TECHNICAL APPENDIX

ESTIMATING THE OPTIMUM MILEAGE FOR VEHICLE RETIREMENT

Regression analysis and calculus provide convenient tools for estimating the optimum point for retiring vehicles from a large fleet. In such large fleets, mileage points derived from these mathematical tools alert fleet management to examine high mileage vehicles. These methods are intended to be used by fleet management in conjunction with mechanical expertise, not as a substitute for it.

This appendix is divided into three parts corresponding to the three equations used in the JLARC vehicle retirement methodology. After an introductory overview of the problem, the operating expense function is detailed. Next, the capital expense function is provided. Finally, these two equations are combined to form the total expense function, and the application of calculus to determine the minimum total expense per mile is discussed.

DATA LIMITATIONS

The primary aim of regression analysis in this exercise is only to provide an improved average for vehicle expenses, not to explain all or most causes of change in the dependent variable. The simple model based on expenses per mile in this example is strongly preferred to a complex model with a somewhat better fit which may include age of vehicle or types of use in addition to mileage. Therefore, the regression results only estimate the component parts of the total average expenses per mile. More complex regression models were examined, as noted in the Appendix, with only limited improvement in the estimate. JLARC staff analysis indicates that the improvement in the model is marginal at best and the tradeoff in confused interpretation and explanation of the final results of the vehicle replacement mileage is judged to be too great.

Inability to predict expenses with greater accuracy are due to two factors: (1) The present system of fleet operations and (2) factors beyond the control of the Central Garage. Errors in the data for expenses are introduced by bookkeeping procedures which debit and credit expenses incurred in an accident and reimbursed by an insurance company. If the debits and credits are correctly entered, then there is no record of a major expense. In some cases the record shows a negative value for a repair expense. JLARC staff determined that other errors are currently beyond the control of the Central Garage management but new technologies will soon allow control. Inability to predict expenses is introduced by unmeasured variables which are theoretically likely to explain some expenses. Variables such as days out of service due to repair, type of maintenance and repair performed, vehicle usage data, regional expense differences, etc., would improve predictions and are expected under the new Equipment Management System now in preparation through VDOT.

OPERATING EXPENSE FUNCTION

The operating expenses for each vehicle in the fleet are maintained on the VDOT performance master automated data system pending conversion to the new vehicle tracking system. The data were used to predict average operating expenses per mile for each mile within the range of current fleet mileage. The predictions are based on regression equations.

Simple, bivariate regression is used to regress maintenance and repair expenses (M&R(x) and fuel and fluid expenses (F(x)) on mileage (x) for each of four vehicle types. The emphasis of this research was prediction of these expenses for the average vehicle.

More accurate predictions may have been produced by controlling for other factors, such as the type of usage or the model year of the vehicle. Model year was found to slightly enhance the predictive powers of the model (see data runs in personal project-related documentation), but only at the expense of complicating interpretation of the results through the use of multiple regression. The marginal gains were not judged to be worth the cost in generalizability and ease of explanation. Also, since the goal was improving on subjective judgment, maximizing the percentage of variance explained through a more complete model was not the first priority.

Further, the central garage data on individual vehicles was not sufficient to explore substantially more sophisticated models. For example, details on the conditions of usage for the vehicles were unavailable. Attempts to proxy types of usage by controlling for the agency to which the vehicle was assigned proved unsuccessful. Another important factor, performance of preventative maintenance, could be not be separated from other maintenance expenses given the current data. The new VDOT ADP system should allow investigation of this factor.

Estimating the Operating Expenses per Mile

The operating expense function is actually the sum of two functions:

$$O(x) = M\&R(x) + F(x)$$

Both functions were derived by regressing operating expenses on vehicle mileage. The result represents an average expense for each mile point. This result represents an improvement over the best guess without regression, and allows computation of a total expense function.

COMPACT SEDANS N=1830

F(x) = .0249 + .00000015 X
M&F(x) = .0249 + .00000019 X

therefore

O(x) = .0428 + .00000034 X

COMPACT STATION WAGONS N=237

F(x) = .0263 + .00000011 X
M&R(x) = .0242 + .00000014 X

therefore

O(x) = .0505 + .00000025 X

LARGE SEDANS N=416

F(x) = .0312 + .00000029 X
M&R(x) = .0167 + .00000037 X

therefore

O(x) = .0479 + .00000066 X

VANS N=113

F(x) = .0777 (Note: Fuel usage was made constant because equation predicted a slight decline over mileage driven, a finding which was judged inaccurate.)

M&R(x) = .0626 + .00000023 X

therefore

O(x) = .1403 + .00000023 X

Note that all coefficients are statistically significant for each vehicle type (see printout). The standard error of the estimate for the fuel expense model is .005, which is acceptable. The standard error for the maintenance and repair model is less good, .02, but still acceptable given our use. The maintenance and repair model for vans is probably the least reliable, probably due to the low number (N=113) and the limited variation. (The van fleet was purchased primarily over a two year period and the data is not as well distributed as compact sedans, which are represented over seven years. Regression results in this case especially should be cautiously interpreted).

Limitations of the Estimates

Several problems were noted with the data set which limit the generalizability of these results beyond the data set. Most problems are not implicit in the methodology but the result of current circumstances. Anticipated improvements in the data processing system should lead to more confidence in future results.

Negative values. Most important, numerous errors were uncovered in the maintenance and repair data. Many negative amounts were present. The central garage administrator explained that these result from the central garage accounting system, and each negative number represents a credit which should have a corresponding debit.

For example, body work on wrecked vehicles is often done by a private body shop. A debit is incurred when the work is completed, and the central garage would normally file an insurance claim. When the insurance company reimburses the state, a corresponding credit will eliminate the debit.

The result of this accounting practice is flawed maintenance and repair records. If both a debit and credit are recorded as intended, no record of a major expense is retained and a vehicle's maintenance and repair expenses are underreported. In addition, a visual review of the database seemed to indicate negative values without apparent offsetting positive values. In this case, the database would be more misleading.

Data Dispersion. These regressions tended to produce a loose fit to the regression line, indicating substantial variation among vehicles. This looseness is reflected in a relatively low coefficient of determination (R^2), especially for the maintenance and repair expenses.

The R^2 value suggests how much our guess about the expense rate per mile is improved by knowing its mileage over taking a simple average of the expense rate per mile of all vehicles in its class. For example, we are able to improve our estimate of fuel rates for compact passenger cars by 33 percent by knowing the mileage. Similarly, we are able to improve our estimate of maintenance and repair rates by about five percent, suggesting that much of the variance remains unexplained. These figures were similar across all vehicle classes.

Both figures are relatively low, especially maintenance and repair expenses. Perhaps the low maintenance and repair percentage is unsurprising when one considers the number of virtually random occurrences, such as accidents or hard usage, which may occur and are not likely to be predicted by presently available fleet-wide statistics. Maximizing the R^2 value is not vital, and any value above zero suggests we are improving on the simple average expense per mile.

Selection Bias and Nonlinear Fit. One problem appears related to the methodology. Vehicles which prove to be particularly poor performers or those involved in serious accidents will probably be removed from the fleet early. This practice was confirmed by the administrator and is proper. However, the result of removing poor performers is that a form of bias, selection bias, may be introduced to the data set.

Selection bias occurs if all vehicles do not have an equal chance of being represented in the data set. Evidence of such bias is present in this data. Examination of the residuals from the bivariate regressions for maintenance and repair expenses reveals a nonlinear relationship between mileage and expenses. The direction of the relationship suggests selection bias.

Using a double-log model, the regression results revealed an initially surprising finding. Rather than maintenance and repair expenses <u>increasing</u> exponentially at higher mileages as expected, the expenses were <u>decreasing</u> exponentially. That is they were increasing at a decreasing rate.

Commonsense suggests that the condition of older, higher mileage vehicles might rapidly decline, resulting in ever higher incidence of repairs. Expenses per mile would then be expected to increase at an increasing rate. However, the data suggests that the expenses increase at a declining rate. The best explanation for this anomaly is selection bias. That is, older performers are no longer in the data set. Therefore, the high mileage vehicles in the data set are unrepresentative of the average fleet – they are better than average performers. These data points would artificially draw the regression line downwards at the end only because average or below average vehicles are missing. Therefore, the expenses per mile increase at a declining rate because the high end of the sample is unreliable.

Few remedies exist. Recent work by Chris Achen (The Statistical Analysis of Quasi-Experiments) discusses censored samples (the missing poor performers) and recommends remedies to model the missing data.

CAPITAL EXPENSE FUNCTION

The purchase of fleet vehicles represents a large capital outlay. Some of the purchase price of a vehicle may be recovered through sale of the vehicle. Therefore capital investment is purchase price less salvage value. Since salvage value can never be expected to exceed purchase price, some positive expense will always be associated with a vehicle.

Capital investment may be thought of as distributed over the number of miles driven. The longer the vehicle is in use, the lower the capital investment per mile. The rate of change in the capital investment per mile can be estimated using a function based on fleet averages.

Average capital investment per mile tends to decrease over the life of the vehicle as the initial investment is spread over more and more miles. Average capital investment per mile may be expressed mathematically as follows:

$$C(x) = \frac{P - S(x)}{x}$$

, where:

C(x)= Average Capital investment per mile at mile x;

P= Purchase price, and;
S(x)=Salvage value per mile at mile x.

Estimating the Purchase Price

The purchase price of the vehicle is a constant. It does not fluctuate with mileage. JLARC staff averaged the actual purchase prices for vehicles within a particular class. The purchase price is listed on the CARS2 file (KSHMILE.RXD) as the variable “VALUE”. The purchase prices for each model year were averaged over all years which are still in the fleet.

Purchase Price Averages (P)

Compact Sedan	$6,479
Compact Station Wagon	7,150
Large Sedan	8,967
Vans	11,972

Estimating the Salvage Value and Capital Expense per Mile

JLARC staff used the National Automobile Dealers Association 1987 Official Used Car Guide in estimating the salvage value of the fleet. Each fleet vehicle required a salvage value to estimate the capital investment in each car, since the purchase price is moderated by the salvage value remaining in a vehicle. Unlike the purchase price, the salvage value of a vehicle does vary with mileage. That is, the greater the mileage, the lower the salvage value.

Each salvage value is expressed as a function of mileage to be consistent with the rest of the optimum retirement point analysis. However, the NADA guide gives salvage value by age of the vehicle. Therefore, JLARC staff converted age to a mileage by multiplying the fleet’s average mileage (14,400 miles) by the age of the vehicle in years as follows:

1986	14,400 miles
1985	28,800 miles
1984	43,200 miles
1983	57,600 miles
1982	72,000 miles

NOTE: This annual mileage is somewhat low. The final average mileage is 15,000 miles per year. However, mathematically, the results are identical for 14,400 and 15,000 miles. This is true because the intercept is the same for each vehicle class. Only the regression coefficient changes when 15,000 miles is substituted. The regression coefficient drops out when the first derivative is taken. Therefore, the first derivative is identical regardless of the mileage used and 14,400 was retained.

Salvage value had to be estimated for each vehicle at the present time to yield a capital expense per mile for each of the fleet's vehicles. This represents the remaining capital expense from the initial purchase which cannot be recovered through resale. The estimate of salvage value is based on the NADA high mileage average loan value because auction prices are not expected to meet retail used vehicle sales and most vehicles are sold with high mileage.

But this process results in salvage values for only one mileage point every 14,400 miles. A continuous salvage function is needed to provide a salvage value for each mile point. JLARC staff regressed salvage value on mileage (a process similar to straight line depreciation). The resulting salvage functions (S(x)) and capital expense functions ((P - S(x))/x) by vehicle class is:

COMPACT SEDANS

YEAR	HIGH MILES LOAN VALUE	ESTIMATED MILEAGE
1982	$ 950	72,000
1983	$1,225	57,600
1984	$1,825	43,200
1985	$2,525	28,800
1986	$3,075	14,400

S(x) = $3,403.33 - 0.033X, where:
X = mileage

C(x) = ($6,479 - (3,403.33 - .033X))/X = 3076.67/X+ .033

COMPACT STATION WAGONS

YEAR	HIGH MILES LOAN VALUE	ESTIMATED MILEAGE
1982	$1,175	72,000
1983	$1,775	57,600
1984	$2,575	43,200
1985	$2,875	28,800
1986	$3,050	14,400

S(x) = $3,682.00 - 0.032X, where:
X = mileage

C(x) = ($7,150 - (3682 - .032X))/X = 3468/X+.032

LARGE SEDANS

YEAR	HIGH MILES LOAN VALUE	ESTIMATED MILEAGE
1981	$1,925	86,400
1984	$1,800	43,200
1985	$3,150	28,800
1986	$3,550	14,400

S(x) = $4,190 - 0.034X, where:
X = mileage

C(x) = ($8,967 - (4,190 - .034X))/X = 4777/X+.034

VANS

YEAR	HIGH MILES LOAN VALUE	ESTIMATED MILEAGE
1983	$3,950	57,600
1984	$5,125	43,200
1985	$5,925	28,800

S(x) = $7,962.50 - 0.069X, where:
X = mileage

C(x) = ($11,972 - ($7,962.50 - .069X))/X = 4009.50/X+.069

The standard errors of the estimates ranged from $73 to $230. The standard errors of the estimates, when doubled (for 95% confidence interval), yield the range within which the estimates are said to vary with a 95 percent level of accuracy. These ranges were taken to be tolerable for prediction, based on the above assumptions and measurements.

ESTIMATING THE END OF THE VEHICLE'S EFFICIENT LIFE

If all expenses are accounted for in the formulas, the optimum point at which to retire a vehicle occurs at the mileage point yielding the lowest total expenses per mile. This is true because, as noted in the text, each additional mile driven after this point will see continuously rising operating costs. These rising costs outweigh the capital savings from distributing the initial cost of the vehicle over more miles.

Therefore, purchase of a new vehicle is prudent. Despite the high initial capital outlay, these capital expenses will eventually be defrayed over many miles also. The new vehicle will have lower operating costs, over time justifying the purchase.

Questions of change in rates may be handled through application of differential calculus. For example, the original equation (T(x)) is expressed in dollar amounts per mile. The rate of change in dollar amounts per mile is derived by taking the first derivative of the original equation (T'(x)) and looking for the stationary point.

A stationary point will be that point in a curve where the curve changes direction. It is found by setting the equation to zero, and solving for x (T'(x)=0, x=?). If there is only one stationary point, then that point represents the maximum or minimum.

It is necessary to calculate the second derivative (T''(x)) to prove that there exists such a point and it is the minimum. In this case, replace x with the result from the first derivative (e.g., for passenger cars, T''(95,000)=?). If a single answer exists and is greater than zero, then the unique minimum has been found.

Calculating the Minimum Total Expense per Mile

T(x) = C(x) + O(x)
T'(x) is the first derivative of T(x) and is used to find the mileage with the minimum total expenses per mile.
T''(x) is the second derivative and simply demonstrates that the mileage point produced by T'(x) is a unique minimum.

COMPACT SEDANS

C(x) = ($6,479-(3,403.33-.033X))/X = 3076.67/X+.033
O(x) = .0428 + .00000034 X
therefore,
T(x) = 3076/X + .0758 + .00000034 X
and
T'(x) = -3076 X^{-2} + .00000034
solving for X at the Stationary point, T'(x) = 0,
0 = -3076 X^{-2} + .00000034
X^2 =9,047,058,824
X = 95,116

and
T''(x) = 3076 X^{-3}
solving for X at the Stationary point, T''(95,116)
X^3 = 3076, therefore X is greater than 0 and 95,000 is the unique minimum.

COMPACT STATION WAGONS

C(x) = ($7,150 - (3682 - .032X))/X = 3468/X + .032
O(x) = .0505 + .00000025 X
therefore,

$T(x) = 3468\ X^{-1} + .0825 + .00000025\ X$

and

$T'(x) = -3468\ X^{-2} + .00000025$

solving for X at the Stationary point, T'(x) = 0,

$0 = -3468\ X^{-2} + .00000025$

$X^2 = 13{,}872{,}000{,}000$

$X = 117{,}779$

and

$T''(x) = 3468\ X^{-3}$

solving for X at the Stationary point, T''(117,779)

$X^3 = 3468$, therefore X is greater than 0 and 117,799 is the unique minimum.

LARGE SEDANS

$C(x) = (\$8{,}967 - (4{,}190 - 0.34X))/X = 4{,}777/X + .034$

$O(x) = .0479 + .00000066\ X$

therefore,

$T(x) = 4777\ X^{-1} + .0819 + .00000066\ X$

and

$T'(x) = -4777\ X^{-2} + .00000066$

solving for X at the Stationary point, T'(x) = 0,

$0 = -4777\ X^{-2} + .00000066$

$X^2 = 7{,}237{,}878{,}787$

$X = 85{,}075$

and

$T''(x) = 4777\ X^{-3}$

solving for X at the Stationary point, T''(85,075)

$X^3 = 4777$, therefore X is greater than 0 and 85,075 is the unique minimum.

VANS

$C(x) = (\$11{,}972 - (\$7{,}962.50 - .069X))/X = (4009.50 + .069X)/X$

$O(x) = .1403 + .00000023\ X$

therefore,

$T(x) = 4010/X + .2089 + .00000023\ X$

and

$T'(x) = -4010\ X^{-2} + .00000023$

solving for X at the Stationary point, T'(x) = 0,

$0 = -4010\ X^{-2} + .00000023$

$X^2 = 17{,}434{,}782{,}609$

$X = 132{,}040$

and

$T''(x) = 4010\ X^{-3}$

solving for X at the Stationary point, T''(132,040)

X^3 = 4010, therefore X is greater than 0 and 132,040 is the unique minimum.

Estimating the Salvage Value and Capital Expense per Mile

JLARC staff used the National Automobile Dealers Association 1987 Official Used Car Guide in estimating the salvage value of the fleet. Each fleet vehicle required a salvage value to estimate the capital investment in each car, since the purchase price is moderated by the salvage value remaining in a vehicle. Each salvage value is expressed as a function of mileage to be consistent with the rest of the optimum retirement point analysis. However, the NADA guide gives salvage value by age of the vehicle. Therefore, JLARC staff converted age to a mileage by multiplying the fleet's average mileage by the age of the vehicle in years. Purchase price of the class was averaged to give a single figure. JLARC staff selected the high mileage average loan value because auction prices are not expected to meet retail used vehicle sales and most vehicles are sold with high mileage.

JLARC staff regressed salvage value on mileage. The resulting function by vehicle class is:

COMPACT SEDANS

YEAR	HIGH MILES LOAN VALUE	ESTIMATED MILEAGE
1982	$ 950	72,000
1983	$1,225	57,600
1984	$1,825	43,200
1985	$2,525	28,800
1986	$3,075	14,400

S(x) = $3,403.33 - 0.033X, where:

X = mileage

C(x) = ($6,479 - (3,403.33 - .033X))/X = 3076.67/X + .033

COMPACT STATION WAGONS

YEAR	HIGH MILES LOAN VALUE	ESTIMATED MILEAGE
1982	$1,175	72,000
1983	$1,775	57,600
1984	$2,575	43,200
1985	$2,875	28,800
1986	$3,050	14,400

S(x) = $3,682.00 - 0.032X, where:

X = mileage

C(x) = ($7,150 - (3682 - .032X))/X = 3468/X + .032

LARGE SEDANS

YEAR	HIGH MILES LOAN VALUE	ESTIMATED MILEAGE
1984	$1,800	43,200
1985	$3,150	28,800
1986	$3,550	14,400

S(x) = $3,833 - 0.022X, where:
X = mileage
C(x) = ($9,653 - (3,833-.022X))/X = 5820/X + .022

VANS

YEAR	HIGH MILES LOAN VALUE	ESTIMATED MILEAGE
1983	$3,950	57,600
1984	$5,125	43,200
1985	$5,925	28,800

S(x) = $7,962.50 - 0.069X, where:
X = mileage
C(x) = ($11,972 - ($7,962.50 - .069X))/X = 4009.50/X + .069

Estimating the Operating Expenses per Mile

O(x) = M&R(x) + F(x)

COMPACT SEDANS N = 1830
F(x) = .0249 + .00000015 X
M&R(x) = .0179 + .00000019 X
therefore,
O(x) = .0428 + .00000034 X

COMPACT STATION WAGONS N = 237
F(x) = .0263 + .00000011 X
M&R(x) = .0242 + .00000014 X
therefore,
O(x) = .0505 + .00000025 X

LARGE SEDANS N= 226
F(x) = .0332 + .00000014 X
M&R(x) = .0192 + .00000014 X
therefore,
O(x) = .0524 + .00000028 X

VANS N = 113
F(x) = .0777

M&R(x) = .0626 + .00000023 X

therefore,

O(x) = .1403 + .00000023 X

Calculating the Minimum Total Expense per Mile

T(x) = C(x) + O(x)

T'(x) is the first derivative of T(x) and is used to find the mileage with the minimum total expenses per mile.

T''(x) is the second derivative and simply demonstrates that the mileage point produced by T'(x) is a unique minimum.

COMPACT SEDANS

C(x) = ($6,479 - (3,403.33 - .033X))/X = 3076.67/X + .033

O(x) = .0428 + .00000034

therefore,

T(x) = 3076/X + .0758 + .00000034

and

$T'(x) = -3076\ X^{-2} + .00000034$

solving for X at the Stationary point, T'(x) = 0,

$0 = -3076\ X^{-2} + .00000034$

$X^2 = 9{,}047{,}058{,}824$

X = 95,116

COMPACT STATION WAGONS

C(x) = ($7,150 - (3682 - .032X))/X = 3468/X + .032

O(x) = .0505 + .00000025

therefore,

$T(x) = 3468\ X^{-1} + .0825 + .00000025$

and

$T'(x) = -3468\ X^{-2} + .00000025$

solving for X at the Stationary point, T'(x) = 0,

$0 = -3468\ X^{-2} + .00000025$

$X^2 = 13{,}872{,}000{,}000$

X = 117,779

LARGE SEDANS

C(x) = ($9,653 - (3,833 - .022X))/X = 5820/X + .022

O(x) = .0524 + .00000028

therefore,

$T(x) = 5820\ X^{-1} + .0744 + .00000028\ X$

and

$T'(x) = -5820\ X^{-2} + .00000028$

solving for X at the Stationary point, T'(x) = 0,

$0 = -5820 X^{-2} + .00000028$

$X^2 = 20{,}785{,}714{,}000$

$X = 144{,}172$

<u>VANS</u>

$C(x) = (\$11{,}972 - (\$7{,}962.50 - .069X))/X = (4009.50 + .069X)/X$

$O(x) = .1403 + .00000023 X$

therefore,

$T(x) = 4010/X + .2089 + .00000023 X$

and

$T'(x) = -4010 X^{-2} + .00000023$

solving for X at the Stationary point, T'(x) = 0,

$0 = -4010 X^{-2} + .00000023$

$X^2 = 17{,}434{,}782{,}609$

$X = 132{,}040$