Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

Emerald Insight

Benchmarking: An International Journal

Performance measures for road managers facing diverse environments

Craig M. Richmond Clemens Kielhauser Bryan T. Adey

Article information:

To cite this document:
Craig M. Richmond Clemens Kielhauser Bryan T. Adey , (2016),"Performance measures for road managers facing diverse environments", Benchmarking: An International Journal, Vol. 23 Iss 7 pp. 1876 - 1891
Permanent link to this document:
http://dx.doi.org/10.1108/BIJ-01-2015-0005

Downloaded on: 14 November 2016, At: 00:36 (PT)
References: this document contains references to 14 other documents.
To copy this document: permissions@emeraldinsight.com
The fulltext of this document has been downloaded 46 times since 2016*

Users who downloaded this article also downloaded:

(2016),"Reliability and validity issues in analysis of IT spending using IT managerial control ratios", Benchmarking: An International Journal, Vol. 23 Iss 7 pp. 1910-1921 http://dx.doi.org/10.1108/BIJ-01-2015-0008

(2016),"Generalized DEA: an approach for supporting input/output factor determination in DEA", Benchmarking: An International Journal, Vol. 23 Iss 7 pp. 1892-1909 http://dx.doi.org/10.1108/BIJ-07-2015-0074

Access to this document was granted through an Emerald subscription provided by emerald-srm:563821 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

The current issue and full text archive of this journal is available on Emerald Insight at:
www.emeraldinsight.com/1463-5771.htm

Received 21 January 2015
Revised 28 October 2015
Accepted 11 November 2015

Performance measures for road managers facing diverse environments

Craig M. Richmond, Clemens Kielhauser and Bryan T. Adey
Department of Civil, Environmental and Geomatic Engineering, Swiss Federal Institute of Technology Zurich, Zurich, Switzerland

Abstract

Purpose – A key difficulty that plagues benchmarking in the public sector is heterogeneity in the production process. The purpose of this paper is to present a strategy for overcoming that difficulty using physical production models and demonstrate it using road renewal management as an example.
Design/methodology/approach – A physical production model is used to linking required prices, inputs and exposures to environmental factors to the desired services to be delivered. A measure is derived from this that adjusts for the additional expected costs from operating in a more difficult environments. A case study is used to present methods for addressing specific parameterization issues that arise in an empirical application.
Findings – The method was found to be implementable and empirically better than naïve ratio measures commonly found in practice.
Research limitations/implications – Data and modeling issues were identified that can be addressed by public supervisors that are expected to greatly improve the quality of the measures.
Social implications – According to the raw data and simple ratios, a very large degree of inefficiency can potentially be eliminated by applying the recommended measures. In all likelihood the real potential is much smaller, but still significant.
Originality/value – Most applied benchmarking exercises use simple ratios as KPI's. These are easily dismissed where environments are heterogeneous. Data envelopment analysis and stochastic frontier analysis are generally difficult to relate to KPI's. The use of an explicit and specific process model with an engineering content is therefore exceptional.

Keywords Performance measurement, Benchmarking, Highway maintenance, KPI, Renewal
Paper type Conceptual paper

Introduction

The proverb goes that "you can't compare apples and oranges" and get a useful result, but the proverb tells only half of the story. In reality, almost every comparison – and certainly those between different public administrations – has to cope with a lack of perfect comparability. Many examples of public benchmarking have an inward focus and thereby avoid the whole question of differential conditions: comparisons are made either against goals one sets for oneself or against one's own history. An advantage of this is that one is quite free in the definition of measures. However, as soon as measures are used to show the relative performance across decision-making units (DMU), the freedom to define is curtailed abruptly. The DMUs know that the formulation of the measure will strongly impact which DMU appears to be more efficient.

The authors would like to thank the Fachverband Infra in Switzerland for financial support that contributed to this research. Fachverband Infra was not involved in any aspect of the research other than facilitating contacts with case study participants, for which the authors are also grateful.

Benchmarking: An International Journal
Vol. 23 No. 7, 2016
pp. 1876-1891
© Emerald Group Publishing Limited
1463-5771
DOI 10.1108/BIJ-01-2015-0005

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

Three challenges need to be overcome in the design of a comparable measure. The environments differ in which the production takes place. Data measurement methods differ both in the definition of the measures and in the details of how the data are collected. Finally, outputs and inputs may not be well-defined objects, but may differ from each other in subtle ways such a quality or timing. This paper essentially describes an example of how to overcome these challenges in a specific case study regarding a specific type of road maintenance: periodic renewal. Of particular interest to readers involved in performance measurement, but not necessarily in applications with respect to road network managers, is the way the proposed performance measure is constructed from physical relationships. Although other industries will require other specific functions, we believe acceptance of the measures by practitioners will be greater if the cost functions can be specifically related to physical realities.

Switzerland is a small country having a wide range of environments that challenge the road network manager in designing and implementing maintenance. On the other hand, the relatively small geographic area makes an assumption of ceteris peribus for many economic variables reasonably acceptable. Personnel derive from the same technical universities, wage rates are competitive, employed technologies are easily transferred across DMUs. Unfortunately, the same cannot be said for differences across administrative processes that have a large impact on data generation. But that same independence is the pre-condition for meaningfully distinct DMUs. Important to the methodological choices made in the study is the fact that nearly all road maintenance of the type of interest is carried out under third-party competitive bid contracts. Therefore most of the questions relating to organizing the physical production process are not in scope. The important trade-off between labor and capital is made by the contractor, not by the road administrator. The central concern of the case study is how various environmental variables, which are referred to as "challenge factors" so as not to suggest that only physical environmental issues are of concern, affect per-unit cost of a given project.

We report here only on the first phase of a three-phase project. The first phase was intentionally defined as exploratory and it therefore considers only a small group of cantons. Its goal was to establish the feasibility of a larger, all-inclusive, benchmarking effort. The included cantons were selected for their relative comparability. Unfortunately the main report of the study was completed under a confidentiality agreement over the participant's data and we are therefore limited in what we can report here. However, for the purposes of this report, this is not a serious defect. Our primary purpose is a discussion of methodologies.

Switzerland maintains statistics on public expenditure on roads with an exceptionally long history and granularity, both in terms of the political units covered (communities, cantons, nation) as well in the division between various cost categories (BFS, Road Accounts[1], various years). For instance "constructive maintenance" is distinguished from "operational maintenance," "improvements" and "new construction" as well as a number of other groupings for other types of road-related costs. Given the independent administrative processes to generate this data, the level of granularity is in practice a curse of riches. Nonetheless, using this data source, Koch and Forster (2010) compared expenditure for constructive maintenance per kilometer road of the cantonal networks as a way to ascertain whether cantonal governments were maintaining their roads sufficiently well. The results showed that some cantons spent up to 22.8 times as much as others per kilometer on constructive maintenance and improvements over a three-year period (p. 16). Surprised by this

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

result, we calculated a 15-year average of expenditure after adjusting for inflation and came up with a similarly dubious ratio of highest to lowest of 10:1.

No doubt, the public officials who found themselves at the low end of such a comparison would have good cause to invoke the apples vs oranges argument: that the comparisons were unfair due to differences in environments. But this only leads to a rephrasing of the question: how much of the difference is inefficiency and how much is due to differing circumstances? To answer that question, methods must be found to compensate for the differences in environments. That in a nutshell is the statement of our research problem, and it surely is a common one.

Our strategy for constructing a performance measure proceeds from a simplified model of the cost function for the production of road services, which is a central purpose of maintenance activity. We add to this function the cost of various challenge factors and explain as we do so what assumptions need to be made to justify the way in which the factors are included. More complex forms could be used in future research that would impose fewer assumptions. But even as they are, the formulations become complex quickly enough. Finally, we add a term to represent inefficiency and isolate it on one side of the equation, away from the cost implications of the challenge factors. That is then the proposed measure. Thereafter, data sources are found to implement the measure. As is usually necessary in practice, the proposed measure must be modified to circumvent missing data elements. The device used here to calculate the measure of output, combining two network inspections together with an expected rate of deterioration, is likely to be useful in other applications related to long-lasting infrastructures.

Literature review

Empirically, total cost for a producer of road quality through periodic maintenance can be represented as the sum of costs of various projects j that occur in a given time period t:

$$TC_t = \sum_{j \in M_t} A_j p_j \left(\Delta C_j, \boldsymbol{z_j} \right) \tag{1}$$

M denotes a long-term maintenance program. The vector $\boldsymbol{z_j}$ represents the level of various challenge factors affecting project j. Examples of challenge factors might be truly environmental, like poor geologic conditions, or they might be abstract, such as the project complexity deriving from having to interact with other infrastructures in an urban environment. ΔC_j is a measure of output per square meter. We refer to ΔC as the "CS jump." It represents the per-unit area improvement of the condition state[2] (CS). p_j is the per square meter unit cost of the project j. It is dependent on both the quantity of output produced, i.e. the amount of CS improvement achieved, and the challenge factors. Finally A_j is the area covered by project j. Summing this over all the projects to do in a year provides the total cost in that year.

The definition of output is not obvious in infrastructure maintenance. Certainly maintenance is about maintaining a given average CS level in the face of deterioration. Two activity archetypes exist for this purpose: those that reduce the speed of deterioration in some time interval following the intervention and those that improve the CS discretely at the moment of intervention. In the latter case, output can be viewed as an amount of CS improvement that can be measured in units of $m^2 \times \Delta C$. Following each project, one could measure the discrete improvement attained. But, as anyone who has driven over a poorly repaired road knows, the improvement is only

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

meaningful if it lasts. Thus output must have a time dimension: it is the attained difference against an ever worsening baseline over a time interval. The worsening baseline is the expected rate of deterioration. A difference against a worsening baseline can be achieved either through discrete improvements or through a slowing of the deterioration process.

An output efficiency measure for road maintenance is therefore a relationship between the total amount of CS improvement achieved in a time period against a deteriorating baseline and the total cost that was required to produce it. As presented in Equation (2) it is still a raw measure because a compensation for challenge factors has not yet been added:

$$Raw\,performance_t = \frac{TC_t}{CS\ improvement_t} = \frac{\sum_{j\in M_t} A_j p_j(\Delta C_j, \boldsymbol{z_j})}{\sum_{j\in M_t} A_j \Delta C_j} \tag{2}$$

What remains is to specify the per-unit area cost function $p(\Delta C, \boldsymbol{z})$ and the relationships that may exist between challenge factors and the amount of area that must be renewed per year. The symbol $p()$ is used rather than $c()$ because the road managers in our case confront project costs as prices paid to third parties.

Cost function specifications available in the literature

A number of studies exist in the academic literature that include cost functions for road maintenance. These differ by factors included in the vector $\boldsymbol{z}$, by the functional form used for $p(\cdot)$ as well as by the methods used to estimate the function's parameters. Alternatively one can turn directly to practitioners or other experts for estimates of $p(\cdot)$. An advantage to the latter approach is that acceptance by those who are to be measured is made more likely than if unfamiliar numeric techniques are used on what inevitably are questionable data sources. A disadvantage of the latter approach is that the practitioners can propose parameters that work to their advantage. In the case study, both approaches were attempted.

The relevant academic literature can be divided into data envelopment analysis (DEA) studies and regression-based studies, of which stochastic frontier analysis (SFA) is the most modern variant. Both approaches have large literatures and many methodological refinements. Ray (2004) provides an introduction to DEA and Kumbhakar and Lovell (2000) give a review of the SFA literature. An excellent comparative coverage of both methodologies is Fried *et al.* (2008).

DEA studies. With respect to determining the efficiency of road network managers, Cook *et al.* (1993), Rouse *et al.* (1997), Rouse and Chiu (2009) and Fallah-Fini *et al.* (2012) are examples that employ a DEA. The DEA methodology does not fundamentally fit our problem. The assumption that factor prices may differ across DMUs, which is basic to the DEA methodology, is not appropriate given the open bidding used to assign contracts. Second, our interest is in economic efficiency rather than simply technical efficiency; the former requires prices to aggregate inputs, the latter does not. Nonetheless DEA studies are of interest in terms of the challenge factors they considered.

Cook *et al.* compare results of Highway Maintenance Patrols in Ontario, Canada. Their measures of outputs are: two-lane length equivalents as a proxy for the network area, average traffic served (interestingly, an adjustment for heavy vehicles is not found to be empirically relevant), the change in the CS measured net of an assumed rate of deterioration, and accident prevention. Their input measures are: expenditure on

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

operation maintenance, expenditure on constructive maintenance, a measure of climate-related difficulty factors. The authors use the derived efficiency measures to conclude that the use of third-party contractors does not clearly change the efficiency ratios.

Rouse *et al.* (1997) apply DEA to operational and constructive maintenance operations managed by a set of 52 regional government units in New Zealand. Only one input measure is considered, total expense, but three output measures are considered: kilometers resealed, kilometers renewed, and vehicle-kilometers travelled. Two outcome measures are considered: roughness and surface condition. Additional controlling variables used in the analysis are total length of the networks, percent of the network in urban vs non-urban areas and a measure of environmental difficulty that included geology and climate factors. In a secondary regression analysis, climate and geology factors were not found to have a statistically significant relation to inefficiency but traffic and percent urban were. Our regression study provided similar findings on geology, heavy vehicles and urbanity. Climate did not enter our study as a variable because it did not differentiate the participants.

Rouse and Chiu (2009) use an improved data set from the same regional government units that newly include measures of road quality. Three different sets of DEAs representing different combinations of observables are run. These represent views on the production process referred to as "efficiency," "effectiveness" and "economy." By assigning the government units into three groups by their three DEA results: those always better than the median, those always worse than the median, and a mixed group; the authors find a basis to define a "best practice" ratio of maintenance intervention types. They find better performers consistently use more routine and less rehabilitative maintenance.

Fallah-Fini *et al.* apply a DEA analysis to the maintenance of 690 km of interstate highways in 19 counties in the state of Virginia, USA. They are interested in the relationship between efficiency and the form of the contracts. Some of the counties in their sample use performance-based contracts to outsource not just construction but the whole maintenance process. To do this, two DEAs are calculated – one for each sub-group – and the resulting frontiers are individually compared to a third frontier calculated with the whole data set. The authors conclude that traditional government provisioning is more efficient than performance-based contracting. Their output variables include a measure for the CS change and a measure for the network size. As in Rouse *et al.*, a set of external challenge factors are regressed against the inefficiency scores to test if statistically significant relationships with the expected signs can be found. Results are only reported for the variables meeting the expectations and significance criteria. High temperature is related to efficiency and the log of snowfall is related to inefficiency. The latter observation is only true for the performance-based contracting. Finally a regression test of a dummy variable to indicate mountainous regions is found to be positively related to inefficiency.

Summarizing, there is a reasonable consensus on the variables to include in an analysis of road maintenance. Expense is the measure of input. The change in the CS and the network's total area are used in all four studies for output. Traffic and/or heavy traffic are used either as output measures or as external challenge factors. Climate or environmental difficulty also appear in each study as external challenge factors. Only accident rates and percent urban appeared singularly. Significance and signs of the external challenge factors are not stable across the studies.

Regression-based studies. Deller *et al.* (1988) and Kalb (2012) use regression methods to specify a relationship between road maintenance cost and various challenge factors.

Deller *et al.* use a data set on county roads for 1,799 townships in the USA. Their goal is to estimate a production function, in this case a translog, to identify whether economies of scale and or scope can be found. Output is simply the length of road maintained without reference to any CS levels, however, differentiation is made by type of road, e.g. paved or gravel. Input is the total expense excluding the cost of new construction. Control variables included in the regression analysis are daily traffic and road subsidies. The study found evidence of economies of scale and scope and the authors concluded that merging smaller townships is likely to result in cost savings. Considering that the average network size was only 61 km in their sample, the finding seems plausible. The smallest cantonal network participating in our study, however, is greater than 400 km.

Kalb (2012) analyzes total expense for road maintenance in 44 German counties (Kreise) in Baden-Württemberg for the period 1990-2004. Of additional interest is that this study uses both an SFA and a DEA approach on the same set of data. In both cases, a production function is estimated and then technical inefficiency is calculated with reference to the frontier. The inefficiencies are then regressed on a second set of variables. Input is total expense including, presumably, all types of maintenance and new construction. A direct CS measure is not available so the authors use as a proxy the number of accidents due to "bad" roads, including accidents due to slippery conditions. The second output is the area in square meter of the road network. The former is not found to have a significant relationship to cost. Unsurprisingly, network size is strongly related to cost. From the set of environmental control variables: two measures of urban/rural, three measures of population, the unemployment rate, the highest altitude in the county, the number of hotels, and the total kilometer-vehicle travelled; only urban (+), hotels (+), population squared (−) and the highest altitude (−) are consistently and statistically significantly related to inefficiency, i.e. higher expense.

The true interest of Kalb, however, is a set of social and political variables that he anticipates will explain the observed inefficiency after environmental conditions have been controlled for. Income is related to higher expense (inefficiency). Outside funding of the county roads (subsidies) also increases expense as does higher tax revenues. For the latter, one can reasonably ask which is cause and which is effect. Finally, a measure of concentration of political power within the local government was not found to be statistically significant. Governments with a left majority were related to higher efficiency in the SFA models and to inefficiency in the DEA models, the latter more significantly. Our study also finds that urbanity and income are significantly related to higher expense per kilometer network. We, however, find that this is also true for population density within settled areas. Kalb's results are mixed with respect to population.

Summarizing, one can certainly conclude that no consensus exists in the literature for a given functional form for the cost function. Indeed, one is impressed by the flexibility in the selection of variables. Kalb's inclusion of the number of hotels per region and a measure of the concentration of political power in the local government are difficult to justify from an engineering perspective. But from a social science perspective it is not difficult to imagine a linkage between expenditure and political concentration or high income. Further, which factors are to be included in "inefficiency," as opposed to being treated as external challenge factors, depends on whose performance is to be measured and which decision variables are under their control. Thus it seems unlikely that either relative prices or functional forms can directly be adopted from the literature in a practical benchmarking exercise. What

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

could be gathered from the literature is a list of candidate challenge factors to be tested as well as examples of values these could take.

Explicit benchmarking examples. To round out the literature review we also note that examples of benchmarking for road maintenance exist in multiple jurisdictions. McNeil *et al.* (2004) offer a useful overview derived from a set of interviews with five different State Departments of Transportations (DOTs) in the USA that were selected for excellent programs. More recent internet searches of published "scorecards" (see Kaplan and Norton, 2008 for the background to this benchmarking approach) of State DOTs leads to interesting examples such as New Hampshire[3] and Pennsylvania[4]. Although examples may exist, we did not observe a case where an explicit effort is made to compare achievement across multiple DOTs. Rather, the comparisons are always against the departments own history and/or a set of department goals. Examples of comparable ratios, such as the percent of highways in poor condition, can be found. But such measures are not the same as an overall measure of output efficiency. Hence no guidance was found from these examples on how to make the output efficiency measures comparable across states.

Proposed model

In absence on a consensus in the literature on a functional form, a linear form having no interactive terms was selected for $p(\Delta C, \boldsymbol{z})$ primarily for reasons of understandability and the limited goals of Phase 1 of our overall project. Each challenge factor can be thought of a causing a mark-up of the price per square meter. Linearity also has the advantage of permitting working in units that express only the mean exposure of each challenge factor for the whole network, which is helpful where differentiated data at the road section level is not available. Second, the model is simplified by assuming a common minimum acceptable CS across all participants. Thus, the potential dependency of $p(\Delta C, \boldsymbol{z})$ on ΔC is not considered. One has from Equation (1) the formal starting point:

$$TC_t = \sum_{j \in M_t} A_j \cdot \sum_{l} p_l \bar{z}_l + \sum_{j \in M_t} A_j \left(\sum_{l} p_l \Delta z_{l,j} \right) \tag{3}$$

where l indexes the challenge factors and $\bar{z}_l$ is the average exposure across all networks to the factor z_l. Thus $\Delta z_{l,j}$ represents the difference in exposure to challenge factor l in the jth project. $p_l \bar{z}_l$ is the cost per square meter of a standard road section across all networks.

Handling challenge factors that increase deterioration

Equation (2) conceals an important relationship between challenge factors and total cost. Where a challenge factor causes quicker deterioration, a larger amount of CS improvement must be achieved per year in order to maintain the same average CS. For example, if the ground under the road is of poor quality, then a given road type will deteriorate more quickly, all else being equal. Assuming for simplicity that interventions always occur when the road reaches an externally specified CS trigger level, so that each intervention results in the same ΔC, then the total amount of CS improvement can be equivalently expressed as a total amount of area to renew. The amount of square meter renewed each year is specified in the definition of the project list for year t, that is in the statement: $j \in M_t$. A method to make this relationship explicit is needed. For this purpose, it is useful to define the concept "the required

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

amount" of area to be renewed within a given time frame. Required can take the meaning that no set of design choices exist that result in a lower cost way of providing the given level of service. More realistically it should take the meaning that if one followed the relevant design codes for the given set of circumstances, then the required amount of renewal will have to be done to maintain the expected level of service.

The required amount of CS improvement

Consider initially a "standard" exposure to challenge factors. Then the minimum life cycle cost design, or perhaps simply the standard design as per applicable codes, will result in some "standard" deterioration function that can be designated $d(\boldsymbol{z})$[5]. The required area to renew can be deduced from the time it takes for this standard road segment to deteriorate by ΔC units of CS to a level $\dot{C}$, the intervention trigger level. Letting $\overline{d}(\Delta C)$ be the average rate of deterioration during this time period, the required area to renew per-unit time $\tilde{A}$ can be written as:

$$\tilde{A} = A\frac{\overline{d}}{\Delta C} \tag{4}$$

A is the total area of the network. For example if the average deterioration rate is 0.1 per year, the trigger level is 2, and the "as new" state is 0, then 1/20th of the network must be renewed per year. $\tilde{A}/A$ is the expected required percent to renew for this standard level of challenge factors.

Consider next a division of the network into subsections indexed by h where the vector of challenge factors $\boldsymbol{z}_h$ is constant in the area A_h resulting in a specific average deterioration level $\overline{d}_h$ in that area; one has the total required area to renew per time unit as:

$$\tilde{A} = \sum_h A_h = \sum_h A_h\frac{\overline{d}_h}{\Delta C} = \sum_h A_h\frac{\lambda_h\overline{d}}{\Delta C} = \frac{\overline{d}}{\Delta C}\sum_h \lambda_h A_h \tag{5}$$

where $\overline{d}_h = \lambda_h\overline{d}$ for some factor of proportionality λ_h. One can see that increasing the average rate of deterioration by a factor λ increases the required area to be renewed per year by the same factor[6]. Thus differences in deterioration enhancing challenge factors can be corrected by multiplying the affected sub-network area by the factor λ.

Compensated performance measure

Recalling Equation (2) we note that the required area enters the performance measure in two ways: one in the numerator as a multiplier of intervention per-unit prices and the other in the denominator as a multiplier of the CS jump. Since each partition h has a constant challenge factor $\boldsymbol{z}_h$, it also has a constant compensated unit price $\boldsymbol{p}'\boldsymbol{z}_h$ which can be divided into a standard level $\boldsymbol{p}'\overline{\boldsymbol{z}}$ and a specific level $\boldsymbol{p}'\Delta\boldsymbol{z}_h$ as in Equation (3). Next one can replace the sum of the area of actual projects $\sum_{j\in M_t} A_j$ with the required amount dependent on the rate of externally imposed deterioration in area h: $\lambda_h A_h$. This results in:

$$\textit{Compensated cost} = \boldsymbol{p}'\overline{\boldsymbol{z}}\sum_h \lambda_h A_h + \sum_h \lambda_h A_h \boldsymbol{p}'\boldsymbol{\Delta z}_h \tag{6}$$

The adjustment to the denominator is:

$$\textit{Required CS improvement} = \Delta C\left(\frac{\overline{d}}{\Delta C}\sum_h \lambda_h A_h\right) = \overline{d}\sum_h \lambda_h A_h \tag{7}$$

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

Substituting these into Equation (2) one arrives at the measure of compensated performance in Equation (8). The interpretation is more specific than it was in Equation (2). By virtue of placing cost in a ratio to the required area to renew – or more correctly – the required amount of CS improvement per-unit time to maintain a given average CS, the measure is relative to a conditional standard. The conditioning necessarily has an engineering component that describes what is achievable given current methods and/or codes:

$$\text{Standard Compensated Perf} = \frac{\boldsymbol{p}'\bar{\boldsymbol{z}}\sum_h \lambda_h A_h + \sum_h \lambda_h A_h \boldsymbol{p}'\boldsymbol{\Delta z}_h}{\bar{d}\sum_h \lambda_h A_h} \tag{8}$$

Adding DMU specific inefficiency to the model

Letting i index DMUs, the empirical form of this measure for each DMU is:

$$\frac{TC_{i,t}}{\Delta\hat{C}_{i,t}} \tag{9}$$

We postpone until the next section a discussion of how $\Delta\hat{C}_{i,t}$ must be measured to reflect the differential rates of deterioration across the network.

One can now write an expression for inefficiency as the difference between the observed DMU performance and the "standard" performance from Equation (8):

$$Inefficiency_{i,t} = \frac{TC_{i,t}}{\Delta\hat{C}_{i,t}} - \frac{\boldsymbol{p}'\bar{\boldsymbol{z}}\sum_h \lambda_h A_{i,h} + \sum_h \lambda_h A_{i,h} \boldsymbol{p}'\boldsymbol{\Delta z}_h}{\bar{d}\sum_h \lambda_h A_{i,h}} \tag{10}$$

A useful feature of this measurement approach is that it allows for different measurement periods to apply to the different DMUs. Although data on spending can usually be found in comparable yearly units, measures of CS improvement depend on intervals of CS inspections and these vary greatly between DMUs. Time enters the relationship linearly through the area to renew $\lambda_h A_h$ which is defined in Equation (5) as a per unit time amount. Provided all the variables are measured for the same time interval, time cancels out of the equation leaving the desired measure: expense per-unit CS improvement. Equation (10) was implemented in the case study and one notes that it is neither purely linear nor purely multiplicative in the cost implications of the challenge factors $\boldsymbol{z}_h$ by virtue of the dependency of λ on $\boldsymbol{z}$.

Specific measurement issues

The only directly observed quantity in the case study for Equation (10) is $TC_{i,t}$. We discuss in this section how the other variables are implemented with available data.

Differential deterioration rates

The factor λ is determined by a "traffic exposure class" which is a proxy for actual traffic exposure. There are five classes in the data set differentiating between minor exposure in residential neighborhoods and heavy exposure on inter-community highways. But super highways that make up the national highway system are excluded. Deterioration rates are taken from Swiss norm VSS 640986 specifying the required financial depreciation of road replacement value so as to estimate the necessary funds to keep the networks in a steady state. Depreciation rates imply a

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

replacement life span which in turn implies an average rate of deterioration. These parameter values differ by the traffic exposure class. λ is the ratio of the implied deterioration rates taking the smallest rate as the base.

Amount of CS improvement

The denominator of the DMUs performance measure is the total amount of CS improvement for the network for a specified time interval. This can be measured using two sequential CS inspections. The difference between these two values for a given road segment is the net change in the CS. It is net because it does not include the impact of interventions that occurred in the time interval. But by postulating a rate of deterioration, the gross amount of CS improvement can be calculated. In fact, given the data from two inspections on section where no improvements took place, the rate of deterioration itself can be estimated. To be consistent with the concept of a "standard rate," the same rate of deterioration must be applied to all participants. As we describe in detail further below, the "standard rate" can thereafter be modified to reflect the exposure to local challenge factors.

A fortuitous feature of this measurement approach is that the net change includes the impact of both types of maintenance intervention impacts: those to improve the CS discretely and those to reduce the speed of deterioration below $d(\boldsymbol{z})$. Adding the net change in CS to the expected deterioration $\Delta C^{d_h}_{t_0,t_1}$ between times t_0 and t_1 gives the measure of gross CS improvement. For a square meter indexed by j in partition h this is:

$$C_{j,t_1} - C_{j,t_0} = \left(\Delta C^{a}_{j,t_0,t_1} + \Delta C^{b}_{j,t_0,t_1}\right) + \Delta C^{d_h}_{j,t_0,t_1} \tag{11}$$

Note that $\Delta C^{d_h}_{j,t_0,t_1}$ will have the opposite sign as $\Delta C^{a}_{j,t_0,t_1}$ and $\Delta C^{b}_{j,t_0,t_1}$. The expectation of $\Delta C^{d_h}_{j,t_0,t_1}$ is:

$$E\left[\Delta C^{\overline{d}_h}_{j,t_0,t_1}\right] = \overline{d}^{h}(t_1 - t_0) = \lambda_h \overline{d}(t_1 - t_0) \tag{12}$$

Reformulating and summing over the total network:

$$\Delta \hat{C}_{i,t} = \sum_h \sum_j \Delta C^{a}_{h,j,t_0,t_1} + \Delta C^{b}_{h,j,t_0,t_1} = \sum_h \sum_j \left(C_{h,j,t_1} - C_{h,j,t_0}\right) - \overline{d}(t_1 - t_0) \sum_h \lambda_h A_h \tag{13}$$

In other words, the gross CS improvement is equal to sum of the observed change of state plus the expected rate of deterioration given the degree of exposure of the network to challenge factors.

Challenge factor exposures

In some cases participants maintained data elements in their road databases on challenge factors. But this was only true for urbanity[7], the fact that a road lies in a built up area, for all participants. In the case of traffic exposure, a road section-level estimate could be obtained for all participants as well. This prevented a direct implementation of the numerator of Equation (8) since, in absence of road-level observations, the amount of impacted road surface could not be calculated. An assumption allows us to precede nonetheless, but the implications of that assumption needs to be considered when interpreting the final values. The assumption is that those challenge factors that require a financial compensation are distributed equally across the network. This allows dropping the subscript h on the vector $\Delta \boldsymbol{z}_h$ in the numerator so that the price can be taken out of the sum. Let an upper case Z represent the average network exposure to challenge factor having a non-zero price:

$$\boldsymbol{p}'\overline{\boldsymbol{z}}\sum_h \lambda_h A_h + \sum_h \lambda_h A_h \boldsymbol{p}'\Delta \boldsymbol{z}_h \cong \boldsymbol{p}'\overline{\boldsymbol{z}}\sum_h \lambda_h A_h + \boldsymbol{p}'\Delta \boldsymbol{Z}_i \sum_h \lambda_h A_h \tag{14}$$

In other words, the benchmark measure for each network manager is adjusted for the extra cost implied by the network average exposure to financially compensated challenge factors. Substituting into Equation (10) and simplifying we get:

$$Inefficiency_{i,t} = \frac{TC_{i,t}}{\Delta \hat{C}_{i,t}} - (\boldsymbol{p}'\overline{\boldsymbol{z}} + \boldsymbol{p}'\Delta \boldsymbol{Z_i}) \tag{15}$$

Financial impact of exposures to challenge factors: $\boldsymbol{p}$

Switzerland has only 26 cantons and it was therefore deemed unlikely that conclusive cost estimates could be estimated from cantonal data given the number of independent factors affecting cost. An alternative data set exists for Swiss communities for which roughly 130 have data published annually by the Swiss Association of Cities (Schweizerischer Städteverband, 1978/2011). A panel including 30 years of data could be acquired but its usefulness as a panel data set is limited by two features. First, for most of the explanatory variables, there is little variability from year to year. Geologic and geometric feature do not change at all. Residents per settled area change very slowly. Second, smaller communities do not have enough road projects to smooth the impact of specific projects in each year's annual budget. For the purposes of demonstrating the viability of the research approach, which is the goal of Phase 1 of our project, pooled weighted least squares on a linear model was seen as sufficient. We do not submit these estimates for use as general measures of the average impacts of the respective factors. For that, too many empirical issues remain unaddressed.

The time series is long enough to provide four periods of roughly eight years that are believed long enough to smooth the effect of individual projects. Additionally, since there is no measure of road quality, it is thought that longer periods will also smooth the effect of differences in the intensity of investment in roads. One of the period observations was reserved to estimate change effects, but none of these were significant and they are not included in the presented results. Thus, 392 observations were used in the regression. Two time dummies were included for the third and fourth periods to capture period effects.

Our selection of challenge factors derived from interviews with participants and the literature review. Table I lists those for which data sources could be found. Unfortunately

Variable	Unit	Mean	SD	Max.	Min.
New and renew expense/kilometer	1,000 CHF	34.21	25.81	250	3.07
Settled area per kilometer	Hectares	9.67	4.11	37.7	2.45
Residents per settled area	Persons/ha	39.35	14.81	124	13.5
Trucks registered per kilometer	Trucks	13.48	9.08	77.5	2.27
Tax per resident	CHF	1,144	963	9,474	421
Average slope	Degrees	5.72	4.94	27.7	0.14
Index of geologic difficulty	[−1, +1]	0.029	0.087	0.50	0.00

Table I.
Regression variables

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

it was not possible to get road-level data for many variables and community-level observation had to be substituted. Therefore in the regression, road networks are equally exposed to challenge factors as are the regions in which they lie. We believe this is an important departure from reality both for the regression study as for the subsequent application to the cantonal networks. In future research we believe this difficulty can be overcome.

A discussion of each factor would be too detailed for the purposes of this presentation but some general issues are worth mentioning. The interview rounds resulted in a reasonably broad consensus on challenge factors. All of those listed except the average wealth, proxied by average taxable income, were mentioned in the interviews. Wealth appears as an exogenous variable in Kalb (2012) and, according to standard economic theory, if road quality is a "normal good" then higher demand for it will result if wealth goes up (see for instance Varian, 1992). "Residents per Settled Area" is a proxy for the complexity of projects due to the fact that road renewal projects can be substantially more difficult to plan and carry out if other infrastructures share the same space or interact in some way. Particularly in densely populated urban areas, the road space is used for many purposes including various utilities as well as other modes of transport like tramways. All the interviews mentioned these factors as important cost drivers.

The results in Table II show that statistically significant relationships between nearly all of the cost factors could be demonstrated. All the variables have the expected signs and in all but one case the *t*-statistics are comfortably high so that the sign of the effect is not in question. The numeric importance is also not at issue. Excepting geology, each of the one standard deviation impacts is larger than 10 percent of the mean value of the independent variable: 34.21. From this we conclude that these

R	R^2	Adjusted R^2	SE of the estimate			
0.613	0.3755	0.3625	328.7			
	Sum of squares	df	Mean square	F statistic		
Regression	24,893,773	8	3,111,721	28.79715675		
Residual	41,385,662	383	108,056			
Total	66,279,435	391				
Variable	Coef.	SD	*t*-Stat.	Signif. level	Mean impact	1 SD impact from mean
(Constant)	−8.444	6.503	−1.298	0.195		–
Dummy_94_01	−24.439	3.776	−6.473	0.000		–
Dummy_02_09	−19.856	3.699	−5.367	0.000		–
Settled area per kilometer	2.455	0.392	6.267	0.000	23.74	10.08
Residents per settled area	0.455	0.091	4.980	0.000	17.90	6.74
Trucks per kilometer	0.581	0.185	3.138	0.002	7.83	5.28
Tax per resident	0.00486	0.001	3.471	0.001	5.57	4.69
Average slope	1.178	0.331	3.561	0.000	6.73	5.82
Geologic difficulty	29.409	17.568	1.674	0.095	0.86	2.55

Source: Richmond *et al.* (2013 unpublished report)

Table II. Estimation of the financial impact of challenge factors on road maintenance expense environmental variables should be included as cost-generating control factors in the calculation of inter-DMU performance measures. The magnitude of the effects of the two population-related variables is noteworthy and consistent with the emphasis placed on these variables in the interviews.

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

Implementation of proposed measure

Table III gives four measures of relative efficiency where the actual results are divided by the mean for easier comparison. A "naïve" efficiency measure that would result from using publicly available data on expense per kilometer, without any adjustments is shown in column 2. Columns 3 and 4 show the results of implementing Equation (15). "Without financial compensation" means that all prices for challenge factors are assumed to be zero. Thus only the compensation working through the different rates of deterioration is operable. Column 4 implements Equation (15) completely using the price estimates from the regression study.

Unfortunately since the actual inefficiency is not observable, there is no way to say which measure predicts it better. But one can expect that measures resulting in lower overall variance between the participants are likely to be eliminating exogenous differences. The following general observations can be made. First, one can see a marked reduction in variance of −33 percent between the naïve measure, column 2 and the other two. Second, the difference between the best and worst cantons still seems dubiously large since the ratio remains greater than 2 (1.51/0.71). Third, the financial adjustment using the regression parameters actually increases the variance. The small sample size of five observations is not large enough to draw any definite conclusion about such a small difference in variance, but clearly the financial compensations are not working as hoped and further research is required. Here we see the main opportunity in obtaining network-level data for the regression data set. Fourth, rank orderings between the naïve and the proposed concepts are very different. Between the two versions of the proposed measure there is almost no difference in ranking. The basic structure of the measure: setting total expenditure in relation to required CS improvement; must therefore be driving the reduction in variance between the naïve and proposed measures.

Given these results, the case study concluded that the general approach is promising and certainly better than the naïve measure. However, the proposed measure in its current form with available data is not sufficiently reliable for cross-DMU comparisons in general. This is not necessarily true for subgroups of cantons that have comparable accounting processes nor for individual cantons vs their own history. Particularly the later aspect recommends that the cantons calculate the measure for internal monitoring purposes.

Participants	Periodic maintenance expense/kilometer (15-year average, real)	Index of relative efficiency: Without financial compensation	Index of relative efficiency: With financial compensation
A	0.98	1.51	1.57
B	0.66	0.80	0.78
C	1.11	0.71	0.73
D	0.54	0.93	1.03
E	1.72	1.04	0.90
SD	0.42	0.28	0.30

Table III. Performance measures with and without financial compensations

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

Conclusions

A case study is carried out with five cantonal road management authorities in Switzerland to assess opportunities to create a comparative performance measure of renewal interventions. Beginning with a model of maintenance cost for a road network, a measure is derived that includes adjustments for differences in the exposure to challenge factors. It could be shown that the functional approach reduced the inter-DMU variance over the naïve measure by one-third but the attempt to estimate the financial impact of challenge factors was not successful. A large source of uncertainty derives from differential accounting practices across cantons and the lack of road network data in the regression study. Due to these constraints, the researchers recommend using the current measure only for comparisons with the canton's own history or with specifically selected cantons having comparable accounting processes.

It has been shown that performance measures for road maintenance can be built from the bottom up proceeding from physical models of how the outputs must relate to inputs and challenge factors. It is believed that this approach makes compensations for challenge factors more transparent and hence potentially more believable to practitioners than alternative approaches for inter-DMU efficiency comparisons. An important avenue for further research, beyond improvements in the data, is a differentiation in the basic model between challenge factors and design responses to them. That aspect is suppressed in the current model. In the current model, challenge factors relate directly to construction costs. But in fact, design choices mitigate between the environmental parameters and the cost of project implementation. In all likelihood, the numerator in the performance measure would then not be linear in the challenge factors.

Notes

1. See www.bfs.admin.ch/bfs/portal/en/index/themen/11/02/blank/02.html for a description.
2. Road quality, or more generally the road's state, is measured by various indexes called condition state indices. This might be the smoothness of the road or the friction of the surface to help vehicles stop. CS indices generally are numeric on a fixed interval such as 1-10. Sometimes large numbers are good and other times they are bad. In Switzerland, 0 is best and 5 is worst.
3. See www.nh.gov/dot/org/commissioner/balanced-scorecard/index.htm
4. See www.dot.state.pa.us/penndot/districts/district12.nsf/Binder2.pdf
5. In Switzerland the applicable code is VSS 640 324 and others referenced in it.
6. This is exactly true for linear deterioration. It is also exactly true for a class of deterioration functions where the relative amount of time spent in each condition state across the life cycle of the road remains constant. But other families of deterioration functions also exist. For the case study and this presentation we maintain proportion preserving changes to the deterioration function.
7. Switzerland has relatively strict laws on where buildings may be constructed making the distinction "built-up" vs "not built-up" quite distinct. Naturally, built-up areas require many features that are not otherwise present such as capturing run-off water, sidewalks, and interaction with utilities.

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)

References

Cook, W.W., Kazakov, A. and Roll, Y. (1993), "On the measurement and monitoring of relative efficiency of highway mainenance patrols", in Charnes, A., Cooper, W.W., Lewin, A.Y. and Seiford, L.M. (Eds), *Data Envelopment Analysis: Theory, Methodoogy and Application*, Kluwer Academic Publishers, London, pp. 195-210.

Deller, S.C., Chicoine, D.L. and Walzer, N. (1988), "Economies of size and scope in rural low-volume roads", *The Review of Economics and Statistics*, Vol. 70 No. 3, pp. 459-465.

Fallah-Fini, S., Triantis, K., de la Garza, J.M. and Seaver, W.L. (2012), "Measuring the efficiency of highway maintenance contracting strategies: a bootstrapped non-parametric meta-frontier approach", *European Journal of Operational Research*, Vol. 219 No. 1, pp. 134-145.

Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (2008), *The Measurement of Productive Efficiency and Productivity Growth*, Oxford University Press, New York, NY.

Kalb, A. (2012), "What determines local governments' cost-efficiency? The case of road maintenance", *Regional Studies*, Vol. 48 No. 9, pp. 1483-1498.

Kaplan, R.S. and Norton, D.P. (2008), *The Execution Premium*, Harvard Business School Press, Boston, MA.

Koch, B. and Forster, M. (2010), *Zustandsanalyse und Werterhaltung bei den Kantonsstrassen in der Schweiz*, Infrastruktur Strasse, Zürich, available at: www.bevoelkerungsschutz.admin.ch/internet/bs/en/home/themen/ski/aktuell.parsys.90593.DownloadFile.tmp/100624infrastrukturstrassed.pdf (accessed April 21, 2011).

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

McNeil, S., Ogard, E., Pagano, A., Pal, S., Morreale, A., Schermann, J. and Berner, J. (2004), "Linking strategic goals and asset management for resource allocation and decision making using pavement management", *6th International Conference on Managing Pavements, Queensland, October 19-24*.

Ray, S.C. (2004), *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*, Cambridge University Press.

Rouse, P. and Chiu, T. (2009), "Towards optimal life cycle management in a road maintenance setting using DEA", *European Journal of Operational Research*, Vol. 196 No. 2, pp. 672-681.

Rouse, P., Putterill, M. and Ryan, D. (1997), "Towards a general managerial framework for performance measurement: a comprehensive highway maintenance application", *Journal of Productivity Analysis*, Vol. 8 No. 2, pp. 127-149.

Schweizerischer Städteverband (1978/2011), "Statistisches Jahrbuch des Schweizerischen Städteverbandes", Schweizerischer Städteverband, Bern.

Varian, H.R. (1992), *Microeconomic Analysis*, 3rd ed., W.W. Norton & Company, New York, NY.

About the authors

Dr Craig M. Richmond received his PhD in Regional Economics from the University of Pittsburgh. He worked in the Asset Management divisions of a number of Swiss Banks in functions related to benchmarking and performance measurement, among other these included the Global Head of Portfolio Analytics for Credit Suisse Asset Management and the Head of Investment Control for Rothschild Private Bank, Zurich. More recently, Dr Richmond has been working as a Senior Research Associate in the Infrastructural Management Group (IMG) at the Institute of Construction and Infrastructure Management of the ETH Zurich. His research focusses on the development of engineering-based performance measures for use in managing large public infrastructures in particular with respect to roads. Dr Craig M. Richmond is the corresponding author and can be contacted at: Richmond@ibi.baug.ethz.ch

Clemens Kielhauser obtained his Diploma in Civil Engineering Sciences from the Vienna University of Technology in November 2011. His Diploma thesis (topic: implementing a waste management knowledge base, based on the example of the implementation of the Landfill Ordinance) was written at the Institute for Water Quality, Resource and Waste Management, Research Center of Waste and Resource Management and awarded the Hans Roth Environmental Award 2012. During his studies, Clemens Kielhauser worked as an intern in the field of project management and controlling. He also has experience as a member of the technical staff in a landfill remediation project. Since March 2012 he has been working as a Research Assistant at the Infrastructural Management Group (IMG) at the Institute of Construction and Infrastructure Management of the ETH Zurich. Here, he is responsible for the lectures in systems engineering. In December 2012 he started his doctoral studies on the topic of interrelated infrastructure networks in urban areas at IMG.

Professor Bryan T. Adey obtained a Bachelor's of Civil Engineering from the Dalhousie University in 1995, a Master's of Science in Structural Engineering from the University of Alberta in 1997 and a PhD in the area of Infrastructure Management from the Swiss Federal Institute of Technology in Lausanne (EPFL), Switzerland in 2002. In 2003, Dr Adey co-founded the consultancy Infrastructure Management Consultants Ltd (IMC) and assumed the role as vice-president. From 2003 to 2009 he evaluated, developed and improved business models and business processes related to the maintenance and operation of infrastructure, developed and implemented infrastructure management systems, and analyzed infrastructure with respect to physical condition, risk and economics. In 2010, Dr Adey left IMC to become the Chair of the Infrastructure Management Group (IMG) in the Institute of Construction and Infrastructure Management (IBI) at the Swiss Federal Institute of Technology in Zürich (ETHZ). His research interests are in improving how infrastructure managers make predictions, determine maintenance programs and construct the processes they use.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 00:36 14 November 2016 (PT)