The Provision of Library Services by English Local Authorities

Philip Andrew Stevens[*]

National Institute of Economic and Social Research
2, Dean Trench Street, Smith Square, London, SW1 3HE, UK.
E-mail: philip.stevens@niesr.ac.uk. Tel: 020 7654 1927. Fax: 020 7654 1900.

Abstract

In this paper we examine the cost and efficiency with which library services are provided by English local authorities using stochastic frontier analysis. We test our results using four commonly used assumptions regarding the distribution of the efficiency scores. We find that costs are higher for authorities operating in regions where a higher proportion of the population have low incomes or are of working age. The efficiency scores net of environmental factors, however, are affected by the choice of whether they are considered as quasi fixed factors in the cost frontier or as determinants of inefficiency.

JEL Classification: D20, H40, L32

Keywords: Local Government, Efficiency, Stochastic Frontier Analysis

July 2005

Discussion Paper No. 259

[*] Acknowledgements: This paper originates in work undertaken at the institute on local authority cost effectiveness on behalf of the DETR, DTLR and ODPM. I would like thank Mary O'Mahony, Hiroko Plant, Michela Vecchi, Martin Weale, Willem deBoer and participants at the NIESR conference on 'Productivity and Performance in the Provision of Public Services' at the British Academy for help and comments. All mistakes remain the author's own

The efficiency with which public sector organisations provide their services has become an issue of increasing interest recently, both in the academic literature and in policy debates. The public sector is responsible for £6 billion in the UK. The performance of the public sector and the effectiveness with which this money is spent is, therefore, an issue of considerable importance. This article uses the method of stochastic frontier analysis to assess the efficiency of English Local Authorities (LAs) in providing library services. We test the robustness of our results to a number of assumptions about the distribution of the efficiency scores and to the influence of environmental factors upon LA service provision. In section 2 we provide some background and a short introduction to stochastic frontier analysis. In section set out our empirical models. Section 4 describes out data and in section 5 we present and discuss our results. Section 6 concludes.

1 Efficiency in the Public Sector

The efficiency of the public sector is an issue attracting increasing interest, particularly in the UK. One reason for this is its size. With such a large proportion of GDP being produced in the public sector, any inefficiency is likely to have a major impact upon national output. Another reason is the belief that additional analysis of performance is required in the public sector because it lacks the strictness of the market to ensure efficiency. In particular, frontier analysis has become a popular tool to assess the performance of local government[1]. Adapting a methodology primarily developed or the study of private sector to the public sector is not without its problems. Identifying inputs and outputs of local government, as well as the technology of producing non-market outputs, is a more difficult exercise than for a standard production process within a firm (Wolf, 1993; Borger and Kerstens, 1996; Stevens and Vecchi, 2002). The set of outputs that are provided by the local government sector is quite diverse: ranging from library books to emptying dustbins. In the absence of prices, this makes

[1] See Worthington and Dollery (2000) for an extensive overview of this topic.

measuring the output, and hence assessing the performance of local government as a whole, extremely difficult (Johnes, 1997; O'Mahony, Stevens and Stokes, 2002). Given these problems, any method that relies on there being no errors in measurement (e.g. data envelopment analysis) is inappropriate for such analysis.

Another factor that affects the measurement of public sector efficiency is the influence of background or environmental factors. Service provision may vary across local authorities because of factors that are beyond the control of these authorities. For example, some operate in areas with high levels of poverty or with a particular demographic profile. Unlike a firm operating in the public sector, a local authority cannot move to a more affluent area to increase its profitability. These factors are likely to impact upon the costs of providing services. The issue of how one accounts for such background factors may well influence the estimates of efficiency one obtains (Stevens, 2004).

Stochastic frontier models date back to Aigner, Lovell and Schmidt (1977) and Meesen and van den Broek (1977), who independently proposed a stochastic frontier production function with a two-part 'composed' error term. In the production function/frontier context, where its use is most common[2], this error is composed of a standard random error term, representing measurement error and other random factors, and a one-sided random variable representing what Farrell (1957) called 'technical inefficiency', i.e. the distance of the observation from the production frontier. This notion of *technical efficiency* reflects the ability of an organisation to obtain maximal output(s) from a given set of inputs. *Allocative efficiency* refers to the ability of the organisation to choose the inputs and outputs in correct proportions, given their prices. The combination of technical and allocative efficiency is called *economic efficiency*.

---

[2] See Lovell and Schmidt (1993), Coelli, Rao and Battese (1999) and Kumbhakar and Lovell (2000) for examples.

Technical efficiency is the subject of the majority of studies in the private sector. This is in part because most studies of production in the private sector consider a single output, usually a composite value-added measure. The lack of prices for outputs is at the heart of the problem of analysing the public sector (O'Mahony, Stevens and Stokes, 2003). It is therefore, impossible to obtain such a composite output measure without the imposition of arbitrary output weights (Johnes, 1997) that are likely to be both contentious and difficult to obtain. Because of this, it is vital to account for the multiple-output nature of production when investigating the provision of services by public sector organisations (Stevens, 2004).

At the first remove, it is overall economic efficiency that is important when one investigates the provision of public service, rather than its decomposition into its technical and allocative components. This economic efficiency represents a real welfare loss. Whether the organisation is operating under a policy that matches appropriately with the preferences of voters is another issue again.

2 Empirical Models

In order to investigate the efficiency of library service provision by English LAs we employ a short-run stochastic cost frontier relating expenditure at LA i, E_i, to a set of k outputs Q_k. Our basic cost frontier is of the form

(1) $$\ln E_i = \alpha_0 + \alpha_1 \ln Q_{1A,i} + \alpha_2 \ln Q_{1B,i} + \alpha_3 \ln Q_{2,i} + \alpha_4 \ln Q_{3,i} + \alpha_5 \ln Q_{4,i} + \eta_i + \varepsilon_i$$

where η is a non-negative random variable and the ε_{it} terms are random errors, assumed to be i.i.d. and have $N(0,\sigma_\varepsilon^2)$-distribution, independent of the η_{it} terms. It is this η term that represents inefficiency. Note that unlike the production frontier SFA approach, this inefficiency represents total economic inefficiency, i.e. technical inefficiency (not getting enough output from the inputs) plus allocative inefficiency (not using the inputs or producing the outputs in the correct proportions).

One of the fundamental assumptions regarding the estimation of (1) by SFA is how η is distributed. We will investigate the impact of four common assumptions regarding its distribution: (*i*) that it has a half-normal distribution, (*ii*) that it has a gamma distribution; (*iii*) that it has a normal distribution truncated at zero and (*iv*) that it has a normal distribution truncated at zero with its mean depending on a set of background or environmental factors.

In common with many studies of local authority performance, we do not have the price of input. Whilst in studies of the private sector, this is likely to be an important omission, in many areas of the public sector nationwide wage setting takes place and authorities pay individuals according to a national pay scale and so there will be little or no cross-sectional variation in the wage rates. This is also the case for capital and intermediate products. One variation that does exist is London weighting for wages, to account for the higher cost of living in the capital. Therefore, we include a dummy variable for London Boroughs to account for this exogenous source of higher costs.

The cost frontier specified in (1) ignores the influence of background or environmental factors. Essentially, we might do this in two ways (or a combination of the two). The first is to include these influences as quasi-fixed factors in the frontier itself. Therefore (1) becomes:

$$(2) \qquad \ln E_i = \alpha_0 + \alpha_1 Q_{1A,i} + \alpha_2 Q_{1B,i} + \alpha_3 Q_{2,i} + \alpha_4 Q_{3,i} + \alpha_5 Q_{4,i} + \beta_m \boldsymbol{Z}_m + \eta_i + \varepsilon_i$$

where $\boldsymbol{Z}_m$ is a vector of m background variables and β is a vector of coefficients to be estimated.

The second method to account for background variables is to include them as determinants of the efficiency terms themselves. Following Battese and Coelli (1995), the inefficiency effect is obtained by a truncation of the $N(\eta_{it}, \sigma^2)$-distribution, where

(3) $$\eta_{it} = \delta_m \boldsymbol{Z}_m$$

where δ is a ($1 \times M$) vector of unknown scalar parameters to be estimated (which includes an intercept parameter).

In this study we will investigate both of these methods and assess their impact upon our estimation of efficiency.

3 Data

Our data on expenditure and outputs come from the 2000/1 Best Value Performance Indicators produced by the Department for Transport, Local Government and the Regions (local government is now the responsibility of the Office of the Deputy Prime Minister). These figures were initially presented as per head of population, so they were converted back into real terms using the resident population at midyear on 30 June 1998. Descriptive statistics of these variables are presented in Table 1.

Table 1 Inputs and outputs

Variable		*Mean*	*Std Dev*	*Min*	*Max*
Q_{1A}	Number of books issued by the authority's libraries	2,678,115	2,335,234	272,322	13,600,000
Q_{1B}	Number of other items issued by the authority's libraries	217,276	196,061	12,849	1,100,455
Q_2	Number of visits to public libraries	2,004,709	1,624,734	170,960	9,735,791
Q_3	Number of books and recordings available in the councils libraries	389,163	264,880	33,193	1,665,005
C	Net expenditure on libraries	4,217,533	2,817,699	372,971	16,200,000

There are four types of LA that provide library services: London Boroughs, Metropolitan Authorities, Unitary Councils and County Councils. We can see how these indicators vary by authority type by looking at Table 2. By far the largest in terms of expenditure are the County Councils, with the London Boroughs and Metropolitan Authorities being of similar size. In our estimation we will include a dummy variable to account for the exogenous influence of local authority type on costs.

Table 2 Mean value of Indicators by Authority Type

	Q_{1A}	Q_{1B}	Q_2	Q_3	C
London borough	1,647,425	203,789	1,579,477	340,814	4,117,981
Metropolitan Authority	2,143,294	163,466	1,611,446	361,885	4,163,234
Unitary council	1,359,996	98,985	981,051	216,383	2,165,319
County council	5,959,378	444,033	4,177,045	691,101	7,063,784
Total	2,678,115	217,276	2,004,709	389,163	4,217,533

The environmental factors that we will take into account in our analysis are presented in Table 3. The first factor is the impact of population density, *KIDDENS*, measured by the population per hectare. The dispersion of the population is likely to affect the ability of the authority to concentrate provision in a small number of central libraries. Conversely, sparsely populated local authorities may have to provide mobile library services to cover large areas. The second factor is the percentage of the population who are income deprived, as defined by the Income Domain of the Index of Deprivation, *INCIID*. The effect of this on the demand for library services is uncertain. Whilst demand for books generally may be higher in more affluent areas, it is possible that lower income households substitute library services for the private purchases. The next two variables account for differences in the age profile of the population, the percentage of the population who are aged five to fifteen years of age, *YOUNGPOP*, and the percentage who are aged over sixty-five years, *OLDPOP*. We include a variable to account for the ethnic diversity of local authority populations, the percentage of the population for whom English is a second language. We also include dummy variables for LA type (the baseline LA type is Unitary Authority).

Table 3 Background variables

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*
KIDDENS	Population per hectare	24.49	26.98	0.62	142.32
INCIID	% of the population who are income deprived	0.24	0.09	0.07	0.51
YOUNGPOP	% population aged 5 - 15	0.14	0.01	0.09	0.18
OLDPOP	% population aged over 65	0.15	0.03	0.10	0.23
E2L	% population with English as a second language	0.09	0.12	0.00	0.49
Local Authority Types					
LONDON	London Borough	0.22	0.41	0	1
METRO	Metropolitan Authority	0.25	0.44	0	1
UNITARY	Unitary Authority	0.30	0.46	0	1
COUNTY	County Council	0.23	0.42	0	1

The London boroughs are the most densely populated local authority type, although there is considerable variation of population density within all authority types (Table 4). Metropolitan and unitary authorities are fairly indistinguishable in this respect, although the variation with the former is greater than that within the later authority type. The differences between authority types with respect to deprivation are much smaller. County councils have, on average, the smallest proportion of people who are classed as income deprived. There is little difference between authority type variation in the percentage of the population who are aged between five and fifteen years of age, but there is a little for the percentage who are over sixty five. London boroughs tend to have slightly less than the average proportion of older people and county councils slightly more. The background variable with the most between-authority type variation is the percentage of the population for whom English is a second language. Almost one quarter of those in London boroughs have English as a second language, whereas less than two percent of those in county councils do. There is, however, considerable within authority type variation in this variable. The highest percentage of population for whom English is a second language among county councils is only 6.7%. In contrast, nine out of ten London boroughs have a figure that is higher than this.

Table 4 Background Variables by Local Authority Type

	KIDDENS	*INCIID*	*YOUNGPOP*	*OLDPOP*	*E2L*
London Borough	61.34 (31.60)	0.248 (0.106)	0.137 (0.019)	0.129 (0.018)	0.244 (0.138)
Metropolitan Authority	20.05 (9.65)	0.296 (0.063)	0.148 (0.008)	0.155 (0.013)	0.063 (0.067)
Unitary Authority	18.6 (15.61)	0.232 (0.082)	0.146 (0.012)	0.151 (0.029)	0.053 (0.081)
County Council	2.555 (1.43)	0.168 (0.036)	0.138 (0.005)	0.171 (0.024)	0.018 (0.018)
Total	24.49 (26.98)	0.237 (0.087)	0.143 (0.013)	0.152 (0.026)	0.089 (0.119)

4 Results

The results of our estimation of the basic cost frontiers (i.e. excluding background variables) are presented in Table 5. The results for all three models are very similar. The γ ratios for all three specifications suggest that traditional average response function is not an adequate representation of the data. The effect of the assumption regarding the distribution of the efficiency terms on the estimates of the α parameters is minimal. The μ term of the truncated normal model (2) is insignificant. The fact that the mean of the half normal is insignificantly different from zero means that one can accept the restricted form (1) over (2).

In all three specifications the coefficients on the four output measures have the expected positive sign, although the coefficient on $\ln Q_{1A}$ is insignificantly different from zero. At first it may seem rather odd that the costs of providing library services do not depend in a significant way upon the number of books issued by the authority's libraries. However, it appears that the more important measures are the number of items made available at the council's libraries (Q_3) and the number of visits made to them (Q_2). It may of course merely be a result of misspecification, i.e. of excluding the background variables from our analysis. There does however appear to be a cost premium attached to 'other items' issued by the authority's libraries (Q_{1B}). These items account for around 10% of issues (Table 1), although they are more common in libraries in London Boroughs than elsewhere (Table 2). Overall, Costs tend to be higher in London Boroughs and Metropolitan Districts than in County Councils and Unitary Authorities, *ceteris paribus*.

Table 5 Excluding background variables

	(1) Half normal	(2) Truncated normal	(3) Exponential
CONSTANT	2.755***	2.622***	2.601***
	(0.698)	(0.661)	(0.777)
$\ln Q_{1A}$	0.106	0.166	0.170
	(0.115)	(0.133)	(0.139)
$\ln Q_{1B}$	0.104**	0.093**	0.093**
	(0.043)	(0.045)	(0.046)
$\ln Q_2$	0.393***	0.366***	0.364***
	(0.095)	(0.098)	(0.100)
$\ln Q_3$	0.289***	0.277***	0.276***
	(0.058)	(0.056)	(0.056)
LONDON	0.185***	0.172***	0.171***
	(0.060)	(0.061)	(0.062)
METRO	0.120**	0.110**	0.109**
	(0.053)	(0.052)	(0.052)
COUNTY	-0.053	-0.070	-0.071
	(0.080)	(0.081)	(0.081)
μ		-0.951	
		(1.203)	
γ	0.762***	0.915***	0.543***
	(0.170)	(0.080)	†
Log likelihood	18.784	19.268	19.229
Observations	139	139	139

Standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*
† *Note that the γ term is not calculated as part of the estimation of the exponential model so we calculate the γ term using the estimates of σ_η and σ_ε. Instead the parameter $\lambda = \sigma_\eta^2 / \sigma_\varepsilon^2$ is estimated. The significance indicated is in fact the significance test based on the estimate of λ.*

Inefficiency, and hence costs, tend to be higher in London Boroughs (Table 6), *ceteris paribus*, although within-authority-type variation in efficiency is higher among London

Boroughs than in other authority types. There is little variation in mean efficiency scores between the other authority types. We will discuss the efficiency scores in more detail below. The implication of the results in Table 6 is that costs in London Boroughs tend, on average, to be around 30 percent above the feasible minimum, given their output mix. In other authority types, costs are between 15 and 25 percent above the feasible minimum.

Table 6 Mean Inefficiency by Authority type

	(1)	(2)	(3)
London Borough	1.305 (0.247)	1.290 (0.284)	1.281 (0.284)
Metropolitan Authority	1.250 (0.180)	1.207 (0.191)	1.197 (0.188)
Unitary Authority	1.229 (0.116)	1.172 (0.103)	1.161 (0.097)
County council	1.212 (0.098)	1.156 (0.083)	1.146 (0.078)
Total	1.247 (0.168)	1.203 (0.182)	1.192 (0.180)

Std. Dev. in parenthesis

4.1 Including Environmental Factors

The immediate question that arises from this result is why there is this variation in efficiency scores. It may be the case that certain local authorities are less efficient in their management and organisation. However, as we noted in the introduction, all local authorities do not operate in the same environment. The variation in the background variables may go some way to explain the differences in efficiency.

Taking account of the environment in which the authority operates affects our results considerably (Table 7). Once again, there is little variation between the results of specifications (4) to (6). Therefore, the assumption made about the distribution of the efficiency scores has little influence on the results. Once we take the characteristics of the authority's population in to account, the relationship between the number of books issued and costs becomes significant. The effects of local authority type become smaller in size and lose some of their significance, suggesting that part of the explanation for higher costs in London Boroughs and Metropolitan authorities is due to the characteristics of their populations. Costs are lower in areas with a larger proportion of the population who are either of school age or are retired. That is to say, costs are higher in areas where a large population of the population are of prime or working age (16-64). The dispersion and ethnic mix of the population does not appear to have any direct influence on the costs of providing library services.

Costs are higher in deprived areas. The result that serving more deprived areas incurs higher costs than more affluent ones, for a given level and mix of output, is an interesting one and worthy of further investigation. This effect is over and above differences in the demand for library services – as measured by the number of visits made and the number of books and other items issued – and their supply, as measured by the number of books and recording available. Therefore, explanations that the higher costs of provision in poorer areas are due to higher demand among people who cannot afford to purchase books and the like do not tell the whole story. The effect may work on the *composition* of demand. Income deprived individuals may demand more expensive items or services. They may place higher demands on library staff. Another potential source is crime; the links (although not necessarily causal) between poverty and crime are well established (e.g. Pudney, Deadman, and Pyle, 2000). It may be the case that libraries in income deprived areas are more susceptible to theft or damage. However, it is not possible to differentiate between these, and other, explanations of this relationship without further, more detailed investigation.

Table 7 Results - Including Background Variables in Frontier

	(4) Half normal	(5) Truncated normal	(6) Exponential
Constant	3.825***	3.871***	3.877***
	(0.595)	(0.597)	(0.599)
lnQ_{1A}	0.373***	0.375***	0.375***
	(0.090)	(0.088)	(0.088)
lnQ_{1B}	0.079**	0.080**	0.080**
	(0.036)	(0.036)	(0.036)
lnQ_2	0.233***	0.231***	0.230***
	(0.078)	(0.078)	(0.077)
lnQ_3	0.183***	0.183***	0.183***
	(0.054)	(0.052)	(0.052)
LONDON	0.103*	0.104*	0.104*
	(0.061)	(0.061)	(0.061)
METRO	0.081*	0.081*	0.080*
	(0.046)	(0.045)	(0.045)
COUNTY	0.149	0.150	0.149
	(0.091)	(0.092)	(0.092)
LNKIDDENS	0.032	0.032	0.032
	(0.025)	(0.025)	(0.025)
YOUNGPOP	-6.057***	-6.060***	-6.068***
	(1.728)	(1.713)	(1.708)
OLDPOP	-3.891***	-3.868***	-3.864***
	(0.874)	(0.869)	(0.867)
E2L	0.078	0.074	0.072
	(0.222)	(0.221)	(0.221)
INCIID	1.675***	1.664***	1.664***
	(0.308)	(0.308)	(0.307)
μ		-0.517	
		(0.587)	
γ	0.608***	0.752***	0.227***
	(0.069)	(0.159)	†
Log likelihood	66.786	66.787	66.787
Observations	139	139	139

Standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*

† *Note that the γ term is not calculated as part of the estimation of the exponential model so we calculate the γ term using the estimates of σ_η and σ_ε. Instead the parameter $\lambda = \sigma_\eta^2 / \sigma_\varepsilon^2$ is estimated. The significance indicated is in fact the significance test based on the estimate of λ.*

Once one accounts for differences in the demographic characteristics of the local authority's population, the variation in mean efficiency scores falls both between and within authority type (Table 8). Once the characteristics of the authorities' populations are taken account of, the mean inefficiency falls from somewhere between nineteen and twenty-five percent (depending on the assumption made about the distribution of the efficiency term) to between nine and thirteen percent. This suggests that some of the variation in costs of providing a given level of service is beyond the control of the local authority.

Table 8 Mean Inefficiency by Authority type

	(4)	(5)	(6)
London Borough	1.141 (0.076)	1.104 (0.074)	1.098 (0.075)
Metropolitan Authority	1.135 (0.068)	1.099 (0.063)	1.092 (0.062)
Unitary Authority	1.133 (0.060)	1.095 (0.050)	1.088 (0.047)
County council	1.126 (0.039)	1.087 (0.027)	1.080 (0.025)
Total	1.134 (0.061)	1.096 (0.055)	1.090 (0.054)

Std. Dev. in parenthesis

The results are similar if we believe that the background variable affects the efficiency of the local authority (i.e. its distance from the minimum cost frontier) rather than the minimum cost frontier itself (Table 9). The direction of the influence of the background variables is similar in both models (4) to (6) and models (7) and (8)[3]. One difference is the effect of the density of the authority's population. Densely populated areas tend to be more inefficient, and hence have higher costs, although the coefficient is only significant at the 10% level.

[3] Note that a reduction in inefficiency – i.e. δ_k is negative – has the same effect as a reduction in direct costs – i.e. β_k is negative.

Table 9 Results - Including background variables in efficiency determinants

	(7) background effects	(8) background effects - LA
Constant	1.710***	2.028***
	(0.483)	(0.512)
$\ln Q_{1A}$	0.467***	0.443***
	(0.079)	(0.076)
$\ln Q_{1B}$	0.069**	0.051**
	(0.032)	(0.031)
$\ln Q_2$	0.198***	0.209**
	(0.075)	(0.070)
$\ln Q_3$	0.216***	0.225***
	(0.046)	(0.045)
LONDON	0.124*	0.069
	(0.063)	(0.053)
METRO	0.032	0.072*
	(0.061)	(0.043)
COUNTY	-0.115*	-0.080
	(0.066)	(0.063)
Efficiency determinants		
CONSTANT (μ)	1.046*	1.319**
	(0.616)	(0.569)
ln(*KIDDENS*)	0.201**	0.134*
	(0.094)	(0.069)
YOUNGPOP	-7.259***	-8.624***
	(2.631)	(2.553)
OLDPOP	-7.965***	-6.812***
	(2.376)	(1.925)
INCIID	2.397***	2.790***
	(0.548)	(0.540)
E2L	-0.486	-0.340
	(0.386)	(0.361)
LONDON		0.079
		(0.117)
METROPOLITAN		0.228**
		(0.102)
COUNTY		0.339
		(0.578)
γ	0.611***	0.519***
	(0.163)	(0.177)
Log likelihood	62.109	65.846
Observations	139	139

Standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*

As with the results for models (4) to (6), the effect of authority type on the cost frontier is reduced once one accounts of for environmental factors. As one would expect, one difference between models (4) to (6) and models (7) and (8) is that the difference in mean inefficiency of authority types returns (Table 10). In model (8), we test the significance of this difference by including authority type in the vector of efficiency determinants, ***Z***. It is interesting to note that the authority type with the highest mean inefficiency (London Boroughs) is not significantly different from that of Unitary Authorities, once one accounts for differences in environmental influences, but Metropolitan Authorities are. Thus the high mean inefficiency score for London Boroughs shown in Table 10 is entirely explicable in terms of the characteristics of their populations; that of Metropolitan Authorities is not.

Table 10 Mean Inefficiency by Authority type

	(7)	(8)
London Borough	1.667 (0.573)	1.626 (0.547)
Metropolitan Authority	1.328 (0.294)	1.329 (0.300)
Unitary Authority	1.181 (0.143)	1.122 (0.113)
County council	1.048 (0.026)	1.046 (0.027)
Total	1.292 (0.381)	1.266 (0.369)

Std. Dev. in parenthesis

4.2 Efficiency Scores

The assumption made about the distribution of efficiency scores has a much smaller effect on the measured inefficiency than the omission of environmental factors (Table 11). The descriptive statistics for models (1) to (3) are very similar. The mean and spread of the efficiency scores falls when we include background variables in our specification of the minimum cost frontier (models (4) to (6)). As one would expect, if we include these background effects as determinants of economic efficiency, the mean value of inefficiency is similar to that obtained from the basic specification. However, note that the dispersion of the scores is much higher in models (7) and (8) that it is for models (1) to (3). This is because these are gross inefficiency scores. We will return to this below.

Table 11 Descriptive Statistics of Efficiency Scores

	Min	*Max*	*Mean*	*s.d.*
Model (1)	1.047	2.034	1.247	0.168
Model (2)	1.038	2.200	1.203	0.182
Model (3)	1.037	2.211	1.192	0.180
Model (4)	1.044	1.424	1.134	0.061
Model (5)	1.034	1.434	1.096	0.055
Model (6)	1.033	1.444	1.090	0.054
Model (7)	1.017	2.753	1.223	0.343
Model (8)	1.010	2.435	1.204	0.306

The descriptive statistics in Table 11 only tell half of the story. Two variables could have similar means, standard deviations, minima and maxima, but still be completely uncorrelated. Therefore, in Table 12 we chart the correlations between our estimates of inefficiency. There is a significant degree of correlation between all of our estimated efficiency scores. This is true if we consider the correlation between the scores themselves (as measured by the Pearson correlations) or the correlation of the implied rank of the authorities in the league table of efficiency (as measured by the Spearman correlation). In what follows we will consider the specifications in three groups: (I) the basic specifications without background variables, models (1) to (3); (II) the equivalent models with background variables included in the frontier, models (4) to (6); and (III) those which include the background variables in the determinants of efficiency, models (7) and (8). If we consider the correlations between these three groups, certain patterns emerge. First, the correlation within these three groups is much higher than that between them. Therefore, the effect of the assumption one makes regarding the distribution of the efficiency scores is less important than the exclusion of relevant factors. A similar result has been found when comparing the effect of function form of a production relationship versus the choice of explanatory variables (Kneller and Stevens, 2003). Second, the correlation between (I) and (III) is much higher than either of them with (II). This is to be expected as (I) and (III) make similar assumptions about the frontier itself. The only difference between them is that (III) model the determinants of the efficiency terms, i.e. the distance of the observations from the frontier. Note that if the variables in the efficiency determinants were uncorrelated with the frontier variables, the β coefficients in groups (I) and (III) would be identical, with the exception of the intercept[4]. The difference in the results is due to the fact that some of the variation in costs that (I) attributes to the outputs or authority type on the production technology, can be attributable to economic efficiency once we model the determinants of efficiency.

[4] See Coelli Rao and Battese (1998) or Kumbhakar and Lovell (2000) for more on this.

Table 12 Correlations of Efficiency Scores

	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)	Model (7)	Model (8)
Pearson Correlations							
Model (1)	0.985	0.979	0.794	0.793	0.788	0.810	0.843
Model (2)		1.000	0.765	0.785	0.785	0.870	0.895
Model (3)			0.760	0.784	0.785	0.873	0.898
Model (4)				0.979	0.967	0.471	0.500
Model (5)					0.998	0.496	0.520
Model (6)						0.500	0.527
Model (7)							0.927
Spearman (Rank) Correlations							
Model (1)	0.998	0.997	0.788	0.791	0.792	0.560	0.638
Model (2)		1.000	0.792	0.795	0.796	0.587	0.654
Model (3)			0.792	0.796	0.797	0.587	0.653
Model (4)				1.000	1.000	0.262	0.245
Model (5)					1.000	0.268	0.251
Model (6)						0.270	0.253
Model (7)							0.830

All correlations significant at the 1% level

The reason why the lower levels of correlation between (III) and (II) are to be expected is that as the measured efficiency in (II) is net of environmental factors and those of (III) are gross measures. We can calculate the net efficiency scores for (III) by recalculating the conditional expectation of η_{it}, replacing $\delta\boldsymbol{Z}_{it}$ with $\min(\delta\boldsymbol{Z}_{it})$[5]. This gives us measures of efficiency levels when all authorities are assumed to face the most favourable conditions. The descriptive statistics of the calculated net efficiency scores for model (7) and (8) are presented in Table 13. The correlations between these scores, presented as (7)* and (8)*, and the other scores are presented in Table 14.

[5] C.f. Coelli, Perelman and Romano (1999) and Stevens (2004)

Table 13 Descriptive statistics of net efficiency scores

	Min	*Max*	*Mean*	*Std Deviation*
*Model (7)**				
London Borough	1.018	1.245	1.047	0.048
Metropolitan Authority	1.016	1.073	1.026	0.013
Unitary Authority	1.012	1.043	1.022	0.005
County council	1.016	1.030	1.020	0.003
Total	1.012	1.245	1.028	0.025
*Model (8)**				
London Borough	1.019	1.232	1.050	0.047
Metropolitan Authority	1.018	1.109	1.033	0.021
Unitary Authority	1.014	1.051	1.025	0.006
County council	1.019	1.040	1.024	0.005
Total	1.014	1.232	1.032	0.026

* *Net efficiency scores*

As one would expect, the *net* efficiency scores are much lower than the respective *gross* efficiency scores. Moreover, much of the variation in authority's efficiency has been explained by the background variables. If one compares these net efficiency scores with those of the other models we find that they are more in fact more correlated with the models that exclude the background variables altogether than they are with those that include them in the minimum cost frontier. The rank correlations are slightly higher, but still lower than those with Models (1) to (3).

Table 14 Correlations with Net Efficiency Scores

	Pearson		*Spearman*	
	Model (7)*	Model (8)*	Model (7)*	Model (8)*
Model (1)	0.803	0.865	0.868	0.916
Model (2)	0.875	0.929	0.894	0.937
Model (3)	0.885	0.938	0.894	0.938
Model (4)	0.613	0.668	0.711	0.743
Model (5)	0.675	0.721	0.716	0.748
Model (6)	0.690	0.733	0.717	0.749
Model (7)	0.875	0.898	0.804	0.722
Model (8)	0.846	0.899	0.727	0.756
Model (7)*		0.983		0.972

** Net efficiency scores*
All correlations significant at the 1% level

This result is similar to the results of Coelli, Perelman and Romano (1999), who perform a similar exercise accounting for environmental influences in their production frontiers for international airlines. Coelli, Perelman and Romano (1999), like us, find that the mean of net efficiency is higher when the environment is included in the efficiency determinants than when it is included in the frontier[6]. Like us, they find a lower rank correlation between the two sets of net efficiency scores than that between the net efficiency scores of the model with the background variables in the efficiency determinants and the standard frontier without background variables[7]. Curiously, Coelli *et al.* do not appear to be troubled by this result. However, this result suggests that the choice of empirical model will influence both the estimate of efficiency (and hence the potential for cost savings) and the ranking of the local authority.

[6] Note that because Coelli, Perelman and Romano (1999) model the (primal) production frontier, their efficiency scores are less than or equal to 100% since inefficiency *reduces* output whereas in our (dual) cost frontier model, inefficiency *increases* costs and hence scores are greater than or equal to 100%.
[7] Note that unlike us, Coelli *et al.* do not estimate an equivalent specification to our Models (1) to (3), but rather calculated gross efficiency scores by recalculating the efficiency term with the assumption that the firm faces the most favourable environment.

As Coelli *et al.* and Stevens (2004) note, the *ex ante* selection of a preferred method is a difficult task. Coelli *et al.* prefer to include the background variables in the set of efficiency determinants. This is because they utilise the (primal) production frontier approach and the frontier thus represents 'the upper boundary of the production possibilities set, irrespective of environmental issues' (p. 267). Like us, Stevens (2004) in his study of institutions of higher education utilises the (dual) cost frontier approach to take into account of the multiple-output nature of university provision. In this situation, the choice is less clear cut. This is because a cost frontier could include these terms as quasi-fixed factors. In this case it would appear that it is more appropriate to include them in the cost frontier. Thus efficiency is calculated as the additional costs over and above those required by the vector of outputs and the caused by the environmental factors beyond the control of the authority[8].

Coelli *et al.* propose a data driven method to overcome the problem of specification choice, comparing both models (i.e. background variables in the frontier and background variables in the efficiency determinants) with an artificial encompassing equation with the background variables in both. Coelli *et al.* find that the null hypothesis associated with the model with background variables in the efficiency determinants against the encompassing equation can be accepted, but that which includes them in the frontier cannot. Unfortunately, we cannot do so for both models (7) and (8). We can accept the null for the former at the 10% level, although not at the 6% level, but not the latter[9]. This ambiguous result means that we cannot appeal to the data to resolve our modelling problem. Therefore, we can feel more satisfied in our preference for including the background variables as quasi fixed factors in the frontier than we would if the data did not support this thesis.

[8] Note that in Stevens (2004), it is less certain that variables are beyond the control of the universities (at least in part). He therefore proposes a statistical method to overcome this.
[9] The likelihood ratio versus the encompassing model for model (7) is 7.75 and for model (8) is 6.51.

5 Conclusions

In this paper we have used the technique of stochastic frontier analysis to investigate the determinants of the costs and efficiency of English local authorities in providing library services. We have found that costs are higher for authorities operating in regions where a higher proportion of the population have low incomes or are of working age (as opposed to being of school or retirement age). There does not appear to be any relationship between costs and the ethnic mix of the region, as measured by the proportion of the population for whom English is a second language. The explanation for this may be that this factor is correlated with our measure of deprivation. There may be some relation between population density and the efficiency with which library services are provided, but this effect is of only marginal statistical significance and is not robust to specification.

These results are robust to the choice of distribution of the efficiency scores. These results (with the exception of the results for population density noted above) are also unaffected by how we chose to model the influence of background or environmental factors. The coefficients on the outputs and the net efficiency scores are effected less by whether we consider their effect on costs do be direct (by moving the minimum cost frontier) or indirect (by affecting the efficiency of authorities) than they are by the exclusion of these factors from the analysis. This suggests that studies of local authority performance that exclude important, exogenous influences on the costs of service provision will overestimate the level of inefficiency in the system.

Bibliography

Battese, G.E., and Coelli, T.J., (1993a), 'A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects', Working Papers in Econometrics and Applied Statistics, No. 69, Department of Econometrics, Armindale, NE: University of New England.

Battese, G.E., and Coelli, T.J., (1993b), 'A Model for Technical Efficiency Effects in a Stochastic Frontier Production Function for Panel Data', *Empirical Economics*, 20, pp. 325-32.

Battese, G.E., and Corra, G.S., (1977), 'Estimation of a Production Frontier Model: With Empirical Applications in Agricultural Economics', *Agricultural Economics*, 7, pp. 185-208.

De Borger, B., and Kerstens, K., (1996), 'Cost efficiency of Belgian local governments: A comparative analysis of FDH, DEA, and econometric approaches', *Regional Science and Urban Economics*, vol. 26(2), pp. 145-170.

Coelli, T., Perelman, S., and Romano, E., (1999), 'Accounting for Environmental Influences in Stochastic Frontier Models: With Application to International Airlines', *Journal of Productivity Analysis*, 11, pp. 251-73.

Coelli, T., Rao, D.S.P., and Battese, G.E., (1999), *An Introduction to Efficiency and Productivity Analysis*, London: Kluwer Academic Publishers.

Farrell, M.J., (1957), 'The Measurement of Productive Efficiency', *Journal of the Royal Statistical Society Series A*, 120, pp. 253-90.

Johnes, G., (1997), 'Costs and Industrial Structure in Contemporary British Higher Education', *Economic Journal*, 107, pp. 727-37.

Kneller, R., and Stevens, P.A., (2003), 'The specification of the aggregate production function in the presence of inefficiency', *Economics Letters*, 81.2, pp. 223-6.

Kumbhakar, S.C., Ghosh, S., and McGuckin, J.T., (1991), 'A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms', *Journal of Business and Economic Statistics*, 9, pp. 249-86.

Kumbhakar, S.C., and Lovell, C.A.K., (2000), *Stochastic Frontier Analysis*, Cambridge: Cambridge University Press.

Lovell, C.A.K., and Schmidt, S.S. (eds.), (1993) *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press.

O'Mahony, M., Plant, H., and Stevens, P.A., (2003), 'Measuring the Cost Effectiveness of Local Authorities', report to the Office of the Deputy Prime Minister.

O'Mahony, M., Stevens, P.A., Vecchi, M., and Weale, M.R., (2001), 'Measuring the Cost Effectiveness of Local Authorities', report to the Department for Transport, Local Government and the Regions.

Pitt, M.M., and Lee, L-F., (1981), 'Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry', *Journal of Development Economics*, 9, pp. 43-64.

Pudney, S., Deadman, D., and Pyle, D., (2000), 'The relationship between crime, punishment and economic conditions: is reliable inference possible when crimes are under-recorded?' *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 163, no. 1, pp. 81-97(17)

Reifschnieder, D., and Stevenson, R., (1991), 'Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency', *International Economic Review*, 32, pp. 715-23.

Stevens, P.A., (2004), 'Accounting for Background Variables in Stochastic Frontier Analysis', National Institute of Economic and Social Research Discussion Paper Number 239.

Stevens, P.A., and Vecchi, M., (2002), 'The effect of environmental variables on public service efficiency', Presentation to the British Academy, 19th November 2002.

Verry, D, and Davies, B., (1976), *University Costs and Outputs*, Elsevier: Amsterdam.

Vitaliano, D.F., (1997), 'X-Inefficiency in the Public Sector: The Case of Libraries', *Public Finance Review*, 25.6, pp. 629-43.

Wolf, C., (1993), *Markets or Governments: Choosing Between Imperfect Alternatives*, Cambridge: MIT Press.

Worthington, A.C., (1999), 'Performance Indicators and Efficiency Measurement in Public Libraries', *Australian Economic Review*, 32, pp. 31–42.

Worthington, A.C. and Dollery, B.E., (2000), 'Efficiency Measurement in the Local Public Sector: Econometric and Mathematical Programming Frontier Techniques', Discussion Papers in Economics, Finance and International Competitiveness No. 78 , Queensland University of Technology, 1-21.