PROCEEDINGS —
Seventeenth Annual Meeting

Theme:

"Beyond The Bicentennial:
The Transportation Challenge"

October 28-29-30, 1976
Sheraton-Boston Hotel
Boston, Massachusetts

Volume XVII • Number 1 1976

T
R
F

TRANSPORTATION RESEARCH FORUM

I. INTRODUCTION

I.1 Rationale

REVENUE/RIDERSHIP DATA reported at regular intervals is generally the most reliable, readily available source of data to an operating agency. Analysis of this data can give a decision-maker valuable insights into day-to-day operating policy as well as provide a basis for long-term planning. One of the reasons this relatively inexpensive source of data has not been used to its full potential is that there is rarely a time series where all factors are constant and the effect of a single intervention or policy can be estimated with confidence. The standard regression-based methods for analyzing time series data and separating out the effects of various interventions have had the effect of putting a complex, expensive, computer-based black box between the decision-maker and the raw data, preventing an intuitive grasp of the real situation. The result of these complex analyses (often done by a consultant) has been either rejection of the analysis, if at variance with the predisposition of the decision-maker, or blind acceptance without checks of verification. Neither is a good response. This paper discusses and demonstrates an alternative to this data analysis/decision-making process.[1]

I.2—Scope and Objectives

A series of simple, easily understood exploratory techniques in the spirit of Tukey (1975) were used to study a time series of monthly Chicago North Western (CNW) commuter rail ridership data spanning a period of time (1963-1975) which included rapid suburban growth, introduction of a competing rapid transit line, repeated fare hikes and the energy crisis. These methods, which allow the investigator to use his experience/intuition to interact with and extract valuable information from the data, have the potential to provide useful and reliable data analyses at a fraction of the time and cost of computer-based methods since they require only graph paper, a hand calculator and a few hours.

Four questions, of major importance to transportation planning in the Chicago metropolitan area and to the transportation field at large were investigated.

(1) Assessment of the impact of the extension of the Chicago Transit Authority (CTA) Kennedy rapid transit line on the competing CNW commuter rail line.

(2) An evaluation of the effect of commuter rail fare increases in the presence of a competing mass transit mode.

(3) An assessment of the time interval of a policy intervention adjustment and geographic extent of impact.

(4) An assessment of the impact of the energy crisis on commuter rail ridership.

I.3 Summary of Results

The initial impact of the CTA extension lasted six months and resulted in a loss of 60,000 riders, about .4% of the total revenue of the line but between 25% and 45% of the riders at the stations involved. This impact was related to the distance from the competing rapid transit station (see Figure 3.8). We found elasticities to be quite high (between .9 and 1.5) in the presence of competition from rapid transit. However, when the automobile was the only mode alternative, losses due to fare increases were small or non-existent. The impact/adjustment period for both the Kennedy and fare intervention was a declining "ramp-like" function of four to six months. Finally, it was found that the energy crisis had a substantial impact on ridership. These results and their policy and planning significance[2] are discussed in Section IV. Section II of this paper provides a background and describes in greater detail the data used for this study, and Section III describes the exploratory methodology, analysis and results.

II. BACKGROUND AND DATA

II.1 Description of Routes

The northwest route of the CNW, which serves a 63-mile, high-intensity corridor, is the most heavily used of the several commuter-rail lines serving the Chicago CBD (see Figure 2.1). During the study period, the middle and outer belt suburbs along this corridor were experiencing growth rates of 6% and 7%, respectively, while the city and inner suburbs maintained a constant population—both in numbers and composition.

Anticipating potential for suburban growth in the late 1950's, CNW made major investments in new equipment with consequent improved levels of service. These improvements were substantially completed by 1963 although the final delivery of new cars was made in 1969. During our study period, there have been no significant changes in service and for our analysis we assume it to be constant.

II.2 Interventions

In 1970, the Chicago Transit Authority (CTA) opened an extension of the

A Time-Series Analysis of the Impact of Mass Competition on Commuter Rail Ridership†

*by Chris Johnson**, *Benjamin Burrows**, *Cherilyn Heramb**, *Robert C. Kunze**, and *Ashish K. Sen**

CNW NORTHWEST LINE INNER STATIONS AND ADJACENT KENNEDY RAPID TRANSIT STATIONS

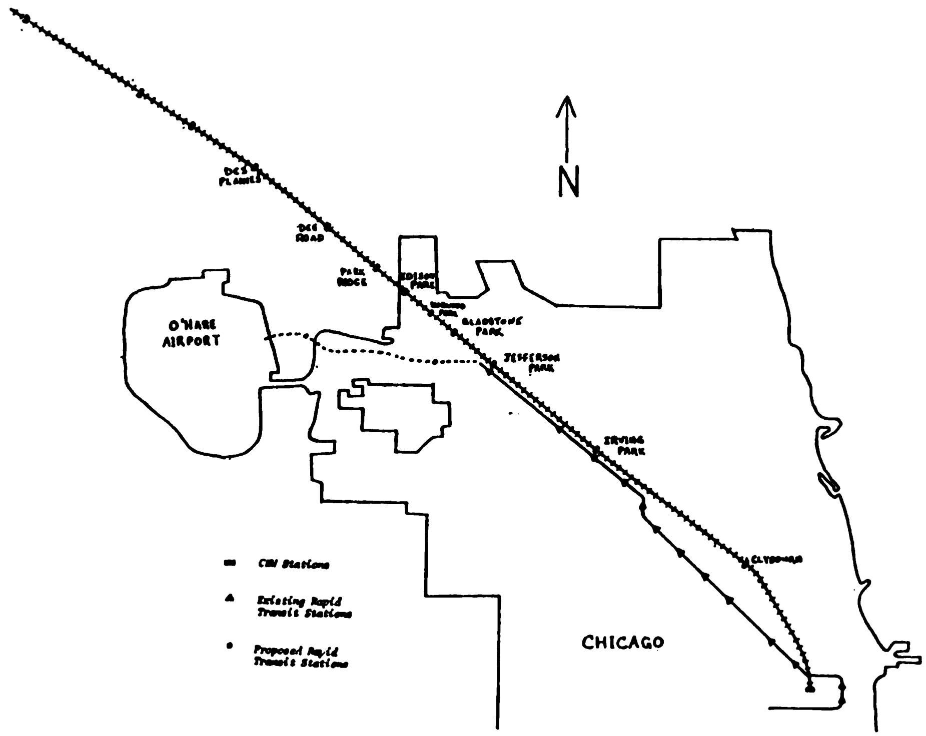


FIGURE 2.1

Kennedy rapid transit line, a modern, comfortable facility extending to within one-half block of the CNW Jefferson Park Station (see map in Figure 2.1) and terminating in the Loop. This facility has recently been augmented with a number of feeder bus routes. There has been discussion of further extension of the Kennedy line which CNW has opposed on the grounds of adverse competition from a publicly subsidized system.

Since 1968, CNW has imposed nine fare increases averaging 7%. Fares on the CNW vary with distance. In 1970, a one-way Loop bound fare varied from 80¢ at Jefferson Park to $1.80 at Crystal Lake; those fares now range between $1.00 and $2.30. Recently, CNW has requested a two step fare increase amounting to 40%.

Other interventions during this time period include major construction on the Kennedy Expressway in 1966 and

**University of Illinois at Chicago Circle, College of Urban Sciences.*

† The writing of this paper was partially supported by an Urban Mass Transportation Administration Research and Training grant. Any views expressed are solely those of the authors and do not represent UMTA's position.

1971, and long-term lane closures in 1968. Further, during the winter of 1973-74 there was a shortage of gasoline and a significant increase in gas prices.

II.3 Description and Preparation of Data

Monthly ridership totals for each of 22 stations along the northwest route of the Chicago North Western commuter railroad were obtained. The figures spanned a thirteen year period between 1963 and 1975. Simple time series plots of the data for the 22 stations revealed three natural groupings of the stations —a city/inner suburb group, a middle suburban group and an outer suburban group. In this paper, only the work on the city/inner suburban group is described. These stations are the ones most affected by the CTA Kennedy line. Analyses of the other stations is given in Burrows, et al. (1976).

III. APPLICATION OF EXPLORATORY METHODS TO DATA

The purpose of this approach was to understand the response of CNW ridership to the interventions mentioned above. Our attempt was not to build a model or test a hypothesis. Rather, we wished to explore and suggest. The primary methods used in the study are those of smoothing, graphing, subtracting out patterns, plotting residuals and doing further subtraction as necessary. All but the smoothing techniques will be obvious from the discussion of the analysis in Section III.1. The details of smoothing are presented in Appendix A. Further details and a rationale are presented in Beaton and Tukey (1974).

Since we had complete data for all 22 stations, we analyzed a single station and then repeated our procedures several times. We will use the same procedures in reporting. Section III.1 presents a detailed analysis of a single station and Section III.2 summarizes the results from analyses of all the city/inner suburban stations.

III.1 Analysis of Edison Park

Figure 3.1 is a plot of the raw time series of Edison Park, a station roughly 3⅓ miles from Jefferson Park CTA terminal which carried 20% of the inner belt riders. Beyond noticing the fundamental shape of the growth and decline, it would be difficult to measure the effects of interventions known to have occurred during this time. Figure 3.2 is a plot of the same data which has been smoothed using the techniques presented in Appendix A.

A number of observations can be made from Figure 3.2 that were not evident in the raw data. Between 1963 and 1970 there appears to have been two growth rates. Fitting a slope through the medians of each year, it appears that ridership grew at a rate of 13% prior to 1965 and about 28% after. We further observe that during this period (1963-1970) there was a strong seasonal fluctuation of at least ±1,000 riders. An exception is 1966. The year seems to have been a transitional point in growth. Yet a conversation with CNW manage-

MONTHLY PASSENGERS AT EDISON PARK STATION

Raw Data

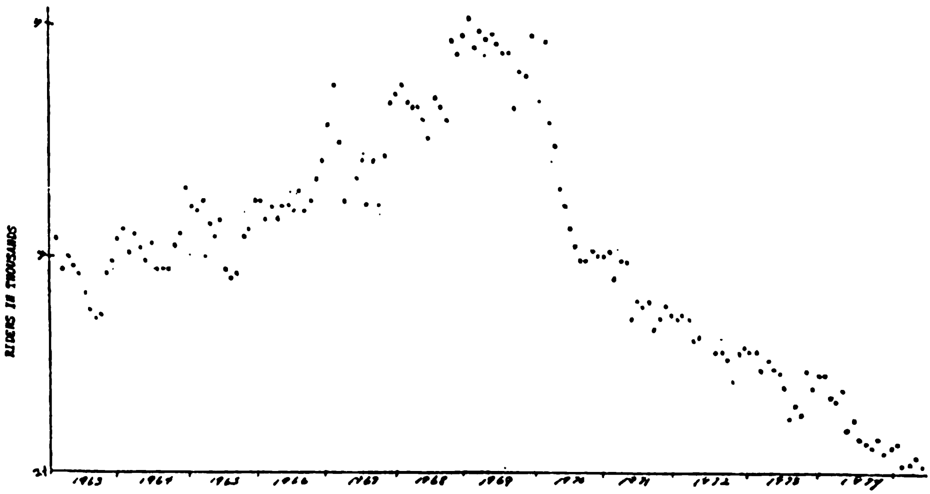


FIGURE 3.1

MONTHLY PASSENGERS AT EDISON PARK STATION
Smooth

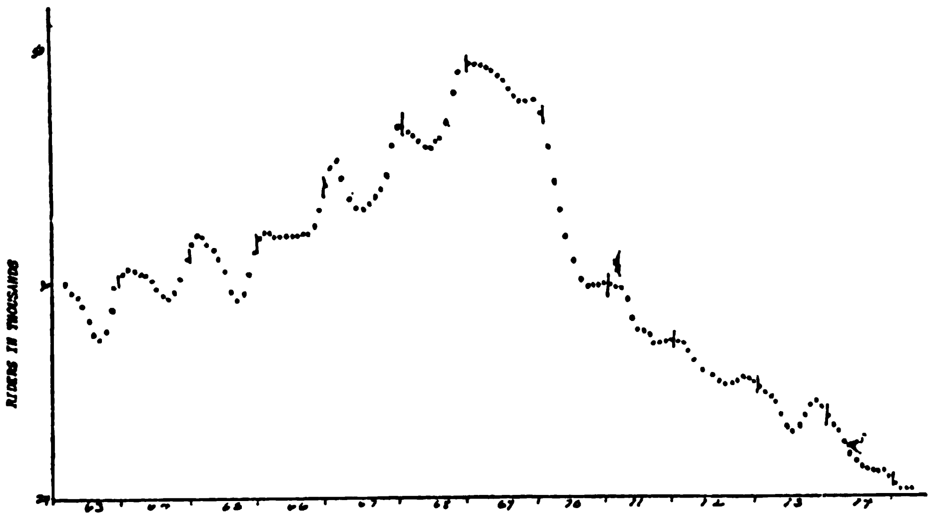


FIGURE 3.2

ment revealed no internal (i.e., service change) cause. We note that during the summer of 1966, the Kennedy Expressway (which lies in the same corridor) underwent major repairs. This may have altered the normal seasonal fluctuation but would not alter the growth rate over a long-term.

As noted previously, population for this area remained constant in number and composition over the 13-year period, indicating a possible increase in market share between 1963 and 1970. Several possible explanations were investigated, including growth in CBD employment. The most promising was traffic congestion. A plot of traffic counts on the Kennedy Expressway versus ridership levels (Figure 3.3) shows a good straight-line fit, suggesting that the annual rates of change between the two series are roughly proportional. Since the plot of ridership against time alone is not particularly straight, this gives a fairly strong hint that congestion on the expressway caused the increase in CNW ridership.

While there are many problems associated with relating increasing functions, it should be pointed out that growth on the inner stations behaved similarly to growth on the mid and outer stations. A quick check of population figures showed the ridership growth of these latter stations to roughly correspond to population growth in their respective service areas. It is reasonable to believe that traffic congestion is roughly proportional to population, and that outer suburban population growth would be reflected in the inner belt as increased congestion.

Impact of the Kennedy

The impact of the Kennedy is obvious from Figure 3.2. The plot shows two distinct types of ridership declines beyond that point. For the first six months of 1970, CNW lost about 12,000 passengers. During this period the impact seemed to take the form of an exponential decay. The second type of decline in ridership after mid 1971 has a "cascade" appearance.

Though CNW has attributed all of its continued loss in ridership of these stations to competition from CTA, the smoothed graph shows that the two declines are fundamentally different in structure, suggesting other factors may be involved. In fact, the steps of the "cascade" in Figure 3.2 correspond roughly to the annual fare increases that occurred. It would appear that each year ridership began an annual seasonal decline, and due to a fare intervention, failed to recover in the fall. It thus seems reasonable to measure only the exponential drop as directly due to CTA competition. (However, since without CTA competition the "cascades" would not be nearly as steep, these declines are also related to CTA competition indirectly).

Fare Increases

To ascertain the effect of fare in-

PLOT OF MEDIAN ANNUAL RIDERS AT EDISON PARK
VS. ANNUAL TRAFFIC COUNTS ON KENNEDY EXPRESSWAY

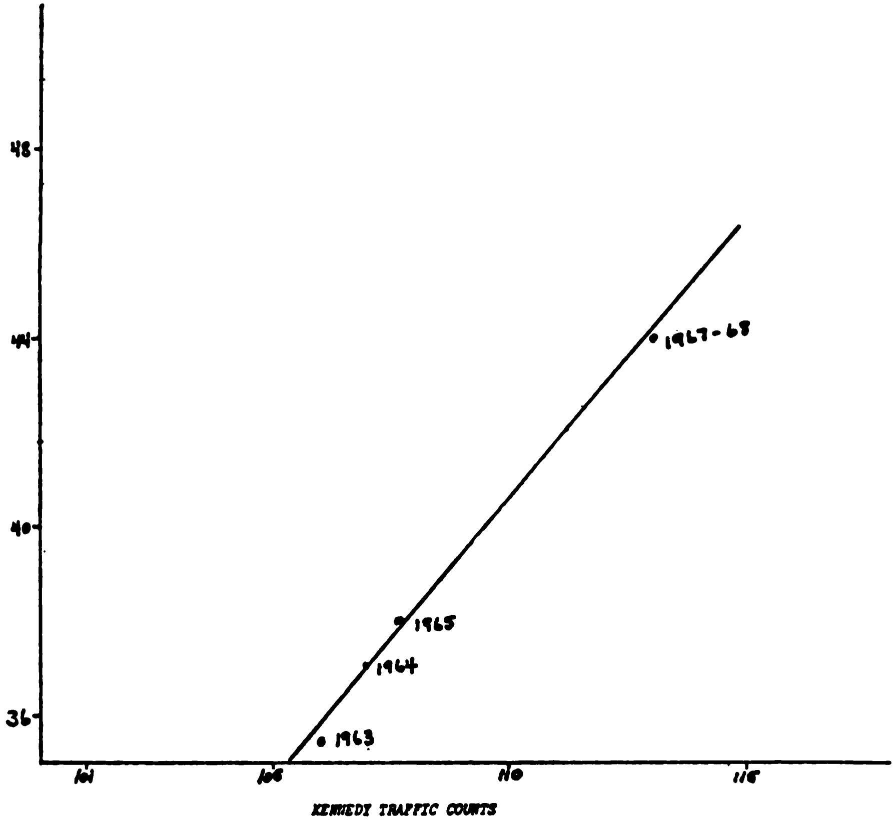


FIGURE 3.3

creases, it was necessary to separate the decreases due to seasonal fluctuation from a true drop in demand. An "average" cycle was estimated by taking medians of the 1963, 1964, 1965, 1967 and 1968 cycles. This cycle was subtracted from the plot shown in Figure 3.2. Figure 3.4 is a plot of those residuals and it clearly shows an over-correction in the post-1970 years since the peaks of the residuals correspond to the seasonal valleys.

After some experimentation, a modification of the 1964 cycle was chosen as representative. It is not as extreme as 1963 or 1965, and it corresponds to a greater degree to the cycles between 1967 and 1970. The modification consisted of lowering December's "recovery" to a level similar to that of the previous January.

The residuals from this subtraction, plotted in Figure 3.5, clearly indicate a ramp-like step function which corresponds to fare increases. Rough elasticities were computed. These are 1.4, 1.5 and 1.1, respectively, for the fare increases of 1971, 1972 and 1973-74. The elasticity for the three fare increases between October, 1973 and May, 1974, were computed as one since we conjectured that the energy crisis that winter may have delayed the full impact of the first two increases. Although there were fare increases in 1968 and 1969, there clearly was little or no drop in demand, nor from surface inspection is there any relative change in rate of growth. Only after the introduction of a competeing mass transit mode serving the same general corridor did fare increase show an appreciable effect.

Preliminary investigation was also made into possible fare increase effects in the mid and outer stations. While there is some evidence that fare increases may have some negative impact, the effect is very small when com-

PLOT OF RESIDUALS AFTER SUBTRACTING SEASONAL CYCLE
Edison Park

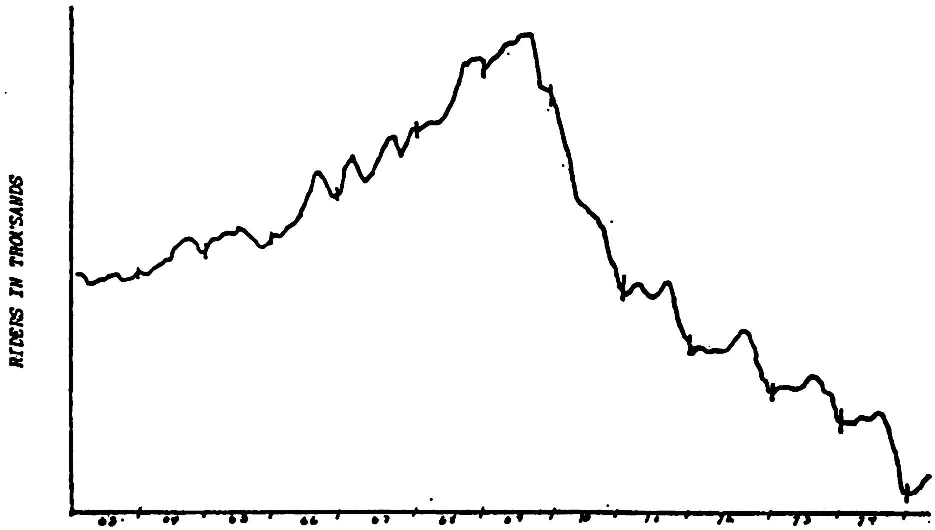


FIGURE 3.4

PLOT OF RESIDUALS AFTER SUBTRACTING SEASONAL CYCLE
Edison Park

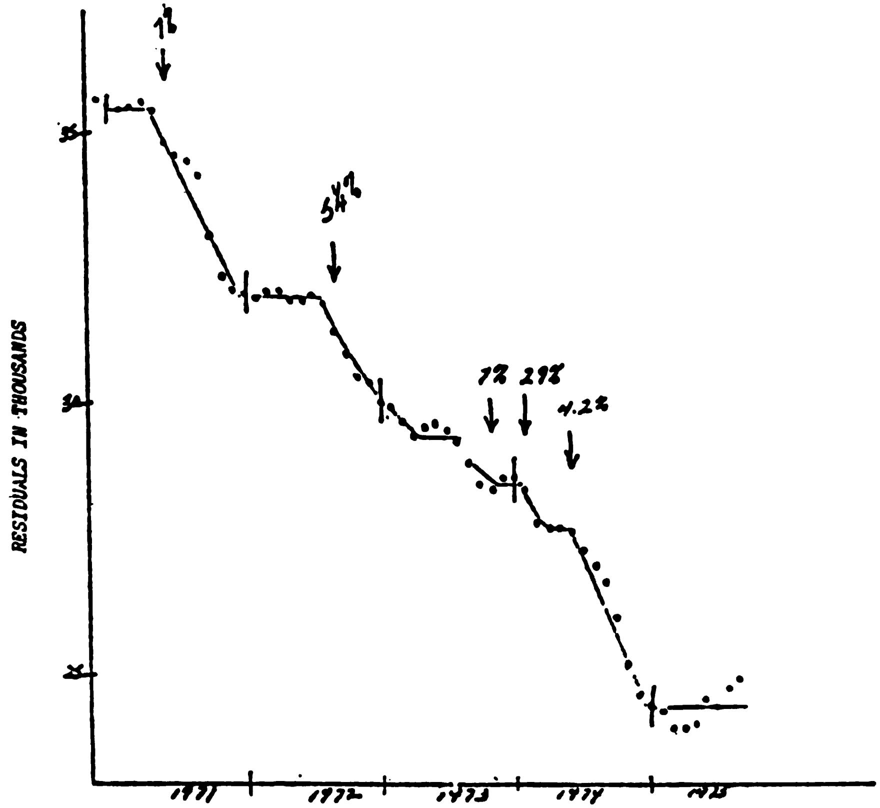
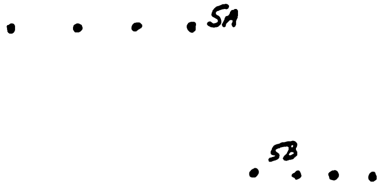


FIGURE 3.5

EXAMPLE OF OVER SMOOTHING



RAW DATA POINTS - STEP FUNCTION

FIGURE 3.6

SMOOTHED DATA POINTS

FIGURE 3.7

pared to that of the inner stations (see Burrows, et al., 1975).

Intervention Period

We notice from Figures 3.5 and 3.2 that the adjust period for both the Kennedy and fare interventions is about six months. These observations should be viewed with some caution since if we planned a true step function illustrated in Figure 3.6, the result would be that illustrated in Figure 3.7. Two points would be altered into a ramp-like pattern; it would be more reasonable, then, to suggest a four to six month impact adjustment period.

III.2 Summary of Results from Exploratory Analysis

Table 3.1 summarizes the overall impact of the Kennedy Rapid Transit extension. Each of the impacted stations showed a nearly identical adaptation period of about six to seven months before ridership stabilized. The implication is that there is about a six month "information" lead time required for commuters to assess time and/or cost savings of a new mode, readjust schedules, learn new feeder routes, etc., and make a reasonably permanent mode split decision. Des Plaines and Jefferson Park are major exceptions. Des Plaines' loss was so small that the loss had to be computed from a change in medians, a clear drop and stabilization could not be seen. Further, over the long run, that loss seemed to be recovered. It is to be expected that Jefferson Park, most heavily impacted, would show a longer term initial impact period.

The impact of the Kennedy on the ridership level at each train station varied with the distance of the station to the nearest CTA station. Figure 3.8 plots the percentage loss of each impacted station against that distance. The curve suggests severe competition or impact when the modes are within four to five miles. The impact of competition is less severe or non-existent beyond this point. It should be pointed out that these impacted areas have been calibrated for modes competing in a high intensity corridor. Caution should be taken in assuming that these impact delimitations will hold true in more diffuse origin and destination patterns.

Fare elasticities were computed for the two major stations of the inner belt and the aggregate of the inner belt. The results are summarized below:

	1971	1972	1973-1974
Edison Park	1.4	1.5	1.1
Gladstone Park	1.4	1.5	1.1
Aggregate	1.4	1.5	.9

Elasticities were not computed for the remaining inner stations, since when the smoothed plots for the stations were carefully examined, it was noticed that the impact of the Kennedy was so severe that after the fare increase of

SUMMARY OF INITIAL KENNEDY IMPACT

Station	Duration of Impact	Number of Passengers Lost	Percentage Loss
Irving Park	6 months	7,400	42%
Jefferson Park	10 months	16,200	43%
Gladstone Park	6 months	5,100	43%
Norwood Park	6 months	10,500	26%
Edison Park	6 months	12,200	25%
Park Ridge	7 months	7,100	8%
Des Plaines	unknown	1,500*	2%

*Loss appeared to be temporary

TABLE 3.1

PLOT OF THE PERCENTAGE PASSENGER LOSS AT INDIVIDUAL TRAIN STATIONS vs. DISTANCE TO NEAREST CTA STATION

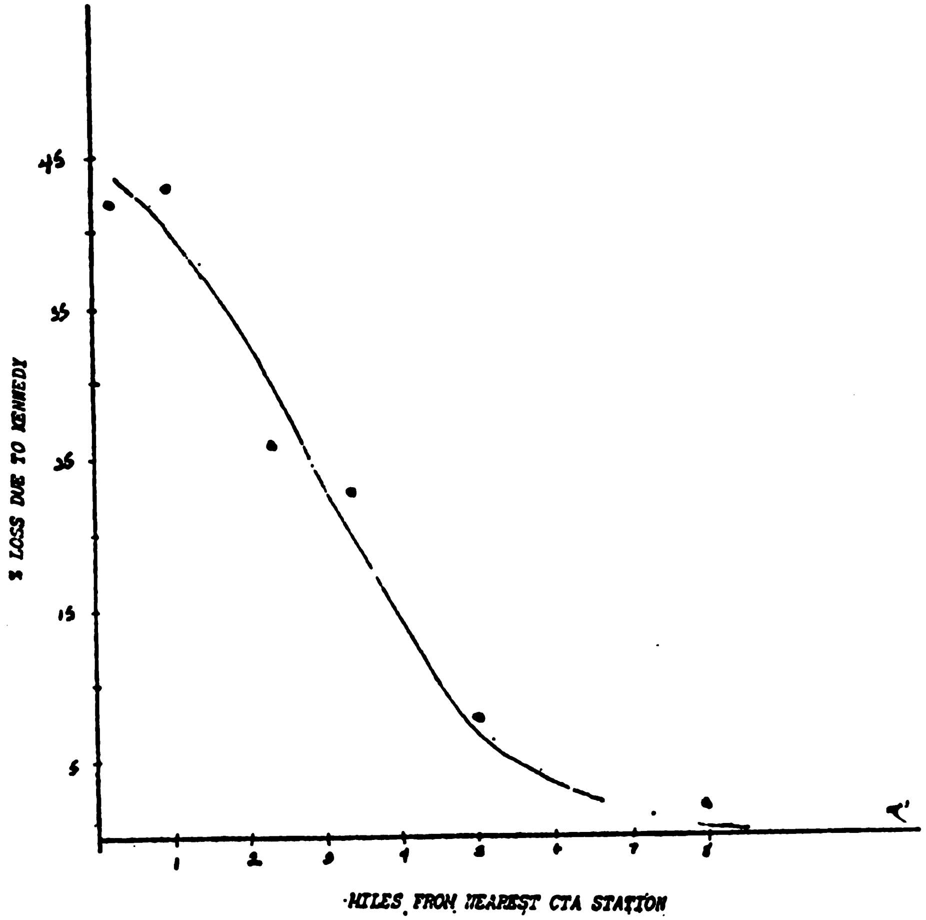


FIGURE 3.8

1971 the demand that was left was fairly inelastic—it stabilized to a fairly constant level and showed little seasonal variation or effect due to fare or energy crisis interventions. The implication is that the passengers who remained, strongly preferred the CNW service to the CTA, either because of the added comfort or because the CNW continued to serve destinations not served by the CTA.

The summary of elasticities shows reasonable consistency. They are decreasing, suggesting that at these stations also, there is a base ridership whose demand for CNW service is essentially inelastic, and that each new fare increase comes closer to approaching that base level.

The aggregate elasticity for 1973-74 would appear at first sight to be inconsistent. If we remember, however, that the aggregate includes stations which showed a demand drop in 1971, and slight demand drop in 1972, but had essentially leveled off by 1973-74, the figure does not seem unrealistic.

In all our analyses of situations where the automobile was the only mode alternative, the drops due to fare were either non-existent or very close to the standard Curtain rule (–.33).

On nearly all of the smoother plots, we have noticed an abnormal fluctuation between mid-1973 and mid-1974. This fluctuation was small enough that it would be difficult once the time series was decycled to measure an effect beyond the range of ordinary cyclical variation. To actually measure the effect would require rather precise estimation of the "true cycles" within the immediate vicinity of the energy crisis. The one effect we were aware of in all three belts was an increase in demand, despite three fare increases during that winter followed by a sharp drop in demand after the May fare increase. It would appear that the gas crisis may have mitigated the effect, temporarily, of these unprecedented fare increases. Beyond that it is sufficient to say that whatever effect the energy crisis had was minor in the context of overall ridership and seasonal variation and certainly temporary.

IV. CONCLUSIONS

At a time when the thrust of transportation planning has been the creation of complex models which require large amounts of data and considerable investments in time, money and personnel, our study has demonstrated that valuable planning information can be obtained quickly and easily from relatively inexpensive revenue/ridership data. These methods we strongly feel have the potential of helping to bridge the much discussed planner-implementer gap.

The remainder of this section is devoted to a discussion of the significance of substantive results.

Impact and Geographical Extent of Inter Modal Competition

In overall revenue, the impact of the Kennedy extension was small primarily because the impacted stations represented only 15% of the total revenue. The percentage loss at individual stations, however, was substantial, indicating that it is difficult for even a high intensity mass transit-oriented corridor such as the northwest corridor in the Chicago area to support two such similar mass transit modes if their service areas substantially overlapped. Nevertheless, it is interesting to note that in a situation such as Jefferson Park, where the train and the elevated rapid-transit stations are within one block of one another, the initial impact was only 45 percent and after the first one to two fare increases, ridership seemed fairly inelastic; even seasonal variations were slight. That the reduction in ridership was not greater is particularly surprising if we consider that there is a relatively small portion of the CBD employment located within four to five blocks of the terminus of the commuter rail station whereas the CTA has several stops right in the heart of the CBD. This seems to suggest that as similar as the two modes are, at least part of their respective passengers are drawn from different market segments.

The S-shaped distance/impact curve of Figure 3.8 indicates a four to five mile service/impact area. This rough delineation suggests the approximate distance at which a second mode becomes a viable mode alternative.

Fare Elasticities in the Presence of Competition

The standard fare elasticity used in transportation planning has been in the area of .33. Numerous other estimates also have been reported. They all generally have been computed with the automobile as the immediate mode alternative. Thus, these elasticities are not sensitive to fare policies which would raise or reduce fares of one mode relative to other (public or private) modes in a given area. With multi-modal transportation systems increasingly coming under the control of regional umbrella agencies, such policies will be considered more and more. Our work shows that passengers are considerably more fare sensitive when there is another comparable alternative than if the only other choice is the automobile. This may suggest that when modelling mode split in a given area, there may be a need for separate models for evaluating the mass transit/automobile and the competing mass transit alternatives.

Impact Intervention Period

To design an evaluation of a policy intervention, it is important to know approximately the "adjustment" period in order to avoid underestimating the effect or unnecessarily dragging out data collecting activities. Even now several fare/level of service evaluations under the auspices of the Department of Transportation are moving into the design phase. The result of our study suggests that any such evaluations should plan for data collection about six months after implementation. That is when the impacts seem to be more or less stabilized.

Energy Crisis

There have been a number of conjectures in the popular media suggesting that the shortage of gas significantly boosted mass transit ridership but very few definitive evaluations. Our study indicated that the impact was very small relative to total ridership and seasonal variations, and at most the effect was temporary.

APPENDIX A

Smoothing

In his limited circulation publication, **Exploratory Data Analyses,** Tukey

(1971) has described and demonstrated an exploratory approach to data analysis, an approach in which the investigator, using a few simple arithmetic and graphical tools and aided by a strong dose of common sense, can directly examine raw data and extract from it useful information. The approach does not formally test hypotheses. Rather, it may lead the investigator to the formulation of hypotheses which can then be tested with more traditional statistical methods.

For purposes of completeness the details of the smoothing techniques are presented here. Further details (of smoothing) and a rationale are presented in Beaton and Tukey (1974).

The first step in smoothing raw data consists of taking running medians. Thus, if the raw data are $x_1, x_2, x_3, \ldots y_n$:

$$y_i = \text{med}\ (x_{i-1}, x_i, x_{i+1})$$

where med(·) stands "for median of" and y, the smoothed data point. Any further application of the procedure causes no changes. The two end points of the graph are without adjacent points to help determine their smoothed position. For such points the following rule may be used:

$$y_n = \text{med}\ (x_n, y_{n-1}, [3y_{n-1} - 2y_{n-2}])$$

Data smoothed by running medians can often result in stabilized peaks: which may be brought into line by treating x_3 as the "end point" of the series x_1, x_2, x_3 and x_4 as the "end point" of x_4, x_5, x_6. After applying the above procedures we took weighed running means, a procedure called "Hanning":

$$z_i = \tfrac{1}{4}\ (y_{i-1} + 2y_i + y_{i+1})$$

The smoothed graph is not the only final product, however. The information (such as unusual peaks—snow storms) can be retained in a plot of residuals obtained from subtracting the smoothed point from the original observation:

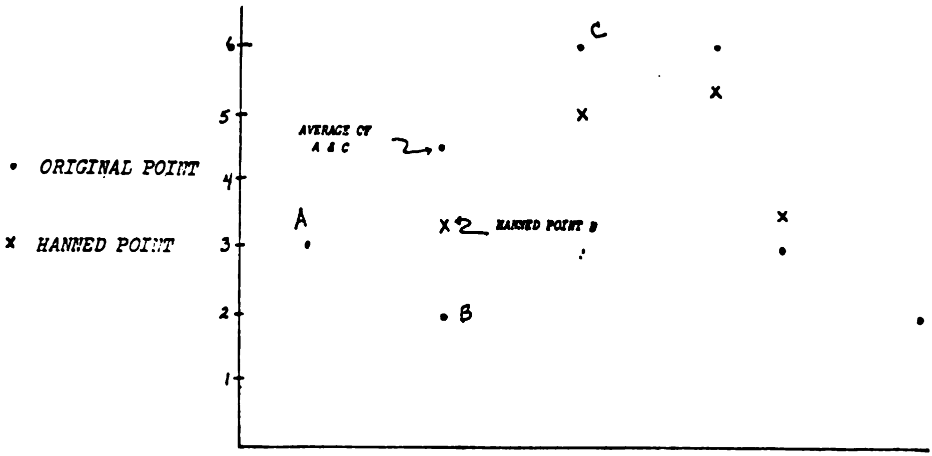
$$r_i = x_i - z_i$$

where r_i is the residual, x_i is the original point, z_i is the smoothed point. This plot can be very helpful in drawing attention to unusual situations that might be otherwise missed.

Brief numerical examples of all three procedures are presented in Table A-1

EXAMPLE OF SMOOTHING FROM GRAPH PAPER

Running Weighted Averages or "Hanning"



HANNED DATA POINT B

= 1/4(A+2(B)+C)

= 1/4(13)

= 3.25

FIGURE A-1

with sample computations. However, it would be misleading to assume any of these methods actually involve repeated arithmetic computation. Completely satisfactory results can be obtained from overlaying tracing paper on a plot of raw data and smoothing "by eye." An example of the tracing paper method of running means is presented in Figure A-1. In all, a complete smoothing, plotting and subtraction process for a data set similar to that presented in this paper should take no more than two hours to complete.

FOOTNOTES

1 This paper is a portion of a larger research effort which is reported in full detail in Burrows, et al, 1976.

2 It would be misleading to draw system-wide operating policy conclusions from the results presented in this paper alone. For greater detail see Burrows, et al, 1976.

BIBLIOGRAPHY

Beaton, Albert E. and Tukey, John W. (1974). "The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data," *Technometrics*, Vol. 16, No. 2, pp. 174-184.

Burrows, B., et al, (1975). *A Time Series Analysis of Commuter Rail Ridership Data*, College of Urban Sciences, University of Illinois Circle Campus, Chicago, Illinois.

Tukey, John W. (1971). *Exploratory Data Analysis*, Limited Preliminary Edition, Addison Wesley Publishing Company.