Article:

Wheat, P orcid.org/0000-0003-0659-5052 (2017) Scale, quality and efficiency in road maintenance: Evidence for English local authorities. Transport Policy, 59. pp. 46-53. ISSN 0967-070X

https://doi.org/10.1016/j.tranpol.2017.06.002

Scale, quality and efficiency in road maintenance: evidence for English Local Authorities

May 2017

Phill Wheat*

Abstract

This paper outlines the first econometric stochastic frontier efficiency analysis of road maintenance costs for local authorities in England in the academic literature. It is motivated by current public sector austerity requiring local authorities to provide efficient highway functions both in terms of learning from best practice (economic efficiency) and potential reorganisation to exploit economies of scale (scale efficiency). The analysis utilises a road condition measure and an end user (public) satisfaction indicator as well as road length and traffic factors. The availability of public satisfaction data is particularly important as incorporation of such a measure into benchmarking is currently in its infancy in economic regulation but is increasing in prominence, such as in regulation of health care. Evidence is found for an optimal road length which has implications for the current trend to merge the delivery of highways services across local authorities. Bigger is not necessarily better. A positive relationship is found between public satisfaction and cost which is strongest for very low or high public satisfaction. Finally, the median cost efficiency is 83% which implies many authorities have the opportunity to save substantial sums by adopting best practice without reducing service quality.

*University of Leeds, p.e.wheat@its.leeds.ac.uk

1. Introduction

This paper outlines econometric analysis of road maintenance costs for local authorities (LAs). It is motivated by the need for LAs to provide their highways functions efficiently, as well as to inform policy makers as to the cost of higher road quality in the presence of greater public concern as to the condition of local highways. For many years and particularly since 2008, LAs in Britain have been under pressure to do more with less. This has been especially challenging for highway departments as, over the same time period, public satisfaction with road conditions has been falling.

To incentivise British LAs to determine and adopt best practice, there have been two key ‘carrot and stick’ initiatives at the national level. Firstly, the Highways Maintenance Efficiency Programme (HMEP) is ‘by the highways sector, for the highways sector’ and aims to “[work] with people and organisations to enable change, so that greater savings and efficiencies can be achieved and the demand for improved roads and services can be met” (HMEP, 2013). HMEP has supported a number of initiatives (including funding early work on this particular piece of work) to help LAs meet these challenges. Secondly, the Department for Transport is moving to make a proportion of highway maintenance funding conditional on demonstrating achievement of several performance-related processes and outcomes (known as the “Incentive Fund”). One criteria is that LAs engage in performance benchmarking and, eventually, use this information to identify and share better practices to raise performance (DfT, 2015 Q15).

The econometric analysis in this paper utilises a data set collected specifically for the project which includes 51 LAs. As well as road length and traffic explanatory variables, the analysis also uses road condition measures as well as end user (public) satisfaction indicators. This enables quantification of the cost of providing a given quality of road as measured both in the physical (the condition of the asset) and perceived sense (by the ultimate customers/users of the LA highway). How to incorporate final user perception of service quality into both benchmarking and price caps is an area of active development in economic regulation, with some regulators, such as Monitor, the National Health Service regulator in the UK, beginning to incorporate such measures into its benchmarking (Monitor, 2016). This analysis aims to demonstrate that final end user measures can enhance cost benchmarking.

The pattern of fiscal austerity has been replicated in many EU countries post the financial crisis and this work, particularly the conclusions on minimum efficient scale, have implications internationally. Furthermore, the analysis seeks to demonstrate that collaborative econometric benchmarking is feasible in a local government setting and that this approach can be extended to other local government functions or into other countries.

Following this introduction, section 2 outlines the literature on cost and performance analysis in highway maintenance. Section 3 outlines the data available for this study and section 4 outlines the methodology. Section 5 discusses the results in four parts, namely scale economies, the cost impact of traffic, the findings on the cost relationship of quality (including public satisfaction) and the predictions of economic efficiency for participating LAs. Section 6 concludes on key findings and discusses the broader national and international applicability of the findings.

2. Literature review

2.1. Performance studies in roads

There exists an established literature which has attempted to develop performance measures for road systems. Many studies focus on key performance indicators, also known as partial performance metrics, to indicate how an organisation is performing across a range of measures e.g. cost per road km and congestion minutes per population. Studies include Pinkney and Marsden (2013), Hartgen et al (2008), Hartgen and Krauss (1993).

Key performance indicators are relatively clear to understand but limited in the sense that they do not provide a single comparative metric to gauge performance across organisations. As such, organisations tend to perform well on some indicators and poor on others, which is of limited value given there is an explicit trade-off between cost, quality and customer satisfaction (Pinkney and Marsden, 2013 p. 15). Similar discussion has been undertaken elsewhere in the road sector literature, for example by Goode et al (1993) for US State highway performance analysis of Hartgen and Krauss (1993) and Litman (2009) for Canadian highway analysis by Hartgen et al (2008). More generally issues of how to use performance measures are discussed across government departments (e.g. GAO, 1999).

One approach when faced with multiple performance indicators is to utilise statistical approaches such as principal component and cluster analyses to try to group organisations into those which appear to have common performance features e.g. Hendren and Niemeier (2008). However, this still has the limitation that it does not quantify the trade-off between the different aspects of performance. To address this issue, this analysis/paper develops a cost frontier which explicitly models the trade-off between costs, outputs, the quality of outputs and the public satisfaction with road provision as a means to provide a comparable 'efficiency' metric between 0 and 1 (see methodology section (4) below).

Pinkney and Marsden (2013) also discuss that there are several situations in transport planning where including public satisfaction alongside physical measures of quality and cost can enhance understanding. They particularly note that certain 'technocratic' quality measures do not correlate well with surveys of public satisfaction (such as 'pavement roughness' and 'public perception of quality' respectively discussed in Kummel et al (2001)). The result is that "[the current state of practice] is not sufficiently developed to connect transportation spend to customer oriented outcomes" (Pinkney and Marsden, 2013 p 3).

Applications in the roads sector aimed at quantifying cost inefficiency outside of partial measures, i.e. through controlling multiple outputs and quality (either physical or perceived) simultaneously, are less common. Braconier et al (2013) use OECD data from 32 countries to undertake Data Envelopment Analysis (DEA) as to the social effectiveness of the countries' strategic roads. Another example is Zang et al (2013) for US interstate highways. The DEA approach is a useful approach to quantify the scope for efficiency savings, but has a number of limitations (Coelli et al, 2005). Firstly, DEA does not estimate explicit relationships between costs and the level of outputs and quality of outputs. This is useful information that the econometric approach explicitly yields. For example, the econometric approach provides evidence on the extent to which road maintenance costs vary with traffic and provides associated measures of statistical confidence for these estimates. Secondly, DEA does not account for statistical noise when quantifying the scope of cost savings. This can result in an over prediction of potential cost savings. DEA does, however, have an advantage that there is less a priori information imposed on the shape of the cost relationship.

2.2. Econometric cost methods

No references to econometric (as opposed to DEA) cost efficiency analysis for local roads specifically can be found in the literature although it is acknowledged that there maybe some studies in the broader highway engineering literature or professional (potentially unpublished) literature not yet uncovered by the author. This review has uncovered that CEDR (2010) collect data on both costs and outputs, which could potentially be used for econometric frontier analysis or on strategic highway networks across European countries, but this was not undertaken in the CEDR study.

The studies by Agbelie et al (2015) and Hendren and Niemeier (2006) do use econometric methods to develop composite performance metrics which each account for more than one characteristic of organisations. For example, Hendren and Niemeier (2006) develop congestion and safety metrics based on two separate regression models; one for each metric. In the context of this paper, these studies are important to illustrate the benefit of econometric methods, however they do not estimate a cost function relationship which is the departure point for the work in this paper. In this paper, the problem of delivering road provision for minimum cost subject to providing a given length of road at a given quality (which could in theory include a measure of congestion and safety) is considered. This is different (but complementary) to the approach in Agbelie et al (2015) and Hendren and Niemeier (2006) who consider that maintenance cost influences a safety metric for example (Hendren and Niemeier 2006, p219).

Whilst there is a lack of econometric work identified specifically for measuring cost efficiency, a related strand of research at the strategic highway sector level (major roads) is an established econometric academic literature on understanding the properties of maintenance and renewal costs associated with highways. Contributions include Small et al (1989), Haraldsson (2007), Nilsson (2014), Link (2014), Jonsson and Haraldsson (2008), Ahmed et al. (2015a), Martin, T. (1994, 1997), Herry and Sedlacek. (2002). These have been motivated primarily in understanding how traffic usage damages roads relative to other damage mechanisms (such as weathering). Two broad statistical approaches are used as characterised in Ahmed et al (2015b), Bossche et al (2001) and Bruzelius. (2004).

Firstly, detailed data over time on multiple road sections are used to estimate the survival rate of assets over time (survival analysis). This approach can be traced back to road economics concepts outlined in Small et al (1989) and recent empirical applications include Haraldsson (2007) and Nilsson (2014)[1]. This approach yields a marginal cost of additional traffic by considering the change in present value of cost resulting from the marginal change in traffic as this is assumed to bring forward the time profile of asset renewal. The approach reconciles well with the fundamental engineering degradation process. However, it is data intensive both in terms of the disaggregation of road sections required and the historical data required. As such there is a degree of approximation required in terms of input data which limits the robustness of results. Ahmed et al. (2015a) develop this further to consider how different intervention approaches result in different marginal cost.

The second approach is to utilise the standard economic cost function with either maintenance cost or maintenance and renewal cost as the cost variable. Such analysis has been popular in Europe through several European Commission-funded projects such as UNITE, GRACE and CATRIN (Link (2002, 2009, 2014), Jonsson and Haraldsson (2008), Schreyer et al (2002), Bak et al (2006), Bak and Borkowski (2009)), in the US (Gibby et al (1990), Hajek et al. (1998), Li and Sinha (2000), Li et al. (2001, 2002) and in Australia (Martin (1994, 1997)).

In this approach, cost is explained by (regressed on) scale measures (e.g. road length or lane length) and traffic usage variables as well as variables to characterise the capability and condition of the assets (for example, age of pavement or number of defects). Input prices should also be considered as cost drivers (and such prices have been available in some studies such as Link (2014) for Germany). However for many studies, such input prices have not been available and so have been assumed uniform over cross sectional units and only to vary through time. As such, time dummy variables have been added to capture such variation (see Jonsson and Haraldsson (2008) for example). This is also common in the railway infrastructure literature (see Wheat and Smith (2008)).

[1] See Ahmed et al (2015a and 2015b) for a full set of study references from this approach.

The analysis in this paper fits within the second category above. It is unique in that it also contains an allowance for cost (economic) inefficiency within the modelling. Cost inefficiency measures the extent to which a LA fails to achieve cost minimisation conditional on the scale of output, network traffic and quality of the output.

3. Data

Table 3.1 describes the data. This study combines both publically available data and also primary sources not in the public domain direct from LAs. The cost data was collected through direct request to participating LAs and data was requested as the sum of structural and reactive carriageway maintenance. Footways were excluded, as was drainage, winter service and street lighting costs. Thus cost items in scope include interventions directly to the carriageway such as patching, inspections and road sealing. This cost data was sourced from a dedicated survey to participating LAs. This was necessary as other data (such as the Section 151 Officer statutory returns to Central Government – RO2 Returns (Department for Communities and Local Government, 2016)) do not provide appropriate cost breakdowns.

Data was requested for five years from 2008/09 to 2012/13 from each LA. Of the data collected, at least one year of data could be used from 51 LAs, however due to missing values of both explanatory variables and dependent variable, the panel was highly unbalanced, with between one and five observations per LA. In total there were 145 observations used in the analysis indicating an average number of observations per LA of 2.8.

Road length and traffic data were sourced from Department for Transport (DfT) statistics (DfT, 2014). The road length data was available by road type. The UK statistics distinguish between five road types: motorways (the vast majority of which are under the direct control of LAs so excluded from analysis), A roads, B roads, C roads and Unclassified (U) roads. The classification system broadly relates to strategic priority (with U roads being the lowest strategic priority) and design standards. Unfortunately, data is not publically available from a common source on lane-km, however the data by road type to some extent mitigates this.

A measure of physical road condition (RDC) was also sourced from central government statistics (DfT, 2014). RDC measures the frequency of defects in the road and corresponds to the proportion of the road network that requires repair. A higher value indicates more defects, therefore a higher value implies lower road condition.

The final piece of data available is from the National Highways & Transport Public Satisfaction Survey undertaken by Ipsos Mori (NHT, 2014). This is a postal survey sent to a representative sample of the populations in participating LAs. The survey covers aspects of public satisfaction with a range of private and public transport. Of use in this work was the measure “HMBI 01-satisfaction with condition of road surfaces”.

Table 3.1 Descriptive statistics of the data

Variable	**Units**	**Symbol**[1]	**Mean**	**Median**	**Maximum**	**Minimum**	**Std. Dev.**
Highway Expenditure	£ per annum	C	16720526	15514000	59000000	1352396	11948193
Length of A, B and C Roads	KM	X_1	1774.054	1628.1	6121.9	60.7	1473.349
Length of U Roads	KM	X_2	2319.087	2245.9	6714.3	183.8	1609.869
RDC	Proportion of network requiring repair	X_3	12.29045	11.89494	31.27044	2.497514	4.775362
Traffic density	Vehicle-km (million)/Road length KM	X_4	1.518813	1.599547	2.603487	0.509837	0.488003
Public Satisfaction Index	0-100 scale	X_5	32.91517	33.4	51.3	12.6	8.190029

Note: [1] Used in subsequent equations

4. Methodology

In this study a cost frontier is estimated using stochastic frontier methods. A cost frontier function relates minimum cost to a set of outputs (and characteristics) and input prices. If firms, or in this case LAs, are successful in cost minimisation, then the function is the dual to the production function (or in the multiple output case, transformation function) in the usual way (see for example Varian, 1999). However, if there is a degree of sub-optimisation (inefficiency), then the minimum cost relationship represents a boundary of the feasible cost set. Thus a frontier function models cost as including both a minimum cost relationship (explained by a set of regressors) and a component (usually a random variable) representing the degree to which an LA is producing above the implied minimum cost. From the perspective of an econometric model, an additional allowance for unobservable 'noise' is required (in recognition that any model will be an abstraction from reality). This yields the stochastic frontier model first proposed simultaneously by Aigner et al (1977) and Meeusen and van den Broeck (1977). For a full introductory treatment of stochastic frontiers see Kumbhakar and Lovell (2000).

Use of stochastic frontier (SF) methods is not the only candidate method for efficiency analysis. There are mathematical programming methods, such as Data Envelopment Analysis (DEA) and Freely Disposable Hull (FDH) approaches, as well as 'deterministic frontier methods' such as Corrected Ordinary Least Squares (COLS). The papers of Zang et al (2013) and Braconier et al (2013), mentioned in the literature review, are examples of the DEA approach in road maintenance. SF methods are chosen over DEA and FDH since the parametric formulation is of particular interest in this case (DEA and FDH are non-parametric). In particular, the scale and density properties (how cost changes with network length and traffic respectively) are of interest here and also quantifying the impact of asset condition and public satisfaction on cost is important as this recognises the cost/quality trade-off. This leaves the choice between SF methods and 'deterministic' methods. Deterministic methods attribute all variation away from the minimum cost frontier to inefficiency. However, this is unlikely to be appropriate in the case of this dataset. Whilst the cost data is a step forward from the published data in this field, there is likely to be a degree of inconsistency (measurement error) between costs reported by individual LAs due to different accounting systems being used. Indeed the statistical results support an approximate

60/40 split between 'noise' and inefficiency. In any case, it goes against general statistical /economic modelling theory to believe that an economic model fully represents the minimum cost generating process (particularly given that there are 'only'[2] five explanatory variables).

The model can be represented as

$$lnC_{it} = \beta_0 + \beta_1 lnX_{1it} + \beta_2 lnX_{2it} + \beta_{22}(lnX_{2it})^2 + \beta_3 X_{3it} + \beta_4 lnX_{4it} + \beta_5 X_{5it} + \beta_{55} {X_{5it}}^2 + \beta_{555} {X_{5it}}^3 + \gamma_2 + \ldots \gamma_T + \varepsilon_{it} \quad (1)$$

Where i identifies each LA i=1,…,51 and t the time period t=1,…,5 with 1 corresponding to the financial year 2008/09. $C, X_1, \ldots, X_5$ are as defined in the data section (see Table 3.1 for correspondence). γ_t are fixed effects in the time dimension i.e. time dummies with the first year excluded for co-linearity reasons (dummy variable trap). $\varepsilon_{it} = v_{it} + u_{it}$, where $v_{it} \sim N(0, \sigma_v^2)$ and $u_{it} \sim |N(0, \sigma_u^2)|$. v_{it} is a random variable which captures random noise. u_{it} is also a random variable which captures positive deviations from the stochastic cost frontier i.e. economic inefficiency. The critical difference between v_{it} and u_{it} is that v_{it} is distributed symmetrically around zero, whilst u_{it} is defined only for non-negative values. Note that to aid estimation in the LIMDEP v10 software used for estimation (Econometric Software, 2010), the model is re-parametrized such that $\sigma = \sqrt{\sigma_v^2 + \sigma_u^2}$ and $\lambda = {}^{\sigma_u}\!/_{\sigma_u}$.

The functional form in (1) requires explanation. All variables except the two percentage variables, namely RDC and public satisfaction, are entered in logarithms rather than in levels. This implies that the coefficients (or combinations thereof[3]) can be interpreted as cost elasticities. So, as an example, the elasticity of cost with respect to traffic (X_4) is given by the estimate of parameter β_4.

For the variables RDC and public satisfaction, the reason these were entered without logarithmic transformation is twofold. Firstly, these variables are percentage measures. Including them in levels

---

[2] The use of '' refers to the observation that many published efficiency studies and studies that have been used for regulatory determinations have less or an equal number of explanatory variables e.g. see Andersson et al (2012), Smith et al (2010) and Smith and Wheat (2012).

[3] The expressions for the cost elasticities are given through partial derivation in the case of where multiple terms of the same variable are included

that yield the partial derivatives of cost with respect to the variables, represents the percentage increase in cost from a 1 unit increase (in these cases a 1% unit increase) in the percentage variables. So for the RDC variable (X_3), the coefficient β_3 represents the percentage change in cost from a 1% unit increase in road defects. This was felt to be an appealing a priori interpretation. Secondly, this formulation fitted the data better relative to a log formulation (as judged by R^2).

The model contains only two second order terms and also a third order (cubic) term in public satisfaction. Originally, a full set of interaction terms between the road length measures and traffic was considered. A mixture of general-to-specific testing down and inspection of the cost elasticities to check they were plausible (in particular checking that all or the majority of computed elasticities were positive as required by economic theory) yielded this parsimonious specification[4].

The third order relationship between cost and public satisfaction was informed through empirical investigation and discussions with stakeholders. In particular, it was thought that there would be a limited relationship between public satisfaction and cost at average levels of public satisfaction. As part of this research, an industry expert committee was formed (comprising of representatives of LA highways departments) which hypothesised that this was due to other factors, such as expenditure on communications strategy being more important for accounting for variation in levels of public satisfaction rather than having any correlation with cost. However, at the extremes, it was hypothesised that variation in public satisfaction would be an influence on cost as low or high public satisfaction could only be achieved through exceptionally poor or excellent service delivery rather than very good or poor 'softer factors'. Thus cost was needed to influence public satisfaction at both extremes. This was confirmed through initial statistical estimations which found that the simpler first order (single

[4] In more detail, the interactions between the traffic variable and road length variables (interactions between X_4 and X_1,X_2) were jointly insignificant (pval 0.1533) and so removed. The same result emerged (pval 0.1842) for the excluded second order terms in X_1 and X_2 themselves. This led to the model as in (1) with the exception of a second order term (square term) in traffic (X_4). The second order term on traffic yielded an elasticity of cost with respect to traffic which was negative for lightly trafficked network (empirically 15% of observations were found to have a negative traffic elasticity). This was deemed implausible and so the second order term was removed (in any case this term was not significant at the 5% level but was significant at the 10% level).

term, not raised to any power) relationship or second order relationship (two terms, one not raised to a power and the other raised to the power 2) could not detect a statistically significant relationship.

Empirically, two functions were considered. The first was a piecewise linear growth rate function i.e. public satisfaction terms interacted with threshold (for public satisfaction) dummy variables. Thus this approach would allow for different growth rates for various (pre-specified) levels of public satisfaction. This approach was found to have an inferior fit relative to the alternative. The alternative (and that which was adopted) was to use a cubic function for growth rate terms. Critically, this allows for a turning point in the relationship in which the hypothesised relationship is nested (i.e. a strong cost relationship for low values, followed by a weak relationship at average values and a strong relationship at large values of public satisfaction). For further details of this function, the reader is directed to the discussion of the empirical implementation and interpretation of this function in section 5.3. This illustrates how the proposed function captures the hypothesised relationship if (as was empirically found) the coefficients on the first order, second order and third order terms are positive, negative and positive respectively.

Finally on methodology, a specific (cross sectional) panel data treatment has not been included into the modelling, such as by including cross sectional (LA) fixed effect dummy variables. This was for two reasons. Firstly, the unbalanced nature of the panel includes several LAs with only one year of data. This means that the fixed effects would dominate explanatory power for several LAs. For those LAs with only one year of data, the fixed effects would completely explain the variation in the data for those LAs. This is not an appropriate means to model efficiency for those LAs given by construction they would be no unexplained gap (the fixed effect would absorb the gap). Secondly, a Hausman test was conducted and this failed to compute due to the variance matrix being non-positive definite. Using established precedent in the econometrics literature e.g. Greene (2012) this result is taken to imply that there is no statistically significant difference between a random and fixed effects treatment. Given that OLS or stochastic frontier maximum likelihood estimation (Aigner et al, 1977) provide consistent parameter estimates under the assumptions of random effects, fixed effects are not included in the stochastic frontier modelling. However, robust standard errors are used to compensate for incorrect computation of standard errors in the standard formulation in the presence of correlation of errors within groups.

Whilst cross sectional fixed effects are not included, fixed effects in time (the γ_t parameters) are included. These proxy for two key omitted variables. Firstly, as is common in the literature on road maintenance (and indeed railway maintenance – see section 2), the time dummy variables proxy for variation in input prices over time. Implicitly this assumes no systematic variation in input prices across the cross section or, at least, any variation will be captured by the inefficiency term. This is a limitation of the analysis given the data available, but not without empirical precedent in the literature. The second set of factors which the time dummy variables capture are winter weather effects such as freeze/thaw. Again, such cross sectional invariant effects are imperfect proxies, however they do cover more extreme and less extreme winters in general. Indeed the large estimated coefficient for γ_3 would seem to capture the harsh winter of 2010/11 (see section 5 and Table 5.1).

5. Results

This section discusses the results under four sub-sections. The first three sub-sections (5.1-5.3) consider the properties of the cost frontier characterising how minimum costs relate to scale, traffic and quality respectively. Sub-section 5.4 then considers the extent that each LA's actual cost departs from the modelled LA specific minimum cost i.e. cost efficiency.

The estimates of the model parameters are given in Table 5.1. Two models are presented, a Preferred Model which includes Public Satisfaction and a Comparator model which does not. Both are presented since public satisfaction is rarely included directly in cost benchmarking models (a notable exception is benchmarking work by Monitor (2016), for the NHS Health Service in Britain). Importantly, Table 5.1 shows that the parameter estimates are similar with and without the inclusion of public satisfaction in the modelling. As such, in the sub-sections 5.1-5.3 on scale, traffic and quality, the empirical findings discussion covers the preferred model only (the comparator model has similar empirical findings).

Table 5.1 Cost frontier parameter estimates

	Preferred Model (equation (1))			Comparator Model (equation (1) less the public satisfaction variables)		
	Coefficient	**Std. Error**	**Prob.**	**Coefficient**	**Std. Error**	**Prob.**
β_0	12.6064	1.484132	0.0000	16.04614	0.136191	0.0000
β_1	0.625307	0.135513	0.0000	0.607676	0.132827	0.0000
β_2	0.170655	0.180644	0.3465	0.220936	0.175152	0.2093
β_{22}	0.205119	0.066772	0.0026	0.208603	0.060277	0.0007
β_3	0.022424	0.008156	0.0068	0.022020	0.008236	0.0084
β_4	0.241204	0.131967	0.0698	0.272175	0.130209	0.0385
β_5	0.341242	0.147382	0.0221			
β_{55}	-0.010498	0.004685	0.0267			
β_{555}	0.000102	4.79E-05	0.0353			
γ_2	-0.068159	0.141489	0.6308	-0.046126	0.138258	0.7392
γ_3	0.013900	0.119311	0.9074	0.051679	0.117124	0.6597
γ_4	0.070376	0.130317	0.5901	0.102136	0.124431	0.4132
γ_5	0.010912	0.126915	0.9316	0.052863	0.122837	0.6676
λ	1.29062	0.26174	0.0000	1.48891	0.291771	0.0000
σ	0. 54269	0.00313	0.0000	0.57508	0.00336	0.0000
Log-Likelihood	-79.76414			-82.66604		

Number of observations =145 , Number of LAs = 51; Dependent Variable=ln(Highway Expenditure)
γ_1 is excluded to avoid the dummy variable trap – implicitly its effect is captured in the constant term.

5.1. Returns to scale

Returns to scale are important to understand and quantify when trying to examine the possibility for cost saving. Firstly, a given LA should not be penalised (or given undue favour) in an efficiency assessment if its size is fixed but it happens to be at a scale which minimises unit costs. Secondly, a current theme in the response of LAs to the cost saving challenge is to examine the possibility to pool service delivery across adjacent LAs (HMEP, 2013). Thus scale can be thought of as quasi fixed.

The model indicates that, at the sample mean of the data, increasing the size of a LA's road network by 1% increases costs by 0.80%. This is statistically significantly different from unity (p val=0.013) indicating that there is indeed increasing returns to scale at the sample mean. Again, this is intuitive given that some costs of road maintenance are fixed irrespective of what work is undertaken. However, for very large LAs a 1% increase in road length increases costs by more than 1%, indicating that at a certain road lengths there are coordination problems. Thus there is an optimal size for a LA's highway network.

Figure 5.1 shows the plot of the cost elasticity with respect to highway length (the scale elasticity) for the (145) observations within the dataset. It clearly shows an upward relationship, with the 'minimum efficient scale point' (MESP) somewhere in the order of 4000 to 8000 km. The exact MESP depends on the mix of U road length and A, B, C road length in each specific LA. The MESP represents the level of scale, measured here by road length, where cost efficient average costs (cost per road-km) are minimum. So it is important to note that an elasticity value less than one does not imply a fall in absolute cost from growing the LA size, only that average cost falls as the LA gets larger (up to the MESP and then average costs start to rise).

Turning to the possible benefits of combining services across LAs, Table 5.2 considers the implied average cost savings (per road km). It considers combining two identical sized LAs into one LA. It shows that there are substantial cost savings for LAs up to the mean (4000km) from merging. For small LAs of half the size of the mean LA, the potential cost saving is 21%. However LAs can be too big (from an average cost optimisation perspective) and this can be seen by the negative entries in the table for LAs above the mean size.

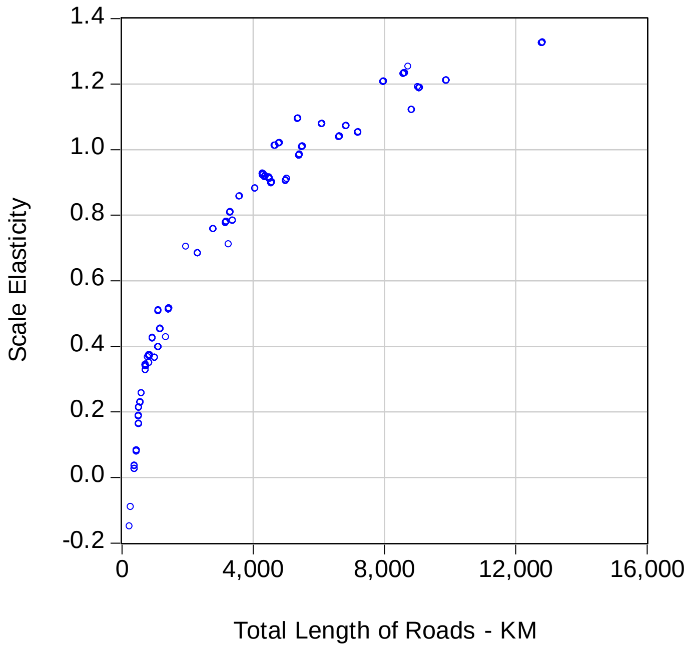
It is doubtful however that all of the cost savings identified in Table 5.2 could be realised in practice, given that it could be the case that LAs would be sharing services (or elements of – such as joint procurement) rather than fully merging to form a new LA. However it does demonstrate that for many LAs, a policy of sharing services across adjacent LAs where feasible could yield substantial savings.

Table 5.2 Potential cost savings from merging LAs of a given size

Size of merging LAs (relative to mean LA)	**Average cost per road km relative to average cost at the sample mean**		**Potential cost saving from merger**
	Pre-combined LA	**Combined LA**	
0.5	1.27	1.00	21%
0.75	1.08	0.95	12%
1	1.00	0.96	4%
1.25	0.97	0.99	-2%
1.5	0.95	1.02	-8%
2	0.96	1.12	-17%

Notes: 1) In each example two LAs of identical size (relative to the sample mean of the data) are considered to merge. So the first line considers two LAs each 0.5 times the size of the mean LA, combining (merging) into a single LA of size 1 relative to the mean LA in the dataset.
2) All other characteristics of the LAs are assumed to be the same i.e. traffic, RDC, public satisfaction.
3) The mix of U roads and A,B,C Roads is assumed to be the same as at the sample mean for both LAs

Figure 5.1 The scale elasticity for the observations used in the highways modelling[5]



[5] It should be noted that there are 2 observations representing the smallest authorities that the model estimates to have a negative scale elasticity. Taken literally this would imply that increasing network length would actually reduce total cost. This is clearly counter intuitive, however such results at the extremes of the sample data are not uncommon in these modelling exercises as such extremes are estimated with a high degree of imprecision.

5.2. The impact of traffic

The model shows that traffic does have a positive relationship with maintenance costs and this is statistically significant at the 7% level. The coefficient, which is a cost elasticity, is 0.24 and this indicates that a 1% increase in traffic results in a 0.24% increase in maintenance cost.

Table 5.3 shows how this estimate compares to that from other studies in the econometric literature. This estimate is below the range from other studies (in Table 5.3 the range is approximately 0.3-0.85 and median 0.425), but this probably reflects the composition of the roads in this study relative to the roads in the comparator studies. In particular the studies in Table 5.3 generally include strategic roads which tend to carry greater traffic per road km and also a greater proportion of HGV traffic which does more damage. The outlier, reported in Link (2009) of 0.85 was for motorways. Thus it is no surprise that the estimate from this study is lower.

The estimated elasticity could be useful for strategic planning purposes in terms of understanding how local highway costs are likely to change in response to long term changes in traffic. Going forward, obtaining traffic data disaggregated by HGV and other vehicles would enhance the cost model and particularly help model cost for those LAs with a large number of A roads which tend to carry the most HGV traffic.

Table 5.3 Cost elasticity with respect to traffic from other studies

Study	**Country**	**Type of Road Network**	**Cost elasticity with respect to traffic**
Schreyer et al (2002)	Switzerland	Motorways and Kanton roads	0.45
Jonsson and Haraldsson (2008)	Sweden	All paved roads	0.39
Bak et al (2006)	Poland	National roads	0.48
Bak and Borkowski (2009)	Poland	National roads	0.3
Link (2009)	Germany	Motorways	0.85
Link (2014)	Germany	Federal roads	0.4
Martin (1994,1997)	Australia	National Roads	0.46 (HGV only – Pavement maintenance)

Source: Amended from Link (2014, Table 7). This table does not include results from all studies identified in the literature review. This is primarily because many studies did not report the cost elasticities and simply focused on marginal cost. Some studies were also excluded e.g. Link (2002) if subsequent studies updated the analysis.

5.3. The impact of quality of service

Increasing RDC i.e. the average number of road defects, increases cost, reflecting the need to do more maintenance to counter the large number of instances of high risk defects in the road network.

As discussed in section 4, there are three terms in the model capturing the influence of public satisfaction on costs. Each of the parameters on the three terms is statistically significant at least at the 5% significance level. The estimated cubic relationship has the coefficient on the first, second and third order terms of positive, negative and positive respectively with the first order term being larger than the second and second larger than the third in absolute value.

This relationship does indeed empirically conform to a priori expectations of the industry expert committee. In particular:

- For low levels of public satisfaction there is a strong positive relationship between cost and public satisfaction, since the positive coefficient on first order term dominates in the cubic function at low satisfaction levels.
- For average levels of public satisfaction, the quadratic term dominates (and this has a negative coefficient), which implies that the relationship is less strong for average public satisfaction levels.
- For large levels of public satisfaction, there is a strong positive relationship between cost and public satisfaction, now because the cubic term dominates the function.

This is evidence that maintenance cost does not vary one to one with public satisfaction. Indeed, the relationship shows that for levels of public satisfaction in proximity to the average, there is little relationship between the two, since in practice when public satisfaction is not very bad or very good, it tends to vary not because of the physical condition of the network, but because of other factors such as variation in public communications strategies across LAs (so called ‘soft factors’). However, at the extremes, there is evidence for a positive relationship i.e. to increase public satisfaction requires more maintenance cost. One possible explanation as to why this is found is that very low levels of public satisfaction are often associated with very poor networks i.e. there has to be something over and above communications strategy that yields very poor satisfaction. This thus requires cost to be incurred to improve satisfaction. Similarly, it is unlikely that even an excellent communications strategy or other soft factors can yield very high public satisfaction without sustained maintenance of the highway network, hence the relationship between cost and public satisfaction at the upper end.

5.4. Cost Efficiency Predictions

Once the cost frontier has been estimated it can be determined how far each of the 51 LAs is from the frontier and thus what scope there is for each LA to potentially make cost savings (subject of course to such an opportunity representing something under control of the LA). Table 5.4 gives descriptive statistics for the distribution of efficiency predictions from the preferred model. Efficiencies are predicted using the conditional expectation formulas in Jondrow et al (1982) which is common practice in stochastic frontier analysis. Importantly, the measure of cost efficiency is the scope for savings after cost effect of variation in road length, traffic, road condition and public satisfaction have been controlled. As such the potential savings captured in the cost efficiency measure are over and above any saving resulting from changes in the explanatory variables such as those arising from the merging of adjacent LAs to form a new LA closer to the minimum efficient scale point (as discussed in section 5.1).

On average (the median efficiency over all observations), LAs are found to be 83% efficient. Literally speaking this means on average LAs can in theory reduce highway maintenance expenditure by 17% (=100% - 83%) and continue to maintain the same network, to the same quality and with the same traffic usage.

Examining the distribution further, it can be seen that over 75% of LAs have efficiency predictions above 75%, which seems intuitive. Only 10% of LAs have efficiency predictions less than 64% and, inevitably with any benchmarking analysis, it is likely that these 10% are probably the LAs which are subject to data issues. Five LAs have efficiency scores above 95% which indicates that the frontier is not defined by a single outlying observation, but a collection of LAs. Thus in general the spread of efficiency predictions seems intuitive.

Of course, what really matters is whether the ranking of LAs makes intuitive sense and such a review was undertaken by an industry expert committee group, comprising representatives from participating LAs. This was an important part of the study given the individual efficiency scores cannot be released for confidentiality reasons, but were available to the industry expert committee. The outcome of the review was that the preferred model was thought more plausible than the comparator model given that there were several LAs with poor public satisfaction scores but very high efficiency scores in the comparator model. In contrast, the preferred model presented some efficiency penalty for these LAs without a substantial change in the distribution of other efficiency scores (the mean in the comparator model was 82% which is not too different to 80% in the preferred model for example).

Table 5.4 Distribution of efficiency predictions for the 51 LAs[6]

Percentile of efficiency distribution	**Efficiency Score**
0%	48%
10%	64%
20%	70%
25%	75%
30%	76%
40%	79%
50% (median)	83%
60%	86%
70%	88%
75%	89%
80%	90%
90%	94%
100%	100%
Mean	80%

6. Conclusions

In this paper a cost stochastic frontier model for English LAs has been developed using a dataset collected from both secondary sources in the public domain and primary sources not in the public

[6]The efficiency scores are predicted for each year and then averaged over the years that each Authority appears as a complete observation in the data, before the distribution is ranked to form the entries in Table 5.3.

domain, using a specific definition of highway maintenance cost. This ensures a greater degree of consistency in cost definition compared to other available cost measures. The paper is the first application of stochastic frontier methods to benchmarking for English LAs.

The model allows conclusions on four key aspects of the cost structure and efficiency of LAs:

1) In terms of the size of LAs (defined by length of road network), there are strong economies of scale for smaller LAs but these diminish to unity as the size of the LA network approaches 4000-8000km (the exact figure depending on the mix road classes in an LA). LAs with network size greater than this suffer from diseconomies of scale. Given the distribution of LA size in the sample, LAs up to and including the sample mean length could reduce unit costs by merging with an LA of similar size. In practice, administrative, political and legacy factors are likely to prevent full merging of LAs and so the cost savings indicated in Table 5.2 are likely to be over estimates. Nonetheless, for small LAs, the potential savings are substantial.

2) It is found that a 1% increase in traffic results in a 0.24% increase in maintenance cost and this agrees with the literature that road maintenance is subject to economies of density. This result is useful for strategic planning in that it allows the cost impact of future traffic growth to be predicted.

3) For the most part, there is little association between public satisfaction with highway condition and the cost of provision of maintenance. The exception is for very low and very high levels of public satisfaction where costs are increasing functions of public satisfaction. Accounting for this relationship is particularly important in terms of the plausibility of the rankings of the efficiency scores. This study highlights the potential benefits for cost analysis and cost benchmarking by including measures of user satisfaction within the analysis.

4) On average (median), LAs are found to be 83% efficient. In theory, this means that LAs can reduce highway maintenance expenditure (100%-83%) or 17% and continue to maintain the same network, to the same quality and with the same traffic usage. This represents a substantial opportunity for the average LA to make savings without trading off the outcomes of activities. Of course, some of the gap

may arise from the influence of unobserved factors outside of a LA's control but this does represent a useful starting point for further process benchmarking analysis.

The implications of the above findings internationally are threefold. Firstly, the study reveals substantial returns to scale for LAs. For all but the largest LAs sharing services across LAs or even formal mergers would seem beneficial from a cost perspective. Secondly, new evidence on the cost variability of traffic for local roads has been established. This can be useful for strategic forecasting of maintenance costs to trends in traffic growth. Thirdly, and more generally, this work has shown that collaborative benchmarking exercises involving LAs which utilise stochastic frontier analysis are feasible and produce useful results. Importantly, this work highlights why costs differ between LAs and, once the cost driving factors have been controlled, what is the remaining cost gap. Such a gap may be able to be eliminated in the future through adoption of best practice.

Further research is required to better understand the influence of public satisfaction on costs and to verify the relationships found in this research. Incorporating public satisfaction (perception of final outputs by users) into benchmarking is relatively new but is gaining traction in such sectors like health care, where Monitor (the economic regulator of the National Health Service in Britain) is including quality of care based on patient satisfaction in their econometric benchmarking (Monitor, 2016). It is suggested that this is an important issue for future regulation in general. Furthermore, there is a need for iterative improvements in both the econometric model and understanding of what constituents the efficiency gap. The former requires a refinement in the definitions and reporting of costs to reduce 'noise' within the modelling as well as identifying and collecting data on more cost drivers. The latter requires case study and other bottom-up investigations with specific LAs. Both activities are being actively taken forward by the CQC Efficiency Network (2016).

References

Aigner, D.J., Lovell, C.A.K, and Schmidt, P. (1977), 'Formulation and Estimation of Stochastic Frontier Production Function Models', Journal of Econometrics, 6 (1), 21-37.

Ahmed, A., Qiang, B.., Lavrenz, S. and Labi, S. (2015a). Estimating the marginal cost of pavement damage by highway users on the basis of practical schedules for pavement maintenance, rehabilitation and reconstruction. Structure and Infrastructure Engineering 11(8), 1069–1082.

Agbelie, B., Labi, S., Noureldin, S., and Sinha, K. (2015). "Linking Interstate Pavement Preservation Investment to Performance: An Exploratory Aggregate Analysis." ASCE Journal of Infrastructure System, 21 (4),

Ahmed, A., Qiang, B., & Labi, S. (2015b). Pavement damage cost estimation – A synthesis of past research. Proceedings of the ICE Transport, 168 (1), 48-58.

Andersson, M., Smith, A.S.J., Wikberg, A., and Wheat, P. (2012). 'Estimating the marginal cost of railway track renewals using corner solution models', Transportation Research Part A, 46 (6), 954–964.

Bossche, V.D.M., Certan, C., Goyal, P., Gommers, M., & Sansom, T. (2001). Marginal cost methodology. Deliverable 3 of UNITE, v. 2.1. Leeds: University of Leeds.

Bruzelius, N. (2004). Measuring the marginal cost of road-use – An international survey. Meddelande 963A. Linkoping: Swedish National Road and Transport Research Institute.

Braconier, H., M. Pisu and D. Bloch (2013), "The Performance of Road Transport Infrastructure and its Links to Policies", OECD Economics Department Working Papers, No. 1016, OECD Publishing. http://dx.doi.org/10.1787/5k4ddxjvpfr5-en

Conference of European Directors of Roads (CEDR) (2010), "Report of BEXPRAC".

Department for Transport (2015). 'Highways maintenance funding: incentive element'. Available at: https://www.gov.uk/government/publications/highways-maintenance-funding-incentive-element.

Coelli, T. J., Rao, D. S. P., O'Donnel, C. J. and Battese, G. E (2005). An Introduction to Efficiency and Productivity Analysis. Second Edition, Springer.

CQC Efficiency Network (2016). “Measuring efficiency and sharing efficient practices to improve”. Website: http://www.nhtnetwork.org/cqc-efficiency-network/home/.

Department for Transport (2014). ‘Road Network size and condition’. Website: https://www.gov.uk/government/collections/road-network-size-and-condition.

Department for Transport (2012). ‘Department for Transport (DfT) Benchmarking Tool’ available at: http://www.dft.gov.uk/publications/local-authority-benchmarking-tool/.

Department for Communities and Local Government (2016). ‘Local authority revenue expenditure and financing’. Website: https://www.gov.uk/government/collections/local-authority-revenue-expenditure-and-financing.

Econometric Software Inc. (2010). LIMDEP, User’s Manual, http://www.limdep.com, Plainview, NY.

Gibby R, Kitamura R and Zhao H (1990) “Evaluation of truck impacts on pavement maintenance cost.” Transportation Research Record 1262: 48–56

Goode, L., Jeff, G., MacGillivray, I., and Pedersen, N. (1993). “Response to ‘Resources vs. results: Comparative Performance of State Highway Systems, 1984–1990.” Policy Studies Review. 21(2), 375–378.

Hajek J, Tighe SL and Hutchinson BG (1998) “Allocation of pavement deterioration due to trucks using a marginal cost method.” Transportation Research Record 1613: 50–56.

Haraldsson, M. (2006) ‘Marginal Cost for Road Maintenance and Operation — A Cost Function Approach’, Annex to Deliverable D3 GRACE, Marginal cost case studies for road and rail transport. Funded by the Sixth Framework Programme. ITS, University of Leeds, Leeds.

Haraldsson, M. (2007) ‘Essays on transport economics’, Economic Studies 104, Department of Economics, Uppsala University.

Hartgen, D.T., Chadwick, C.G., and Fields, M.G. (2008). Transportation Performance of the Canadian Provinces, Fraser Institute. http://www.fraserinstitute.org/research-news/display.aspx?id=13549 (Accessed October 28, 2016).

Hartgen, D., and Krauss, K. (1993). "Resources vs. Results: Comparative Performance of State Highway Systems, 1984–1990." Policy Studies Review. 21(2), 357–374.

Hendren, P., and Niemeier, D. A. (2008). "Identifying Peer States for Transportation System Evaluation & Policy Analyses" . Transportation, 35(4), 445–465.

Hendren, P., and Niemeier, D. A. (2006). "Evaluating the Effectiveness of State Department of Transportation Investment Decisions: Linking Performance Measures to Resource Allocation." Journal of Infrastructure Systems (12)4, 216–229.

Herry, M., & Sedlacek, N. (2002). Road econometrics: Case study motorways Austria. Annex A-l(b) of Deliverable 10 of UNITE, v. 2.2. Leeds: University of Leeds.

Highways Maintenance Efficiency programme (HMEP) (2013). 'HMEP Annual plan 2014/15'. Available at: http://www.highwaysefficiency.org.uk/about-us/hmeps-vision-and-annual-plan.html.

Jonsson, L. and M. Haraldsson (2008). Marginal Costs of Road Maintenance in Sweden, Annex 1 to CATRIN (Cost Allocation of Transport Infrastructure cost), Deliverable D6 Road Cost Allocation for Europe: Recommendations and open questions. Funded by the Sixth Framework Programme. VTI, Stockholm.

Kuemmel, D.A., Robinson, R.K., Sonntag, R.C., Griffin, R.J. and Giese, J.K. (2001) 'Public Perceptions of the Midwest's Pavements: Explaining the Relationship between Pavement Quality and Driver Satisfaction', Journal Transportation Research Record: Journal of the Transportation Research Board, 1769, 61-70.

Li Z and Sinha KC (2000) A Methodology to Determine the Load and Non-Load Shares of Pavement Repair Expenditure. Joint Transportation Research Program, Purdue University, West Lafayette, IN, USA, Report IN/JTRP-2000–09.

Li Z, Sinha KC and McCarthy PS (2001) “Methodology to determine the load and non-load shares of highway pavement rehabilitation expenditures”. Transportation Research Record. 1747: 79–88.

Li Z, Sinha KC and McCarthy PS (2002) “A determination of load and non-load shares of highway pavement routine maintenance expenditures”. Road and Transport Research 11(2): 3–13.

Link, H. (2014). ‘A Cost Function Approach for Measuring the Marginal Cost of Road Maintenance’, Journal of Transport Economics and Policy, 48(1), 15-33.

Link, H. (2006): ‘An econometric analysis of motorway renewal costs in Germany’, Transportation Research Part A, 40, 19–34.

Link, H. (2009): Marginal Costs of Road Maintenance in Germany, Annex 3 to Deliverable D6, Road Cost Allocation for Europe: Recommendations and open questions. CATRIN (Cost Allocation of Transport Infrastructure). EU-Project funded by the Sixth Framework Programme. VTI, Stockholm.

Litman, T. (2009). A Good Example of Bad Transportation Performance Evaluation: A Critique of the Fraser Institute Report. Transportation Performance of the Canadian Provinces. Victoria Transport Policy Institute, Victoria, BC, Canada

Martin, T. (1994). Estimating Australian’s attributable road track cost. Australian Road Research 254 Victoria, Australia: Australian Road Research Board Ltd.

Martin, T. (1997). ‘Australian road track cost recovery: some recent findings.’ Transport Reviews, 17:3, 223-243.

Meeusen, W. and van Den Broeck, J. (1977). ‘Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error’, International Economic Review, 18 (2), 435-444.

Monitor (2016), '2016/17 National Tariff Payment System: A consultation notice', February, p. 34, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/499594/2016-17_national_tariff_statutory_consultation.pdf.

NHT (2014). 'NHT Public Satisfaction Survey'. Website: http://www.nhtnetwork.org/nht-public-satisfaction-survey/home/.

Nilsson, J.E. (2014). 'The marginal costs of road wear,' Paper presented to the hEART conference. 2014 University of Leeds, UK.

Pinkney, S. and Marsden, G. (2013). 'Measuring and Benchmarking User Satisfaction with Transportation' TRB 92nd Annual Meeting, Washington, United States.

Small, K. A., Winston, C. and Evans, C. A. (1989), Road work: A new highway pricing and investment policy, The Bookings Institution, Washington.

Smith A and Wheat P (2012) Estimation of cost inefficiency in panel data models with firm specific and sub-company specific effects, Journal of Productivity Analysis, 37, pp.27-40.

Smith A; Wheat P; Smith G (2010) The role of international benchmarking in developing rail infrastructure efficiency estimates, Utilities Policy, 18, 86-93.

Varian, H. (1999). Intermediate Microeconomics: A Modern Approach. 5th Ed. W. W. Norton & Co.; US.

Wheat, P. and A. Smith (2008) 'Assessing the Marginal Infrastructure Maintenance Wear and Tear Costs for Britain's Railway Network', Journal of Transport Economics and Policy, 42(2), 189-224.

Zhang, Z., Arman, M., Labi, and S. (2013). Efficiency Measurement of Interstate Highway Pavement Maintenance Using Data Envelopment Analysis, Procs., Annual Meeting of the Transp. Research Board, Washington, DC.

Acknowledgements

The author acknowledges funding from the Highways Maintenance Efficiency Programme (HMEP) to support this work.