tude of the cross-subsidy will not be reduced. Thus, as the policy has been established, the misallocation of resources caused by the external subsidy will be added to the misallocation of resources caused by the current practice of cross-subsidization. The substitution of external subsidies for cross-subsidies, which would have decreased the misallocation of resources, all other things being equal, was not permitted by the new legislation and regulations.

The external subsidy program itself was established in a way that creates a misallocation of resources. First, the size of the federal commitment is out of proportion to the benefits the nation receives from maintaining branch lines. The use of subsidies, if properly implemented, improves the allocation of resources. In the case of local rail service, however, the external factors are largely local in nature. Therefore, most of the subsidy should be financed by local taxpayers.

Second, if the subsidy program was established to help save the local communities from economic disaster, the approach is inefficient. Past studies indicate that, in the majority of cases, the communities suffer little from the loss of rail service (2). A more productive approach to helping these communities economically would be to make this conditional grant less conditional. The communities should be allowed to invest the money in projects that would produce a larger benefit-cost ratio.

Third, more emphasis should be put on uses of the money other than continuation subsidies, as outlined by section 1a of the Interstate Commerce Act. In many cases, efficiency in allocation would be improved by devoting more money to help shippers make the switch to other modes of transport. Presumably, the rail continuation subsidy provisions provide only transitory help and thus will aid shippers. If these subsidies turn out not to be transitory in nature, however, then the shippers will continue to use a mode that, based on cost-revenue considerations, should not be used.

Fourth, as argued by Baumel, Drinka, and Miller (3), the nature of the branch-line subsidy program will not increase the efficiency of the local rail service and, thus, will not help the national constituency. By switching a portion of this aid from these lines to lines that are still able to provide valuable service to the rural and agricultural communities but need rehabilitating, a much larger return on investment of these public funds would be realized.

CONCLUSIONS

Though the full impact of the new branch-line policy will take time to fully reveal itself, two important but tentative conclusions can be reached.

First, unless the new legislation and regulations will have some indirect and unexpected effects on the decision processes of the ICC, the railroads will not be helped by the new policy. The legislation and regulations give incentives to the communities and shippers affected to continue using the uneconomic lines at least in the short run. The railroads, on the other hand, were given no tangible incentives and little encouragement to abandon burdensome branch lines.

The policy developed reflects a "political pareto optimal" solution to the light-density branch-line problem. By moving to the new policy, some were helped (communities, but mainly individual shippers), some were not hurt but not helped (the railroads), and some were hurt by such a small amount as to create no political problem (the general tax-paying public).

Second, Congress incorrectly assumed that the program that was appropriate for the 17-state region under the 3R Act was appropriate for the entire nation. The largely federally financed subsidy program under the 3R Act was undoubtedly a correct approach, given the situation where one particular region was facing massive, widespread abandonments occurring in a short period of time.

Under the 4R Act, however, the ICC will act on abandonment of lines on a case-by-case basis, and thus the impact of an abandonment decision on a large multistate area will be minimal. Therefore, more of the subsidy should have been financed by the state and local governments. Because of this and other characteristics of the rail continuation subsidy program, the new branch-line policy may cause a greater misallocation of resources than the old policy.

REFERENCES

1. Improving Railroad Productivity. Task Force on Railroad Productivity, Washington, DC, 1973.
2. B. J. Allen and J. F. Due. Railway Abandonments: Effects Upon the Communities Served. Growth and Change, Vol. 8, No. 2, Apr. 1977.
3. P. Baumel, T. P. Drinka, and J. J. Miller. Implications of the Local Rail Assistance Section of the Railroad Revitalization and Regulatory Reform Act of 1976. The Logistics and Transportation Review, Vol. 12, No. 5, 1976.

Publication of this paper sponsored by Committee on State Role in Rail Transport.

**This paper was written while Mr. Allen was a Brookings Economic Policy Fellow with the U.S. Department of Transportation.*

Class 2 Railroad Operating Costs

C. John Langley, Jr., and Edwin P. Patton, Transportation Center, University of Tennessee, Knoxville

Multiple regression analysis was used to develop predictive equations for the estimation of operating costs associated with the provision of class 2 railroad service. Annual report data for 102 carriers was the basis for the construction of five equations, each of which pertained to estimation of a specific type of operating cost. Categories included were maintenance of way, maintenance of equipment, traffic, transportation, and general. Five specific predictor variables were included in the analyses: carrier geographic location, ownership, main trackage, traffic volume, and one other depending on the particular type of cost being estimated. In addition, an equation was developed for the prediction of the sum in dollars of the individual costs. All equations appeared to be correctly specified, and each exhibited an acceptable explanatory ability. The research

findings from this study should provide significant insight into the expected magnitudes of the costs of operating a light-density line independent of class 1 ownership. The results will be of specific interest to states involved in developing and updating their state rail plans. Areas of primary application include branch-line economic viability analyses and efforts to rank branch lines in order to determine the best candidates for federal or state assistance or both.

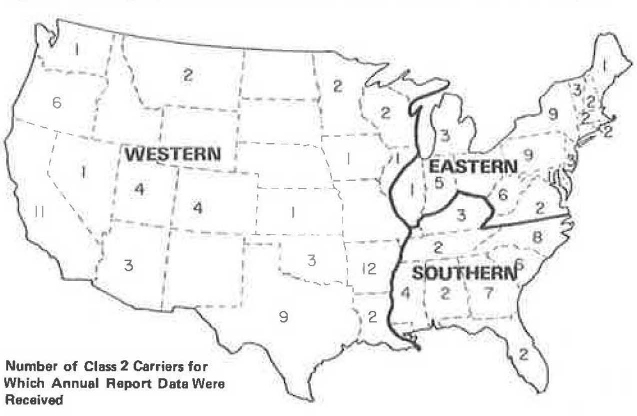
Section 803(j) of the Rail Revitalization and Regulatory Reform (4R) Act of 1976 provides for a federally sponsored program of financial assistance to the states for the provision and continuation of local rail service. There are, however, several criteria that must be met by the respective states in order to receive such assistance from the Secretary of Transportation. The major prerequisite is the development by each state of an adequate plan for rail services as part of an overall planning process for all transportation services, including a suitable process for updating, revising, and amending such a plan. In addition to a number of other components that must be included in an acceptable state rail plan, the planning methodology must include procedures for branch-line economic viability analyses and for prioritization of branch lines in order to determine which are the best candidates for federal or state assistance or both.

The research reported in this paper is designed to provide information of use in estimating the cost of providing rail freight service over light-density lines. Specifically, a study was undertaken of operating cost data for the population of class 2 rail carriers that were unrelated in terms of ownership to a class 1 railroad. Using a variety of data concerning the characteristics of the specific rail line operations, a series of equations was developed, each designed to explain a particular type of operating cost that was seen to be relevant. In addition to the five equations developed for each type of operating cost, an overall relationship was developed that permits the single-equation estimation of total operating expenses for light-density lines being operated independent of class 1 ownership.

ACQUISITION OF DATA

In order to study the factors affecting costs associated with a short-haul railroad operation, it was first necessary to define the types of information that would be important to such an investigation. A listing of the various categories of information follows. It should be noted that our original calculations in customary units have been retained throughout this paper.

Figure 1. Geographical locations of carriers included in data analyses.



Category	Subject
Railway operating expenses	Maintenance of way and structures Maintenance of equipment Traffic Transportation rail line General
Rail line characteristics	Location Ownership type Distance operated Connections with other carriers, etc.
Physical characteristics of the operation	Weight of rail Number of crossties Gauge of track Types of power equipment Ownership of freight cars, etc.
Employees, service, and compensation	Executives, officials, and staff assistants Professional, clerical, and general Maintenance of way and stores Transportation: yardmasters, switch tenders, and hostlers Transportation: train and engine
Rail line operation	Train-miles Locomotive unit-miles Car-miles Revenue and nonrevenue freight traffic Traffic by commodity types, etc.

Included are those related to railway operating expenses; rail line characteristics; physical characteristics of the operation; employees, service, and compensation; and rail line operation. The main objective of the analyses was to explain the various types of railway operating expenses associated with rail line operations.

Implementation of the methodology required that the above types of data be made available for a number of existing short-haul carrier operations. A majority of the information needs were found to be available for such carriers on selected pages of their annual reports to the Interstate Commerce Commission (ICC) for the year 1974.

Having thus determined the source of such information, it was necessary to identify the carriers for which the data would be acquired. There were a total of 196 class 2 line-haul carriers (defined in 1974 as those having less than $5 million in gross revenues) that were neither controlled nor operated by class 1 carriers. Elimination of the larger carrier bias was considered essential in selecting the short-haul carriers for study. Thus, each of the 196 carriers was subjected to a rigorous scrutiny to determine whether data on it were particularly desirable. For several reasons, 26 of the 196 were considered unsuitable for analysis; 17 were involved in passenger operations; 4 did not operate throughout the entire year of 1974 (the latest year for which annual report data were available from the ICC); 1 was a switching road; 1 had evidence of class 1 ownership; and 3 were dropped for other reasons.

Thus, the ICC was requested to provide the pertinent annual report data for 170 carriers for the year ending December 31, 1974. It was found subsequently that the data were available for only 148 of the 170 carriers requested. Figure 1 shows the number of carriers by state for which data were received. The information received was considered as the primary data base for use in the analysis that followed.

It should be acknowledged that some of the additional information needs were fulfilled by reference to American Short Line Railway Guide (1). This source provided a substantial amount of current data on all existing class 2 railroad operations.

ANALYSIS OF THE DATA

This particular phase of the research methodology required that each of the relevant variables be studied intensively and that efforts be made to understand the relationships among various sets of those variables. The overall objective was to produce a number of cost-estimating equations that would be of value when we attempted to predict likely magnitudes of costs associated with the operation of a light-density line independent of class 1 ownership. As previously mentioned, "historical" data were obtained for the operations of 148 class 2 line-haul carriers that were neither owned nor operated by a large carrier. It was hoped that such a study of past experiences would provide significant insight into the potential levels of each of the relevant types of costs.

Multiple regression was the statistical technique selected for developing the equations. As used here, the technique may be best described as one that provides descriptive ability; that is, it allows a study of the linear dependence of one variable on others. In addition, use of this approach permits the efficient computation of quantitative measures that may be used to assess the predictive accuracy of the entire equation as well as measures that evaluate the individual contributions of each independent variable toward an explanation of variation in the dependent variable. The six dependent variables used in this study are shown below.

Variable	Description
TOTMOWS	Total maintenance of way and structures, $000
TOTMOE	Total maintenance of equipment, $000
TRAFFIC	Traffic expenses, $000
TOTTRANS	Total transportation rail line, $000
TOTGEN	Total general expenses, $000
GRANDTOT	Grand total railway operating expenses, $000

More specific information regarding the particular types of expense items included in each is available in the annual reports to the ICC.

Preliminary Considerations

The first major step taken was to construct simple frequency distributions for each of the study variables. Considerable attention was then devoted to a study of those distributions, and the notion of using all data for each of the 148 carriers was reconsidered. A decision was made at that point to eliminate from further consideration all carriers having gross freight revenues in 1974 of $1 million or more.

Because the analysis was aimed at developing cost equations that would be appropriate to the independent provision of service over light-density lines, it was felt that very few of such potential new lines would be capable of generating revenues of such magnitude. The elimination of all carriers having gross freight revenues equal to or greater than $1 million reduced the list of carriers by 43 to a new total of 105.

Examination of the frequency distributions constructed for variables of 105 carriers revealed three additional carriers that were eliminated from further study. One was dropped because its main trackage of 259 km (161 miles) was substantially greater than any of the other carriers; the next longest was 107 km (67 miles). Two other carriers having respective operating ratios (operating costs as a percentage of operating revenues) of 10 and 973 were likewise deleted from further consideration. Such ratios were considered to be rather unusual in light of the range of the 35-315 exhibited by others of the 105 carriers. The remaining portions of the analysis were then based on the resulting sample of 102 class 2 line-haul carriers having operating revenues in 1974 of less than $1 million.

Methodology

Following the initial phases of analysis, attention was devoted to the functional form of the actual cost-estimating equations. Several objectives were necessary to consider. First, each equation should contain a variety of variables that would help to explain variation exhibited by the respective dependent variables. Second, all independent variables should represent measures that can be estimated with some degree of accuracy before an independent carrier operation is implemented. Third, it was desired that the regression analyses for each dependent (cost) variable be conducted in a consistent manner. This was necessary in order to ensure that the various cost-estimating equations all would require similar, perhaps identical in some cases, data inputs. Finally, each of the resulting equations was to be evaluated in terms of its statistical validity.

Five types of independent variables were ultimately selected for inclusion in the multiple regression analyses: carrier location, ownership, main trackage, traffic volume, and one other depending on the particular type of cost being estimated. Each of these will be discussed in turn.

Carrier Location

The geographic location for each of the 102 carriers studied was determined by reference to the map of ICC districts included in Figure 1. Of that number of carriers, 38 were located in the eastern district, 25 in the southern, and 39 in the western. There were other methods of locational segmentation that were considered, but each was either too arbitrary or categorized the carriers into so many areas that any resulting statistical analysis would lose much of its validity.

For purposes of the regression analysis, three dummy variables were constructed for a carrier's location. They were

1. EASTERN: carrier located in eastern district,
2. SOUTHERN: carrier located in southern district, and
3. WESTERN: carrier located in western district.

Ownership

Three different types of ownership were identified: independent, shipper/industry, and government unit. There were 40 carriers owned independently, 57 owned at least in part by shippers and/or local industry, and 5 owned by local or state government units. Ownership status, like location possibilities, required the creation of three dummy variables:

1. INDEP: carrier ownership independent,
2. INDUSTRY: carrier ownership shipper/industry, and
3. GOVT: carrier ownership government unit.

Main Trackage

Early stages of the analysis provided a strong indication that this variable would be helpful in explaining a variety of cost data. The relevant measure computed for each carrier was the total distance of single or first main track plus that of second and additional main tracks. Excluded was trackage associated with passing tracks, crossovers, turnouts, and way and yard switching tracks. The variable included in the analysis was TRAKMAIN, main trackage, in miles. The sample of 102 carriers revealed that distances ranged from 1 to 67, and the mean and median were 14.08 and 10.25, respectively.

Traffic Volume

There are several ways in which this category of variable may be measured, the most appropriate of which is carloads moved. Unfortunately, the ICC does not require that carriers filing annual reports indicate the magnitudes of such a variable. Although such carriers are required to submit data concerning car-miles, it is difficult to transform such a measure into the actual number of loaded freight cars handled without making a perhaps unreliable assumption regarding average length of haul. Since all cars do not necessarily travel the entire length of a given line, an assumption that they did so would introduce an unnecessary bias. To overcome such shortcomings, the total amount of revenue freight carried by the respective railroads in 1974 was used as the measure of traffic volume: TOTTNSRV, total tons revenue freight carried. Although this variable was coded in thousands for input to the computer analysis, its actual value was seen to range from as low as 1000 tons to as high as 1 039 000 tons. The mean revenue freight was 212 592 tons and the median was 120 500 tons.

Other Variable

Depending on the particular dependent variable being investigated, a fifth independent variable was introduced from the following list:

1. XTIESREP: number of ties replaced per mile,
2. LOCMILES: total locomotive unit-miles,
3. NUMCONNS: number of connecting carriers, and
4. ADMIN: administration costs.

Results of correlation analyses indicated that the number of ties replaced per mile of track in 1974 would be helpful in explaining expenditures for maintenance of way and structures, and also for the grand total of railway operating expenses. Such a variable is interesting in that once a predictive equation has been developed value inputs for estimation purposes may be based on the category of track class desired as indicated by Federal Railroad Administration (FRA) track standards. Table 1 provides a summary of such standards for FRA class 1 (10-mph maximum speed) and FRA class 2 (25-mph maximum speed), and normalized track. As can be seen, these three classes will require yearly replacements of 21, 33, and 93 ties/km (34, 54, and 150 ties/mile), respectively, assuming a conservative average tie life of 20 years.

Table 1. FRA track standards and necessary tie renewal rates.

Item	Class of Track		
	1	2	Normalized[a]
Maximum distance between nondefective ties, center to center (in)[a]	100	70	21
Minimum number of nondefective ties/ 39 ft of track	5	8	22
Minimum number of good ties required per mile	677	1083	3000
Average necessary tie replacements per mile per year to maintain track standard by average installed tie life			
10 years	68	108	300
20 years	34	54	150
30 years	23	36	100

Note: Assuming a total of 3000 ties/mile of track.

[a]Track is maintained on a normalized basis when one-half of the useful life of the track components remain. Theoretically, this standard of maintenance will preserve the entire capital investment in perpetuity.

Locomotive unit-miles was used as an additional independent variable in the estimation of costs associated with maintenance of equipment and transportation-rail line. The values of this variable were determined by adding together the number of unit-miles reported for road service and train and yard switching. Once the predictive equations have been developed, an estimate of locomotive unit-miles may be developed for a proposed operation by using the following computational formula:

$$\begin{aligned} \text{LOCMILES} &= \text{unit-miles road service} + \text{unit-miles switching} \\ &= 2 \times L \times F \times 52 + (0.35) \times 2 \times L \times F \times 52 \\ &= (1.35) \times 2 \times L \times F \times 52 \end{aligned} \quad (1)$$

where

L = length of line in miles,
2L = round-trip distance,
F = service frequency (round trips per week), and
52 = weeks per year.

Such an estimate is based on operating frequency, length of line, and an adjustment (0.35) for average switching miles as developed from the data available.

The number of other carriers with which connections were made varied from one to five for the carriers investigated. It was found that 74 of the 102 had only one carrier with which freight was interchanged and that there were only three carriers having more than three connections. This variable was incorporated into the multiple regression analysis that pertained to the estimation of traffic expenses.

Finally, there was strong evidence to indicate that administrative expenditures would help substantially in explaining general expenses. Components of the administrative expense are salaries and expenses of general officers, clerks, and attendants; general office supplies; and legal expenses. Other notable general expenses are insurance and other expenses such as employee health and welfare benefits, pensions, and stationery and printing.

In summary, the above discussions of independent variable categories represent end results of rather extensive preliminary investigations. Although the number of variable candidates that could have been included in the above categories was large, the results of the preliminary analyses strongly indicated that those discussed above were likely to be consistently valuable in explaining variation in each of the respective dependent variables.

Intermediate Results of Regression Analyses

This section presents the empirical results of application of the technique of multiple regression to the task of deriving cost-estimating equations. The general form of the regression model used is

$$Y' = A + B_1 X_1 + B_2 X_2 + \ldots + B_k X_k \quad (2)$$

In this equation, Y′ represents the estimated value for Y, the dependent variable, A is the Y intercept, and the B's are the regression coefficients respective to each independent variable. The functional form implies that, once such a relationship has been derived by empirical analysis, the value of the dependent variable may be estimated in a given case by adding to the constant term (A) the sum of a number of products, each resulting from the multiplication of an independent variable value and the corresponding regression coefficient. The general form may be expressed more concisely as

$$Y' = A + \sum_{i=1}^{k} B_i X_i \tag{3}$$

The independent variables to be included in the intermediate analyses are listed below (the dependent variables were listed previously). Such variables are in the order in which they were discussed in the immediately preceding discussion of methodology.

Variable	Description
EASTERN	Carrier located in eastern district
SOUTHERN	Carrier located in southern district
INDEP	Carrier ownership, independent
INDUSTRY	Carrier ownership, shipper/industry
TRAKMAIN	Main track mileage, miles
TOTTNSRV	Total tons freight carried, $000
XTIESREP	Number of ties replaced per mile
LOCMILES	Total locomotive unit-miles, $000
NUMCONNS	Number of connecting carriers
ADMIN	Administration, $000

Table 2 presents the results of the individual multiple regression analyses for 86 observations. Information relating to the estimation of each respective dependent variable appears in the appropriately labeled column of that table. For example, the first column shows the results of the analysis for which total maintenance of way and structures was the dependent variable, the second for total maintenance of equipment, etc. While the rows are identified by the independent variables of interest, the body of the table includes information in each cell regarding the coefficients that were computed for the values of independent variables in the respective regressions. The lower part of the table provides data of a summary nature pertaining to each equation.

The first of the equations produced will be described in detail to allow more meaningful interpretation of the results achieved. The form of the cost-estimating equation is

$$\begin{aligned} \text{TOTMOWS} = {} & 19.619 - \underset{(10.656)}{8.138}\ \text{EASTERN} - \underset{(11.667)}{9.066}\ \text{SOUTHERN} \\ & - \underset{(20.029)}{26.731}\ \text{INDEP} - \underset{(19.743)}{30.036}\ \text{INDUSTRY} \\ & + \underset{(0.378)}{2.679}\ \text{TRAKMAIN} + \underset{(0.019)}{0.114}\ \text{TOTTNSRV} \\ & + \underset{(0.053)}{0.182}\ \text{XTIESREP} \end{aligned} \tag{4}$$

Keeping in mind that TOTMOWS is expressed in thousands of dollars, any estimate derived from use of this equation should be adjusted accordingly. Referring to this equation and the first column of Table 2, the constant or intercept term is 19.619. The coefficients for the independent variables, EASTERN, SOUTHERN, etc., are -8.138, -9.066, etc., respectively. The numbers in parentheses are the standard errors for each calculated regression coefficient. They are helpful in evaluating whether the coefficient values are significantly different from zero in a statistical sense.

Independent variable coefficients that are in fact significantly different from zero (as evaluated by use of the partial F-value) are appropriately noted in the particular cells of Table 2.

Included for TOTMOWS are those respective to TRAKMAIN, TOTTNSRV, and XTIESREP, all of which were significant at the 0.01 level. The other variable coefficients (for EASTERN, SOUTHERN, INDEP, and INDUSTRY) are not significant, and hence their in-

Table 2. Intermediate results of regression analyses.

Independent Variable	Dependent Variable					
	TOTMOWS	TOTMOE	TRAFFIC	TOTTRANS	TOTGEN	GRANDTOT
CONSTANT	19.619	5.328	-7.289	29.292	4.562	81.505
EASTERN	-8.138 (10.656)	4.565 (7.372)	-0.025 (1.867)	-12.221 (9.611)	2.802 (2.820)	-14.140 (29.712)
SOUTHERN	-9.066 (11.667)	-3.337 (8.051)	4.444 (2.039) SIG.05	-20.943 (10.497) SIG.05	3.262 (3.099)	-24.549 (32.529)
INDEP	-26.731 (20.029)	0.957 (13.740)	5.182 (3.458)	-10.725 (17.913)	-1.579 (5.226)	-57.791 (55.844)
INDUSTRY	-30.036 (19.743)	6.969 (13.565)	1.948 (3.426)	1.839 (17.686)	-1.589 (5.163)	-45.652 (55.047)
TRAKMAIN	2.679 (0.378) SIG.01	-0.383 (0.280)	0.176 (0.069) SIG.05	-1.007 (0.365) SIG.01	0.059 (0.100)	4.634 (1.055) SIG.01
TOTTNSRV	0.114 (0.019) SIG.01	0.028 (0.017)	0.006 (0.003) SIG.10	0.137 (0.022) SIG.01	0.013 (0.005) SIG.025	0.587 (0.052) SIG.01
XTIESREP	0.182 (0.053) SIG.01					0.194 (0.149)
LOCMILES		1.941 (0.254) SIG.01		3.683 (0.331) SIG.01		
NUMCONNS			4.492 (1.231) SIG.01			
ADMIN					1.069 (0.047) SIG.01	
Summary						
R^2	0.613	0.660	0.409	0.864	0.904	0.701
F-value	17.661	21.600	7.723	71.079	105.098	26.067
Standard error	40.923	28.345	7.176	36.954	10.828	114.098
Dependent variable mean	67.695	37.507	6.811	81.034	37.035	230.438

clusion in the equation constitutes an error in specification of the functional form of the equation. Because the results currently being discussed were "intermediate" in nature, there was ample opportunity to make necessary changes before the development of final results.

The lower part of Table 2 gives a variety of information important to an understanding of the statistical validity of the various equations. In the case of estimation of total maintenance of way and structures expenses, there were 86 carriers for which data were used to develop the relationship described above. Although this is somewhat short of the 102 carriers for which data were available, the difference was due to a failure by some carriers (16 to be exact) to report all the necessary information in their annual reports. The R^2 value of 0.613 indicates that the equation was successful in explaining 61.3 percent of the variation in the dependent variable TOTMOWS. The entry in the next row shows the computed F-value of 17.661, which is significant at the 0.01 level, indicating that the set of particular independent variables selected for inclusion was significant in explaining variation in the dependent variable. Finally, the standard error 40.923 is actually the standard deviation of actual values of the dependent variable (TOTMOWS) from the values predicted by the equation. For reference purposes, the mean value of the dependent variable computed for the sample of 86 carriers is shown on the bottom line of Table 2. The table below has been included to show the means and standard deviations for all variables included in the analyses.

Variable	Mean	Standard Deviation
TOTMOWS	67.6953	63.0269
TOTMOE	37.5070	46.5453
TRAFFIC	6.8105	8.9444
TOTTRANS	81.0337	96.1611
TOTGEN	37.0349	33.5025
GRANDTOT	230.4383	199.7320
EASTERN	0.3721	0.4862
SOUTHERN	0.2558	0.4389
INDEP	0.3953	0.4918
INDUSTRY	0.5465	0.5008
TRAKMAIN	14.5465	12.2586
TOTTNSRV	207.7791	246.3794
XTIESREP	96.9397	87.0024
LOCMILES	13.8721	17.5836
NUMCONNS	1.3605	0.7180
ADMIN	26.7046	28.3680

The other columns of Table 2 should be reviewed in a similar manner. The most striking observation to be made is that, overall, only a small number of variable coefficients were statistically significant. In fact, of the 42 coefficients (exclusive of the constant term) appearing in the table, only 16 were significant at the 0.10 level or more. Thus, any attempt to reduce this specification error definitely would affect the consistency of variable inclusion desired among the various equations. The existence of such a number of nonsignificant values required the omission of certain variables (i.e., those least significant) from the final regression analysis.

On the positive side, all equations were significant in their ability to explain dependent variable behavior, and the R^2 values that ranged from 0.409 to 0.904 are generally acceptable for this type of research. It was interesting to note that the dummy variables constructed for location and ownership were not very helpful in general, but that the remaining selected independent variables proved to significantly contribute to the explanation.

In summary, the intermediate results discussed in this section provide sufficient indication that the relationships studied were in fact valid, but that their functional forms should be subjected to a thorough reevaluation. It was necessary to incorporate certain changes to ensure that each equation was specified properly.

Final Results of Regression Analyses

A variety of approaches were used in order to improve the intermediate results. The correlation matrices originally computed for all relevant variable pairs were reviewed, as were the frequency distributions for all variables. It was concluded that the regression results previously discussed had taken advantage of the major types of variables related to the particular dependent variables of interest.

It would have been possible to add more variables of the types already considered, but the possibility of extreme multicollinearity among independent variables was a chief deterrent. If variables highly related to variables already in the equation were added, it would have been almost impossible to separate the influence of each on the dependent variables. R^2 values would have been likely to increase, but the presence of this type of specification error would have provided results with greatly reduced meaning. Other attempts at producing more valuable and efficient relationships included using stepwise (both step-up and step-down) regression procedures and experimenting with various combinations of independent variables in the equations.

The results of the above efforts are shown in Table 3. Although the format of the table is identical to that of Table 2, there are considerably fewer entries, so it is obvious that the number of independent variables included in the final cost-estimating equations has been reduced. Reference to the R^2 value for each equation, however, indicates that such simplification did not reduce appreciably the explanatory ability of the remaining independent variables. A comparison of the F-values and standard errors of the intermediate and final results indicates that the overall statistical validity of each equation has increased and that estimation based on the use of the final results may be made with greater precision. F-values were all significant at the 0.01 level. Each of the particular equations developed will be examined in detail and interpreted as is appropriate.

Total Maintenance of Way and Structures

The form of the equation that may be used to predict this type of expenditure is

$$\text{TOTMOWS} = -14.916 + \underset{(0.371)}{2.738}\ \text{TRAKMAIN} + \underset{(0.018)}{0.114}\ \text{TOTTNSRV} + \underset{(0.052)}{0.198}\ \text{XTIESREP} \qquad (5)$$

As would be expected, larger expenditures are suggested for operations that have longer main trackage, haul more freight, and pursue more intensive policies regarding maintenance of way. Because TOTMOWS is expressed in thousands of dollars, an increase of one main track mile will increase expected expenses by $2738, and the carrying of each additional thousand tons of freight will add $114 (at 1974 price levels). In addition, the setting of track standards at FRA class 1 (average renewal of 34 ties/mile per year) will contribute $6732 (32 × 0.198 expressed in thousands) to the total. FRA class 2 (54 ties) will add $10 692, and normalized

(150 ties) $29 700. It should also be noted that each of the independent variables in the equation was seen to be highly signficant on the basis of partial F-tests.

In terms of overall explanatory ability, the R^2 value indicated that the relationship shown accounted for 59.9 percent of the variation in total expenses for maintenance of way and structures. Although the F-value indicates significance of the equation at the 0.01 level, the practice of deferring maintenance as pursued by some carriers may have resulted in less valid results than would have been obtained otherwise. Also, extraordinary costs incurred by some carriers, particularly with regard to maintenance of structures, surely affected the variation that remained unexplained by the equation. The equation presented, however, should prove to be of value for estimation purposes in its present form.

Total Maintenance of Equipment

The relationship derived for estimating the magnitude of this category of expense contains fewer independent variables:

$$\text{TOTMOE} = 5.888 + \underset{(0.016)}{0.031}\ \text{TOTTNSRV} + \underset{(0.228)}{1.817}\ \text{LOCMILES} \qquad (6)$$

The total revenue freight carried is also included in this equation (as it is in each of those remaining to be discussed). Its significance was measured at the 0.10 level, and its meaning for estimation purposes is that expenditures for maintenance of equipment are expected to increase by $31 for each additional thousand tons carried. The total locomotive unit-miles variable, however, contributes $1817 for every thousand miles to predicted maintenance of equipment expenditures.

Although only two independent variables (and the constant term) were included in this regression, it still explained approximately 64 percent of variation in expenditures for maintenance of equipment. It had been anticipated that data regarding number of locomotives and their total horsepower per line would have contributed significantly to an explanation. Investigations, however, indicated that most of their respective abilities to contribute had been captured by the two variables included in the equation.

Traffic

This particular variable was perhaps the most difficult to explain, largely because it exhibited the least variability of all dependent variables. In addition, expenditures in this category generally represent a very small percentage of total operating costs. The average for 86 carriers was $6811, but about 10 percent of those studied indicated no traffic expense at all. The following equation was developed for estimating this type of expense:

$$\begin{aligned}\text{TRAFFIC} = -5.554 &+ \underset{(1.775)}{4.469}\ \text{SOUTHERN} + \underset{(1.590)}{3.427}\ \text{INDEP} \\ &+ \underset{(0.068)}{0.176}\ \text{TRAKMAIN} + \underset{(0.003)}{0.006}\ \text{TOTTNSRV} \\ &+ \underset{(1.216)}{4.457}\ \text{NUMCONNS}\end{aligned} \qquad (7)$$

The inclusion of two dummy variables, SOUTHERN and INDEP, indicates that both location and ownership type are important variables for prediction of traffic expenses. The relationship suggests that $4469 would be added to the constant term of -$5554 if the line is in the ICC's southern district, while $3427 would be added if the carrier is independently owned.

This information indicates that carriers having such characteristics are likely to spend greater sums for activities such as advertising, soliciting and securing traffic, and preparing and distributing tariffs governing such traffic. It is understandable that carriers that are independently owned would find it necessary to place greater emphasis on such expenditures than carriers owned, for example, by an on-line shipper.

Table 3. Final results of regression analyses.

Independent Variable	Dependent Variable TOTMOWS	TOTMOE	TRAFFIC	TOTTRANS	TOTGEN	GRANDTOT
CONSTANT	-14.916	5.888	-5.554	7.108	5.806	15.318
EASTERN						
SOUTHERN			4.469 (1.775) SIG.025	-15.590 (9.546) SIG.15		
INDEP			3.427 (1.590) SIG.05			
INDUSTRY						
TRAKMAIN	2.738 (0.371) SIG.01		0.176 (0.068) SIG.025			4.830 (1.028) SIG.01
TOTTNSRV	0.114 (0.018) SIG.01	0.031 (0.016) SIG.10	0.006 (0.003) SIG.10	0.150 (0.022) SIG.01	0.012 (0.005) SIG.025	0.588 (0.051) SIG.01
XTIESREP	0.198 (0.052) SIG.01					0.234 (0.144) SIG.15
LOCMILES		1.817 (0.228) SIG.01		3.367 (0.313) SIG.01		
NUMCONNS			4.457 (1.216) SIG.01			
ADMIN					1.076 (0.044) SIG.01	
Summary						
R^2	0.599	0.643	0.407	0.845	0.902	0.694
F-value	40.834	74.776	10.973	149.182	381.442	61.924
Standard error	40.634	28.140	7.101	38.526	10.620	112.532
Dependent variable mean	67.695	37.507	6.811	81.034	37.035	230.438

Two measures of the scope of operations, TRAKMAIN and TOTTNSRV (see Table 3) also have significant explanatory ability. Although their coefficients are respectively smaller in magnitude than those included in other cost-estimating equations, average traffic expenditures are also smaller.

Lines that have greater numbers of connecting carriers also tend to spend more for this expense category. The average figure of $4457 per connection tends to support the notion that those lines having a variety of connecting carriers, and hence a greater range of services to offer, incur greater costs to secure traffic volumes.

Total Transportation for Rail Line

This type of expense is extremely important. The analysis indicated that, for the 86 carriers studied, transportation expense averaged approximately 35 percent of total operating expenses. The explanatory relationship developed was

$$\underset{}{\text{TOTTRANS}} = 7.108 - \underset{(9.546)}{15.590\ \text{SOUTHERN}} + \underset{(0.022)}{0.150\ \text{TOTTNSRV}} + \underset{(0.313)}{3.367\ \text{LOCMILES}} \quad (8)$$

The presence of the dummy variable SOUTHERN was due in large part to the fact that prevailing wages for train and engine employees were substantially lower for southern district roads than for those located in other parts of the country. In 1974, average hourly wages for class 2 operating personnel were $3.54, $4.21, and $4.82 for carriers studied in the southern, eastern, and western districts, respectively.

Also important to consider is that labor costs represent approximately 50 percent of transportation expenditures, and the variable SOUTHERN actually represents a surrogate for such costs. If data had been available for the study carriers regarding the degree of labor organization, perhaps even greater insight would have been provided.

Also contributing to an explanation was total revenue freight carried and locomotive unit-miles, each significant at the 0.01 level. Transportation expenses are estimated to increase by $150/1000 tons freight carried, and by $3367/1000 locomotive unit-miles. The inclusion of these variables provides strong evidence that measures of the scope of operations are extremely important when attempting to explain expenditures for the provision of transportation service.

The relationship derived was quite acceptable. The computed R^2 value of 0.845 and an accompanying high level of significance as measured by the F-value indicate that the equation was responsible for a great deal of explanation of variation in transportation expenses.

Total General Expenses

Because such expenses are largely composed of those related to administration, the following equation was able to explain over 90 percent of variation in the dependent variable relating to general expenses.

$$\text{TOTGEN} = 5.806 + \underset{(0.005)}{0.012\ \text{TOTTNSRV}} + \underset{(0.044)}{1.076\ \text{ADMIN}} \quad (9)$$

The estimation procedure requires that $12/1000 tons of freight and $1076 for each $1000 of administration expense be added to the constant term of $5806. Once again, TOTTNSRV provides an indication of the scope of operations, implying that more intensive operations incur greater levels of general expenses.

Grand Total of Railway Operating Expenses

Although estimates derived from the preceding five equations could be added together to construct an estimate of the grand total, it was felt that separate treatment of the total would provide information of additional interest. The use of multiple regression analysis resulted in the following equation:

$$\text{GRANDTOT} = 15.318 + \underset{(1.028)}{4.830\ \text{TRAKMAIN}} + \underset{(0.051)}{0.588\ \text{TOTTNSRV}} + \underset{(0.144)}{0.234\ \text{XTIESREP}} \quad (10)$$

Incorporated are variables related to main trackage, total revenue freight carried, and number of ties replaced per mile per year. Each of these has been discussed previously with respect to estimation equations for individual types of costs.

The ability of this equation to explain the grand total of railway operating expenses was quite satisfactory, as measured by the R^2 value of 0.694 and an associated high level of significance. Additional variables could have been included, but the possibility of introducing extreme multicollinearity kept those included to a minimum.

SUMMARY

The preceding describes the development of equations appropriate to the task of estimating both individual types and total of railway operating expenses for class 2 railroad operations. Each relationship shown was seen to be statistically significant and properly specified with regard to particular variables included. It is strongly suggested that the actual use of the estimating procedures be accompanied by a keen sense for special characteristics of individual lines, implying that unusual costs are sometimes incurred. If such a conscientious effort is pursued, the equations developed are likely to provide significant insight into the expected magnitudes of the costs of operating a light-density line independent of class 1 ownership.

ACKNOWLEDGMENTS

This document is based on research supported by the Program of University Research of the U.S. Department of Transportation. The contents reflect our views, and we only are responsible for the facts and accuracy of the data presented here. The contents do not necessarily reflect the official views or policies of the U.S. Department of Transportation.

REFERENCE

1. E. A. Lewis. American Short Line Railway Guide. Baggage Car, Strasbourg, PA, 1975.

Publication of this paper sponsored by Committee on State Role in Rail Transport.