Adding Realism to Cost Functions for Road Network Renewal to Identify Better Management

Working paper submitted to the Industry Studies Association Conference, Portland Oregon May, 2014

Version 02.05.2014.

Adding Realism to Cost Functions for Road Network Renewal to Identify Better Management

Craig Richmond [a *], [1] Meike Akveld [b], [2] Clemens Kielhauser [c], [3] and Bryan T. Adey [d], [4]

[a,c,d] *Institute of Construction and Infrastructure Management, Swiss Federal Institute of Technology (ETHZ), 8093, Zurich, Switzerland*;

[b] *Department of Mathematics, Swiss Federal Institute of Technology (ETHZ), 8092, Zurich, Switzerland*;

ABSTRACT

A cost function for road network renewal is derived analytically that is consistent with a basic model of road deterioration under the influence of multiple modes of use and a set of cost influencing variables. The model is intended for use in benchmarking public works departments that operate under unique conditions and are therefore not readily comparable. Deterioration is modeled as a conservative vector field. The average annual network renewal costs are modeled as a function of the time to renewal and project costs. The model allows pavement design to vary in face of environmental challenge and intensity of use in order to minimize steady state cost.

Keywords: Highway Renewal, Cost Functions, Deterioration Modelling, Benchmarking, Performance Measures.

1 INTRODUCTION

This paper is best understood as a case study in adding engineering 'micro-foundations' to what is in principle a economist's standard estimation of a cost function for a production process. The final goal of the research is the calculation of comparable efficiency measures, benchmarks, with which to judge managers who maintain public road networks. One key obstacle to overcome in the development of the efficiency measure is the requirement that it be realistic enough to obtain the endorsement of a set of practicing road network managers. A fair respect for the reasonable skepticism of such managers suggests that the proponents of the measure should be able to demonstrate, through the logic of mathematical modelling, that the measure would actually measure management efficiency in some sufficiently simplistic hypothetical world. It is only through the definition of such a model, that its realism can be challenged and defended. This paper presents such a model for renewal activity of road networks.

With respect to discussing methods to define cost functions with engineering foundations, road network maintenance has the advantage of being familiar to most readers. Roads are omnipresent as is their production,

[1] * Corresponding author. Email: richmond@ibi.baug.ethz.ch

[2] Co-author. Email: meike.akveld@math.ethz.ch

[3] Co-author. Email: kielhauser@ibi.baug.ethz.ch

[4] Co-author. Email: adey@ibi.baug.ethz.ch

or rather, their re-production as we all know from summer driving. *Maintenance* as a category includes multiple activities. This paper is concerned with renewal maintenance. It does not consider either the construction of new roads or the more or less continuous activities that can be subsumed under the heading 'operational maintenance'. Both of these would require a very different modelling approach. Even if the delineation between these categories of activity is not black and white, statistics on expenditure for roads, as well as the structure of road management organizations, tend to respect this high level categorization.

Three salient features characterize renewal maintenance and these motivate the modelling approach. First, renewal is 'lumpy' in time, that is periodic, so that the length of the period between successive renewals is a key feature to model. Second, renewal is lumpy in space. The whole surface of the road over a significantly long stretch is done at once. The replacement or repair of less than the full width of a road section is relegated to the *operational maintenance* category. Third, most physical construction is done by third parties. Therefore the actual physical production including trucks, asphalt and hardhats, is largely absent from the model. This is also true for most recorded data. Whereas it is not unusual to know how much labor input goes into operational maintenance, since this is frequently carried out by the public works department staff, it is unusual to know this for renewal maintenance. Obviously, these delineations are idealized and are not true everywhere and always.

The assumed institutional setting for the model is a public road administration, referred to as a decision making unit (DMU). We assume the network is large enough so that individual projects are unimportant. The primary activity of the road management organization with respect to renewal management consists of monitoring the condition of the roads and of defining and supervising renewal projects.

An important motivation for modelling this activity is that competition does not exist for DMU's in the sense that it exists for producers of apples when they sell nearly identical products in nearly identical markets. Competition can exist through substitute goods: boats, rail, trolleys etc. But no two DMU managers provide the same set of transportation services in the same geographic area. That fact strongly limits the ability to use one DMU to benchmark another. They are always different. And unless someone can demonstrate the cost impact of those differences, any comparison across DMU's is easily rebutted. This research is aimed at answering that rebuttal.

The literature review, which starts the paper, shows that there is a gap between bottom-up modelling approaches, provided largely by engineers, and top-down modelling approaches, provided largely by economists. It further shows that this gap limits practitioners' ability to benchmark DMU's in important ways. In the modelling section that follows, we demonstrate with a mathematically tractable model, how a bridge can be built between the engineering models and the econometric models. Unfortunately in the interest of tractability, concessions had to be made to realism. The model is therefore only offered as a *proof of concept* of the approach, not as a serious proposal of how renewal maintenance should in fact be modelled. It is our intention to improve the realism in future research. Conclusions and a discussion of further research directions end the paper.

A comment on terminology

Given that this paper is intended for an interdisciplinary audience a few introductory words on terminology are useful. Activity to repair, change or improve a road is generically called an 'intervention'. A renewal intervention rebuilds some or all of the various pavement layers. Depending on the pavement's design, different materials are used and these give the new pavement specific characteristics that affect, e.g., how long it lasts. In the economist's jargon these materials are 'inputs' into a production function. Certainly no engineer would consider a pavement's design as well specified solely by a list of inputs quantities, but that is how an economist's model would typically model the cost of the project: the sum of a list of quantities multiplied by their price. And, since anything that can be paid for, can be put on a list, the price list approach does not lack generality. In fact, for the purposes of the current model, it has too much generality.

If a model of costs were linear, it would be natural to create two categories of independent variables: those that change the intercept and those that change the slope. Similarly, in the current model we need to distinguish between factors related to project costs that determine the longevity of the road section on the one hand, from those that do not affect longevity but do raise project costs on the other. To understand the distinction, one needs to imagine a specific road design, and all the quantities it would take to construct it, first in an 'easy'

environment and then in a difficult environment. One can imagine a rural environment with little traffic versus a highly travelled urban environment where construction has to be interrupted to allow trolleys to pass. Quantities of inputs that would apply in both cases can be thought of as one category of variable which will be referred to as 'input variables'. Environmental factors, the exposures to which generate additional project costs but have no direct influence on longevity, will be collected in a category called 'non-deterioration related'. In reality, some factors could reasonably be placed in both. The 'exposure' to buried water lines can make project planning more complex, but it can also result in frequent *utility cuts* being made into the pavement to enact repairs, and these cuts weaken the integrity of the road section's structure and thereby amplify deterioration. Such complexities will not be addressed in the current model. One can however imagine two variables, one to measure the project cost aspects related to utilities sharing the same road space and one related to the likelihood of utility cuts occurring during the lifespan of the pavement.

Yet another category of variable can be found in the features of the environment that affect how quickly a specific pavement design deteriorates, but which do not affect project costs. Holding a road section's design constant, it would deteriorate more rapidly in a curve than on a straight, and more rapidly if built on soft ground rather than solid ground. Environmental features like 'poor geology' increase the cost of maintaining a network, even when they have no direct effect of project cost. They do this by making the time interval between successive renewals shorter. The objection that 'poor geology' does of course affect project costs through the selection of more expensive designs, is not being overlooked. Rather, one should expect in any internal solution to an optimization problem, that the 'pain' is taken in multiple ways. There should be an adjustment to a deterioration enhancing environments through the selected design, but there should also be an adjustment through the optimal renewal interval. A benefit for engineers to be gained from an analytic formulation of the cost function is the ability to describe such substitution behaviors, but it can only be done if categories of variables distinguish between exposures themselves and the design responses to them.

A forth category is required to capture variables that measure how a road is used. The intensity of transits by heavy vehicles is the standard example, but other vehicle types also use roads. However, because we are modelling road renewal and not the use of roads, our interest is less in use per say, than it is in the 'using up' of the road. Variables that measure factors that use-up the road will be called 'use' variables or use factors.

When we need to refer jointly to all three categories of variables other than the use variables, we call these generically 'challenge factors'. A likely alternative term 'environmental variables' can be confusing in the context of roads. The concept *challenge factor* is intended to be more general than things in the physical environment. It should capture anything that has a relevance to the cost of maintaining a network and is not within the control of the road manager. Challenge factors are always measured in the direction of increasing cost, e.g. ground weakness rather than ground strength.

2 LITERATURE REVIEW

Three relatively distinct sets of literature are relevant and we refer to these for convenience as 'engineering', 'duration' and 'econometric' literatures. Our interest is in the employed modelling approaches that relate to cost functions for renewal maintenance of road networks. In keeping with the economic literature, the cost function expresses the minimal cost of producing a given output and not simply the cost, which could be much higher, if the planning or execution were done poorly. The derivation of a cost function is not necessarily the goal of the studies themselves. It is therefore useful at the outset to summarize what we are looking for.

Schematically, the derivation of a cost function begins with the statement of cost in general that would result if various inputs were used to produce a given level of output. That statement is the objective to minimize by selecting a set of choices subject to a set of constraints and a context described by levels of exogenous variables. The relationship between the selected inputs and the resulting outputs is expressed by a *production function*. The next step, which is not common to all three literatures, is to solve the cost minimization problem for a set of *demand functions.* These specify the choices of inputs that must be made under all the various constellation of the exogenous variables in order to generate the required output at minimal cost. The context includes prices and challenge factors. The *cost function* only results when the demand functions are substituted back into the original cost minimization problem. The cost function expresses then the minimum cost of producing a given level of

output. The input quantities are no longer independent variables because they have been internalized through the demand functions. The cost function relates prices and challenge factors directly to cost. It is easy to see that cost functions are in fact benchmarks. Benchmarks also express what an activity should cost if done in the cost minimizing way. Three topics are of particular concern: 1) How is the production function modelled?; 2) Which functionally distinct categories of variables are employed?; and 3) How are the models solved to arrive at the minimum cost of renewal?

Engineers are dedicated to finding real solutions to real problems and it is not surprising that the relevant studies to an estimation of a cost function generally conclude with an algorithm to derive an optimal renewal strategy. That is, they conclude with the demand function and do not go on to substitute it back into the cost minimization objective to derive a cost function. The lack of need to derive anything beyond the optimal strategy makes an analytic solution unnecessary. Numeric solutions are sufficient. In practical terms, they can be implemented in computer based planning tools for pavement management.

None the less, cost functions are implicitly present in the routines to value and select an optimal strategy. We discuss here two examples. Santos and Ferreira (2013) formulate the road manager's problem as a numeric optimization model. The choice variable is the selection from a set of pavement designs that are to be applied to specific road contexts. The specific designs form a set of discrete choices. As such, an analytic solution to the cost minimization problem is precluded from the outset. However, the underlying relationship linking pavement design and longevity, the structural number, is a continuous function of pavement thickness. Through the structural number, there is in principle a continuous trade-off relationship between various design elements with respect to their construction costs and longevity benefits. In the vocabulary of economists this would be marginal cost versus marginal benefit. But in Santos and Ferreira's model, this trade-off behavior is condensed into the selection from a set of discrete pavement designs. A lower bound on road quality is imposed by rising user costs as quality declines. The challenge factors include variables on traffic levels, sub-base strength[5] (how rigid the supporting layers are, where the pavement is to be renewed) and drainage. Traffic per annum determines how long a road will last given its design. The sub-base strength and drainage are parameters in the pavement design equation and modify how much traffic the road will withstand before depletion. The pavement design equation employed is taken from AASHTO (1993). Total demand for the quantities of inputs within the planning horizon, and hence total cost, is a function of the thicknesses, pavement type and how often the road must be renewed.

From this brief description one can see that *output* of the renewal activity is modelled as the creation of a given *capacity* to carry traffic. The production function is the pavement design equation. Three categories of variables are employed. The *input* variable is the selection of the pavement design. *Use* variables are limited to ESALs, Equivalent Standard Axel Loads. ESAL is function relating a vehicle's actual axel-load-transits to an amount of transits of a standard axel load that would cause an equivalent amount of damage (AASHTO, 1993). Third, sub-base strength and drainage are *deterioration related* variables. The model is shown to be solvable numerically for a fixed time horizon. It can be used in simulation exercises or to develop renewal plans for the network. Since the model yields the minimum cost way of producing a given amount of traffic carrying capacity it can be seen as a cost function. Demand functions for pavement types are not derived but are, none the less, implicit in the model.

A second example of a family of engineering models is (Ouyang & Madanat, 2006) that does solve for an analytic solution to a road renewal problem. To do so, it must overcome the discreteness of pavement design in Santos and Ferreira by defining design continuously as 'intensity', which represents the pavement's thickness. The economic behavioral choice is therefore limited to the trade-off between intensity and the renewal interval. As above, user costs provide a lower bound on quality. The production function actually comprises two functions. Jointly they express how much traffic a pavement section will carry based on the intensity of the previous intervention. That is, output in the model is again the capacity to carry traffic. The presented model does not explicitly include deterioration challenge variables, but parameters in the deterioration function can be varied to reflect different exogenous conditions. The model is solved using dynamic programing under constraints on the beginning and ending condition states. A generalization of the model in Sathaye and Madanat (2012) uses a budget constraint instead of boundary conditions to limit the model. A closed form of the analytic

[5] Resilient modulas

solution, which would be the demand function, is not provided in either paper. Rather an implementable algorithm is provided in both cases.

In general it can be said, that these models are developed for a single physical discrete pavement section. They employ a physical design equation to relate design choices and challenge factors to derive an amount of road capacity. The level of annual traffic determines how quickly that capacity is used up, and hence the renewal interval. Jointly, the renewal interval and the selected pavement's construction cost determine the owner's cost. By determining the pavement design and the renewal interval, the minimum cost method of providing the required road capacity is determined. The models do not provide closed form solutions that can be substituted into the original cost model. Because the starting point of the model is a discrete decision on a discrete pavement section, it seems immediately reasonable to formulate the cost minimization problem as the discounted present value of various cash flows. Even when the modelling context is changed from a fixed horizon to the more tractable infinite horizon, the discounted present value is what is being minimized. This turns out to be neither necessary nor the easiest route to take. This is shown by two papers of the 'duration' type to be discussed next. They take a middle ground between the bottom-up engineering models just discussed and the top-down econometric models to be discussed further below.

As a rule, a manager does not manage a single road section; he or she manages a road network and what is being managed has neither a begin date nor an end date. A relative innocuous abstraction is therefore to work at a level of annual averages for an assumed steady state rather than include the details of aggregating from individual pavement sections. This can provide the important dividend of taking the interest rate out of the problem. This is shown in both Newbery (1988) Equation 9 and Lindberg (2002) Equation 6. In both cases it is shown mathematically what is intuitively relatively straightforward. Regardless how one comes up with an optimal strategy for each pavement section, in a steady state it will repeat every T years. Consider a partitioning of the network so that all the sections having the same interval T_i are grouped together. Consider a partition P_j with interval T_j. In this case, the renewal of P_j can be paid for via a sinking fund financed with annuities, one for each road section. The size of the individual annuities clearly will depend on the interest rate. But, if P_j is large enough so that it can be further divided into $1/T_j$ subdivisions such that each subdivision has the same cost of renewal, then one could also finance the renewal maintenance of P_j with a constant annual payment out of the running budget covering current period renewal projects. That payment is dependent only on current period expenditures and therefore has no dependency on the interest rate. If the interval T_j were to change, for instance because a general weight limit on heavy vehicles was raised, then the impact of that change on required annual spending has no dependency on the interest rate[6].

In many respects the model we present below follows the approach taken in Newbery and later in Lindberg. Newbery adopts a uniform length of pavement as his modelling environment. Importantly, the age of the pavement is uniformly distributed between *0* and *T*. The time to renewal, *T*, is then the key variable in the model and the reason why these models are referred to as duration models (Link, 2014). However, unlike the engineering models, the ability of road managers to adjust the design to minimize cost is not the focus. Cost of an overlay is a constant. Their interest lies in estimating the marginal social cost of heavy traffic and formulating the relationship between traffic and T. Of particular interest is that both authors formulate a version of their models where the consumption of road capacity is affected by both weather and traffic. The importance of the inclusion of weather is not completely obvious and not explainable from a desire to make the models more realistic; there are numerous other ways in which the models could be made more realistic. If weather is included, as in Newbery, as a means by which road capacity is consumed even when no traffic passes[7], then this

[6] This has the very useful property that shares of cost to assign to types of users can be calculated from the amount that their use of the road shortens the interval of renewal. A simplified example: If truck usage reduces the life expectancy from 20 years to 10 years, then trucks should be charged the difference between 1/20 and 1/10 of the cost of renewal. However, for this approach to be implementable, one must have an estimate of how long the road would last if heavy vehicles did not use the road.

[7] The alternative is to include weather as a parameter that increases the damage done by traffic, but has no affect if traffic is zero.

can have an important impact on the calculation of marginal cost and/or cost-shares to assign to various users of roads, both of which are key objectives of this literature.

The relevance of the inclusion of weather is particularly important since it has been shown experimentally (Highway Research Board, (1962)) that heavy vehicles cause far more damage to roads than light ones[8]. If however light vehicles are assigned the cost of maintaining road capacity in face of damage done by the effects of time – effects that would occur even if no heavy vehicles travelled – then the assigned shares of road costs might be quite different. That issue has considerable importance. Newbery reports the shares of road renewal cost covered by an efficient marginal cost based vehicle charge as varying between 30 and 90% depending on weather and traffic intensity [Newbery (1988), Table 2].

Econometric studies, usually done by economists, is a third set of models relating to specifying cost functions. We discuss three cases that are similar to studies of other transport infrastructures Deller, Chicoine, and Walzer (1988); or Kalb (2009) and Link (2006). These studies do not build their models from a characterization of the physical process of renewal, the functional forms are selected from the set of functional forms commonly used in economics. With these functions, how one can derive the cost function from the production function is well established so that the cost minimizing solution is available in analytic form. Numeric procedures are unnecessary. In the three mentioned studies the translog function form was assumed. The primary downside of this approach is, even if the functional form could in fact be related to the physical productions process in a convincing way, it has not been done. A skeptic can reasonably shrug off the results of as irrelevant to practice. One question that held the first author in suspense throughout the derivation which follows in Section 3 was: to what extent the cost function derived from engineering considerations would resemble one of the standard functional forms used in econometric functions

Differences in the modelling strategy to the previous two sets, beyond the selection of functional form, can be found in the introduction of additional categories of variables and differences in the definition of *output*. The freedom from the need to derive the equation to be estimate from underlying engineering behavior permits variables to be employed with far less restrictions than in the two previous sets of models. This freedom also extends to the concrete empirical definition of the variables. The availability of data of course plays a large role. Link (2006) has detailed data at the road section level over a long time span. She presents two models, the second of these addresses the total renewal cost of the network. Her variable for output is traffic (ESALs), which is consistent with the two previous sets of models. In both Kalb and Deller et al., output is simply the size of the respective networks, however this must be understood as a proxy for work done to maintain the network. It is assumed that maintenance work is proportionate to the size of the network.

With respect to categories of variables in comparison to the engineering models above, the primary additional categories are: 1) prices for non-material inputs such as labor and capital and 2) variables to capture non-deterioration related cost factors. Kalb includes a number of these such as the unemployment rate, degree of exposure to urban areas, population density and average disposable income. Link's detailed data set allows her to include quantity measures of materials used. Kalb and Deller et al. do not have measures for quantities used in production.

Studies have suggested that non-deterioration related variables can affect total expense (Cook, Kazakov, & Roll, 1993), (Rouse, Putterill, & Ryan, 1997), (Kalb, 2009). In a recent study (Richmond, Kielhauser, & Adey, 2013), a series of interviews with road network managers in Switzerland led to a short-list of cost relevant factors of this type. Urbanity and 'project complexity' relating to densely populated areas headed that list.

Summarizing the state of the literature relative to developing a cost function for road renewal, one finds elements but not a consolidated whole. In the middle there is a set of models that identify the renewal interval and its consumption through traffic and weathering as the key determinants of network renewal cost. But these models are at a high level of abstraction and do not include either the choice variables of the network managers: the timing and design of repaving interventions; or important challenge factors. At the bottom-up end of the spectrum, one finds models by engineers that relate specific design choices to the renewal interval. These do not

[8] The order of magnitude is between 10^3 and 10^4 times the difference in axel loads.

include non-deterioration related cost factors and they are often specified in a way that prevents analytic solutions and thus, the derivation of demand functions in a closed form. At the top-down end of the spectrum, one finds convenient functional forms and a broad approach to capturing factors that affect cost but an absence of connection to the physical choices that road managers make.

In the next section we present a model that begins with engineering behavioral functions, even if these are simplified, and proceeds to the derivation of a cost function having an analytical form that in principle should be estimable. We intended to close the paper with an empirical example but unfortunately the combination of data and model did not lead to stable parameter estimates. Very similar r-square values resulted from very different parameter estimates. Thus we leave the example to future research and perhaps better data and a more realistic model. In addition to the demonstrating feasibility of deriving a cost function with engineering foundations for road renewal, which to the best of the author's knowledge is new, we find the specification of deterioration as a vector field through which, so to speak, the road is driven, noteworthy. This is not new, but it is also not common. Third, it is noteworthy how much simpler a model of renewal cost is, if it begins with average annual cost as a function of the renewal interval T , than if it begins with the discounted present value of the maintenance of individual road sections over an infinite horizon, or worse, a finite horizon. That is also not new, but not common, in this literature.

3 DEVELOPMENT OF THE MODEL

The Theory of Production as described in graduate level textbooks, e.g. (Varian, 1992) give the steps schematically that we will follow.

1. Define a production function and the respective cost of any possible production plan given various conditioning variables.
2. Derive demand functions that represent the minimum cost way to produce any level of output given prices and the conditioning variables.
3. Substitute the demand function back into the representation of cost so that a cost function results. That function relates prices, conditioning variables to the minimum cost of producing a given level of output.

Since it is unrealistic to attempt to define all the specific parameters required in Step 1 from engineering knowledge, a compromise is sought. That compromise is a functional form that can be estimated in exactly the same way this would be done if a translog functional form were assumed. The functional form should be specific enough to contain all the relevant categories of variables and represent their respective interactions in a way that is consistent with engineering knowledge. Choice is about trade-offs, so minimally each category should have at least two representatives in the model. What follows is essentially a proof of concept and no claim to realism of detailed model assumptions is made. The production function has been selected to be mathematically manageable while still having behavior that includes the essential physical relationships.

The central feature of the model is the time to renewal T .. It is specified without reference to the interest rate or discounting over an infinite horizon. The variables are: the time to renewal T , the road condition triggering renewal $\hat{S}$, the inputs employed $\mathbf{x}$, the amounts of use $\mathbf{u}$, various non-deterioration related cost factors $\mathbf{v}$ and various cost factors that affect the rate of deterioration $\mathbf{z}$. Each will be described in more detail shortly. The road service level is defined as the remaining amounts of service *available* for consumption by each of the means by which the road service can be consumed (usage). The *output* of road renewal activity, is modelled here as the affected change in the road service level. The only difficulty with this definition is that the road service level is not a scalar, it being the difference between to vectors. However, for the purposes of this paper, which only concerns the derivation of a functional form, this complexity can be overlooked. As discussed above, the two primary modes of use are weather and vehicle transits.

3.1 RELATION BETWEEN CONDITION STATE, DETERIORATION AND USE

The condition state of a road, $\mathbf{S}$, is in general a vector of indicators expressing important physical characteristic(s) that determines the ability of the road to provide current and future road service. The amount of

surface cracks is a key condition state measure that all road users will be familiar with. In practice, there is no universal agreement on the condition state scale, nor even on whether values increase or decrease as a road ages. Here, the condition state is modeled on a non-negative interval $[a,b]$ where $a \le S \le b$ and a is the 'as new' condition state. Worse conditions have higher values. In what follows a=0 and b=5. We also simplify the exposition by assuming that there is exactly one relevant physical characteristic that determines serviceability. We therefore use the scalar S rather than $\mathbf{S}$.

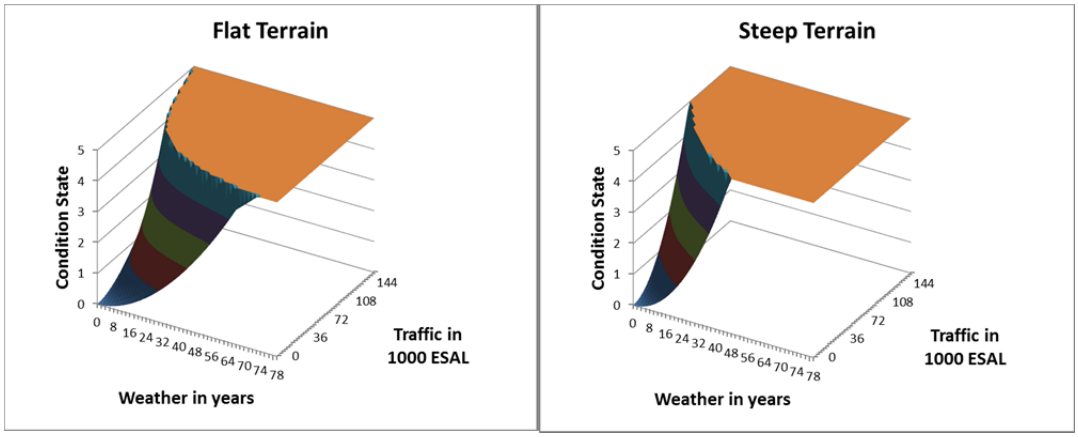

FIGURE 1 ROAD SERVICE LEVELS FOR STANDARD AND MOUNTAINOUS GEOMETRIES

As shown in Figure 1, the model of deterioration not yet fully depicted. Additionally a path of usage is required. Road service use $\mathbf{u}$ is a vector of the amounts of different modes by which a road can be ***used-up***. A specific amount, u_m, might be heavy vehicle transits. $\mathbf{u}^*(\tau)$ is a specific path of use and is depicted in Figure 2, first panel. $\mathbf{u}'(\tau)$ is the rate of usage at time τ, measured as time since the last renewal to the 'as new' state. It is assumed that the rate of usage is always non-negative and that the rate of usage for at least one type of use is always positive. Since time always passes, this is not unreasonable.

$$u_m'(\tau) \ge 0 \quad \forall m,\tau \text{ and for every } \tau^* \ \exists\, u_i'(\tau^*) \text{ s.t. } u_i'(\tau^*) > 0 \tag{1}$$

We use the notation $\breve{}$ and $\hat{}$ to indicate the beginning and endpoint of a path respectively. Since it is unusual in practice to know the time path of usage and even less likely to be able to steer it as a choice variable, assuming constant usage over a path is an empirically reasonable simplification to make. Path-independence, introduced momentarily as a maintained hypothesis, goes beyond this assumed empirical simplification. A constant usage path is designated as $\bar{\mathbf{u}}(\tau,\hat{\mathbf{u}})$ and shown in Figure 2, second panel. The rates of usage along such a path are designated by $\bar{\mathbf{u}}'(\tau,\hat{\mathbf{u}})$.

$$\bar{u}_m' = \frac{\hat{u}_m}{\hat{\tau}} \tag{2}$$

It will be assumed from here forward that the path $\mathbf{u}^*(\tau)$ has constant rates of use over time.

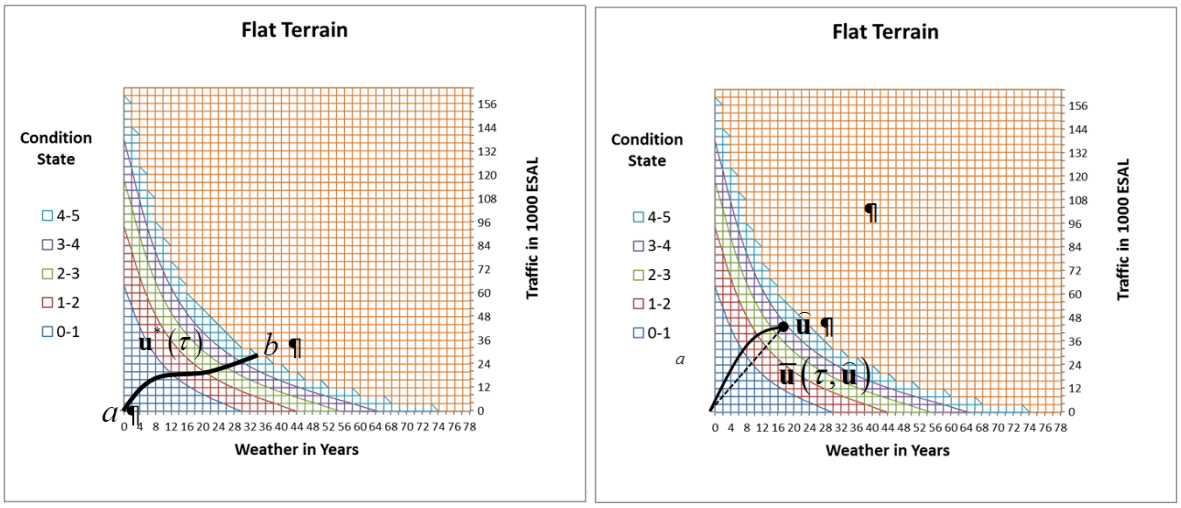

FIGURE 2: PATH OF USAGE

The physical road type is defined by a vector of physical input quantities $\mathbf{x}$ per m^2 that was used in the most recent renewal. The point of measuring it in units of purchasable input quantities rather than construction dimensions like thickness is so that cost can be represented by the sum $\sum_l p_l x_l$. The first element of the vector $\mathbf{x}$, x_0, is always 1 so that p_0 can capture fixed costs.

Challenge factors that affect the rate of deterioration are represented by a vector $\mathbf{z}$. Challenge factors that do not affect the rate of deterioration are represented by a vector $\mathbf{v}$. Examples of the former are steep terrains or soft ground. Examples of the latter are residential areas where the road space is shared with other service providers or densely populated areas where construction logistics are difficult.

Deterioration defines the change in the condition state $\Delta\mathbf{S}\left(\mathbf{u}^*(\tau), \mathbf{v}, \mathbf{x}, \mathbf{z}, S(\tau)\right)$. In order to depict the distinct impacts of different kinds of use, we follow Newbery (1988) and model deterioration as a vector field. See Geltner and Ramaswamy (1987) p. 23 for a similar modelling approach.

Corresponding to the empirical assumption that paths of usage exhibit constant rates of use, we will extend this to assume that the total change in condition state ΔS does not depend on the specific use path, but only on the total amount of use by each of the modes of use. That is, the deterioration vector field is a *conservative* vector field. Deterioration is therefore path independent. Paths having the same beginning and end points are equivalent. This assumption is unlikely to be physically accurate even if in practice it can still be a reasonable approximation. In any case, road policies that intentionally steer the relative rate of usage between different kinds of use over time are uncommon so that the potential gain from a more realistic model probably has few applications. The simplification allows the depiction of the state as a smooth, continuous surface lying over a coordinate system of amounts of use as in Figure 1.

The condition state is a function of the elapsed time since the last renewal, τ, the path which defines the rates of usage and the various conditioning variables. Obviously, operational maintenance that slows deterioration also matters but its effect is suppressed to avoid unnecessary complexity.

$$S(\mathbf{u}, \mathbf{x}, \mathbf{v}, \mathbf{z}, T) = \int_{\bar{\mathbf{u}}} \nabla S(\mathbf{u}, \tau, \mathbf{x}, \mathbf{v}, \mathbf{z}) \cdot \mathbf{u}' d\tau$$
$$\tau \in [0, T] \tag{3}$$

We assume that the condition state always declines in usage

$$\frac{\partial S}{\partial u_i} > 0 \quad \forall i \tag{4}$$

From Equation (1) therefore, the change in the condition state is always strictly positive (getting worse). Therefore, once a specific path of usage is fixed, the relationship between use and condition state becomes one to one. If the relation $S\left(\mathbf{u}(\tau)\right)$ is one to one along a given path $\mathbf{u}^*(\tau)$, then its inverse $\mathbf{u}\left(S, \mathbf{u}^*(\tau)\right)$ exists. Expressed in terms of the bound of integration, this also implies that $S(T)$ is monotonically increasing, and therefore $T(S)$ exists given an assumed path of usage. Invertability of $S(T)$ is crucial to deriving demand functions.

3.2 Required area to renew

Road network i has a total road surface area A_i. The age of the roads in the network are assumed to be uniformly distributed. Further, even within a partition of the network selected to have homogeneous exposure levels to challenge factors, the distribution of road age is assumed to be uniform. The alternative method to make models tractable is to work in expectations. For reasonably large county or state networks, the uniform distribution assumption is not unduly unrealistic.

The required area to renew, $\tilde{A}$, is defined implicitly by the assumption of a uniform distribution of age and a time to renewal T. The time to renewal itself is determined by the externally specified trigger condition state $\hat{S}$ and the deterioration function through the function $T(S)$. To maintain a steady state on a road network having a uniform distribution of age, employing road design $\mathbf{x}$ and subject to challenges $\mathbf{z}$ one must renew a fixed percentage of the network area each year. That percentage is

$$\tilde{A} = \frac{A}{T\left(\hat{S}, \mathbf{u}, \mathbf{x}, \mathbf{z}\right)} \tag{5}$$

3.3 Derivation of the Cost Function

The road network manager is assumed to solve the minimization problem (6) by choosing the road design $\mathbf{x}$ and implicitly the time to renewal, T. User costs and budget concerns enter the problem through the externally set, minimum condition state $\hat{S}^*$. At least in a Swiss context, it is not uncommon for road network managers to have explicit targets limiting the percentage of the total road network that may be less than a specified quality level, an example from the US is e.g. New Hampshire DOT (2013).

$$\begin{aligned} &\underset{\mathbf{x}}{Min}\ TC = \tilde{A}\left(\hat{\mathrm{S}}, \mathbf{u}, \mathbf{x}, \mathbf{z}, T\right) \sum_i p_i x_i \qquad \text{Given } \mathbf{p}, \mathbf{u}, \mathbf{z} \\ &\text{s.t. } \hat{\mathrm{S}}\left(\mathbf{u}, \mathbf{x}, \mathbf{z}, T\right) \geq \hat{\mathrm{S}}^* \end{aligned} \tag{6}$$

3.3.1 Assumed form of the condition state function

Our primary interest is the derivation of a cost function for a realistic deterioration function. An initial attempt based on the AASHTO road design equation as used in Santos and Ferreira (2013) was not successful but the functional form in (7) fits general intuition on how the variables should relate to deterioration and can be solved for the cost function. A goal for future research is to identify a functional form that is both mathematically tractable and reasonably realistic from a pavement engineering standpoint.

$$S = \frac{\beta_0\left(\beta_1 u_1^{\ 2} + \beta_2 u_2^{\ 2} + 2\beta_{12} u_1 u_2\right)}{\delta_1 z_1^{\ -1} + \delta_2 z_2^{\ -1} + \phi_1 x_1 + \phi_2 x_2} \tag{7}$$

The numerator of (7) defines the interaction between types of usage. Greater amounts of either mode of usage lead to worse condition states at an increasing rate. Combined usage is worse than separate usage as can be seen from the positive cross partials. In words, the effect of an additional transit on the condition state increases as total previous transits increase. However, holding previous transits constant, an increase of age also increases the damage done by an additional transit. Specifying use parametrically as a function of time, one has Equation (8) which shows that the proposed damage relationship is proportional to the square of time.

$$\begin{aligned} &\mathbf{u}(\tau) = [\alpha_1\tau, \alpha_2\tau] = [\text{Total ESAL, Total Weathering}] \\ &\beta_0\left(\beta_1\alpha_1^2 + \beta_2\alpha_2^2 + 2\beta_{12}\alpha_1\alpha_2\right)\tau^2 = \beta_0\gamma^2\tau^2 \end{aligned} \tag{8}$$

Substituting (8) into (7) one has

$$S = \frac{\beta_0}{\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1 x_1 + \phi_2 x_2}\gamma^2\tau^2 \tag{9}$$

The denominator of (7) defines the interaction between design choices **x** and challenge factors **z** as non-interactive. Relative to one another, their impacts on deterioration $\frac{\partial S}{\partial \tau}$ have no scale effects; they can be substituted for one another at constant rates. With respect to deterioration, there are diminishing marginal returns to increasing x_i and increasing marginal returns to making the environment more difficult by increasing z_i .All parameters γ, δ_i, ϕ_i are assumed to be positive so that the signs of the derivatives are evident from the signs in front of the derivatives.

$$S' = \frac{\partial S}{\partial \tau} = \frac{2\tau\gamma^2}{\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1 x_1 + \phi_2 x_2\right)} \tag{10}$$

$$^{*}. \tag{11}$$

3.3.2 SOLVING FOR THE TIME TO RENEWAL BY INVERTING THE CONDITION STATE FUNCTION

The condition state at the moment of renewal is

$$\hat{S} = \frac{\beta_0}{\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1 x_1 + \phi_2 x_2}\gamma^2 T^2 \tag{12}$$

Which can be inverted to arrive at $T\left(\hat{S}\right)$.

$$T = \left(\frac{\hat{S}}{\beta_0\gamma^2}\right)^{\frac{1}{2}}\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1 x_1 + \phi_2 x_2\right)^{\frac{1}{2}} \tag{13}$$

The first factor is dependent on the externally specified trigger condition state and rate of usage, one can replace it with $1/\rho_0$ for notational convenience and think of ρ_0 as a reference portion of the network to renew.

$$T = \frac{1}{\rho_0}\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1 x_1 + \phi_2 x_2\right)^{\frac{1}{2}} \tag{14}$$

3.3.3 DERIVING DEMAND FUNCTIONS FOR MATERIAL INPUTS

Substitute (18) into (6) to arrive at total cost which is to be minimized.

$$TC = \tilde{A}\sum_{l} p_i x_i \; = \left(\frac{A\rho_0\left(\mathbf{u}, \hat{S}\right)}{\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1 x_1 + \phi_2 x_2\right)^{\frac{1}{2}}} \right) \sum_{l} p_i x_i \tag{15}$$

The FONC with respect to x_1 is

$$\frac{\partial TC}{\partial x_1} = \tilde{A}\left(\frac{\phi_1}{-2\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1 x_1 + \phi_2 x_2\right)} \sum_{i} p_i x_i + p_1 \right) \tag{16}$$

Setting this equal to zero and assuming an internal solution gives, after a few simplifications, a demand function that is still conditional on x_2.

$$x_1\left(\mathbf{p}, \mathbf{z}, \mathbf{x}_2\right) = \sum_{i \neq 1} \frac{p_i}{p_1} x_i - \frac{2\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_2 x_2\right)}{\phi_1} \tag{17}$$

By symmetry one also has

$$x_2\left(\mathbf{p}, \mathbf{z}, \mathbf{x}_1\right) = \sum_{i \neq 2} \frac{p_i}{p_2} x_i - \frac{2\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1 x_1\right)}{\phi_2} \tag{18}$$

Substituting for x_2 and setting $x_0 = 1$, one can derive, after a few simplification steps[9], the demand function $\hat{x}_1(\mathbf{p}, \mathbf{z})$ for x_1 dependent on prices and challenge factors only.

$$\hat{x}_1\left(\mathbf{p}, \mathbf{z}\right) = \frac{p_0\phi_2 - p_2\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1}\right)}{\left(p_2\phi_1 - p_1\phi_2\right)} \tag{19}$$

By symmetry we also have

$$\hat{x}_2\left(\mathbf{p}, \mathbf{z}\right) = \frac{p_0\phi_1 - p_1\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1}\right)}{\left(p_2\phi_1 - p_1\phi_2\right)} \tag{20}$$

3.3.4 DERIVING THE COST FUNCTION BY SUBSTITUTING DEMAND FUNCTIONS INTO THE COST MINIMIZATION PROBLEM

The cost function is found by substituting the input demand functions (19) and (20) into Equation (15). This of course replaces the road manager with an optimization algorithm that selects input materials – and hence the interval of renewal – given the levels of variables that the road manager does not control so as to minimize the resulting annual costs. In other words, Equation (21) is a benchmark for the cost of road renewal across networks having different external conditions.

[9] All derivations are available from the authors on request.

$$TC(\mathbf{p},\mathbf{z}) = \left(\frac{A\rho_0(\mathbf{x}_0,\mathbf{z}_0,\mathbf{u},\bar{S})}{\left(\delta_1 z_1^{-1} + \delta_2 z_2^{-1} + \phi_1\left(\frac{p_0\phi_2 - p_2(\delta_1 z_1^{-1} + \delta_2 z_2^{-1})}{(p_2\phi_1 - p_1\phi_2)}\right) + \phi_2\left(\frac{p_0\phi_1 - p_1(\delta_1 z_1^{-1} + \delta_2 z_2^{-1})}{(p_2\phi_1 - p_1\phi_2)}\right)\right)^{\frac{1}{2}}} \right) \left(p_0 + p_1\left(\frac{p_0\phi_2 - p_2(\delta_1 z_1^{-1} + \delta_2 z_2^{-1})}{(p_2\phi_1 - p_1\phi_2)}\right) + p_2\left(\frac{p_0\phi_1 - p_1(\delta_1 z_1^{-1} + \delta_2 z_2^{-1})}{(p_2\phi_1 - p_1\phi_2)}\right)\right) \quad (21)$$

Thankfully, Equation (21) simplifies to a much more concise form. If done by hand, it is quite involved, but with symbolic math software it is quite simple. The simplified form is

$$\left(\frac{TC}{A}\right)^2 = \frac{2\rho_0^2 p_2^2}{\phi_2(p_2\phi_1 - p_1\phi_2)}\left(p_0\phi_1 - p_1\delta_1 z_1^{-1} - p_1\delta_2 z_2^{-1}\right) \quad (22)$$

3.3.5 ADDING IN VARIABLES TO CAPTURE NON-DETERIORATION RELATED CHALLENGE FACTORS

Equation (22) already has a dependency on **u** and $\bar{S}$ through ρ_0^2. The dependency on **z** and **p** is evident. What is missing are the non-deterioration related variables, **v**. Looking back to Equation (15), one can imagine building the exact same road under the same deterioration affecting environmental conditions but subject to other conditions that affect cost. Carrying out a project in the tight conditions of an urban location versus a non-urban location would be a typical example. Others would be the interaction with tramways that share the road space or the requirement to build sidewalks and rain runoff drainage systems.

If one models these as additively separable mark-ups on project costs where the degree of exposure is either 0 or 1, then one can adapt (15) as follows. Let v_j be the percent of the network exposed to the factor j and κ_j be the mark-up. Assuming as before that the age of the roads are spread uniformly relative to challenge factors, then each year one will face $v_j\tilde{A}$ meters squared of challenge of type j. One can write the extended model of total cost as

$$TC = \tilde{A}\left(\sum_i p_i x_i + \sum_h v_h\kappa_h\right) \quad (23)$$

Since .. is independent of **z** and **x** one can see that $\sum_h v_h\kappa_h$ will behave as a fixed mark-up on average per-unit cost. That is, one could redefine p_0 – which already represents fixed per-unit cost – to include this additional mark-up and nothing in the previous derivations would change. Therefore Equation (22) becomes

$$\left(\frac{TC}{A}\right)^2 = \frac{2\rho_0^2 p_2^2}{\phi_2(p_2\phi_1 - p_1\phi_2)}\left(p_0\phi_1 - p_1\delta_1 z_1^{-1} - p_1\delta_2 z_2^{-1} + \phi_1\sum_h \kappa_h v_h\right) \quad (24)$$

Reintroducing **u** and $\bar{S}$ is accomplished by substituting $\rho_0 = \left(\frac{\beta_0\gamma^2}{\bar{S}}\right)^{\frac{1}{2}}$ and $\gamma^2 = \beta_1 u_1^{\,2} + \beta_2 u_2^{\,2} + 2\beta_{12}u_1u_2$ into (24) to get

$$\left(\frac{TC}{A}\right)^2 = \left(\frac{2\beta_0}{\bar{S}}\frac{p_2^2}{\phi_2(p_2\phi_1 - p_1\phi_2)}\right)\left(\beta_1 u_1^{\,2} + \beta_2 u_2^{\,2} + 2\beta_{12}u_1u_2\right)\left(p_0\phi_1 - p_1\delta_1 z_1^{-1} - p_1\delta_2 z_2^{-1} + \phi_1\sum_h \kappa_h v_h\right) \quad (25)$$

This is the cost function to be estimated. It is a cost function that results if one applies standard economic methods to functional relationships that intentionally approximate physical realities of road renewal, even if the

approximation is admittedly highly stylized. A person who wishes to defend, criticize or improve the realism has something concrete to work with. That would not be the case if the functional form were simply assumed.

It is also evident that (25) is not similar to a translog function. Categories of variable relate to one another in specific and distinguishable ways. Use variables have an interaction amongst themselves that is different from their interaction to deterioration related and non-deterioration related challenge factors. Deterioration related variables have a relationship with relative prices of inputs. That is not true for non-deterioration related cost factors. Perhaps most important is the multiplicative structure between the two terms in parenthesis. Even Equation (15) has the basic structure of a term representing the cost of a unit of renewal and a term representing how often that cost will have to be paid. That seems to be a fundamental trade-off in road renewal.

4 CONCLUSIONS

The goal of this research was to find a way to benchmark the renewal activity of road network managers. This is difficult to do because road network managers operate under unique circumstances and each network area has only one supplier. Thus, no natural comparators exist. An alternative route is to develop a calculated benchmark, a cost function. However, to be convincing, it should be possible to demonstrate that the proposed cost function would in fact represent the cost of maintaining a network, in some hypothetical world. To the best knowledge of the authors, such a cost function does not exist in the literature, even if elements needed to build one do exist.

It has been possible to derive a cost function for a hypothetical network, one that has a uniform distribution of age of road sections. The challenge factors are limited in number to two in each of four categories of variable. The four categories cover many possible types of interaction that an external factor could have with road renewal cost. The effect of 'use' variables on deterioration as well as the mitigating effects of stronger design or the enhancing effects of difficult environmental conditions are included in the model in specific and distinguishable ways. No claim is made that the assumed functional relationships determining deterioration are realistic. The fact that they are present and can be challenged concretely is key.

Certain important mathematical features have been shown to be required of the functional forms to represent deterioration. These would not normally be apparent to the engineers who understand the physical processes best, nor required of their models of deterioration. Hence intentional cooperation with pavement engineers is needed to find appropriate functional forms. The relationship between time and condition state must be invertible. It must be possible to solve for demand equations. Re-substitution into the cost minimization objective is no doubt always feasible, but the resulting cost function should, if possible, not be unduly complex. From an empirical point of view, two functional forms may fit the data equally well and yet be very different in their mathematical tractability.

Future research is needed to develop empirically realistic functional forms for road deterioration behavior with respect to the various challenge factors that also fulfil the mathematical requirements on tractability. The current model does not specifically address aggregation topics and this extension is a prerequisite to calculating empirical benchmarks. Finally, empirical estimation of the model's parameters are needed. Another interesting direction of research relates to the bias that would result if a relationship such as presented in Equation (25) were estimated with functional forms typically used in econometric studies of cost functions.

5 REFRENCES

AASHTO. (1993). *Guide for design of pavement structures* (4 ed.). Washington D.C.: American Association of State Highway and Transportation Officials.

Cook, W. W., Kazakov, A., & Roll, Y. (1993). On the Measurement and Monitoring of Relative Efficiency of Highway Mainenance Patrols. In A. Charnes, W. W. Cooper, A. Y. Lewin & L. M. Seiford (Eds.), *Data Envelopment Analysis: Theory, Methodoogy and Application*. London: Kluwer Academic Publishers.

Deller, S. C., Chicoine, D. L., & Walzer, N. (1988). Economies of Size and Scope in Rural Low-Volume Roads. *The Review of Economics and Statistics, 70*(3), 459-465. doi: 10.2307/1926784

Geltner, D., & Ramaswamy, R. (1987). Economic efficiency implications of optimal highway maintenance policies for private vs. public highway owners. *Transportation Research Record, 1116*, 9.

Highway Research Board. (1962). The AASHTO Road Test: Pavement Research *HRB Special Report* (Vol. 61E). Washington DC.

Kalb, A. (2009). What Determines Local Governments' Technical Efficiency? The Case of Road Maintenance *Discussion Paper* (Vol. 09, pp. 34).

Lindberg, G. (2002). Deliverable 10: Infrastructure Cost Case Studies, Annex A2: Marginal

cost of road maintenance for heavy goods vehicles on Swedish roads.

Link, H. (2006). An econometric analysis of motorway renewal costs in Germany. *Transportation Research Part A: Policy and Practice, 40*(1), 19-34. doi: DOI: 10.1016/j.tra.2005.04.005

Link, H. (2014). A Cost Function Approach for Measuring the Marginal Cost of Road Maintenance. *Journal of Transport Economics and Policy, 48*(1), 18.

New Hampshire DOT. (2013). Performance 2012, State Highway Pavement in Good or Fair Condition (pp. 2) Available at http://www.nh.gov/dot/org/commissioner/balanced-scorecard/department/documents/2012_bs_performance_pavecond.pdf..

Newbery, D. M. (1988). Road Damage Externalities and Road User Charges. *Econometrica, 56*(2), 295-316.

Ouyang, Y., & Madanat, S. (2006). An analytical solution for the finite-horizon pavement resurfacing planning problem. *Transportation Research Part B: Methodological, 40*(9), 767-778. doi: http://dx.doi.org/10.1016/j.trb.2005.11.001

Richmond, C., Kielhauser, C., & Adey, B. (2013). KPI - Key Performance Indicators for Cantonal Road Management (pp. 160). Zurich, Switzerland: Eidgenösssischer Technischer Hochschule, IBI.

Rouse, P., Putterill, M., & Ryan, D. (1997). Towards a General Managerial Framework for Performance Measurement: A Comprehensive Highway Maintenance Application. *Journal of Productivity Analysis, 8*(2), 127-149. doi: 10.1023/a:1007743606303

Santos, J., & Ferreira, A. (2013). Life-cycle cost analysis system for pavement management at project level. *International Journal of Pavement Engineering, 14*(1), 71-84. doi: 10.1080/10298436.2011.618535

Sathaye, N., & Madanat, S. (2012). A bottom-up optimal pavement resurfacing solution approach for large-scale networks. *Transportation Research Part B: Methodological, 46*(4), 520-528. doi: http://dx.doi.org/10.1016/j.trb.2011.12.001

Varian, H. R. (1992). *Microeconomic Analysis* (Third ed.). New York: W.W. Norton & Company.