Statistical Controls in Ridesharing Demonstration Programs

DAVID T. HARTGEN AND JOANNA M. BRUNSO

The application of scientific experimental designs in ridesharing demonstrations is discussed. A review of typical designs, particularly those that use test and control groups and over-time observations of behavior, suggests that the numerous problems that jeopardize the validity of studies could be reduced or eliminated. Nine possible outcomes of demonstrations are reviewed and interpreted against the need for experimental designs. Two applications in the ridesharing area are then described: one conducted during a period of rapid background change (1979 energy crisis) and the other in a recent period of stability (1981). In both cases (conducted at employer and community sites in the Albany, New York, area), the use of a control group and before-and-after background surveys permitted isolation of the true effects of the demonstration. In the first case (1979 energy crisis), this included the direct effect of the program (from coordinator records), indirect effects (from the existence of the ridesharing program itself), and external effects (from the energy crisis). In the second case (stable background), the indirect and external effects were found to be negligible. From this study it is concluded that the use of scientific designs in ridesharing analysis should be increased and expanded.

Government-sponsored carpooling programs were begun as a response to the 1973-1974 energy crisis and focused on matching services by using grid systems and computerized match-ups for interested employees (1). But consumer interest fell sharply as the crisis abated, and two-thirds of the programs initiated were discontinued. For those programs that did continue, promotional campaigns were expanded and the focus was on consumer economic savings. Public interest again increased during the 1979 energy crisis, but again subsided. Although this suggests that consumer interest in carpooling was closely related to the energy crises, the precise nature of this relation was not determined, and subsequent federal policy treated carpooling as a viable transporation system management (TSM) option.

A basic problem in carpool program evaluation is that most programs are not set up with careful evaluation in mind. As a result, most programs contain numerous technical problems that prevent a fair assessment of their impacts. Few programs separate existing and newly created carpoolers or follow up on carpools actually formed from inquiries. In addition, high failure rates have prevented a careful look at many programs.

Basic problems with carpool evaluations include failure to (a) sort out background (e.g., energy crisis), (b) separate created from discovered carpoolers, (c) consider carpool breakups, (d) account for additional travel by cars left at home or circuity of carpool trips, and (d) generalize to the appropriate population. It is recognized that these programs suffer from such lack of control that evaluation of true effects is generally not possible. Considering that the effect of such failures is to overstate the impact of the programs, taxpayers would be better served by a more careful assessment of the data.

The purpose of this paper is to suggest that through the use of statistical controls, such an assessment is feasible and possible and does not necessarily reflect negatively on carpool programs. A number of straight-forward carpool designs are described, which are based on experimental principles that have been found to be effective in assisting in these assessments.

PRINCIPLES OF STATISTICAL DESIGNS

Statistical designs evolved from the tradition of scientific experiments and are intended to isolate and quantify the causal linkages in analytical relations. The designs usually contain the following elements:

1. A test group (or individuals) selected to receive the service or treatment;
2. A control group that does not receive treatment but is monitored over time;
3. Before-and-after observations of behavior, attitude, status, and so on, of members of the test and control groups; and
4. Internal observations (records) that permit reporting and evaluation of the direct effects of the treatment.

In classic experiments, identical units are obtained, but only one is treated. In the social sciences we cannot obtain identical individuals, so units are randomly selected (or randomly assigned). Randomly selected (or assigned) individuals or groups are then treated with services or policies, with background factors allowed to vary; the resulting causal linkage is inferred from the differences in responses from differently treated groups. Basic statistical designs involve the use of a test service or treatment (X) and a series of observations (O) of the behavior of the tested (or other) entities. Basic common designs in the transportation literature are

1.	X 0	One-shot case study;
2.	0_1 X 0_2	One group pretest and posttest;
3.	R 0_1 X R 0_2 / R 0_3 R 0_4	Pretest and posttest with control groups and random assignment (r) of observation;
4.	R|0_1 X 0_2 / R|0_3 0_4	Nonequivalent control group; entire group rather than individual groups is assigned (randomly) to test or control; and
5.	0_1 X 0_2 / 0_3 0_4	Predetermined nonequivalent control group

Because designs 4 and 5 are often conducted in real-world settings rather than in laboratories, and the nature of the control is inexact, they are often called quasi-experiments.

The extent of the causal inferences that can be drawn depends on the nature of the design and the strengths of the controls. Campbell and Stanley (2) review the designs most often used and describe their limitations. They describe two kinds of validity of the study: (a) internal validity, which refers to conclusions drawn about the experiment itself, and (b) external validity, which refers to conclusions drawn (from the experiment) about a larger population. In each case, many factors can mask the design and threaten validity. The primary concerns that jeopardize internal validity are

1. History--events occur between the first and second measurements;

Table 1. Properties of some common experimental designs.

			Internal							
No.	Design	Name	History	Maturation	Testing	Instrumentation	Statistical Regression	Differential Selection	Mortality	Selection and Maturation
1	X 0	Case study	–	–				–	–	
2	0_1 X 0_2	Pretest and posttest	–	–	–	–	?	+	+	–
3	R 0_1 X R 0_2 / R 0_3 R 0_4	Random pretest and posttest with control	+	+	+	+	+	+	+	+
4	R|0_1 X 0_2 / R|0_3 0_4	Nonequivalent control group (random group)	+	+	+	+	?	+	+	–
5'	0_1 X 0_2 / 0_3 0_4	Predetermined nonequivalent control group[a]	+	+	+	+	+	+	+	–

Note: – = weakness, + = factor is controlled, and ? = possible problem.

[a]Assumes equivalent response by test and control groups.

Table 2. Effect of statistical designs on carpool program evaluation.

	Design				
Problem	1	2	3	4	5
Changes in background that encourage carpooling			X	X	X
Carpool impact in a flat background		X	X	X	X
Magnitude of uncovered versus created carpools			X	X	X
Changes in questionnaire format in before versus after surveys			X	X	X
Carpool breakups		–[a]	X	X	X
Differential impact of program by user group		–[a]	X		
Differential impact of program by site			X	X	
Changes in background that decrease carpooling			X	X	X

Note: X = design accounts for these problems.

[a]From internal records.

2. Maturation--subjects age or otherwise change naturally, thus changing behavior or sensitivity to the experiment;
3. Testing--test takers better understand or become more familiar with the questions;
4. Instrumentation--changes in test procedure, questions, and observers;
5. Statistical regression--tendency for extreme points to drift toward the mean on repeat observations;
6. Differential selection--differences of respondents or subjects for test and control groups;
7. Mortality--subjects die, resign, or cannot be recontacted differentially between test and control groups; and
8. Selection and maturation interaction--subjects in certain behavioral groups change or age more rapidly.

The primary threats to external validity are

1. Reactive effect of testing--questionnaire itself causes a change in behavior or inclination in subjects;
2. Selection and experiment interaction--subjects self-choose to participate from interest;
3. Reactive effects of experiment--service or test itself causes changes in behavior; and
4. Multiple treatment--effects of multiple testing or treatments on subjects cannot be erased.

The above designs only partly control for basic internal threats and some external threats to validity. The data in Table 1 (2) summarize the capabilities of each design. The data in Table 2 indicate how each design handles typical problems concerning carpool program evaluation. It is clear from these tables that the designs currently in most common use (case study and pretest and posttest) do not adequately address most validation problems because no control group is available for isolation of most effects. But even the more complex designs do not remove threats to external validity.

Transportation policy studies rarely permit random assignment of individuals to receive treatment (e.g., a new service or lower fare), so designs 3 and 4, which involve random assignment, are not often conducted in real-world settings, although they have been conducted in laboratory or classroom settings.

A particularly useful feature of designs is that the external impact of the test can be separated from its internal impact. For example, in the random pretest and posttest design (number 3 in Table 1), 0 represents observations of behavior, attitude, and so on; these are usually determined from a sample drawn from a larger population, and means ($\bar{0}$) are calculated to estimate average values.

As an example, from the population there will be a random assignment sample, which can be set up as follows:

$\bar{0}_1$ X $\bar{0}_2$ (test)

$\bar{0}_3$ $\bar{0}_4$ (control)

where (a) internal effects are effects caused by the program (X), which consist of direct effects (i.e., effects caused directly by the program) and indirect effects (i.e., effects caused indirectly by the existence of the program); and (b) external effects, which are effects caused by changes in the background. Total program effects are isolated by comparing the differences [i.e., total program effects = internal effects - external effects, or $TOT_x = (\bar{0}_2 - \bar{0}_1) - (\bar{0}_3 - \bar{0}_4)$].

The internal effect consists of two parts: the direct impact of the program ($Dir._x$), which can be determined directly from the internal records of the study (e.g., number of new carpoolers attracted, new transit riders), and the indirect (additional halo) effects ($Ind._x$), i.e., $TOT_x = Ind._x + Dir._x$.

The null hypothesis is that, if the program has no effect, there should be no difference in the changes observed in the test versus control group statistics; that is $(\bar{0}_2 - \bar{0}_1) - (\bar{0}_4 - \bar{0}_3) = 0$. Standard statistical procedures for the significance of these differences are readily available.

Interpretation of results from such studies depend generally on the strength and direction of

External			
Reactive Testing	Selection and Experiment	Reactive Experiment	Multiple Treatment
	–		
–	–	?	
–	?	?	
–	?	?	
–	?	?	

changes observed in the test and control groups. Nine possible results are shown in Figure 1. The results can be most confidently interpreted when test and control groups diverge in changes in behavior (Δtest +, Δcontrol –; or Δtest –, Δcontrol +). However, most studies do not yield such clear results. Particular care should be taken in situations in which the test appears to have little or no effect but, when compared with the control, it is seen to slow a declining process (e.g., Δtest 0, Δcontrol –; Δtest –, Δcontrol –). An example from the transportation field would be transit fare-saver programs that halt or slow declines in transit ridership. Without a control group, it is difficult to estimate what the ridership would be if the fare had not been saved.

APPLICATIONS

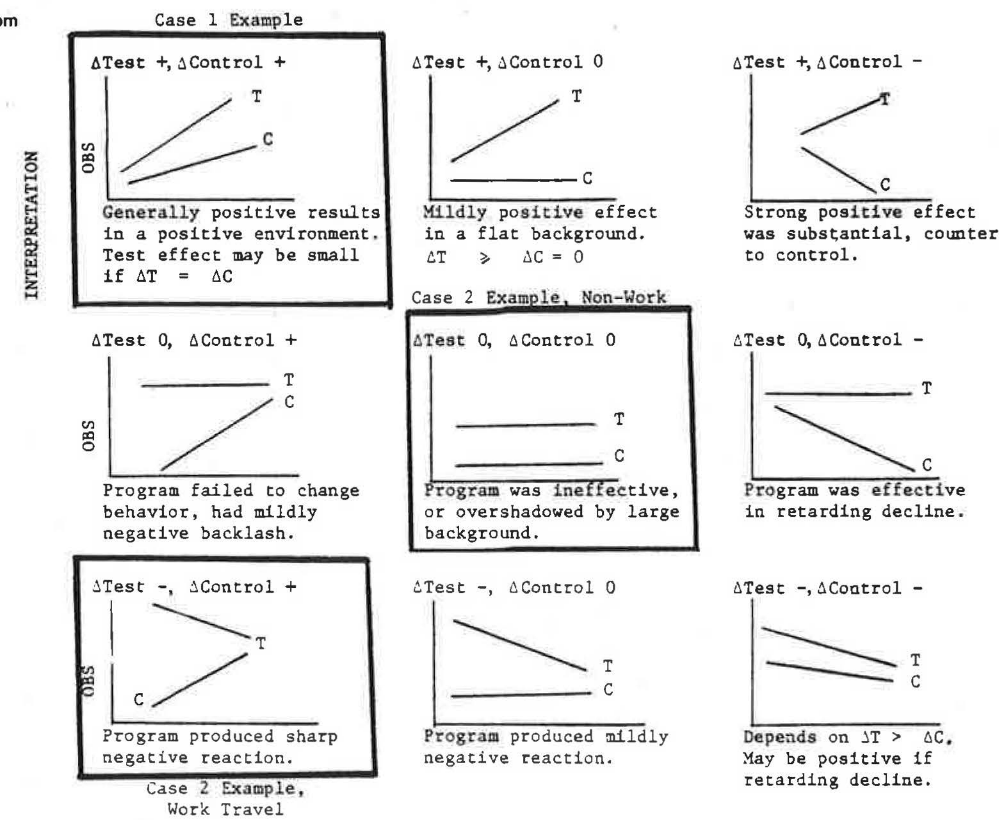
Although the applications of these principles are widespread in the literature (education, psychology, and sociology), the use of these principles in transportation studies is limited. Dunbar (3) describes the errors in model coefficients that occur from using cross-sectional data and calls for the integrated use of before-and-after data and cross-sectional data in model building. Although literally thousands of examples of the use of before-and-after data sets to infer internal validity exist in transportation, most do not have controls or the context structured so that the effect of the policy can be isolated. Louviere and others (4) describe a number of experiments (using college students) that deal with the rating of shopping sites in the Laramie, Wyoming, area, but generalization to the behavior of nonstudents shows marked differences in the laboratory and real-world models.

Tischer and Phillips (5) describe the use of a similar technique--the cross-lagged panel--which involves repeat observations of behavior and attitudes of a group of commuters over time, during which the treatment (in this case, a carpool lane) is introduced. In a later report Tischer (6) used the same data set to test whether structural models of modal use changed as a result of the introduction of the carpool lane service (the conclusion was negative). Finally, McClelland and others (7) describe the pitfalls of relying solely on aggregate changes in behavior over time without having detailed positive and negative switching behavior.

In addition to the above studies, several applications to ridesharing, conducted by the New York State Department of Transportation (NYSDOT), are discussed below. These examples demonstrate how the concepts described above have been used to evaluate carpooling programs. The cases are drawn from recent NYSDOT studies of carpooling services in employer and community-based settings.

These two demonstrations took place within economic and energy settings that were radically different from each other. The employer-based ridesharing demonstration took place in a year of

Figure 1. Interpretation of results from statistical designs.

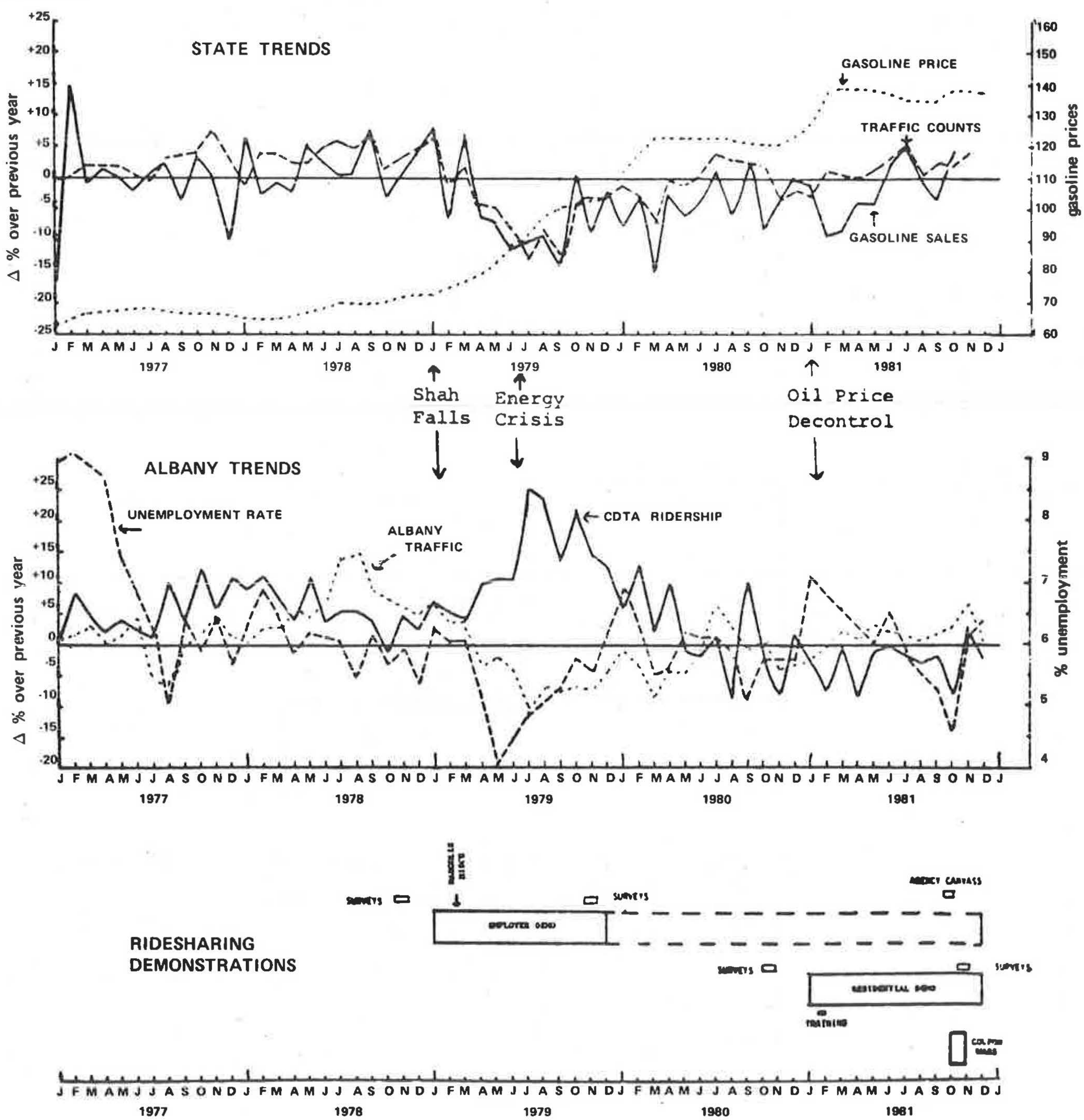


uncertainty and adversity that generally favored carpooling, whereas the neighborhood ridesharing demonstration took place in a year of general optimism and economic growth--trends that likely retarded increases in ridesharing. Further, the sharpest part of the decline in 1979 was much greater than the greatest rise of optimism in 1981; therefore, it was expected that more moderate changes in travel behavior would occur during the 1981 demonstration relative to the 1979 demonstration.

The chronology of national and local events associated with each of these demonstrations is shown in Figure 2. The employer-based demonstration was coincidentally initiated at the time of the fall of the Shah of Iran in January 1979 and spanned the difficult period of the 1979 energy crisis. During that period, gasoline prices rose sharply, gasoline supplies dropped by a maximum of 13 percent, and traffic declined substantially compared with previous years. Unemployment, which had been quite low in the early months of 1979, rose sharply, as did transit ridership. These events encouraged ridesharing by workers and precipitated a major increase in ridesharing within the state agencies surveyed.

On the other hand, the circumstances surrounding the neighborhood ridesharing demonstration, which began in January 1981, were substantially different. This was a period of moderate stability in gasoline prices after an initial rise following oil

Figure 2. Events associated with NYSDOT ridesharing demonstration.



price decontrol. Traffic, which had been down in 1979 and 1980, rose steadily through this period, as did gasoline sales. Conversely, transit ridership dropped and unemployment declined (until the end of the period when it began to rise as a result of the 1981 recession). These trends indicated increased flexibility on the part of drivers and a generally rising economy; there was also a predictable rise in traffic and gasoline sales and a decline in transit use. All of these trends favor declining or stable ridesharing behavior.

Case 1: Employer-Based Ridesharing Coordination Program During an Energy Crisis

The employer-based ridesharing demonstration formally began in January 1979. Originally funded by the New York State Energy Office, the demonstration evaluated the effectiveness of coordinators in work sites over a 1-year period. Six New York State agencies participated: three as test and three as control agencies.

In order to measure the effectiveness of the program in forming and maintaining carpools, a before-and-after study with control and test groups was planned. This study was undertaken on both the uptown (Campus) and downtown (Nelson A. Rockefeller Plaza) sites of the New York State government in Albany. These areas represent two extremes in their respective accessibilities and parking situations. The Campus area has generally ample parking and is easily accessible through several major routes. The downtown area, however, has severe parking restrictions and suffers from the congestion of the Albany central business district (CBD).

The test designers were concerned that the incidence of carpooling might also be affected by other events, including another oil embargo or changing transportation service to downtown Albany.

To ensure a strong test, New York State agencies at both the Campus and downtown Albany locations were surveyed before and after the demonstration program. These included nonparticipating (control) agencies, against which change in carpooling--independent of the carpool coordinator project--could be measured. Other agencies (test agencies) were provided with carpool coordinators.

The design used for this experiment (8,9) was a version of design 5 (predetermined, nonequivalent control group). Agencies were assigned to test or control status primarily on the basis of willingness to participate and general demographic similarity. As with all such designs, the choice must be a compromise between statistical appropriateness and administrative and institutional concerns. The resulting design attempts to control for background effects as well as agency location (see Table 3).

A random sample of approximately 150 employees from each agency was surveyed at the beginning of the demonstration, and a separate random sample was surveyed again at the end of the demonstration project. Comparison of the before (0_a) and after (0_b) data on both the test (X) and control agencies would uncover any significant changes in carpool formation, method of travel to work, and attitudes toward carpooling. The effect of various demographic characteristics (e.g., age, sex, income, family size, and automobile ownership) on carpool formation and continuance were also investigated. The initial survey was distributed in November 1978, and the project was initiated in early December 1978. The follow-up survey was undertaken again for each agency in October 1979.

Agency changes in mode to work are given in Table 4. The results indicated that in test agencies the carpool coordinators increased ridesharing substantially (10 percentage points), whereas ridesharing among control agencies rose 3.5 percentage points

Table 3. Evaluation design for employer-based demonstration.

			External Data		
Site	Agency	Status	Before (November 1978)	Carpool Coordinator Demonstration[a]	After (September 1979)
State Campus	Transportation	Test	O_a	X	O_b
	Labor	Control	O_a		O_b
Rockefeller Plaza	Motor Vehicle	Test	O_a	X	O_b
	General Services	Test	O_a	X	O_b
	Health	Control	O_a		O_b
	Public Service	Control	O_a		O_b

Note: O's represent observations of carpooling (i.e., an employee survey) and X's represent carpool coordinator activities.
[a]Internal records.

Table 4. Changes in mode to work for employer-based coordinator program.

	Changes in Mode to Work (%)																	
	Drive Alone			Drive with Another Employee			Drive with Family Member			Transit			Walk			Other		
Item	1978	1979	Δ	1978	1979	Δ	1978	1979	Δ	1978	1979	Δ	1978	1979	Δ	1978	1979	Δ
Agency																		
Transportation[a]	54	42	-12	27	39	12	11	16	5	5	2	-3	3	0	-3	1	1	0
Labor	50	45	-5	25	32	7	18	18	0	6	4	-2	0	0	0	2	2	0
Motor Vehicles[a]	43	33	-10	27	41	14	13	8	-3	12	11	-1	3	4	1	3	4	1
General Services[a]	43	40	-3	28	32	4	14	11	-5	9	11	2	6	4	-2	0	2	-1
Health	41	37	6	26	23	-3	16	16	0	11	18	7	5	4	-1	1	3	2
Public Service	42	39	-3	34	35	1	10	5	-5	10	15	5	5	3	0	2	2	0
Overall total	46	40	-6	27	33	5	14	13	-1	9	10	1	3	2	-1	1	2	1
Avg test			-8			10			-1			-2			-2			0
Avg control			-3			3.5			-2			3			0			1

[a]Test group.

over the same period of time. Thus the coordinator project was able to effect an increase of 6.5 percentage points because of its activities. Approximately 195,000 gal of gasoline was conserved by new carpoolers in all six agencies--an average of 283 gal of gasoline per year per carpooler. Of this, 101,000 gal was attributable to the carpool coordinator program.

Comparison of internal records (data reports by the coordinators themselves on new carpoolers) isolates the direct effect of the coordinator program from the indirect (halo) effect of the program and the effect of background changes. Halo program effect refers to the inducement of additional positive behavior by individuals not actually registered in the program but merely encouraged by it. The data in the table below indicate that, of the 530 new carpoolers in the test agencies, 150 came directly from the program, 233 from the halo effects of the program, and 147 from the energy crisis (external effect):

	Agencies	
Item	Test	Control
Agency population	4,207	4,365
Change in carpool (%)	+10	+3.5
Effects		
External (nonprogram)	147	162
Internal (program)		
Direct (from program records)	150	-
Indirect (halo)	233	-
Total	383	-
Total	530	162

[Note that 147 (external effect) was derived by multiplying 4,207 x 0.035; 150 (direct internal effect) includes dropouts; and 233 (indirect internal effect) is the total of program and energy crisis carpoolers.] The 233 employees were encouraged to carpool by the spirit of the program in these agencies, but would not have done so solely because of the energy crisis. [Note that the estimate of the 147 carpoolers from the energy crisis is determined by applying the percentage change in the control group (+3.5) to the test population (4,207).]

Thus, although the energy crisis itself clearly accounted for an increase in carpooling, the carpool coordinator program in the test agencies was capable of taking advantage of that event by directly assisting some employees and indirectly encouraging others to participate on their own. Furthermore, the use of the test and control design with internal and external observations allowed the quantification of these effects. Without this design (relying on program data records only), the program would have registered only 150 new carpoolers and missed entirely the 233 indirect carpoolers. Thus the strong design, instead of hurting the program, actually strengthened it.

These results are an example of the Δ test +, Δ control + results in Figure 1. The positive interpretation stems from the fact that ΔT is much greater than ΔC and both are positive.

Case 2: Neighborhood-Based Program During a Stable Period

The neighborhood-based demonstration began in January 1981 and ended in December 1981. As shown in Figure 2, the period was generally one of economic and travel stability, and therefore a surge in carpooling would not be expected. The goal of this study was to determine the effectiveness of home-end carpooling success. The study featured the use of ridesharing coordinators operating within neighborhood sites (community based), a quasi-experimental design in which the effects of coordinator services are compared with the behavior of control sites, the use of before-and-after surveys, and careful periodic monitoring. Comparison of results was made on the basis of changes in travel behavior, carpool formation and retention, vehicle miles of travel (VMT), energy savings, and the effect of marketing materials.

The evaluation design may be represented as follows (note that the internal effect includes $T_1, T_2, \ldots$):

	Overall Effect				
	Before Survey				After Survey
Site	(T_B)	T_1	T_2	...	(T_A)
Cohoes	O_B				O_A
Clifton Park	O_B	M_1 O_1	M_2	...	O_A
Glenville	O_B				O_A
Albany-South Side	O_B				O_A
Control area	O_B				O_A

Four sites from the Albany standard metropolitan statistical area (SMSA) were selected and provided with community-based coordinators for a 1-year period. The control site was the remainder of the Albany SMSA. Sites varied in density and accessibility to the urban cores and in location of the coordinators' offices:

Type of	Location of Office	
Setting	Town Hall	Home
Suburb	Clifton Park	Glenville
City	Cohoes	Albany-South Side

The general hypothesis tested (10) is that if the marketing and coordinators' services are effective, the change in behavior in the test sites should be different (significantly) from the change in behavior in the control site. Similarly, tests between specific sites, or groups of sites, can be arranged to evaluate the effectiveness of suburban versus city programs, town hall versus home offices, and each site's program relative to the others.

The before-and-after surveys consisted of random sample telephone surveys of residents in each site and in the control area. Each household contacted was given a brief questionnaire on travel patterns, work and nonwork ridesharing, reasons for ridesharing (before survey only), and program awareness (after survey only). Persons selected in the before survey were recontacted in the after survey, thus constituting a panel of observations. The use of the panel approach allowed a determination of changes in travel over time within each site. Tests for the significance of changes in travel over time were conducted by using paired observations for nonwork travel and modal-shift behavior of work travel. Tests for significance between sites were conducted by using standard tests for the difference in mean change.

Direct effects were measured through a review of records kept by each carpool coordinator. Basic information included

1. Number of applications attracted;
2. Number of ridesharers attracted;
3. Number of new carpoolers per coordinator hour of effort; and
4. VMT reductions associated with the number of new carpoolers, the number of coordinator hours, and the number of applications received.

During the program weekly records were kept of the hours and type of coordinator effort, number of applications received, number of applications attributed to each marketing strategy, and number of carpools formed. The results of this effort were summarized on a quarterly basis. Each marketing strategy was evaluated by the number of applications generated, the cost, and the resulting VMT savings.

Analysis of the external data is given in Table 5. Work-based carpooling activity declined in the test sites and increased in the control sites, although the changes were small. Further analysis suggested that these results were caused by inordinately large drops in the number of reported workers per household, which suggested reporting problems in the panel data that were possibly caused by carpool dropouts in the not-working group rather than the drive-alone group.

Nonwork carpooling--already high in these sites--was found not to have changed in the 1-year period. The data in Table 6 indicate that, although some increases in nonwork travel did occur, these appear to be related to shopping travel increases in Clifton Park and Glenville and are probably the result of widescale supermarket coupon wars in those towns (see Figure 2). Carpool nonwork travel, however, showed no changes (11), which implies that the coordinator services had no significant effect on overall community travel.

Surprisingly, the internal records indicated that the community-based service was equally as effective as the employer-based service described above. In the four sites, 176 new carpoolers were formed from 396 applications for a savings of 18,000 VMT/week. When reduced to the same time period as the employer demonstration, the results are almost identical (Table 7). Taking into account that these results did not have the benefit of an energy crisis, the study concluded that the neighborhood approach has considerable potential.

However, in contrast to the employer demonstration, the overall effect of the program on community travel was small. The data in Table 8 indicate that the program saved less than about 0.3 percent of community VMT, and this saving was concentrated in the suburban sites of Clifton Park and Glenville.

Table 5. Changes in work travel (1981-1982) for neighborhood ridesharing demonstration project.

	Drive Alone		Rideshare		Transit		Other[a]		Total		
Area	Percent	Avg Distance (miles)	Percent	Avg Distance (miles)	Percent	Avg Distance (miles)	Percent	Avg Distance (miles)	Percent	Avg NST	Private Vehicle Avg Distance (miles)
Cohoes											
1981	62	8.0	26	8.5	4	5.5	7.9	1.6	100	7.5	8.1
1980	56	9.0	32	9.2	4	5.5	8.0	1.5	100	8.3	9.0
Δ	+6	-1.0	-6	-0.7	0	0	0.1	0.1		-0.8	-0.9
Clifton Park											
1981	70	14.9	29	14.6	0.5	22.0	0.5	1.0	100	14.8	14.8
1980	69	14.3	30	14.4	0.5	22.0	0.5	1.0	100	14.3	14.3
Δ	+1	+0.6	-1	+0.2	0	0	0	0		+0.5	+0.5
Glenville											
1981	76	9.2	17	8.9	3	11.5	4	3.0	100	8.9	9.2
1980	76	10.1	16	8.8	4	8.4	4	3.0	100	9.5	9.9
Δ	0	-0.9	+1	+0.1	-1	3.1	0	0		-0.6	-0.7
Albany-South Side											
1981	57	6.6	15	6.9	16	4.1	12	1.8	100	5.7	6.7
1980	58	8.2	15	7.5	17	4.4	10	1.8	100	7.2	8.1
Δ	-1	-1.6	0	+0.6	-1	-0.3	+2	0		-1.5	-1.4
Total test areas											
1981	67	10.3	22	10.8	5	5.3	6	1.6	100	9.7	10.4
1980	65	10.8	24	10.7	6	5.6	5	1.6	100	10.1	10.8
Δ	+2[b]	-0.5	-2[b]	+0.1	-1	-0.3	1	0		-0.4	-0.4
Capital District control											
1981	68	10.7	18	9.0	8	4.6	6	3.0	100	9.5	10.3
1980	71	10.7	17	8.8	6	4.4	6	3.0	100	9.5	10.3
Δ	-3[b]	0	+1[b]	+0.2	+2	+0.2	0	0		0	0

[a] Includes walk and bicycle. [b] Statistically significant.

Table 6. Summary of tests for significance of differences in changes in nonwork travel between sites (neighborhood ridesharing demonstration).

	Summary of Tests (t-statistics)				
Significance of Δ VMT for	Shop, Grocery	Shop, Nonwork	School	Church and Civic	Visit and Social
Test as a whole	2.1	2.3			4.5
Cohoes				2.1	2.5
Clifton Park	3.6	2.8			4.9
Glenville	2.1	2.1			3.8
Albany-South Side					2.9
City versus control					3.2
Suburban versus control	3.3	2.8			5.0
Town hall versus control	2.3				4.3
Home and office versus control		2.3			3.9
Suburb versus city	3.2				2.0
Town hall versus home and office					

Notes: There were no significant changes in the t-statistics for Δ carpool VMT or Δ carpool person miles of travel (PMT). x = t ≥ 2.0; n = 821 (reduced panel: used only female respondents in second year of survey).

Table 7. Comparative summary of direct effects for employer versus neighborhood ridesharing demonstrations.

Phase	Item	Demonstration	
		Neighborhood	Employer[a]
Target	When	January-November 1982	January-October 1979
	Target population	101,723	4,207
Effort (input).	No. of coordinators	4	4
	No. of hours	3,755	2,230
Results (output)	Applications received	346	624
	New carpoolers attracted	154	150
	Total VMT saved per week	16,447	16,335
	Total gallons of gasoline saved per week	1,097	1,126

[a]Initial period.

Table 8. Direct program effects as a percentage of community VMT.

Item	Cohoes	Clifton Park	Glenville	Albany-South Side
Work				
Community VMT per week[a]	639,000	1,067,000	974,000	974,000
VMT saved per week	1,282	2,454	1,739	401
Percentage saved	0.20	0.29	0.18	0.04
School and recreation				
Community VMT per week[a]	16,000	58,000	44,000	34,000
VMT saved per week	536	7,876	2,956	1,533
Percentage saved	3.35	13.58	6.72	4.51
Total				
Community VMT per week[a]	1,085,000	1,501,000	1,662,000	1,516,000
VMT saved per week	1,818	10,330	4,695	1,954
Percentage saved	0.19	0.67	0.28	0.13

[a]Estimated from external survey data.

Therefore the use of a strong design in this case had the effect of placing the generally positive results into perspective to remind the researcher that the apparent positive news from the internal records should not be presumed to have generated large benefits to the community as a whole.

The test results in this study fall into two interpretation groups of Figure 1. The work travel results are Δtest −, Δcontrol +, which suggests a negative backlash on the program or (in this case) methodological problems with the data. On the other hand, the nonwork results are best interpreted as Δtest 0, Δcontrol 0, which suggests an ineffectual program in a large stable background.

CONCLUSIONS

It has been demonstrated that the use of statistical designs can be helpful, and not necessarily negative, in evaluating carpooling programs. Among the benefits of such designs are the following:

1. Isolation of background changes: In the face of major background changes, the test and control design prevents inappropriate attribution of results entirely to the program. It also permits quantification of the direct and indirect effects of the program as part of its positive impact. This would not be possible without before-and-after data.
2. Perspective setting: In both studies the design permits the results to be placed into a broader perspective. In cases where the community is large and the program small, this naturally means a dampening of apparent positive results. Although ridesharing agencies might therefore be reluctant to include such findings, it is believed that the people (as taxpayers) are better served by them.
3. Insurance: No one knows, of course, when rapid changes in background might occur. Given the relatively long lead time necessary to plan and implement transportation services, the prudent researcher should include the possibility that such an analytical disaster might occur during the study.

ACKNOWLEDGMENT

This paper was prepared from the findings of the neighborhood ridesharing demonstration study funded by FHWA. However, the views expressed in this paper are ours and should not necessarily be attributed to FHWA or NYSDOT. Insightful comments on an earlier draft were provided by Mary Lynn Tischer. We acknowledge the work of Linda Unangst and Donna Brown for manuscript preparation. We, of course, remain solely responsible for errors of fact or omission.

REFERENCES

1. F.A. Wagner. Evaluation of Carpool Demonstration Projects. U.S. Department of Transportation, Oct. 1979.
2. D.T. Campbell and J.C. Stanley. Experimental and Quasi-Experimental Designs for Research. Rand McNally and Company, Chicago, 1963.
3. F.C. Dunbar. Use of Before-and-After Data to Improve Forecasting Methods. TRB, Transportation Research Record 723, 1979, pp. 39-45.
4. J.J. Louviere and others. Application of Psychological Measurement and Modeling to Behavioral Travel Demand Analysis. Center for Behavioral Studies, Univ. of Wyoming, Laramie, March 1977.
5. M.L. Tischer and R.V. Phillips. The Relationship Between Transportation Attitudes and Behavior Over Time. Transportation, Vol. 8, No. 1, 1979.
6. M.L. Tischer. Attitude-Behavior Models: An Evaluation of the Effectiveness of Transportation System Management to Change Commuter Mode Choice and Meet Public Policy Goals. Univ. of Maryland, College Park, Ph.D. dissertation, 1981.
7. L. McClelland and others. Evaluating Employer Programs Encouraging the Use of Alternate Travel Modes. Transportation, Vol. 10, No. 3, 1981, pp. 247-256.

8. J.M. Brunso and D.T. Hartgen. Carpool Coordinator Demonstration Study: Overview and Analysis of Before Data. New York State Department of Transportation, Albany, Prelim. Res. Rept. 150, March 1979.
9. J.M. Brunso and D.T. Hartgen. Can Employer-Based Carpool Coordinators Increase Ridesharing? TRB, Transportation Research Record 823, 1981, pp. 45-50.
10. J.M. Brunso, W.R. Ugolik, and D.T. Hartgen. Research Design for the Neighborhood Ridesharing Demonstration. New York State Department of Transportation, Albany, Prelim. Res. Rept. 189, Aug. 1980.
11. D.T. Hartgen and J.M. Brunso. Neighborhood Ridesharing Demonstration Study: Final Report. New York State Department of Transportation, Albany, March 1982.

Publication of this paper sponsored by Committee on Ridesharing.

Impact of Flexitime Work Schedules on an Employer-Based Ridesharing Program

FREDERICK J. WEGMANN AND STANLEY R. STOKEY

The impact on commuting behavior of employees when flexitime is superimposed on a large employer-based ridesharing program is discussed. The case study uses the Tennessee Valley Authority (TVA) program in downtown Knoxville, Tennessee. Based on the first 6 months of experience with TVA's Knoxville flexitime program, it is shown that giving employees greater choice in working hours can serve to upset an established ridesharing program. It must be noted that the TVA ridesharing program is unique in that it provides a high level of consumer-oriented services. Buses operate equivalent to a subscription program and, along with vans, arrive just before the work day starts and leave immediately at the end of the work day. The element of choice then adds complexity to the operations. With shifting demands for different starting and leaving times, it becomes difficult to balance the services with the demand. Also, it is difficult for 35 to 40 people who use the same vehicle to reach a mutually agreed on schedule. Van operations are easier to adapt to flexitime because the decisions involve a smaller number of individuals and decisions can be made at the decentralized level of the van. However, when individuals are accustomed to receiving a high level of commuter service, and events take place to spread that demand over a longer time period, readjustments in travel behavior and accompanying services will be required. These adjustments will require the provision of additional commuter services. As TVA's experience indicates, without service adjustments, people will make use of the flexitime opportunities by carpooling or by driving alone. Both ridesharing and flexitime are important concepts for energy conservation. However, when flexitime is added to a large customized ridesharing program, the net energy savings will not equal the sum of both energy conservation actions taken singularly.

Two critical issues that confront transportation planners are increased concern over the cost and availability of energy and the ability of the government to undertake large-scale capital investment programs to increase the capacity of transportation facilities. Increasingly, it is becoming apparent that many transportation problems are related to the peaking of trips. Work trips tend to cluster during about 4 hr of the day, which necessitates the sizing of transportation facilities to accommodate the travel demand concentrated in these hours. Peaking problems create travel delays and cause inconvenience to users of the transportation systems; these delays are also costly in terms of excess pollution and energy use.

Rather than building excess transportation capacity that is only used for a few hours per week, a philosophy of peak-period demand management is evolving as a transportation system management (TSM) strategy. Attempts are being made to reduce peak-hour demands through such concepts as staggered work hours, flexitime, or the 4-day work week.

Flexitime, in particular, is receiving increased attention as a peak-period demand management technique. Flexitime differs from staggered work hours in that it does not formally assign work arrival and departure times to groups of employees. For example, in a firm in which all employees worked from 8:00 a.m. to 4:30 p.m., the work force could be divided into three groups by initiating staggered work hours. The first group might work from 7:30 a.m. to 4:00 p.m., the second from 8:00 a.m. to 4:30 p.m., and the last from 8:30 a.m. to 5:00 p.m. Some employees will benefit from improved transportation because of less congestion, but each employee's arrival and departure time remains fixed.

Flexitime is different. One popular variation is to designate certain hours as flexible or core hours within the span of a work week. An employee must work a set number of hours, but there is more latitude in choosing working hours within an established range. Typically, all employees must be available for a core time (e.g., 9:00 a.m. to 3:00 p.m.); within a certain number of flexible hours employees may choose their own arrival and departure times. In some programs, lunch breaks may also be defined as flexible time (1).

Flexitime is a relatively new idea that is receiving increased attention in the United States. Historically, flexitime is generally attributed to a program initiated in 1967 by the Messerschmidt-Boelkow-Blohm aerospace firm in West Germany. Since that date, flexitime has spread rapidly through Europe; but, until recently, it has received only limited attention in the United States (2). It is estimated that more than 3,000 West German companies have extended the flexitime concept to more than 50 percent of the labor force (3). Similar acceptance rates have been achieved in other European countries. Projections made from a 1977 survey conducted by the American Management Association estimated the use of flexitime in the United States as follows (4):

1. Almost 13 percent of all nongovernment organizations with 50 or more employees use flexitime,
2. More than 5 percent of all employees are on flexitime, and
3. Between 2.5 and 3.5 million employees are on flexitime, not counting self-employed persons and many professionals, managers, and sales people who