CHAPTER 6

Data, Sample, and Variables

6.1. Introduction

This chapter describes the data base developed to provide a test estimation of the model theory. Before describing the data preparation efforts, however, it is useful to discuss the objectives of the empirical analysis in order to set the data preparation efforts in perspective.

Primarily, the objective of this empirical effort is not to produce numerical estimates of model parameters for their own sake, but instead to illustrate and test the application of the theory. In keeping with this objective, we shall not attempt to estimate equations for all types of trips observed in the sample, but rather shall focus our efforts on modeling those trip types where the decision-making process is relatively clear and direct and where reasonably accurate data can be developed for statistical estimation.

This reasoning has led us to concentrate our effort on two trip purposes, work and shopping trips. For work trips we will limit our empirical analysis to choice of mode. This is done for two reasons. First, the choices of trip destination and frequency of travel are highly constrained for work trips. Accordingly, these choices are of less interest than the choices of mode or time of day of travel for work trips or the choice of destination or trip frequency for non-work purposes. Second, the analysis of the choice of time of day of travel to work, while highly interesting because of its policy implications, is confounded by the lack of available data on travel times in off-peak hours. Estimates of off-peak travel times can be developed by making assumptions about the relationships between peak and off-peak speeds, but the resulting parameter estimates may be unreliable because of the dependence on assumed data.

For shopping trips, all four aspects of the decision process are of interest, since the choices of where to shop and how frequently to shop are open to the trip-maker. We plan to analyze all four aspects of the shopping trip decision, in part because all four aspects are of interest, and in part because it is desirable to illustrate the complete workings of the model for at least one trip purpose. Doing this will require making up off-peak travel times, but we have discovered that such a small proportion of shopping trips (especially transit shopping trips) both originate and terminate in the peak hours that we would probably have to do this anyway. As a result, some of the empirical results will be confounded by the nature of the data used in the analysis, but nevertheless, the analysis should provide a good illustration of the complete workings of the theory.

To simplify the analysis at this initial estimation stage we have also decided to limit the number of modes to two, auto driver and transit rider. Moreover, we have deliberately sought to select a city for analysis where the transit system is largely limited to a single transit mode so that the estimation is not complicated by differences between transit modes. Auto passenger trips have been excluded from the analysis for several reasons. To avoid confounding the parameter estimates for auto driver trips, auto passenger trips should be separated rather than simply lumped in with the former. Moreover, we have decided to forego analyzing auto passenger trips in this initial analysis both because the motivation for these trips is more complex than that for auto driver or transit rider trips, and because the data needed to analyze these trips properly are not available for many auto passenger trips.

As a result of these decisions, the data preparation efforts have concentrated on developing as accurate a data base as possible for a relatively few—but highly important—trip types.

The first step in obtaining and preparing data for a test estimation of the demand model was selecting a metropolitan area with appropriate data on individual trip records and household characteristics, and well-developed data on travel times and costs. After comparing a number of metropolitan areas, the Pittsburgh area was selected for reasons that are described in the section immediately below. Following this is a discussion and evaluation of the available Pittsburgh data. The next section describes briefly the selection of the areas within the Pittsburgh region which were sampled to get a suitable cross-section of households and trips. The final section of the chapter describes the actual variables to be used in the statistical analysis.

6.2. Selection of metropolitan area

The important criteria in the selection of a city for empirical analysis are that it should have the following:

(1) Good home interview data from a recent survey. This means that the questionnaires should be designed to provide appropriate, detailed information on the individual trips and on the socioeconomic characteristics of the trip-makers. Further, the individual trip records must be accessible rather than already buried in zonal aggregates.

(2) Good transportation network data. Ideally, the city would have good coverage on highway driving speeds and transit system operations so that accurate travel times can be computed. Data should be available for both the peak and off-peak times of day.

(3) To simplify the analysis at this initial stage of estimation, we want to pick a city with only one CBD and that is basically a two-mode city, preferably with only bus and auto. In addition, it would be desirable for the city to have a fairly typical socioeconomic profile.

We originally obtained a list of cities which reportedly have good data from recent surveys. A number of these were eliminated because they have two CBDs—e.g., Minneapolis–St. Paul and San Francisco–Oakland–Berkeley—or their socioeconomic make-up is unusual—e.g., Wilmington, Miami, San Diego—or they had more than two basic modes—e.g., Cleveland.

A number of the remaining cities were eliminated because of the limited data available on the household questionnaire—e.g., Baltimore, Syracuse, and Rochester—or because the city offered no apparent data advantages and its home interview survey was less recent than that of some of the alternatives—e.g., Baltimore (1962), Toledo (1964).

Of the list of cities, the two which were most seriously considered were Washington, D.C. and Pittsburgh. Even though Washington is obviously an atypical city as the nation's capital, it has a relatively good data base. After making field trips to both Washington and Pittsburgh, Pittsburgh was chosen. Both cities have relatively good home interview data but both are lacking in coverage on highway driving speeds. Thus, there was no obvious data advantage in selecting Washington and it has some potential disadvantages in terms of atypicality. The concentration of government office employment is obviously atypical. Other atypical aspects of Washington are the large black population, the existence of much work travel from outside the central city into the central city (relative to other cities), age (median age is low), and transientness. In addition, Washington's CBD is very large in land area and car pooling and taxi riding are unusually extensive.

A general discussion of Pittsburgh's transport system and socio-economic characteristics is useful to understanding the implications of using Pittsburgh for this study. The big corridors go to the east and south and all the streetcar lines, some of which are on their own right-of-way, go to the south. Streetcars run all the way into downtown. There is significant cross-town traffic from the east corridor to the south corridor. There is only one superhighway, the Penn-Lincoln Parkway, going east and southwest. Pittsburgh is mostly a dual-mode city, auto and bus. A few streetcar lines operate in the southern corridor. The railroad commuter service is very small so that a fairly large universe of two-mode choice trips is available.

The old pre-automobile towns in the Pittsburgh area were located in flatlands along the rivers and valleys, and are widely dispersed. These areas are now generally poor as the automobile and modern building techniques have permitted newer residential developments in hillier, previously inaccessible areas. The more affluent and mobile families live in these areas. The region is extremely hilly and thus the pattern of development has not been one of simple outward growth, but also one of upward implosion, we might say, leading to a very heterogeneous land-use pattern and a jumbled and complicated street pattern.

Table 6.1 gives a comparison of selected characteristics between Pittsburgh and all SMSAs of over 250,000 population. From this it can be seen that the population is larger and considerably more dense in Pittsburgh than in the average SMSA of over 250,000. Of the socio-economic characteristics, the principal differences appear to be the relatively lower median family income in Pittsburgh, the relatively high concentration in manufacturing employment, and the relatively low percentage of non-whites in the total SMSA. However, the percent of non-whites in the central city is the same as that for the average of SMSA's of over 150,000 population.

6.3. *Review and evaluation of the Pittsburgh data*

The basic data sources for this study were the existing home interview and network information available from the files of the Southwestern Pennsylvania Regional Planning Commission (SPRPC). Their information is compiled from a number of sources, but a large part of it is based on a 4 percent sample home interview survey of the six counties in the

Table 6.1

Comparative profiles of Pittsburgh SMSA and the average of all SMSA's over 250,000 population.*

	Average	Pittsburgh
Population, 1960 (thousands)	1005	2406
Population/square mile	443	789
Median school years (persons over 25)	11.1	10.6
Occupation group (percent)		
Prof., tech.	12	12
Man.,prop.	9	7
Clerical	17	15
Sales	7	8
Craftsmen	13	16
Operatives	17	17
Service	8	8
Laborers	4	8
Industry group (percent)		
Manuf.	29	37
Transp., util.	7	7
Wholesale & retail	18	17
Finance, insurance	5	3
Bus. & rep.	2	2
Personal service	5	4
Prof.	11	11
Public admin.	5	3
Median family income	$6,439	$5,954
Percent of population in central city	51	25
Percent non-white population in SMSA	10.8	6.7
Percent non-white population in central city	16.7	16.7

* *Source:* U.S. Census of Population, 1960.

Pittsburgh metropolitan area. A second large body of compiled data had to do with the transportation network itself. In addition to the coded computer network there exist speed and delay data for portions of the highway system and considerable information regarding public transit.

We did not undertake extensive field work, but instead confined our field work to checking the existing SPRPC data and developing limited supplementary information where needed. Our review and evaluation of the home interview and network data are given below.

6.4. The home interview survey

The Home Interview Survey (HIS) was done in September, October, and November of 1967. Two basic questionnaires were compiled: a person and trip questionnaire which collected information on all trips taken by the sample household in the previous 24-hour period, and a household questionnaire which obtained a wide range of information on the household's socioeconomic characteristics. All the data have been put on tape. The HIS of person and trip data consists of three magnetic tapes containing about 30,000 records each. It is sorted by survey number, person number, and trip number, with survey number equivalent to family number. The HIS of household data is one magnetic tape of about 30,000 records, sorted by survey number. The survey number of both files refer to the same family so that the household data and the trip data can be readily combined. The individual trip records are accessible, unlinked trip data are available, and the data are sorted so that it is fairly easy to follow the household's trip behavior for a 24-hour period.

The two questionnaires are shown in fig. 6.1 and 6.2. It is clear from a review of these forms that they generate a wealth of information on trip behavior and household characteristics. Nevertheless, there are a number of shortcomings in the Home Interview Survey when used for our purposes in calibrating and testing an urban travel demand and modal split model. Perhaps the greatest omission is the lack of information on alternatives. No questions are asked about alternative modes, destinations, or travel times.[1] Thus, it is difficult to identify the alter-

[1] The 1965 Chicago Area Transportation Study (CATS), for example, questioned households on their alternative modes of transportation and collected portal-to-portal travel times and costs for both the mode selected and the best alternative means of transportation. This information was used effectively in two important studies of modal choice: Warner (1962) and Moses and Williamson (1963).

TYPE OF INTERVIEW: ——————

SAMPLE NUMBER: DIST. WK. DAY LIST LINE

DISTRICT NO.

1. 2. 3. 4. 5. 6. 7. 8. 9. 10.

WEEK OF SURVEY

0. 1. 2. 3. 4. 5. 6. 7. 8. 9.

0. 1. 2. 3. 4. 5. 6. 7. 8. 9.

TRAVEL DAY: MON. TUE. WED. THU. FRI.

LIST NO. GUIDE LINE →

0. 1. 2. 3. 4. 5. 6. 7. 8. 9.

0. 1. 2. 3. 4. 5. 6. 7. 8. 9.

LINE NO.

0. 1. 2. 3. 4. 5. 6. 7. 8. 9.

0. 1. 2. 3. 4. 5. 6. 7. 8. 9.

TYPE OF INTERVIEW: SPECIAL REGULAR

PART I

A. SAMPLE LOCATION: BLOCK ——————

STREET ADDRESS

CITY ZIP

B. TRAVEL DAY—DATE ——————

C. STRUCTURE TYPE: .... ——————

D. NUMBER OF HOUSING UNITS AT SAMPLE ADDRESS: .... ——————

E. INTERVIEW HOUSING UNIT NUMBER: .... ——————

PART II STRUCTURE AND RELATED INFORMATION

A. STRUCTURE TYPE (observation): 1 FMLY DETACHED, 2 FMLY, ROW, MULTI-FMLY, CONVERT

MOTEL-HOTEL, RESID HOTEL, RESORT, LESS 5, 5–9, ROOMING HOUSE, 10 OVER, 1 TRL., TRL. CT., INSTITU-TIONAL, OTHER

MARK IF HOUSING UNIT IS MIXED WITH NON-RESID ESTAB. →

B. STRUCTURE CONDITION (observation): SOUND, DETERIORATED, DILAPIDATED

C. NUMBER OF FLOORS IN STRUCTURE (observation): IF LESS THAN 3 FLOORS MARK

WITH ELEVATOR: 3–5, 6–7, 8–10, 11–15, 16 OR OVER

WALK-UP: 3, 4, 5 OR OVER, LESS 3

D. AGE OF STRUCTURE (ask): YEARS → 7 OR LESS, 8–17, 18–27, 28–37, OLDER 37

DATES → SINCE 1960, 1950–60, 1940–50, 1930–40, BEFORE 1930

E. LOT SIZE (ask) 1 FMLY DETACHED ONLY: AREA IN ACRES: LESS 1/4, 1/4–1/2, 1/2–1, 1–3, OVER 3

F. NUMBER OF BEDROOMS AND BATHROOMS IN HOUSING UNIT (ask):

1. BEDROOMS: 1, 2, 3, 4, 5, 6 OR OVER, EFFIC'Y, DORM

2. BATHROOMS (INCL. POWDER ROOM): OUTHSE., SHARE, 1, 2, 3, 4, 5 OR OVER

G. ARE THERE ADEQUATE PARKING SPACES AVAILABLE FOR FAMILY: YES, NO, NO AUTO

H. NUMBER OF HOUSING UNITS IN MULTI-FMLY STRUC. (ask): 0, 100, 200, 300

0, 10, 20, 30, 40, 50, 60, 70, 80, 90

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

PART III HOUSEHOLD INFORMATION (ask):

A. NUMBER OF RESIDENTS IN HOUSEHOLD (ask): 0, 10, 20

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

B. NUMBER OF RESIDENTS 5 YEARS OF AGE OR OVER IN HOUSEHOLD (ask): 0, 10, 20

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

C. NUMBER OF OUT-OF-REGION VISITORS AT HOUSEHOLD ON TRAVEL DAY (ask): 0, 10, 20

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

D. AUTOS AVAILABLE TO HOUSEHOLD MEMBERS (incl. company & rental autos): (ask)

1. TOTAL AUTOS AVAILABLE: 0, 1, 2, 3, 4, 5 OR MORE

a. AUTOS AVAILABLE FOR UNLIMITED USE: 1, 2, 3, 4, 5 OR MORE

b. AUTOS AVAILABLE FOR RESTRICTED USE (Incl. Visitors Autos): 1, 2, 3, 4, 5 OR MORE

E. OTHER MOTOR VEHICLE AVAILABLE (motorcycles, motor bikes, or motor scooters): 1, 2, 3, 4, 5 OR MORE

PART IV MOBILITY STATUS OF HOUSEHOLD (ask):

A. NUMBER OF YEARS AT PRESENT ADDRESS (ask): 0, 10, 20, 30, 40

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

B. DID YOU LIVE IN THE SIX-COUNTY REGION ON JAN. 1, 1960 (ask): YES, NO — INTERVIEWER! NOW GO TO TRIP REPORT

PART V PERSONAL INFORMATION—ASK AT END OF INTERVIEW!

A. OWNERSHIP STATUS (ask): OWNED, RENT, COOP OR LEASED

B. IF OWNED (OR COOP-LEASE) APPROXIMATE MARKET VALUE OF HOUSE & GROUNDS ($000):

(1) UNDER 5.0, (2) 5.0–7.4, (3) 7.5–9.9, (4) 10.0–12.4, (5) 12.5–14.9, (6) 15.0–17.4, (7) 17.5–19.9, (8) 20.0–24.9, (9) 25.0–34.9, (10) 35.0 OVER

C. IF RENTED (OR COOP-LEASE) WHAT IS THE MONTHLY RENT, EXCLUDING UTILITIES ($): INTERVIEWER ESTIMATE — 1. PVT. OWNED, 2. PUBLIC HOUSING

(1) UNDER 40, (2) 40–49, (3) 50–59, (4) 60–69, (5) 70–79, (6) 80–99, (7) 100–119, (8) 120–149, (9) 150–199, (10) 200 OVER

D. HIGHEST EDUCATION ATTAINED BY HEAD OF HOUSEHOLD: — REFUSAL

(YEARS) (1) 6 OR LESS, (2) 7–8, (3) 9–10, (4) 11–12, (5) SOME COLLEGE; RECEIVED DEGREE OR CERT → (6) LESS 4 YR SCH. OR (7) COLLEGE, (8) GRAD. SCH.

E. TOTAL HOUSEHOLD GROSS INCOME IN 1966 ($000): ESTIMATE

(1) UNDER 3.0, (2) 3.0–4.9, (3) 5.0–6.9, (4) 7.0–9.9, (5) 10.0–14.9, (6) 15.0–19.9, (7) 20.0–24.9, (8) 25.0–29.9, (9) 30.0–34.9, (10) 35.0 OVER

F. NUMBER OF WORKERS CONTRIBUTING TO HOUSEHOLD GROSS INCOME: 1, 2, 3, 4, 5 MORE

G. RACE OF HEAD OF HOUSEHOLD (observation): WHITE, NON-WHITE, OTHER

■ TIME REQUIRED TO COMPLETE INTERVIEW (hrs.): 1/2–1, 1–1 1/4, 1 1/4–1 1/2, 1 1/2–2

PART VI SUMMARY DATA—COMPLETE AFTER YOU LEAVE HOUSEHOLD

A. NUMBER PERSONS 5 YEARS OF AGE OR OVER MAKING TRIPS: 0, 10, 20

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

B. NUMBER PERSONS 5 YEARS OF AGE OR OVER WITH NO TRIPS: 0, 10, 20

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

C. NUMBER PERSONS 5 YEARS OF AGE OR OVER WITH TRIPS UNKNOWN: 0, 10, 20

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

D. NUMBER OF LICENSE DRIVERS IN HOUSEHOLD: 0, 10, 20

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

E. TOTAL NUMBER OF TRIPS REPORTED: 0, 10, 20, 30, 40, 50

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

F. PERSONNEL NUMBER: 500, 600, 700, 800, 900

00, 10, 20, 30, 40, 50, 60, 70, 80, 90

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

INTERVIEWER'S REMARKS:

RESPONDENT'S TELEPHONE NO. ——————

TELEPHONE CALLS MADE		PERSONAL CALLS MADE	
DATE	TIME	DATE	TIME
/ / 67	AM PM	/ / 67	AM PM
/ / 67	AM PM	/ / 67	AM PM
/ / 67	AM PM	/ / 67	AM PM

"I CERTIFY THAT THE INFORMATION LISTED ON THIS INTERVIEW FORM IS ACCURATELY AND COMPLETELY RECORDED AS REPORTED TO ME BY THE RESPONDENTS."

OTHER REMARKS:

INTERVIEWER'S SIGNATURE —— DATE

FINAL DISPOSITION CODE — COMPLETED INTERVIEW, PARTIALLY COMPLETED, NONINTERVIEW (SPECIFY TYPE)

TYPE A—TRIPS POSSIBLE: REFUSAL, AWAY, OTHER | TYPE B—NO TRIPS POSSIBLE: OUT REGION, VACANT, OTHER | TYPE C—NO SAMPLE POSSIBLE: DEMOLISHED, NONRESID., OTHER

FINAL EDIT: PASS, SOME ERRORS, FAIL — DATE — INITIALS

QUALITY CONTROL CALL-BACK MADE: GOOD, FAIR, POOR — DATE — INITIALS

Fig. 6.1. Southwestern Pennsylvania Regional Planning Commission household data.

SAMPLE NUMBER — DIST. — WEEK — DAY — LBT — LINE

FOR EACH PERSON SHEET ____ OF ____ SHEETS

TOTAL TRIPS REPORTED ________

A. PERSON INFORMATION (for each person 5 years of age or over)

1. PERSON NUMBER (USE "01" FOR HEAD OF HOUSEHOLD) (Punch for all cards)

2. RELATIONSHIP TO HEAD OF HOUSEHOLD (CODE ONE OF THE FOLLOWING)
   1. HEAD OF HOUSEHOLD
   2. SPOUSE
   3. CHILD
   4. PARENT
   5. BROTHER or SISTER
   6. GRANDPARENT
   7. OUT-OF-REGION VISITORS
   8. BOARDER
   9. ROOM-MATE
   0. OTHER

3. SEX
   1. MALE
   2. FEMALE

4. AGE (RECORD ACTUAL AGE)

5. AUTO DRIVING STATUS
   1. LICENSED–DROVE WITHIN LAST MONTH
   2. LICENSED–DID NOT DRIVE WITHIN LAST MONTH
   3. LEARNERS PERMIT
   4. NOT LICENSED (ALL PERSONS UNDER 16)

6. OCCUPATION STATUS (WORKER MUST BE 14 YEARS OF AGE OR OLDER)
   1. WORK – ONE JOB FULL TIME (6 HOURS OR MORE)
   2. WORK – ONE JOB PART TIME
   3. WORK – TWO OR MORE JOBS
   4. STUDENT
   5. STUDENT – WORK PART TIME (LESS THAN 6 HOURS PER DAY)
   6. HOUSEWIFE
   7. HOUSEWIFE – PART TIME JOB (LESS THAN 6 HOURS PER DAY)
   8. RETIRED
   9. UNEMPLOYED
   0. OTHER (specify) ________

   NOTE: IF PERSON WORKS 6 HOURS OR MORE PER DAY CONSIDER THEM A FULL TIME EMPLOYED PERSON.

B. WORK PLACE AND OCCUPATION DATA (record for person if employed part or full time)

JOB STATUS	OCCUPATION OR POSITION TITLE (Give in Detail)	NAME AND TYPE OF BUSINESS OF EMPLOYER	STREET ADDRESS OF EMPLOYER	DID THIS PERSON WORK AT THIS LOCATION ON THE TRAVEL DAY?
1. PRIMARY JOB STATUS (TRIP CODE 01) 1. FULL TIME 2. PART TIME	IF A PERSON WORKS 6 OR MORE HOURS PER DAY CONSIDER THEM AS WORKERS–NOT HOUSEWIFE OR STUDENT	(NAME OF EMPLOYER – COMPANY NAME) (DESCRIBE BUSINESS, INDUSTRY OR SERVICE OF COMPANY) CHECK ONE ○ MANUF. ○ WHOLESALE ○ RESID. ○ RETAIL ○ SERVICE ○ OTHER	(STREET NO.) (DIR.) (STREET NAME AND SUFFIX) LOCATION STATUS — CIVIL DIVISION: ○ CITY ○ BORO ○ TWP. (CHECK ONE) RECORD COUNTY CODE — (NAME)	1. YES – WENT TO WORK IN VEHICLE 2. YES – DROVE A TRUCK TO WORK 3. YES – WALKED OR WORKED AT HOME 4. NO – WORKED SOMEPLACE ELSE 5. NO – DAY-OFF, HOLIDAY OR VACATION 6. NO – SICK 7. NO – TEMPORARILY LAYED-OFF 8. NO – STRIKE 9. NO – OTHER
2. SECONDARY JOB STATUS (TRIP CODE 02) 1. FULL TIME 2. PART TIME		(NAME OF EMPLOYER – COMPANY NAME) (DESCRIBE BUSINESS, INDUSTRY, OR SERVICE OF COMPANY) CHECK ONE ○ MANUF. ○ WHOLESALE ○ RESID. ○ RETAIL ○ SERVICE ○ OTHER	(STREET NO.) (DIR.) (STREET NAME AND SUFFIX) LOCATION STATUS — CIVIL DIVISION: ○ CITY ○ BORO ○ TWP. (CHECK ONE) RECORD COUNTY CODE — (NAME)	

TO-DUP L. 1-16

C. TRIP DATA

TRIP PURPOSE CODES
- 00. HOME
- 01. WORK – PRIMARY JOB LOCATION
- 02. WORK – SECONDARY JOB LOCATION
- 03. WORK – OTHER JOB LOCATION
- 04. SHOP
- 05. SCHOOL
- 06. SOCIAL (AT PRIVATE HOME)
- 07. RECREATIONAL (AT NON-RESID. LOCATION)
- 08. PERSONAL BUSINESS
- 09. SERVE PASSENGER
- 10. CHANGE TRAVEL MODE
- 11. RIDE (NO PARTICULAR DESTINATION)
- 12. WALK TO NEXT TRIP ORIGIN

COUNTY CODES
1. ALLEGHENY
2. ARMSTRONG
3. BEAVER
4. BUTLER
5. WASHINGTON
6. WESTMORELAND

MODE OF TRAVEL CODES
1. AUTO DRIVER
2. AUTO PASSENGER
3. RAILROAD PASSENGER
4. COMMERCIAL BUS PASSENGER
5. STREET CAR OR INCLINE PASSENGER
6. TAXI PASSENGER
7. TRUCK PASSENGER
8. SCHOOL BUS PASSENGER
9. MOTORCYCLE
0. WALKED TO NEXT TRIP ORIGIN

TYPE OF PARKING CODES
1. STREET FREE
2. STREET METER
3. LOT FREE
4. LOT PAID
5. GARAGE FREE
6. GARAGE PAID
7. LEFT FOR SERVICE
8. RESIDENTIAL PROPERTY
9. NOT PARKED (PICKUP OR DROP-OFF)
0. NOT PARKED (RIDE, NO PARTICULAR DESTINATION)

NOTE: FOR CHOICE OF MODE UNDER TRANSIT TRIP ASK: "COULD THIS TRANSIT TRIP HAVE BEEN MADE CONVENIENTLY AS AN AUTO DRIVER OR AUTO PASSENGER." CODE
1. NO CHOICE
2. YES, AS AUTO DRIVER
3. YES, AS AUTO PASS.

1 TRIP NO.	2 TRIP PURPOSE	3 TRIP END LOCATION (RECORD FULL STREET ADDRESS)	4 CIVIL DIVISION	5 LAND USE ACTIVITY (DESCRIBE ESTABLISHMENT)
	(IF PURPOSE IS 00, 01 OR 02 OMIT ITEMS 3 TO 5)	(STREET NO.) (DIR.) (STREET NAME AND SUFFIX) LOCATION STATUS	(NAME) ○ CITY ○ BORO ○ TWP. RECORD COUNTY CODE	○ MANUF. ○ WHOLESALE ○ RESID. ○ RETAIL ○ SERVICE ○ RECREATION
	(IF PURPOSE IS 00, 01 OR 02 OMIT ITEMS 3 TO 5)	A N D	(NAME) ○ CITY ○ BORO ○ TWP. RECORD COUNTY CODE	○ MANUF. ○ WHOLESALE ○ RESID. ○ RETAIL ○ SERVICE ○ RECREATION
	(IF PURPOSE IS 00, 01 OR 02 OMIT ITEMS 3 TO 5)	A N D	(NAME) ○ CITY ○ BORO ○ TWP. RECORD COUNTY CODE	○ MANUF. ○ WHOLESALE ○ RESID. ○ RETAIL ○ SERVICE ○ RECREATION

RECORD FOR ALL TRIPS					FOR ALL AUTO DRIVER TRIPS			FOR ALL TRANSIT TRIPS		
6 MODE OF TRAVEL	7 TIME ARRIVED AT DESTINATION	8 BRIDGES CROSSED DURING TRIP – RIVER	9 BRIDGES	10 BLOCKS WALKED	11 TYPE OF PARKING	12 DURATION OF PARKING	13 COST OF PARKING	14 FARE PAID	15 CHOICE OF MODE	16 WAIT TIME AT ORIGIN & TRANSF. WALK AND WAIT TIME
(SEE CODES ABOVE)	HOUR MINUTE AM PM (RECORD TIME TO NEAREST MINUTE)	○ OHIO ○ ALLEGH. ○ MONON.		AT ORIGIN AT DESTINATION	(SEE CODES ABOVE)	HOURS MINUTES	DOLLARS CENTS	DOLLARS CENTS		WAIT TIME (MINUTES) TRANSF. TIME (MINUTES)
(SEE CODES ABOVE)	HOUR MINUTE AM PM (RECORD TIME TO NEAREST MINUTE)	○ OHIO ○ ALLEGH. ○ MONON.		AT ORIGIN AT DESTINATION	(SEE CODES ABOVE)	HOURS MINUTES	DOLLARS CENTS	DOLLARS CENTS		WAIT TIME (MINUTES) TRANSF. TIME (MINUTES)

Fig. 6.2. Southwestern Pennsylvania Regional Planning Commission person and trip information.

natives that are relevant to the traveler, and any data on travel times and costs for alternatives must be generated by the researcher.

One specific result of this omission is the difficulty caused in studying trips in which a traveler parks his automobile at a transit station ("park-ride") or is driven to a transit station and dropped off ("kiss-ride"). In an area where the transit trip involves park-ride or kiss-ride, the required data are available for transit trips employing these modes. It is easy to determine that the relevant alternative is the auto trip. For auto trips in this area, however, it is not clear whether the relevant transit trip is a park-ride or kiss-ride.

Data are not available from the Home Interview Survey on car occupants from households other than those included in the survey. Thus, it is impossible to analyze car pools insofar as they involve riders from other households. In this context, it is important to recognize that some 30 percent of all trips (based on the analysis of a sample of 150 households presented in appendix A) are auto passenger or serve-passenger trips. Thus, the fact that kiss-ride, park-ride, and car pooling are difficult or impossible to analyze with the available data is a serious shortcoming.

Another important omission for both auto and transit trips is lack of information on the highway or bus route taken. Thus, in this study, the route had to be determined by the researcher on the basis of his judgment. This is particularly troublesome for transit trips when there are choices between infrequently scheduled express buses and frequently scheduled local buses.

A major omission in the Home Interview Survey was "time travel originated". The forms did contain "time of arrival" but this is useless by itself, as total perceived travel time could not be computed. The existence of "time travel originated" would have been very useful although it would have clearly been perceived rather than actual travel time. "Desired time of arrival" would also be extremely useful as a means of estimating the effects of the reliability of the transportation alternatives.

On the transit trips a number of items were missing that would have improved the usefulness of the data. In particular, there was no measure of comfort or convenience. Therefore, the existence of crowding, the necessity to stand, lack of air conditioning, association with undesirable people, etc. cannot be taken into account in evaluating the choice of mode. It would also have been interesting and useful to know whether the traveler normally used public transit or normally went by another mode.

6.5. Network data

The network data actually used in this study came from a variety of sources. The principle source of topological data was the coded networks compiled by the SPRPC. SPRPC actually has two highway networks. The first and older one employed a five-digit node numbering scheme, and the links drawn on the maps corresponded to actual road segments. The second and newer network was a computer-plotted straight line (spider web) network utilizing a four-digit node numbering scheme. This latter network was designed for eventual use as a multimode network; however, SPRPC is currently employing it in their forecasting work. It seemed clear that this four-digit network would be hard to relate to the base maps, owing to the straightline links, and so we selected the older, not entirely updated maps of the five-digit network. The two networks are convertible, as all nodes are supposedly in the same places.

No written documentation on the data sources for the highway network exist; however, two data banks were used as sources. One was a speed and delay study done in 1967 by the Bureau of Traffic Planning of the City of Pittsburgh. This study recorded actual peak-hour time trials on major and intermediate arteries within the city limits but outside of the CBD (Golden Triangle). The second source was called the Rational Priority File and was done by another state agency. The file essentially consists of zero-volume safe speeds for state numbered highways. The remaining roads were assigned speeds determined strictly by judgment, and the two original data banks have also been modified. A minimum speed of 10 mph was required on all links, so all the CBD streets ended up with that speed.

The speeds actually coded on the highway network were called assignment speeds because the network was used for traffic assignment purposes. However, realistic peak-hour speeds resulted in an underassignment to downtown streets, so the speeds were frequently raised or lowered to bring assigned traffic volumes into closer accord with ground-count traffic volume. These network data sources are summarized below.

Area	Data available	Comments
CBD	Assignment speeds; uniform 10 mph because of assumed minimum speed.	Based on assumptions rather than evidence. Probably inaccurate.
City of Pittsburgh—outside CBD	Assignment speeds and speed and delay study.	Speed and delay data appear to be as good as we can expect. Assignment speeds will probably not be useful for our purposes.
Outside Pittsburgh	Assignment speeds.	Probably not adequate. Assignment speeds may be speed limits in many cases.

The principal problem with the coded networks is the lack of underlying data for any part of the region except the city of Pittsburgh outside the CBD. Moreover, the available underlying data—the speed and delay studies—only give peak-hour speeds, so assumptions about speeds are needed to analyze any non-peak-hour trips. The networks are useful in defining the physical characteristics of the street network and identifying major arteries. Because of the importance of travel time to overall results we carefully examined the underlying speed and delay data generated by the Bureau of Traffic Planning. We attempted to check and supplement this data base with a number of travel times obtained by the Charles River Associates' staff actually driving the principal streets. All three areas were actually driven with a view to the following items.

(1) Establish an average peak period speed for the CBD. The Golden Triangle is very densely built up and covers around 100 blocks. The boundaries are the rivers and the Crosstown Boulevard–Civic Center area on the east. The streets are small and congested. Trolley tracks, passenger islands, and crowded bus stops are on just about every block. It thus seemed reasonable to treat all the streets as having equal speeds. A possible exception may be Liberty Street, which is the main travel street downtown.
(2) Check the speed and delay study times on some main arterials within the city limits in the eastern corridor.
(3) Check the assignment speeds on main roads outside the city limits.

To establish an average CBD speed, two different loops in the CBD were driven a number of times. The loops were selected to cover a variety of streets, that is, both two-way and one-way. One loop was driven several times in both morning and evening peak hours. A second loop was only driven in the morning peak hour. Driving each loop several times should eliminate effects of local or specific congestion, so that a reasonable average speed is calculated. The speed on heavily congested Liberty Street was 4 mph and on moderately congested Grant and Forbes and Seventh Avenue was 6 mph. The speed on lightly congested Smithfield and Wood Street was 9 mph. However, it should be noted that we did not experiment enough to establish meaningful confidence limits. In addition, it became clear that the travel time pattern downtown was complex; for example, Boulevard of the Allies seemed to have high speed, over 20 mph, low volume, and synchronized traffic lights; but no formal time trials were taken. Another single trial covering random streets downtown was taken and the average speed was 8 mph.

The speed and delay study was checked by driving four different routes in the Eastern Corridor which were all included in the speed and delay study. Three of the routes were on main radials and cross-town routes, scattered so as to obtain a representative sample. The fourth was on the parkway.

The congestion on the parkway appears to be heavy and erratic. As a result, the time trials from one day may not be indicative of average travel. The situation is complicated by the fact that inbound on-ramps are often closed in the morning peak hours. The travel times observed are substantially lower (faster) than the speed and delay elapsed travel times which were based on more trials spread out over several days. Thus, we do not feel that the speed and delay parkway data were adequately validated. This is unfortunate because the parkway is easily the most important artery in the city.

The other three routes provided a reasonable check on the speed and delay data. Observed times for individual road segments were often quite different from the speed and delay mean elapsed times. This is often due to the effects of stoplights. However, when the mean observed times were added over several road segments, they often showed good agreement with similarly aggregated times from the speed and delay data. In addition, the variations in aggregated mean observed times from the aggregated speed and delay mean times did not seem to show a bias. That is, some observed aggregate mean times were higher and some were lower.

Table 6.2

Summary comparison of Charles River Associates' (CRA) empirical highway travel times with existing data sources.

Link	CRA mean test travel time (CRA)	Speed and delay mean and travel time (S&D)	Difference (S&D − CRA)	S&D − CRA / CRA
Source: City of Pittsburgh's Speed and Delay Study				
Area: Within the city limits of Pittsburgh, but outside the downtown				
Shady (5–Forbes)	4:21	4:35	0:14	5%
Forbes (Shady–Halket)	10:18	9:10	−1:08	−9%
5 Ave. (Halket–Shady)	11:59	11:49	−0:10	−1%
Baum (Euclid–N. Craig)	3:48	5:49	+1:55	50%
Bigelow (N. Craig–7 Ave.)	3:28	5:49	2:21	68%
Liberty (12 St.–Centre)	10:25	11:29	1:04	+10%
Centre (Liberty–S. Negley)	4:15	1:47	−2:28	−58%
Negley (Centre–Stanton)	6:10	5:04	−1:06	−18%
Stanton (Negley–Butler)	6:35	5:54	−0:41	−10%
Butler (Stanton–Highland)	7:05	7:39	0:34	8%
Parkway E. (Grant–SWSSVL)	9:[illegible]	13:59	4:30	47%
Source: Southwest Pennsylvania Regional Planning Commission (SPRPC) Highway Network Assignment Speeds				
Area: Outside the city limits of Pittsburgh, but within the limits of Allegheny County				
Wm. Penn (Akers Cr.–Penn. Trnpk)	2:20	2:10	−0:10	−7%
Various roads	11:20	12:39	1:19	12%
Demmer/Ardmore	2:55	1:29	−1:26	−49%
Graham (Penn–Laketon)	2:35	1:28	−1:07	−43%
Frnkstwn (Rbnsn–Slzbrg)	11:47	7:17	−4:30	−38%
Slzbrg (Frnkstwn–Frvw)	13:[illegible]	12:43	−0:37	−5%

Comparisons of the driving times obtained by the Charles River Associates' staff with both the Pittsburgh speed and delay times and the highway assignment speeds outside the city of Pittsburgh are presented in table 6.2.[2] It should be noticed from an examination of that table that even the speed and delay study shows large variations from the mean of our empirical times for individual road segments. With the exception of the trials on the Parkway, the city generally carried out ten different peak-hour trials on each road link using the "average car" method [Walker (1967)]. Theoretically, then, the speed and delay study gives data as accurate as are economically feasible. For example, for a 5 percent maximum error, 95 percent degree of confidence, same type of road, forty time trials are required [Walker (1967)].

Travel time variations are known to be large in urban areas. They increase rapidly as the system becomes congested. It is these fluctuations that account for the differences between the city's and CRA's observed times. Percentage variation can be expected to go down as trip length goes up. Since the speed and delay study gives data by road links, usually stoplight to stoplight, the variations on each link were felt to be quite large.

After observing that the travel times were so highly variable we decided to prepare "variability of travel time" for use as an explicit variable in the model. This was measured as the variance in line-haul travel time and was computed for each trip along with total travel times. Where possible, the variable was computed from the different time trials on each link obtained from the speed and delay study. For links not covered by the speed and delay study, the variable was inferred from similar links on the speed and delay study.

The source data for the transit networks built by SPRPC were an internal set of operating schedules from 1967 provided by PAT and the other independent bus companies. These schedules reflect operating rather than scheduled speeds.

The PAT system is the only authorized bus system within Allegheny County. Independents running into downtown are permitted to drop off anywhere in the county but not pick up on inbound trips, and pick up anywhere but not drop off on outbound trips. The vast majority of the

[2] A long route outside of the city limits, but primarily within Allegheny County, was driven in the evening peak period to check the assignment speeds on roads outside Pittsburgh.

PAT routes are radially oriented with respect to downtown. They start somewhere, run into downtown, make a loop, and then run out again. Many routes are express from downtown to an outlying neighborhood where they become local routes making frequent stops. There are, however, a few circumferential routes plus a downtown distribution bus that loops around the CBD. In addition, the fare is on a zone basis; zones are different for each bus route. The fare in 1967 was a base charge of thirty cents, plus five cents to buy a transfer, plus five cents for each zone boundary crossed. Zone 1 is essentially the City of Pittsburgh, with the other zone boundaries located approximately every two miles going radially away from the city boundaries.

The bus loops in downtown were not coded into the transit network, but are represented schematically. In order to tell exactly where the bus routes go, a detailed map of the PAT system as it existed in 1967 was obtained from SPRPC and a complete set of published schedules from each bus route from 1967 was obtained. These schedules located the fare zone boundaries exactly.

In summary, the highway and transit networks were closely analyzed to determine whether they accurately reflect peak-period speeds. The highway network is insufficient but the speed and delay study is helpful. We have calculated average CBD peak period speeds and plan to use the speed and delay study data wherever possible. Outside Pittsburgh, the only available highway data are the assignment speeds. Since the speed and delay study provides peak speeds only, assumptions are needed about off-peak speeds to be able to analyze off-peak trips. The transit data are fairly complete. Both the schematic transit networks prepared by SPRPC and the underlying route maps and operating schedules are available.

6.6. *Selecting the areas to be sampled*

Experience with both the problem and the data led us to the conclusion that it would be more important to select a small sample and prepare the data for each observation meticulously than to sample broadly but prepare data less carefully. An examination of the data from one tape suggested that it would be possible to obtain an adequate sample size by taking data only from two major corridors: one east of the CBD to the city limits, and the other to the suburbs in the south. Restricting the

1 Highland and East Liberty
2 The Hill
3 Golden Triangle (CBD)
4 Oakland
5 Squirrel Hill
6 Jones-Laughlin Steel Mill
7 South Side
8 Homestead Steel Mill
9 Dormont
10 Mt. Lebanon
11 Castle Shannon
12 Bridgeville
13 Upper Saint Clair
14 Bethel Park

Fig. 6.3. Areas included in the sample.

observations to two corridors, we felt, would make the preparation of network data reasonably efficient. The areas of the region which are included in the sample are shown in fig. 6.3.

In selecting the sample from which the observations were taken, an effort was made to include the sections of Pittsburgh which are job centers and shopping centers, as well as a broad cross-section of inner-city and suburban residential neighborhoods. The following were selected as work and shopping areas: the CBD, Oakland, the Homestead and Jones-Laughlin steel mills, East Liberty, and Bethel Park. The CBD and Oakland double as both employment and shopping centers. The CBD houses major companies, government offices, department stores, and general services. Oakland houses the universities and specialty shops. The steel mills are a major source of blue-collar employment in Pittsburgh and are the main industry of the metropolitan area. The East Liberty Mall and Bethel Park's South Hills Village are, respectively, inner-city and suburban shopping centers, both of which provide alternatives to downtown shopping.

Four inner-city residential sections were chosen: the Hill, Highland Park, Squirrel Hill, and the South Side. The Hill is a large ghetto area just outside the CBD, with an 88 percent non-white population. Median income in the Hill is near the poverty level. Highland Park is a blue-collar neighborhood with a 91 percent white population. Squirrel Hill is a wealthy inner-city neighborhood with an almost exclusively white population made up of professionals and white-collar workers. The South Side is a white, blue-collar, low-income community located immediately across the river from the CBD.

The selected suburban communities form the commuting corridor to the south of the city. Furthest out are Bridgeville and Upper Saint Clair. Bridgeville was selected because it is a small town relatively more self-sufficient than some of the others. Of the suburbs chosen, Bridgeville has the highest non-white population (3 percent). Its workforce, like Highland Park's, is blue collar. Upper Saint Clair is a newer low-density, middle-class suburb that extends to the border of Allegheny County and lies at the end of transit routes. Mt. Lebanon is an older, wealthy suburb, which has reached the limits of its development. Castle Shannon and Dormont are the last two suburbs included in the sample. Both are essentially blue-collar, lower middle-class communities. Castle Shannon was chosen because it is the hub of the southern corridor's streetcar lines. Dormont was the first of the suburbs developed, as it is at the mouth of the tunnels leading downtown.

Once the residential, job, and shopping areas were selected, the trip data were taken from the tapes. The criteria used for selecting the trips to be put in the file were: (1) that the home location of a household had to be in one of the selected areas, and (2) that somewhere in the trip record an origin and a destination had to fall within the selected area. For each trip record that met these criteria, the trip data for that person's whole household was put in the file.

6.7. Variables for analysis of individual trips

The remainder of this chapter describes the variables developed for the empirical analysis. It starts with a discussion of the computations of travel times and costs for the individual observations. Data are developed for the complete door-to-door trip for both the mode used and the alternative mode. For shopping trips, data are also developed on the door-to-door travel times and costs for the actual and the alternative times of day (peak and off-peak), for the actual and alternative destinations, and for the potential destinations for households that did not make trips during the 24-hour sample period. Following this is a brief description of the available socioeconomic data.

6.8. Auto and transit line-haul times for work trips

To obtain data as accurate as feasible on individual point-to-point travel times and costs, estimates of times and costs were made manually using the underlying network information as a data base. Large and detailed maps of the Pittsburgh area were used to generate data. Both highway and transit networks were displayed. Data were generated one trip at a time working from trips taken directly from a printout of the Home Interview Survey. Since the underlying speed and delay data are restricted to peak times only, the work trips selected were limited to peak-hour trips.

Once the block face of the trip origin and destination had been identified on the map, the best route between the two was determined according to the judgment of the researcher. The auto route was selected giving priority to freeways and major arterials. For some trips information on bridges crossed could be utilized to select the appropriate route. If the route were a transit trip in the survey, then a logical minimum time path was picked manually based on the judgment of the researcher.

Once the route was chosen, all links of the route were analyzed to see if they were in the speed and delay study. If they were, the mean travel time and the variance for each such link was recorded. As mentioned earlier, the variances for these links were computed from the different time trials taken for each link in the speed and delay studies.

For links not covered by the speed and delay study, a travel time was assigned manually, utilizing the following guidelines:

(1) The highway network assignment speeds;
(2) Local land use conditions;
(3) Type of road;
(4) Topographical conditions;
(5) General knowledge of the area.

Since variances were not available for these links, they were estimated from similar links on the speed and delay study. Since it is probably easier to develop judgmental estimates of speeds than variances of speeds, it is likely that the resulting variances will be less reliable than the travel times. Once each link had been assigned a travel time, all links on the route were added up to obtain the highway travel time and variances for that particular observation.[3]

For transit trips, line-haul times were calculated from schematic network schedules, supplemented where necessary by the transit system's operating schedules.

The PAT schedule division timetables are the internal schedules used by the transit company itself and represent as complete a description of the transit system as exists. CRA obtained a set of these schedules for each bus route from 1967 or 1968. The transit network was coded originally by SPRPC using these schedules. Determining times was somewhat simplified by working basically from the SPRPC network because the schedules are voluminous and complex.

Further validation of the SPRPC data source did not seem economical. The schedules allow for make-up time at strategic points so that drivers

[3] Summing the variances on each link to get the total trip variances requires the further assumption that all the covariances between links are zero. That is, it is assumed that if travel time is higher than average on one link on a given time trial, it is not likely to be higher on successive links. If the actual covariances are positive, the assumption of zero covariances results in an underestimate of the trip variances, and conversely if the covariances are actually negative. Positive covariances would occur, for example, on a rainy day where the entire system is slowed down. Negative covariances might occur with an auto accident, where the speeds on links approaching the location of the accident are lower than usual. Since no data were available on covariances, we simply assumed they were zero. However, this assumption probably adds to the unreliability of the variance estimates.

of buses can get back on schedule if they are held up by congestion. In addition, running times are assigned according to different time periods during the day. That is, a bus route will be expected to cover a certain route segment in, say, six minutes before 6:00 A.M. and, say, eight minutes between 6:00 A.M. and 9:00 A.M. Both PAT and SPRPC felt that the schedules, while not perfect, are very accurate and do represent the real operating conditions. This was, in fact, partly verified by CRA as part of another data collection task.

The procedure followed in estimating transit line-haul times was roughly the same as that for highway line-haul times. First, the route was selected given the origin and destination. The available information was "bridges crossed" and "transfer walk and wait time", which indicated whether a transfer was made at all during a transit trip. Once the route was selected, the line-haul time from the schedules was added up. When a transfer was involved, the times from all routes were added together. As we suspected, the only problem here was that of determining boarding and alighting points. Other than that, no conceptual problems complicated the data collection, although the work was laborious and time-consuming.

6.9. Auto and transit line-haul travel times for shopping trips

Shopping trip line-haul times were basically estimated in the same way as work trips. That is, point-to-point travel times were developed manually using the same source materials and methods as for work trips. However, shopping trips presented some additional problems in calculating line-haul times. First, whereas work trips normally are made during the peak, shopping trips are not. In fact, shoppers generally seek to avoid the peak if at all possible. The result was that our sample yielded only about 10 percent of the shopping trips in which both legs of the trip occurred during the peak hours, about 25 percent of the shopping trips in which one leg of the trip occurred in the peak hours, and about two-thirds of the trips in non-peak hours.

In order to analyze shopping trips, we had to use some mixed-peak and off-peak trips which required adjustment of travel times. For the southern corridor, the assumption was made that the highway times in the suburban areas would change significantly only on those streets serving as arteries to the parkway and tunnels leading downtown. Speed and delay study times were altered on these roads according to the judgment of the researcher. The assignment speeds and rational priority speeds were compared with the peak-hour data; and, both land use and type of road were considered as well. The method used in calculating off-peak auto times in the city was to add up the peak times on the links used in the trip and cut the travel times on arterials by an arbitrary constant percentage. For the CBD, speeds were kept at the peak level as downtown Pittsburgh is highly congested at both peak and off-peak times. No changes were made in speeds on the collector and distributor system. It was assumed that arterial speeds would be 25 percent faster in the off-peak. These gross adjustments were decided upon because the highway system in the city is so much more complex than that of the suburbs. In addition, the travel time variance was dropped for shopping trips because of the lack of any genuine data base for this variable for off-peak times.

For transit trips, off-peak line-haul times were developed by using operating schedules to adjust the basic transit network. Since the transit system in the city is extremely complex, however, it was also necessary to develop rules of thumb for some transit trips. In the more complex areas, we simply assumed that off-peak transit line haul times were 85 percent of the comparable peak hour transit network times.

Finally, for shopping trips, it was necessary to develop travel times for alternative destinations in order to analyze choice of destination. Travel times were also needed for households which made no shopping trips, in order to examine shopping trip frequency. Since the latter is comparable to selecting alternative destinations, the two are discussed together.

The method used to select alternative destinations was to prepare a matrix of shopping origins and destinations for the actual trips on the sample. From this, the different shopping destinations actually used by each residential area in the sample could readily be determined. This matrix was used to specify alternative destinations for both the households actually making trips and the households which did not record shopping trips in the 24-hour sample period. For most households, two to four alternative shopping destinations were identified, in addition to the actual destinations.

6.10. Park and unpark times

Parking data requested in the Home Interview was "type of parking". For surveyed auto trips this information basically identified whether the traveler parked on the street, in a lot, or in a garage. An inventory of available parking spaces was made for downtown Pittsburgh in 1969 by the Parking Authority of Pittsburgh. This inventory gives the following breakdown of parking spaces by type.

Type	Number of facilities	Total capacity	Average capacity	Downtown capacity
Authority garage	8	6,100	763	29.0%
Public garage	10	5,800	580	27.6%
Large public lot (capacity 100+)	27	6,500	241	31.0%
Small public lot (capacity under 100)	54	2,600	48	12.4%
	99	21,000	212	100.0%

On-street spaces account for only 3 percent of capacity in downtown and so were ignored.

CRA collected empirical data on park and unpark times for various types of downtown parking facilities. They are given below.

Parking times in downtown Pittsburgh.

Type	Place	Capacity	Park	Unpark
Garage (Authority)	Sixth and Ft. Duq.	938	3:30	6:25
	Mellon Sq.	1,040	3:00	5:45
	Smithfield and Forbes	867	3:00	3:13
Garage (Public)	Liberty and Sixth	300	2:30	2:50
	Washington Pl. and Fifth	2,200	2:40	3:45
	Forbes and Diamond	325	2:45	2:45
	Liberty and Comm.	750	3:02	2:42
		750	4:50	3:42
		750	3:18	3:55
		750	3:47	
		750	4:06	
Lot (Public)	Grant and Fourth	139	1:35	1:50

Based on these figures, we assumed that all downtown auto trips took three minutes to park and four minutes to unpark.

Since no data were available for other trip destinations, it was necessary to rely on arbitrary assumptions. The assumptions used were as follows.

Type of area	Park (minutes)	Unpark (minutes)
Industrial	2	2
Oakland	2	2
High-density shopping area	2	2
Low-density shopping area	1	1
High-density residential	1	1
Low-density residential	0	0

6.11. Walking times—auto and transit

Walking is the normal access and egress mode for auto trips, but it need not be for transit trips. In fact, transit access and egress has traditionally been one of the most complicated variables to develop. Much of the difficulty has arisen in the past from including different modes like railroad and streetcar in the same transit network. Where access to the transit system by automobile has been prevalent, additional difficulties have been created. Typically, the time to drive to the transit stop has been included as part of access time, which has led to problems in specifying the model or interpreting the result.

The Pittsburgh data does not present these particular problems because only one mode (bus) is predominant. The railroad was not being considered and the streetcar lines are limited. In addition, auto access to bus routes is negligible and therefore we are only allowing for walk access to transit. The few trips which use auto to access transit were deliberately excluded from our sample. Thus, in contrast to some models of mode choice, a very homogeneous transit mode is being evaluated.

The only data source that existed on walking is the number of blocks walked at origin and destination for all types of reported trips from the Home Interview Survey. For alternative trips the number of blocks walked had to be estimated from the base maps by the researcher. By utilizing an average walking speed and an average block size, the walking time was calculated. The SPRPC suggested a speed of 2 mph in downtown Pittsburgh and 3 mph elsewhere in the area. These speeds were based on their own empirical evidence. They are probably a reasonable reflection of conditions, and they agree with typical speeds in other cities. To get data on typical block sizes, a large number of blocks from different areas were actually measured from the highway network base maps, resulting in the following table.

Area type	Walk speed	Block size	Block speed
Downtown Pittsburgh; congested streets and sidewalks; delay at every intersection	2 mph	322 ft/blk	1.83 min/blk
Homestead; high-density residential and commercial suburban district	3 mph	368 ft/blk	1.4 min/blk
Squirrel Hill; medium-density residential and commercial suburban district	3 mph	546 ft/blk	2.0 min/blk

In low-density districts it seems reasonable to assume that "blocks walked" will either be zero or will have been estimated by the trip-maker at no more than about ten blocks to the mile. Unfortunately, the SPRPC Home Interview Survey Manual does not define a block.

The results of these computations were somewhat unusual. Actual auto trips almost always reported zero blocks walked. Either auto drivers, in fact, park very near their residences and find parking places very nearby their final destinations, or they assign zero disutility to walking time. In computing the auto walk time for the alternative trip when the actual trip was transit, the researcher was often unable to find a parking facility at the final destination, so alternative auto trips frequently had positive walk times. Because of this discrepancy, two alternative assumptions were made about auto walk times: one assumption was that all auto trips had zero walk times; the other was that the combination of reported and estimated walk times as described above provided a closer approximation to true auto walk times.

For transit trips, the same discrepancy exists between perceived blocks walked for actual transit trips and estimated blocks walked when transit is the alternative mode, but the disparity is less clear-cut since a substantial number of transit users actually report some number of blocks walked.

6.12. Transit wait and transfer times

The estimation of transfer times was straightforward. For actual transit trips, transfer time is reported on the Home Interview Survey. For alternative transit trips, transfers were determined from the network. Then the transfer time was taken to be half the headway on the subsequent route, following the assumption that transit schedules are not coordinated between routes.

Estimating wait times was complicated and involved a number of alternative assumptions. Wait time is related to the schedule frequency and is usually defined as half the headway up to some maximum. For actual transit trips, wait times are reported in the trip questionnaire.

The basic data sources are estimates of headways from the transit network and the operating schedules, the reported wait times for actual transit trips, and a sample of actual wait times and headways collected by CRA in Pittsburgh in the morning peak for work trips. From this data, reported wait times were plotted against estimated headways for the actual transit trips to see if a discernible relationship could be developed to assist in estimating wait times for the alternative transit trips. The plot resulted in a wide scatter of observations with a number of the reported wait times exceeding the estimated headways and no discernible relationship between the two. In fact, the scatter diagram casts doubt on the accuracy of either the reported wait times, the estimated headways, or both.

Following this, the CRA data were plotted. This is shown as fig. 6.4. This diagram also shows considerable scatter, but waiting times appear to be one-third to one-half the headway. However, the data do not show any obvious maximum wait time.

Based on these comparisons, we decided to develop a range of alternative measures of wait times, based on a variety of different assumptions. One measure assumed that the wait time was the reported wait time for actual transit trips and one-half the headway up to a limit of five minutes (per one-way trip) for alternative transit trips. This is the conventional assumption. Another set of measures based the wait times for both actual and alternative transit trips on the estimated headways. Wait times were assumed alternatively to be one-third and one-half the headway, with no limit and with limits of five, ten, and fifteen minutes (per one-way trip).

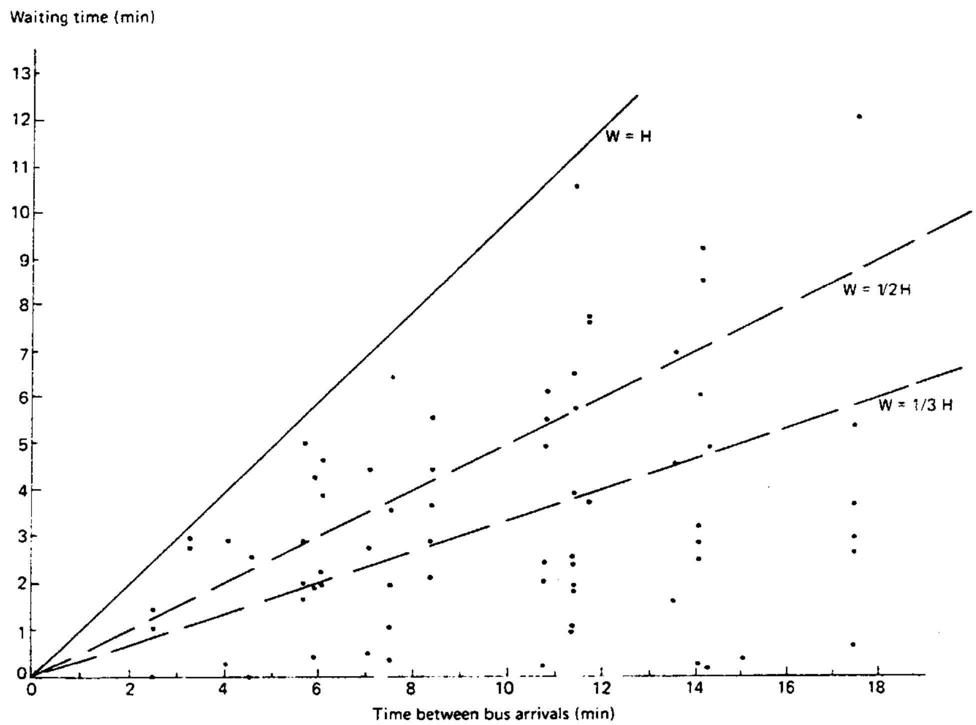

Fig. 6.4. Measured wait times vs. measured headways (Rt. 73, Highland and Stanton, Pittsburgh).

6.13. Travel costs—auto and transit

Mileage figures were given for all links in the highway network in addition to the speed and delay study. The data source was engineering blueprints provided by various public works departments. No reason to doubt these mileage figures existed, and CRA did not attempt to validate them, although it would have been simple to do so.

Gasoline, oil, tires, and maintenance costs for the Pittsburgh area were obtained from a major oil company. From these, operating cost per mile was calculated (three cents per mile in 1967) and multiplied by the mileage of each observation's highway route, the result being an estimate of automobile operating line-haul cost. Capital costs and insurance were not included in these figures. These are typically common costs over the trip purposes and individual trip-makers in the household rather than marginal costs of the given trip. Thus their allocation to a given trip is typically arbitrary. They were accordingly excluded from the analysis.

Auto out-of-pocket costs include those items a trip-maker must pay for during the actual trip—tolls and parking charges for the relevant cost items. The only major toll facility in the Pittsburgh area is the Pennsylvania Turnpike. This was not used by any of the trip-makers in our sample, so out-of-pocket costs are restricted to parking charges.

Parking charges are requested on the Home Interview Survey. For actual auto trips we used the reported data. When auto was the non-selected mode, an average parking charge was calculated, given trip purpose and destination. Two sources of data were available. The first is a tabulation of average cost per hour by downtown zone by two purposes (work and other) compiled from all reported trips in the Home Interview Survey. These data are given below. Note that these data are for downtown zones only.

Parking charges in downtown Pittsburgh calculated from all trip records in 1967 Home Interview Survey, SPRPC.

Destination zone (SPRPC analysis zone)	Long-term purpose = work	Short-term purpose = other
1	$0.078/hr.	$0.122/hr.
2	0.089/hr.	0.133/hr.
3	0.099/hr.	0.207/hr.
4	0.094/hr.	0.281/hr.
5	0.071/hr.	0.210/hr.
6	0.050/hr.	0.093/hr.
7	0.102/hr.	0.250/hr.
8	0.094/hr.	0.219/hr.
9	0.037/hr.	0.194/hr.
10	0.078/hr.	0.308/hr.
968	0.044/hr.	0.152/hr.

The second source is the 1969 Parking Authority Study and contains average rates by type of facility for one-, two-, three-hour periods and an all-day period. Weighted averages of rates by size and type of facility were calculated and are presented below. They agree fairly closely with the SPRPC parking cost estimates.

All day	$1.94
One hour	0.56
Two hours	0.83
Three hours	1.10
Additional hours	0.27

Again, it should be noted that these data apply only to the downtown. However, the survey results overwhelmingly showed free lots as the type of parking facility used by auto travelers at destinations outside the CBD. It was assumed, accordingly, that the same would follow for reported transit users if they were to use auto. So parking charges for all non-downtown destinations were assumed to be zero, unless parking charges were reported for an actual surveyed auto trip.

For downtown trips, the reported charges were used for actual auto trips and the SPRPC hourly cost estimates, applied by downtown destination zone, were used for trips where auto was the alternative mode. The duration of parking was determined from the trip record.

The transit cost is simply the fare on the transit trip. For a reported transit trip, these data were taken from the survey. For actual auto trips, the alternative transit fare was determined from the PAT fare structure on the basis of fare zone boundaries given by the actual route maps.

6.14. Socioeconomic variables

As can be seen from fig. 6.1 and 6.2, the Home Interview Survey provides a fairly broad list of socioeconomic variables. The following were determined to be the most useful to this study:

(1) Number of residents in the household, and number 5 years and older;
(2) Autos per household;
(3) Family gross income—1966;
(4) Number of workers;
(5) Race—white, non-white;
(6) Number of licensed drivers;
(7) Occupation of head of household;
(8) Occupation status of trip-maker and head of household (e.g., employed, unemployed, student, etc.);
(9) Sex of head of household and trip-maker;

(10) Age of head of household and trip-maker;
(11) Employment by industry by traffic analysis zone.

Although most of these variables need no explanation, a few merit some discussion. The family gross income was either reported by the respondent, or when the respondent refused, was estimated by the survey worker. A check of the household print-out showed that about half the incomes in our sample were estimated by the survey worker. Thus, it is likely that the incomes are subject to substantial reporting error.

A number of variables required some transformations to make them applicable to the study. Occupation of head of household was converted to a simple white-collar vs. blue-collar variable. Occupation status was converted to an "employed full time" vs. "other" dummy variable. Age was converted to a 65 and older vs. less than 65 dummy variable.

The auto ownership variable was transformed into a number of different measures of auto ownership in an attempt to provide measures of the number of autos relative to the number of users. Thus, autos per person in the household was computed, as were autos per person 5 years and older, autos per worker, and autos per licensed driver. In addition the following four dummy variables were constructed: $D_1 = 0$ if no auto, 1 otherwise; $D_2 = 0$ if autos < workers, 1 otherwise; $D_3 = 0$ if autos < licensed drivers, 1 otherwise; $D_4 = D_2$ multiplied by D_3.

The first identifies whether or not the family owns at least one auto, the remaining variables are measures of the extent to which the autos must be allocated within the family.

The remaining variables developed for the study measured weather conditions on the travel day. Data were obtained on noontime temperatures and the amount of rainfall by hour of the day. The latter variable was converted to the following dummy variables. For work trips, the variable was 1 if any rain occurred in the morning peak, and zero otherwise. For shopping trips, the variable was 1 if rain occurred by midday, and zero otherwise.