%
27	MICHIGAN BELL	0.02	27%	16%	2%	2%
28	VERIZON NEW JERSEY	0.02	28%	17%	2%	3%
29	CENTRAL OF VIRGINIA	0.02	29%	17%	2%	3%
30	VERIZON NEW ENGLAND	0.03	29%	18%	3%	4%
31	VERIZON PENNSYLVANIA	0.04	31%	20%	4%	5%
32	VERIZON FLORIDA	0.04	31%	20%	4%	5%
33	VERIZON CALIFORNIA	0.04	31%	20%	5%	5%
34	BELLSOUTH	0.05	31%	20%	5%	5%
35	VERIZON VIRGINIA	0.05	32%	20%	5%	6%
36	ALLIANT	0.05	32%	21%	5%	6%
37	QWEST	0.06	33%	22%	6%	7%
38	CINCINNATI BELL	0.07	34%	22%	7%	7%
39	PACIFIC BELL	0.07	35%	23%	7%	8%
40	NEVADA BELL	0.07	35%	23%	8%	8%
41	WESTERN RESERVE	0.11	40%	27%	11%	12%
42	SW BELL	0.11	40%	28%	12%	12%
43	VERIZON WEST VIRGINIA	0.11	40%	28%	12%	13%
44	VERIZON SOUTHWEST	0.13	42%	30%	13%	14%
45	BA NEW YORK	0.18	50%	37%	19%	20%
46	UNITED OF TEXAS	0.33	74%	59%	39%	39%
47	UNITED OF NORTHWEST	0.43	92%	76%	53%	54%

Source: NERA analysis

Table 5.8
Single-year OLS Large LECs Sample Results

Rank	Company	Residual	% Inefficiency from Best Performer	% Inefficiency from Decile	% Inefficiency from Average	% Inefficiency from Median
1	S NEW ENGLAND TEL CO	-0.17	0%	-4%	-16%	-16%
2	CENTRAL TEL CO	-0.16	1%	-3%	-15%	-15%
3	VERIZON NORTH	-0.13	4%	0%	-12%	-12%
4	INDIANA BELL	-0.12	5%	0%	-12%	-12%
5	WISCONSIN BELL	-0.10	7%	3%	-10%	-10%
6	BT	-0.10	7%	3%	-10%	-10%
7	SPRINT FLORIDA	-0.09	9%	4%	-9%	-9%
8	CAROLINA TEL CO	-0.04	14%	9%	-4%	-4%
9	OHIO BELL	-0.04	14%	9%	-4%	-4%
10	BA MARYLAND	-0.01	18%	13%	-1%	-1%
11	VERIZON SOUTH	-0.01	18%	13%	-1%	-1%
12	ILLINOIS BELL	-0.01	18%	13%	-1%	-1%
13	PACIFIC BELL	-0.01	18%	13%	-1%	-1%
14	VERIZON NEW JERSEY	0.00	18%	13%	0%	0%
15	VERIZON CALIFORNIA	0.00	19%	14%	0%	0%
16	BELLSOUTH	0.00	19%	13%	0%	0%
17	QWEST	0.00	19%	14%	0%	0%
18	BA DELAWARE	0.01	21%	15%	1%	1%
19	VERIZON PENNSYLVANIA	0.02	21%	16%	2%	2%
20	SW BELL	0.03	23%	18%	4%	3%
21	VERIZON NEW ENGLAND	0.04	23%	18%	4%	4%
22	VERIZON FLORIDA	0.04	24%	19%	5%	4%
23	VERIZON NORTHWEST	0.05	25%	19%	5%	5%
24	MICHIGAN BELL	0.05	25%	20%	5%	5%
25	UNITED OF OHIO	0.06	26%	20%	6%	6%
26	VERIZON VIRGINIA	0.06	26%	21%	6%	6%
27	CINCINNATI BELL	0.07	27%	22%	7%	7%
28	BA NEW YORK	0.13	36%	30%	14%	14%
29	VERIZON SOUTHWEST	0.16	39%	33%	17%	17%
30	VERIZON WEST VIRGINIA	0.16	40%	34%	18%	18%

Source: NERA analysis

Note a negative result indicates efficiency compared to the reference point and a positive result indicated inefficiency. Therefore, in the above table Carolina Tel Co is 4% efficient when compared to the average and is 9% inefficient when compared to the upper decile point.

The results of the single-year OLS (large LECs only sample) regression show BT to be in the region of 3% inefficient compared to the upper decile. This is different to the results obtained under both the multi-year techniques, which show BT's inefficiency to be in the region of 9% to 10%. However, the rankings of BT under these different techniques are broadly similar, as BT is ranked 11th out of the 30 firms under the single-year OLS (large LECs only sample) and is ranked 9th out of the 30 firms under both of the multi-year techniques.

The results for BT under the single-year OLS (full sample) regression are broadly consistent with the efficiency results obtained under the multi-year techniques. BT is 9% inefficient (compared to the upper decile points) and is ranked 18th out of the 47 firms in the full sample.

It is worth noting at this point, however, that the results of both of the single-year regressions are based on considerably smaller samples than the multi-year regressions. The sample size used for the single-year OLS (large LECs only sample) is only 29, and that for the single-year OLS (full sample) is only 46, compared to 174 for the multi-year techniques. Hence the results of both of the single-year OLS regressions will be less reliable than those of the multi-year techniques.

5.2.4. Single-year SFA

The single-year SFA analysis was completed for the year 2001/02. As mentioned above, testing for structural stability of the regression coefficients between the smaller and larger US LECs in this one year showed that there was structural stability between the two groups of LECs and therefore that it is valid to include the smaller LECs in the sample for single-year analysis.

Therefore single-year SFA models were estimated for both the full sample, and for comparison purposes, with only the larger LECs included. With both samples, estimation was completed separately under the assumptions that the inefficiency term follows a half-normal, exponential and lastly truncated normal distribution.

With the full sample there were estimation problems[51] with the truncated-normal and half-normal distributions but SFA was able to decompose the residual using the exponential distribution. With the large LECs only sample there were estimation problems with all three distributions, therefore SFA was unable to decompose the residual for the large LECs only sample.

Therefore, presented below are the results of the SFA for 2001/02 using the full sample for the exponential distribution.

---

51 Estimation problems encountered include non-concavity of the likelihood function and the inability of the regression to fully estimate the parameters of the distribution of the inefficiency term.

Table 5.9
Single-year SFA Full Sample - Exponential

	Coefficient	Standard Error	T Statistic	Significance
Log of lines	0.88	0.08	11.25	0.00
Log of switch minutes	0.13	0.06	1.96	0.05
Log of average total sheath length per switch line	0.10	0.04	2.49	0.01
North East US dummy	0.20	0.06	3.13	0.00
Log of leased lines per switch line	-0.26	0.03	-7.83	0.00
Log of national calls per other calls	0.00	0.01	0.13	0.90
Constant term	4.64	0.50	9.33	0.00
Gamma (s^2_U/s^2)	0.66			
Log Likelihood	36.36			

Source: NERA analysis

The gamma reported for the exponential distribution indicates that this distribution attributes 66% of the residual to inefficiency.

The efficiency results using single-year SFA are shown below.

Table 5.10
Single-year SFA Full Sample – Exponential Results

Rank	Company	Residual	% Inefficiency from Frontier	% Inefficiency from Best Performer	% Inefficiency form Decile	% Inefficiency from Median
1	CENTRAL TEL CO	0.02	2%	0%	-1%	-4%
2	INDIANA BELL	0.03	3%	0%	-1%	-4%
3	WISCONSIN BELL	0.03	3%	0%	-1%	-4%
4	ALLTEL GEORGIA	0.03	3%	1%	0%	-4%
5	S NEW ENGLAND TEL CO	0.03	3%	1%	0%	-3%
6	VERIZON NORTH	0.03	3%	1%	0%	-3%
7	VERIZON MIDWEST	0.03	3%	1%	0%	-3%
8	SPRINT FLORIDA	0.04	4%	1%	0%	-3%
9	BT	0.04	4%	1%	1%	-3%
10	UNITED OF NEW JERSEY	0.04	4%	1%	1%	-3%
11	CONTEL OF THE SOUTH	0.04	4%	2%	1%	-3%
12	UNITED SOUTHEAST	0.04	4%	2%	1%	-2%
13	OHIO BELL	0.04	4%	2%	1%	-2%
14	CAROLINA TEL CO	0.04	5%	2%	1%	-2%
15	BA DELAWARE	0.05	5%	2%	1%	-2%
16	FRONTIER OF ROCHESTER	0.05	6%	3%	2%	-1%
17	UNITED OF PENNSYLVANIA	0.06	6%	3%	3%	-1%
18	ILLINOIS BELL	0.06	6%	4%	3%	-1%
19	UNITED OF OHIO	0.06	6%	4%	3%	0%
20	BA MARYLAND	0.06	7%	4%	3%	0%
21	VERIZON NEW JERSEY	0.07	7%	4%	3%	0%
22	VERIZON SOUTH	0.07	7%	4%	4%	0%
23	MICHIGAN BELL	0.07	7%	4%	4%	0%
24	VERIZON NORTHWEST	0.07	7%	4%	4%	0%
25	PACIFIC BELL	0.07	7%	5%	4%	1%
26	VERIZON PENNSYLVANIA	0.08	8%	5%	5%	1%
27	VERIZON CALIFORNIA	0.08	8%	5%	5%	1%
28	BELLSOUTH	0.08	8%	6%	5%	1%
29	ALLTEL PENNSYLVANIA	0.08	8%	6%	5%	1%
30	ALLTEL CAROLINA	0.08	8%	6%	5%	1%
31	UNITED OF INDIANA	0.08	9%	6%	5%	2%
32	SPRINT MISSOURI	0.09	9%	7%	6%	2%
33	VERIZON NEW ENGLAND	0.09	9%	7%	6%	2%
34	QWEST	0.10	10%	8%	7%	3%
35	VERIZON VIRGINIA	0.10	10%	8%	7%	3%
36	VERIZON FLORIDA	0.10	11%	8%	7%	3%
37	SW BELL	0.11	12%	9%	8%	4%
38	NEVADA BELL	0.12	13%	10%	10%	6%
39	CENTRAL OF VIRGINIA	0.12	13%	10%	10%	6%
40	CINCINNATI BELL	0.13	14%	11%	10%	6%
41	ALLIANT	0.13	14%	12%	11%	7%
42	VERIZON WEST VIRGINIA	0.16	17%	15%	14%	10%
43	VERIZON SOUTHWEST	0.17	19%	16%	15%	11%
44	WESTERN RESERVE	0.19	21%	18%	17%	13%
45	BA NEW YORK	0.23	26%	23%	22%	18%
46	UNITED OF TEXAS	0.43	53%	49%	48%	43%
47	UNITED OF NORTHWEST	0.51	66%	62%	61%	55%

Source: NERA analysis

A negative result indicates efficiency compared to the reference point and a positive result indicates inefficiency. Therefore, in the above table Indiana Bell is 3% inefficient when compared to the frontier and is 1% efficient when compared to the upper decile point.

BT's efficiency score is very different under this regression. BT is 1% inefficient under the single-year SFA (full sample) regression and is ranked 9th in the sample, this compares to 9% and a rank of 18th under the comparable OLS regression.

5.2.5. Single-year DEA

DEA analysis involves using a mathematical programme, rather than statistical, techniques to assess comparative efficiency. DEA operates by searching for a 'least cost peer group' of comparator companies for BT. 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 BT. If such a 'peer group' exists, and the linear combination of their costs is lower than that of BT, this cost difference is assumed to be attributable to inefficiency on the part of BT.

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.

DEA analysis was completed for the year 2001/02 both for the full sample of LECs and with the smaller LECs excluded.

BT's efficiency result and its identified peers for both samples are set out in the table below.

Table 5.11
Single-year DEA

Sample	BT Inefficiency Score	Peers
Full sample	17%	South Western Bell, Alltel Pennsylvania, Central Tel Co
Large LECs only	13%	South Western Bell, Indiana Bell

Source: NERA analysis

There are two particular aspects of the DEA analysis which warrant further investigation. The first is the weight DEA attributes to the different explanatory variables included in the analysis. The second is they way the DEA specification accounts for both constant and variable returns to scale. Both of these aspects of the analysis are discussed below.

To assess the impact that each of the explanatory variables from the regression analysis has upon the DEA results, we completed DEA analysis with the non-output variables excluded (i.e. without the total sheath length, leased lines per switch line, national calls per other calls and North East US dummy variables). This analysis produced identical results to those presented in Table 5.11 above. This indicates that, when assessing BT's efficiency, DEA adopts a weighting structure which places no importance on the environmental variables. As the DEA analysis seeks to identify the particular inputs and outputs which show an observation in the best possible light, this means that, in the case of BT, accounting for environmental differences does not show BT to be more efficient.

Independent of the variables included in the analysis, and the weights DEA places on these variables, DEA analysis can also be completed assuming either constant or variable returns to scale. As analysis of the regression results suggests the presence of constant returns to scale in the sample (as tests for the sum of the coefficients on total lines and switch minutes summing to one are generally accepted) we first completed the DEA analysis using a constant returns to scale specification. The results of this specification are those presented in Table 5.11 above. However, to assess the impact of this assumption we also completed DEA analysis assuming variable returns to scale.

DEA analysis was unable to find any peers for BT under a variable returns to scale specification. The conclusions that can be drawn from this are uncertain, as there are two possible explanations for why DEA cannot find peers for a company:

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

One method for helping to distinguishing between these two conclusions is to identify whether a firm that has no peers is itself within the lowest cost peer group of another firm within the sample. If a firm is included within the lowest cost peer group of another firm, then it must be possible to draw a comparison between that firm and other firms within the sample. Hence, it is likely that the firm is actually 100% efficient[52].

Analysis of whether BT is used as a peer for any of the other firms within the sample found that BT is never used as a peer for another firm. Therefore, it is not possible to use this information to distinguish between the two possible explanations for BT's DEA score under the variable returns to scale specification. However, as the results of the regression analysis generally suggest that BT is not one of the most efficient firms in the sample, it is reasonable to assume that the explanation that DEA was unable to find comparable firms is the most likely explanation.

---

[52] However, please note that the failure to find a firm within the lowest cost peer group of other firms does not mean that it is not 100% efficient, it just means that no assessment of its comparative efficiency can be made.

Therefore, on can conclude that the DEA analysis shows BT to be in the region of 13% to 17% inefficient.

5.2.6. Summary of Efficiency Results

In the five sub-sections above, the results of the comparative efficiency assessment of BT have been presented for a variety of different techniques and, for some techniques, using different samples. In this section we compare these different results and identify any consistency between the results of the different techniques.

Table 5.12
Summary of BT's Comparative Efficiency

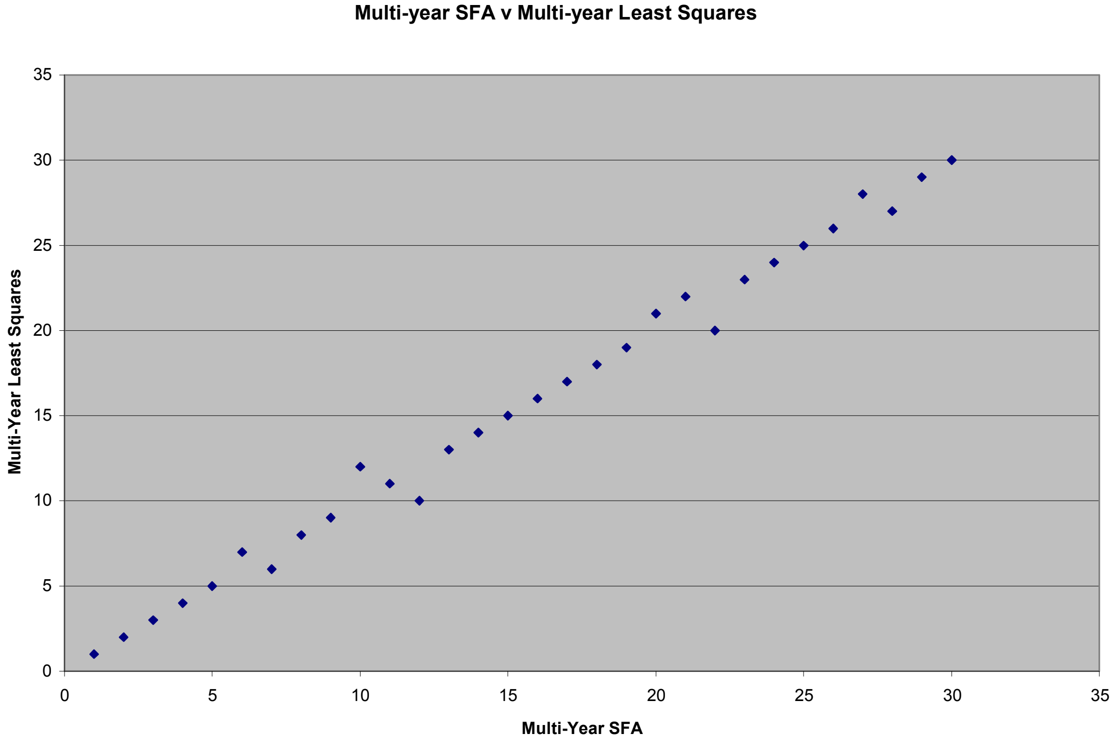
Point of Comparison	Full Sample		Large LECs only	
	Best Performer	Upper Decile	Best Performer	Upper Decile
Multi-year SFA	n.a.	n.a.	16%	9%
Multi-year least squares	n.a.	n.a.	16%	10%
Single-year OLS	20%	9%	7%	3%
Single-year SFA - Exponential	1%	1%	n.a.	n.a.
Single-year DEA (compared against peer group)	17%	n.a.	13%	n.a.
Range of All Results	1%-20%	1%-9%	7%-16%	3%-10%
Range Excluding Single-year Regressions	**17%**	**n.a.**	**13%-16%**	**9%-10%**

Source: NERA analysis

The summary table above shows that there is variation between the results of the different comparative efficiency measurement techniques. In order to analyse the reasons behind this variation, we first compared BT's efficiency ranking under each of these techniques. This analysis is limited to comparison of the rankings under the regression analysis techniques as it is not possible to construct comparable rankings for the DEA results, because the DEA technique identifies different peer groups for different operators and therefore it is not possible to produce an overall ranking of all the operators.

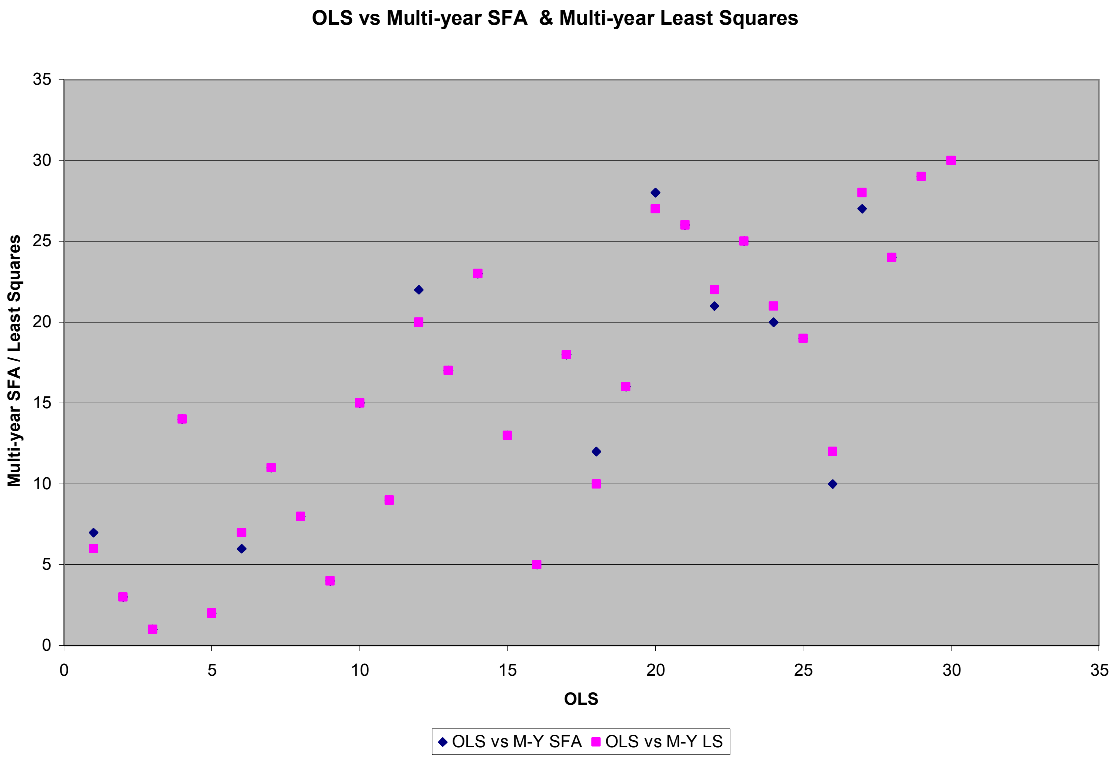
First of all, in the diagram below, the rankings of BT and the US LECs under the two multi-year techniques have been plotted. This diagram shows that there is a very high degree of correlation between the rankings of the operators under the two multi-year techniques. The correlation coefficient of the rankings under these two techniques is 0.99.

Figure 5.1
Multi-year SFA vs Multi-year Least Square



We next compared the rankings of the two different multi year techniques and the OLS single-year (large LECs sample) regression.

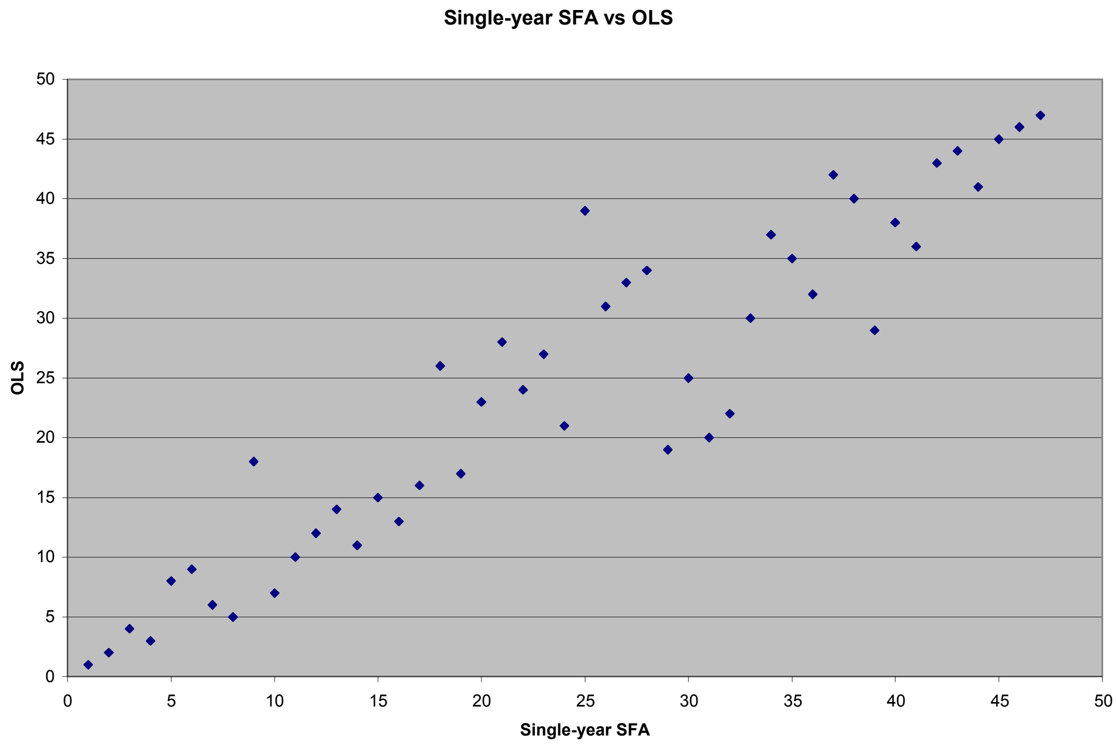
Figure 5.2
Multi-year SFA and Multi-year Least Squares vs. Single-year OLS (Large LECs)



As can be seen from the charts above, there is far less correlation between the rankings under the OLS (Large LECs) regression and those under the multi-year techniques than there is between the rankings of the two multi-year techniques. The correlation between the OLS (Large LECs) regression rankings and the multi-year SFA rankings is only 0.78 whilst the correlation between the OLS (Large LECs) and multi-year least squares rankings is only 0.80. As mentioned above, however, the correlation between the rankings of the two multi-year techniques is 0.99. This suggests that, overall, the OLS (Large LECs) regression presents a somewhat different picture of relative efficiency levels to that presented by the two multi-year techniques.

Finally we compared the rankings under the single-year SFA and the single-year OLS (both of which were estimated using the full sample of the US LECs).

Figure 5.3
Single-year SFA vs Single-year OLS Full Sample



This shows that there is a relatively high degree of correlation between the efficiency rankings produced by these two techniques, albeit still a lower degree of correlation than between the two multi-year techniques. The correlation coefficient between the two rankings is 0.93 and BT's ranking is 9th under the single-year SFA and 18th under the single-year OLS.

This analysis suggests that whilst the rankings of the two single-year full sample regression techniques are broadly consistent for most firms, there is significant variation in their assessment of BT. For example, the single-year SFA and single-year OLS full sample results rank BT as 9th and 18th respectively out of the 47 companies in the full sample. Additionally, BT's level of inefficiency is 1% and 9% respectively (compared to the upper decile) using these two techniques.

The overall picture of the comparison of the rankings under the different techniques is that the highest degree of correlation is between the two multi-year techniques, but that the single-year SFA and single-year OLS (both estimated using all of the US LECs) also exhibit a high degree of correlation.

Whilst there is consistency between the rankings of the majority of the firms under the two multi-year techniques and under the single-year SFA and single-year OLS (both estimated using all of the US LECs) there is little consistency between these two sets of rankings (i.e.

between the rankings of the two multi-year techniques and the rankings of the two single-year full sample techniques). However, this inconsistency could be due to the fact that these two sets of techniques are estimated using different samples, one using only the large LECs and the other using the full sample of LECs.

The results of all the single-year techniques should be treated with some caution as these regressions were estimated using relatively small samples. The single-year OLS (estimated using only the large LECs) is based on a sample size of 29, the single-year SFA and single-year OLS (estimated using the full sample) are based on a sample size of 46, however, the multi-year regressions are both based on a sample of 174. Hence, the results of the single-year regressions should be considered to be less reliable than those estimated using the multi-year techniques.

Comparing the results of the two multi-year techniques, the efficiency rankings are almost identical and the efficiency scores for BT are very similar. In addition, when BT is compared against the best performing firm, its efficiency scores under the two multi-year techniques are close to those produced by DEA (using the sample of large LECs). This consistency, coupled with the doubts about the reliability of the single year techniques, suggests that more reliance should be placed on the results of these techniques, and that therefore it is reasonable to conclude that BT is in the region of 9% to 10% less efficient than upper decile point of the large LECs and in the region of 13% to 16% less efficient than the best performing large LECs[53].

When the full sample of LECs is assessed there is some consistency between the rankings of most of the firms under the single-year SFA and single-year OLS techniques. However, the assessment of BT under these two techniques is very different. The results of the single-year OLS, when BT is compared against the best performing firm, is broadly consistent with the results from the DEA technique when it is estimated using the full sample of LECs. This consistency, coupled with the doubts regarding the single-year SFA, because it only works for one of the three possible inefficiency distribution assumptions, suggests that greater reliance should be placed on the results from single-year OLS full sample and full sample DEA. This in turn suggests that BT is in the region of 9% less efficient than the upper decile point of the full sample of the LECs and between 17% and 20% less efficient that the best performing LECs[54].

53 The results of the single-year OLS are not included in the range as the very limited sample size used to estimate this regression may have limited the robustness of these results.

54 The results of the single-year SFA are not included in the range as the comparison of the rankings of BT under the two single-year full sample regressions suggests that these two techniques do not present a consistent assessment of BT's comparative efficiency. Therefore, given the consistency between the results of the DEA analysis using the full sample and the OLS single-year full sample results, and the doubts about the reliability of the SFA results given the difficulties experienced in fitting the efficiency distributions to the residuals when only a single-year of data is used, this suggests that 17% to 20% is a reasonable estimate of BT's relative efficiency (when compared to the best firm) under the full sample of LECs.

This indicates that BT is more inefficient when the smaller US LECs are included in the sample. For example, under the single-year OLS regression estimated using the large LECs BT is 3% behind the upper decile point however, under the comparable regression estimated using the full sample of LECs, BT is 9% behind the upper decile point. One possible explanation for this is that there are some environmental characteristics which impact upon the costs of either the larger or smaller firms in the sample. However, from this analysis it is not possible to identify whether this is the cause of the result.

In addition, as mentioned above, the results of the single-year regression techniques should be treated with some caution as these regressions were estimated using relatively small samples. Therefore, given the greater robustness of the results of the multi-year analysis, it is reasonable to conclude that BT's efficiency is in the region of 9% to 10% lower than the upper decile point and 13% to 16% lower than that of the most efficient firm(s).

5.3. BT's Alternative Model Specification

In this section NERA summarises its understanding of the alternative model specification developed by BT.

BT expressed concerns over the use of a total lines variable, and suggested that switched lines and leased lines should enter the regression as separate variables. This allows for substitution between the two types of lines in the cost function. As a result, the following model specification was developed by BT:

$$log(C) = \beta_0 + \beta_1 log(LL) + \beta_1 log(SL) + \beta_2 log(M) + \beta_3 log(TS/SL) + \beta_4 NED + \beta_5 T + \beta_7 log(NM/OM) + \beta_8 D9698*log(TS/SL) + \beta_9 D9698*log(SL) + \beta_9 D9698*log(LL) + \beta_{10} D9698$$

Where:

- *C* = total cost
- *LL* = leased lines
- *SL* = switched lines
- *M* = switch minutes
- *TS/SL* = average total sheath length per switch line
- *NED* = North East US dummy
- *T* = Time
- *NM/OM* = national minutes / other minutes
- *D9698*log(TS/SL)* = interaction dummy for average total sheath length per switch line in 1996-1998
- *D9698*log(SL)* = interaction dummy for switch lines in 1996-1998
- *D9698*log(LL)* = interaction dummy for switch lines in 1996-1998

- *D9698* = shift dummy for the period 1996-1998.

When developing its model specification NERA investigated using separate switched lines and leased line variables. However, decided to use a combined lines variable for two reasons. Firstly, this would be consistent with the approach used in previous comparative efficiency studies, and would therefore allow an assessment of how BT's efficiency had changed since the earlier studies. Secondly, the multicollinearity which would be introduced by having separate line variables could impact upon the estimation of the model, particularly when using the single-year techniques (the problem of multicollinearity is discussed in section 4.2 above).

The results of estimating BT's alternative model specification are presented below.

Table 5.13
Regression Results BT Model

	Coefficient	Standard Error	T Statistic	Significance
Log of leased lines	0.05	0.01	3.30	0.00
Log of switched lines	0.86	0.05	18.91	0.00
Log of switch minutes	0.08	0.03	2.34	0.02
Log of average total sheath length per switch line	0.16	0.04	3.53	0.00
North East US dummy	0.21	0.06	3.34	0.00
Time	0.01	0.00	2.48	0.01
Log of national calls per other calls	0.03	0.01	3.24	0.00
Interaction dummy for log of average total sheath per switch line in the period 1996-1998	0.07	0.02	4.49	0.00
Interaction dummy for log of leased lines in the period 1996-1998	-0.07	0.01	-5.05	0.00
Interaction dummy for log of switched lines in the period 1996-1998	0.07	0.02	4.61	0.00
Shift dummy for the period 1996-1998	0.10	0.11	0.89	0.38
Constant term	6.04	0.40	15.26	0.00
Log Likelihood	**257.39**			

Source: NERA analysis

In comparison to the NERA model the elasticity associated with the number of switched lines is somewhat higher. The elasticity with respect to the number of leased lines is small but not zero, while the elasticity with respect to the number of switched minutes is significantly lower than in the NERA model. Additionally, the BT specification yields a much higher value for the log-likelihood function, implying that it fits the data better.

Table 5.14
Efficiency Results BT Model

Rank	Company	Residual	% Inefficiency from Frontier	% Inefficiency from Best Performer	% Inefficiency form Decile	% Inefficiency from Median
1	WISCONSIN BELL	0.03	3%	0%	-8%	-19%
2	INDIANA BELL	0.07	7%	4%	-4%	-16%
3	BA DELAWARE	0.11	11%	8%	-1%	-12%
4	S NEW ENGLAND TEL CO	0.11	12%	9%	0%	-12%
5	MICHIGAN BELL	0.11	12%	9%	0%	-12%
6	VERIZON NORTH	0.12	13%	9%	1%	-11%
7	BT	0.12	13%	10%	1%	-11%
8	VERIZON NEW ENGLAND	0.17	18%	15%	6%	-7%
9	OHIO BELL	0.17	18%	15%	6%	-7%
10	CENTRAL TEL CO	0.19	21%	18%	8%	-4%
11	CAROLINA TEL CO	0.21	23%	20%	10%	-3%
12	PACIFIC BELL	0.21	24%	20%	11%	-2%
13	ILLINOIS BELL	0.23	25%	22%	12%	-1%
14	VERIZON PENNSYLVANIA	0.23	26%	23%	13%	-1%
15	SPRINT FLORIDA	0.23	26%	23%	13%	0%
16	BA MARYLAND	0.24	28%	24%	14%	0%
17	VERIZON NEW JERSEY	0.25	28%	24%	14%	1%
18	UNITED OF OHIO	0.25	28%	25%	15%	1%
19	BELLSOUTH	0.27	31%	27%	17%	3%
20	VERIZON SOUTH	0.28	32%	29%	18%	4%
21	QWEST	0.29	33%	30%	19%	5%
22	VERIZON NORTHWEST	0.30	35%	31%	21%	6%
23	VERIZON VIRGINIA	0.30	35%	31%	21%	6%
24	VERIZON WEST VIRGINIA	0.32	37%	33%	22%	8%
25	CINCINNATI BELL	0.32	37%	33%	23%	8%
26	VERIZON CALIFORNIA	0.33	39%	35%	24%	9%
27	VERIZON FLORIDA	0.34	40%	36%	25%	10%
28	SW BELL	0.34	41%	37%	26%	11%
29	VERIZON SOUTHWEST	0.35	42%	38%	27%	12%
30	BA NEW YORK	0.41	51%	47%	35%	19%

Source: NERA analysis

Comparison of the rankings of the companies under the two models shows a high degree of correlation (the correlation coefficient between the rankings under the two models is 0.98). However, the efficiency scores for some of the companies, including BT, are significantly different. For example BT is ranked as 7th under the BT model and 9th under the NERA model. However, BT's efficiency scores under these two models, when compared to the upper decile point of the sample, are 1% and 9% respectively. This suggests that there are significant differences in the way the two models measure inefficiency.

In the next section of this report the performance of the BT and NERA models is assessed using a number of criteria. The aim of this assessment is to identify the model which is most appropriate for assessing BT's comparative efficiency for the purposes of Ofcom's price cap modelling.

5.4. Model Comparison

While statistical fit is an important criterion for judging the performance of a model, there are other considerations that need to be taken into account. For example, the reasonableness of the model from the perspective of *a priori* expectations of the key drivers of the costs of telecommunications networks, and the importance of these drivers in explaining costs (i.e. their magnitude). After discussions with Ofcom, NERA identified the following additional criteria by which the suitability of the different cost models could be assessed:

- ***Continuity with the models previously used to estimate BT's comparative efficiency:*** the model should allow an assessment of how BT's efficiency had changed since the previous comparative efficiency study.
- ***Consistency with the movements in costs over time:*** the results of the model should be sensible when compared with movements in the costs and outputs of the firms analysed.
- ***Consistency with BT's regulatory cost base:*** the results of the comparative efficiency assessment are to be used by Ofcom in a financial price-cap model and, therefore, the cost function implied by the model should not be inconsistent with BT's cost base, as this cost base will be the source of the underlying data included in the price-cap model.

In the following sections the performance of the NERA and BT model specifications are assessed under each of the above criteria.

5.4.1. Continuity with previous studies

Models used in previous comparative efficiency studies have been estimated using a simplified version of the model specified by NERA as part of this current study. They have estimated total costs as a function of total lines, minutes of switch use, and sheath length variables. Therefore, the use of separate switched and leased lines variables represents a significant departure from the models used in previous studies.

To assess the impact of this departure upon the comparability of the efficiency assessments, NERA estimated the model used in the previous comparative efficiency study using 2001 US LEC data, and then assessed how the results of this model compared to those obtained using the NERA and BT models in the current study. Additionally, to inform this comparison further, a model was estimated with the same explanatory variables as the model used in the previous study with the exception of the lines variable, which was spilt into separate switched and leased lines variables.

The table below shows the results of OLS regressions for 2001 using the full sample of US LECs for: the model from the previous study; the previous study model but with separate lines variables; the NERA model from the current study; and the BT model from the current study.

Table 5.15
Comparison using OLS 2001 Results

	Model from Previous Study		Model from Previous Study (with separate lines variables)		NERA Model from Current Study		BT Model from Current Study	
	Coeff	**P>t**	**Coeff**	**P>t**	**Coeff**	**P>t**	**Coeff**	**P>t**
Lines	0.65	0.00	n.a.	n.a.	0.91	0.00	n.a.	n.a.
Switched Lines	n.a.	n.a.	0.99	0.00	n.a.	n.a.	1.03	0.00
Leased Lines	n.a.	n.a.	0.02	0.52	n.a.	n.a.	-0.01	0.72
Switch Minutes	0.31	0.00	0.01	0.91	0.08	0.29	0.00	0.97
Aerial Sheath per line	0.13	0.20	0.06	0.09	n.a.	n.a.	n.a.	n.a.
Non aerial Sheath per line	0.06	0.01	0.09	0.01	n.a.	n.a.	n.a.	n.a.
Total Sheath per switch line	n.a.	n.a.	n.a.	n.a.	0.12	0.02	0.15	0.00
Leased Lines per Switched Line	n.a.	n.a.	n.a.	n.a.	-0.28	0.00	n.a.	n.a.
National Calls per Other Calls	n.a.	n.a.	n.a.	n.a.	-0.01	0.63	-0.02	0.17
North East US Dummy	n.a.	n.a.	n.a.	n.a.	0.22	0.01	0.18	0.01
Constant	4.42	0.00	6.60	0.00	5.23	0.00	6.32	0.00
R^2	0.98		0.99		0.99		0.99	
BT efficiency (compared to upper decile)	32.5%		0%		9.4%		1.82%	

Source: NERA analysis

The table shows the impact on the efficiency scores of two departures from the model used in the previous study:

- the separation of the total lines variable into switched lines and leased lines; and
- the introduction of additional environmental variables (i.e. leased lines per switch line, national calls per other calls and a North East US dummy).

The results indicate that the separation of the total lines variable represents a more significant departure from the model used in the previous study than the inclusion of the additional explanatory variables. The separation of the lines variable on its own reduces BT's inefficiency in 2001 compared to the decile from 32.5% to 0%. The impact of including the additional explanatory variables (but not separating the lines variable) is to reduce BT's inefficiency relative to the decile from 32.5% to 9.4%.

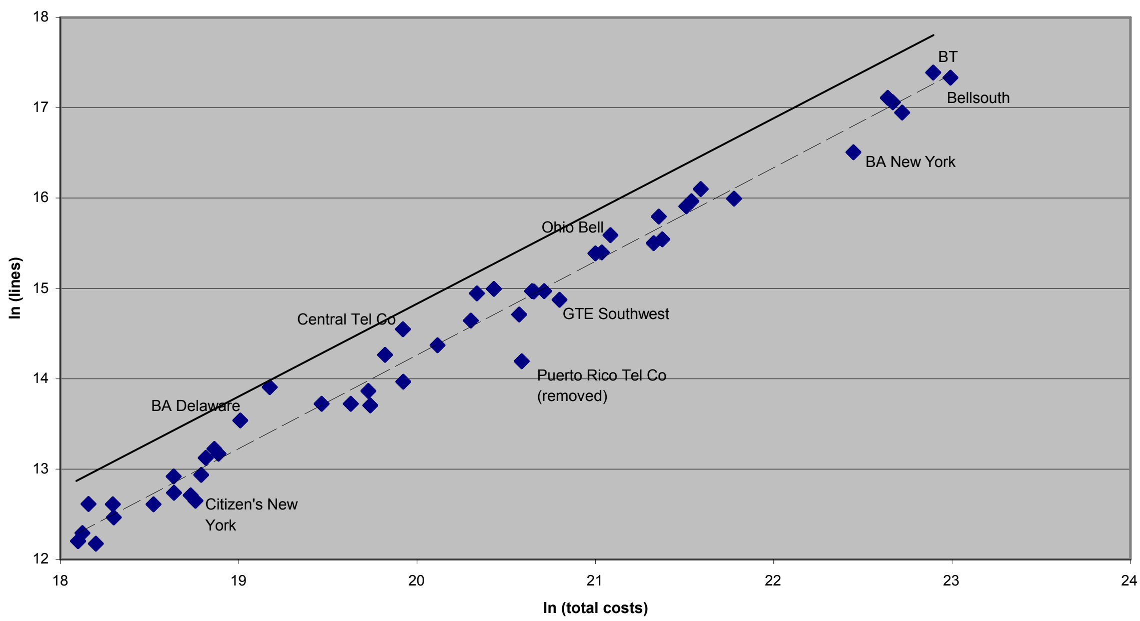
Both of these departures from the previous model result in an improvement in the fit of the model, as shown by the higher R^2 for these models. However, based on the results of the 2001 OLS regression, the current NERA model, which does not involve separation of the lines variable, fits the data as well as the BT model, which involves separation of the lines variable as well as the inclusion of environmental variables.[55]

The conclusion regarding consistency with the previous model is therefore that both the NERA and BT models represent a departure from the model used in the previous study. However, the BT model represents a more significant departure from the previous model than the NERA model.

5.4.2. Consistency with the movements in costs overtime

Analysis of BT's network cost per line compared with that of the US LECs in both 1999/00 and 2001/02 indicates that BT's relative performance deteriorated over this period. The two diagrams below show plots of the log of total cost (horizontal axis) against the log of total lines (vertical axis) in each of the two years.

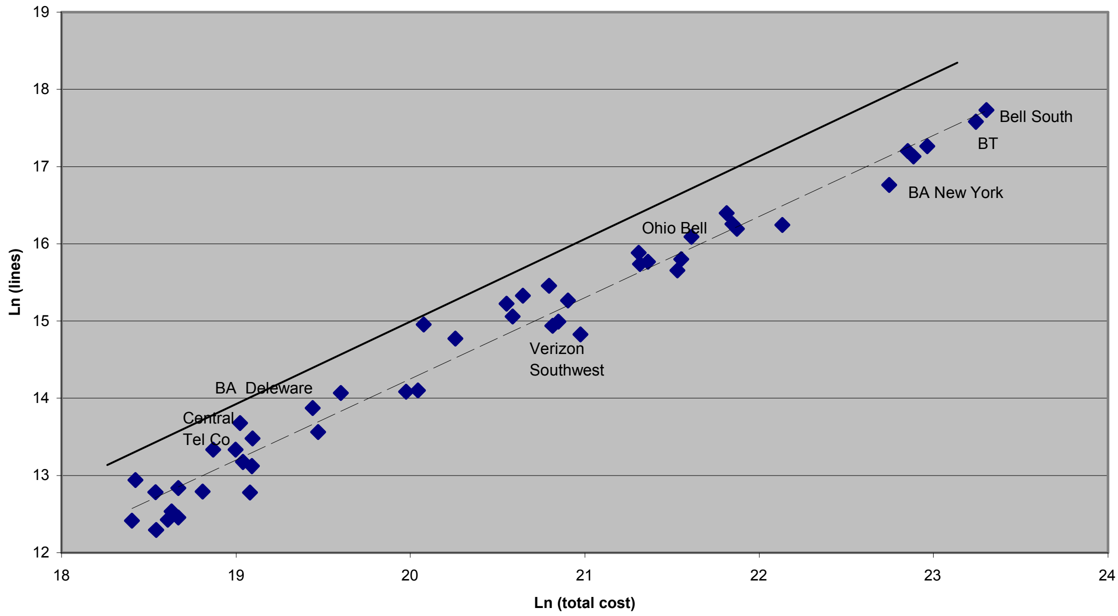
Figure 5.4
Costs vs Lines (1999)



Source: "The Comparative Efficiency of BT: an Update", report prepared for Oftel by NERA, January 2001. (A linear trend line and frontier have been superimposed).

55 However, when estimated using panel SFA over the period 1996 to 2001 the statistical fit of the BT model is superior to that of the NERA model.

Figure 5.5
Cost vs Lines (2001)



Source: NERA Analysis

These diagrams show that BT was above the line of best fit between costs and lines in 1999 but is below this line in 2001.

To investigate this further, the percentage change in BT's output(s) over the period, as expressed in terms of the total number of lines (and the total number of switched lines and leased lines), total number of call minutes, leased lines per switch line and national calls per other call, were used, along with the coefficients on these variables in the NERA and BT models, to estimate the change in network costs that would be expected as a result of the change in outputs.

Given the use of a Cobb-Douglas functional form in the regression analysis, expressed in a log-log format, the regression coefficients can be interpreted as the expected percentage change in network costs that would result if the explanatory variable (e.g. the number of lines) were to increase by 1%. Therefore, for example, the coefficient on lines of 0.81 indicates that as the total number of lines increases by 1%, total network costs would be expected to increase by 0.81%. Furthermore, for example, the coefficient on switch minutes of 0.15 would imply an increase in total network costs of 0.15% for a 1% increase in switch minutes.

Table 5.16
Review of Change in Costs and Outputs

	All Calls	Switched Lines	Leased Lines	All Lines	Leased Lines per Switched Line	National Calls per Other Calls
2001/2	202,434	29,087	549	30,185	0.037753	0.15035
2000/1	170,689	28,881	707	30,294	0.048938	0.197481
1999/00	144,432	28,485	638	29,760	0.044769	0.269955
1998/9	129,245	28,049	607	29,263	0.043268	0.334238
	Annual % change	**Annual % change**	**Annual % change**	**Annual % change**	**Annual % change**	**Annual % change**
2001/2	19%	1%	-22%	-0.4%	-23%	-24%
2000/1	18%	1%	11%	2%	9%	-27%
1999/00	12%	2%	5%	2%	3%	-19%
1998/9	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Source: NERA analysis

The data in this table are based on the network costs reported in BT's regulatory financial statements[56] and the numbers of BT lines and calls reported in "Annual Market Information 1996/97 to 2001/02" published by Oftel in March 2003. The total number of lines is the number of exchange lines plus the number of inland private circuits (expressed in 64kb equivalents)[57].

Taking all of these changes into account, together with the regression coefficient on the time variable, the NERA model specification suggests that BT's total network costs would have been expected to increase by 9% between 1999/00 and 2001/02. The expected increase implied by the BT model over the same period is only 3%. The actual change in BT's network costs over this period was 19%. Therefore, on the basis of coefficients in both the BT and NERA models, the actual increase in BT's costs has been substantially greater than that predicted by both the models, implying that BT's inefficiency increased between 1999/00 to 2001/02.[58]

We also looked at how the change in BT's costs relative to output compared with that for the US LECs. The table below shows how BT's network costs, total lines and call minutes changed between 1999/00 and 2001/02 and compares these changes to the average change for the US LECs.

56 For 1998/99 and 1999/00 BT's network costs are estimated by applying the percentage of Access and Network costs reported for 2000/01 which are re-classified as Network in the 2001/02 financial statements

57 The numbers of digital private circuits were reported in categorised by capacity (sub 2Mb, 2Mb and greater than 2Mb). To express these circuit numbers in 64Kb equivalents the sub 2Mb circuits were assumed to be evenly distributed between 64Kb and 2Mb, the greater than 2Mb circuits were assumed to be, on average 5Mb circuits (this assumption is unlikely to have a significant impact upon the analysis as BT reported only 1 circuit in the greater than 2Mb category).

58 In carrying out this calculation we were obliged to use call minutes as a proxy for switch minutes as no switch minute data was provided by BT for 1999/00. Between 2000/1 and 2001/2 the volume of BT switch minutes grew less fast than the volume of call minutes so that, if anything, our analysis is likely to overstate BT's growth in output and hence understate its increase in inefficiency.

Table 5.17
Analysis of Cost And Output Changes 1999/00 to 2001/02

	Percentage change in network costs	Percentage change in total line numbers	Percentage change in call minutes
BT	19%	2%	40%
Average of US LECs	0%	11%[59]	-1%

Source: NERA analysis

It can be seen that, while BT's network costs increased by 19% over this period, the LECs' costs, on average, remained more or less constant. Whilst BT's call minutes increased significantly faster than the LECs' call minutes between 1999/00 and 2001/02, its line numbers increased more slowly than those of the LECs. If we assume that lines account for 70% of total network costs and calls for 30% (which was the case for BT in 2001/2), then, between 1999/00 and 2001/02, BT experienced a 19% increase in costs in the face of a 13.4% increase in output while the LECs experienced no increase in costs while their output grew by 7.4%. In other words, BT costs grew faster than its output while the opposite was true for the LECs.

These figures support the view that BT has become less efficient since the previous comparative efficiency assessment carried out in 2000 (using data for the financial year 1999/00). In fact they suggest that BT's inefficiency could have increased by something in the region of 13 percentage points since 1999/00. The difference between BT's inefficiency in the previous study (approximately 3%) and the range for BT's inefficiency identified by the NERA model in this study, of between 9% and 10% relative to the upper deciles, is around 7 percentage points. The conclusion therefore is that the efficiency decline implied by the NERA model is broadly consistent with the pattern of changes in costs and output volumes over time.

In contrast, the efficiency scores obtained from the BT model are not consistent with this pattern. The BT model suggests that, compared to the upper decile, BT's performance has improved since the previous study. This is difficult to square with actual changes in costs and output. However, this implied improvement may be largely due to the inconsistency of the BT model with the model estimated in the previous study.

5.4.3. Consistency with BT's regulatory cost base

In order to assess the consistency with BT's regulatory cost base, the implied cost-volume elasticities and the implied costs shares of the BT and NERA models have been assessed. The results of these analyses are presented below.

---

59 A significant driver of the increase in the number of total lines for the US LECs is the growth in the number of leased lines.

5.4.3.1. *Analysis of Implied Cost-Volume Elasticities*

To assess the different implications for BT of the NERA and BT models, NERA has investigated the implied cost volume elasticities (CVEs) under both models. This involved the calculation of the following cost volume elasticities:

- The elasticity of switch line costs with respect to switch lines;
- The elasticity of leased line costs with respect to leased lines; and
- The elasticity of the costs of calls with respect to switch minutes.

The results of these calculations are presented in the table below. The calculations are set out in detail in Appendix C.

Table 5.18
Cost-Volume Elasticities Implied by the BT and NERA Models

	BT Model	NERA Model
Exchange lines (E)	1.31	1.10
Leased lines (L)	0.31	0.62
Switch minutes (M)	0.26	0.50

Source: NERA analysis

It can be seen that (for BT) the BT model implies a CVE for exchange lines of 1.31. This suggests very pronounced diseconomies of scale and seems unlikely to be a representation of reality. The CVE for switch minutes implied by the BT model is consistent with the CVE for the conveyance network calculated by Ofcom of between 0.2 and 0.3[60]. However, the CVE for leased lines is significantly below Ofcom's CVE for the access network of 0.5[61].

The NERA model also indicates diseconomies of scale in the case of exchange lines, although the extent of such diseconomies is not as large. Whilst the CVE for switch minutes implied by the NERA model is high compared to Ofcom's CVE for the conveyance network, the CVE for leased lines is significantly closer to Ofcom's access network CVE of 0.5.

Analysis of the results above indicates that neither model is consistently superior to the other in terms of the proximity of the implied CVEs to those previously estimated by Ofcom. The NERA model is closer in the case of both exchange lines and leased lines but the BT model is closer in the case of switch minutes. The BT model implies a cost volume elasticity

60 This is the range of CVEs for the conveyance network assumed by Oftel in its network charge control review statement in February 2001.

61 This is the CVE used by Oftel in its price cap modelling of the wholesale line rental product.

for exchange lines that appears to be at odds with reality. The NERA model suffers from the same problem but to a lesser degree.

5.4.3.2. Analysis of Implied Cost Weights

To facilitate further the comparison between these different models NERA has estimated the implied cost shares for BT under each model. This involves calculating the proportion of the costs forecast for BT by each model that can be attributed to exchange lines, leased lines and switch minutes[62]. The results of these calculations are presented in the table below, along with the actual cost shares for BT[63].

Table 5.19
Cost Shares Implied by the BT and NERA Models

	BT Model	**NERA Model**	**Actual BT Cost Shares (2001/02)**
Exchange lines (E)	81%	66%	54%
Leased lines (L)	6%	11%	16%
Switch minutes (M)	13%	23%	30%

Source: NERA analysis

These results show that the cost shares implied by the NERA model are significantly closer to the actual cost shares in BT's regulatory accounts than those implied by the BT model. Therefore, as BT's regulatory costs will form the base of Ofcom's price cap model, an efficiency score which was based on the BT model would appear to be inconsistent with the costs included in the price cap model. In particular, as Ofcom is modelling BT's provision of partial private circuits, it would be misleading to use an efficiency score in this price cap model that placed so little weight on leased lines in predicting the efficient level of BT's network costs. Therefore, analysis of the implied costs shares suggests that the NERA model is more appropriate for Ofcom's purposes (i.e. assessing BT's comparative efficiency for inclusion in a price cap model for partial private circuits).

5.4.4. Conclusion of model comparison

In this section the performance of the BT and NERA models has been assessed under three criteria. Firstly the extent to which they provide continuity with the models previously used to estimate BT's comparative efficiency has been analysed. Secondly, their consistency with

62 The forecast costs attributed to the environmental variables and the constant have been spread over the three outputs as follows: the costs due to the constant and time variable have been spread over all the outputs in proportion to their forecast costs, the costs due to the national calls per other calls variable has been allocated to switch minutes, and the costs due to the total sheath length variable have been split over switch lines and leased lines in proportion to their relative costs.

63 The method used to calculate BT's cost shares is set out in Appendix C.

the movements in costs overtime has been assessed, and finally their consistency with BT's regulatory cost base has been reviewed.

The conclusions reached in this assessment are that the NERA model achieves a greater degree of continuity with the previous models used to estimate BT's comparative efficiency, the results obtained by the NERA model are more consistent with actual movements in costs[64], and the NERA model is more consistent with BT's regulatory cost base.

This suggests that, whilst the BT model is superior in terms of its statistical fit, the NERA model is more appropriate for assessing BT's comparative efficiency, particularly given that the comparative efficiency results will be included in Ofcom's price cap model of BT's partial private circuits which uses the costs in BT's regulatory accounts.

5.4.5. Sensitivities

In this section the results of some sensitivity tests are provided for the NERA model. These tests assess the impact of variations in the assumptions made in the analysis.

The sensitivity tests cover:

- The move to weighting special access lines and leased lines as 1.5 and 2 times an exchange line respectively (rather than the equal weights of 1 and 1 used in the previous comparative efficiency study). We have also completed additional sensitivities to assess the impact of weighting special access lines as 1.25 and 1.75, rather than 1.5, times an exchange line.
- The allocation of the LECs' indirect costs (including the depreciation and capital costs of indirect assets) between retail and network activities. In the core analysis presented above (hereafter called the base case) this allocation is based upon the proportions of direct costs which are allocated to each activity. This allocation mechanism results in, on average, 80% of the indirect costs being allocated to network and 20% being allocated to retail. The impact of this allocation was assessed by increasing, and decreasing, the amount allocated to network activities by 10 percentage points. In other words, BT's efficiency was estimated when, on average, 90% of the LECs' indirect costs were allocated to network activities, and when, on average, 70% of the LECs' indirect costs were allocated to network activities.
- The sensitivity of the results to changes in the routing factors used for the US LECs was also investigated. In particular sensitivities were completed to assess the impact of:

64 The consistency of the results of the BT model with movements in costs overtime is difficult to establish. The results of the BT model indicate an improvement in inefficiency over time that is not consistent with the increase in BT's costs over the period. However, this improvement may be largely due to the lack of continuity with models estimated in previous comparative efficiency studies.

  - using local switch routing factors for local calls of 1.2, and of 1.7, rather than the 1.54 used in the base case; and
  - using main switch routing factors for Inter- and Intra-LATA calls of 0.5 and 0.7, and of 0.9 and 1.5, rather than the 0.2 and 0.3 used in the base case.

- The impact of using a main switch routing factor for Inter-LATA calls of 0.2 rather than 0.3. BT suggested that the main switch routing factor for this call type should be 0.2, as this is the value used in the Hatfield model. Therefore, a sensitivity was completed to assess the impact of this suggestion.
- The impact of using different weights on main switch minutes in the total switch minutes variable. Sensitivities have been completed for weights of 0.3 and 0.7. These are compared to the weight of 0.5 used in the base case.
- The sensitivity of the results to the NRC/GRC ratio used to adjust the US LECs' assets was also assessed. For comparability purposes a common NRC/GRC ratio was assumed for all of the companies in the sample; in the base case this common ratio was assumed to be BT's ratio of 0.48. To assess the impact of assuming that the US LECs have the same NRC/GRC ratio as BT, a sensitivity was completed which adjusted the US LECs asset base using an NRC/GRC ratio of 0.50 (rather than 0.48), whilst BT's asset base remained un-changed.
- The impact of moving from expressing the US LECs' costs in pounds sterling to expressing BT in US dollars was assessed.
- The use of different sheath length variables. Here the impact of using average total sheath per line or just total sheath length, rather than average total sheath per switch line, is assessed.
- The impact of excluding data network costs from BT's total cost figure was also assessed. These costs were excluded in the base case as the FCC had indicated that such costs are not supposed to be included in the data reported for the US LECs. The FCC cannot however, confirm that none of the US LECs have included such costs in their data. Therefore, a sensitivity test was completed which assessed the impact on BT's efficiency of including data network costs in BT's total cost figure.
- The impact of using smoothed exchange rates rather than market exchange rates to convert certain of BT's costs from UK£ to US$.

The table below presents BT's efficiency score compared to the upper decile point in the sample under both the multi-year SFA and multi-year least squares techniques under the base case (i.e. the results presented in Table 5.12 above) and under each of the sensitivities set out above. This analysis is completed using only the multi-year techniques, as the purpose of the sensitivities is merely to illustrate the magnitude of the impact upon BT's efficiency score of these changes.

Table 5.20
Sensitivity Analysis

	Multi-year SFA Inefficiency compared to upper decile	Multi-year Least Squares Inefficiency compared to upper decile
Base case	**9%**	**10%**
Use of equal weights of 1 and 1 on special access and leased lines (rather than 1.5 and 2)	1%	0%
Use of weights of 1.25 and 2 on special access and leased lines (rather than 1.5 and 2)	7%	7%
Use of weights of 1.75 and 2 on special access and leased lines (rather than 1.5 and 2)	11%	11%
Allocation of LEC common costs (additional 10% to Network)	8%	8%
Allocation of LEC common costs (additional 10% to Retail)	12%	12%
LEC local switch routing factor for local calls set to 1.2 rather than 1.54	7%	7%
LEC local switch routing factor for local calls set to 1.7 rather than 1.54	11%	11%
LEC main switch routing factors for inter-LATA calls set to 0.2, rather than 0.3	9%	8%
LEC main switch routing factors for intra-LATA and inter-LATA calls set to 0.5 and 0.7, rather than 0.2 and 0.3	11%	10%
LEC main switch routing factors for intra-LATA and inter-LATA calls set to 0.9 and 1.5, rather than 0.2 and 0.3	12%	12%
Use of weight of 0.3 on main switch minutes (rather than 0.5)	10%	10%
Use of weight of 0.7 on main switch minutes (rather than 0.5)	10%	10%
Use of UK £s rather than US $s	9%	10%
Use of average total sheath *per line* rather than average total sheath *per switch line*	11%	11%
Use of total sheath length (not averaged) rather than average total sheath per switch line	10%	10%
Inclusion of data networks costs in BT's total costs	15%	16%
Increasing NRC/GRC ratio for the US LECs from 0.48 to 0.5 (Note: BT's NRC/GRC ratio has not been adjusted)	9%	8%
Using smoothed exchange rates rather than the market rates to convert from UK£s to US$s	10%	10%

Source: NERA analysis

It can be seen that the changes in the currency used, the use of different sheath length variables and the different weights on main switch minutes all have a limited impact upon the inefficiency score. The maximum impact of any of these changes being only 4 percentage points.

The change in the weight used on the different line types in the total lines variable has a much greater impact upon the efficiency score, decreasing BT's inefficiency to 1% and 0% under the different multi-year techniques. However, there are reasons to consider the weights of 1.5 and 2 used on special access lines and leased lines respectively to be superior to the weights of 1 and 1 used in this sensitivity (see section 4.2 above for further discussion of this issue). Further to this point, increasing the weight on special access lines from 1.5 to 1.75 (whilst keeping the other weights the same), and decreasing it from 1.5 to 1.25, both have a limited impact upon the efficiency score. These change BT's inefficiency scores under multi-year SFA from 9% to 11% and 7% respectively.

The only other sensitivity test that has a significant impact upon BT's efficiency score is the inclusion of data networks costs in BT's cost base, which increases BT's inefficiency scores from 9% and 10% to 15% and 16% respectively.

Given the view of the FCC, namely that data network costs are not supposed to be included in the US LECs' submitted cost information, we believe it to be prudent to remove these costs from BT's cost base. However, it is worth noting that the exclusion of these costs does produce an appreciable improvement in BT's comparative efficiency.

6. CONCLUSION

This study has assessed BT's comparative efficiency in 2000/01 and 2001/02 relative to the US LECs using a number of different techniques. The results of these different techniques are summarised in the table below.

Table 6.1
Summary of BT's Comparative Efficiency

Point of Comparison	Full Sample		Large LECs only	
	Best Performer	Upper Decile	Best Performer	Upper Decile
Multi-year SFA	n.a.	n.a.	16%	9%
Multi-year least squares	n.a.	n.a.	16%	10%
Single-year OLS	20%	9%	7%	3%
Single-year SFA - Exponential	1%	1%	n.a.	n.a.
Single-year DEA (compared against peer group)	17%	n.a.	13%	n.a.
Range of All Results	1%-20%	1%-9%	7%-16%	3%-10%
Range Excluding Single-year Regressions	**17%**	**n.a.**	**13%-16%**	**9%-10%**

Source: NERA analysis

The multi-year techniques were estimated using the larger US LECs only, as tests showed that there were structural differences between the regression coefficients estimated for the smaller and larger US LECs over the period 1996-2001[65]. The single-year techniques were estimated, wherever possible, using both the full sample and the larger US LECs only, as comparable stability tests for the period 2001 showed that, when this year is considered in isolation, there are no significant differences between the regression coefficients estimated for the smaller and larger US LECs.

The summary table above shows that there is some variation between the results of the different comparative efficiency measurement techniques. In order to analyse the reasons behind this variation, we compared BT's efficiency ranking under each of these techniques. This analysis (which is limited to comparison of the rankings under the regression analysis techniques)[66] showed that the highest degree of correlation is between the two multi-year techniques, but that the single-year SFA and single-year OLS (both estimated using all of the US LECs) also exhibit a high degree of correlation.

[65] Data was included for the US LECs for the period 1996 to 2001, as this substantially increases the size of the data set used to estimate the regressions. This increase in the sample size would be expected to significantly increase the reliability of the results. BT's efficiency result, however, is estimated using data for the period 2000/01 to 2001/02.

[66] It is not possible to construct comparable rankings for the DEA results, because the DEA technique identifies different peer groups for different operators and therefore it is not possible to produce an overall ranking of all the operators.

The results of all the single-year techniques should be treated with some caution as these regressions were estimated using relatively small samples. The single-year OLS (estimated using only the large LECs) is based on a sample size of 29, the single-year SFA and OLS (estimated using the full sample) are both based on a sample size of 46, however, the multi-year regressions are both based on a sample of 174. Hence, the results of the single-year regressions should be considered to be less reliable than those estimated using the multi-year techniques.

Comparing the results of the two multi-year techniques, the efficiency rankings are almost identical and the efficiency scores for BT are very similar. In addition, when BT is compared against the best performing firm, its efficiency scores under the two multi-year techniques are close to those produced by DEA (using the sample of large LECs). This consistency, coupled with the doubts about the reliability of the single year techniques, suggests that more reliance should be placed on the results of these techniques, and that therefore it is reasonable to conclude that BT is in the region of 9% to 10% less efficient than upper decile point of the large LECs and in the region of 13% to 16% less efficient than the best performing large LECs[67].

When the full sample of LECs is assessed there is some consistency between the overall rankings under the single-year SFA and single-year OLS techniques. However, the assessment of BT under these two techniques is very different. The results of the single-year OLS, when BT is compared against the best performing firm, is broadly consistent with the results from DEA technique when it is estimated using the full sample of LECs. This consistency, coupled with the doubts regarding the single-year SFA, because it only works for one of the three possible inefficiency distribution assumptions, suggests that greater reliance should be placed on the results from single-year OLS full sample and full sample DEA. This in turn suggests that BT is in the region of 9% less efficient than the upper decile point of the full sample of the LECs and between 17% and 20% less efficient that the best performing LECs[68].

However, as mentioned above, the results of the single-year techniques should be treated with some caution as these regressions were estimated using relatively small samples.

Reflecting this, it is reasonable to conclude that BTs inefficiency score in 2001 is:

---

67 The results of the single-year OLS are not included in the range as the very limited sample size used to estimate this regression may have limited the robustness of these results.

68 The results of the single-year SFA are not included in the range as the comparison of the rankings of BT under the two single-year full sample regressions suggests that these two techniques do not present a consistent assessment of BT's comparative efficiency. Therefore, given the consistency between the results of the DEA analysis using the full sample and the OLS single-year full sample results, and the doubts about the reliability of the SFA results given the difficulties experienced in fitting the efficiency distributions to the residuals when only a single-year of data is used, this suggests that 17% to 20% is a reasonable estimate of BT's relative efficiency (when compared to the best firm) under the full sample of LECs.

- between 9% and 10% when compared against the upper decile point of the large US LECs, and
- between 13% and 16% based on the more demanding comparison against the best performance (i.e. the most efficient large US LEC).

This represents a substantial increase in BT's comparative inefficiency since the previous study undertaken in 2000, which assessed BT's comparative efficiency in the financial year 1999/00. To further assess whether BT's efficiency has deteriorated over the period 1999/00 to 2001/02, analysis was completed to reconcile the change in BT's efficiency score to changes in BT's costs and outputs.

The results of this reconciliation (which is set out in section 5.4.2 above) support the conclusion arising from the comparative efficiency analysis, namely that BT has become more inefficient over this period. Indeed the analysis of the change in BT's costs that would be expected given the change in its outputs over the period suggests that BT's inefficiency could have increased by up to 13 percentage points. This magnitude of efficiency change is broadly in line with the difference between BT's inefficiency in the previous study (approximately 3%) and the range for BT's inefficiency identified in this study (between 9% and 10% or between 13% and 16% if the best performing company is used as the reference point).

The results discussed above are those estimated using the NERA model specification. BT suggested an alternative model specification, which yields a much higher value for the log-likelihood function, implying that it fits the data better. Additionally, the BT model produces significantly different efficiency results. Using the multi-year SFA technique BT's inefficiency compared to the upper decile is 1% under the BT model compared to 9% under the NERA model.

NERA assessed the suitability of the two alternative model specifications using a number of criteria (see section 5.4 for a discussion of this analysis). The aim of this assessment is to identify the model which is most appropriate for assessing BT's comparative efficiency for the purposes of Ofcom's price cap modelling. The assessment indicated that the NERA model is more appropriate for assessing BT's comparative efficiency for the purposes of deriving an efficiency factor for inclusion in a price cap model. The main reasons for this are that the NERA model is more consistent with BT's regulatory cost base, the results of the NERA model are consistent with analysis of the relative movements overtime in the costs and outputs of BT and the US LECs, and the NERA model allows for a comparison of how BT's efficiency has changed since the previous comparative efficiency study completed in 2000.

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.

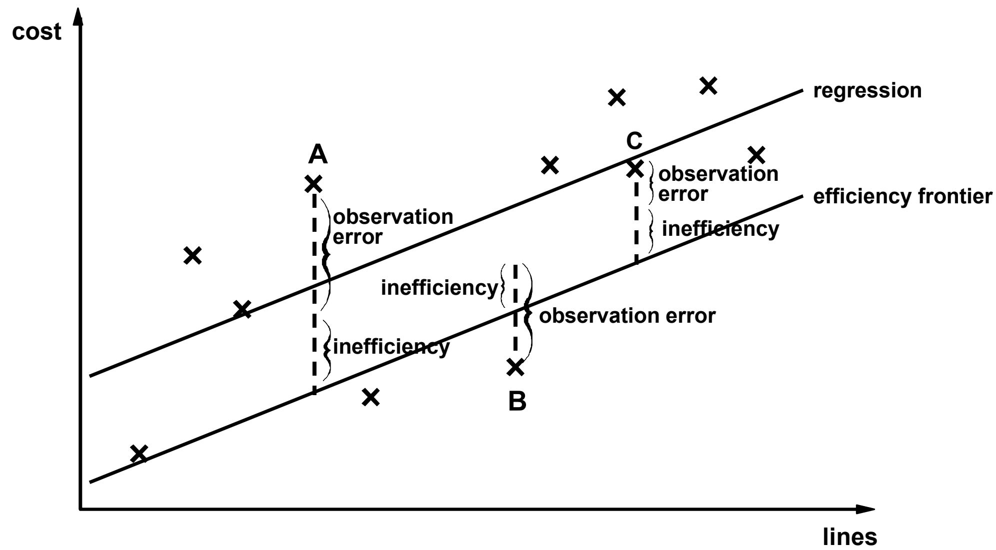
Under the simplest implementation of SFA, an Ordinary Least Squares (OLS) regression is first estimated to capture the sensitivity of costs to differences in operating environments, scale of production or mix of outputs. This model will leave residuals that can be attributed to the two remaining causes 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).[69] 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).

[69] See Jondrow, Knox Lovell, Materov and Schmidt, On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model, Journal of Econometrics, Volume 16 (1982).

APPENDIX B. DATA ENVELOPMENT ANALYSIS (DEA)

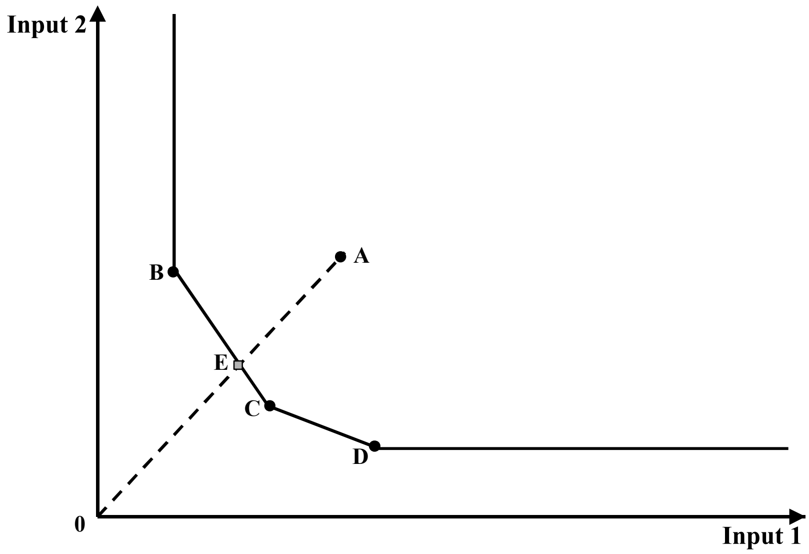
Data Envelopment Analysis (DEA) can be used as an alternative to regression-based techniques. It is not a statistical estimation technique, but rather makes use of mathematical programming methods, without the need to rely on a precise parametric cost function. This, in fact, is its main advantage; it 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.[70]

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.

70 Note that multicollinearity is a potential problem in our dataset.

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;[71]
- 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.

[71] See Charnes A., Cooper W. and Rhodes E., 1978, Measuring the efficiency of decision making units, European Journal of Operations Research, Volume 2, pages 429-444.

APPENDIX C. IMPLIED CVES IN THE BT AND NERA MODELS

The general form of the cost function assumed by BT is:

$$C = BE^{\alpha} L^{\beta} M^{\gamma} \left(S/E\right)^{\delta} X^{\varepsilon}$$

- C is total cost;
- B is a constant;
- E is the number of exchange lines ("switched lines");
- L is the number of leased lines;
- M is the number of switch minutes;
- S is sheath length;
- X is a variable representing other factors that determine costs; and
- α, β, γ, δ, ε are elasticity coefficients, which measure the impact of a change in the volume of the variable concerned on total costs. For example, if α had a value of 0.8, this would mean that a 1% increases in the volume of exchange lines led to a 0.8% increase in total costs.

For the purposes of producing a model that can be used for assessing comparative efficiency, the above equation is converted into logarithmic form:

$$\log C = \log B + \alpha \log E + \beta \log L + \gamma \log M + \delta \log(S/E) + \varepsilon \log X$$

The resulting linear regression equation then provides estimates of the constant term (log B) and the elasticity coefficients (α, β, … ε).

The form of equation used by NERA is a similar function:

$$C = A(L+E)^{a} \left(\tfrac{L}{E}\right)^{b} M^{c} \left(\tfrac{S}{E}\right)^{d} X^{e}$$

This can be expanded to give the same form as BT's model, although it is not possible to directly calculate the size of many of the coefficients. The estimated coefficients for exchange lines (E), leased lines (L) and switch minutes (M), which can be derived from the NERA and BT models, are shown in Table C.1 below. For the BT model the coefficient shown for exchange lines (E) is given by (α - δ), reflecting the fact that exchange lines appear in the equation as a variable in their own right and also in the average sheath length per line variable. NERA's model is specified in a slightly different form: instead of separate exchange line and leased line variables, there is a total line variable; and there is also a variable for the ratio of leased lines to switched lines. In order to derive the coefficients on an equivalent basis to those in the BT model, an estimate was made of the impact on predicted total costs (using the model) of varying (a) the number of exchange lines and (b) the number of leased lines by 1%. For switch minutes the coefficient in the estimated NERA equation could be taken directly.

Table C.1
Comparison of E, L and M Coefficients in the AC and NERA Models

	AC Model	NERA Model
Exchange lines (E)	0.70	0.59
Leased lines (L)	0.05	0.10
Switch minutes (M)	0.08	0.15

Source: NERA analysis

Therefore, the two models can be expressed in a common form as follows[72]:

- $\log C = \log B + 0.70 \log E + 0.05 \log L + 0.08 \log M + \delta \log S + \varepsilon \log X$
- $\log C = \log B + 0.59 \log E + 0.10 \log L + 0.15 \log M + \delta \log S + \varepsilon \log X$

Based upon the above equations we have been identified the following cost-volume elasticities:

- The elasticity of switch line costs with respect to switch lines;
- The elasticity of leased line costs with respect to leased lines; and
- The elasticity of the costs of calls with respect to switch minutes.

To show how this was achieved, take the general form of the cost function and rearrange it so that there is just one term involving exchange lines; therefore, using the BT form of equation:

$$C = BE^{\alpha-\delta} L^{\beta} M^{\gamma} S^{\delta} X^{\varepsilon}$$

Differentiating with respect to exchange lines (E) gives:

$$\partial C / \partial E = (\alpha - \delta) BE^{\alpha-\delta-1} L^{\beta} M^{\gamma} S^{\delta} X^{\varepsilon}$$

$$\partial C / \partial E = (\alpha - \delta) C / E$$

72 As stated above, the coefficients in the NERA model have been identified by expanding the model around BT.

$(\alpha - \delta) = \partial C / \partial E \times E / C$

Given that we are differentiating with respect to exchange lines (i.e. holding all other variables constant), only the costs of exchange lines change, so in the above equation $\partial C = \partial C^E$ where the superscript E denotes exchange line. This in turn means that:

$(\alpha - \delta) = \partial C^E / \partial E \times E / C = \partial C^E / \partial E \times E / C^E \times C^E / C$

where, given that only the costs of exchange lines vary when differentiating with respect o exchange lines, $\partial C^E / \partial E \times E / C$ is the cost-volume elasticity (CVE) of <u>total costs</u> with respect to exchange lines, $\partial C^E / \partial E \times E / C^E$ is the CVE of <u>exchange lines costs</u> with respect to exchange lines, and C^E / C is the share of exchange lines costs in total network costs. In other words, the elasticity of <u>total costs</u> with respect to the number of exchange lines (α - δ) is equal to the cost-volume elasticity of exchange lines costs to exchange lines multiplied by the share of exchange lines in total network costs.

Similarly, β is equal to the CVE for leased lines costs with respect to leased lines multiplied by the share of leased lines in total network costs, and γ is equal to the CVE for calls costs with respect to calls multiplied by the share of calls in total network costs.

Using these relationships, it is possible to derive the CVEs for each output-cost category (e.g. exchange lines costs) implied by the combination of the estimated coefficients in the BT and NERA models (which are shown in Table C.1) and the shares of exchange lines, leased lines and calls in BT network costs (which are shown below).

Examination of BT's regulatory accounts for 2001/2 reveals that, of the total network costs that are included in the efficiency study, exchange lines account for 53.5%, leased lines for 16.2% and core network costs (excluding leased lines) 30.3%. Details of this calculation are provided in the section C.1 below.

Based on these cost shares the implied CVEs for BT are shown in Table C.2.

Table C.2
Cost-Volume Elasticities Implied by the BT and NERA Models

	BT Model	NERA Model
CVE of exchange line costs to exchange lines	1.31	1.10
CVE of leased line costs to leased lines	0.31	0.62
CVE of switch minutes costs to switch minutes	0.26	0.50

Source: NERA analysis

It can be seen that the BT model implies a CVE for exchange lines of 1.31. This suggests very pronounced diseconomies of scale and seems unlikely to be a representation of reality. At the same time, the CVEs for leased lines and switch minutes seem low given that they are long-run CVEs (having been derived from cross-sectional data).[73]

The NERA model also indicates diseconomies of scale in the case of exchange lines, although the extent of such diseconomies is not as large. At the same time, the CVEs for leased lines and switch minutes are not as low as those for the BT model and appear to be more plausible.

The evidence would appear, therefore, to support the view that the NERA model coefficients are more plausible than those in the BT model.

73 Reference to BT's Current Cost Financial Statements for the Businesses and Activities 2003 and 2002 indicates that in 2001/2 LRIC (without hybrid adjustments) was 74% of CCA fully allocated cost in the case of inland private circuits and 97% of fully allocated cost in the case of the core network (excluding inland private circuits) as defined in section C.1. This latter figure is derived after excluding intra core common fixed costs from LRIC. The high value is partly the result of the fact that outpayments are included and, for these, LRIC is the same as fully allocated cost.

C.1. BT Network Cost Shares – 2001/2

<u>**Switched Lines**</u>	**£M**	
Local exchange concentrators	246	
Access network	4024	
Public payphone lines	18	
	4288	*53.5%*
<u>**Leased Lines**</u>	**1295**	*16.2%*
<u>**Core Network (excluding leased lines)**</u>		
Local exchange processor	338	
Main exchange switching	193	
Transmission	386	
Product management, policy and planning	202	
Interconnection connections/rentals	113	
Number portability, ASU, SMDS, DMS 100 etc.	177	
Carrier pre-selection	20	
DDNS and other multifunction	24	
Outpayments (excluding international and calls to mobile)	976	
	2429	*30.3%*

Source: BT Current Cost Financial Statements for the Businesses and Activities 2003 and 2002, pages 21-23