Mode Choice: Implications for Planning

JEFFREY M. ZUPAN, Peat, Marwick, Livingston & Co.

To appraise land-use and transportation plans for the future, U. S. Census Journey-to-Work survey data were used to develop distribution and modal split models for work trips in the New York region.

Gravity distribution models were developed for each of three income groups. Using multiple regression techniques, an equation was derived that related walk-to-work trips to measures of a zone's self-containment. An equation to allocate work trips to transit and automobile modes was developed. It was found that the percent of transit trips for an origin-destination pair depends on employment density at the work end of the trip, residential density at the home end of the trip, availability of adequate rail service between origin and destination, relative times of the automobile and transit modes, the cost of tolls, and the cost of parking. Analyses of the root-mean-square error, geographical and transportation biases, and sensitivity to changes in the variables showed the equation to be a reasonable forecasting tool.

For each of three income levels, modal split equations were derived, including all of the same variables. Comparisons were made of the parameters of the three stratified equations and it was found that as the income level rose, the significance of transportation-related variables increased the significance of land-use variables decreased. Analyses of the stratified equations showed them to be reasonable forecasting tools. Examination was made of sensitivities for the assumption of a radically different distribution of income levels, and it was found that the sensitivities as well as predictive values would change appreciably. It was concluded that only if income level distribution were not changing radically would the unstratified equation be adequate for modal split forecasts.

The model calibration process required no trip data collection beyond that provided by the census survey. Future refinements have been suggested, particularly the use of a second point in time. Also, a recognition of the dynamic nature of the social structure is necessary to insure proper urban planning.

•AT a time when the entire structure of the urban environment is being reevaluated and large metropolitan areas are contemplating dramatic improvements in their mass transportation systems, rational means to forecast transportation demands are needed. Since it is the work trip that places the greatest burden on the transportation system, the characteristics of the work trip constitute a vital area of study for the transportation planner.

Paper sponsored by Committee on Future Concepts and presented at the 47th Annual Meeting.

To better evaluate transportation and land-use plans in the New York region, Peat, Marwick, Livingston & Co. (PML) developed a mathematical model for the Regional Plan Association (RPA). The model was designed to forecast work trips on an origin-destination (O-D) basis and to assign these trips to either the transit or the automobile mode. The source of trip data was the 1960 U. S. Census Bureau Journey-to-Work survey. The study area consisted of 31 counties in New York, New Jersey, and Connecticut, totaling 12,750 square miles. The area is inhabited by almost 20 million people, including 8 million workers.

The development of the distribution and modal split models represents a significant breakthrough in the transportation planning field in at least two respects. First, the trip data for model development were obtained from U. S. Census data only. Second, the model calibration process was completed using a system of large and relatively heterogeneous zones. In other words, a work trip model was calibrated without resorting to costly data collection procedures or complex zone systems.

A gravity distribution model was calibrated for each of three income levels. Using multiple regression techniques, we attempted to derive statistically significant relationships for walk-to-work trips and for work-at-home trips. No acceptable equations resulted for the work-at-home trips, but a sound equation was derived that related walk-to-work trips to an area's employment density and to its self-containment (the likelihood of finding employment within the worker's residential zone).

A modal split equation was developed that did not consider income as a mode choice determinant. The equation related the percent of transit work trips in an O-D pair to:

- employment density at the work end of the trip;
- residential density at the home end of the trip;
- availability of adequate rail service between origin and destination;
- relative times of the auto and transit trips;
- cost of tolls; and
- cost of parking.

The equation proved to be more than adequate when evaluated for the root-mean-square (RMS) error (the square root of the mean of the sums of the squares of the differences between the actual values of the dependent variable and those predicted by the equation), geographical and transportation biases, and sensitivity to changes in the values of the variables.

Modal split equations were also derived for each of three income levels. These equations included most of the variables cited. When the parameters of the three stratified equations were compared, we found that as income level rose, the significance of the transportation variables increased and the significance of the environmental or land-use variables decreased.

The RMS error, geographical and transportation biases, and sensitivity to changes in values of the variables were analyzed for the stratified equations. The stratified equations also proved to be good forecasting tools. Examining the modal split and the sensitivities of the variables to changes for a radically different distribution of income levels, we found that the sensitivities as well as the predicted values would change appreciably under such an assumption.

Therefore, if the purpose of the modal split forecast is to study the influence of transportation and land-use plans in the near future, the income-stratified modal split equations are unnecessary. However, if the purpose of the forecast is to study the influence of transportation and land-use plans for a time when income level distribution will be radically different, the income-stratified modal split equations should be applied to obtain a higher level of accuracy and proper sensitivities.

ABBREVIATIONS

CA = the cost of the automobile trip from origin to destination, including tolls, parking cost, and an over-the-road cost of 2.5 cents per mile (cents).
CT = the fare of the transit trip from origin to destination (cents).
D = over-the-road distance from origin to destination (miles).

ED = employment density (thousands of employees per developed square mile).
L = tolls from origin to destination (cents).
ln = natural logarithm.
N = number of data points.
P = parking costs at destination (cents).
R = coefficient of multiple correlation.
RD = residential density (thousands of residents in labor force per residential square mile).
SF = combined rail service factor for both no-transfer and one-transfer service.
S_N = rail service factor for no transfers.
S_T = service factor for one transfer.
TA = total travel time by automobile between origin and destination (minutes).
TT = total travel time by transit between origin and destination (minutes).
$\bar{y}$ = mean value of dependent variable.

THE MODEL AND IMPLICATIONS FOR THE FUTURE

Undoubtedly, the best way to develop a model to determine future mode choice is to simulate all factors that influence each traveler's decision. Much research remains to be done before we will be able to quantify all of these factors and forecast them. The model development described in this paper attempts to incorporate those characteristics of the trip and trip-maker that can be accurately and easily evaluated at the present time.

There are some limitations in the modal split model. The rail service factor is less than perfect because many levels of service are represented by identical values of this variable. The absence of a direct measure of transit cost could also be a serious problem. Our analysis showed that present transit fares do not significantly affect mode choice, but this might not be the case if new systems with radically different fares come into existence. The absence of a transit fare variable raises the question of whether we can properly evaluate exotic transportation systems of the future by applying a model based on the present transportation system. These limitations are the cause of legitimate concern, and investigation should be undertaken to remove them.

Perhaps the greatest criticism leveled at most transportation models is that they are based on one point in time. The assumption is usually made that the significant variables and the values of their parameters for the base year remain constant for future years. There can be no doubt that this assumption is open to serious question. To remove this objection, data sources for other years should be employed. If we can measure the changes in the significance of variables and in the size of the parameters over time, we could extrapolate these changes into the future. This approach could be attempted in the RPA region by using the data from the home interview survey that was taken by the Tri-State Transportation Commission in 1963. Another data source to consider will be the 1970 U.S. Census Bureau Journey-to-Work survey. Since transportation planning is a continuous process, it is vital that models incorporate new data as they become available.

Changes in the characteristics of the worker and the work trip may influence transportation planning in many ways. For instance, a growing number of married women are entering the nation's labor force. From 1940 to 1966, the rate of working wives under age 35 doubled, and the rate of working wives over 35 tripled (1). A second worker in the family may have a significant effect on mode choice. Since it is doubtful that a transportation planner in 1940 could have foreseen such a dramatic change in the nation's work characteristics, can we place much faith in our abilities to anticipate such changes?

The length of the work week is another factor that may influence model development. Thirty years ago a 6-day work week was in effect. The decline in the economic viability of our public transportation systems can be traced in part to the present reduction to a 5-day work week. We can expect that a possible further reduction to four or three days might further hamper the service and thus the attractiveness of public transportation. The shorter work week may also make the worker more tolerant of longer trips and less comfortable service, since the trip is made less frequently. Adopting a system of staggered working hours might have a major effect on mode selection, with the trip-maker reacting to changing congestion conditions. Increasing numbers of people in higher education programs for longer periods, coupled with earlier retirements, will serve to significantly alter the age distribution of workers. These factors may, in turn, alter their travel habits. A continued increase in the proportion of white-collar jobs also requires careful examination by the model-builder.

These are samples of changes in worker characteristics that must revise some of our basic assumptions in transportation planning.

The impact of far-reaching changes in the social order must not be overlooked. The application of new communication techniques and new energy sources and the ascendance of a semicomputerized society will have wide implications for planning. What might be the effect on society of changes in size, structure, or function of the family unit? And what will be the impact of full opportunity for the Negro in America? If heart disease and cancer are controlled, can we properly plan for the resulting changes in the age distribution of our population? Can we measure the values and ideals of an emerging class of leisure and affluence? In short, there are many complex ways in which a changing society can have implications for planning, and they are often difficult to identify and evaluate. It is clear that the planner must be aware of the dynamic nature of our social structure and he must relate this dynamism to his plans for society.

DISTRIBUTION MODEL STRUCTURE

The distribution process was accomplished by the gravity distribution model. That is, the total work trips produced in an origin zone, which are equal to the zone's resident labor force, were distributed to the various destination zones in proportion to the employment in each destination zone, and in proportion to an empirically derived friction factor that measures the propensity to travel between origin zones and destination zones. The definition of the gravity model distribution, stated mathematically is

$$T_{ij} = P_i \frac{A_j \cdot F_{(t_{ij})}}{\sum_{x=1}^{n} A_x \cdot F_{(t_{ix})}}$$

where

T_{ij} = trips produced in i that are attracted to j;
P_i = trips produced by i;
A_j = trips attracted by j; and
$F_{(t_{ij})}$ = empirically derived friction factor measuring propensity to travel between i and j.

The source of trip data was the U. S. Census Bureau Journey-to-Work survey, which is based on a relatively coarse-zone system (Fig. 1). For example, the entire Borough of Brooklyn was considered one zone. We used only census data, and the collection of additional trip data was not required to develop the model.

We attempted to calibrate the gravity model by the usual trial-and-error procedure of adjusting the friction factors, but the model could not be adequately calibrated. We decided to stratify the trip-makers by income level. We felt that if workers were separated by income at both residence and employment ends of their work trips, the gravity models would be able to match workers with jobs. Therefore, workers living in each zone and employees in each zone were separated into low-, medium-, and high-income groups, corresponding to annual incomes of \$0 to \$5,000, \$5,000 to \$10,000, and greater than \$10,000.

Using the census data, the three gravity models were calibrated for the coarse-zone system. To determine whether the coarse-zone system reflected the distribution of

Figure 1. RPA study area coarse zones.

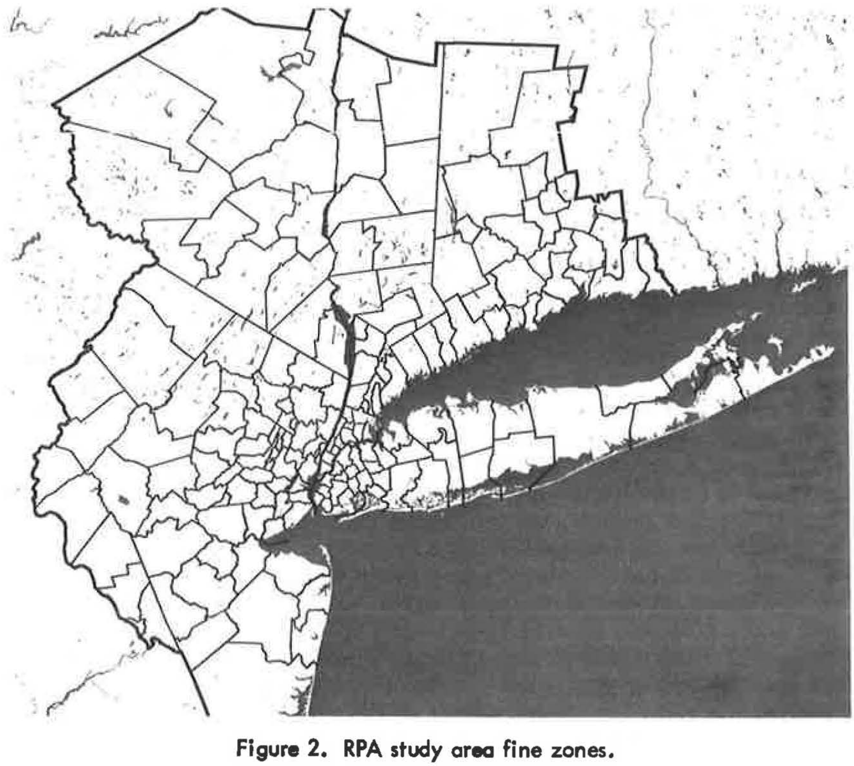

Figure 2. RPA study area fine zones.

work trips, the three gravity models were applied to a fine-zone system. (RPA has developed a 177-zone system, shown in Figure 2, which permits analysis of the region using relatively homogeneous zones.) The three resulting trip distributions were then compressed to the coarse-zone system for comparison with the source data. The two distributions agreed well, indicating that the gravity models were able to reproduce closely the zonal interchange trip volumes. The three sets of friction factors are shown in Figure 3.

MODAL SPLIT ANALYSIS

Since trip distribution required income stratification and since income level is intuitively relevant to mode choice, we decided to consider income stratification in developing the modal split model. The trip data used for the modal split analysis were taken from the 1960 Bureau of the Census survey of the Journey-to-Work. Data on the origin and destination of trips by mode, the number of trip-makers in each income group, and the number using automobiles or some other means of travel were available on the coarse-zone basis only.

Before developing the modal split equations, it was necessary to determine the number of people who did not use the transportation system, that is, those who either worked at home or walked to work. Data on income level were not available for these two categories. Since all other data were available for the fine-zone system, there was no ad-

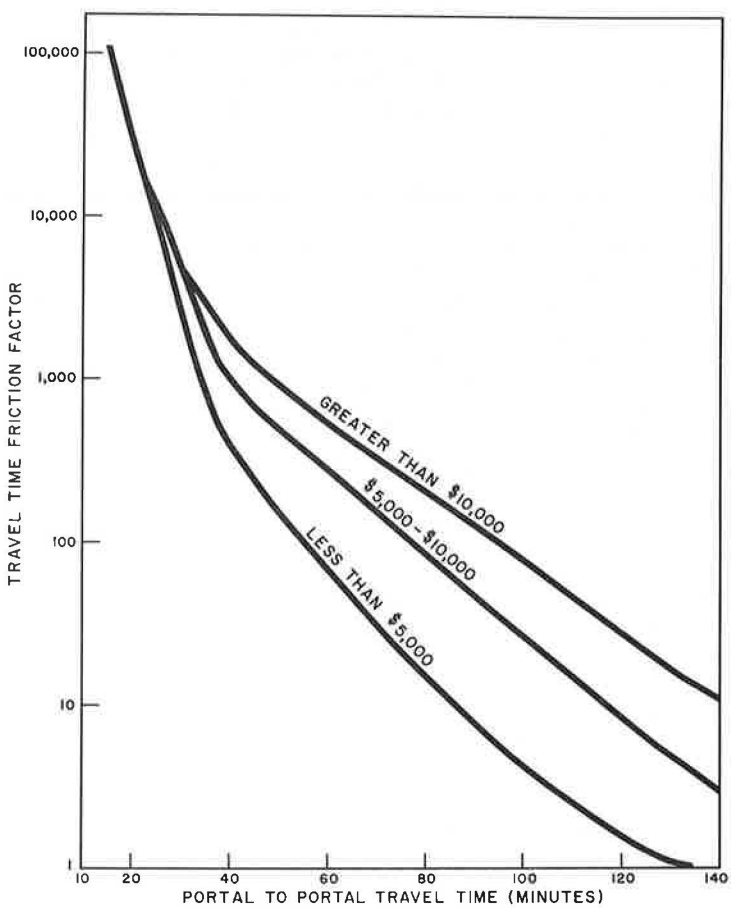

Figure 3. Gravity model propensity curves.

vantage to using the coarse-zone system for the walk-to-work and work-at-home analyses. Because the origin zone was of prime importance for these types of trips, the analysis was made by looking at the origin zones only. (For the person who works at home, origin and destination are identical; the destination of walk-to-work trips was identical to the origin, on the coarse-zone level, in 97 percent of the walk-to-work trips.)

Work-at-Home and Walk-to-Work Analysis

For regression analysis of work-at-home trips, the dependent variable used was the percent of the resident labor force that works at home. The independent variables were residential density (the resident labor force per net residential square mile) and employment density (number of employees per developed square mile). We felt that positive correlations would be found between densities and work-at-home percentages.

RPA calculated the values of residential and employment density. These values, as well as their natural logarithms and square roots, were tried in the regression analysis. The best results produced a correlation coefficient of 0.558 with a standard error of 2.67 percent. The mean value of the work-at-home trips was 3.75 percent with a standard deviation of 3.71 percent.

When we examined the residuals (the difference between the actual values and those predicted by the equation) of these results, we could find no additional variables that might explain the remaining variance. Apparently, many factors determine the portion of residential labor force that works at home. High percentages appeared in extremely high-income zones, where many professionals probably have offices in their homes; in agricultural zones, where farms serve as both home and place of employment; and in zones that have a large number of retail store owners who live on the store premises. Adding variables that reflected the peculiar zonal characteristics might have improved the results somewhat, but the net effect would have been minimal. We decided that none of the equations developed for work-at-home volumes was adequate for use as a forecasting tool. Therefore, we decided to assume that the work-at-home percentage remained constant for each zone, unless there were good reasons to assume otherwise for particular zones in future years.

Walk-to-work trips were also analyzed as a percent of the resident labor force. The independent variables were residential and employment densities and their natural logarithms. Next, a new independent variable, the percent of the resident labor force that made internal trips, was used. We thought that this variable, which measures a zone's self-containment, would reflect the likelihood of a worker's finding a job within walking distance of his home. Independent variables entering the equation, in the order of their significance, were employment density, percent internal trips of resident labor force, and the natural logarithm of employment density. This equation resulted:

$$WTW/RLF = 0.044\ ED + 8.349\ (INTRAS/RLF) + 0.907\ (\ln ED) + 4.109$$

where

WTW = walk-to-work trips,
ED = employees (thousands) per developed square mile,
$INTRAS$ = internal trips, and
RLF = resident labor force.

The correlation coefficient of this equation was 0.737 and the standard error was 3.72 percent. The standard deviation was 5.38 percent and the mean value of the dependent variable was 10.06 percent.

Development of Modal Split Equations

Next we focused attention on the trips that place a burden on the transportation system, i.e., automobile and transit trips. We decided to develop the modal split by devising relationships both for work trips unstratified by income level and for work trips stratified by the three income levels.

Before developing an unstratified modal split equation, we examined the income-stratified modal split, without using variables that measure characteristics of the auto or transit systems, in order to determine to what extent mode choice could be explained by residential and employment densities. If these variables explained mode choice completely, the implication would be that modal split would be affected only by the degree to which the transportation system could affect densities. The residential and employment densities and various transformations of them were used as independent variables for each of three income groups. Only O-D pairs where transit trips occurred were considered. Correlation coefficients ranged from 0.540 to 0.605 for the three groups. While these results were not suitable for use as forecasting equations, it was interesting to note the large amount of variance in modal split that can be explained without giving direct consideration to the characteristics of the transportation system.

Development of the Unstratified Modal Split Equation—Next, we developed an unstratified equation that included transportation system variables. Table 1 gives the regression results for these trials. The analysis contained 629 O-D pairs. Many O-D pairs were not used because density data were lacking. O-D pairs were also excluded if the number of work trips was represented by fewer than ten interviews from the census survey. This was necessary because the relatively large sampling errors that resulted from a small number of interviews produced an unreliable split between modes. O-D pairs were also excluded if no transit trips occurred, since the large number of such zonal pairs would tend to distort the equation at the lower end of the percent transit scale.

RPA developed average zone-to-zone travel times by transit for the 629 zonal pairs. Automobile times were determined by skimming the minimum time path from the auto network and adding access and egress times to and from each centroid. Zone-to-zone transit costs were determined by examining the fare structure of railroads, buses, ferries, PATH (Port Authority Trans-Hudson system), and the New York City subway system.

The determination of average zone-to-zone auto costs was a more difficult problem. It was important to consider how the trip-maker perceives his costs; after all, this is

TABLE 1

MODAL SPLIT REGRESSION ANALYSIS—UNSTRATIFIED BY INCOME

Dependent Variable	Variables Attempted	Variables Entered	Standard Deviation	Standard Error	$\overline{Y}$	R
% Transit	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, CA/CT, TA/TT	ln ED, $\sqrt{RD}$, TA/TT, CA/CT, ED	0.2550	0.1452	0.3341	0.824
% Transit	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, CA/CT $(TA/TT)^2$	ln ED, $\sqrt{RD}$, $(TA/TT)^2$, CA/CT ED	0.2550	0.1457	0.3341	0.822
% Transit	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, CA/CT, TA/TT, D	$\sqrt{ED}$, $\sqrt{RD}$, TA/TT, CA/CT	0.2453	0.1445	0.3368	0.812
% Transit	RD, ED, $\sqrt{RD}$, ln RD, CA/CT, TA/TT, $\sqrt[3]{ED}$, D	$\sqrt[3]{ED}$, RD, TA/TT, CA/CT	0.2543	0.1445	0.3368	0.824
% Transit	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, CA/CT, TA/TT	$\sqrt{ED}$, $\sqrt{RD}$, TA/TT CA/CT	0.2543	0.1483	0.3368	0.814
% Transit	$\sqrt{RD}$, ln ED, TA/TT, L + P	ln ED, $\sqrt{RD}$ TA/TT, L + P	0.2544	0.1430	0.3363	0.828
% Transit	$\sqrt{RD}$, ln ED (L + P)/CT, TA/TT	ln ED, $\sqrt{RD}$, TA/TT, (L + P)/CT	0.2544	0.1441	0.3363	0.825
% Transit	$\sqrt{RD}$, ln ED, L + P, TA/TT, S_T, S_N	ln ED, $\sqrt{RD}$, TA/TT, S_T, S_N, L + P	0.2543	0.1330	0.3360	0.854
% Transit	ln RD, ln ED, L + P, TA/TT, S_T, S_N	ln ED, ln RD, TA/TT, L + P	0.2543	0.1350	0.3360	0.849
% Transit	$\sqrt{RD}$, ln ED, L + P TA/TT, SF	ln ED, $\sqrt{RD}$, SF TA/TT, L + P	0.2543	0.1256	0.3360	0.871

his basis for choosing the way he will travel. For example, if the trip-maker does not view automobile depreciation as a commuting cost, it seems logical that the model-maker should not consider it. It seemed to us that the automobile costs that the commuter perceives are gasoline, oil, tolls, and average parking costs at his destination. Therefore, these costs were used in computing the cost variable. Gasoline and oil costs were assumed to be 2.5 cents per mile.

It should be recognized that these transportation system characteristics represent an "averaging" effect. Zone-to-zone travel times and costs reflect only an approximation to the actual times and costs experienced by all the trip-makers between residence and job.

Only variables that can be supported by intuitively sound arguments can be used in any regression analysis. It makes little sense to throw all possible variables into the pot in a shotgun approach merely to obtain high correlation coefficients. The unstratified modal split analysis considered the two density variables, their square roots and natural logarithms, the ratio of auto costs to transit cost, and the ratio of auto time to transit time. The density variables measure the trip-maker's environment at the home and work ends of his trip. A positive correlation of transit usage and density can be explained by assuming that higher densities will support more mass transportation service and increased service will attract greater usage. Time and cost ratios represent a comparison of two transportation system characteristics that are of major importance to the person choosing between alternate modes. Automobile ownership was not used as an independent variable because we felt that the high correlation it had with residential density would permit the density variable to serve in its place. Income, which is highly correlated with automobile ownership, was considered later for the stratified modal split model.

Use of the independent variables produced a regression equation with a correlation coefficient of 0.824. In the order of entry, the variables were natural logarithm of employment density, square root of residential density, time ratio, cost ratio, and employment density. Since we theorized that the effect of the time ratio would increase as the difference between the two times increased, the time ratio squared was tried in lieu of the time ratio in another trial; however, this trial produced no improvement in the results.

All the unstratified modal split equations developed to this point predicted internal Manhattan trips as over 100 percent transit. In an attempt to produce an equation that would predict transit usage for this O-D pair more realistically, we decided to bar the natural logarithm of employment density from the equation. This variable contributed most heavily to the high percent predicted for this particular interchange. Also, the square root and the cube root of employment density were tried instead of the natural logarithm, but the resulting equation predicted internal Manhattan transit trips no better than the previous one. We also tried the zone-to-zone distance of the O-D pairs to see if the length of the trip would affect the modal split. No significant correlation was found.

We therefore concluded that the time ratio and the natural logarithm of employment density were suitable to use for the modal split equation. However, we decided to examine the cost term further.

As mentioned earlier, it is the commuter's perception of cost that should be considered. Lansing and Hendricks (2) point out that less than one-third of automobile commuters actually compute their driving costs. Those that do compute the cost tend to overestimate the over-the-road cost per mile. This study also indicated that parking costs greatly affect the number of trip-makers using the automobile. Therefore, we decided to consider only tolls and parking costs as perceived auto costs, and to use the ratio of the sum of tolls and parking to transit costs as the variable. The sum of tolls and parking was also used directly, without relation to transit costs. When taken with the previously accepted variables, the best cost variable proved to be the sum of tolls and parking.

Analysis of residuals produced by this equation showed clearly that those zone-to-zone O-D pairs in which transit ridership was being underpredicted almost invariably had good rapid transit rail service. Two new variables were incorporated in the analysis to reflect this: a no-transfer rail service factor and a one-transfer rail service factor. If an O-D pair possessed rail transit service that allowed the trip-maker to travel without transferring, that O-D pair was rated as 1 for the no-transfer factor. Other O-D pairs received a 0. If the O-D pair possessed one-transfer service, it was rated a 1 for the one-transfer factor, the other O-D pairs received a 0. These two variables were tried with the four variables already determined. A correlation coefficient of 0.854 resulted; however, both rail service factors entered with approximately the same coefficients and significance. This indicated that one-transfer service did not have a significantly different effect on transit usage from no-transfer service. Accordingly, the two variables were combined and the analysis was performed again. The new result showed a correlation coefficient of 0.871 and a standard error of 12.56 percent. The mean value of the dependent variable was 33.6 percent with a standard deviation of 25.43 percent. The final unstratified modal split equation was:

$$\begin{aligned}\%\ \text{TRANSIT} = {} & 7.756\,(\ln \text{ED}) + 2.723\sqrt{\text{RD}} + 17.884\text{SF} \\ & + 20.474\,(\text{TA}/\text{TT}) + 0.112\,(\text{L}+\text{P}) - 14.50\end{aligned}$$

where

ED = employees (thousands) per developed square mile at destination zone;
RD = resident workers (thousands) per net residential square mile at origin zone;
SF = rail service factor;
TA = zone-to-zone time by auto, in minutes;
TT = zone-to-zone time by transit, in minutes;
L = zone-to-zone toll cost, in cents; and
P = parking cost at destination, in cents.

These variables entered in the order given in the equation. Table 2 shows the statistical characteristics of the variables.

Analysis of the Unstratified Modal Split Equation—The root-mean-square error of the trips predicted by the unstratified modal split equation measures the predictive ability of the equation. Analyzing the residuals for particular geographic sectors or transportation systems determines whether geographic or transportation biases result from use of the equation. It is also necessary to test the sensitivity of the variables. Do reasonable changes in transit demand occur when one or more variables in the equation are altered?

For the RMS error analysis, the transit trips were computed using the predicted percent of transit, which was compared with the actual percent of transit usage. Since a few large, poorly predicted O-D pairs would distort the analysis, we focused attention on them. We discovered two major O-D pairs, Manhattan-to-Manhattan and Bronx-to-Manhattan, that would require special treatment.

As mentioned previously, the predicted percent of Manhattan-to-Manhattan transit usage was well over 100 percent, which is obviously impossible. What caused this poor prediction? The values of density and parking cost variables used in the modal split equation for this O-D pair weigh heavily against any auto usage. However, the equation is not sensitive to those trip-makers who must use their cars during the working day,

TABLE 2

STATISTICAL CHARACTERISTICS OF THE VARIABLES—UNSTRATIFIED MODAL SPLIT EQUATION

Variable	F-level Entering	t-Value Entering	Final Coefficient	Standard Error of Coefficient	Final t-Value
ln ED	492.2	22.2	7.756	0.445	17.4
$\sqrt{\text{RD}}$	298.5	17.3	2.723	0.169	16.1
SF	190.4	13.8	17.844	1.310	13.6
TA/TT	75.5	8.7	20.474	2.723	7.5
L + P	44.2	6.6	0.112	0.017	6.6

Note: The complete list of all regression trials is given in Table 1.

regardless of densities or costs, or those trip-makers who are rich enough to be unaffected by high costs. In other words, if the variables used were truly the only ones to be considered, absolutely no one would use an automobile for Manhattan-to-Manhattan work trips. We decided to eliminate this O-D pair from the RMS error analysis. The Bronx-to-Manhattan O-D pair also showed an impossibly high percent transit prediction. Here again, the value (99.9 percent) did not reflect trip-makers who must use their cars during the working day. This O-D pair was also removed from the RMS error analysis.

With these two data points omitted, the RMS error was 30.1 trips. This number can be interpreted as follows: the predicted volume of transit trips between any origin and destination is in error by less than 30.1 trips in 68.3 percent (one standard deviation) of the cases. It is in error by less than 60.2 trips in 95 percent of the cases and by less than 90.3 trips in 99 percent of the cases.

Next, we checked the predicted transit usage for various travel corridors. Table 3 gives the results. Because of the coarse-zone structure, the base used to evaluate the percent errors did not necessarily correspond to the actual volumes experienced in the corridors under study. For example, to calculate the percent error for the Erie-Lackawanna Railroad, a base of 22,300 trips was used instead of the 35,000 trips that actually occur, since only certain O-D pairs could be definitely identified as prime Erie-Lackawanna territory. For the geographic and transportation corridors analyzed, the ability of the unstratified modal split equation to predict transit trips appears to be excellent. The predictions for two important categories, "All transit trips to Manhattan" and "Subway trips in New York City," are particularly good.

Of course, the ability to reproduce only the past or present is not of any particular value. Since it is impossible to confirm the model's forecast of the future now, the best the analyst can do is examine the equation for reasonableness. Valid use of the modal split equation depends on its sensitivity to changes in the variables. Therefore, we examined the change in the dependent variable when the independent variables were allowed to vary.

Table 4 gives the changes in transit usage that will occur for five typical O-D cases. Case 1, travel from a high-density urban zone to the central business district (CBD), illustrates the change in transit usage for six variations:

- Ten minutes was added to auto time to simulate increased congestion on the highways into Manhattan; transit usage would increase from 91.6 percent to 95.4 percent, according to the model.
- Transit time was reduced by 10 min to simulate a widespread improvement in subway service; transit usage would increase to 96.2 percent. This variation and the preceding one indicate the sensitivity to travel times when auto and transit are competitive.

TABLE 3

UNSTRATIFIED MODAL SPLIT BIAS ANALYSIS

Category	Error of Predicted Transit Trips	% Error
All transit trips to Manhattan	+ 654	+ 0.07
All transit trips to Brooklyn	- 8,202	- 2.13
All transit trips to Queens	+15,889	+ 9.43
All transit trips to the Bronx	+ 7,714	+ 5.56
All transit trips to Newark	+ 8,339	+15.59
Bus trips to N. Y. C.—eastbound	+ 1,364	+ 2.31
Trans-Hudson transit trips—eastbound	- 2,280	- 1.64
Trans-Hudson transit trips—westbound	- 7,746	- 3.11
Subway trips in N. Y. C.	-13,236	- 0.67
New York Central R. R. trips to N. Y. C.	- 494	- 0.81
New Haven R. R. trips to N. Y. C.	- 1,539	- 9.17
Long Island R. R. trips to N. Y. C.	+10,951	- 9.41
Staten Island ferries to N. Y. C.	+ 4,554	+15.81
Pennsylvania R. R. to N. Y. C.—Mainline	+ 843	+ 4.22
PATH to N. Y. C.	- 963	- 4.59
Erie-Lackawanna R. R. to N. Y. C.	- 1,320	- 5.92
Central R. R. of N. Y. to N. Y. C.	- 297	- 2.54
Pennsylvania R. R. to N. Y. C.—Shore Branch	- 202	- 3.54

TABLE 4

SENSITIVITIES OF UNSTRATIFIED MODAL SPLIT EQUATION

CASE 1—HIGH DENSITY URBAN TO CBD (Brooklyn to Manhattan)

Present Values: ED = 230.860, RD = 55.553, TA = 50, TT = 52, SF = 1, L = 5, P = 50

Actual Transit Percent = 91.6
Predicted Transit Percent = 91.6

Variations	Δ%	New %
Add 10 minutes to TA	+ 4.0	95.6
Reduce TT by 10 minutes	+ 4.6	96.2
Double parking cost	+ 5.6	97.2
Provide free parking	- 5.6	86.0
Add 15 cents to tolls	+ 1.7	93.3
Provide free parking, halve employment density, and eliminate tolls	-11.6	80.0

CASE 2—SUBCENTER TO CBD (Trenton to Manhattan)

Present Values: ED = 230.800, RD = 2.127, TA = 100, TT = 99, SF = 1, L = 100, P = 50

Actual Transit Percent = 94.2
Predicted Transit Percent = 95.2

Variations	Δ%	New %
Double parking cost	+ 5.6	100.0+
Reduce TT by 15 minutes	+ 3.7	98.9
Double residential density	+ 1.9	97.1
Triple residential density	+ 3.0	98.2

CASE 3—SUBURBAN TO CBD (Suffolk to Manhattan)

Present Values: ED = 230.860, RD = 2.330, TA = 99, TT = 105, SF = 1, L = 9, P = 50

Actual Transit Percent = 74.6
Predicted Transit Percent = 75.6

Variations	Δ%	New %
Double parking cost	+ 5.6	81.2
Reduce TT by 25 minutes	+ 6.1	81.7
Double residential density	+ 1.7	77.3

CASE 4—SUBURBAN TO SUBCENTER (Milford to Bridgeport)

Present Values: ED = 3.904, RD = 3.164, TA = 19, TT = 39, SF = 0, L = 8, P = 7.5

Actual Transit Percent = 10.9
Predicted Transit Percent = 12.5

Variations	Δ%	New %
Triple parking cost	+ 1.7	14.2
Double employment density	+ 4.7	17.2
Reduce TT by 10 minutes	+ 3.4	15.9
Provide good rail service	+17.8	30.3
All 4	+27.6	40.1

CASE 5—SUBURBAN TO SUBURBAN (Western Essex Co. to Morris Co.)

Present Values: ED = 1.465, RD = 4.267, TA = 28, TT = 54, SF = 0, L = 0, P = 0

Actual Transit Percent = 4.3
Predicted Transit Percent = 4.7

Variations	Δ%	New %
Double residential density	+ 1.7	6.4
Double employment density	+ 5.3	10.0

- Parking costs were alternately doubled and eliminated, and this caused a variation of 5. 6 percent in each direction, illustrating the model's sensitivity to this variable.
- A $0.15 toll increase, simulating imposition of tolls on free East River crossings, was tried. This produced only a slight gain in the proportion of transit users.
- A combination of changes was tested to simulate an emphasis on private automobiles. Free parking was introduced, tolls were removed, and employment density was halved to provide for increased parking. Auto travel times were held constant, on the assumption that the increased automobile trips will be handled by new facilities. The result was a drop in transit usage from 91.6 percent to 80.0 percent. Auto trips would increase by 240 percent, from 8.4 percent to 20.0 percent of the total.

Other cases examined included travel from a subcenter to the CBD and from a suburban zone to the CBD. Of current interest is the anticipated 100-mph commuter train. The model shows that the accompanying 25-min improvement in transit time from a suburban zone would reduce auto usage by 25 percent.

Case 4, suburban zone to subcenter, was particularly important since an accurate forecast of transit usage into the subcenters would be of great value in evaluating the concept of satellite cities. To test the sensitivities in this O-D pair, we tripled parking costs, doubled employment density, improved transit time by 10 min and provided rapid rail transit service. These combined variations increased transit usage from 12.5 to 40.1 percent of all work trips.

The final case tested, suburban zone to suburban zone, showed very little increase of transit usage when densities were increased. This is not surprising for such interchanges.

Although we cannot be certain that the variations examined will produce the calculated reactions, the sensitivities of the equation do seem to be intuitively reasonable.

Modal Split Stratified by Income—We next developed equations for each of the three income groups ($0 to $5,000, $5,000 to $10,000, and greater than $10,000 per year). Although the unstratified equation is quite adequate, we wanted to investigate the possibility of even finer results. This can be done by examining income-stratified modal split equations. When such equations were explored previously with no system variables, the results were promising.

Low-Income Modal Split—For the low-income group, the same O-D pairs used for the unstratified equation were examined after those pairs with no low-income transit trips were eliminated. The regressions were run using the density and cost variations, the rail service factor, and the time ratio. Table 5 gives the results of these trials. The statistical results of the first four trials were not as good as they were for the unstratified equation. Since these results were based on a smaller number of trips, we felt that sampling errors resulted for a number of O-D pairs. Accordingly, as with the unstratified modal split analysis, O-D pairs were removed from consideration if their sample size was less than ten.

When the regressions were rerun, the cost ratio terms were omitted from the analysis, since they showed up poorly in the first four runs. This time, the resulting equation did not even include those cost terms that were tried. To examine the effect of the cost term in the equation, the F-level-to-enter[1] was lowered from 9. 00 to 4.00 so that the toll and parking cost variable could enter. Parking alone was also tried as the cost term, but it never entered the equation. The only decision remaining was whether to accept the equation with the tolls and parking cost value. The t-value[2] was 2. 2, indicating a considerable likelihood that the true coefficient was not the calculated one. While the statistical evidence for keeping the variable was not overwhelming,

[1]The F-level is a statistical index that enables the analyst to determine the probability that only chance factors cause the improvement in correlation resulting from adding a variable. Lowering the F-level increases the probability that the improvement in correlation is caused by chance factors.

[2]The t-value is a standard that enables the analyst to determine the probability that the coefficient computed using the sample has a value close to the coefficient that would be computed if the entire universe were considered.

TABLE 5

MODAL SPLIT REGRESSION ANALYSIS

Trial No.	Variables Attempted	Variables Entered	Standard Deviation	Standard Error	$\bar{Y}$	R	N	F-level to enter
		(a) Low Income						
1	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, L + P SF, TA/TT	ln ED, $\sqrt{RD}$, SF TA/TT, L + P	0.2636	0.1511	0.4057	0.821	613	9.00
2	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, (L + P)/CT	ln ED, $\sqrt{RD}$, SF TA/TT, (L + P)/CT	0.2636	0.1529	0.4057	0.816	613	9.00
3	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT CA	ln ED, $\sqrt{RD}$, SF TA/TT, CA	0.2636	0.1509	0.4057	0.821	613	9.00
4	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$ ln RD, ln ED, TA/TT, CA/CT	ln ED, $\sqrt{RD}$, SF TA/TT, CA/CT	0.2636	0.1519	0.4057	0.819	613	9.00
5	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, L + P, D, SF	ln ED, $\sqrt{RD}$, SF TA/TT	0.2602	0.1284	0.3876	0.871	557	9.00
6	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, CA, D, SF	ln ED, $\sqrt{RD}$, SF TA/TT	0.2602	0.1284	0.3876	0.871	557	9.00
7	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, L + P, SF	ln ED, $\sqrt{RD}$, SF TA/TT, L + P	0.2602	0.1274	0.3876	0.873	557	4.00
8	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$ ln RD, ln ED, TA/TT, L, SF	ln ED, $\sqrt{RD}$, SF, TA/TT, -RD, $\sqrt{ED}$, -ED, ln ED removed	0.2602	0.1263	0.3876	0.876	557	4.00
		(b) Middle Income						
1	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, L + P, SF	ln ED, SF, RD, L + P, TA/TT	0.2612	0.1161	0.2771	0.897	538	
2	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, CA, SF	ln ED, SF, RD, CA, TA/TT, $\sqrt{RD}$, RD removed	0.2612	0.1138	0.2771	0.901	538	
3	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, CA/CT, SF	ln ED, SF, RD, TA/TT, CA/CT	0.2612	0.1209	0.2771	0.888	538	
4	ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, L + P, SF	ln ED, SF, $\sqrt{RD}$, L + P, TA/TT	0.2608	0.1124	0.2763	0.903	537	
5	ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, L + P S_N, S_T	ln ED, RD, TA/TT, L + P, S_T, S_N	0.2608	0.1122	0.2763	0.904	537	
6	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, CA, SF	ln ED, SF, RD, CA, TA/TT, $\sqrt{RD}$, RD removed	0.2608	0.1108	0.2763	0.906	537	
7	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, CA, SF	ln ED, RD, TA/TT, SF, CA, $\sqrt{RD}$, RD removed	0.2608	0.1107	0.2763	0.907	537	
8	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT L + P	ln ED, RD, TA/TT, L + P, $\sqrt{RD}$	0.2608	1.1318	0.2763	0.864	537	
		(c) High Income						
1	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, L + P, SF	$\sqrt{ED}$, L + P, TA/TT, SF, -ED	0.2892	0.1341	0.2867	0.888	286	
2	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, L + P	$\sqrt{ED}$, L + P, TA/TT, ED, RD	0.2892	0.1368	0.2867	0.883	286	
3	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED TA/TT, CA, SF	$\sqrt{ED}$, CA, SF, TA/TT, RD	0.2892	0.1291	0.2867	0.897	286	
4	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, CA/CT, SF	$\sqrt{ED}$, TA/TT, SF, CA/CT	0.2892	0.1508	0.2867	0.855	286	
5	RD, $\sqrt{RD}$, ln RD, ln ED, TA/TT, L + P, SF	ln ED, L + P, TA/TT, SF	0.2892	0.1357	0.2877	0.885	285	
6	RD, ED, $\sqrt{RD}$, $\sqrt{ED}$, ln RD, ln ED, TA/TT, L + P, SF	$\sqrt{ED}$, L + P, TA/TT, SF	0.2892	0.1362	0.2877	0.884	285	
7	RD, $\sqrt{RD}$, ln RD, ln ED, TA/TT, L + P	ln ED, L + P, TA/TT, RD	0.2892	0.1391	0.2877	0.879	285	

some sensitivity to automobile costs in the low-income equation seemed desirable. Therefore, trial 7 was adopted, and the low-income modal split equation became:

$$\% \text{ TRANSIT} = 9.289 (\ln ED) + 2.978 \sqrt{RD} + 16.431SF + 17.447 (TA/TT) + 0.043 (L + P) - 8.997$$

Table 6 gives the statistical characteristics of the variables.

Middle-Income Modal Split—The trials for the modal split analysis of the middle-income group are also given in Table 5. As with the low-income group, O-D pairs were removed when the census survey showed no transit trip-makers or when the total number of trips were represented by fewer than ten interviews.

Essentially the same variables that were tried for the unstratified and low-income analyses were tried with the middle-income analysis. Each of three cost variables was alternately tried with the density, time, and rail service factor variables. Trial 4 was selected even though trials 5 through 7 gave slightly better statistical results. We rejected these trials because the marginal improvement afforded by including over-the-road costs and two rail service factors would not warrant the added labor of determining their values for forecast years. Furthermore, using two service factors presented the additional difficulty of defining them and distinguishing between them for a future system. Consequently, this equation was selected for the middle-income modal split:

$$\% \text{ TRANSIT} = 7.251 (\ln ED) + 20.572SF + 2.067 \sqrt{RD} + 0.167 (L + P) + 21.875 (TA/TT) - 19.584$$

Table 6 gives the statistical characteristics of the variables.

High-Income Modal Split—The high-income modal split equation was developed in the same way as the low- and middle-income equations (Table 5). Trial 5 was preferred over trial 3 for the same reasons mentioned in discussing the middle-income equation. The high-income modal split equation was:

$$\% \text{ TRANSIT} = 7.010 (\ln ED) + 0.307 (L + P) + 25.840 (TA/TT) + 11.399SF - 20.413$$

Note that no residential density term appears in this equation. Table 6 gives the statistical characteristics of the equation variables.

Comparisons of the Income-Stratified Equations—To understand mode choice, it is useful to compare the three income-stratified equations. The crucial question is

TABLE 6

STATISTICAL CHARACTERISTICS OF THE VARIABLES

Variable	F-level Entering	t-Value Entering	Final Coefficient	Standard Error of Coefficient	Final t-Value
(a) Low-Income Equation					
ln ED	556.8	23.6	9.289	0.511	18.1
RD	307.3	17.5	2.978	0.189	15.8
SF	137.8	11.7	16.431	1.476	11.1
TA/TT	40.0	6.3	17.447	2.948	5.9
L + P	4.8	2.2	0.043	0.020	2.2
(b) Middle-Income Equation					
ln ED	679.6	26.1	7.251	0.482	15.1
SF	229.3	15.1	20.572	1.452	14.1
RD	139.8	11.8	2.067	0.173	11.9
L + P	93.7	9.7	0.167	0.018	9.3
TA/TT	68.4	8.3	21.875	2.645	8.3
(c) High-Income Equation					
ln ED	495.6	22.3	7.010	0.826	8.5
L + P	76.5	8.7	0.307	0.032	9.6
TA/TT	57.7	7.6	25.840	3.988	6.5
SF	25.9	5.1	11.399	2.241	5.1

whether the comparative values of the variables' coefficients are reasonable when related to our intuitive understanding of the three income groups. Table 7 should be helpful in answering that question. For four out of five cases, a pattern is readily discernible. As the income level rises, the employment and residential density variables carry less weight and the time ratio and cost variables carry more weight. The service factor follows an erratic pattern, increasing in value from low- to middle-income level, but decreasing for the high-income level. The constant term of the equations drops as the income level rises.

What can we infer from these comparisons? Low-income groups are more susceptible to their environment; that is, the choice to travel to work by automobile or transit is determined most often by the variables that measure the characteristics of the home and work locations, and not by the characteristics of the transportation system between travel points. The relative insignificance of the cost term for the low-income group reflects the fact that the poor person rarely considers automobile costs, probably because he does not have access to an automobile. He is the "captive rider" of the transit system and usually locates his home and job accordingly, i.e., in high-density locations. The high-income person, on the other hand, is likely to have a choice—to be able to consider freely the merits of the alternate transportation modes; hence the great significance of the cost variable for the high-income individual. People in the middle-income group are a combination of captive and choice riders: less beholden to the environment than the low-income group, yet not as free as the high-income group to choose a mode. The values of the constant terms suggest that as income rises the likelihood of using transit declines.

The one inconsistency in Table 7 is the behavior of the rail service factor coefficient, which decreases from the middle- to the high-income group. This result contradicts the thesis that the higher the income is, the more significant are the transportation system characteristics. This inconsistency might be explained by the contrast between the New York City subway system and the commuter railroads. Both warrant service factors, but for very different service. The high-income individual is not likely to be greatly influenced by the service factor for O-D pairs served by the subway. The subway is probably less attractive than his automobile, while the commuter railroad is probably a satisfactory alternative to driving. People in the low- and middle-income groups are likely to find the subway cost more attractive than the high cost of owning and using a car to drive to work; hence the higher coefficients.

Therefore, the variations in relative parameter values for the three income-stratified equations appear to be adequately explained by sound intuitive reasoning.

The Stratified Equations Versus the Unstratified Equation

The stratified equations must pass the same tests that the unstratified equation passed: they must have a predictive ability as measured by the RMS error; they must not introduce geographic or transportation biases; and their sensitivity toward changes in the variables must be reasonable. These analyses can be used both to examine the validity of the equations and to compare the unstratified equation with the income-stratified equations.

TABLE 7

PARAMETER COMPARISONS OF INCOME-STRATIFIED EQUATIONS

Parameter	Low		Middle		High	
	Coeff	Rank	Coeff	Rank	Coeff	Rank
ln ED	9.289	1	7.251	1	7.010	1
RD	2.978	2	2.067	3	—	—
SF	16.431	3	20.572	2	11.399	4
TA/TT	17.447	4	21.875	5	25.840	3
L + P	0.043	5	0.167	4	0.307	
the constant	-8.997		-19.584		-20.413	

The RMS error for the stratified equations is 29.7 trips. Since the unstratified equation has an RMS error of 30.1, the equations show a similar level of predictive ability.

Analysis of geographical and transportation bias for the stratified equations was carried out in the same manner as for the unstratified equation (Table 8). Stratified equations produced better results in 10 of the 18 trip categories examined, but the unstratified equation showed up better for the important categories, "All transit trips to Manhattan" and "Subway trips in New York City." Examination for biases showed that both the unstratified equation and the stratified equations produce reasonable results.

The third means of evaluating the validity of the stratified equations is to test their sensitivities to changes in the independent variables. There were really three separate but related analyses to be made. First, the sensitivities of the stratified equations were examined. Second, we compared the effect of changes in the variables for both the stratified equations and the unstratified equations, based on the present income level distribution. Third, a comparison was made of the effects of changes in the variables based on a radically different income level distribution.

Table 9 illustrates the sensitivities of the income-stratified equations to changes in the variables. Comparable data for the unstratified equation are reproduced from Table 6. Case 1 shows that changes favoring transit did not produce as great an increase in transit percent usage for the stratified equations as for the unstratified equation. This was because the low-income equation produced a transit percent usage of over 100 percent when the improvement was made, and this had to be adjusted back to less than 100 percent. The same phenomenon is seen in Case 2, where the transit usage increases that accompany a twofold increase in parking cost are not as great for the stratified set of equations. Cases 3, 4, and 5 illustrate that the sensitivities of the stratified equations are reasonable. The unstratified equation has similar sensitivities for these O-D pairs and the variety of variable changes, indicating that similar results would be obtained using either the stratified or unstratified equations when testing a proposed land-use and transportation plan.

However, similar results would not necessarily occur if there are sweeping changes in the distribution of income levels. The coefficients of the three income-stratified modal split equations vary significantly for each of the variables. Therefore, any major change in the income level distribution may affect the modal split results.

The Regional Plan Association (3) has forecast that for the metropolitan region the percent of households earning less than $5,000 (in 1960 dollars) will decrease from

TABLE 8

COMPARISON OF UNSTRATIFIED AND STRATIFIED MODAL SPLIT BIAS ANALYSIS

Category	Unstratified		Stratified	
	Error	% Error	Error	% Error
All transit trips to Manhattan	+ 654	+ 0.07	+ 9,460	+ 0.96
All transit trips to Brooklyn	- 8,202	- 2.13	- 4,479	- 1.16
All transit trips to Queens	+15,889	- 9.43	+14,352	+ 8.52
All transit trips to the Bronx	+ 7,714	+ 5.56	+ 9,347	+ 6.77
All transit trips to Newark	+ 8,339	+15.59	+ 9,664	+18.06
Bus trips to N. Y. C.—eastbound	+ 1,364	+ 2.31	+ 1,052	+ 1.78
Trans-Hudson transit trips—eastbound	- 2,280	- 1.64	- 1,564	- 1.08
Trans-Hudson transit trips—westbound	- 746	- 3.11	- 1,335	- 5.55
Subway trips in N. Y. C.	-13,236	- 0.67	-30,569	- 1.55
New York Central R. R. trips to N. Y. C.	- 494	- 0.81	- 1,191	- 1.96
New Haven R. R. trips to N. Y. C.	- 1,539	- 9.17	- 1,099	- 6.55
Long Island R. R. trips to N. Y. C.	+10,951	+ 9.41	+ 5,773	+ 4.96
Staten Island ferries to N. Y. C.	- 4,554	-15.81	- 4,168	-14.47
Pennsylvania R. R. to N. Y. C.—Mainline	+ 843	+ 4.22	+ 1,055	+ 5.28
PATH to N. Y. C.	- 963	- 4.59	- 607	- 2.89
Erie-Lackawanna R. R. to N. Y. C.	- 1,320	- 5.92	- 1,489	- 6.68
Central R. R. of N. Y. to N. Y. C.	- 297	- 2.54	- 276	- 2.36
Pennsylvania R. R. to N. Y. C.—Shore Branch	- 202	- 3.54	+ 120	+ 2.11

TABLE 9

SENSITIVITIES OF STRATIFIED AND UNSTRATIFIED MODAL SPLIT EQUATIONS

CASE 1—HIGH DENSITY URBAN TO CBD (Brooklyn to Manhattan)

Actual Percent Transit = 91.6
Predicted Pereent Transit—Stratified = 94.5
Predicted Percent Transit—Unstratified = 91.6

Variations	Stratified		Unstratified	
	Δ%	New %	Δ%	New %
Add 10 minutes to TA	+ 1.5	96.0	+ 4.0	95.6
Reduce TT by 10 mintues	+ 1.6	96.1	+ 4.6	96.2
Double parking costs	+ 3.2	97.7	+ 5.6	97.2
Provide free parking	- 6.0	88.5	- 5.6	86.0
Add 15 cents to tolls	+ 0.8	95.3	+ 1.7	93.3
Provide free parking, halve employment densities, and eliminate tolls	-10.7	83.8	-11.6	80.0

CASE 2—SUBCENTER TO CBD (Trenton to Manhattan)

Actual Percent Transit = 94.2
Predicted Percent Transit—Stratified = 95.7
Predicted Percent Transit—Unstratified = 95.2

Variations	Stratified		Unstratified	
	Δ%	New %	Δ%	New %
Double parking cost	+ 2.3	98.0	+ 5.6	100.0
Reduce TT by 15 minutes	+ 3.3	99.0	+ 3.7	98.9
Double residential density	+ 1.3	97.0	+ 1.9	97.1
Triple residential density	+ 2.3	98.0	+ 3.0	98.2

CASE 3—SUBURBAN TO CBD (Suffolk to Manhattan)

Actual Percent Transit = 74.6
Predicted Percent Transit—Stratified = 75.6
Predicted Percent Transit—Unstratified = 75.7

Variations	Stratified		Unstratified	
	Δ%	New %	Δ%	New %
Double parking cost	+ 8.1	83.8	+ 5.6	81.2
Reduce TT by 25 minutes	+ 6.1	81.8	+ 6.1	81.7
Double residential density	+ 0.8	76.5	+ 1.7	77.3

CASE 4—SUBURBAN TO SUBCENTER (Milford to Bridgeport)

Actual Percent Transit = 10.9
Predicted Percent Transit—Stratified = 11.6
Predicted Percent Transit—Unstratified = 12.5

Variations	Stratified		Unstratified	
	Δ%	New %	Δ%	New %
Triple parking cost	+ 1.4	13.0	+ 1.7	14.2
Double employment density	+ 5.4	17.0	+ 4.7	17.2
Reduce TT by 10 minutes	+ 3.2	14.8	+ 3.4	15.9
Provide good rail service	+17.7	29.3	+17.8	30.3
All 4	+27.7	39.3	+27.6	40.1

CASE 5—SUBURBAN TO SUBURBAN (Western Essex Co. to Morris Co.)

Actual Percent Transit = 4.3
Predicted Percent Transit—Stratified = 4.4
Predicted Percent Transit—Unstratified = 4.7

Variations	Stratified		Unstratified	
	Δ%	New %	Δ%	New %
Double residential density	+ 1.3	5.7	+ 1.7	6.4
Double employment density	+ 4.4	8.8	+ 5.3	10.0

TABLE 10

SENSITIVITIES OF STRATIFIED MODAL SPLIT EQUATIONS FOR PRESENT AND FUTURE INCOME LEVEL DISTRIBUTIONS

CASE 1—HIGH DENSITY URBAN TO CBD (Brooklyn to Manhattan)

Actual Percent Transit = 91.6
Predicted Percent Transit—Stratified—Present = 94.5
Predicted Percent Transit—Stratified—Future = 83.9

	Present Income Distribution (Low = 63, Mid = 32, High = 5)		Future Income Distribution (Low = 20, Mid = 50, High = 30)	
Variation	Δ%	New %	Δ%	New %
Add 10 minutes to TA	+ 1.5	96.0	+ 4.4	88.3
Reduce TT by 10 minutes	+ 1.6	96.1	+ 4.8	88.7
Double parking costs	+ 3.2	97.7	+ 9.2	93.1
Provide free parking	- 6.0	88.5	- 8.9	75.0
Add 15 cents to tolls	+ 0.8	95.3	+ 3.1	87.0
Provide free parking, halve employment density, and eliminate tolls	-10.7	83.8	-14.4	69.5

CASE 2—URBAN TO CBD (Trenton to Manhattan)

Actual Percent Transit = 94.2
Predicted Percent Transit—Stratified—Present = 94.5
Predicted Percent Transit—Stratified—Future = 97.8

	Present Income Distribution (Low = 45, Mid = 46, High = 9)		Future Income Distribution (Low = 10, Mid = 40, High = 50)	
Variation	Δ%	New %	Δ%	New %
Double parking cost	+ 2.3	98.0	+ 1.8	99.6
Reduce TT by 15 minutes	+ 3.3	99.0	+ 2.0	99.8
Double residential density	+ 1.3	97.0	+ 0.7	98.5
Triple residential density	+ 2.3	98.0	+ 1.2	99.0

CASE 3—SUBURBAN TO CBD (Suffolk to Manhattan)

Actual Percent Transit = 74.6
Predicted Percent Transit—Stratified—Present = 75.6
Predicted Percent Transit—Stratified—Future = 72.4

	Present Income Distribution (Low = 24, Mid = 56, High = 20)		Future Income Distribution (Low = 5, Mid = 15, High = 80)	
Variation	Δ%	New %	Δ%	New %
Double parking cost	+ 8.1	83.8	+13.6	86.0
Reduce TT by 25 minutes	+ 6.1	81.8	+ 7.2	79.6
Double residential density	+ 0.8	76.5	+ 0.3	72.7

CASE 4—SUBURBAN TO SUBCENTER (Milford to Bridgeport)

Actual Percent Transit = 10.9
Predicted Percent Transit—Stratified—Present = 11.6
Predicted Percent Transit—Stratified—Future = 8.0

	Present Income Distribution (Low = 41, Mid = 51, High = 8)		Future Income Distribution (Low = 10, Mid = 50, High = 40)	
Variation	Δ%	New %	Δ%	New %
Triple parking cost	+ 1.4	13.0	+ 3.2	11.2
Double employment density	+ 5.4	17.0	+ 5.1	13.1
Reduce TT by 10 minutes	+ 3.2	14.8	+ 3.9	11.9
Provide good rail service	+17.7	29.3	+16.5	24.5
All 4	+27.7	39.3	+28.7	36.7

CASE 5—SUBURBAN TO SUBURBAN (Western Essex Co. to Morris Co.)

Actual Percent Transit = 4.3
Predicted Percent Transit—Stratified—Present = 4.4
Predicted Percent Transit—Stratified—Future = 1.1

	Present Income Distribution (Low = 42, Mid = 42, High = 16		Future Income Distribution (Low = 10, Mid = 30, High = 60)	
Variation	Δ%	New %	Δ%	New %
Double residential density	+ 1.3	5.7	+ 0.5	1.6
Double employment density	+ 4.4	8.8	+ 2.2	3.3

37.2 percent to 12.8 percent between 1960 and 2000, and that the percent of households earning more than \$10,000 (in 1960 dollars) will increase from 19.8 percent to 71.3 percent in the same period. We therefore felt it was advisable to compare the modal split sensitivities for the present income level distribution with those for a hypothetical distribution having a greater percent of high-income individuals. All other variables were held constant for this analysis. Table 10 illustrates the comparison, using the five cases discussed previously.

For all five cases studied, the predicted percent transit usage for a hypothetical future income distribution was appreciably different from what it was for the present income distribution. It is obvious that a major shift of the population to higher income levels can alter the modal split.

Cases 3, 4, and 5 show that the projected higher income distribution reacts more strongly to cost changes than does the present income distribution. It is also apparent that changes in employment densities or residential densities produce a weaker reaction for the future income distribution. These results are not surprising in light of the previous comparisons of the coefficients. Cases 1 and 2 do not permit effective comparisons of the two income distributions because of the limit of 100 percent transit discussed earlier.

The population shift into higher income levels will show a strong tendency to alter the modal split in favor of the automobile. As the trip-maker acquires the ability to support more automobiles for his family, he will also have a freer choice between transit and automobile. He will no longer choose transit because he does not have an automobile available; he will choose it only if it truly is the preferable mode of travel. He will become increasingly sensitive to transportation system characteristics and less sensitive to the characteristics of his home and work locations.

Comparison of the RMS error and the geographical and transportation biases of the unstratified equation with the stratified equations produced no clear-cut choice between them. When the existing distribution of income levels was applied, the sensitivities of the variables to changes in their values showed that both unstratified and stratified equations produced reasonable results that were not dissimilar. However, when future income level distributions were applied using the stratified equations, significant differences occurred both in the forecast modal split and in the sensitivities of the changing variables to transportation system and land-use changes.

ACKNOWLEDGMENTS

The work described in this paper was developed jointly by the Regional Plan Association and Peat, Marwick, Livingston & Co. A detailed presentation of the findings derived from the research and development described in this paper will be published by the Regional Plan Association in the Transportation Volume of the Second Regional Plan Series.

The author gratefully acknowledges the comments of Boris Pushkarev and Ronald Greenwald of the Regional Plan Association and of Donald Hill and Arnold Kupferman of Peat, Marwick, Livingston & Co., who have reviewed the technical content of this paper.

REFERENCES

1. Unites States Department of Labor, Bureau of Labor Statistics. Special Labor Force Report No. 80, Marital and Family Characteristics of Workers. March 1966.
2. Lansing, John B., and Hendricks, Gary. How People Perceive the Cost of the Journey-to-Work. Survey Research Center, Institute for Social Research, Univ. of Michigan, Jan. 1967.
3. Regional Plan Association. The Region's Growth—A Report of the Second Regional Plan. May 1967.