Estimation of Highway Pavement Deterioration from In-Service Pavement Data

ROHIT RAMASWAMY AND MOSHE BEN-AKIVA

An accurate deterioration prediction model should include, as completely as possible, all the factors that affect deterioration and the effects of past maintenance. Many of the deterioration models that currently exist in the literature are simple models that include only a few explanatory variables, and most do not incorporate the effects of maintenance. The literature has reported that the performance of many of these models has been poor. Even in cases where more complex models have been formulated, researchers have reported that the estimates of important explanatory variables have counterintuitive signs. The usefulness of these models in the planning of future maintenance is therefore questionable. The reason for the unexpected signs of the parameters is that models use data collected from in-service pavements. Maintenance on in-service pavements is not carried out by highway agencies exogenously; instead, it is carried out in response to factors such as condition and traffic. This implies that the condition of a highway pavement at any point in time is determined by the point of intersection of two simultaneous processes: (*a*) the process of deterioration by which the pavement quality degrades and (*b*) the behavior of the agency performing maintenance in response to this deterioration. An appropriate model for predicting condition over time is therefore a simultaneous equation specification that reflects both the processes described above. In this paper, a simultaneous-equation deterioration model is formulated for highway pavements and estimated from data collected in the field. On comparison with a traditional single-equation model, it is seen that the simultaneous-equation approach produces realistic parameter estimates.

Over the last 30 years, many attempts have been made to develop models that predict the deterioration of a highway pavement over time. These models express the condition or performance of a pavement as a function of explanatory variables, such as maintenance activities performed, traffic loads, age, and environmental variables. Many different models of this sort have been estimated both by researchers and by highway agencies (*1–5*). These models are used to predict the deterioration of a pavement under many different maintenance scenarios, and can therefore be used by highway agencies for the planning of maintenance strategies.

Maintenance planning for highway agencies involves the allocation of limited resources for maintenance and rehabilitation to a number of pavements that make up the highway network. Since maintenance planning depends upon the ability of the highway agency to accurately predict future pavement condition, it is very important to be able to specify and estimate an accurate deterioration prediction model. Failure to do so will result in incorrect choices of maintenance strategies and consequently inefficient utilization of resources. From the above discussion, it is clear that a deterioration prediction model should (*a*) include, as completely as possible, all the factors that affect deterioration of the pavement and (*b*) accurately represent the effect of maintenance on pavement condition.

The factors that affect the deterioration of a highway pavement can be categorized as follows:

1. Pavement characteristics: pavement strength, layer thicknesses, base type, surface type.
2. Pavement history: time since last rehabilitation, total pavement age.
3. Traffic characteristics: average daily traffic, cumulative traffic, traffic mix (percentage of trucks).
4. Environmental variables: average monthly precipitation, number of freeze-thaw cycles, average annual minimum temperature, and so on.

The maintenance information that is required for the estimation of deterioration prediction models is the extent of maintenance activities, by activity, for each year in which the pavement has been in operation. For example, this information can be measured in terms of square yards of crack filling, seal coating, and so on, per section or in terms of the amount of labor or materials used during the performance of the maintenance activities.

Many of the deterioration models estimated in the literature do not satisfy the two requirements laid out in the preceding paragraphs. Many of the models (*1,2*) do not use more than a few explanatory variables, usually just age or traffic. This is because of the difficulties associated with the measurement of environmental effects and other intangibles associated with the deterioration of the facility, for example, the quality of initial construction. Also, maintenance information has not been readily available in the past, and so few models have included the effects of maintenance on deterioration.

As a result of these incomplete specifications, many researchers have reported that deterioration prediction models have poor fit to data and do not produce accurate predictions of future condition. Over the past few years, studies have been conducted specifying more complete models that include the effects of maintenance as explanatory variables (*5,6*). Unfortunately, these models have not reportedly performed any better than the previous models. Researchers such as

---

R. Ramaswamy, AT&T Bell Laboratories, Room 3K-312, Holmdel, N.J. 07733. M. Ben-Akiva, Department of Civil Engineering, Room 1-181, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

Butler et al. (*6*) have reported that the parameter estimates of important explanatory variables have counterintuitive signs.

The poor performance of even the more realistic specifications indicates that there is a problem with the traditional approaches used to model the deterioration of highway pavements. This problem pertains to the way in which the effect of maintenance is specified in the existing models. All the models in the literature that include the effects of maintenance have collected data from in-service pavements. In these pavements, maintenance is performed in response to the condition of the pavement. In other words, more maintenance has been performed over time on pavements that are in a worse condition or that deteriorate faster. Typically, such pavements tend to carry higher traffic loads. Because of increased maintenance on high-traffic pavements, it is conceivable that higher-traffic pavements are in a better condition than lower-traffic pavements. If a deterioration model is estimated with condition as the dependent variable and traffic as an independent variable, such a situation will clearly produce a counterintuitive sign for the parameter of the traffic variable.

A proper specification of models that use maintenance information from in-service pavements is therefore to define maintenance as an endogenous variable that is affected by pavement condition, traffic, and other explanatory variables. The deterioration model, then, is specified as a system of simultaneous equations, with one equation representing condition as a function of maintenance and other explanatory variables affecting deterioration and a second equation expressing the extent of maintenance performed as a function of condition and traffic.

In this paper, the specification and estimation of a simultaneous-equation system for modeling highway pavement deterioration are described. The following section describes the specification of the model in greater detail and discusses identification issues. In the section on the case study, the application of the model to data collected in the field is described.

SIMULTANEOUS-EQUATION MODEL SYSTEM

Specification of a Pavement Deterioration Model

Consider a pavement that is not in service but has been chosen as a test section for developing pavement deterioration models. If a controlled experiment is performed on the test section where the pavement condition is monitored by running different traffic loads on pavements maintained to different levels, the condition in some time period can be exactly specified in terms of the events of the past and the future deterioration of the pavement can be predicted from a time-series model. The maintenance specification in this experiment is exogenous; that is, the activities and extents of maintenance to be performed are prespecified and do not depend on the condition of the pavement. In such a case, a single-equation deterioration model is entirely appropriate for predicting the condition of the pavement in a given time period.

There is clearly no natural or obvious manner in which to define the condition of the pavement in a given time period. The condition, however, is correlated with the extent of damage on the pavement. In other words, the higher the extent of damage, the worse is the condition of the pavement. Damage data are usually available on the extents of several damage types, such as alligator cracking, linear cracking, patching, and potholes. Traditionally, the models that have been estimated follow two different approaches. The first approach is to estimate an independent deterioration model for each of several damage types. For example, such a model system has a model that predicts the propagation of cracking as a function of explanatory variables and another model that predicts the propagation of potholes, and so on. Typically, data on a large number of damage types are collected by different highway agencies, and so the number of potential models of this kind is large. From among the large number of candidate damage measurements, a few important damage types are selected, and models that predict the propagation of these damage types are used to plan maintenance strategies. The basis on which the important damage types are chosen is developed from engineering experience (*5,6*). The second approach derives aggregate measures of pavement performance from observations on different damage types using a previously calibrated performance index formula that determines the weights to be allotted to each damage type used in the calculation of the performance index. Examples of such measures are the two most widely used indices of pavement performance: the present serviceability index, or the PSI (*7*), and the pavement condition index, or the PCI (*4*).

The PSI is calculated for flexible pavements from observations on rut depth (RD, calculated in inches), slope variance (SV), and the extents of cracking (C), and patching (P) (in square feet per 1,000 ft^2) using the following formula:

$$\begin{aligned} \mathrm{PSI} = {} & 5.03 - 1.91 \log_{10} (1 + \mathrm{SV}) \\ & - 1.38\ \mathrm{RD}^2 - 0.01(C + P)^{0.5} \end{aligned} \tag{1}$$

The PSI was developed as a result of the AASHO Road Test in 1958–1960 (*7*) as part of a study to tie serviceability (i.e., the user's perspective of the quality of the pavement) to measurable surface damage factors. A group of evaluators was asked to provide their opinions on the quality of various ride surfaces and rate their opinion on an integer scale from 0 to 5, with 5 representing a newly resurfaced or constructed pavement and 0 representing a completely disintegrated surface. The terminology used for this rating was the "present serviceability rating" (PSR), and the average of the evaluations by each rater was taken to be the PSR of a particular pavement. This rating was intended to simulate users' points of view and was performed according to some ground rules discussed in a previous work (*8*). The PSR was then regressed against damage variables transformed in different ways and the fitted value of the PSR for each pavement was termed the present serviceability index, or the PSI. Obviously, the PSI is also on a scale from 0 to 5. Subsequent to the calibration, the PSI can be calculated from future observations of damage simply from Equation 1.

The PCI is calculated in a similar manner as a combination of 19 damage types. A new pavement with no damage has a PCI of 100. Each damage component contributes a certain number of "deduct points," which is the number of PSI points to be deducted from 100 as a result of the damage. The deduct points increase as the extent of damage increases. The appropriate number of deduct points, depending on the particular damage types present, is calculated from calibration graphs developed by Shahin and Kohn (*4*).

Once the PCI or the PSI has been calculated, a deterioration model can be estimated by regressing the PCI or the PSI against a set of explanatory variables [see, for example, previous methods (*1–3*)]. The deterioration model then predicts the progression of the performance index as a result of changes in the explanatory variables.

In the case of the test pavement section, the deterioration model can be specified as follows for a stationary situation, where traffic and maintenance levels are stable over time:

$$S = f(A, X_1) + \eta_1 \tag{2}$$

where

S = condition in any time period, measured in terms of either the extent of a particular damage type or an aggregate index such as the PCI or the PSI;
A = average extent of maintenance performed per time period;
X_1 = vector of exogenously determined explanatory variables affecting pavement condition, for example, traffic, age, environmental factors, etc.; and
η_1 = error term accounting for unobserved explanatory variables and random effects.

The case of a test section is not typical of the manner in which data are collected for deterioration modeling. Most of the data used for the estimation of deterioration models are collected on in-service pavements. For such pavements, as mentioned before, maintenance activities are performed in response to the condition of the pavement as a result of previous condition assessments. Thus, historical maintenance activities must be treated as endogenous variables, and every data point obtained from in-service pavements represents the outcome of two processes taking place simultaneously over time. The first process is the process of deterioration as a result of traffic loads, environmental factors, and past maintenance. The second process is maintenance that is performed over the years by the highway agency and is a function of past deterioration and the factors that might cause future deterioration, for example, traffic loads. These processes can be summarized by the following addition to Equation 2:

$$S = f(A, X_1) + \eta_1 \tag{3a}$$

$$A = g(S, X_2) + \eta_2 \tag{3b}$$

where X_2 is a vector of exogenous explanatory variables for the maintenance equation, and the other variables are the same as those in Equation 2. η_2 is the measurement equation error term.

The system of Equations 3 is a simultaneous system, since the condition S depends upon the extent of maintenance A, and the amount of maintenance performed depends upon the condition. Equation (3a) represents the effect that the explanatory variables affecting deterioration have on pavement condition, and can be called the deterioration model. Equation 3b represents the maintenance performed by the agency in response to the condition and traffic, and can be called the maintenance model or the agency behavior model. This system is represented in Figure 1, where links 1 and 3 represent the deterioration model and links 2 and 4 represent the maintenance model. It is not possible to estimate both equations in the system (3) unless both equations are identified. Consider the system in Equation 3 for the case where X_1 and X_2 are identical vectors. In order for the deterioration equation to be identified, at least one independent variable that is contained in the maintenance equation should be omitted from the deterioration equation. If X_1 and X_2 have identical elements, then both equations are not identified. If X_1 should contain some variables that are not found in X_2, then the maintenance equation is identified, but not the deterioration equation. If a single-equation model such as the one specified in Equation 2 is estimated when the deterioration equation is not identified, then in actuality the maintenance model or some uninterpretable combination of the deterioration and maintenance equations is being estimated rather than the deterioration model. Therefore, it is not possible to predict the signs of the parameters, and counterintuitive signs for some of the parameters of the model could be obtained.

Identification of a Simultaneous-Equation Model

In this section, the identification issues mentioned in the previous section are discussed in greater detail. Consider a linear specification of Equation 3, where it is assumed that the vector of explanatory variables X_1 for the deterioration equation has only two elements, average daily traffic (ADT) and age, and that the explanatory variable vector X_2 for the maintenance equation has a single element, age. Also assume that the vector A measures the extent of a single activity, crack filling. Equations 3a and 3b are then a system of two simultaneous equations that can be written as shown below:

$$S = \beta_1 A + \beta_{21}\,\text{ADT} + \beta_{22}\,\text{Age} + \eta_1 \tag{4a}$$

$$A = \gamma_1 S + \Gamma_1\,\text{Age} + \eta_2 \tag{4b}$$

The reduced form of Equations 4, which expresses the endogenous variables in terms of the exogenous variables, can be obtained by substituting for A and S, respectively, in the right-hand sides of Equations 4a and 4b. The reduced-form equations can be written as follows:

$$(1 - \beta_1\gamma_1)S = (\beta_1\Gamma_1 + \beta_{22})\,\text{Age} + \beta_{21}\,\text{ADT} + (\eta_1 + \beta_1\eta_2)$$

$$(1 - \beta_1\gamma_1)A = (\Gamma_1 + \gamma_1\beta_{22})\,\text{Age} + (\gamma_1\beta_{21})\,\text{ADT} + (\gamma_1\eta_1 + \eta_2)$$

Equations 5a and 5b can be rewritten as follows:

$$S = \Pi_{11}\,\text{Age} + \Pi_{12}\,\text{ADT} + \mu_1 \tag{6a}$$

$$A = \Pi_{21}\,\text{Age} + \Pi_{22}\,\text{ADT} + \mu_2 \tag{6b}$$

where $\Pi_{11} = (\beta_1\Gamma_1 + \beta_{22})/(1 - \beta_1\gamma_1)$ and the other composite terms are similarly defined.

Also,

$$\eta_1 = \mu_1 - \beta_1\mu_2 \tag{6c}$$

$$\eta_2 = \mu_2 - \gamma_1\mu_1 \tag{6d}$$

If Equations 6a and 6b are estimated separately using ordinary least squares, four parameter estimates, for Π_{11}, Π_{12}, Π_{21}, and Π_{22}, are obtained. These four parameters are expressed in terms of the five parameters of Equations 4, that is, β_{21}, β_{22}, β_1, γ_1, and Γ_1. Therefore, it is not possible to estimate all the parameters of Equation 4.

The fact that the necessary condition is not satisfied does not mean that none of the parameters are estimatable. The parameters that can be estimated can be derived by substituting Equations 6 into Equations 4. The following equations are then obtained:

$$\Pi_{11} = \beta_{22} + \beta_1\Pi_{21} \tag{7a}$$

$$\Pi_{12} = \beta_{21} + \beta_1\Pi_{22} \tag{7b}$$

$$\Pi_{21} = \Gamma_1 + \gamma_1\Pi_{11} \tag{7c}$$

$$\Pi_{22} = \gamma_1\Pi_{12} \tag{7d}$$

From Equation 7d, $\gamma_1 = \Pi_{22}/\Pi_{12}$ and $\Gamma_1 = \Pi_{21} - \gamma_1\Pi_{11}$. The two parameters for the maintenance equation can therefore be solved for. However, only two equations exist for the three parameters of the deterioration Equation 4a, and so these parameters are not identified. The maintenance equation is identified and the deterioration equation is not.

The identification problem is shown in Figure 2. Because the reduced-form Equations 6 are obtained by solving for the endogenous variables in Equation System 4, these equations simply describe the point of intersection of the deterioration Equation 4a and the maintenance Equation 4b. Each observation on a pavement provides one such intersection point. The objective of the estimation procedure is to determine the slopes of the maintenance and the deterioration equations from these intersection points. In general, it is not possible to uniquely identify the slope of two lines simply from their intersection points, because there is an infinite number of lines that intersect at any given point. However, in the system of Equation 4, there is enough information about the pattern of intersection points that allows the determination of one equation. In Equation 4a, pavement condition depends upon traffic and age. If there is enough variability in the data, then different intersection points between the deterioration and maintenance equations are observed for different values of age. For each value of age, therefore, a deterioration equation can be plotted as shown in Figure 2, where the lines labeled 1 are the deterioration equation and the line labeled 2 is the maintenance model. The maintenance equation can therefore

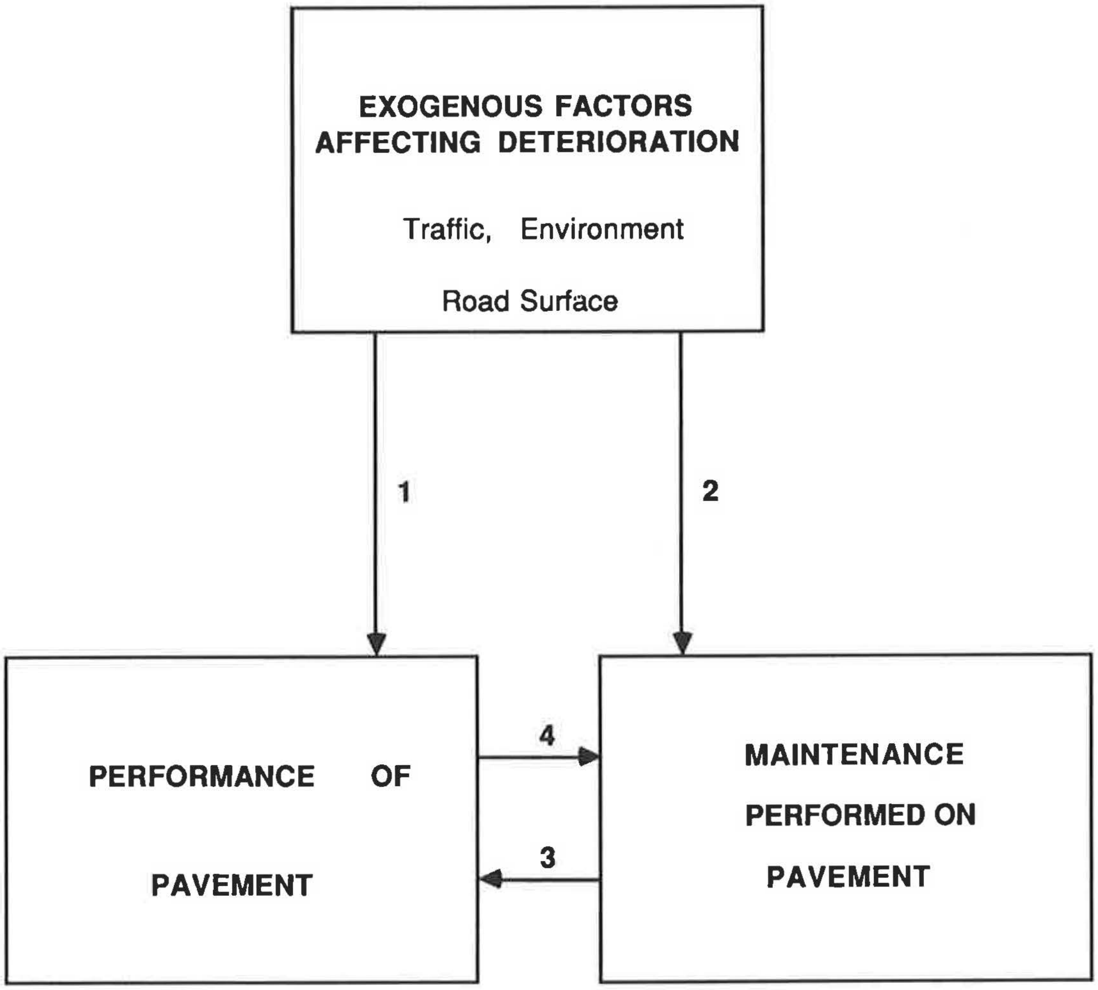

FIGURE 1 Simultaneity of deterioration and maintenance.

be traced out from these different deterioration lines as shown in Figure 2. Because there is no explanatory variable on which a similar exercise can be carried out with the maintenance equation, the deterioration equation cannot be identified in this fashion. In order for both equations to be identified, each equation in the simultaneous-equation system should contain at least one explanatory variable not found in the other equations as mentioned in the previous section.

In order to identify the deterioration model as well, one additional restriction is required in the model system that will generate one more equation. This restriction can be obtained by assuming that the error terms of the structural and the maintenance equations η_1, η_2 are uncorrelated. The physical processes that cause deterioration and trigger maintenance activities are different, so there is no reason why the two error terms should be correlated. This assumption implies that $E(\eta_1\eta_2)$ is zero, where $E(.)$ stands for the expected value. Making this assumption produces the following additional equation:

$$E(\eta_1\eta_2) = E(\mu_1 - \beta_1\mu_2)(\mu_2 - \gamma_1\mu_1)$$

from Equations 6c and 6d. Therefore, $E(\eta_1\eta_2) = 0$ implies:

$$(1 + \beta_1\gamma_1)\mu_{12} - \gamma_1\mu_{11} - \beta_1\mu_{22} = 0 \qquad (7e)$$

where $\mu_{12} = E(\mu_1\mu_2)$, $\mu_{11} = E(\mu_1^2)$, and $\mu_{22} = E(\mu_2^2)$. These values can be obtained from the residuals after ordinary least-squares (OLS) estimation of Equations 6 is performed. Equations 7a, 7b, 7c, 7d, and 7e together form a system of five equations from which the five unknown parameters can be

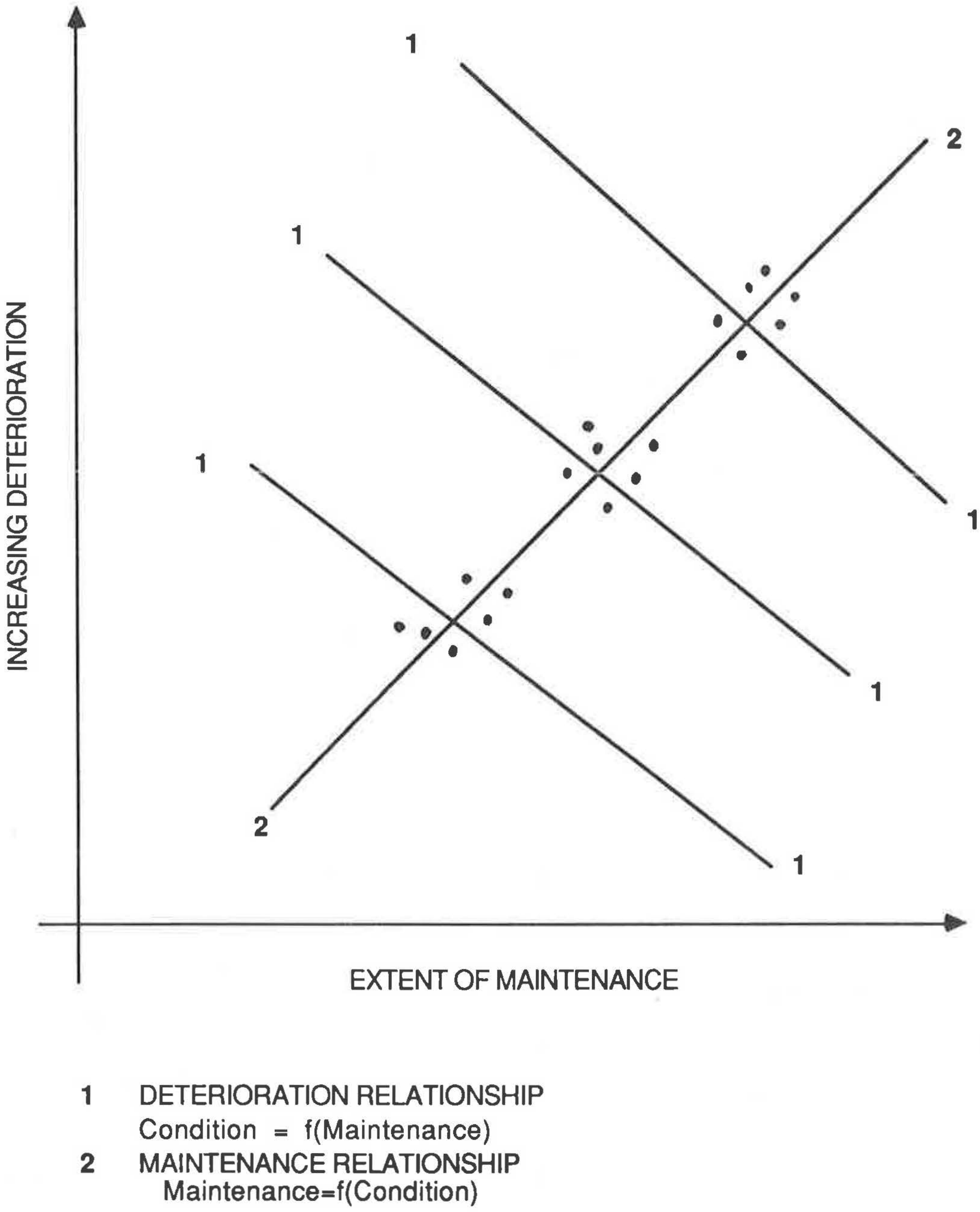

FIGURE 2 Relation between deterioration and agency behavior.

estimated. The system is now exactly identified.

From the above discussion on identification, the following points can be summarized:

1. If the deterioration model is specified as a simultaneous-equation system, both equations need to be identified in order to estimate all parameters of the system.
2. Estimation of a single unidentified equation when a simultaneous-equation system is appropriate results in unpredictable signs of the parameter estimates, because it is not clear which equation is being estimated.
3. In order for an equation in a simultaneous-equation system to be identified with respect to another, it is necessary that the equation omit at least one explanatory variable that is contained in the other equation.
4. Additional restrictions, such as assumed independence of error terms, are sometimes required to identify all equations in a simultaneous-equation system.

With these points in mind, the results of this section will now be applied to a case study using data collected in the field. First, a single-equation model similar to a traditional deterioration model found in the literature is specified. Then a set of maintenance equations is added to the deterioration specification, and all the equations are estimated simultaneously. It is observed that the problems with incorrect signs on the parameters are solved when the simultaneous-equation specification is estimated.

CASE STUDY

Description of Data

The data set used for the case study was obtained from the state of Nevada. The unit of observation is characterized as a "rural, flexible, full unit mile pavement section." The information contained in the data set for each year and each pavement section can be summarized as follows:

1. *Damage measurements*: square feet of alligator and linear cracking, widths of alligator and linear cracking, square feet of patching, and average rut depth and slope variance.
2. *Environment and traffic variables*: ADT, average daily truck traffic, age of section since last rehabilitation, minimum temperature, annual precipitation, number of freeze/thaw cycles, and number of wet days.
3. *Maintenance activities*: extent of maintenance performed on 10 activities, the number of labor hours, and the total material cost.

The data were collected for 4,871 pavement sections over a period of 5 years from 1980 through 1984. Each year several teams of inspectors drove the entire system, evaluating each mile for damage. The procedure was to examine the first two-thirds of the mile and then select a typical 1,000-ft^2 section in the last third of the mile to evaluate in detail. The condition of these 1,000 ft^2 is taken as representative of the condition of the section as a whole. These data were originally used by ARE, Inc. for the specification and estimation of deterioration models. The results of this research have been reported elsewhere (*6*).

The Nevada data set was chosen because it is one of the few available data sets that has substantial maintenance information, which is critical for the development of models predicting deterioration. In the analysis performed by ARE, Inc. with the data, four preventive maintenance activities were isolated from among the 10 activities available in the data set. These were the activities that were performed regularly on the pavement and in amounts exceeding more than a few square yards.

The obvious deterioration model that can be specified with time-series data links the performance of the facility at some time *t* to a sequence of all events that occurred in the past. For example, the condition of a highway pavement at time *t* can be linked to the traffic in past years, to the maintenance performed, and to the condition before and after maintenance in the past. Such a specification would entail the definition of a latent variable for each performance characteristic for each year, and indicators of the performance characteristics for each year. However, in order to estimate such a model it is necessary to know the exact sequence of past events, such as at what point in the year the maintenance was performed, what the condition was before and after maintenance, and so on. At a minimum, for each year in which a routine maintenance activity is performed, it would be necessary to know what was done since the inspection or the maintenance activity. It would make sense to assume that the annual inspection always precedes the maintenance activity for the year.

However, the elapsed time between the most recent maintenance activity and the subsequent inspection can only be estimated grossly. Conversations with the Nevada Department of Transportation revealed that it was not possible to extract this necessary information from the data set. This made it difficult to specify a time-series model from the Nevada data. As a result, the 5 years of data had to be pooled for each pavement section. The data were averaged over the 5 years to produce average values of damage, condition, maintenance, and the explanatory variables over a 5-year period. The 5-year averages represent the best method for representing the long-term effects of maintenance; on the average, a better-maintained pavement section will be in a better condition. Rather than predict the condition of the pavement every year, the models presented in this section predict the average condition over a 5-year period. This information will be useful to highway agencies for medium- to long-range maintenance planning.

A large number of the 4,871 pavement sections had errors, miscodings, and missing data; cleaning of the data and removal of outliers brought the number of observations down to 3,837. The variables used in the analysis for each section are summarized in Table 1 (reference names are given in parentheses). The PSI is used as the dependent variable for the deterioration models. Each variable is a 5-year average; for variables that are nonlinear functions of other variables (such as the PSI), the value of the variable was calculated for each year and then averaged over the 5 years. For example, the average daily truck traffic as a percentage of ADT was calculated by computing the percentage truck traffic for each of the 5 years and then taking their average. The ranges and mean values of the data are summarized in Table 2.

The data were collected from three districts, labeled districts 1, 2, and 3. Conversations with the Nevada Highway

Administration revealed that district 1 is in the south of Nevada, with a very dry climate. The major arterial highway through Nevada goes through this district, so there is higher car and truck traffic in district 1 than in the other districts. Districts 2 and 3 are in the mountains with more precipitation than district 1, but with less traffic. Analysis of the data, however, indicates that there are no substantial differences between the traffic in district 1 and district 2, and the district 2 traffic even appears to be higher. Some summary traffic and climate statistics are presented in Table 3 for each district (for the 3,837 observations).

The fact that some of the high-traffic roads are in areas that are environmentally less susceptible to deterioration creates an ambiguity in the role that traffic plays in deterioration. This ambiguity is compounded by the fact that overall traffic volumes, except in a few heavily traveled sections, are very low. Fifty percent of the pavement sections have volumes of less than 200 vehicles per day, and 90 percent of the sections have less than 2,000 vehicles per day. In addition, as can be seen from Table 2, even the areas that have rainfall receive relatively small amounts of precipitation; the average precipitation is only 8.5 in., and the average number of wet days (i.e., days with any precipitation at all) is only 44, which corresponds to less than 4 days of precipitation a month.

TABLE 1 DESCRIPTION OF VARIABLES USED IN ANALYSIS

a) average PSI;measure of pavement condition (PSI)
b) average sq. yds. of sand seal coat per year (SANDSEAL)
c) average lbs. of filler used for crack filling per year (CR. FILL.)
d) average sq. yds. of chip seal coat per year (CHIPSEAL)
e) average sq. yds. of flush seal coat per year (FLUSHSEAL)
f) average daily traffic (ADT)
g) average daily truck traffic as percentage of ADT (TRUCKS)
h) average age of section (AGE)
i) square of average age (AGE2)
j) district 1 dummy (DIST1)
k) district 2 dummy (DIST2)
l) road surface type dummy (SURF. TYPE)
m) average annual precipitation (in.) (PRCPTN)
n) average annual number of freeze-thaw cycles (FRZE/THAW)
o) average annual minimum temperature (MIN TEMP)
p) average annual number of days with precipitation (WETDAYS)
q) average extent of linear or alligator cracking on section (sq. ft.)(C)
r) average extent of patching on section (sq. ft.) (P)
s) average rut depth on section (inches) (RD)
t) average slope variance on section (in^2) (SV)
u) average width of linear cracking on section (in.) (Al. Width)
v) average width of alligator cracking on section(in.) (Lin. Width)

Model Specifications

Single-Equation Specification

Table 4 shows the simplest model that can be estimated on the data set, referred to as MODEL0. This is a single-equation OLS regression model with the PSI as the dependent variable. This is the type of model that has traditionally been estimated

TABLE 2 SUMMARY STATISTICS OF VARIABLES

Variable	Mean	Minimum	Maximum
-PSI	-2.55	-1.25	-4.14
ADT	690.00	5.00	8990.00
NO. OF TRUCKS	126.00	0.00	1680.00
AGE	14.40	3.00	18.00
MIN. TEMP.(degrees F)	36.30	20.00	56.00
FREEZE-THAW CYCLES	147.00	0.00	230.00
PRECIPITATION(in.)	8.50	2.00	35.00
WETDAYS	44.20	11.00	81.00
AREA OF SAND SEALCOAT (yds^2)	1600.00	0.00	7250.00
AMOUNT OF CRACK FILLING (lbs.)	807.00	0.44	5300.00
AREA OF CHIP SEALCOAT (yds^2)	570.00	0.00	8017.00
(CRACKING + PATCHING)$^{0.5}$ (ft.)	11.68	0.00	35.10
(RUT DEPTH)^2 in^2	0.02	0.00	0.50
log$_{10}$ (1 + SLOPE VARIANCE)	0.92	0.00	1.59

TABLE 3 MEAN VALUES OF KEY VARIABLES BY DISTRICT

District	1	2	3
average daily traffic	765	1050	350
average daily truck traffic	125	170	90
average precipitation(in./yr)	6.5	8.0	10.5
average annual min. temperature	42	35.2	31.9

in the literature. All the variables are in deviation form (i.e., for each observation, each variable has been subtracted from its sample mean). There is therefore no constant term for the model. It is not required to estimate the intercepts, because the PSI is measured on an arbitrary scale. There is no "natural" physical interpretation for the intercept, and so even if it were estimated, it could be absorbed by rescaling the PSI.

Four preventive maintenance activities performed on the pavements are sand seal coat, flush seal coat, chip seal coat, and crack filling. Because there are very few observations on which flush seal coating has been performed, this activity was not used in the analysis.

As can be seen from Table 4, many significant parameter estimates have the wrong signs. For example, the model shows that the condition of the pavement deteriorates as the extent of crack filling increases. Similarly, traffic has the wrong sign, with the pavement condition improving as the ADT increases. Some environmental variables have the wrong sign as well. The two most significant variables are the road surface type dummy and age. The road surface type dummy has a value of 1 if the road surface is concrete mix, and a value of 0 if the surface is bituminous mix. Finally, the district dummy parameters are not significant: with district 3 as the reference, pavements in district 1 have somewhat higher PSIs, whereas those in district 2 have somewhat lower PSIs. As shown by the table of summary statistics for each district, district 2 has the highest traffic and more precipitation than district 1. District 3 has more precipitation than the other two sections, but substantially less traffic. This might be the reason for the difference between the districts.

As discussed in the section on the Simultaneous-Equation Model System, the incorrect signs on the traffic and maintenance parameters are likely to be due to the fact that the correct specification is a simultaneous-equation model. In order to investigate this further, Table 5 plots traffic and the PSI aggregated into categories. The categories are defined as fol-

TABLE 4 SINGLE-EQUATION PSI MODEL: MODEL0

Dependent Variable: (PSI)

Independent Variable	Estimate	t-statistic
Maintenance		
Sand Seal Coat Area	0.151	6.74
Crack Filling Amount	-0.316	-3.72
Chip Seal Coat Area	0.179	5.94
Traffic and Age		
Av. Daily Traffic	1.550	4.12
Percentage Trucks	-2.110	-5.11
Age	-1.447	-8.98
Age2	0.142	6.64
Environmental		
Av. Annual Precipitation	0.255	2.59
Av. Annual Min. Temp.	0.064	0.82
Av. Annual Freeze/Thaw Cycles	0.223	2.39
Av. No. of Wet days	-0.675	-5.52
Dummies		
District 1 Dummy	0.173	1.06
District 2 Dummy	-0.283	-2.02
Road Surface Dummy	-2.334	-23.16

No. of observations: 3837
R^2: 0.38

TABLE 5 PERCENTAGE OF PAVEMENT IN PSI CATEGORY BY TRAFFIC CATEGORY

		PAVEMENT CONDITION (PSI)			
		Low	Medium	High	Total
ADT	Low	30.1	59.0	10.9	100.0
	Medium	1.90	62.7	28.4	100.0
	High	0.90	54.8	44.3	100.0

lows: for ADT, low = 0 to 1,000; medium = 1,000 to 4,000; and high = >4,000. For PSI, low = 0 to 2.5; medium = 2.5 to 3.5; and high = >3.5. The table gives the percentage of pavements belonging in each category. As can be seen from Table 5, the percentage of pavements in the low-PSI category decreases from 30.1 to 0.9 percent as the traffic increases. Correspondingly, the percentage of pavements in the high-PSI category increases from 10.9 to 44.3 percent as traffic increases. This indicates that high-traffic roads are better maintained, and in better condition than low-traffic pavements. The same fact is also illustrated by Table 6, where the average traffic is plotted by PSI category. As the PSI increases, the average traffic increases from 293 cars per day to 1,504 cars per day.

The effect of maintenance is seen in Table 7, where the extents of the three maintenance activities are tabulated against traffic categorized as low, medium, and high. For almost all activities, the extent of maintenance performed is more for the high-traffic pavements than for the low-traffic pavements.

Examination of Tables 5 through 8 shows that high-traffic pavements are better maintained and in a better condition than low-traffic pavements. A simultaneous-equation specification is therefore appropriate. This specification is now described.

Simultaneous-Equation Specification

The simultaneous-equation model is specified as an extension of MODEL0, referred to as MODEL1. The deterioration model is identical to the specification of MODEL0 (with PSI as the dependent variable, and the independent variables listed in Table 4). In addition, three maintenance equations are specified as follows:

$$\mathrm{PSI} = \beta_1 A_1 + \beta_2 A_2 + \beta_3 A_3 + \beta_4 X_1 + \eta_1 \qquad (8a)$$

$$A_1 = \gamma_1 (\mathrm{PSI}) + \Gamma_1 X_2 + \eta_2 \qquad (8b)$$

$$A_2 = \gamma_2 (\mathrm{PSI}) + \Gamma_2 X_2 + \eta_3 \qquad (8c)$$

$$A_3 = \gamma_3 (\mathrm{PSI}) + \Gamma_3 X_2 + \eta_4 \qquad (8d)$$

where X_1 is a 10-element vector of explanatory variables from Table 4 that includes all the listed variables except the three maintenance variables and X_2 is a vector of explanatory variables for the maintenance models consisting of average daily traffic, percentage of trucks, and road surface type. A_1 refers to the sandseal maintenance variable, A_2 to crack filling, and A_3 to the chip sealing activity, and Γ_1, Γ_2, and Γ_3 are each (1×3) vectors of parameters.

Equation 8a specifies the deterioration model, whereas the other three equations specify the maintenance model for each of the three routine maintenance activities considered. The specification of the explanatory variables was arrived at after conversations with the Nevada Highway Administration. In this model system the maintenance equations (8b–d) are identified relative to the deterioration equation (8a) because they do not include a number of exogenous variables that are included in Equation 8a. However, as discussed in the section on the Simultaneous-Equation Model System, the deterioration equation is not identified without further restrictions, because there are no explanatory variables in the maintenance equations (8b, 8c, and 8d) that do not appear in the deterioration equation (8a). Fortunately, the restrictions that $E(\eta_1,$

TABLE 6 MEAN VALUES OF ADT BY PSI CATEGORY

Condition (PSI)	**Low**	**Medium**	**High**
average daily traffic	295	670	1504

TABLE 7 EXTENT OF MAINTENANCE BY TRAFFIC CATEGORY

	AVERAGE DAILY TRAFFIC (ADT)		
MAINTENANCE EXTENT	**Low**	**Medium**	**High**
ACTIVITY			
Sand Seal Coat(sq. yds.)	559	1887	2049
Crack Filling(lbs. of filler)	117	379	726
Chip Seal Coat(sq. yds.)	536	738	598

TABLE 8 MEAN VALUES OF PSI BY CRACK-FILLING EXTENT

Crack Filling Extent	**Low**	**Medium**	**High**
Average PSI	2.86	2.86	2.96

$\eta_2)$, $E(\eta_1, \eta_3)$, and $E(\eta_1, \eta_4)$ are all zero satisfy the necessary conditions for the identification of Equation 8a. These restrictions are based on the assumption that the physical process of deterioration and the maintenance behavior of the agency are two different processes, and there is no reason why their error terms should be correlated.

The maximum likelihood estimates of the parameters of MODEL1 are given in Tables 9 and 10. Table 9 shows the deterioration equation parameters, and Table 10 the parameters of the maintenance equations. As is apparent from these tables, all the parameters have the expected sign. From Table 9, the condition of the pavement improves as the maintenance increases, and gets worse as the traffic and the percentage of trucks increase, and as the pavement gets older. The significant environmental variables also have the expected signs. As in MODEL0, age and the road surface dummy are the most significant variables. Traffic and most of the environmental effects except for the number of wet days are all not significant. This is a reasonable result since in the data set, traffic levels are low, and there is not much precipitation. In Table 10 as well, all parameters have the expected signs. Maintenance increases as the pavement condition gets worse, reflected by the positive sign of the coefficient of condition in the maintenance equation. Similarly, the maintenance activity increases as traffic and the percent of trucks increase.

For MODEL1, the fit of the deterioration equation, as measured by the R^2-statistic is 0.27, from Table 9. This value is less than that obtained from MODEL0, which is only to be expected since MODEL1 is a simultaneous-equation formulation of the single regression equation of MODEL0. A

TABLE 9 DETERIORATION MODEL: MODEL1

Dependent Variable: (PSI)

Independent Variable	**Estimate**	**t-statistic**
Maintenance		
Sand Seal Coat Area	0.434	3.01
Crack Filling Amount	1.676	6.46
Chip Seal Coat Area	0.042	0.16
Traffic and Age		
Av. Daily Traffic	-0.067	-0.76
Percentage Trucks	-2.649	-6.65
Age	-1.743	-10.83
Age2	0.172	5.46
Environmental		
Av. Annual Precipitation	0.205	1.43
Av. Annual Min. Temp.	0.141	1.72
Av. Annual Freeze/Thaw Cycles	0.308	3.30
Av. No. of Wet days	-0.453	-3.30
Dummies		
District 1 Dummy	-0.618	-2.54
District 2 Dummy	-0.209	-0.60
Road Surface Dummy	-1.700	-13.58

No. of observations: 3837
R^2: 0.28

TABLE 10 MAINTENANCE MODEL: MODEL1

Dependent Variable: Maintenance Activity

Independent Variable	**Sand Seal Coat Estimate/t-statistic**	**Chip Seal Coat Estimate/t-statistic**	**Crack Filling Estimate/t-statistic**
PSI	-0.154/-2.40	0.049/0.75	-0.072/-7.83
Av. Daily Traffic	0.147/3.67	0.047/1.25	0.113/8.
Percentage Trucks	0.206/5.23	0.828/2.48	-0.228/-2.79
Road Surface Dummy	-1.471/-7.64	0.532/2.74	-0.311/-9.83
R^2	**0.05**	**0.05**	**0.03**

single equation OLS would yield the highest possible value of R^2 for that equation. The simultaneous-equation estimator gets rid of the bias, but in the process worsens the fit. In some situations one may accept some level of bias to achieve a better fit. However, in this case MODEL0 is unacceptable because key parameters were estimated to have counterintuitive signs.

Summarizing the results obtained from the case study, the simultaneous-equation specification, MODEL1, solves the problems associated with the single-equation deterioration model. All the significant parameters have the expected signs. The simultaneous-equation specification therefore appears to be a more realistic model for predicting the deterioration of highway pavements.

CONCLUSION

This paper has described an approach to modeling highway pavement deterioration from data collected on in-service pavements that improves on methods that are traditionally used. The deterioration model is estimated simultaneously with a set of maintenance equations, which describe the behavior of the agency performing maintenance on the pavement. This specification might solve some of the problems with unreasonable signs for parameter estimates that have been reported in the literature.

Finally, a case study was presented that applied the concepts to data collected in the field. The results from the case study indicated that the simultaneous-equation specification was indeed appropriate and produced sensible results. It is very important to estimate an accurate deterioration prediction model that realistically takes into account the role of maintenance for purposes of maintenance planning and strategy selection. The approach presented in this paper aids in the formulation and estimation of such a model.

REFERENCES

1. A. Garcia-Diaz and M. Riggins. Serviceability and Distress Methodology for Predicting Pavement Performance. *Transportation Research Record 997*, TRB, National Research Council, Washington, D.C., 1984.
2. R. K. Kher and W. D. Cook. PARS, The MTC Model for Program and Financial Planning in Pavement Rehabilitation. North American Management Conference, Vol. 2. Ontario Ministry of Transport and Communications, 1985.
3. F. T. Fwa and K. C. Sinha. A Study of Pavement Routine Maintenance Effects. An investigation conducted by Purdue University and Indiana DOT, FHWA, U.S. Department of Transportation, 1986.
4. M. Y. Shahin and S. D. Kohn. Pavement Maintenance Management for Roads and Parking Lots. Technical Report M-294. Construction Engineering Research Laboratory, U.S. Army Corps of Engineers, Champaign, Ill., 1981.
5. W. D. O. Paterson. *Road Deterioration and Maintenance Effects: Models for Planning and Management*. Johns Hopkins University Press, Baltimore, Md., 1987.
6. B. C. Butler, R. F. Carmichael III, and P. R. Flanagan. *Impact of Pavement Maintenance on Damage Rate*, Vol. 2. Final Report DTFH61-83-C-00125. ARE, Inc., FHWA, U.S. Department of Transportation, 1985.
7. *Pavement Performance*. Special Report No. 61E, HRB, National Research Council, Washington, D.C., 1962.
8. W. N. Carey and P. E. Irick. The Pavement Serviceability Performance Concept. *Highway Research Board Bulletin 250*, HRB, National Research Council, Washington, D.C., 1960.

Publication of this paper sponsored by Committee on Pavement Management Systems.