THE COST OF OPERATING BUSES IN U.S. CITIES

By Hoyt G. Wilson*

INTRODUCTION
The expansion and revitalisation of the U.S. transit industry in recent years has naturally been accompanied by an increase in transit planning. A necessary part of transit planning is to estimate the costs of proposed transit systems, whether those systems be entirely new or extensions of existing ones. This paper deals with estimating the costs of operating buses in U.S. cities. Capital costs and fixed guideway modes are not considered. The cost models developed are intended for use rather late in the planning process. Model inputs include operating parameters, such as bus miles and bus hours (operated in a year), which are available only after routes and schedules have been planned in some detail.

Cost models—based on data taken from a number of U.S. transit properties—are developed for each of five components of total operating cost. Model outputs are in terms of cost per unit, an output unit being either a bus mile or a bus hour. If costs per mile or per hour did not vary substantially from one property to the next, industry averages would suffice and no models would be necessary. Unfortunately, the observed variations are great, so that some refinements over raw averages are called for in order to relate unit costs to the circumstances of operation. The objective of the paper is to present a practically useful tool for making these cost forecasts.

COST RELATIONSHIPS
The results of some earlier studies of bus operating costs will be discussed briefly here. The reader who is interested in more detail is referred to the research report on which this article is based [16]. Most of the important studies have relied heavily upon data supplied by the American Public Transit Association1 (APTA) in an annual publication titled Transit Operating Report. The Transit Operating Report gives a summary of financial and operating statistics for each member bus system

* Visiting Assistant Professor, Faculty of Commerce and Business Administration, University of British Columbia. This research was carried out at the Pennsylvania Transportation Institute of The Pennsylvania State University and was financed by the Urban Mass Transportation Administration. Thanks are due to Trevor Heaver, David Maister, and an anonymous referee for their helpful comments on earlier drafts of the paper.

1 APTA was formed by the merger of the American Transit Association (ATA) and the Institute for Rapid Transit in the fall of 1974. To avoid confusion the APTA designation will be used throughout this article, even though much of the information cited was published by the ATA before the merger.
that supplies data to APTA. Entries include operating expenses (for the year) broken down into a number of categories, and various operating statistics such as bus miles operated, gallons of fuel consumed, and passengers carried. Table 1 shows the breakdown of operating expenses, and also the distribution of costs among the various expense categories. The figures given are averages calculated from the 1973 Transit Operating Report.

Although several past studies have developed formulas which give operating cost as a function of other variables, the interest of most of the researchers was not in cost prediction per se, but rather in studying the structure of the cost function. For example, Miller [6,7], Veatch [13], and Wells et al. [15] carried out studies concerned largely with detecting economies or diseconomies of scale in bus transit operations. The approach in all three studies was to fit regression equations to cross-sectional APTA data. The approach of Foster [4,5] was similar to that of Miller and Veatch, but he carried the analysis further by examining the changes in the cost functions over time.

Two other studies [1,3] were based on data from individual systems. These did consider individual cost components, but in each case the model parameters were found to be specific to the particular system under study. The formulas were of little value for cost estimation in general.
In summary, the earlier work that was based on industry cross-sectional data concentrated on modelling total operating cost rather than individual components. The studies that did deal with components of operating cost were based on data for particular cities, so their models were not applicable to other systems. The primary contribution of this past research to the current study is a group of potentially useful predictor variables which were found to correlate with operating cost.

Since the primary focus of the current study is on forecasting rather than analysis of cost structures, no formal hypotheses will be stated and tested. It is to be observed, however, that hypotheses are implicit in the choice of explanatory variables for inclusion in the models. With each proposed variable is associated a tacit suggestion that that variable influences costs, and the complementary null hypothesis that it does not. The variable screening process may be viewed as a series of tests of such hypotheses.

The proposed variables fall into two general classes: (1) those which describe the bus system itself and (2) those which describe the environment in which the system operates. Past studies have generally found both kinds of variables to have significant influences on operating costs. One system descriptor which is of special political interest is system ownership (whether public or private). The typical private bus system in the U.S. suffered through an extended period of declining revenues and increasing costs, with resultant cost-cutting in every possible quarter. Deferral of maintenance of equipment is a natural result of such a process. In contrast, public takeover of a private bus property has often been accompanied by an infusion of new funds—capital equipment grants from the federal government and operating subsidies from local governments. If these observations are valid, one might reasonably expect to find a relationship between ownership and operating costs.

On the basis of data for the years 1960 and 1970, Foster [5] found that ownership had no significant effect on total operating cost. Wells et al. [15], however, in a study based on 1960 and 1968 data, found that private ownership was associated with lower total operating cost. None of the other works cited above considered system ownership at all. The Wells result would seem to be consistent with the cost-cutting syndrome described above. In the face of such meagre and mixed evidence, however, no conclusions can be reached with any confidence. A priori expectations of the effects of other explanatory variables on operating costs will be discussed when the variables are defined.

All the studies (including the work to be reported here) ignore the important, but difficult to measure, factor of service quality. Even if one accepts the proposition that privately-owned systems operate at lower per-mile costs, it does not necessarily follow that private ownership is to be preferred to public. Differences in service offered must certainly be considered as well.

This study will be concerned with forecasting individual components of operating cost. Because the important past studies dealt primarily with total operating cost, the value of their results as guides for this investigation is diminished. Since any variable is expected to affect various cost components in different ways, the individual effects are masked in an analysis which deals only with total cost. If the correlations have opposite signs, one might detect no relationship at all with total cost, even when effects on individual components are strong.
THE COST OF OPERATING BUSES IN U.S. CITIES

Hoyt G. Wilson

STUDY STRUCTURE

All the modelling efforts mentioned earlier had one thing in common: the cost being modelled, whether it was total operating cost or a component thereof, was always measured in monetary units. This is a natural approach, since most accounting records are in dollars, and monetary data are the easiest to obtain. There is a drawback, however, to modelling dollar expenditures directly. As price levels and wage rates change, the model parameters necessarily change. It is possible conceptually to dissect many elements of operating cost into a quantity of real resource and a rate or price for the use of that resource. For example, drivers' wages may be thought of as the product of the number of man-hours worked times the applicable wage rate. Expenditures on fuel are the product of the number of gallons consumed times the price per gallon.

The approach in the work described here will be, where possible, to model real resource usage, leaving the task of applying the appropriate rate to the user of the model. This has the advantage of producing cost models which, if they are valid at the outset, may remain valid over a period of time, even in the face of considerable inflation. The advantage accrues from the fact that such models purport to represent underlying relationships that are basic to a particular technology and mode of operation, but which are relatively unaffected by changes in price levels. Of course in the long run levels of resource usage are expected to adjust to changes in the relative resource prices, although the extent of the adjustments which can be made in practice may be limited.

The scheme will not be applicable to all cost components, for two reasons. First, for many types of expenditures it is difficult to conceive of a measurement unit other than dollars. Examples of this are advertising expenses and legal fees. Second, even in cases in which one can identify meaningful units for measuring resource usage, acquisition of data in those units may be difficult or virtually impossible. Although many transit systems may record the desired quantities for internal use, extracting and aggregating the data for modelling purposes is often time-consuming and expensive.

As stated earlier, the models will be framed in terms of cost per mile or cost per hour. It has been observed by others that such scaling of the data prior to statistical analysis reduces heteroscedasticity problems (see, for example, Miller [6] or Wabe and Coles [14]).

The system of operating cost components used in this study will be consistent with the APTA definitions (Table 1). To report costs in the APTA format, bus properties must combine accounts from their own accounting systems. Individual properties use a variety of charts of accounts, the commonest of which are those defined by the Interstate Commerce Commission (ICC) [8] and the American Electric Railway Accountants' Association [2] (often referred to as “the ATA system”). There are minor inconsistencies in the account definitions from one chart of accounts to another; hence the data reported to APTA by properties using different accounting systems are not strictly comparable. However, such inconsistencies are judged to be relatively minor when compared with the variations that arise from liberties in the interpretations of account definitions at individual properties. A standard chart of accounts offers a convenient basis for an accounting system, but many properties
may make significant alterations to the original account definitions to suit their particular needs. Especially in small systems, assignment of expenses among accounts may reflect the personality of the bookkeeper. Discovery of this phenomenon resulted from observing extreme outliers in regression runs, and following up by telephone conversations with bookkeepers at the bus systems in question.

A large part of the data used in the study was provided by APTA. The most important source was the Transit Operating Report, but other APTA publications were used as well: these will be described later. Primary operating data on bus systems were also collected by mail questionnaires and telephone interviews. Only APTA member bus systems were included in the survey, since the survey data served as supplements to those found in the APTA reports.

The APTA membership represents the entire range of sizes of bus systems, from very large to very small; but the distribution is clearly distorted toward the large side. That is, the proportion of larger properties among the APTA membership is greater than the proportion of larger properties in the total population of U.S. bus systems. This is certainly undesirable, and may constitute a ground for questioning the general validity of results based on the data. Indeed, the very fact of membership in APTA may mark a transit system as atypical. What excuse can be given, then, for selecting a sample that is so likely to be unrepresentative of the population? The answer is simply data availability. Data on U.S. bus systems are notoriously difficult to come by. This unhappy state of affairs for researchers is largely attributable to the neglect under which the urban transit industry suffered for so many years. Despite their shortcomings, then, the records compiled by APTA have the distinction of being the only available source of data of any consequence on the U.S. bus transit industry.

Besides APTA reports, some additional secondary data sources were used. Data on population were extracted from publications of the U.S. Bureau of the Census, and on weather from those of the National Oceanic and Atmospheric Administration.

Conceptual considerations relating to the identification of resource units for various categories of expenses, as well as practical considerations of data availability, led finally to modelling five separate elements of operating cost: driver wages, bus repair labour, fuel, oil, and all other. Models for each of the first four are in real resource units. The last is simply the remainder of operating cost that is left over after accounting for the first four: it is that portion of operating cost which, for one reason or another, could not be modelled in terms of real resource usage. The units of the last model, then, are dollars. Each of the five models will be discussed in turn in later sections.

**METHODOLOGY**

As in the past studies, the statistical technique that served as the workhorse in this research was least squares regression analysis. For variable selection, a stepwise computer algorithm was used which steps variables both into and out of the equation based on values of the partial F statistic.

The selections made by the computer program were not accepted blindly as the
THE COST OF OPERATING BUSES IN U.S. CITIES
Hoyt G. Wilson

"best" models, however. Human intervention was used, sometimes forcing variables into the models, in order to allow for three additional considerations:

1. *Multicollinearity.* In the interest of the efficiency of parameter estimates, models with highly collinear independent variables were avoided where possible.

2. *Parsimony.* Especially where relatively few data points were available, a model with few variables was considered preferable to one with many, even at the cost of decreasing the coefficient of determination.

3. *Intuition.* A model that seemed reasonable, based on personal preconceptions of what factors ought to influence a particular cost component, was favoured over one that was less intuitively pleasing.

In short, the process of developing the model, as carried out in this study, embraced significant portions of both art and science.

Incompleteness of data complicated the process of selecting variables. The approach that was used was to identify a tentative model based on a nucleus of points (bus systems) with relatively complete data. Such a model typically contained only a few of the many possible independent variables. This model was then refitted to the larger group of properties which had complete data on those particular independent variables. In some cases the model did not fit the expanded sample at all well; this called for a return to the small sample of complete data points and consideration of alternative models. Through this process, each of the possible variables was considered, at some point, for inclusion in the final model.

Although the central focus of the study is on forecasting, it is natural to speculate on the underlying causes of individual model coefficients. However, most of the models developed in this study are based on rather limited samples, and there is thus considerable doubt whether the coefficients are capable of being interpreted. They may best be regarded as suggestive of the true cost structures. The discussion accompanying each of the ensuing statistical analyses is presented in this spirit.

In a similar vein, it should be recognised that the models developed here are intended to be used as planning tools, but not as design tools. Once a bus transit system is designed, they should do a reasonably good job of predicting the expected operating costs. However, it would be a serious mistake to use the coefficients in the forecasting models as a basis for making cost tradeoffs in the design of the system.

In the presence of multicollinear regressors, even in a good forecasting model, the regression coefficients may be poor estimates of the true marginal costs associated with individual regressors. Further, even if the parameter estimates were accurate, the statistical model does not imply a causal relationship; there is therefore no certainty that adjusting the value of an independent variable would change the costs in proportion.

The remainder of the article will be devoted to descriptions and analyses of the forecasting models developed for the five components of bus operating cost.

**Driver Wages**
The objective of modelling drivers' wages, the largest component of operating cost, is to provide the transit planner with a tool with which he can predict the level of expenditure for driver wages in a proposed bus system. It must be kept in mind that
all the values that go into the forecasting formula should be readily available. In other words, the model user should not have to forecast the independent variables.

A natural product of route and schedule planning for a transit system is the expected total number of bus hours. "Bus hours", as used here, if often defined as platform hours, or vehicle hours in revenue service, including layover time and travel time to and from the garage before and after runs. The quantity to be forecasted, then, is the number of driver pay hours per platform hour. Pay hours could be termed equivalent hours paid for. Thus an hour worked at time-and-a-half represents 1.5 pay hours. The ratio of pay hours to platform hours varies from system to system over a range of about 1.0 to 1.5. The variations were assumed to arise largely from differences in provision for paying overtime and assigning work time credits.

A sizeable group of properties have reported to APTA, on a standard form, a summary of their driver union contracts. These summaries indicate the presence or absence of various overtime provisions, driver wage rates, provisions for paying for time not worked, and information on the distribution of run spreads (the elapsed time from the beginning of the first piece of work to the end of the last piece of work in a day, including dead time if the run is split). A host of independent variables — many of the dummy (zero-one) variety — were synthesised from the information on the union contract summaries; these are listed in the Appendix.

Data on the dependent variable — the ratio of pay hours to platform hours — proved to be more difficult to obtain. It required the combining of data from two different reports. First, the average wage rate was estimated as follows:

$$\bar{W} = \alpha W_{\text{max}} + (1 - \alpha)W_{\text{start}}$$

$$\alpha = e^{-R/60}$$

where $W_{\text{max}}$ is the maximum driver wage rate,
$W_{\text{start}}$ is the starting driver wage rate, and
$R$ is the number of months required to reach the maximum rate.

Values for $W_{\text{max}}$, $W_{\text{start}}$, and $R$ are given in the contract summaries. The formula for $\alpha$ was established arbitrarily in order to weight the upper and lower limits of the wage range. The analysis was not very sensitive to values of $\alpha$, since the spread between $W_{\text{start}}$ and $W_{\text{max}}$ was typically quite small (of the order of 15c per hour).

Total driver wages paid during the year and platform hours operated were extracted from the Transit Operating Report. The ratio of interest was then calculated as:

$$Y = (D/\bar{W})/P$$

where $D$ is the total driver wages paid in dollars,
$\bar{W}$ is the average rate as estimated above, and
$P$ is the number of platform hours operated.

The driver wages, $D$, includes overtime pay, but does not include vacation pay, sick pay, or other benefits.

For any property that had submitted summaries of more than one driver union contract, only the most recent version was retained. They bore dates from 1967
through 1974. Issues of the *Transit Operating Report* were available for the years 1969 through 1973. Many properties were represented in one of the two sets of reports but not in the other. The exercise of matching data from the two sources finally yielded 33 useable points (bus systems).

The forecasting model developed from these data is summarised in Table 2. The first independent variable, \(X_1\), reflects the fact that some contract provisions grant "special" overtime allowances, such as paying overtime for work performed on Sundays or on the driver's regular day off. The variable is set equal to one if overtime is paid on overtime when the special provisions overlap with the regular provisions. For example, suppose that a driver has already worked forty hours during a week, entitling him to time-and-a-half for any further hours worked. Now if he is called in on his day off and the contract allows time-and-a-half for that, he will be paid double-time if the overtime-on-overtime clause is in effect. The coefficient of the variable \(X_1\) has the expected positive sign.

The second variable, \(X_2\), probably reflects labour scheduling efficiencies. The positive coefficient indicates that as the number of drivers per bus hour increases, the number of pay hours per bus hour also increases. This may be due to the daily and weekly guaranteed minimums specified in many contracts. As more extra men are added, the work is spread among more drivers and more drivers qualify for the minimum pay.

The third variable, \(X_3\), is an estimate of the average number of spread overtime
minutes per run. Spread overtime is paid for any work performed beyond a specified spread limit. For example, suppose that the spread limit is eleven hours, and a particular run begins at 7 a.m. All work after 6 p.m. on that run qualifies for spread overtime. The contract summaries give distributions of run spreads by one-hour intervals. Say a particular property has twenty runs with spreads of twelve to thirteen hours. It is assumed that the average spread in that group is 12.5 hours, so that the average number of spread overtime minutes per run (given an eleven-hour spread limit) is ninety for those twenty runs. For the complete distribution of run spreads, the overall average number of spread overtime minutes per run is calculated for all runs. As expected, the dependent variable increases as this spread index increases.

The last variable, $X_4$, is simply a dummy variable indicating whether the contract provides for any special time allowances in calculating daily and weekly overtime. For example, fifteen minutes per day for completion of a daily report may be added to platform hours for purposes of calculating overtime. The positive coefficient indicates that systems with such provisions tend to have a higher ratio of pay hours to platform hours.

DEFINITIONS OF VARIABLES
The development of each of the four remaining models was based on the same set of candidates for independent variables. It is convenient, therefore, to define the variables here in a single location. Each variable in the list is assigned a mnemonic acronym by which it will be identified in later references.

**POPDEN**, Population Density. (City population/1000)/(city area in square miles).

**GROWTH**, City Growth Index. A weighted index of the population growth in the five decades since 1920:

$$\sum_{i=2}^{6} W_i \max \{ 0, P_i - P_{i-1} \}$$

$$\max_i (P_i)$$

where $P_1, P_2, \ldots, P_6$ are the city population levels in 1920, 1930, \ldots, 1970, and

$W_2 = 0.1523$ $W_5 = 0.3678$

$W_3 = 0.2317$ $W_6 = 0.4715$

$W_4 = 0.2851$ (see explanation below).

**JANTEM**, Mean January Temperature. (Mean January Temperature in degrees F)/10.

**JULTEM**, Mean July Temperature. (Mean July temperature in degrees F)/10.


---

3 The only exception is that the last model, "all other costs", did not embrace the vehicle-specific variables, such as seating capacity and bus age.
SNOW,  *Annual Snowfall.* (Annual snowfall in mm)/100.

HEATDA,  *Heating Degree Days.* Annual heating degree days based on deviations from 18.5° C.

COOLDA,  *Cooling Degree Days.* Annual cooling degree days based on deviations from 18.5° C.

PRIVATE,  *Privately Owned.* Equal to 1 if the property is privately owned, 0 otherwise.

SMALL,  *Small System.* Equal to 1 if the annual bus miles operated is less than 1.5 million, 0 otherwise.

BIG,  *Large System.* Equal to 1 if the annual bus miles operated is greater than 10 million, 0 otherwise.

MIPBUS,  *Miles Per Bus.* (Bus miles/10,000)/active buses.


RTM11,  *One way Route Miles.* One way route miles/100.

RTM12,  *Round Trip Route Miles.* Round trip route miles/100.

BUSOWN,  *Buses Owned.* Buses owned/100.

BUSSACT,  *Active Buses.* Active Buses/100.

PEAKBS,  *Peak Buses.* Buses in normal weekday peak schedule/100.


EMPLOY,  *Employees.* Number of employees/100.

BUSMI,  *Bus Miles.* Bus miles/10^3.

SRBM,  *Square Root of Bus Miles.* (BUSMI)^1/2

BUSHRS,  *Bus Hours.* Bus Hours/10,000.


SLOWNS,  *Slowness.* 100x hours/bus miles.

WAGE,  *Wage Rate.* Top driver wage rate as of 1 July in dollars per hour.

AREA,  *City Land Area.* Land area of city in square miles.

FREQ,  *Frequency.* 10 x active buses/one way route miles.

BUSAGE,  *Bus Age.* Age of the bus in years.

CAPAC,  *Seating Capacity.*

V6,  *V-6 Engine.* Equal to 1 if the bus has a V-6 engine, 0 otherwise.

(VMost buses in the study had "straight 6" diesel engines.)

V8,  *V-8 Engine.* Equal to 1 if the bus has a V-8 engine, 0 otherwise.

4CYL,  *4 Cylinder Engine.* Equal to 1 if the bus has a 4 cylinder engine, 0 otherwise.

GAS,  *Gasoline Engine.* Equal to 1 if the bus has a gasoline engine, 0 otherwise.

FLXIBL,  *Flxible Bus.* Equal to 1 if the bus was made by Flxible, 0 otherwise.

AMC,  *AMC Bus.* Equal to 1 if the bus was made by AMC, 0 otherwise.

TWIN,  *Twin Coach.* Equal to 1 if the bus was made by Twin Coach, 0 otherwise.

OTHER,  *Other Make.* Equal to 0 if the bus was made by GMC, Flxible, AMC, or Twin Coach, 1 otherwise.

AIR,  *Air Conditioned.* Equal to 1 if the bus is air conditioned, 0 otherwise.

AIRDAY,  *Air Conditioning Days.* The product of the variables AIR and COOLDA.
Each of the variables has been scaled appropriately so that the regression coefficients will all be of the same order of magnitude. This practice reduces the danger of serious rounding errors in the computer calculations.

The city growth index, GROWTH, is an extension of a "city age" variable first suggested by Miller [6]. It is an index of the pattern of population growth of the city between 1920 and 1970. The five ten-year absolute growth figures are weighted by the average per capita motor vehicle registrations in the United States during the corresponding periods (\(W_i\) through \(W_5\) above). The index is normalised by dividing by the maximum reported population of the city. Only decades of growth are represented in the index; periods of population decline are omitted. A large value of the index corresponds to a "newer" city which has experienced a large share of its growth more recently—particularly during periods of widespread use of automobiles. Such a city is expected to be less compact, with less intensive land use. Also, since it was built with the automobile in mind, traffic congestion should be less severe and vehicle operation should be generally easier.

The population density (POPDEN) and slowness (SLOWNS) variables represent further attempts to measure similar attributes. SLOWNS is the reciprocal of schedule speed. It is conjectured that a slower schedule is a result of heavier traffic and/or more frequent stops to load and discharge passengers—conditions which are associated with more costly operation on a per-mile basis. As expected, GROWTH is negatively correlated with both POPDEN and SLOWNS (correlation coefficients —0.52 and —0.80, respectively).

All the variables that describe the transit system itself are for the year 1973. POPTOT, POPDEN, and AREA are based on the 1970 Census of Population [11]. JANTEM, JULTEM, and PRECIP are actual measurements for the year 1972 [9]. SNOW, HEATDA, and COOLDA are actual measurements for the year 1973 [12].

BUS REPAIR LABOUR

The second cost category, bus repair labour, comprises only a part of the APTA item labelled Repairs to Revenue Equipment (see Table 1). The remainder of that account is comprised of non-labour repair costs—largely parts.

Data on bus repair labour, in man hours, are not available from any secondary source. Thus a direct industry survey had to be relied upon. Most systems record hours worked in one form or another, but aggregating the data as desired (annual total) is often very difficult and expensive. For example, labour may be accumulated weekly by individual buses, or garage janitorial labour may be combined with mechanics' labour. Managers were often understandably reluctant to donate staff members' time to do the necessary data retrieval, collating, and accumulating. Some agreed to provide the information but simply never got around to it. Every one of the 60 systems which had contributed cost data to the Transit Operating Report was contacted, but, as a result of these practical difficulties, the sample of responses was of modest size (20). There were no apparent biases in the sample, but, as always, one could only hope that the omission of non-responses from the data did not detract too seriously from the validity of the analysis.

At the outset it was felt that vehicle descriptors, such as seating capacity and
engine configuration, may be important determinants of the need for repair labour. It would be reasonable, then, to consider each fleet\(^1\) as a data point rather than to combine all the fleets in a system. Unfortunately this approach was not feasible, since repair labour could (in most cases) only be obtained by total system.

To accommodate this data format, an aggregate model had to be devised.

Define the following variables:

- \( h_i \), repair man hours for bus \( i \);
- \( H \), total repairs man hours for all buses in the system;
- \( m_i \), miles operated by bus \( i \) (in thousands);
- \( M \), total bus miles for all buses in the system (in thousands);
- \( X_{ij} \), the value of the \( j \)th independent variable for bus \( i \);
- \( B_j \), the coefficient of the \( j \)th independent variable;
- \( N_i \), the number of the buses in the system.

A model of the following form was assumed for bus \( i \):

\[
h_i/m_i = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i, \quad i = 1, 2, \ldots, N,
\]

which may be rewritten

\[
h_i = m_i \beta_0 + \sum_{j=1}^{p} m_i \beta_j X_{ij} + m_i \epsilon_i.
\]

Summing over all buses in the system,

\[
\sum_{i=1}^{N} h_i = \beta_0 \sum_{i=1}^{N} m_i + \sum_{j=1}^{p} \beta_j \sum_{i=1}^{N} m_i X_{ij} + \sum_{i=1}^{N} m_i \epsilon_i,
\]

and dividing through by total system mileage,

\[
H/M = \beta_0 + \sum_{j=1}^{p} \beta_j \sum_{i=1}^{N} m_i X_{ij} + \sum_{i=1}^{N} m_i \epsilon_i / M,
\]

(1)

gives a form (1) which allows parameter estimation with the available data.

Define

\[
X_j = \sum_{i=1}^{N} \frac{m_i}{M} X_{ij}.
\]

Many of the fleet-descriptor variables are dummy (0 or 1) variables. Say, for example, that \( X_{na} \) equals 1 if bus \( i \) is a GMC bus and 0 otherwise. The variable \( X_s \) for the system is then the weighted fraction of fleets in the system that are manufactured by GMC. The weights are proportional to the bus miles operated by each fleet.

Any system- or city-descriptor variable is constant over all fleets in the system, so that

\[\text{1 The term "fleets", as used here, refers to a group of buses of the same make, model, and year of manufacture. The word "series" is often used with the same meaning. Thus there is typically more than one fleet or series of buses in a single system.}\]
\[ X_i = X_{ij} \sum_{i=1}^{N} \frac{m_i}{M} = X_{ij}. \]

That is, such variables (say, snowfall or city age) go into the regression analysis unchanged. Equation (1) becomes

\[ H/M = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \frac{1}{M} \sum_{i=1}^{N} m_i \epsilon_i. \]  

(2)

Since the \( m_i \) are constants,

\[ E \left[ \frac{1}{M} \sum_{i=1}^{N} m_i \epsilon_i \right] = \frac{1}{M} \sum_{i=1}^{N} m_i E[\epsilon_i] = 0, \]

and

\[ \text{Var} \left[ \frac{1}{M} \sum_{i=1}^{N} m_i \epsilon_i \right] = \frac{1}{M^2} \sum_{i=1}^{N} m_i^2 \text{Var}[\epsilon_i] \]

\[ = \sigma^2 \frac{\sum_{i=1}^{N} m_i^2}{M^2}. \]  

(3)

Thus the residuals have zero expectation but non-constant variance. In such a case, ordinary least squares still produces unbiased parameter estimates, but the estimates are not minimum-variance. Since the relative magnitudes of the variances are known (approximately) and the covariances are zero, application of weighted least squares is straightforward. For each observation (each property), the dependent variables and all independent variables must be multiplied by the factor

\[ \sqrt{\frac{M^2}{\sum_{i=1}^{N} m_i^2}} \]

for that system. Ordinary least squares can then be applied to the transformed data to obtain efficient parameter estimates. This was the procedure followed in modelling repair man hours per bus mile.\(^4\) (Note that the exact form of the dependent variable is repair man hours per thousand bus miles.)

Since the \( m_i \) (miles for bus \( i \)) in equation (3) were not known, the assumption was made that each bus in a fleet operated the same number of miles. The weighting factors were calculated accordingly for each system and applied to the data before running regressions. This tended to weight the larger systems more heavily, since larger systems typically operate more buses than small ones do.

The final model for bus repair labour is given in Table 3. The negative coefficient for SRBM (square root of bus miles) may be indicative of some economies of

\(^4\) In applying weighted least squares, one must exercise care in the handling of the intercept. The constant-valued regressor corresponding to the intercept must be weighted along with the rest of the data.
scale in this particular cost item. The model also suggests that privately-owned systems employ more repair labour per bus mile than do public systems. This finding is consistent with the earlier description of the financial crunch experienced by private systems over a period of years, and the fact that federal subsidies have made capital relatively less expensive for public systems. Older, poorly maintained vehicles are expected to require more repairs. There also may be a tendency to opt for the apparent economy of "fixing", rather than replacing, faulty parts. For example, in one small private system, mechanics chose to tape leaky cooling system hoses rather than to replace them. Over time, such a policy amounts to a substitution of labour for capital. Lastly, repair labour requirements appear to increase with increasing snowfall, as would be expected.

The conspicuous absence of other variables which are expected to affect repair labour is explained largely in terms of correlations with variables already in the model. For example, BUSAGE undoubtedly has some influence on the need for repair labour, but BUSAGE is correlated with PRIVAT (correlation coefficient 0.38), so that the latter variable accounts for much of the effect of running older buses.

FUEL

Data on fuel usage were collected in a direct survey along with the bus repair man hours. Both fuel usage (in gallons) and bus miles were given for each series of buses (rather than for system totals), thus providing a much greater sample from which to draw inferences. The dependent variable was gallons of fuel per ten bus miles. Of course the common way of thinking of fuel usage is in miles per gallon; but gallons per ten bus miles was chosen for modelling purposes for the sake of consistency with the framework of the other resource models, which are all in terms of units of real resource per unit of output.

Building from an underlying model, similar to that in the previous section, in which the variance of the disturbance term for each bus is \( \sigma^2 \), and assuming equal mileage for each bus in a fleet, leads to a fleet model with variance given by
where \( N \) is the number of buses in the fleet. Thus, in applying weighted least squares estimation, the data for each fleet must be weighted by the factor \( \sqrt{\frac{1}{N}} \).

The fuel consumption model is given in Table 4. The negative coefficient associated with GROWTH is consistent with the interpretation of the growth index given earlier; that is, fuel mileage is better in newer, less compact cities. The positive coefficient of PRIVAT indicates that buses in privately-owned systems consume more fuel per mile than those in public systems. This finding is in agreement with the earlier speculation regarding deferral of maintenance in private systems. Signs of other coefficients are as would be expected: increased bus age, increased bus size, gasoline engines, and air conditioning are all associated with increased fuel consumption.

LUBRICATION OIL

As for repair labour, oil consumption data were available only by system—not by fleet. Many properties do record oil consumption by fleet, so that it could have been collected at that level along with fuel consumption data. Oil is such a small portion of operating cost, however, that the added burden on the personnel filling out the questionnaires did not seem to be justified. It was difficult enough to collect the data on mileage, fuel, and repair labour. The source of oil consumption data, therefore, was the Transit Operating Report.

The form of the dependent variable was gallons of oil per thousand bus miles. The aggregating of the data for parameter estimation was exactly the same as in the case of repair labour, so the same weighting scheme was applied in an effort to achieve constant variance.

The oil consumption model is summarised in Table 5. The first two coefficients suggest that buses in private systems consume oil at a lower than average rate and that those in small systems consume oil at a greater than average rate. The underlying reasons are not clear. A factor which clouds the picture is that oil is consumed in two ways: the first is actual consumption by the engine due to leakage or burning; the second is in regularly scheduled oil changes. The frequency of oil changes is largely a matter of managerial discretion; thus the variation in frequency could mask the differences in actual engine oil consumption, which are likely to be a function of engine age and driving conditions.

The negative coefficient of MIPBUS (miles per bus) indicates that more heavily utilised buses require less oil per mile. This would be expected if oil is changed after a maximum time interval irrespective of mileage. The coefficient of SLOWNS suggests that buses operating on a slower schedule (more hours per mile) use more
THE COST OF OPERATING BUSES IN U.S. CITIES

Hoyt G. Wilson

TABLE 4
Forecasting Model for Fuel Consumption in Gallons per Ten Bus Miles

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t Statistic</th>
<th>Probability Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.7197</td>
<td>-5.315</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GROWTH</td>
<td>-1.581</td>
<td>2.708</td>
<td>0.008</td>
</tr>
<tr>
<td>PRIVAT</td>
<td>0.0883</td>
<td>3.918</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BUSAGE</td>
<td>0.0618</td>
<td>9.514</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAPACITY</td>
<td>0.6569</td>
<td>5.216</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GAS</td>
<td>0.2118</td>
<td>2.651</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Sample size = 142
$R^2 = 0.924$
Adjusted $R^2 = 0.921$

$F$ statistic for equation = 771.8

TABLE 5
Forecasting Model for Gallons of Oil per Thousand Bus Miles

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t Statistic</th>
<th>Probability Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.5486</td>
<td>-3.402</td>
<td>0.004</td>
</tr>
<tr>
<td>PRIVAT</td>
<td>-1.1310</td>
<td>3.081</td>
<td>0.007</td>
</tr>
<tr>
<td>SMALL</td>
<td>1.1484</td>
<td>-4.261</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MIPBUS</td>
<td>0.0812</td>
<td>6.794</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Sample size = 21
$R^2 = 0.869$
Adjusted $R^2 = 0.837$

$F$ statistic for equation = 104.0

oil. It was suggested earlier that a slower schedule is a result of heavier traffic and/or more frequent stops to load and discharge passengers. Operating under such conditions would be expected to increase oil consumption.

OTHER OPERATING COSTS

The final category to be modelled is comprised of all operating costs not embraced in one of the four resource models just discussed. Thus this “other” category includes the non-labour portion of repair costs, tyres and tubes, station expense (little or nothing for most bus companies), traffic and advertising, insurance and safety, and administrative and general expenses (see Table 1). Although included in the APTA accounts, depreciation, amortisation, taxes, and licence fees were not considered to be part of operating cost for purposes of this study. To be sure, taxes and licence fees are very real expenditures for many bus properties. However, they are transfer payments that do not represent the consumption of real resources.
Similarly, depreciation and amortisation are more or less arbitrary accounting calculations; they do not generally correspond to avoidable resource consumption. Thus none of the excluded items qualifies as a true economic cost. Further, brief reflection will reveal that such "costs" can be forecast in a straightforward manner on the basis of tax rates and the planned asset base. They do not lend themselves to statistical modelling.

Since the "all other" cost category is made up of those costs that could not be modelled in terms of real resource usage, the units must necessarily be dollars. The dependent variable to be forecast is in dollars per bus mile. The cost data were drawn directly from the Transit Operating Report.

The model for all other costs is given in Table 6. Interpretation of the presence of the two independent variables is apparently straightforward. The coefficient of GROWTH indicates that transit can be operated more cheaply (per mile) in cities which experienced a large share of their growth since the rise in automobile use. This is consistent with earlier discussion. The coefficient of WAGE suggests that per-mile costs rise with the driver wage rate. Certainly some of this effect is due to drivers' fringe benefits, which are buried in the administrative and general account and thus are included in the all other cost category. It is likely also that all or most of the wage rates at a property are geared to the driver rate, either directly or indirectly. Thus the driver wage rate is a barometer for the general level of wages in a system.

**FORECASTING ERROR**

As was stated at the outset, the purpose of the modelling exercises reported here was to produce a usable tool for forecasting bus operating costs. Concern, then, must focus on the accuracy of the final forecast. It should be observed at this point that the level of confidence associated with individual regression coefficients is quite a separate matter from expected forecasting errors. Especially in the presence of multicollinear regressors, a model may contain substantial errors in regression coefficients and still be a good forecasting tool. Since varying degrees of
multicollinearity were present in the several sets of regressors utilised in the component cost models, attention must be turned to the error in the final forecast rather than to individual regression coefficients or $R^2$ values.

Because the models forecast resource usage in different units, the error analysis is somewhat involved. The forecast of total operating cost is a composite of the five individual component forecasts. In order to add the quantities, we must convert each to dollars by applying the appropriate rate for the use of the resource.

The error in each component forecast is attributable to random factors which affect cost at various bus transit properties. Even though these effects are random, however, that is not to say that they are necessarily independent. For example, it may be that a system which pays for a higher than expected number of driver hours is also likely to pay for more bus repair man-hours than would be predicted (by the model). On the other hand, there may be tradeoffs between cost components, so that a system with above average costs in one category will have below average costs in another. There is no way to predict the directions of the relationships beforehand; the point is that an overall error analysis should allow for correlations among the residuals of the submodels.

The general formula for the variance of a linear combination of random variables is

$$\text{Var}\left( \sum_{i=1}^{N} a_i X_i \right) = \sum_{i=1}^{N} a_i^2 \text{Cov}(X_i, X_j)$$

where $\text{Cov}(X_i, X_j) = \text{Var}(X_j)$. Also,

$$\text{Cov}(X_i, X_j) = \rho_{ij} [\text{Var}(X_i)\text{Var}(X_j)]^{1/2}$$

where $\rho_{ij}$ is the coefficient of correlation between $X_i$ and $X_j$. Since three of the forecasting models assumed non-constant variances, a constant matrix of covariances cannot be calculated. It is possible, however, to find a matrix of correlations, since correlation coefficients are dimension-free; that is, they do not depend upon the scale of the variables involved. The calculation of correlation coefficients was achieved by weighting the residuals from the three models in question so as to produce constant variances. For the repair labour and oil consumption models, this simply amounted to using the residuals from the weighted least squares regression runs. For fuel consumption, the calculation was somewhat more involved, since each data point corresponded to a fleet, not to a system. The predicted fuel consumption had to be calculated for each fleet in a system (in gallons) and summed to obtain a forecast for the entire system. The constant-variance residual was then computed as

$$\left[ (\text{actual gallons}) - (\text{forecasted gallons}) \right] \sqrt{\frac{1}{F} \sum_{j=1}^{F} \frac{m_j}{N_j}},$$

where $F$ is the number of fleets in the system, $m_j$ is the bus miles operated by fleet $j$ (in thousands), and $N_j$ is the number of buses in fleet $j$. All the models except that for driver labour were based solely on 1973 data, and the samples overlapped
enough for correlations among the residuals to be estimated. The data for the driver labour model, however, came from a different source and covered several years. As a result there was no overlap with the other samples, and correlations with residuals from other models could not be calculated.

For ease of exposition, the following notation will be adopted for the residuals of the cost component models. The units of each residual are also indicated to avoid confusion.

\[ e_d \] Driver wages: man hours per bus hour.
\[ e_r \] Repair labour: man hours per thousand bus miles (weighted).
\[ e_f \] Fuel consumption: gallons of fuel (weighted).
\[ e_o \] Oil consumption: gallons of oil per thousand bus miles (weighted).
\[ e_a \] All other operating costs: dollars per bus mile.

Correlations among these residuals are shown in Table 7. As noted above, no correlations could be calculated between \( e_d \) and other residuals, so they must be assumed to be zero in the error analysis that follows. Table 7 also gives the number of observations from which each correlation coefficient was calculated. It is interesting that there are fairly high correlations in both positive and negative directions.

The next step is to create a "typical" hypothetical transit system, plug its associated variables into the formulas, and examine the magnitude of the composite standard error. The following values will be used for the regressors:

- \( X_1 = 0 \) (overtime not paid on overtime)
- \( X_2 = 0.5 \) (5 drivers per 10,000 bus hours)
- \( X_3 = 12 \) (spread overtime index)
- \( S_4 = 1.0 \) (special overtime allowances are present)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUSMI</td>
<td>64</td>
<td>(6,400,000 annual bus miles)</td>
</tr>
<tr>
<td>SRBM</td>
<td>8.0</td>
<td>(square root of BUSMI)</td>
</tr>
<tr>
<td>PRIVAT</td>
<td>0</td>
<td>(publicly owned system)</td>
</tr>
<tr>
<td>SNOW</td>
<td>12</td>
<td>(1200 mm. annual snowfall)</td>
</tr>
<tr>
<td>GROWTH</td>
<td>0.140</td>
<td>(city growth pattern index)</td>
</tr>
<tr>
<td>SMALL</td>
<td>0</td>
<td>(annual bus miles greater than 1.5 million)</td>
</tr>
<tr>
<td>MIPBUS</td>
<td>2.712</td>
<td>(27.120 miles per bus per year)</td>
</tr>
<tr>
<td>SLOWNS</td>
<td>8.526</td>
<td>(11.75 miles per hour average operating speed)</td>
</tr>
<tr>
<td>BUSACT</td>
<td>2.36</td>
<td>(236 active buses)</td>
</tr>
<tr>
<td>WAGE</td>
<td>5.00</td>
<td>(top driver wage rate $5 per hour)</td>
</tr>
<tr>
<td>BUSHRS</td>
<td>54.57</td>
<td>(545,700 bus hours operated per year).</td>
</tr>
</tbody>
</table>

The system operates five fleets of buses, as shown in Table 8. Some additional information will be needed for the calculations:

- Mechanic's wage rate: $5.50 per hour
- Cost of fuel: $0.32 per gallon
- Cost of oil: $0.80 per gallon.
TABLE 7  
Correlations among Residuals of the Cost Component Models

<table>
<thead>
<tr>
<th>Correlation Coefficients</th>
<th>$e_D$</th>
<th>$e_F$</th>
<th>$e_O$</th>
<th>$e_A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_D$</td>
<td></td>
<td>0.0233</td>
<td>0.3952</td>
<td>-0.1010</td>
</tr>
<tr>
<td>$e_F$</td>
<td>0.0233</td>
<td></td>
<td>0.3312</td>
<td>-0.3696</td>
</tr>
<tr>
<td>$e_O$</td>
<td>-0.0462</td>
<td>0.3312</td>
<td></td>
<td>0.0802</td>
</tr>
</tbody>
</table>

Number of Joint Observations

<table>
<thead>
<tr>
<th>Number of Joint Observations</th>
<th>$e_D$</th>
<th>$e_F$</th>
<th>$e_O$</th>
<th>$e_A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_D$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$e_F$</td>
<td>16</td>
<td>18</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>$e_O$</td>
<td>18</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$e_A$</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 8  
Fleet Data for a Hypothetical System

<table>
<thead>
<tr>
<th>Fleet</th>
<th>Number of Buses</th>
<th>Seating Capacity</th>
<th>Age</th>
<th>Air Conditioned</th>
<th>Miles Operated</th>
<th>Fuel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80</td>
<td>55</td>
<td>1</td>
<td>Yes</td>
<td>2,400,000</td>
<td>Diesel</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>45</td>
<td>4</td>
<td>Yes</td>
<td>1,740,000</td>
<td>Diesel</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>45</td>
<td>10</td>
<td>No</td>
<td>810,000</td>
<td>Diesel</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>33</td>
<td>12</td>
<td>No</td>
<td>810,000</td>
<td>Diesel</td>
</tr>
<tr>
<td>5</td>
<td>36</td>
<td>33</td>
<td>20</td>
<td>No</td>
<td>640,000</td>
<td>Diesel</td>
</tr>
<tr>
<td>Total</td>
<td>236</td>
<td></td>
<td></td>
<td></td>
<td>6,400,000</td>
<td></td>
</tr>
</tbody>
</table>

Now the expected annual cost in each category can be calculated from the model, yielding the following results:

Driver wages 3,404,000  
Bus repair wages 500,000  
Fuel 433,500  
Oil 16,000  
All other 2,588,000  

Total operating cost 6,942,000

This corresponds to a cost of $1.085 per mile. Calculated values of the standard deviations of residuals for the hypothetical system are given below:
January 1977

JOURNAL OF TRANSPORT ECONOMICS AND POLICY

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Driver wages</td>
<td>0.0624 pay hours/bus hour</td>
<td></td>
</tr>
<tr>
<td>Bus repair wages</td>
<td>2.291 man hours/thousand bus miles</td>
<td></td>
</tr>
<tr>
<td>Fuel</td>
<td>53,940 gallons</td>
<td></td>
</tr>
<tr>
<td>Oil</td>
<td>0.3594 gallons/thousand bus miles</td>
<td></td>
</tr>
<tr>
<td>All other</td>
<td>0.0645 dollars per bus mile</td>
<td></td>
</tr>
</tbody>
</table>

The units of the standard deviation for the fuel model are gallons rather than gallons per mile, because of the computations involved in combining fleet forecasts to form a system-wide forecast, as described previously. Now these standard deviations, along with the correlation coefficients given in Table 7, can be substituted into equations (4) and (5) to produce a composite standard error of $443,000, which for this system is about 6.4 percent of the forecast operating cost. Although the percentage will vary with the parameters of the system under consideration, the result for this more-or-less typical system gives an indication of the order of magnitude of the forecasting error involved.

The preceding analysis assumes that the prices to be applied to the resources are known with certainty. To the extent that there are forecasting errors in the prices, the error associated with the total dollar forecast will be inflated. Also, the correlations between residuals of the driver wages model and those of other models were assumed to be zero. Non-zero correlations would either increase or decrease the error estimate, depending upon their signs. In summary, given the unknown factors just mentioned and the limited sample sizes, the error estimate must be regarded only as an indication of the order of magnitude of the error.

CONCLUSION

As stated at the outset, the primary purpose of this work was the practical one of developing a useful forecasting tool. Concern for detecting economies of scale or the importance of certain variables was only secondary. Despite the pragmatic focus, however, some information of the latter type certainly comes out as a by-product of the study. In this vein, it should be pointed out that more can be inferred from observing which variables are in the models than which variables are not in the models. That is, the presence of a particular variable in a cost component model is a good indication that the variable has a significant relationship with that cost component; but the absence of another variable does not mean that that variable is not strongly related to the cost component. In the modelling of every cost component there was room for personal discretion as to the set of regressors that would be chosen. An analyst trying to make a case for the importance of a particular variable could very likely justify its presence in one or more of the models. The objective of this paper is to present a forecasting tool—not to study the importance of variables per se. Therefore those variables not found in any of the models will not be discussed.

Each of the cost component models, except that for the “all other” category, dealt in terms of real resource usage, and therefore should not be affected strongly
by price fluctuations. A natural question that arises is what effect inflation is likely
to have on the "all other" cost model. The presence of the driver wage rate in the
model may serve to account in some degree for increases in general price levels,
depending on how good an index the driver wage rate is of other prices encountered
by bus properties. It must be admitted frankly, however, that, since that com-
ponent model is based exclusively on 1973 data, there is no way to project it to other
years with confidence. Some further work is needed to find how well this model fits
other years' data—particularly the more recent experience.

Some past researchers were concerned with establishing the kinds of variables
that are related to operating cost. The analysis carried out here clearly suggests that
both system-descriptor variables and city-descriptor variables are important as cost
predictors.

The would-be user of this forecasting tool is reminded that the total forecast
operating cost does not include depreciation or taxes and licence fees. Statistical
modelling is probably not an appropriate method of estimating such expenses. Tax
rates and licence fees are public information; their calculation should be straightforward.
Depreciation expense is a function of the asset base and the property's
selected accounting methods. Again, the computations are straightforward.

While the cost models developed in this research are less than perfect, it is felt
that they offer a better basis for predicting operating costs than has been available
heretofore. In particular, they provide a generally applicable tool for estimating in-
dividual cost components.

APPENDIX
Items from Union contracts affecting wage payments

Source: [16]

1. Maximum hourly wage rate. The top rate paid to any driver. In most cases the rates are
given for several years, along with the effective dates for each.
2. Starting rate. The lowest driver wage rate. A starting rate is given for each maximum
rate.
3. Time required to reach maximum rate. The number of months of service required for a
driver to reach the maximum rate.
4. Night shift premium. Additional pay, if any, for driving the night shift or "owl run."
5. Average total operator's wages per platform hour. Total driver wages paid during a pay
period divided by total platform hours during the period. Platform hours are made up of the
total time the driver is in charge of the bus, including running time, layover time, and travel
time to and from the garage.
6. Average pay hours in regular runs (both daily and weekly). The average of the pay hours
of men with regular runs. This includes all time paid for, so that an hour at time-and-a-half
counts as one and one-half pay hours.
7. Average platform hours in regular runs (both daily and weekly). The average number of
platform hours of men with regular runs.
8. Overtime rate. Overtime pay rate as a multiple of the regular rate, e.g., 1.5.
9. Overtime before or after scheduled run. Whether overtime is paid for any work before the
start of or after the completion of a regular run (yes or no).
10. Overtime after specified hours per day. The number of hours of work in a day after
which overtime is paid. Separate figures are given for regular men and extra men.
11. Overtime after specified hours per week. As in the previous item, except on a weekly basis.
12. Time allowances included in calculating daily and weekly overtime. Allowances such as turn-in time, report time, and travel time that are added to platform hours for the calculation of overtime payments.
13. Number of hours of spread after which overtime is paid. Spread is defined as the total elapsed time from the start of the first piece of work in a run to the completion of the last piece of work in that run. This provision is aimed at compensating for swing (split) runs.
14. Time allowances included in calculating spread overtime. The same kinds of allowances that may be made in item 12.
15. Is spread overtime paid in addition to daily or weekly overtime? (Yes or no.) Indicates whether a double penalty is paid on the same hours. If the answer is no, only the larger of the two amounts is paid.
16. Overtime on day off. (Yes or no.) Indicates whether a man is paid overtime for working on his scheduled day off.
17. Overtime on Sundays when part of regular scheduled work week. (Yes or no.)
18. Overtime on specified number of holidays when part of regular scheduled work week.

The number of holidays with this provision.
19. Other special overtime provisions. Any other special basis for paying overtime.
20. When regular and special overtime provisions overlap is overtime paid on overtime? (Yes or no.)
21. Percentage of straight runs on weekdays. The percentage of straight, as opposed to swing, runs in the schedule.
22. Minimum pay hours of regular runs on weekdays. This number of pay hours is guaranteed for all regular runs.
23. Number of regular runs having outside spread of hours of . . . (Includes both straight and swing runs.) Frequency distribution of outside spreads, with the following six counts: less than 9 hours; 9 to 10 hours; 10 to 11 hours; 11 to 12 hours; 12 to 13 hours; over 13 hours. For weekdays only.
24. Percentage of straight runs on Saturdays, Sundays, and holidays.
25. Average minutes of relief for meal purposes.
26. Is operator paid during meal period? (Yes or no.)
27. Allowance for reporting, sign-on, or pullout time—first report. (In minutes.)
28. Allowance for reporting, sign-on, or pull-out time—subsequent report. (In minutes.)
29. Allowance for cash-in time. (In minutes.)
30. Allowance for travel time. (In minutes.)
31. Number of holidays paid for but not worked.
32. Number of operators.

REFERENCES

THE COST OF OPERATING BUSES IN U.S. CITIES

Hoyt G. Wilson


