DWELL TIME RELATIONSHIPS FOR URBAN RAIL SYSTEMS

by

Tyh-ming Lin

B.S.C.E. of National Taiwan University (1976)
M.E. of Asian Institute of Technology (1981)

SUBMITTED TO THE DEPARTMENT OF CIVIL ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN TRANSPORTATION

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1990

Signature of Author _____
Department of Civil Engineering
May 18, 1990

Certified by _____
Professor Nigel H. M. Wilson
Thesis Supervisor

Accepted by _____
Ole S. Madsen
Chairman, Departmental Committee on Graduate Studies

DWELL TIME RELATIONSHIPS FOR URBAN RAIL SYSTEMS

by

Tyh-ming Lin

Submitted to the Department of Civil Engineering on May 18, 1990 in partial fulfillment of the requirements for the degree of Master of Science in Transportation.

Abstract

Traditionally, for an urban rail system, the passenger dwell time at each station enroute is often assumed a constant in scheduling and in operations. In fact, however, the dwell time may be affected by the numbers of passengers boarding and alighting, the level of congestion on the platform and on the train, and other unmeasurable factors. Neglecting the dwell time factor in operations, may affect reliability of service, reduce line capacity, and increase passenger wait time.

In this thesis, multiple linear regression models are estimated to deal with the passenger dwell time relationships, in which dwell times may be viewed as a function of the numbers of passengers boarding and alighting, and some alternative variables reflecting crowding. In order to exploit differences in the passenger boarding and alighting process between one and two-car trains, separate models were estimated based on one and two-car train data sets. The data, and the estimated models, relate to the Green Line light rail operation of the Massachusetts Bay Transportation Authority, but the principles and underlying theory apply equally well to heavy rail systems. The MBTA case study demonstrates that linear regression models can explain about 70% of the observed variation of the dwell times by using three explanatory variables; the crowding effect is statistically significant in both one and two-car train models; and, in many model specifications, adding crowding variables reflecting the congestion level on board significantly improves the explanatory power of the models.

Models with nonlinear forms of crowding variables were also estimated to compare with the linear forms. It appears in many cases, the models with nonlinear forms of crowding variables are a significant improvement over those with linear forms. Lastly, checks are made of the key linear regression assumptions to confirm that the promising models do not violate any of these assumptions.

Thesis Supervisor: Professor Nigel H. M. Wilson
Title: Professor of Civil Engineering

Dedication

It would be a prodigious task to name and thank everyone who contributed to this research effort. My special thanks must go to a few, however, for without their guidance and assistance this work would have been impossible.

I would like to greatly appreciate the help and guidance I received from my supervisor, Professor Nigel H. M. Wilson, in preparing this thesis. Through many discussions and suggestion, he encouraged me during the different stages of this work.

I would also like to acknowledge my directors in Institute of Transportation, Dr. Jia-juch Chang and Mr. Chia-hou Huang, who recommended and encouraged me to MIT.

Great thanks are due to my wife Shao-mei whose devotion, sacrifices, and encouragement helped make this thesis possible.

Most of all, I am greatly indebted to my parents for leading me through the different stages of my education. Their sacrifices were uncountable and their help inmeasuable. To them I dedicate this piece of work.

Finally, I would like to acknowledge the National Science Council of Republic of China for the financial support during my study at MIT.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Problem Description

Urban rail and bus systems are the major forms of public transit service in urban areas throughout the world. Generally, urban rail systems are either heavy rail or light rail systems with one of the major difference between them being right-of-way. While heavy rail systems have exclusive rights-of-way, light rail systems may have some shared right of way, and indeed may not have any exclusive right-of-way. In addition, high station platforms are most common in heavy rail systems in contrast with low platforms generally used in light rail systems. Because high platforms make it easier for passengers boarding and alighting, the passenger dwell time per passenger in a heavy rail system will generally be shorter than that in a light rail system.

For an urban rail system, the passenger dwell time at each station enroute is often assumed a constant for scheduling purposes, however, because passenger dwell time may be affected by the number of passengers boarding and alighting, the levels of congestion on the station platform and on the vehicle, and other factors, the passenger dwell time for different trains at stations may vary. Neglecting the variable factors affecting dwell time in real operation, may directly affect reliability of service, and indirectly affect line capacity.

1.2 Motivation

Several studies exist in the literature analyzing the boarding and alighting process of a mass transit vehicle berthed at a station (stop) in an urban area. But, most of these studies were limited to predicting dwell time for buses requiring on board fare payment and were based on counts of total passengers boarding and alighting. Little attention has been given to dwell time for light or heavy rail systems. In most rail systems, passengers make their fare payment at the station, which means they are not required to pay a fare when they board or alight from the vehicle, therefore the boarding and alighting characteristics are expected to be quite different from those of bus.

Generally, an urban rail system operates according to a timetable, but the actual headways, defined as the time interval between two successive trains at a station, depend on several factors and therefore are not typically constant. In real operation, the scheduled headways are set as a function of the passenger demand rates at different time periods, that is, the headways are shorter in the peak periods and are longer in the off-peak periods. The actual headway at a specific station is related to the headway at the preceding station as well as the dwell times of the preceding train at the station, the running times of both trains from the preceding station, and the dwell time of the following train at the preceding station. Assuming the running times for all trains from the preceding station to a specific station are constant under the identical headways, the passenger dwell times at the specific station for different trains may affect the actual headways implying that the headways are no longer identical. Assuming the passenger arrival rates are uniform within a time period, then the variable headways at the specific station are likely to result in uneven passenger loadings for the trains and higher probability of vehicle bunching. Consequently, the reliability of operation and line capacity are dramatically affected, and the passenger wait times are expected to increase.

In view of the importance of the potential impacts of passenger dwell times on scheduling, operations, reliability, line-capacity, and service quality (e.g. reducing average passenger wait times), it is desirable to derive appropriate dwell time functions and apply them in the scheduling process and in operations control. However, in past studies little attention has been given to this topic, and none has focussed on multiple car train dwell times. To fill this research gap and to improve the aforementioned weak points in urban rail operations, this study will be devoted to investigate possible impacts of passenger-vehicle interaction upon passenger dwell time, for both one and two-car trains.

1.3 Literature Review

A major objective of the literature review was to identify other studies that shed light on passenger dwell time functions. Research into the literature found studies that discuss one quantity that will be shown to be related in this study: passenger service times. The passenger service time is defined in the HCM (1985) as the amount of time requirement by each boarding and alighting passenger, in contrast with dwell times which will be governed by boarding and alighting demand, as well as other factors. Dwell times are simply the product of boarding and/or alighting volumes and the service time per passenger. The literature review also found that ordinary least squares regression was widely used by several researchers to estimate dwell time relationships.

Boardman and Kraft (1970) analyzed bus passenger service time data collected in downtown Louisville, Kentucky. These data were analyzed in the following categories: alighting-only; boarding-only; and combined boarding and alighting. Boarding occurred only through the front door, while alighting occurred from the front or rear door or both. Separate linear regression equations were estimated for each category and showed that bus stop passenger dwell times could be predicted accurately using the number of passengers

boarding and/or alighting. Two fare systems were evaluated in Louisville, Kentucky, and the "exact fare" system was shown to be significantly faster when passengers were boarding.

Cundill and Watts (1973) analyzed the stop time characteristics of a variety of different bus types, including both one and two-man-operated systems. Analysis showed that on an urban route with an average of 3 people boarding and 3 people alighting at each stop, the average stop-time of a traditional two-man-operated bus with an open-rear-platform would be about 8 seconds, whereas on the same route the average stop-time of one-man-operated buses would be from 11 to 20 seconds. For some bus types, a significant portion of stop-time was taken up by door operation.

Kraft and Bergen (1974) investigated the effects on passenger service time of various vehicles, different methods of fare collection, combinations of boarding and alighting through the front and rear doors, and time of day. The method of ordinary least squares regression was used to develop equations to predict passenger dwell time based on the number of passengers boarding and alighting. The exact-fare method of fare collection resulted in lower passenger service times than did the traditional cash-and-change method. Trolleybuses with double doors had lower service times than did those with single doors. In addition, intercity passenger service times were found to be greater than those for local transit service.

Kraft (1975) developed the term PVI (Passenger Vehicle Interface), measured in terms of passenger service time, to denote the interaction between passengers and transit system elements while passengers board or alight. Seven factors were labelled as either affecting PVI or being in turn affected by PVI, as shown in Table 1.1.

Table 1.1

FACTORS AFFECTING PASSENGER VEHICLE INTERFACE (PVI)

1. Human
   - type of passenger
   - physical attributes
   - passenger preferences
   - baggage
   - passenger demand

2. Modal
   - type of vehicle
   - service
   - physical characteristics

3. Operating Policies
   - vehicle procurement
   - fare structure

4. Operating Practices
   - type of fare
   - driver practice variations

5. Mobility
   - different system users affect service times, system costs, and travel times

6. Climate/Weather
   - varying composition of system users with varying mobility potentials

7. Other System Elements
   - orderliness of the queue
   - terminal arrangements

Kraft was not able to test the influence of all factors on PVI, but was able to test the presence and impact of PVI on both bus and light rail modes under several service conditions, such as: time of day; whether a particular observed vehicle had only alighting, or boarding passenger, or both boarding and alighting; type of fare collection system; type of passenger; and varying door geometries.

The major methodology employed in Kraft's paper was multivariate regression. Independent variables used to predict total passenger service time included not only those used previously in British studies, counts of boarding and alighting passengers, but also included a cross-term for interactive effects. The latter term, occuring only for mixed processing calibrations of total boarding plus alighting service time, is designed to quantify either or both of the following:

-- the overlap between passengers being serviced in opposite directions simultaneously (a negative coefficient is possible if overlap occurs).

-- the interaction between boarding and alighting passengers at the same doorway area (positive coefficient likely).

In order to develop simulation models to evaluate the street transit operations, distributions of passenger service times through bus doors (the rates at which passengers entered, passed through, and departed from the bus) have been analyzed by photographic studies and simulated by an Erlang function. Kraft and Deutschman (1977) applied these mathematical expressions to simulate the passenger rates of flow entering and departing from a bus and compared them with the observed times, finding that the differences were not significant at 5 percent significance level.

Fritz (1981, 1983) did research on boarding and alighting times of passengers on light rail vehicles based on sampling rush-hour operations on the Presidents' Conference

Committee (PCC) vehicles of the Massachusetts Bay Transportation Authority's (MBTA) Green Line, a high-volume, light rail subway-surface line. Linear regression relations were calibrated between the number of passengers boarding per unit time and concurrent passenger counts (or density) on board the vehicle and on the platform. These models reflected the trends in the raw data that the boarding rates declined markedly under increasing passenger congestion, especially as the space per standee fell below the often used nominal standee space level of $2.7ft^2$/standee and approached crush-capacity density of $1.5ft^2$/standee. On the other hand, at freer circulation levels, those models provided predictions quite similar to predictions from constant-service-time models frequently formulated in earlier research.

Levinson (1983) analyzed bus transit speeds, delays and dwell times based on surveys conducted in a cross section of U.S. cities. He found that bus dwell times (including door opening and closing) were approximatly 5 seconds plus 2.75 times the number of passengers, in addition, both fare-collection policies and door configurations and widths were important determinants of dwell time, especially along high-density routes.

Guenthner and Sinha (1983) examined two causes of bus delay: the delay from the stopping and starting at passenger stops, and the dwell time as the passengers boarded and alighted from the bus. Their evaluation of data on the number of passengers boarding and alighting at stops along a route showed that the negative binomial is a good descriptor of this distribution. Additional data were used to determine dwell time per passenger as a function of the number of passengers boarding and alighting.

The HCM (1985) shows that dwell time may be governed by boarding demand (e.g., when a relatively empty bus arrives at a heavily used stop), or by alighting demand (e.g., at a major transfer point on the system). In all cases, dwell times are proportional to boarding and/or alighting volumes times the service time per passenger. Suggested service times for typical operating conditions -- single door loading, pay on bus-- are: boarding, 2.6 seconds

for single coin; 3.0 seconds for exact fare; 3.5 seconds for exact fare with standees on bus; alighting, 1.7 to 2.0 seconds.

Zografos and Levinson (1986) observed passenger service times for a no-fare bus system and tried to find how the service time per boarding passenger varied with the size of the boarding group and the number of passengers already on the bus. Those relationships were developed for two different occupancy conditions: (a) when the number of passengers on the bus before reaching a stop was less than or equal to the seating capacity of the bus (about 30), and (b) when the number of passengers on board was greater than the seating capacity of the bus (over 30). Simple and multiple regression analyses were performed to examine the effects of bus occupancy and the rank of boarding passengers on the service time per passenger. Both factors were found to influence passenger boarding times when the number of passengers on the bus exceeded the seating capacity, with a service time of more than 2 seconds per passenger. When the number of passengers already on the bus was less than the seating capacity, the service times was approximately 2 seconds per passenger.

Koffman et al. (1984) collected two Boston light rail line data sets (on the Riverside Line and the Boston College Line) on the Green Line service operated by the Massachusetts Bay Transportation Authority (MBTA), and the San Diego Trolley to estimate dwell time model using ordinary least squares regression. In all the models estimated, the dependent variable was dwell time, and total boardings, total deboarding, and total passengers on-board for each car at each stop were used as independent variables. The summary of boarding time statistics used in their regression analysis is shown in Table 1.2. They found that the marginal times for boarding, alighting, and on-board crowding are: 0.67, 0.59, and 0.034 on San Diego Trolley; 0.65, 0.61, and 0.040 on the Riverside Line; and 0.84, 0.52, and 0.029 on Boston College Line. The intercepts for their models are 8.14, 3.04, and 2.96 seconds on San Diego Trolley, the Riverside Line and Boston College Line respectively, and the R^2 values for these models ranged from 0.43 to 0.84.

Table 1.2 Summary of Boarding Time Statistics (a)

	Loading Time (seconds)	Passengers Boarding	Passengers Deboarding	Passengers On-board
San Diego (SSFC)				
Mean	15.7(b)	3.4	3.5	48.3
Std. Dev.	7.7	5.0	4.8	24.5
n = 1,078				
Boston Outbound (Fare Free)				
Mean	10.0(c)	3.1	6.3	46.7
Std. Dev.	6.7	4.5	6.5	27.2
n = 211				
Boston Inbound (Conventional)				
Mean	15.4(d)	4.0	0.8	41.9
Std. Dev.	15.7	4.3	1.5	26.1
n = 550				

Notes:

(a) Excludes data not used in regression analysis due to missing values for any variable. In particular, ends of the line are not included. Also excludes the driver relief point in San Diego.

(b) Time from first door open to last door shut. Data from front and back cars combined.

(c) Maximum of measurements of duration of passenger boarding and deboarding at each of three doors. Excludes time door open but without passenger activity.

(d) Duration of passenger boarding and deboarding activity at single driver-operated door.

1.4 Approach to Be Taken

The approach to be taken in this research includes three major aspects, which are briefly stated below. First, to collect a appropriate data set for model estimation. The data will be used first for preliminary analysis of dwell time relationships based on several key factors and then for model estimation. Second, to develop theory that can explain the dwell time relationships for urban rail systems based on a set of specific assumptions about the relationships between dwell times and the explanatory variable. Third, to integrate the available data and the developed theory to formulate model specifications. It is hoped that the resulting models can be applied for dwell time prediction purposes.

1.5 Thesis Contents

In Chapter 2 a mathematical formulation of the rail dwell time problem is presented. Chapter 3 presents a case study in which multiple linear regression is applied to one and two-car dwell time observations collected on the Masschusetts Bay Transportation Authority Green Line. Chapter 3 also discusses a number of issues that arise in applying multiple linear regression including key assumptions in using any linear regression model and in adopting specific model forms. Chapter 4 summarizes the thesis results and recommends future research needs in this area.

Chapter 2

Mathematical Model

2.1 Introduction

For a transit system, basic operating units and composition of trains are determined by several factors, such as: train sizes, number of driver control sets per car, passenger demand, and so forth. While one to four-car trains are common in light rail systems, heavy rail systems may have four to ten-car trains in operation. Traditionally, there are four to eight doors in any rail car, that is two to four doors in each side, which provides two to eight doors for passenger alighting and boarding at any station, depending whether the station has a single or double platform configuration.

The dwell time of any train at any station may be affected by several factors, including the congestion level on the station platform, type of fare collection, number of passengers alighting and boarding, and passenger crowding level on board. Kraft (1975) developed seven key categories of factors affecting passenger vehicle interface (PVI), as presented in Table 1.1. The key factors that may affect dwell time of urban rail systems are discussed in the following section.

2.2 Key Factors Affecting Dwell Time

The number of passengers alighting and boarding are assumed to affect the dwell time most crucially since it takes time for each alighting and boarding passenger. Apart from these two factors, the following factors may affect dwell time: the arriving passenger load (APL) reflecting the passengers crowding on board; passenger density on station

platforms reflecting the congestion level on station platforms; the number, width, and configuration of doors used and type of station platform reflecting the ease of passengers boarding and alighting; the fare collection system reflecting the time required for each passenger alighting or boarding; the operator behavior; and passenger characteristics.

Assuming the fare collection system, rail car design, and station platform type are identical in an urban rail system, then the variable factors affecting dwell time are reduced, however, there still exist practical difficulties in observing all these remaining factors simultaneously that almost always prevent us from formulating a model specification using all these variables. Thus, the only feasible approach is to collect data on the key factors for further analysis.

Among the stated factors, the number of passengers alighting (OFFS) and boarding (ONS), and arriving passenger load (APL) are assumed most likely to affect the dwell time, furthermore, these quantitative data could be obtained simultaneously at any station by a team of data collectors. In the following sections, underlying dwell time models are expressed in both general and detailed formulations, using these three aforementioned explanatory variables.

2.3 General Formulation

To formulate the dwell time problem, first, the dwell time for each car is defined as the time periods between the first door opening and the last door closing. Thus, the dwell time for an n-car train can be taken as the longest dwell time for cars 1, 2, ...n. Thus:

$$DT_n = \max (DT_1, DT_2, \ldots\ldots, DT_i, \ldots\ldots\ldots, DT_n) \qquad (2.1)$$

where

DT_i: dwell time for the i^{th} car of the n-car train, $1 \leq i \leq n$, and i and n are integers.

This assumes that doors open at the same time on all cars, although this is not always true.

Similarly, the dwell time for any car i is the longest door open time for doors 1, 2, ..., m of that car, that is,

$$DT_i = \max (DOT_1, DOT_2, \ldots\ldots, DOT_j, \ldots, DOT_m) \qquad (2.2)$$

where

DOT_j : the door open time for the j^{th} door.

Based on Eqs. (2.1) and (2.2), the dwell time for an n-car train is taken as the longest door open time for that train.

As stated earlier, for a given rail car design and fare collection system, the number of alighting and boarding passengers, and arriving passenger load are believed to be the main factors affecting the dwell time. Thus the relationships between these variables can be expressed as the general form:

$$DT = f (ONS, OFFS, APL) \qquad (2.3)$$

where

ONS: the number of boarding passengers.

OFFS: the number of alighting passengers.

APL: arriving passenger load.

Several alternative forms can be derived, according to different definitions of the

number of alighting and boarding passengers, and arriving passenger load. For example: treating the train as an entity, the relationships between these variables is:

$$DT = f\left(\sum_{i=1}^{n} ONS_i, \sum_{i=1}^{n} OFFS_i, APL_{train}\right) \qquad (2.4)$$

Similarly, dwell time for a single car can be defined based on the number of passengers alighting and boarding, and passenger crowding level of that car. Then, applying equation (2.1), the train dwell time will be based on the single car with the largest dwell time (LDT), and the relationship between these variables can be expressed as:

$$DT = f(ONS_{LDT}, OFFS_{LDT}, APL_{LDT}) \qquad (2.5)$$

Similarly, dwell time for a single door can be defined based on the number of passengers alighting and boarding, and passenger crowding on that car. Then, applying equation (2.2), the train dwell time will be based on the door with the longest door opening time (LDOT), and the relationship between these variables can be expressed as:

$$DT = f(ONS_{LDOT}, OFFS_{LDOT}, APL_{LDOT}) \qquad (2.6)$$

These three approaches all related to the dwell time but require different data both for estimation and application. Therefore, the core problem of applying these model forms is the data available and the accuracy of the resulting model estimation. For example: the APL_{LDOT} is difficult to measure because of the difficulty of identifying the APL close to the door with the LDOT; and there may be no passenger alighting and/or boarding through the single door with the LDOT resulting in lacking of accuracy of model estimation. Finally the problems of using such a model for forecasting even if it could be estimated, appear in surmountable.

2.4 Detailed Formulation

Based on the preceding general formulation, Eq. (2.6) can be developed at a detailed level by deriving each of the independent variables.

The passengers wait to board at any station r (ONS_r) can be expressed as:

$$\begin{aligned} ONS_r &= \sum_{i=1}^{n}\sum_{j=1}^{m} ONS_{rij} \\ &= f(H_{rs}) \\ &= \sum_{s=r+1}^{t}\sum_{i=1}^{n}\sum_{j=1}^{m} R_r \times H_{rs} \times P_{rs} \times P_{rsij} \end{aligned} \qquad (2.7)$$

where

ONS_r : passengers wait to board at station r.

ONS_{rij} : passengers boarding through the j^{th} door of the i^{th} car of that train.

H_{rs} : the previous headways (minutes) of trains operating between stations r and s.

R_r : passenger arrival rate per minute at station r.

P_{rs} : percentage of all passengers boarding at station r who travel to station s.

P_{rsij} : percentage of passengers travelling from r to s who board through the j^{th} door of the i^{th} car.

Similarly, passengers alighting at station r ($OFFS_r$) can be expressed as follows:

$$
\begin{aligned}
OFFS_r &= \sum_{i=1}^{n}\sum_{j=1}^{m} OFFS_{rij} \\
&= f(APL,\ P_r) \\
&= f(f(H_p),\ P_r) \\
&= \sum_{p=l}^{r-1}\sum_{i=1}^{n}\sum_{j=1}^{m} ONS_{pij} \times P_{prij} \\
&= \sum_{p=1}^{r-1}\sum_{i=1}^{n}\sum_{j=1}^{m} R_p \times H_{ps} \times P_{ps} \times P_{psij} \times P_{prij} \qquad (2.8)
\end{aligned}
$$

where

$OFFS_{rij}$: passengers alighting from the j^{th} door of the i^{th} car of at station r.

APL : arriving passenger load.

ONS_{pij} : passengers boarding from the j^{th} door of the i^{th} car at previous stations p.

P_r : percentage of on board passengers alighting at station r.

H_p : the previous headway (minutes) of train at preceding station p.

P_{prij} : percentage of passengers boarding from the j^{th} door of the i^{th} car at previous station p, and alighting at station r.

R_p : passenger arrival rate per minute at station p.

H_{ps} : the previous headway (minutes) of train operating between stations p and s.

P_{ps} : percentage of all passengers boarding at station p who travel to station s.

P_{psij} : percentage of passengers travelling from p to s who board through the j^{th} door of the i^{th} car.

Substituting Eqs. (2.7) and (2.8) into Eq. (2.6), then

$$DOT_{rji} = f(ONS_{rji}, OFFS_{rji}, APL)$$

$$= f\Big(\sum_{s=r+1}^{t}\sum_{i=1}^{n}\sum_{j=1}^{m} R_r \times H_{rs} \times P_{rs} \times P_{rsij},$$

$$\sum_{p=l}^{r-1}\sum_{i=1}^{n}\sum_{j=1}^{m} R_p \times H_{ps} P_{ps} \times P_{psij} \times P_{prij},$$

$$\sum_{p=l}^{r-1} ONS_{pij} - \sum_{p=l+1}^{r-1} OFFS_{pij}\Big) \qquad (2.9)$$

where

ONS_{pij} : passengers boarding through the j^{th} door of the i^{th} car at station p.

$OFFS_{pij}$: passengers alighting from the j^{th} door of the i^{th} car at station p.

As indicated by the above expressions, it is more difficult to utilize the detailed formulation than the general formulation since a unreasonable amount of data would be required for dwell time analysis based on the detailed formulation. Therefore, it appears that the general formulation of dwell time relationships with several explanatory variables is more feasible.

The preceding knowledge is reflected in a set of specific assumptions about the relationships between the dwell times and the explanatory variables. To explain or predict the dwell time using these variables, multiple linear regression method is applied (multiple linear regression is discussed in detail in Greene (1990), Johnston (1984) and Never et al. (1983)).

Chapter 3

Model Estimation -- MBTA Green Line Case

In this chapter dwell time functions are estimated for one and two-car trains on the MBTA Green Line, a light rail line operating with articulated light rail vehicles. Because of the heavy ridership, close station spacing, and high frequencies of trains, the dwell time function is particularly important in determining both Green Line capacity and operating performance, and is a critical element in developing both the operations plan and operations control options. Moreover, no dwell time function has been developed for the Green Line, or for any similar high frequency and high ridership light rail system using articulated vehicles. The analysis described in this chapter focusses on estimating dwell time functions, based on a set of disaggregate dwell time observations.

3.1 MBTA Rail Transit Network

The MBTA rail transit network, excluding commuter rail, consists of four lines, the Red, Orange, Blue, and Green, with a total length of 69 miles and 126 stations of which 4 are central transfer stations (CTS) between the lines (see Figures 3.1 and 3.2).

Figure 3.1 **MBTA Subway System**

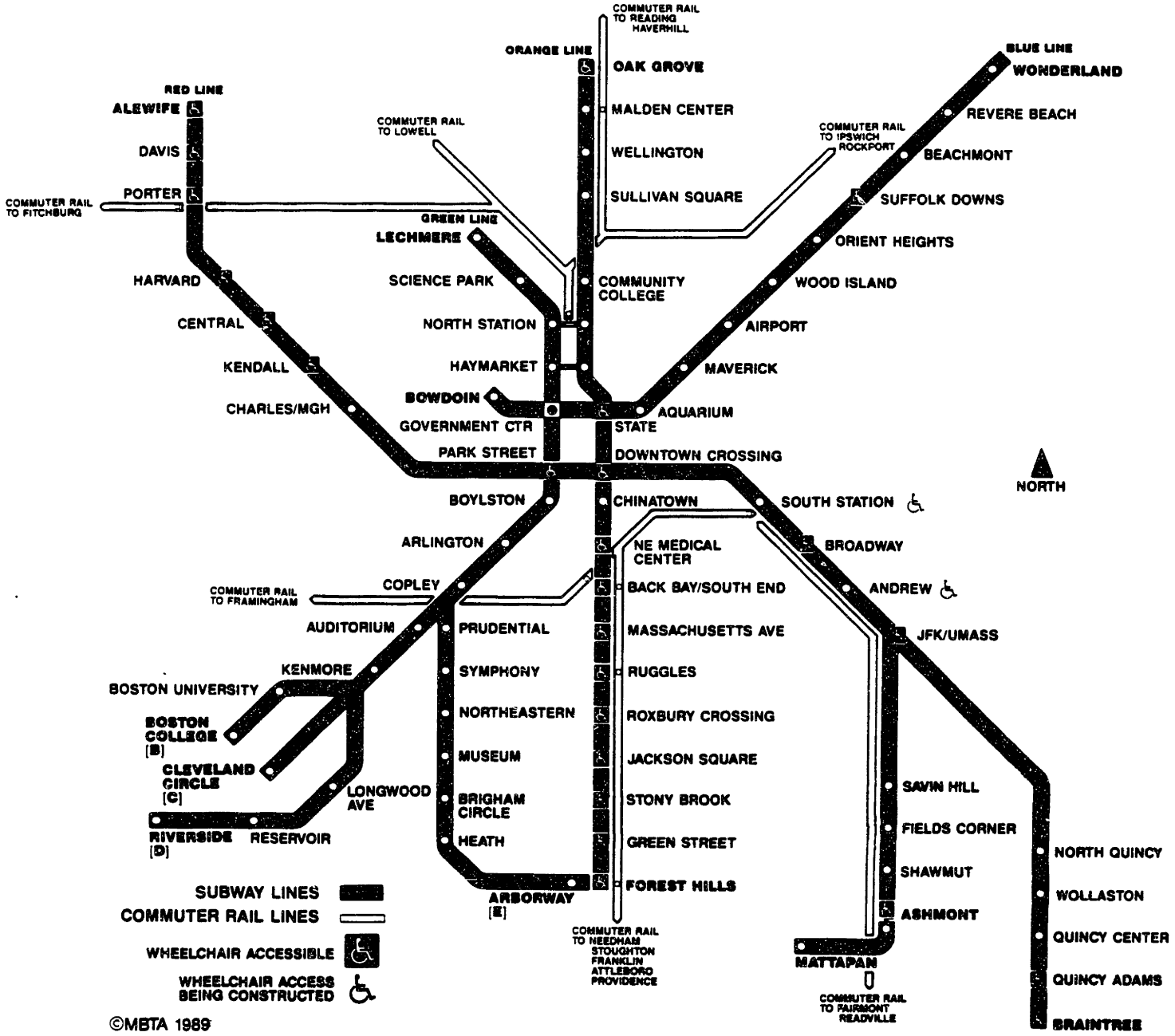
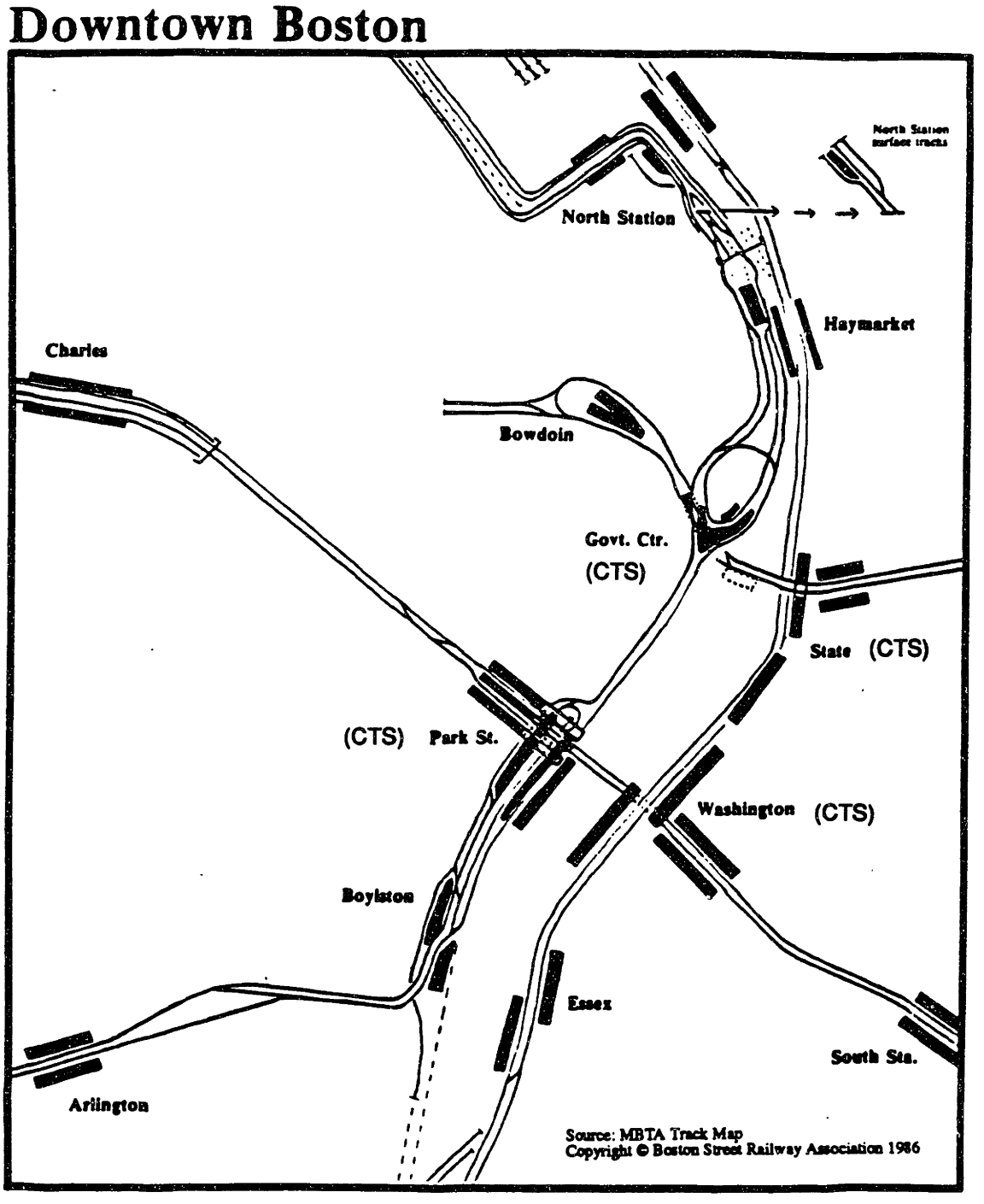


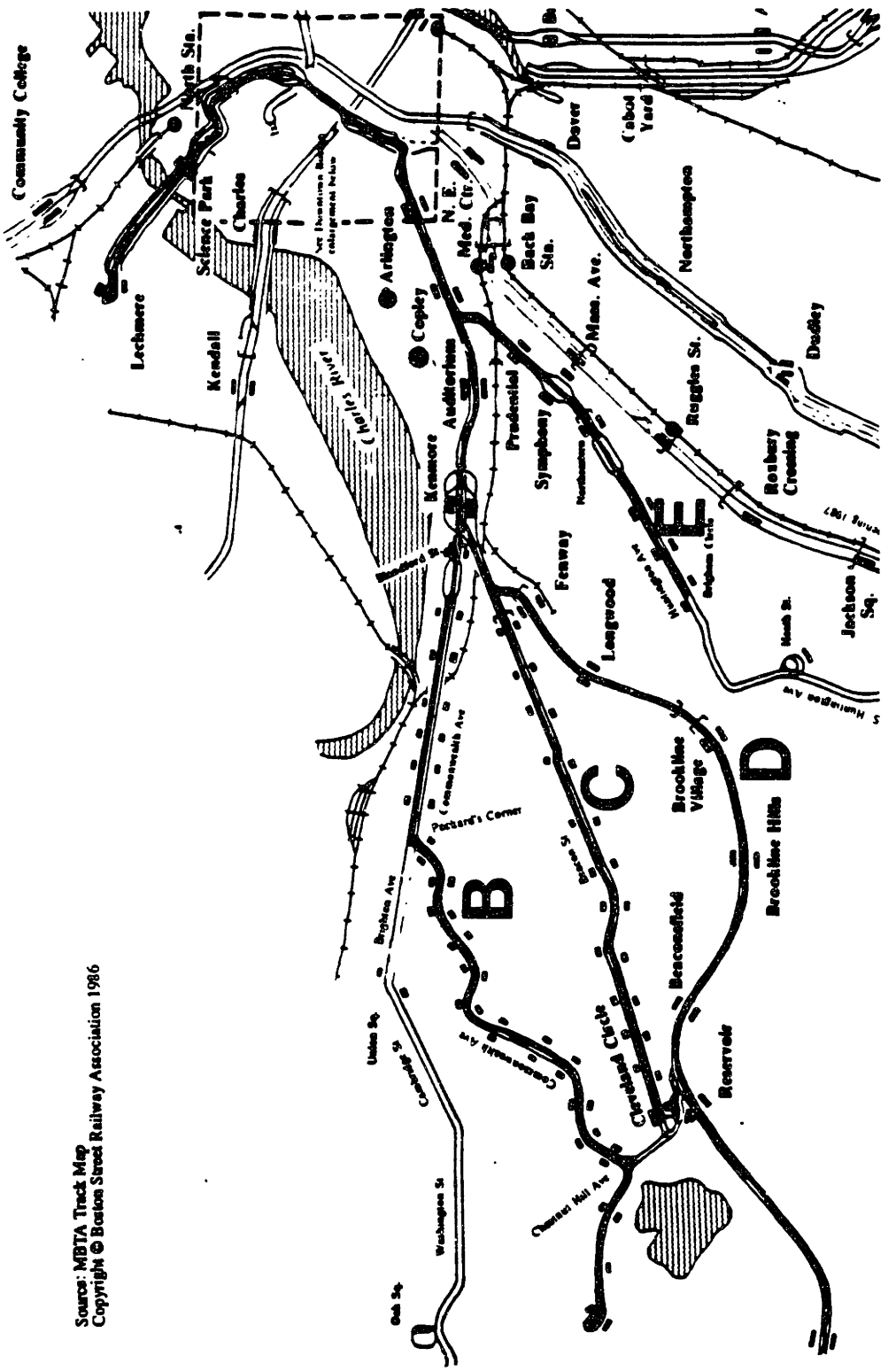
Figure 3.2 Downtown Subway System



Two different rail technologies co-exist in the MBTA rail transit system. The Red, Orange and Blue Lines are Rail Rapid Transit lines, with a total length of 41 miles and 60 stations, using conventional steel wheel on steel rail technology operating on an exclusive right-of-way with third rail power pickup, while the Green Line uses Light Rail technology over a branching network of 28 miles and 70 stations. Although portions of the Green Line right-of-way include at-grade crossings, most of the line is fully grade separated, including the central portion which operates in a tunnel. In fact the original section of Boston's Light Rail system operates in the nation's oldest subway (the first two stations opened in 1897). Today, the Green Line is made up of four routes (see Figure 3.3), B, C, D, and E, which join in one central subway tunnel (from Lechmere Station to Kenmore Station) with trains from all four routes operating on the same tracks. Within this section, fares are paid upon entering a station rather than on board the train, which is the rule on the surface branches of the line.

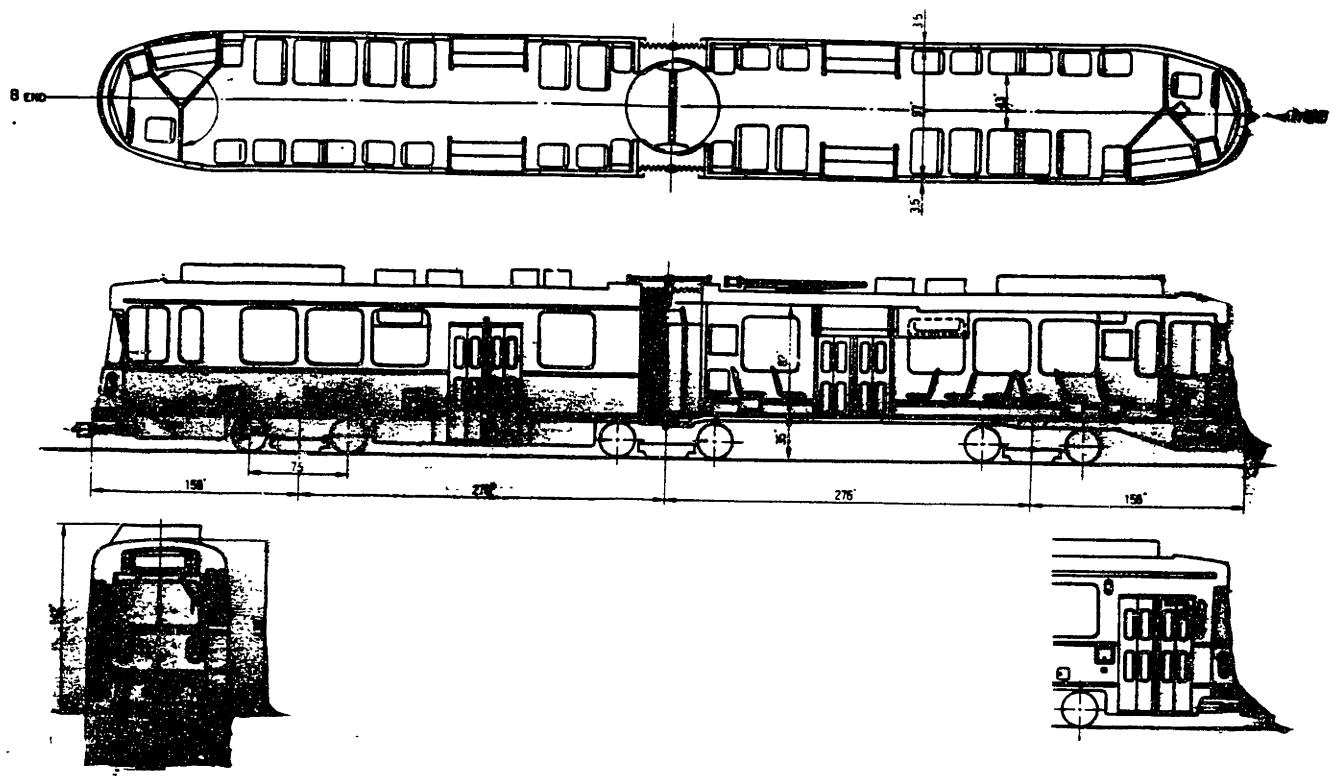
In the 1970's, The President's Conference Committee (PCC) cars were the major vehicles to run on the Green Line; they have given way to the new Light Rail Vehicles (LRVs) (see Figure 3.4) today. All LRVs of the Green Line have 52 seats, 35-inch high non-slip floors, highlighted stair edges, hand/grab rails, and priority seating decals which encourage passengers to offer a seat to disabled and elderly persons. There are 6 doors per car, 3 in each side, the middle and rear doors are 35-inch wide while the front door is 32-inch wide. The great majority of trains are composed of either one or two cars, depending on passenger demands in the peak and off-peak periods. However, some three-car trains are also in daily operation. Therefore, there are three doors available for passengers alighting and boarding in any one-car train while it dwells at any single (low) platform station, while there are six doors available in any two-car train. In addition, these LRVs have a public address systems over which destinations and stops are announced.

Figure 3.3 Green Line Subway and Branch Lines



Source: MBTA Track Map
Copyright © Boston Street Railway Association 1986

Figure 3.4 General Arrangement of the LRV



3.2 Data Collection

As discussed earlier in Chapters 1 and 2, there are many factors which might affect dwell times, including the number of passengers boarding and alighting, congestion on the station platform and in vehicles, passenger and operator behavior and vehicle design. Among these factors, the number of passengers boarding and alighting are expected to be the most important determinants of dwell time, and these data can be obtained by direct observation. For this analysis, a special detailed data set was gathered, with each observation including the following data: the number of passengers boarding and alighting through each door, the time the front door was opened and closed for each car, and the departing passenger load for each car. Because of the unusual level of detail required, it was necessary to have a two person team per car, or a four person team for a two car train, to collect the data.

One-car train data was collected in the westbound direction at the Copley and Arlington Stations by teams of two data collectors. Each collector had responsibility for one half of the train. Recordings were made of the number of passenger boarding and alighting through each door as well as of the time the front door was opened and closed. In addition, the departing passenger load was estimated by each collector for his half of the train. Lastly, each train was identified according to its route and the number of its first car.

Two-car train data similar to the one-car train data was collected in the westbound direction at Arlington Station by a team of four data collectors. Once again each collector had responsibility for one half of a car.

These data was collected in April of 1988 and 1989, and resulted in 122 observations of one-car train dwell times and 51 samples of two-car train dwell times.

3.3 Preliminary Analysis

The purpose of the preliminary analysis was to examine whether the dwell time is related to the number of passengers boarding and alighting as well as to the leaving passenger load. The dwell time for the two-car train data set was taken as the longer of the dwell times for each car (LDT), as described in Chapter 2.

Because both the number of passengers boarding and alighting and passenger load may be expected to affect the dwell time, two preliminary analysis were conducted based on the sum of boardings and alightings (ONOFFS) and on leaving passenger load (LPL), both defined on a per car basis.

3.3.1 One-Car Train Observations

The one-car train data set was collected from two stations (Copley and Arlington). Because the main purpose of the study is to determine if dwell time is a function of passenger load and the numbers of passengers alighting and boarding, a dummy variable was used in the regression analysis to include the station variable in model formulations. As will be shown later it appears appropriate to perform further analysis based on the pooled data. The pooled data permitted model estimation with the largest possible sample size, to produce the best possible models.

There are 52 seats in each car and it was expected that increasing the numbers of standees in a car would likely result in increased dwell time due to interference effects between standees and passengers attempting to board and alight. Accordingly, observations were classified into four groups for the preliminary load analysis, based on the following ranges for LPL: LPL<53, 53≤LPL<81, 81≤LPL<109, and LPL≥109. The larger the LPL, the more significant the crowding impact was expected to be.

From Table 3.1, it can be seen that the mean dwell times are 16.8, 20.6, 24.0 and 36.0 seconds for the four groups implying that the mean dwell time is positively related to the LPL. The standard deviations of the dwell times for these four groups are 5.65, 8.35, 6.68, and 13.31 seconds respectively implying the variability of the dwell time also increases with LPL. Analysis of variance is applied to test the null hypothesis that the mean dwell time is equal for these four groups. As indicated by the F-statistics and P-values shown in Table 3.2, the null hypothesis is rejected at 0.05 significance level. Therefore, the mean dwell time appears significantly related to LPL.

Similarly, it was expected that increasing the number of passengers boarding and alighting would result in increased dwell time due to the fact that each passenger boarding and alighting affects the dwell time. Accordingly, observations were classified into four groups for the boarding and alighting analysis, based on the following ranges for ONOFFS: ONOFFS$\leq$9, 9<ONOFFS$\leq$17, 17<ONOFFS$\leq$25, and ONOFFS>25.

As indicated by Table 3.1, the mean dwell times are 15.8, 20.0, 27.1, and 41.0 seconds, with standard deviations of 6.65, 6.32, 5.90, and 14.96 seconds for the four groups implying that the mean dwell time increases with the ONOFFS count. It also appears that the variability of the dwell time increases significantly when ONOFFS is greater than 25. Once again, analysis of variance is used to test the statistical significance of the mean dwell time between the four groups. The null hypothesis is that the mean dwell time is no different between the four groups. As indicated by the F-statistics and P-values shown in Table 3.3, the null hypothesis is rejected, which implies, as expected, that the mean dwell time appears positively related to the ONOFFS variable.

In conclusion, the mean dwell time is positively related to both LPL and ONOFFS, for the one-car train observations.

Table 3.1 One-Car Train Dwell Times

Total Sample:	n = 122	Mean = 23.31		Standard Deviation = 11.41	
a) Analysis based on LPL					
LPL		< 53	53-80	81-108	> 108
Sample Size		41	37	16	28
Mean LPL		32	65	94	132
Mean ONOFFS		10	15	20	21
Mean (Dwell Time)		16.83	20.60	24.00	36.00
Std. Dev. (Dwell Time)		5.65	8.35	6.68	13.31
b) Analysis based on ONOFFS					
ONOFFS		< 10	10-17	18-25	>25
Sample Size		37	39	30	16
Mean LPL		47	75	89	101
Mean ONOFFS		6	13	21	32
Mean (Dwell Time)		15.81	20.03	27.10	41.56
Std. Dev. (Dwell Time)		6.65	6.32	5.90	14.98

Table 3.2 F Test for Mean Dwell Time on One-Car Train (Based on LPL)

SOURCE	DF	SS	MS	F	P
GROUP	3	6511.4	2170.5	27.73	0.000
ERROR	118	9234.7	78.3		
TOTAL	121	15746.2			

Note: Critical F Value = F(0.05,3,118) = 2.68

Table 3.3 F Test for Mean Dwell Time on One-Car Train (Based on ONOFFS)

SOURCE	DF	SS	MS	F	P
GROUP	3	8262.9	2754.3	43.43	0.000
ERROR	118	7483.3	63.4		
TOTAL	121	15746.2			

Note: Critical F Value = F(0.05,3,118) = 2.68

3.3.2 Two-Car Train Observations

A similar preliminary analysis was conducted for the set of two-car train dwell times based on the ONOFFS and LPL variables. The values of the ONOFFS and LPL variables were taken as the ONOFFS and LPL for the car with the larger dwell time.

First, the observations were classified into the same four groups for the load analysis as in section 3.3.1. Table 3.4 presents the mean dwell times of 20.4, 23.2, 27.5, and 35.5 seconds with standard deviations of 5.68, 7.39, 6.81, and 6.31 seconds for the four groups. It appears that the mean dwell time increases as the mean LPL increases but the variability of the dwell time does not differ significantly. Analysis of variance is applied to test the null hypothesis that the mean dwell time is equal for these four groups. As indicated by the statistics shown in Table 3.5, the null hypothesis is rejected at 0.05 significance level implying that the mean dwell time is positively related to the LPL.

Second, the observations were analysed based on the ONOFFS variable. Table 3.4 presents the mean dwell times of 19.3, 22.8, 28.7, and 34.9 seconds with standard deviations of 5.69, 4.87, 6.81, and 6.56 seconds respectively for the four groups. It appears that, as expected, the mean dwell time increases with ONOFFS and the variability of dwell time does not differ significantly between groups. To test the statistical significance of the mean dwell time between groups, analysis of variance is applied. As indicated by the F statistic and P-value shown in Table 3.6, the null hypothesis is rejected at 0.05 level, which implies that the mean dwell time is positively related to ONOFFS.

Table 3.4 Two-Car Train Dwell Times

Total Sample:	n = 51	Mean = 26.57		Standard Deviation = 8.40	
a) Analysis based on LPL for LDT car					
LPL		< 53	53-80	81-108	> 108
Sample Size		11	13	16	11
Mean LPL		41	69	98	132
Mean ONOFFS		11	15	21	27
Mean (Dwell Time)		20.36	23.15	27.50	35.46
Std. Dev. (Dwell Time)		5.68	7.39	6.81	6.31
b) Analysis based on ONOFFS for LDT car					
ONOFFS		< 10	10-17	18-25	>25
Sample Size		12	14	11	14
Mean LPL		61	74	97	109
Mean ONOFFS		6	14	21	32
Mean (Dwell Time)		19.33	22.79	28.73	34.87
Std. Dev. (Dwell Time)		5.69	4.87	6.81	6.56

Table 3.5 F Test for Mean Dwell Time on Two-Car Train (Based on LPL)

SOURCE	DF	SS	MS	F	P
GROUP	3	1457.5	485.8	11.02	0.000
ERROR	47	2073.0	44.1		
TOTAL	50	3530.5			

Note: Critical F Value = F(0.05,3,47) = 2.80

Table 3.6 F Test for Mean Dwell Time on Two-Car Train (Based on ONOFFS)

SOURCE	DF	SS	MS	F	P
GROUP	3	1841.6	613.9	17.08	0.000
ERROR	47	1688.9	35.9		
TOTAL	50	3530.5			

Note: Critical F Value = F(0.05,3,47) = 2.80

3.3.3 Comparisons between One-Car and Two-Car Train Observations

From Tables 3.1 and 3.4, as indicated by the statistics shown based on the LPL, it can be seen that in the three less-crowded groups, the mean dwell time for the two-car train data set, with 20.4, 23.2, and 27.5 seconds respectively, appears larger than that for the corresponding one-car train data set, 16.8, 20.6, and 24.0 seconds. For the heaviest passenger load group with LPL$\geq$109, the mean dwell time for two data sets was virtually the same at about 36.0 seconds. The t-test is applied to test the statistical significance of the mean dwell time difference between the two data sets for the same LPL groups. The null hypothesis is that the mean dwell time is equal between the two data sets for each group with the same range. Table 3.7 shows that the t-statistics are less than the critical t-values, therefore it is concluded that at 0.05 level there is insufficient evidence to reject the null hypothesis from these observations.

Similarly, as indicated by the statistics based on ONOFFS (see Tables 3.1 and 3.4), the mean dwell times are 15.8, 20.0, 27.1, and 41.6 seconds for the one-car train data set versus 19.3, 22.8, 28.7, and 34.9 seconds for the four groups of the two car data set. The dwell times appear longer for the two-car train data set than for the one-car train data set for the lowest three ONOFFS groups. Once again, the t-test is applied to test the null hypothesis that the mean dwell times are equal for the two sets of data with the same ranges. The t-statistics in Table 3.8 show that the test is insignificant at 0.05 level implying there is insufficient evidence to reject the null hypothesis in all cases.

Table 3.7 t test for Mean Dwell Time between One-Car and Two-Car Trains (Based on LPL)

Group	DF	t	Critical t-value
LPL<53	50	1.84	2.01
53<=LPL<81	48	0.97	2.01
81<=LPL<109	30	1.47	2.04
LPL>=109	37	-0.13	2.03

Table 3.8 t test for Mean Dwell Time between One-Car and Two-Car Trains (Based on ONOFFS)

Group	DF	t	Critical t-value
ONOFFS<=9	47	1.65	2.01
9<ONOFFS<=17	51	1.48	2.01
17<ONOFFS<=25	39	0.75	2.02
ONOFFS>25	28	1.32	2.05

3.3.4 Summary

From the preliminary analysis, the following conclusions can be drawn:

(1) Mean dwell time is positively related to both LPL and ONOFFS for the two sets of data for one-car and two-car train observations at the 0.05 level.

(2) There is no significant difference in the mean dwell times between the one-car and two-car data sets for groups with the same ranges of LPL and ONOFFS at the 0.05 level.

3.4 Model Variables and Estimation Procedure

Based on the preliminary analysis, two major factors, the number of passengers boarding and alighting, and the effect of crowding on board, were expected to enter into the dwell time function. However each factor can be represented in different forms and may interact in different ways. Accordingly a series of linear regression models of passenger processing were estimated to identify the strongest functional form. In the following discussion of the estimation results, the significant variables that were used to explain the variation in the dependent variable DT (dwell time measured in seconds) are as follows:

ONS: number of passengers boarding per car

OFFS: number of passengers alighting per car

ONOFFS: sum of ONS and OFFS

AS: number of arriving standees per car

LS: number of departing standees per car

ABAS: product of ONOFFS and AS, i.e. ONOFFS*AS

ABLS: product of ONOFFS and LS, i.e. ONOFFS*LS

MAXASLS: maximum of ABAS and ABLS, i.e. MAX(ABAS, ABLS)

OFFAS: product of OFFS and AS, i.e. OFFS*AS

ONLS: product of ONS and LS, i.e. ONS*LS

SUMASLS: sum of OFFAS and ONLS

As discussed in Chapter 2, the dwell time for a one-car train is the maximum door open time for doors 1, 2 and 3 of that car, i.e.

DT = max {boarding/alighting time through door 1,
boarding/alighting time through door 2,
boarding/alighting time through door 3}

Similarly, the dwell time for a two-car train is the longer of the dwell times for car 1 and car 2 of that train, i.e.

DT = max {dwell time for car 1, dwell time for car 2}

Because the number of passengers boarding and alighting for each car of a two-car train will not generally be the same, and because the processes are different as shown above, separate models were estimated for the one-car train data set and the two-car train data set.

The statistical packages SST (1986) and MINITAB (1989) were used for the regression analysis, as well as for checking the results. The resulting models shown below

include (t-statistic), [p-value] (defined as the significance level at which the hypothesis test procedure changes conclusions, that is, the level at which the test becomes significant), corrected coefficient of determination (R^2), and Durbin-Watson statistic (DW) (this statistic ranges from zero to four, with a value near zero indicating strong positive autocorrelation and a value close to four meaning that there is a significant negative autocorrelation. A value near two indicates that there is very little autocorrelation -- the ideal situation). The t-statistic and p-value are used to determine the contribution of each variable used in model estimation; the corrected R^2 is used to measure how well the model estimation fits the sample data; and the Durbin-Watson statistic is used to test for autocorrelation in the regression residuals. An example with these statistics is shown below:

```
DT=9.07+1.15*ONS+0.63*OFFS,  (R2=0.48)
  (5.96)  (8.46)     (5.58)    (DW=1.69)
  [0.00]  [0.00]     [0.00]
```

All coefficients are strongly significant as indicated by the t-statistics and p-values. The t-statistics for these variables are 5.96, 8.46, and 5.58 respectively, while p-values are all approximately zero. The corrected coefficient of determination (R^2) for the model is 0.48, while the Durbin-Watson statistic (DW) is 1.69. The DW statistic is used later to check the assumption of independent residuals for the most promising models.

3.5 One-Car Train Models

Since the one-car train data set was collected from two stations, a dummy variable is introduced in the regression analysis below to include the qualitative station variable. If the coefficient of the dummy variable is not statistically significant, then it is possible to omit it to produce the best possible models.

The models were estimated based on three approaches: all data together, the data set with ONS being equal to, or greater than, OFFS (ONS≥ OFFS) and that with OFFS being greater than, ONS (OFFS>ONS). The available sample points for these three approaches are 122, 83, and 39 respectively. In the following analysis, model estimations are conducted based on these three approaches, and named One_M1, One_M1a, and One_M1b models in the subsequent sections.

3.5.1 Correlation Analysis

Appendix A presents the relationships between the dwell time and the key potential independent variables, all of which appear to show some correlation.

The correlation coefficients (r) between all variables used are shown in Appendix B. It is clear from Tables B.1 to B.3 that except for the dummy variable, all other variables are highly correlated with the dependent variable dwell time (DT) as well as with some other independent variables. In any multivariate linear regression estimation, it is desirable to avoid strong correlations between independent variables. The highest correlations among the independent variables are between those reflecting crowding, suggesting that these variables should not be included together in any model formulation.

3.5.2 Model A: DT = f(ONS, OFFS)

This model assumes that only the numbers of passengers boarding and alighting affect the dwell time, so there is no effect of passenger crowding on board. Since the one-car train data set was collected from two stations, a dummy variable is introduced in the model estimation below to include the qualitative station variable, one is set for the observations from Copley Station and zero for the observations from Arlington Station.

The resulting models based on the three approaches discussed in section 3.5 are expressed by 1, 1a, and 1b respectively in the specifications below:

Model A1P:

$$DT = 9.33 + 1.14*ONS + 0.64*OFFS - 0.44*DUMMY, \quad (R^2 = 0.48)$$
$$(5.22) \quad (8.01) \quad (5.48) \quad (-0.28) \quad (DW = 1.69)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.78]$$

Model A1aP:

$$DT = 8.61 + 0.90*ONS + 1.41*OFFS + 0.10*DUMMY, \quad (R^2 = 0.51)$$
$$(3.46) \quad (3.97) \quad (5.24) \quad (0.05) \quad (DW = 1.52)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.96]$$

Model A1bP:

$$DT = 12.67 + 0.81*ONS + 0.47*OFFS - 1.27*DUMMY, \quad (R^2 = 0.64)$$
$$(7.26) \quad (3.73) \quad (3.69) \quad (-0.68) \quad (DW = 1.97)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.50]$$

The coefficients for the variable DUMMY are insignificant in all three models, as indicated by the t-statistics and p-values, implying there is no significant difference between the two sets of data due to station-specific factors. Therefore, the variable DUMMY is dropped from these (and subsequent) models, producing the following models:

Model A1:

$$DT = 9.07 + 1.15*ONS + 0.63*OFFS, \quad (R^2 = 0.48)$$
$$(5.96) \quad (8.46) \quad (5.58) \quad (DW = 1.69)$$
$$[0.00] \quad [0.00] \quad [0.00]$$

Model A1a:

$$DT = 8.67 + 0.90*ONS + 1.41*OFFS, \quad (R^2 = 0.52)$$
$$(3.91) \quad (4.03) \quad (5.28) \quad (DW = 1.52)$$
$$[0.00] \quad [0.00] \quad [0.00]$$

Model A1b:

$DT=11.98+0.88*ONS+0.43*OFFS$, $(R^2=0.64)$
(8.51) (4.61) (3.82) (DW=1.90)
[0.00] [0.00] [0.00]

The coefficients of the variable ONS is about twice that of variable OFFS in models A1 and A1b implying that the marginal dwell time for boarding is almost twice that for alighting, while that for boarding is about 0.6 times that for alighting in model A1a. All coefficients are very strongly significant in all three models, as indicated by their t-statistics and p-values, with coefficients of determination (corrected R^2) of 0.48, 0.52, and 0.64 in model A1, A1a, and A1b respectively.

It appears from these results, that models A1a and A1b using two data sets based on the relative magnitude of ONS and OFFS are a significant improvement over A1.

3.5.3 Model B: DT = f(ONS, OFFS, ONOFFS * Standees)

This model recognizes the effect on dwell time of crowding on board as well as the number of passengers boarding and alighting. The crowding effect was examined using three alternative variables: MAXASLS, ABLS, and ABAS, assuming movement of boarding and alighting passengers would be affected by arriving standees and departing standees respectively. The resulting models are:

Model B1:

B1-1 $DT=12.59+0.55*ONS+0.22*OFFS+0.0076*MAXASLS$, $(R^2=0.62)$
(8.87) (3.66) (1.89) (6.49) (DW=2.05)
[0.00] [0.00] [0.06] [0.00]

B1-2 $DT=12.34+0.67*ONS+0.20*OFFS+0.0072*ABAS$, $(R^2=0.61)$
(8.64) (4.76) (1.65) (6.20) (DW=1.99)
[0.00] [0.00] [0.10] [0.00]

```
B1-3 DT=12.28+0.53*ONS+0.33*OFFS+0.0075*ABLS,  (R2=0.62)
       (8.89)  (3.57)    (3.10)       (6.76)      (DW=2.07)
       [0.00]  [0.00]    [0.00]       [0.00]
```

It is clear that adding any of these three variables to reflect the effect of crowding on board significantly improves the explanatory power of the One_M1 model. There also appears to be little to choose between the MAXASLS, ABAS and ABLS forms, although the ABLS form is slightly preferred since all coefficients are significant at 0.05 level. This implies that the on board crowding term is more related to departing passenger load than arriving passenger load.

Compared to model A1, the constant for the model B1 is larger, with lower marginal dwell times for boarding and alighting. The coefficient of the variables ONS and OFFS is 0.55 and 0.22 respectively in model B1-1, which is about 70 times and 30 times that of the variable MAXASLS implying that the marginal dwell time for boarding is about 2.5 times that for alighting, as well as 70 times that for MAXASLS.

Similarly, model One_M1a was estimated based on the MAXASLS variable, since this is the only relevant crowding variable for this model form:

```
B1a  DT=12.08+0.52*ONS+0.26*OFFS+0.009*MAXASLS,  (R2=0.64)
       (6.00)   (2.53)    (0.81)        (5.36)      (DW=1.93)
       [0.00]   [0.01]    [0.42]        [0.00]
```

It is clear that the coefficient of OFFS is insignificant in model B1a suggesting the variable OFFS can be dropped from the model, producing the following results:

```
DT=12.43+0.54*ONS+0.01*MAXASLS,  (R2=0.64)
  (6.33)   (2.65)       (8.12)      (DW=1.93)
  [0.00]   [0.01]       [0.00]
```

As indicated by the t-statistics and p-values, all coefficients are significant at 0.05

level, with R^2 about 0.64. It is clear that using the variables ONS and MAXASLS, to reflect the effect of crowding on board is a significant improvement over model A1a. The coefficients of the variables ONS and MAXASLS are 0.54 and 0.01 respectively in model B1a implying that the marginal dwell time for a single boarding is equivalent to about 50 additional standees.

Model estimations for form One_M1b produced the following results:

```
B1b  DT=12.47+0.67*ONS+0.39*OFFS+0.002*MAXASLS,  (R2=0.64)
       (8.47)   (2.51)    (3.34)       (1.11)       (DW=2.08)
       [0.00]   [0.02]    [0.00]       [0.28]
```

The coefficient of the variable MAXASLS, is less significant than in model A1b, implying that adding this crowding variable does not significantly improve the explanatory power of the model.

These results suggest that the on board crowding effect is most important when the boarding process dominates alighting.

3.5.4 Model C: DT = f(ONS, OFFS, ONLS, OFFAS)

This model recognizes that movement of alighting passengers would be affected by arriving standees, while movement of boarding passengers would be affected by departing standees. Therefore, the crowding effect may be represented by three alternative variables: OFFAS, ONLS, and SUMASLS, and can be represented in several different forms. Similar to model B, this model assumes the effect on dwell time of crowding on board as well as the number of passengers boarding and alighting. The first form of this model introduces the variable SUMASLS to express the marginal effect on the dwell time from the crowding on board, which is the sum of OFFAS and ONLS with the following results:

Model C1-1:

```
DT=12.50+0.55*ONS+0.23*OFFS+0.0078*SUMASLS,  (R2=0.62)
  (8.94)  (3.76)    (2.03)       (6.70)        (DW=2.06)
  [0.00]  [0.00]    [0.04]       [0.00]
```

All coefficients are strongly significant in C1-1 model with an R^2 of 0.62. Therefore, similar to model B1, adding the variable SUMASLS to reflect the effect of crowding on board significantly improves the explanatory power of the model. All coefficients are very similar to those in the MAXASLS form of model B1.

Model C1a1:

```
DT=12.01+0.54*ONS+0.24*OFFS+0.009*SUMASLS,  (R2=0.64)
  (6.02)   (2.65)    (0.76)       (5.49)       (DW=1.93)
  [0.00]   [0.01]    [0.45]       [0.00]
```

As indicated by the t-statistics and p-values, once again, the coefficient of the variable OFFS is insignificant in this model, therefore the variable OFFS can be removed to produce the following results:

```
DT=12.32+0.56*ONS+0.01*SUMASLS,  (R2=0.65)
  (6.33)   (2.78)      (8.25)      (DW=1.93)
  [0.00]   [0.00]      [0.00]
```

All coefficients are significant at 0.05 level with an R^2 of 0.65 which implies that adding the variable SUMASLS to reflect the effect of crowding on board is a significant improvement over model A1a. The coefficients of the variables ONS and SUMASLS are 0.56 and 0.01 respectively implying that the marginal dwell time for boarding is about 55 times greater than that of the SUMASLS.

Model C1b1:

$$DT = 12.46 + 0.65*ONS + 0.39*OFFS + 0.002*SUMASLS, \quad (R^2=0.65)$$
$$(8.60) \quad (2.43) \quad (3.43) \quad (1.25) \quad (DW=2.10)$$
$$[0.00] \quad [0.02] \quad [0.00] \quad [0.22]$$

Similar to model B1b, the coefficient of the variable SUMASLS reflecting crowding effect on board is less significant in this model. As indicated by the corrected R^2, it is clear that adding the variable SUMASLS does not significantly improve the explanatory power.

The second form of model C presents the separate effects of crowding on alighting and boarding passengers with two variables OFFAS and ONLS. The variable OFFAS represents the interaction between alightings and arriving standees while the variable ONLS reflects the interaction between boardings and departing standees. The model with the OFFAS and ONLS form produces the following results:

Model C1-2:

$$DT = 12.33 + 0.49*ONS + 0.33*OFFS + 0.0044*OFFAS + 0.010*ONLS, \quad (R^2=0.62)$$
$$(8.75) \quad (3.07) \quad (2.16) \quad (1.21) \quad (3.53) \quad (DW=2.10)$$
$$[0.00] \quad [0.00] \quad [0.03] \quad [0.24] \quad [1.00]$$

The coefficient of OFFAS is not significant as indicated by the p-value of 0.24. This may result from the multicollinearity of the variables. The correlation coefficient for the variables OFFAS and ONLS was 0.737 while that for the variables OFFAS and OFFS was 0.684 (see Table B.1). As previous discussed in section 3.5.1, it is desirable to avoid high correlations between independent variables. The rather high correlation between the variables OFFAS and ONLS which both reflect crowding, suggests that these variables should not be included together in any model formulation. One way of dealing with this problem is to introduce either variable ONLS or OFFAS to represent the crowding effect.

Model C1a2:

$$DT = 11.78 + 0.85\,ONS - 0.33\,OFFS + 0.025\,OFFAS - 0.00014\,ONLS, \quad (R^2 = 0.66)$$
$$(6.07) \quad (3.58) \quad (-0.85) \quad (3.64) \quad (-0.03) \quad (DW = 1.84)$$
$$[0.00] \quad [0.00] \quad [0.40] \quad [0.00] \quad [0.98]$$

As indicated by its t-statistics and p-values, the coefficients of OFFS and ONLS are insignificant and negative, implying that marginal dwell time is negatively related to the OFFS and ONLS which doesn't make sense. This may result from high correlation between independent variables. The correlation coefficient for the variables OFFS and OFFAS was 0.834 while that for the variables OFFAS and ONLS was 0.869 (see Table B.2).

Model C1b2:

$$DT = 11.66 + 0.42\,ONS + 0.57\,OFFS - 0.003\,OFFAS + 0.012\,ONLS, \quad (R^2 = 0.66)$$
$$(7.58) \quad (1.38) \quad (3.34) \quad (-0.78) \quad (1.68) \quad (DW = 2.14)$$
$$[0.00] \quad [0.17] \quad [0.00] \quad [0.44] \quad [0.10]$$

The coefficients of the variables ONS and OFFAS are insignificant in model C1b2. The negative coefficient of the variable OFFAS implying that the marginal dwell time is negatively related to the OFFAS which is not reasonable. Thus, the variable OFFAS can be dropped for further model estimation.

The variable ONLS or OFFAS is introduced in the third form of model C which assumes that either the departing standees affect the passengers boarding or the arriving standees affect the passengers alighting with the following results:

Model C1-3:

$$DT=11.94+0.45*ONS+0.47*OFFS+0.013*ONLS, \quad (R^2=0.62)$$
$$(8.70) \quad (2.88) \quad (4.66) \quad (6.64) \quad (DW=2.12)$$
$$[0.00] \quad [0.01] \quad [0.00] \quad [0.00]$$

As indicated by the t-statistics and p-values, all coefficients are strongly significant with an R^2 of 0.62. The coefficients for the variable ONS and OFFS are 0.45 and 0.47 respectively, while that for ONLS is 0.013 implying the marginal dwell times for boarding and alighting are similar and about 35 times greater than the crowding variable ONLS. Similar to the first form of model C1, it is clear that adding the variable ONLS to reflect the effect of crowding on board significantly improves the explanatory power of the model. The constant is about 12 seconds in this model.

Since the coefficients of the variables OFFS and ONLS are insignificant and negative in model C1a2, it is possible to omit these variables with the following results:

Model C1a3:

$$DT=11.43+0.80*ONS+0.022*OFFAS, \quad (R^2=0.67)$$
$$(6.15) \quad (4.51) \quad (8.84) \quad (DW=1.85)$$
$$[0.00] \quad [0.00] \quad [0.00]$$

All coefficients are highly significant in model C1a3 with an R^2 of 0.67. The coefficient of the variables ONS and OFFAS are 0.80 and 0.022 respectively implying that the marginal dwell times for boarding is about 35 times greater than for the variable OFFAS. It also appears that adding the variable OFFAS to reflect the effect of crowding on board is a significant improvement over the the model A1a. The constant is about 11 seconds in this model.

The third form of model C1b is to drop the variable OFFAS from model C1b2 and produce the following results:

Model C1b3:

$$DT=12.19+0.48*ONS+0.47*OFFS+0.007*ONLS, \quad (R^2=0.66)$$
$$(8.86) \quad (1.61) \quad (4.23) \quad (1.74) \quad (DW=2.17)$$
$$[0.00] \quad [0.12] \quad [0.00] \quad [0.09]$$

As indicated by the t-statistics and p-values, the coefficients of the variables ONS and ONLS are only marginally significant. In this model, the coefficients of the variables ONS, OFFS, and ONLS are 0.48, 0.47, and 0.007 respectively, implying that the marginal dwell time for boarding and alighting is about 65 times greater than the ONLS. The constant is about 12 seconds in model C1b3.

In conclusion, it is clear from these model estimations that adding either the variables SUMASLS, or ONLS in model C1 to reflect the effect of crowding on board significantly improves the explanatory power of the model. There also appears to be little to choose between these two forms. However, because the SUMASLS form includes possible effects of both AS on OFFS and LS on ONS, it appears better able to reflect all possible crowding effects. Based on this viewpoint, the SUMASLS form is preferred among these three model forms.

Similarly, in model C1a, adding either the variables SUMASLS and OFFAS to reflect the crowding effect is a significant improvement over model A1a which does not include the effect of passenger crowding on board.

As to model C1b, the marginal dwell time using the variable SUMASLS to reflect the effect of crowding is negligible, while that from the variable ONLS is only marginally significant.

3.5.5 Model D: DT = f(ONS, OFFS, AS, LS, SUMASLS)

This model assumes the effect on dwell time of crowding on board as well as the number of passengers boarding and alighting. The crowding effect was examined using three independent variables: AS, LS, and SUMASLS, with the following results:

Model D1-1:

```
DT=10.93+0.52*ONS+0.44*OFFS-0.11*AS+0.17*LS
  (6.69)   (3.36)    (3.20)   (-1.70)   (2.75)
  [0.00]   [0.00]    [0.00]   [0.09]    [0.01]

                                  +0.0051*SUMASLS,  (R2=0.64)
                                       (1.93)       (DW=2.15)
                                       [0.06]
```

As indicated by the t-statistics and p-values, the coefficient of the variable AS is insignificant at the 0.05 level and is negative, implying that marginal dwell time is negatively related to the AS which is not reasonable. Once again, this unexpected result may be due to multicollinearity between independent variables. Table B.1 shows the correlation coefficients between the variables AS, LS, and SUMASLS are between 0.87 and 0.92, implying that these variables should not be included together in any model.

Model D1a1:

```
DY=10.35+0.51*ONS+0.56*OFFS-0.097*AS+0.155*LS
  (4.55)   (2.35)    (1.64)     (-1.20)   (2.20)
  [0.00]   [0.02]    [0.10]     [0.25]    [0.03]

                                 +0.006*SUMASLS,  (R2=0.66)
                                     (1.74)       (DW=2.05)
                                     [0.08]
```

As indicated by the t-statistics and p-values, the coefficients of the variable AS is insignificant and negative, implying that marginal dwell time is negatively related to the AS which doesn't make sense. As discussed in earlier D1-1 model, the variable AS should not be included in further model estimation.

Model D1b1:

```
DT=10.20+0.79*ONS+0.56*OFFS+0.057*AS+0.074*LS,
  (5.48)   (2.36)    (2.76)     (0.24)    (0.32)
  [0.00]   [0.02]    [0.01]     [0.81]    [0.76]

                                -0.004*SUMASLS,  (R2=0.66)
                                    (-1.07)      (DW=2.16)
                                    [0.29]
```

The coefficients of the variables AS, LS, and SUMASLS are insignificant in model D1b1. Once again, this may be due to high correlation between these variables. The correlation coefficients for the variables AS and LS, AS and SUMASLS, and LS and SUMASLS were 0.978, 0.891, and 0.836 respectively (see Table B.3), once again, which suggests that this model form is not worth persuing.

Because the coefficient of the variable AS is not reasonable and insignificant in the first form of model D1-1, the variable AS is dropped from the model to produce the following model:

Model D1-2:

```
DT=10.60+0.62*ONS+0.39*OFFS+0.10*LS
  (6.48)   (4.18)   (2.90)     (2.14)
  [0.00]   [0.00]   [0.00]     [0.03]

                              +0.0032*SUMASLS,  (R2=0.63)
                                   (1.33)       (DW=2.10)
                                   [0.19]
```

As indicated by its t-statistic and p-value, the coefficient of the variable SUMASLS is insignificant in the second form of this model implying it contributes little to the prediction of the dwell time, given that the variable LS is included in the model. Therefore, it appears possible to drop the variable SUMASLS for further model estimation.

Similarly, the variable AS is omitted from model D1a1 producing the following results:

Model D1a2:

```
DT=9.96+0.60*ONS+0.44*OFFS+0.11*LS
  (4.41) (2.96)    (1.35)     (1.84)
  [0.00] [0.01]    [0.18]     [1.84]

                                +0.0043*SUMASLS,  (R2=0.65)
                                      (1.38)      (DW=1.97)
                                      [0.17]
```

As indicated by the t-statistics and p-values, the coefficients of the variables OFFS and SUMASLS are less significant in this model suggesting that it is possible to omit these variables for further model estimation.

In the third form of model D, either LS and AS is introduced to reflect the effect of crowding on board. The variable SUMASLS is dropped from model D1-2 producing the following form of model D1:

Model D1-3:

```
DT=9.24+0.71*ONS+0.52*OFFS+0.16*LS,  (R2=0.63)
  (7.19) (5.40)    (5.35)    (6.98)  (DW=2.10)
  [0.00] [0.00]    [0.00]    [0.00]
```

As indicated by the t-statistics and p-values, all coefficients are strongly significant in the LS form, with an R^2 of 0.63. The constant is about 9 seconds in this model, with coefficients of 0.71, 0.52, and 0.16 for the variables ONS, OFFS, and LS, which imply the marginal dwell time for boarding is approximately 1.4 times that for alighting, and 4.4 times that for leaving standees.

The variables SUMASLS and OFFS are dropped from the model D1a2 producing the following results:

Model D1a3:

```
DT=8.10+0.88*ONS+0.22*LS,  (R2=0.62)
  (4.13)  (4.65)    (7.61)   (DW=1.98)
  [0.00]  [0.00]    [0.00]
```

Similarly, all coefficients are strongly significant in this model, with an R^2 of 0.62. The coefficients of the variables ONS and LS are about 0.88 and 0.22 respectively, implying that the marginal dwell time for boarding is about 4 times that for leaving standees. The constant is about 8 seconds.

As discussed earlier in model D1b1, the variables AS, LS, and SUMASLS should not be included together in any model. In the final form of model D1b, the variable LS is introduced to reflect the effect of crowding, producing the following results:

Model D1b3:

```
DT=11.46+0.60*ONS+0.48*OFFS+0.066*LS,  (R2=0.67)
  (8.37)   (2.64)    (4.38)     (2.09)    (DW=2.20)
  [0.00]   [0.01]    [0.00]     [0.05]
```

All coefficients are significant at 0.05 level with an R^2 of 0.67 implying that adding the variable LS reflecting the crowding effect does improve the explanatory power of the model. The coefficients of the variables ONS, OFFS, and LS are 0.60, 0.48, and 0.066

respectively which imply that the marginal dwell time for boarding is about 1.2 times that for alighting, and about 9 times greater than the leaving standees.

In conclusion, it is obvious that adding the variable LS in these three models to reflect the effect the crowding on board significantly improves the explanatory power of the models. However, if there were standees, but no passengers boarding and alighting, the number of standees should have no significant impact on the dwell time. Further, the coefficient for the variable LS is rather large compared with that of the SUMASLS and MAXASLS forms. Therefore, in the above situation, the model may not accurately predict the dwell time. For example, provided the passengers on board are close to the LRV's capacity, and no passenger boarding and alighting occurs, then the prediction of the dwell time may be dramatically affected by the marginal contribution from the LS term. Based on this standpoint, models B1-1 and B1a1 with the MAXASLS form and models C1-1 and C1a1 with the SUMASLS form may be preferred over this LS model form.

Finally, the equality of individual parameter from the two data sets (ONS≥OFFS, and OFFS>ONS) was examined by t-statistics. The test results for A1a-A1b, B1a1-B1b1, C1a1-C1b1, and D1a3-D1b3 model pairs are presented in Table 3.9, which shows that estimates for the variable OFFS from data set 1a are significantly different from those for data set 1b when only the variables ONS and OFFS are included in the model. It also appears that the estimates for the variables reflecting the crowding effect are significantly different between two data sets, while the parameters of the constant and ONS are insignificant at the 0.05 level.

Table 3.9 t test for Equality of Individual Coefficient between One-Car Train Data Sets (t-statistics shown)

Variable	A1a & A1b models	B1a & B1b models	C1a1 & C1b1 models	D1a3 & D1b3 models
Constant	-1.26	-0.02	-0.06	-1.39
ONS	0.07	-0.39	0.27	0.30
OFFS	3.38*	-	-	-
MAXASLS	-	3.67*	-	-
SUMASLS	-	-	3.74*	-
LS	-	-	-	2.54*

* significant at 0.05 level.

3.5.6 Model Estimation with Alternative Forms

The previous model estimations, have assumed the effect on dwell time of crowding is linear, however, it may well be nonlinear. To investigate this possibility, a nonlinear form for the variables reflecting crowding is introduced in the following model formulations. As discussed in the sections 3.5.3-3.5.5, the crowding effect is examined with respect to five alternative variables: MAXASLS, SUMASLS, ONLS, OFFAS and LS. Two examples with the MAXASLS and ONS*LS^E forms are presented below:

$$DT = b_0 + b_1*ONS + b_2*OFFS + b_3*(MAXASLS)^E$$

$$DT = b_0 + b_1*ONS + b_2*OFFS + b_3*ONS*LS^E$$

These models were estimated by changing the value of the exponent (E) over the range of 0.0 to 5.0. Surprisingly, there are wide ranges for optimal values of E in these models.

A. One_M1 Model:

Figures 3.5 to 3.9 plot the corrected R^2 values as a function of the exponent of the variables MAXASLS, SUMASLS, ONLS and LS for model estimations for each model form. As indicated in Figure 3.5, the corrected R^2 is a maximum with E in the range of 1.0 to 1.3 in model B1-1 with the MAXASLS form. Corresponding ranges of E for maximum R^2 values are 0.8-2.0, 0.8-1.3, 2.5 and 2.5-3.0 respectively in the model with the SUMASLS, ONLS, ONS*LS^E and LS forms.

Figure 3.5 R-SQUARE vs. E of MAXASLS

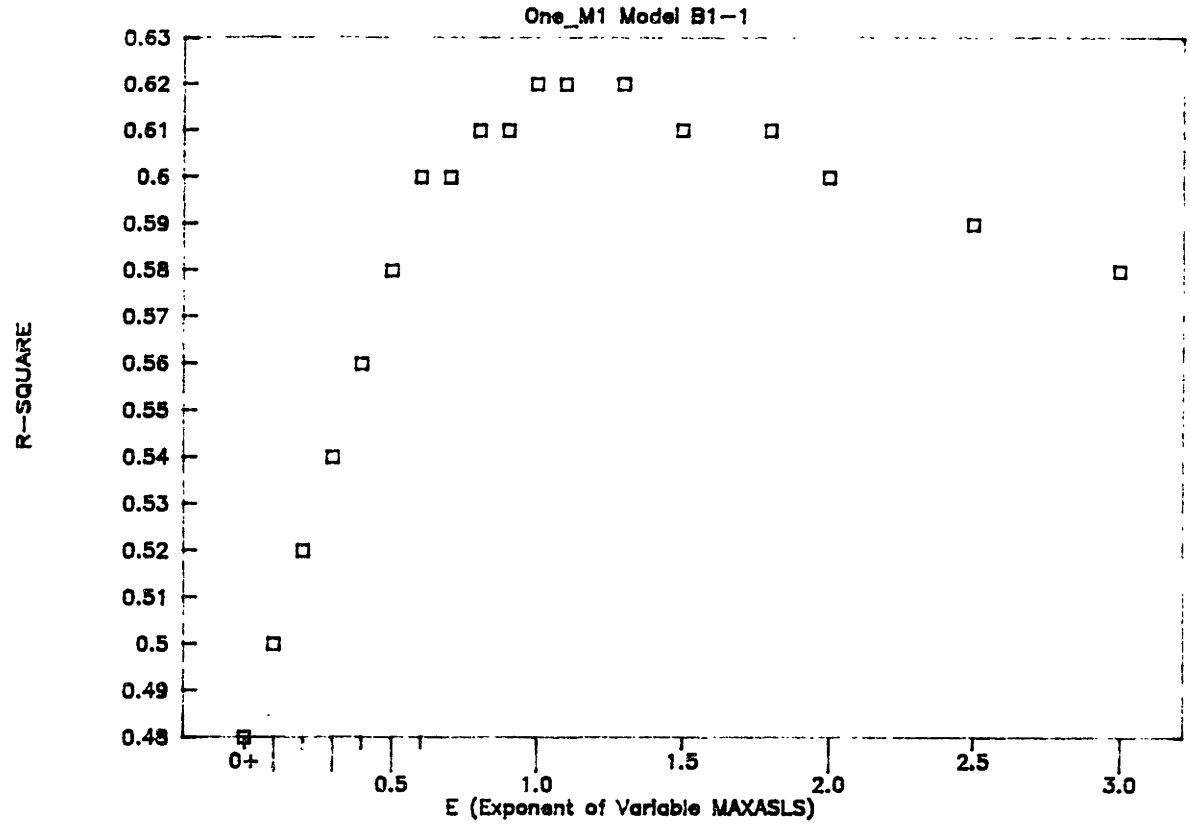


Figure 3.6 R-SQUARE vs. E of SUMASLS

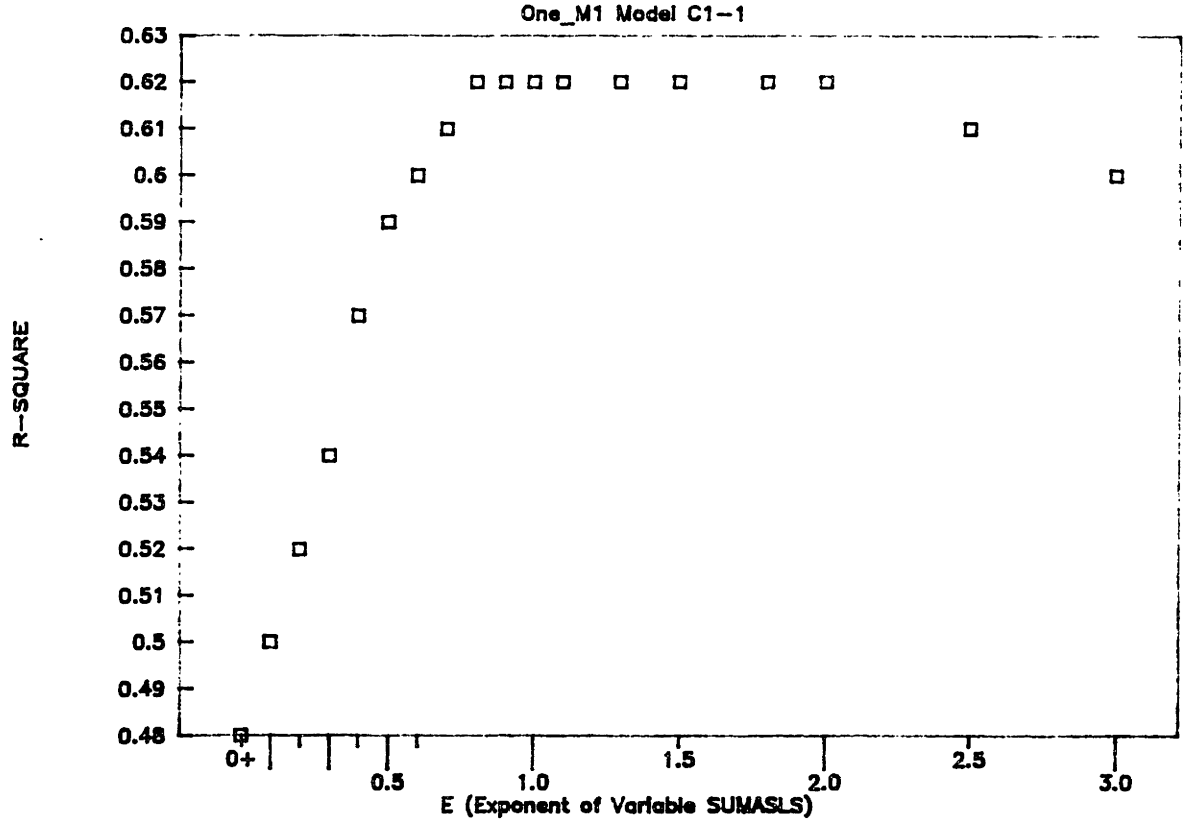


Figure 3.7 R-SQUARE vs. E of ONLS

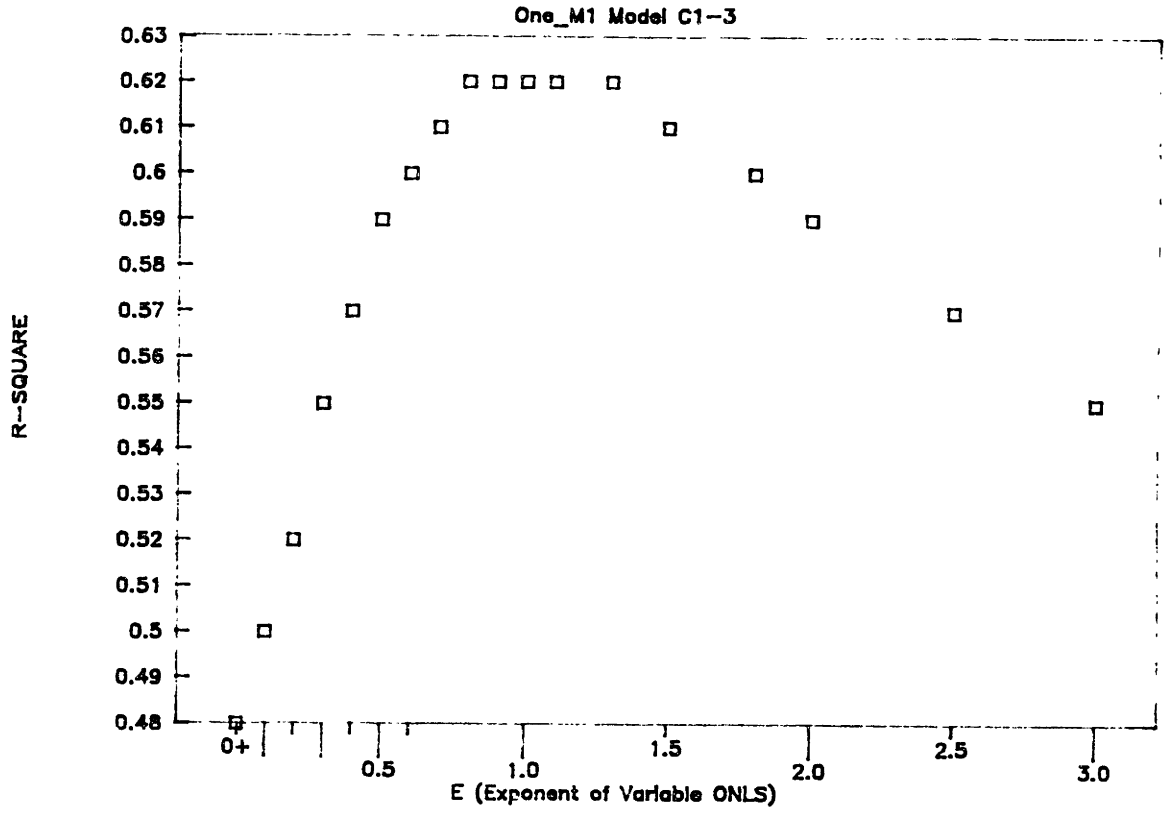


Figure 3.8 R-SQUARE vs. E of LS

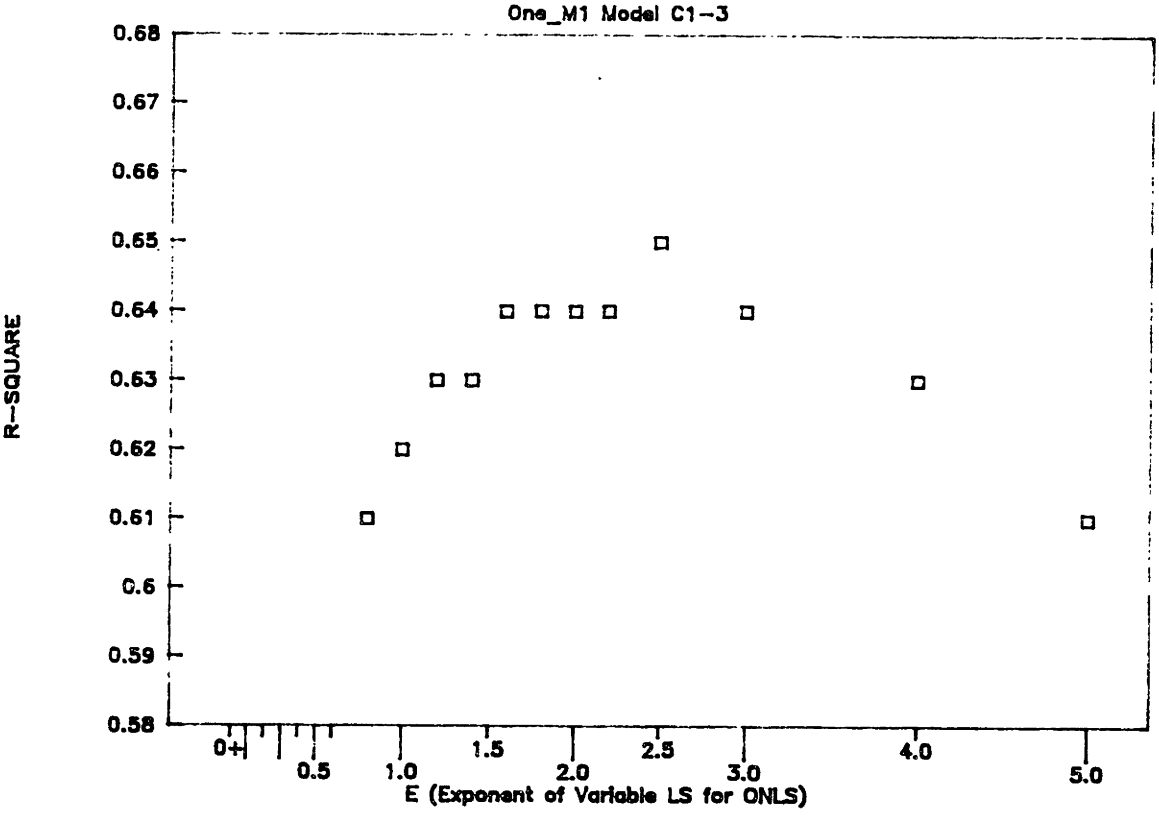


Figure 3.9 R-SQUARE vs. E of LS

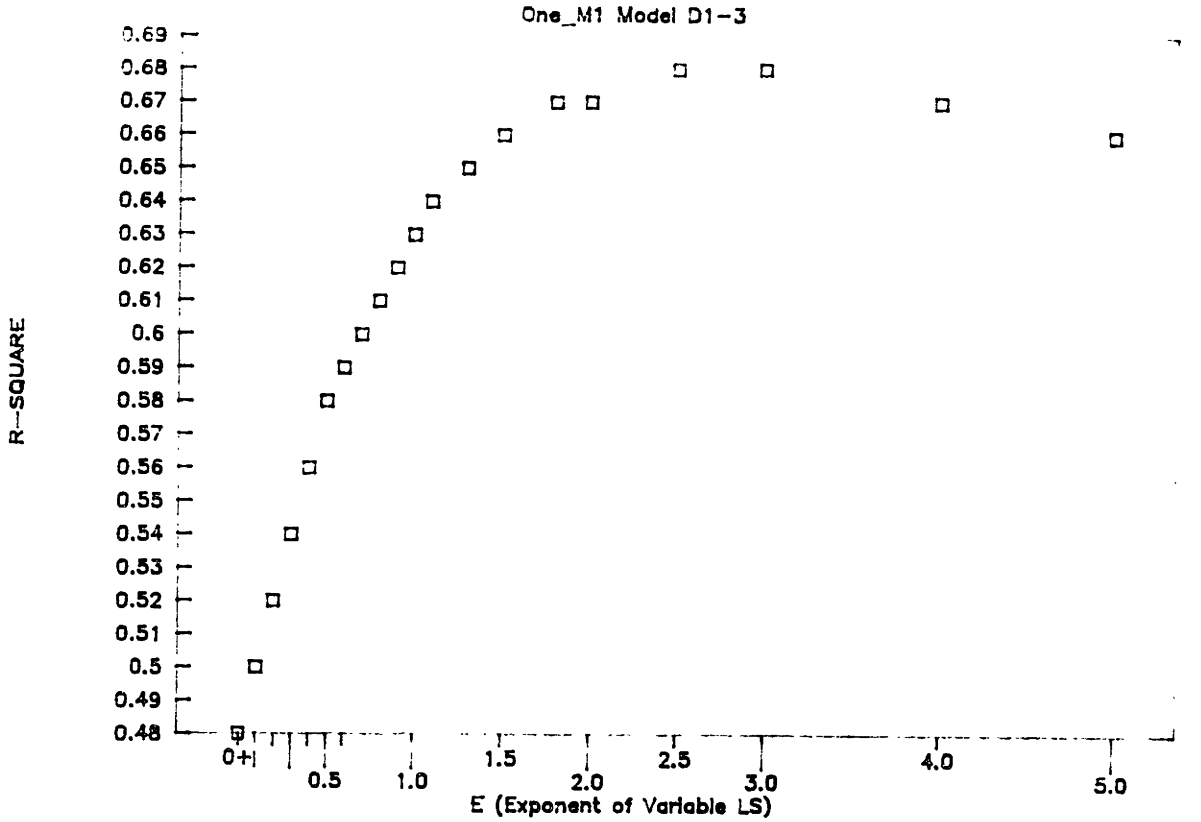
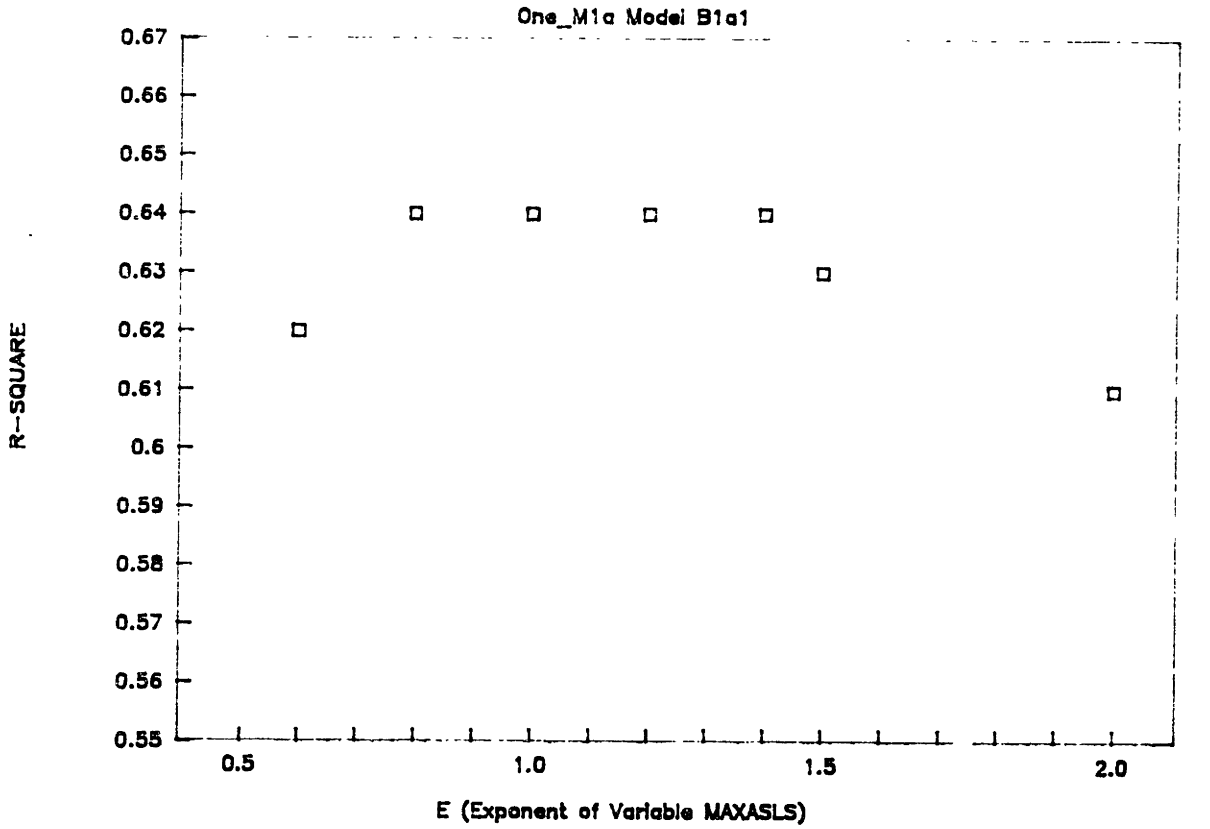


Figure 3.10 R-SQUARE vs. E of MAXASLS



The wide ranges for E imply that different exponential forms of the variables can be applied in each model. Because the explanatory power is similar with ranges of E of 1.0-1.3, 0.8-2.0, and 0.8-1.3 respectively for the models B1-1, C1-1, and C1-3, but the linear model is more easily interpreted, it is suggested that the models B1-1, C1-1, and C1-3 discussed in sections 3.5.3-3.5.4 are appropriate to explain variations in dwell time. Model C1-3 with ONS*LS^E form with E of 2.5 and model D1-3 with nonlinear form of LS with E in the range of 2.5 to 3.0, do offer better explanatory power, implying that the variables reflecting on board crowding for these two models are more related to departing passenger load with an exponent of 2.5. However, the LS form still has the drawback discussed in the section 3.5.5. The estimated models C1-3 and D1-3 with an exponent of 2.5 for the LS term are presented below:

Model C1-3(B):

$$DT = 11.43 + 0.69*ONS + 0.48*OFFS + 1.35*10^{-5}*ONS*LS^{2.5}, \quad (R^2=0.65)$$
(8.78) (5.38) (4.99) (7.41) (DW=2.16)
[0.00] [0.00] [0.00] [0.00]

Model D1-3:

$$DT = 10.05 + 0.78*ONS + 0.50*OFFS + 2.0*10^{-4}*LS^{2.5}, \quad (R^2=0.68)$$
(8.32) (6.70) (5.51) (8.50) (DW=2.10)
[0.00] [0.00] [0.00] [0.00]

B. One_M1a Model:

Figures 3.10-3.14 plot the corrected R^2 values as a function of the exponent of the variables MAXASLS, SUMASLS, OFFAS, and LS for estimations for each model form. As indicated in Figure 3.10, the corrected R^2 is highest over the range of E from 0.8 to 1.4 in the model B1a1 with the MAXASLS form. Corresponding ranges of E for maximum corrected R^2 values are 1.0-1.2, 1.2-1.8, 1.6-2.7, and 2.5-3.0 respectively in the models with SUMASLS, OFFAS, OFFS*AS^E and LS forms (see Figures 3.11-3.14).

Figure 3.11 R–SQUARE vs. E of SUMASLS

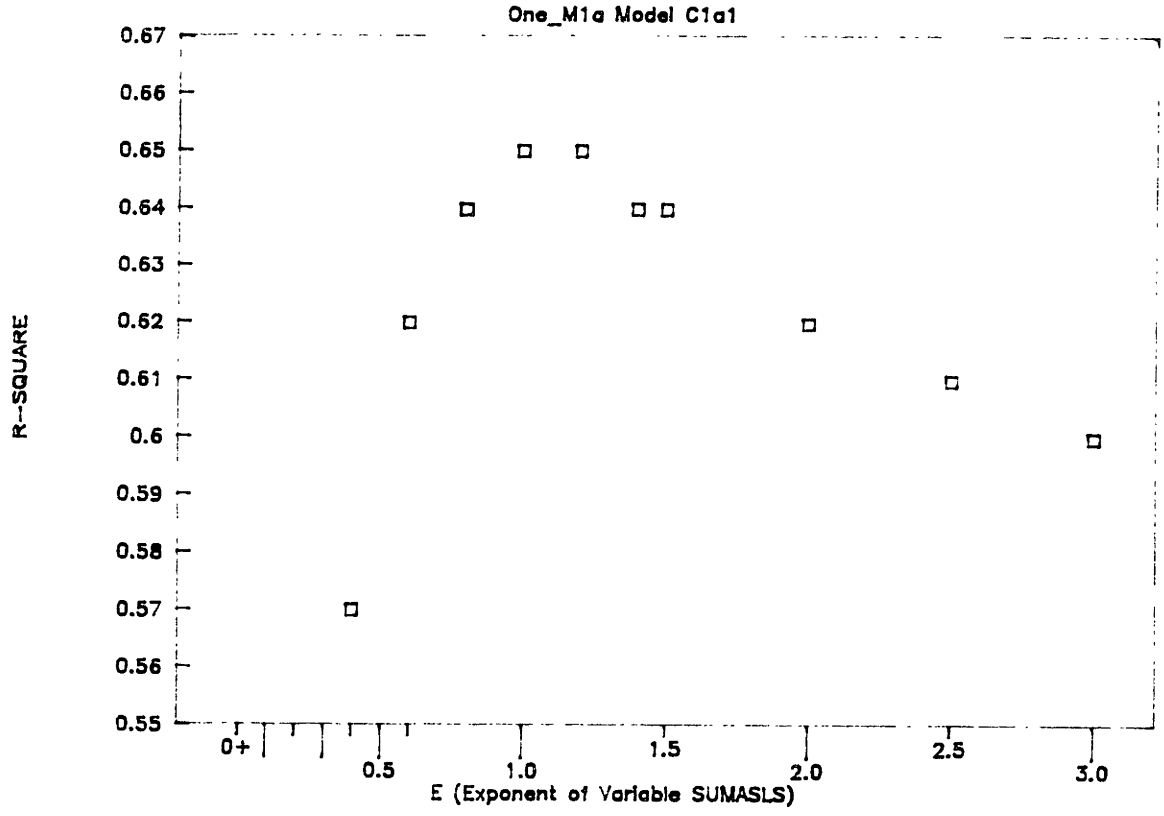


Figure 3.12 R–SQUARE vs. E of OFFAS

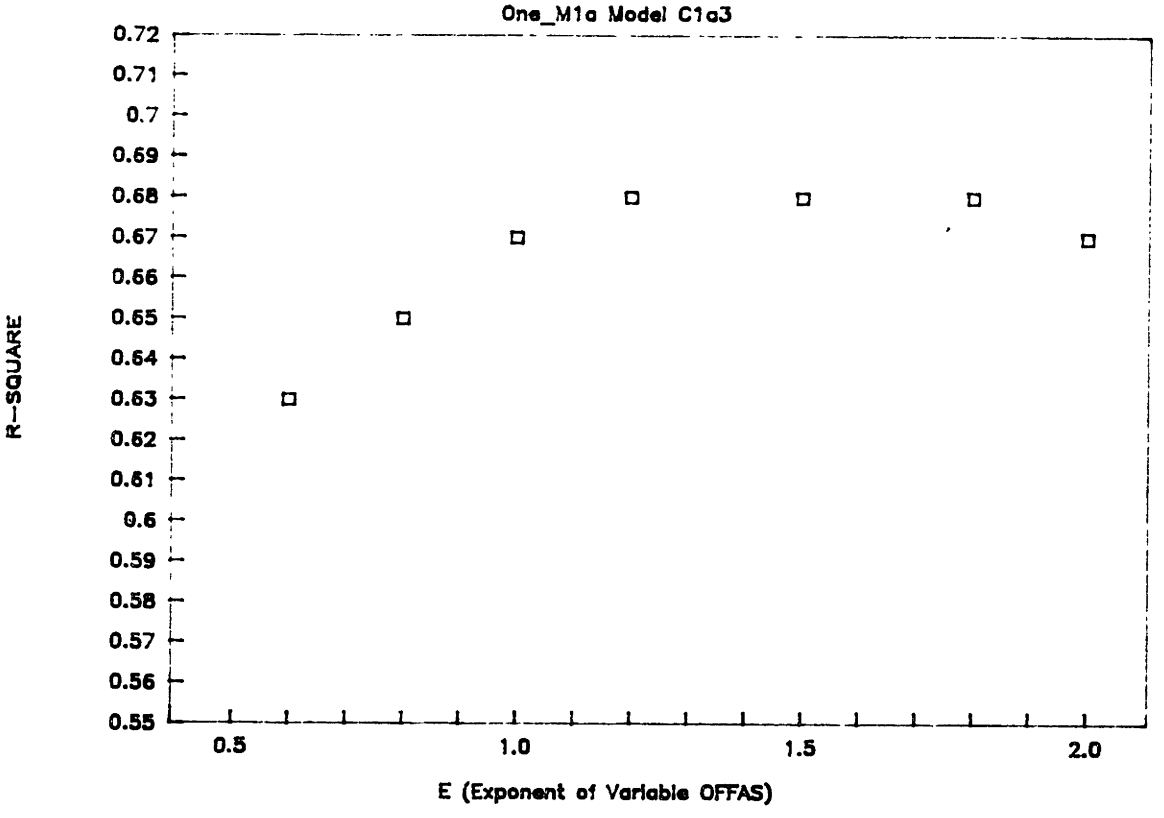


Figure 3.13 R-SQUARE vs. E of AS

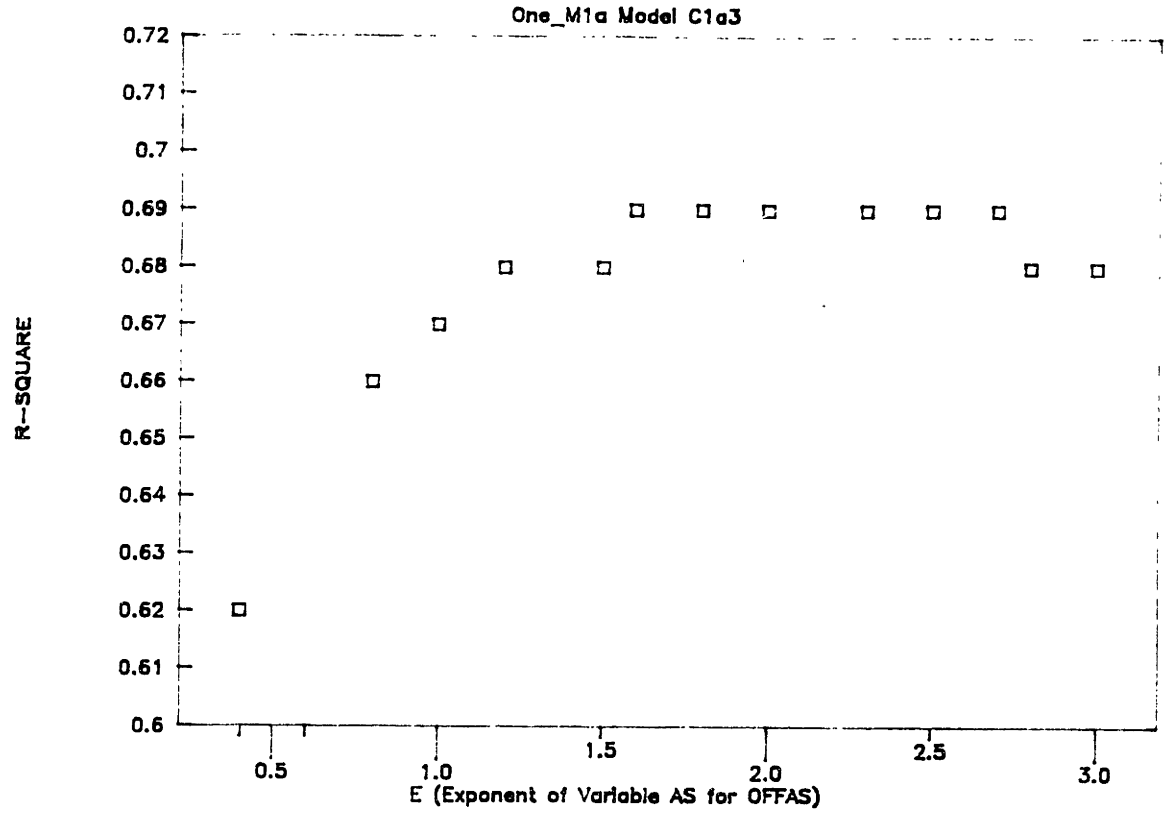
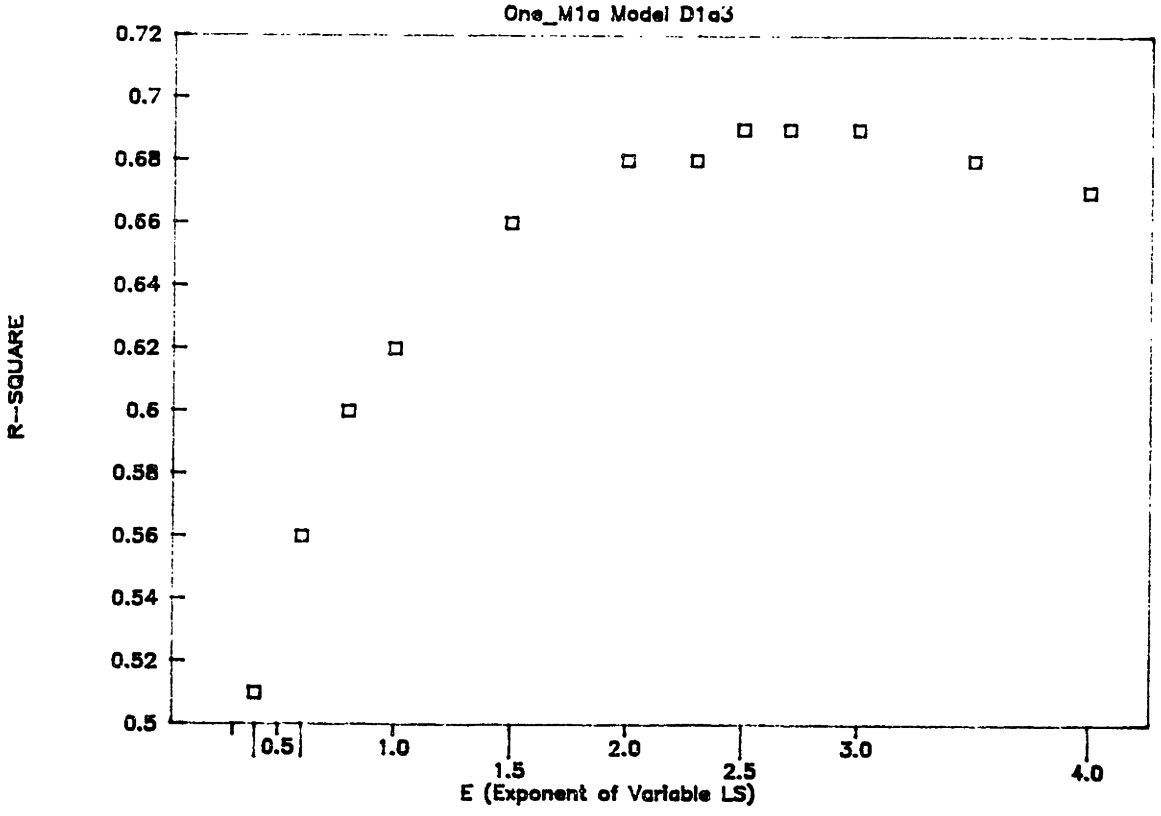


Figure 3.14 R-SQUARE vs. E of LS



The wide ranges for E suggests that different exponential forms of the variables can be applied in each model. As indicated by these results, models B1a1 and C1a1 discussed earlier are appropriate to explain variations in dwell time. But model C1a3 in the OFFAS form with E of 1.2-1.8 and in the $OFFS*AS^E$ form with E in the range of 1.6-2.7 does offer a better explanatory power. Similarly, the model D1a3 with nonlinear form of LS with E in the range of 2.5-3.0 offers improvement over the linear model. The estimated models C1a3 and D1a3 with an exponent of 1.5, 2.0, and 2.7 respectively for the OFFAS, AS, and LS terms are presented below:

Model C1a3(A):

$$DT=12.0+0.83*ONS+6.32*10^{-4}*OFFAS^{1.5}, \quad (R^2=0.68)$$
(6.52) (4.81) (9.13) (DW=1.90)
[0.00] [0.00] [0.00]

Model C1a3(B):

$$DT=11.58+0.85*ONS+2.44*10^{-4}*OFFS*AS^{2.0}, \quad (R^2=0.69)$$
(6.42) (5.08) (9.38) (DW=1.91)
[0.00] [0.00] [0.00]

Model D1a3:

$$DT=9.71+0.94*ONS+1.1*10^{-4}*LS^{2.7}, \quad (R^2=0.69)$$
(5.44) (5.69) (9.34) (DW=2.05)
[0.00] [0.00] [0.00]

C. One_M1b Model

The nonlinear form for four alternative variables, MAXASLS, SUMASLS, ONLS, and LS, are introduced as with the previous model formulations, with estimations for values of the exponent from 0.0 to 5.0. The model estimations yield the same results as for the linear forms, the coefficients of crowding variables MAXASLS, SUMASLS, and ONLS are insignificant at 0.05 level in the nonlinear forms. However, in models C1b3 and D1b3, the crowding variables either with ONS*LS^E form and E of 1.3-2.0 or with LS and E in the range of 0.8 to 4.0 are significant at the 0.05 level. Figures 3.15 and 3.16 plot the corrected R^2 values as a function of the variables LS. It is clear from Figure 3.15 that the model C1b3, the ONS*LS^E form, with E in the range of 1.4-2.8 offers highest R^2 values. Similarly, the model D1b3, the LS form, with E in the range 1.2-2.7 offers highest explanatory power. The estimated models C1b3 and D1b3 with an exponent of 2.0 for the LS term are presented below:

Model C1b3(B);

$$DT = \underset{(8.73)\ [0.00]}{11.85} + \underset{(2.78)\ [0.01]}{0.63}*ONS + \underset{(4.30)\ [0.00]}{0.48}*OFFS + \underset{(1.94)\ [0.06]}{7.7*10^{-5}}*ONS*LS^{2.0}, \quad (R^2=0.67)\ (DW=2.19)$$

Model D1b3:

$$DT = \underset{(8.47)\ [0.00]}{11.46} + \underset{(3.17)\ [0.00]}{0.66}*ONS + \underset{(4.46)\ [0.00]}{0.49}*OFFS + \underset{(2.26)\ [0.03]}{7.7*10^{-4}}*LS^{2.0}, \quad (R^2=0.68)\ (DW=2.21)$$

In conclusion, it is clear that the variables reflecting on board crowding for models C and D are more related to departing or arriving standees with an exponent greater than 1.0. The comparison of exponents for variables reflecting crowding effect is summarized in Table 3.10.

Figure 3.15 R-SQUARE vs. E of LS

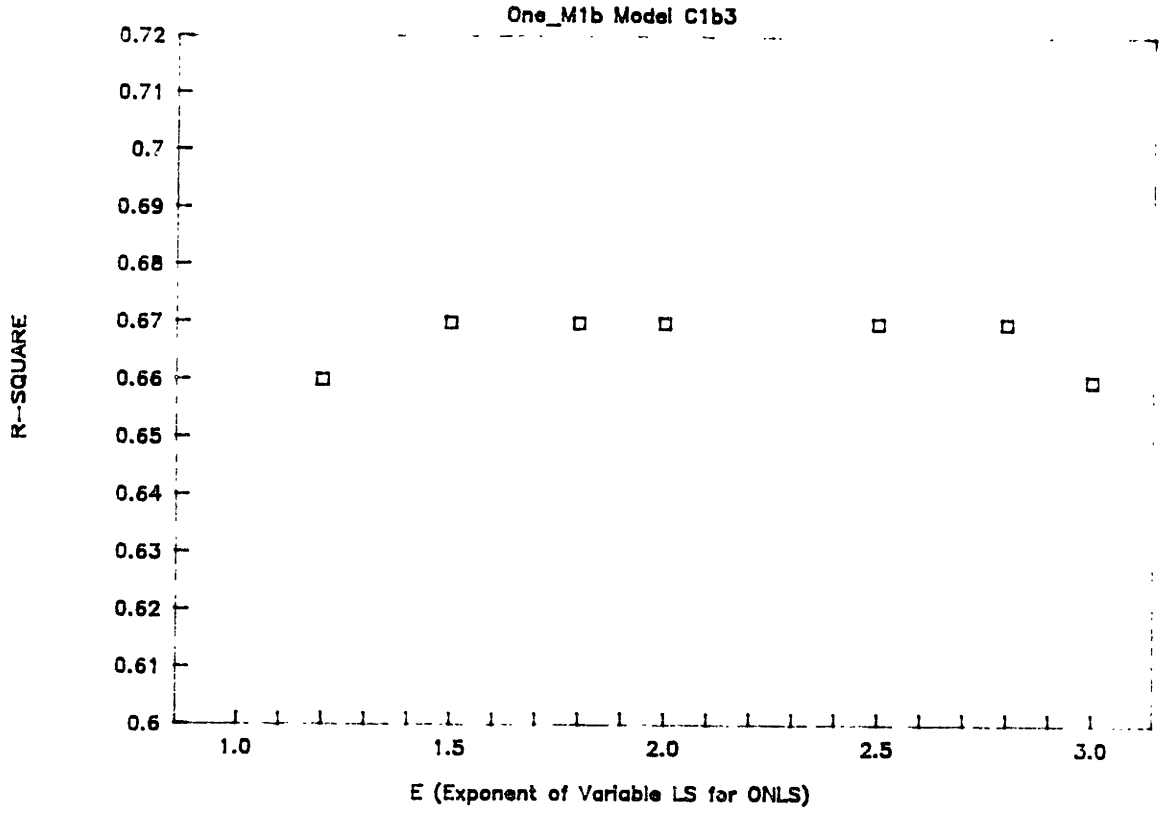


Figure 3.16 R-SQUARE vs. E of LS

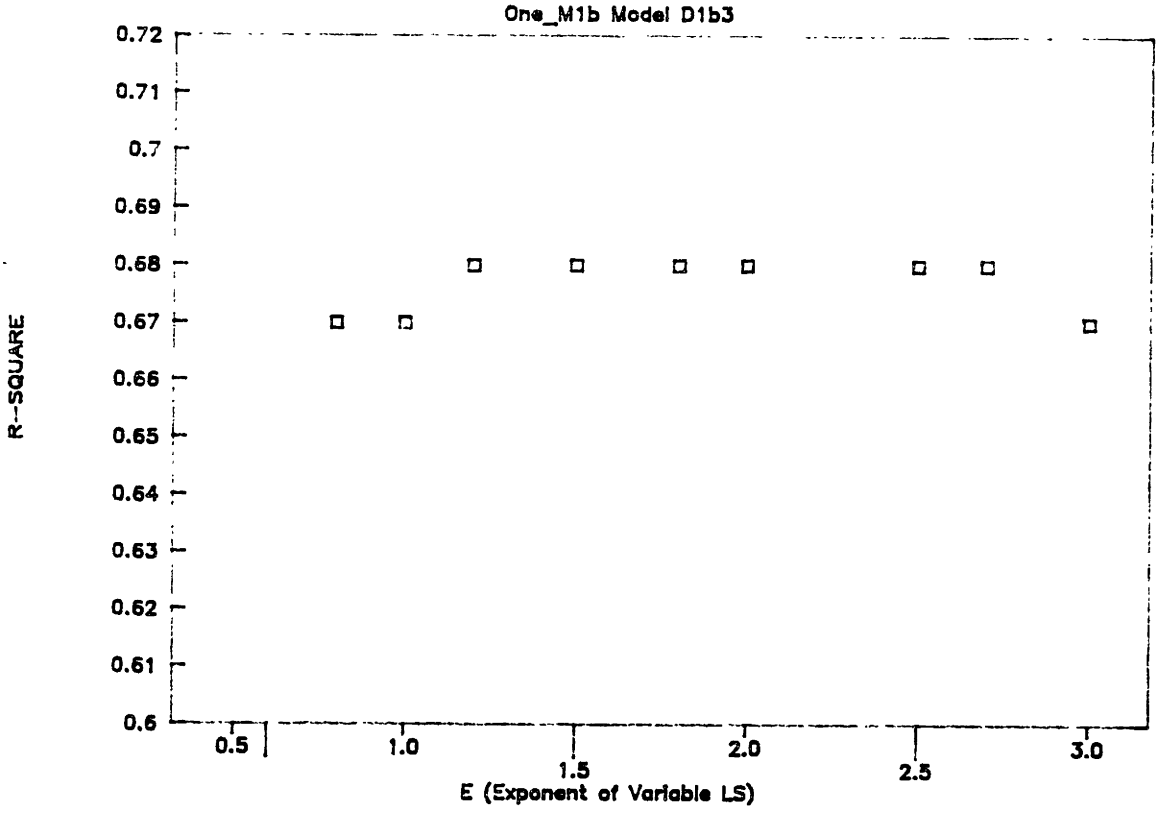


Table 3.10 Comparison of Exponents for Variables Reflecting Crowding Effect

Model Form	One_M1 model	One_M1a model	One_M1b model
MAXASLS	1.0-1.3	0.8-1.4	-
SUMASLS	0.8-2.0	1.0-1.2	-
ONLS	0.8-1.3	-	-
ONS*LS(EXP)	2.5	-	1.4-2.8
LS	2.5-3.0	2.5-3.0	1.2-2.7
OFFAS	-	1.2-1.8	-
OFFS*AS(EXP)	-	1.6-2.7	-
AS	-	-	-

3.5.7 Checking the Assumptions for Linear Regression

There are four key assumptions which must be met for a linear regression model to be appropriate:

1: The mean of each error component is zero.

2: The error components are approximately normal.

3: The variance of the error component is the same for each predicted value.

4: The errors are independent of each other.

These four key assumptions are required to construct confidence intervals and to perform hypothesis tests. If these assumptions are violated, the resulting models may still provide accurate predictions of DT, but the validity of any inferences from the model will be questionable. Therefore, before conclusions are drawn from this analysis the assumptions for those most promising models discussed earlier (Models B1, B1a, C1-1, C1a1, C1-3, C1a3, C1b3, D1-3, D1a3 and D1b3) should be checked to make sure that none are violated.

Because an exactly normal distribution is not necessary for assumption 2, and problems arise only when the distribution is severely skewed and does not resemble a normal distribution, assumptions 1 and 2 are checked by plotting histograms of residuals. Assumption 3 is checked by plotting standardized residuals against predicted dwell times, and Assumption 4 is checked using the Durbin-Watson statistic (DW). As discussed earlier, the ideal value of DW is 2, in which case, the errors are completely uncorrelated, and there is no violation of the independent errors assumption.

Appendix C presents the results for those promising models with a linear form for the variables reflecting crowding, along with several examples for the variables in nonlinear

form. It is clear from the Figures C.1 to C.6, that the error component is approximately bell-shape, with the mode at zero and the variance of the error component almost the same for predicted values below 40 seconds. Further, the DW statistics are between 1.85 and 2.15 for these desirable models with linear form of MAXASLS, SUMASLS, ONLS, OFFAS and LS, while the DW statistics are between 1.90 and 2.21 for these models with nonlinear forms. Because all DW statistics are close to 2 no significant autocorrelation exists. Thus, it appears that these desirable models do not violate the assumptions of the multiple linear regression model. Therefore, models B1, B1a, C1-1, C1a1, C1-3, C1a3, C1b3, D1-3, D1a3 and D1b3 look promising as dwell time functions.

3.5.8 Summary of One-Car Models

From the preceding analysis, the following conclusions can be drawn:

(1) It appears that linear regression models can accurately describe dwell time using several variables, moreover, the dwell time model should consider the effects of passenger crowding. The most interesting models are summarized below:

Model B1-1:

$$DT = 12.59 + 0.55*ONS + 0.22*OFFS + 0.0076*MAXASLS, \quad (R^2 = 0.62)$$
$$(8.87) \quad (3.66) \quad (1.89) \quad (6.49) \quad (DW = 2.05)$$
$$[0.00] \quad [0.00] \quad [0.06] \quad [0.00]$$

Model B1a1:

$$DT = 12.43 + 0.54*ONS + 0.01*MAXASLS, \quad (R^2 = 0.64)$$
$$(6.33) \quad (2.65) \quad (8.12) \quad (DW = 1.93)$$
$$[0.00] \quad [0.01] \quad [0.00]$$

Model C1-1:

$DT=12.50+0.55*ONS+0.23*OFFS+0.0078*SUMASLS$, $(R^2=0.62)$
(8.94) (3.76) (2.03) (6.70) (DW=2.06)
[0.00] [0.00] [0.04] [0.00]

Model C1a1:

$DT=12.32+0.56*ONS+0.01*SUMASLS$, $(R^2=0.65)$
(6.33) (2.78) (8.25) (DW=1.93)
[0.00] [0.01] [0.00]

Model C1-3:

$DT=11.94+0.45*ONS+0.47*OFFS+0.013*ONLS$, $(R^2=0.62)$
(8.70) (2.88) (4.66) (6.64) (DW=2.12)
[0.00] [0.01] [0.00] [0.00]

Model C1a3:

$DT=11.43+0.80*ONS+0.022*OFFAS$, $(R^2=0.67)$
(6.15) (4.51) (8.84) (DW=1.85)
[0.00] [0.00] [0.00]

Model D1-3:

$DT=9.24+0.71*ONS+0.52*OFFS+0.16*LS$, $(R^2=0.63)$
(7.19) (5.40) (5.35) (6.98) (DW=2.10)
[0.00] [0.00] [0.00] [0.00]

Model D1a3:

$DT=8.10+0.88*ONS+0.22*LS$, $(R^2=0.62)$
(4.13) (4.56) (7.61) (DW=1.98)
[0.00] [0.00] [0.00]

Model D1b3:

$$DT=11.46+0.60*ONS+0.48*OFFS+0.066*LS, \quad (R^2=0.67)$$

(8.37)	(2.64)	(4.38)	(2.09)	(DW=2.20)
[0.00]	[0.01]	[0.00]	[0.05]	

(2) The boarding and alighting process appears different between two sets of data: ONS≥OFFS, and OFFS>ONS. In the former data set, it is clear that adding any alternative form of the variables proposed to reflect the effect of passengers crowding on board does significantly improve the explanatory power of the model. It also appears that the marginal dwell time for alighting is negligible in this data set. In the latter data set, the linear regression model explains dwell times using just two variables ONS and OFFS. However, in some cases, adding one variable to reflect the effects of passengers crowding may be a slight improvement over the model without the crowding effect. T tests show that coefficients for the two sets of data, ONS≥OFFS and OFFS>ONS, are significantly different for those variables reflecting crowding which implies that aggregation may obscure those particular characteristics related to the two sets of data.

(3) A variable reflecting the product of passenger movements and standees, i.e. the variable MAXASLS ,SUMASLS, or ONLS makes more sense than using the standee variable alone, (LS or AS), because if there were standees, but no passengers boarding or alighting, the number of passengers standing should have no significant impact on dwell time. Based on this viewpoint, models B1-1, C1-1 and C1-3 may be preferred over model D1-3. Similarly, models B1a1, C1a1 and C1a3 are preferred over model D1a3. Since models C1-1 and C1a1 with the SUMASLS form assume possible effect of AS on OFFS and of LS on ONS, it appears better able to reflect all possible crowding effects. Based on this viewpoint, the SUMASLS form is most preferred among these model forms.

(4) The constant in all models (about 10 seconds) was reasonable, since dwell time always includes some time for the doors of the train to open and close, and some time for passengers who may want to alight, even if there is no one waiting to board the train.

(5) In One_M1 models, it appears appropriate to choose a linear form for the variables MAXASLS, and SUMASLS reflecting crowding effect in the models B1-1, and C1-1. In model C1-3 with the ONS*LS^E form and model D1-3 with the LS form, a nonlinear form with an exponent of 2.5 for the variable LS significantly improves the explanatory power, implying that the on board crowding term is more related to departing passenger load with an exponent of 2.5. Similarly, in One_M1a model, the model C1a3 with OFFAS form with E in the range of 1.2 to 1.8 and that with OFFS*AS^E form with E of 1.6-2.7 do offer better explanations. It also appears that model D1a3 with nonlinear form of LS with E in the range of 2.5-3.0 offers maximum improvement over the linear form. While in One_M1b models, model C1b3 with E in the range of 1.4 to 2.8, and model D1b3 with E in the range of 1.2-2.7 offer highest explanatory power. These results are shown below:

Model C1-3:

$DT = 11.43 + 0.69*ONS + 0.48*OFFS + 1.35*10^{-5}*ONS*LS^{2.5}$, $(R^2=0.65)$
(8.78) (5.38) (4.99) (7.41) (DW=2.16)
[0.00] [0.00] [0.00] [0.00]

Model C1a3(A):

$DT = 12.0 + 0.83*ONS + 6.32*10^{-4}*OFFAS^{1.5}$, $(R^2=0.68)$
(6.52) (4.81) (9.13) (DW=1.90)
[0.00] [0.00] [0.00]

Model C1a3(B):

$DT = 11.58 + 0.85*ONS + 2.44*10^{-4}*OFFS*AS^{2.0}$, $(R^2=0.69)$
(6.42) (5.08) (9.38) (DW=1.91)
[0.00] [0.00] [0.00]

Model C1b3:

$DT = 11.85 + 0.63*ONS + 0.48*OFFS + 7.7*10^{-5}*ONS*LS^{2.0}$, $(R^2=0.67)$
(8.73) (2.78) (4.30) (1.94) (DW=2.19)
[0.00] [0.00] [0.00] [0.06]

Model D1-3:

$$DT=10.05+0.78*ONS+0.50*OFFS+2.0*10^{-4}*LS^{2.5}, \quad (R^2=0.68)$$
(8.32) (6.70) (5.51) (8.50) (DW=2.10)
[0.00] [0.00] [0.00] [0.00]

Model D1a3:

$$DT=9.71+0.94*ONS+1.1*10^{-4}*LS^{2.7}, \quad (R^2=0.69)$$
(5.44) (5.69) (9.34) (DW=2.05)
[0.00] [0.00] [0.00]

Model D1b3:

$$DT=11.46+0.66*ONS+0.49*OFFS+7.7*10^{-4}*LS^{2.0}, \quad (R^2=0.68)$$
(8.47) (3.17) (4.46) (2.26) (DW=2.21)
[0.00] [0.00] [0.00] [0.03]

(6) It appears that the desirable models do not violate the assumptions of the multiple linear regression model. Therefore, these models look promising as dwell time functions. These recommended models explain between 60% and 70% of the variation in dwell times, implying that some factors affecting the dwell time were not included, most importantly operator behavior and passenger characteristics, however, the most significant factors have been captured in these model forms.

3.6 Two-Car Train Models

As discussed in Chapter 2, initially, three different approaches were used to build two-car train dwell time models. These approaches were:

1. DT = f(sum of ONS, sum of OFFS for that train,

(sum of ONOFFS)*Standees for that train)

where DT is the longer dwell time (LDT) car for that train

2. DT for each car = f(sum of ONS, sum of OFFS for that car, (sum of ONOFFS)*Standees for that car)

where DT is the dwell time for car 1 and car 2 of that train

3. DT for car having longer dwell time (LDT) = f(ONS and OFFS for LDT car, (sum of ONOFFS)*standees for LDT car)

The first approach was based on the theory that the dwell time for a train depended on the total numbers of passengers boarding, and alighting, and the effect of crowding on board, treating the train as an entity.

The second approach dealt with a two-car train as two single cars, and found the relationships between the dwell time and total boarding passengers, total alighting passengers, and crowding effect, for each car separately.

The third approach was based on the theory that the dwell time for a two-car train is the longer of the dwell time for car 1 and car 2 of that train, i.e.

DT = MAX{Dwell Time for car 1, Dwell time for car 2},

so that the dwell time for a two-car train is taken as the longer dwell time (LDT) of the two dwell times (one for each car). So, it was expected that the number of passengers boarding and alighting and crowding level of the car with LDT would be more related to the observed dwell time.

Apart from the above approaches, alternative models may be estimated based on dividing the data into two sets: the data set with ONS being equal to, or greater than, OFFS (ONS≥OFFS) and that with OFFS being greater than ONS (OFFS>ONS). 32 and 19 sample points are obtained for the former group (ONS≥OFFS) and latter group (OFFS>ONS) respectively, which enable the following two alternative models to be estimated adopting the first approach above:

1a. DT = f(sum of ONS, sum of OFFS for that train,
ONOFFS * Standees for that train),
where (ONS≥OFFS)

1b. DT = f(sum of ONS, sum of OFFS for that train,
ONOFFS * Standees for that train),
where (OFFS>ONS)

In the following analysis, models were estimated under these five approaches, named Two_M1, Two_M1a, Two_M1b, Two_M2, and Two_M3 models in the subsequent sections. Similar to the one-car models, the statistical packages SST (1986) and MINITAB (1989) were used for the regression analysis, as well as for checking the results.

3.6.1 Correlation Analysis

Appendix A (Figures A.11 to A.20) presents the relationships between the dwell time and the key potential independent variables for the data set used in the Two_M3 models, all of which appear to show some correlation. The correlation coefficients (r) between all variables in these five models are shown in Appendix B (Tables B.4 to B.8). It is clear from Tables B.4 to B.8 that each variable is highly correlated with the dependent variable (DT) as well as with some other independent variables. As section 3.5.1 discussed, in

multivariate linear regression estimation, it is desirable to avoid strong correlation between independent variables. The high est correlations among the independent variables are between those reflecting crowding, as with the one-car data set, suggesting that these variables should not be included together in model formulations.

3.6.2 Model A: DT = f(ONS, OFFS)

This model assumes that only the numbers of passenger boarding and alighting affect the dwell time, so there is no effect of passenger crowding on board. The resulting models based on the five approaches discussed in section 3.6 are expressed by 1, 1a, 1b, 2, and 3 respectively in this (and subsequent) specifications, producing the following results:

Model A1:

$$DT = 11.73 + 0.42 \cdot ONS + 0.49 \cdot OFFS, \quad (R^2 = 0.68)$$
$$(7.44) \quad (7.59) \quad (6.22) \quad (DW = 2.08)$$
$$[0.00] \quad [0.00] \quad [0.00]$$

As indicated by the t-statistics and p-values, all coefficients are strongly significant in model A1, with a corrected R^2 of 0.68. The coefficient of the variable ONS is close to that of the variable OFFS implying that the marginal dwell time for boarding is about equal to that for alighting. The constant for the model is about 12 seconds.

Compared with the corresponding one-car train model, the constant term is about 3 seconds greater, while the coefficients for the variables ONS and OFFS are lower, implying that the marginal dwell times for boarding and alighting are lower in aggregated two-car data. It also appears that this two-car model better explains the dwell times using two variables ONS and OFFS only, than the one-car model, implying that the crowding effect is more significant in the one-car data set.

Model A1a:

$$DT=9.69+0.42*ONS+0.66*OFFS, \quad (R^2=0.71)$$
$$(4.32) \quad (4.49) \quad (3.99) \quad (DW=2.17)$$
$$[0.00] \quad [0.00] \quad [0.00]$$

Similar to model A1, in model A1a, in which ONS≥OFFS, all coefficients are strongly significant, with R^2 of 0.71. The coefficient of the variables ONS and OFFS is 0.42 and 0.66 respectively implying that the marginal dwell time for boarding is about 0.63 times that for alighting. The constant for this model is about 10 seconds.

Compared with the corresponding one car model, the constant term is slightly greater with lower marginal dwell time for boarding and alighting. Similar to A1 model, this two-car model better explains dwell times using the variables ONS and OFFS only.

Model A1b:

$$DT=14.31+0.13*ONS+0.50*OFFS, \quad (R^2=0.67)$$
$$(7.38) \quad (0.76) \quad (3.97) \quad (DW=2.24)$$
$$[0.00] \quad [0.46] \quad [0.00]$$

In model A1b, in which OFFS>ONS, as indicated by the t-statistics and p-values, the coefficient of the variable ONS is insignificant implying that the marginal dwell time for boarding is negligible. Consequently, the variable ONS is dropped from this (and subsequent) models, producing the following model A1bR:

Model A1bR:

$$DT=14.39+0.56*OFFS, \quad (R^2=0.68)$$
$$(7.46) \quad (6.29) \quad (DW=2.08)$$
$$[0.00] \quad [0.00]$$

All coefficients are strongly significant and the corrected R^2 is high in model A1bR which implies that the dwell time is more strongly related to alighting passengers than to

boarding passengers. In this model, the constant is about 14 seconds, with the marginal dwell time for alighting being about 0.56 seconds.

Compared with the corresponding one car train model, it is clear that the constant term is greater and the marginal dwell time for boarding is negligible in this two-car model.

Also, it is clear from these results, as indicated by the R^2, that models A1a and A1bR using two data sets based on the relative magnitude of ONS and OFFS are a slight improvement over model A1.

Model A2:

```
DT=13.27+0.70*ONS+0.56*OFFS,  (R2=0.57)
  (12.39)  (8.87)     (5.76)    (DW=1.51)
  [0.00]   [0.00]     [0.00]
```

In model A2, all coefficients are strongly significant, but the R^2 is low compared to models A1, A1a, and A1bR. The coefficient of the variables ONS and OFFS is 0.70 and 0.56 which implies that the marginal dwell time for boarding is about 1.3 times that for alighting. The constant is about 13 seconds.

Model A3:

```
DT=14.37+0.73*ONS+0.56*OFFS,  (R2=0.65)
   (9.85)   (7.88)    (4.73)    (DW=1.92)
   [0.00]   [0.00]    [0.00]
```

In model A3, all coefficients are strongly significant, as indicated by t-statistics and p-values, and with an R^2 of 0.65. The results show that the coefficient of the variable ONS is about 1.3 times that for alighting, implying that the marginal dwell time for boarding is about 1.3 times that for alighting. The constant is about 14 seconds.

3.6.3 Model B: DT = f(ONS, OFFS, ONOFFS*Standees)

Compared to model A, this model recognizes the effect on dwell time of crowding on board as well as the number of passengers boarding and alighting. The crowding effect was examined using three alternative variables: MAXASLS, ABLS, and ABAS, assuming movement of passenger boarding and alighting would be affected by arriving standees, or departing standees. Model estimations using these three alternative variables produce the following results:

Model B1:

B1-1	**DT=14.04+0.26*ONS+0.36*OFFS+0.0008*MAXASLS,**				**(R^2=0.70)**
	(7.43)	**(2.78)**	**(3.71)**	**(2.07)**	**(DW=2.15)**
	[0.00]	**[0.01]**	**[0.00]**	**[0.04]**	
B1-2	**DT=13.65+0.30*ONS+0.36*OFFS+0.0007*ABAS,**				**(R^2=0.70)**
	(7.40)	**(3.67)**	**(3.61)**	**(1.88)**	**(DW=2.16)**
	[0.00]	**[0.00]**	**[0.00]**	**[0.07]**	
B1-3	**DT=13.89+0.26*ONS+0.38*OFFS+0.0008*ABLS,**				**(R^2=0.70)**
	(7.44)	**(2.77)**	**(4.15)**	**(2.02)**	**(DW=2.15)**
	[0.00]	**[0.01]**	**[0.00]**	**[0.05]**	

As indicated by the R^2, it is clear that adding any of these three variables to reflect the effect of crowding on board does improve the explanatory power of the model. There also appears to be little to choose between the MAXASLS, ABAS and ABLS forms, although the MAXASLS and ABLS forms are slightly preferred since all coefficients are significant. The coefficients of the variables ONS, OFFS and MAXASLS are 0.26, 0.36 and 0.0008 in the MAXASLS form implying that the marginal dwell time for boarding is about 0.7 times that for alighting, and about 320 times that for MAXASLS.

Compared with the corresponding one car train model, the constant term is slightly greater with smaller marginal dwell time for boarding and greater marginal dwell time for alighting. The coefficients for the variables reflecting effects of passenger crowding is about one-tenth that for a one-car model.

As in the one car case, only single model based on MAXASLS are presented here:

Model B1a:

DT=11.38+0.34*ONS+0.52*OFFS+0.0005*MAXASLS, (R^2=0.70)
(3.79) (2.52) (2.23) (0.85) (DW=2.22)
[0.00] [0.02] [0.03] [0.40]

In model B1a, the coefficient of variable MAXASLS is insignificant at the 0.05 level, with an R^2 of 0.70. Since the coefficient of the crowding variables is insignificant, compared to model A1a, it is clear that adding MAXASLS to reflect the effect of crowding on board doesn't improve the explanatory power of the model.

Model B1b:

DT=15.75+0.40*OFFS+0.0008*MAXASLS, (R^2=0.72)
(8.11) (3.39) (1.90) (DW=2.46)
[0.00] [0.00] [0.08]

As indicated by the R^2, it is clear from the model B1b that adding MAXASLS to reflect the effect of crowding on board is a slight improvement over model A1b. The coefficient of the variable OFFS and MAXASLS is about 0.40 and 0.0008 respectively, which implies that the marginal dwell time for alighting is about 500 times greater than the crowding variable MAXASLS. The constant for this model is about 15 seconds.

Compared with the corresponding one car train model, the constant term is greater with similar marginal dwell time for alighting. While the marginal dwell time for crowding is insignificant in the one-car model, that for boarding is insignificant in this two-car model.

Model B2:

B2-1 DT=15.34+0.33*ONS+0.30*OFFS+0.004*MAXASLS, (R^2=0.63)
(13.72) (2.77) (2.72) (4.06) (DW=1.72)
[0.00] [0.01] [0.01] [0.00]

```
B2-2 DT=14.95+0.42*ONS+0.29*OFFS+0.004*ABAS,  (R2=0.63)
       (13.72)  (4.08)    (2.54)     (3.91)    (DW=1.74)
       [0.00]   [0.00]    [0.01]     [0.00]

B2-3 DT=15.20+0.32*ONS+0.37*OFFS+0.004*ABLS,  (R2=0.63)
       (13.73)  (2.67)    (3.59)     (4.02)    (DW=1.72)
       [0.00]   [0.01]    [0.00]     [0.00]
```

In model B2, all coefficients are strongly significant, as indicated by the t-statistics and p-values, with an R^2 of 0.63. Compared to model A2, it is clear that adding any of these three variables to reflect the effect of crowding on board significantly improves the explanatory power of the model. There also appears to be little to choose between the MAXASLS, ABAS, and ABLS forms.

The constant is larger in model B2 than in A2, with a lower marginal dwell time for boarding and alighting. The coefficient of the variables OFFS is about 0.30, which is about 80 times that of the variable MAXASLS and also somewhat below the coefficient for ONS.

Model B3:

```
B3-1 DT=16.14+0.47*ONS+0.37*OFFS+0.0027*MAXASLS,  (R2=0.67)
       (9.87)   (3.12)    (2.59)      (2.13)       (DW=1.89)
       [0.00]   [0.00]    [0.01]      [0.00]

B3-2 DT=15.79+0.55*ONS+0.37*OFFS+0.0026*ABAS,  (R2=0.67)
       (10.0)   (4.22)    (2.52)      (2.03)    (DW=1.92)
       [0.00]   [0.00]    [0.02]      [0.05]

B3-3 DT=16.01+0.47*ONS+0.42*OFFS+0.0026*ABLS,  (R2=0.67)
       (9.98)   (3.06)    (3.20)      (2.13)    (DW=1.88)
       [0.00]   [0.00]    [0.00]      [0.04]
```

In model B3, all coefficients are significant at the 0.05 level, as indicated by the t-statistics and p-values. Similar to model B2, adding MAXASLS, ABAS or ABLS to express the effect of crowding on board does improve the explanatory power of the model. There also appears to be little to choose between the MAXASLS, ABAS and ABLS forms.

The constant is larger in model B3 than in A3, with a lower marginal dwell time for boarding and alighting. The coefficients of the variables ONS, OFFS, and MAXASLS are 0.47, 0.37, and 0.0027 respectively, which implies that the marginal dwell time for boarding is about 1.3 times that for alighting, as well as 170 times that for MAXASLS.

3.6.4 Model C: DT = f(ONS,OFFS,ONLS,OFFAS)

Once again, like Model C for one-car trains, this model recognizes that movement of alighting passengers would be affected by arriving standees, while movement of boarding passengers would be affected by departing standees. Therefore, the crowding effect may be examined in three alternative variables: OFFAS, ONLS, and SUMASLS, and can be represented in several different forms. Similar to model B, this model assumes the effect on dwell time of crowding on board as well as the number of passengers boarding and alighting. The first form of this model introduces the variable SUMASLS to express the marginal effect on dwell time of crowding on board, which is the sum of OFFAS and ONLS:

Model C1-1:

```
DT=13.93+0.27*ONS+0.36*OFFS+0.0008*SUMASLS,  (R2=0.70)
   (7.43)   (2.92)   (3.79)       (2.03)        (DW=2.15)
   [0.00]   [0.01]   [0.00]       [0.05]
```

All coefficients are significant at the 0.05 level, as indicated by the t-statistics and p-values. Similar to model B1, adding the variable SUMASLS in model C1-1 is a slight improvement over model A1. The coefficients of the variables ONS, OFFS, and SUMASLS are 0.27, 0.36, and 0.0008 implying that the marginal dwell time for boarding is about 0.7 times that for alighting, as well as about 320 times that for the crowding variable SUMASLS. It is clear that all coefficients are almost identical with those for the MAXASLS form of model B1.

Compared with the corresponding one car train model, the constant term is slightly greater with smaller marginal dwell time for boarding and greater marginal dwell time for alighting. Since the coefficient of the variable SUMASLS in the two-car model is only one-sixth of that in the one-car model implying once again that the crowding effect is less significant in two-car train dwell times.

Model C1a1:

$$DT = 11.31 + 0.34*ONS + 0.52*OFFS + 0.0005*SUMASLS, \quad (R^2=0.70)$$

(3.83)	(2.62)	(2.23)	(0.85)	(DW=2.22)
[0.00]	[0.01]	[0.03]	[0.40]	

As indicated by the t-statistics and p-values, the coefficient of the variable SUMASLS is insignificant in model C1a1. Once again, similar to model B1a, adding the crowding variable SUMASLS doesn't improve the explanatory power of the model since the coefficient of the variable SUMASLS is insignificant at the 0.05 level.

Compared with the corresponding one car train model, the constant term is similar, but the marginal dwell time for boarding is smaller. The marginal dwell time for the crowding effect is insignificant in this two-car model while that for alighting is insignificant in the one-car model.

Model C1b1:

$$DT=15.69+0.41*OFFS+0.0008*SUMASLS, \quad (R^2=0.72)$$
$$(8.10) \quad (3.50) \quad (1.88) \quad (DW=2.46)$$
$$[0.00] \quad [0.00] \quad [0.08]$$

In model C1b1, all coefficients are significant at 0.10 level. Compared with model A1bR, it is clear that adding the crowding variable SUMASLS in model C1b1 significantly improves the explanatory power of the model. The coefficients of the variables OFFS and SUMASLS are about 0.43 and 0.0008 implying that the marginal dwell time for boarding is about 540 times greater than that for SUMASLS.

Similar to models C1-1 and C1a1, the constant term is greater than that in the corresponding one car train model. The marginal dwell time for alighting is similar in the one and two-car models, but the marginal dwell time for boarding is insignificant in the two-car model while that for the crowding effect is insignificant in the one-car model.

Model C2-1:

$$DT=15.25+0.33*ONS+0.31*OFFS+0.004*SUMASLS, \quad (R^2=0.63)$$
$$(13.79) \quad (2.91) \quad (2.84) \quad (4.10) \quad (DW=1.73)$$
$$[0.00] \quad [0.00] \quad [0.01] \quad [0.00]$$

As indicated by the t-statistics and p-values, all coefficients are strongly significant in model C2-1 with an R^2 of 0.63. Compared to model A2, it is clear that adding the crowding variable SUMASLS significantly improves the explanatory power of the model. The coefficients of variables ONS, OFFS, and SUMASLS are 0.33, 0.31, and 0.004 which implies that the marginal dwell time for boarding is about the same as that for alighting, and is about 80 times greater than the crowding variable SUMASLS. The constant is about 15 seconds in this model.

Model C3-1:

$$DT = 16.08 + 0.48*ONS + 0.38*OFFS + 0.0028*SUMASLS, \quad (R^2=0.68)$$
$$(9.98) \quad (3.25) \quad (2.68) \quad (2.17) \quad (DW=1.89)$$
$$[0.00] \quad [0.00] \quad [0.01] \quad [0.04]$$

Similar to model C2-1, all coefficients are significant at the 0.05 level in model C3-1 with an R^2 of 0.68. As indicated by the R^2, it is clear that adding the crowding variable SUMASLS does improve the explanatory power over model A3. The coefficients of variables ONS, OFFS, and SUMASLS are 0.48, 0.38, and 0.0028 respectively implying the marginal dwell time for boarding is about 1.3 times that for alighting, as well as 140 times greater than the crowding variable SUMASLS.

The second form of this model presents the separate effects of crowding on alighting and boarding passengers with two variables OFFAS and ONLS. The variable OFFAS represents the interaction between alightings and arriving standees while ONLS reflects the interaction between boardings and departing standees. Models with OFFAS and ONLS are as follows:

Model C1-2:

$$DT = 13.99 + 0.23*ONS + 0.41*OFFS + 0.0003*OFFAS$$
$$(7.39) \quad (1.92) \quad (3.16) \quad (0.23)$$
$$[0.00] \quad [0.06] \quad [0.00] \quad [0.82]$$

$$+0.001*ONLS, \quad (R^2=0.70)$$
$$(1.38) \quad (DW=2.13)$$
$$[0.18]$$

Model C1a2:

$$DT = 11.15 + 0.39*ONS + 0.45*OFFS + 0.0013*OFFAS + 0.00014*ONLS, \quad (R^2=0.69)$$
$$(3.63) \quad (1.84) \quad (1.28) \quad (0.43) \quad (0.092) \quad (DW=2.19)$$
$$[0.00] \quad [0.08] \quad [0.21] \quad [0.67] \quad [0.93]$$

Model C1b2:

$$DT = 15.88 + 0.38*OFFS + 0.0016*OFFAS - 0.0001*ONLS, \quad (R^2=0.71)$$
$$(7.88) \quad (2.90) \quad (1.06) \quad (-0.06) \quad (DW=2.44)$$
$$[0.00] \quad [0.01] \quad [0.31] \quad [0.95]$$

Model C2-2:

$$DT = 15.28 + 0.32*ONS + 0.33*OFFS + 0.004*OFFAS + 0.005*ONLS, \quad (R^2=0.63)$$
$$(13.60) \quad (2.05) \quad (2.10) \quad (1.29) \quad (2.10) \quad (DW=1.73)$$
$$[0.00] \quad [0.04] \quad [0.04] \quad [0.20] \quad [0.04]$$

Model C3-2:

$$DT = 16.21 + 0.42*ONS + 0.44*OFFS + 0.001*OFFAS + 0.004*ONLS, \quad (R^2=0.67)$$
$$(9.84) \quad (2.07) \quad (2.23) \quad (0.30) \quad (1.45) \quad (DW=1.87)$$
$$[0.00] \quad [0.04] \quad [0.03] \quad [0.76] \quad [0.16]$$

As indicated by the t-statistics and p-values, the coefficients of variables ONLS and OFFAS are insignificant in models C1-2, C1a2, C1b2, and C3-2 while that of OFFAS is insignificant in model C2-2. These results may be due to multicollinearity in the variables:

the correlation coefficients are between 0.58 and 0.90 for variables OFFAS and ONLS, 0.75 to 0.88 for variables OFFS and OFFAS, and between 0.85 and 0.87 for variables ONS and ONLS (see Tables B.4-B.8). The high correlation between variables OFFAS and ONLS which both reflect crowding, suggests that these variables should not be included together in model formulations. One way of dealing with this problem is to introduce either variable ONLS or OFFAS in any model formulation.

The variable ONLS or OFFAS is introduced in the third form of this model which assumes that departing standees affect passenger boarding, and arriving standees affect passengers alighting with the following results:

Model C1-3:

$$DT = 13.92 + 0.22*ONS + 0.43*OFFS + 0.0013*ONLS, \quad (R^2=0.70)$$

(7.51)	(2.00)	(5.37)	(2.08)	(DW=2.12)
[0.00]	[0.05]	[0.00]	[0.04]	

In model C1-3, all coefficients are significant at 0.05 level, as indicated by the t-statistics and p-values. Compared with model A1, as indicated by the R^2, it is clear that adding the crowding variable ONLS does improve the explanatory power. The coefficients of variables ONS, OFFS, and ONLS are 0.22, 0.43, and 0.0013 respectively, which implies the marginal dwell time for boarding is about half that for alighting, and 140 times greater than the crowding variable ONLS.

Compared with the corresponding one car train model, the constant item is slightly greater, but the marginal dwell times are smaller for boarding and greater for alighting. The coefficient of variable ONLS is about one-tenth that in the one-car model implying that the marginal dwell time effect of the variable ONLS is less significant in the two-car model.

Model C1a3:

$DT=11.24+0.33*ONS+0.58*OFFS+0.00067*ONLS$, $(R^2=0.70)$
(3.72) (2.12) (2.91) (0.78) (DW=2.24)
[0.00] [0.04] [0.01] [0.44]

As indicated by the t-statistics and p-values, the coefficient of variable ONLS is insignificant implying that the marginal dwell time effect of variable ONLS is negligible. The R^2 also shows that adding the crowding variable ONLS doesn't result in any improvement over model A1a.

In model C1b2, the coefficient of variable ONLS is both insignificant and negative, implying that marginal dwell time is negatively related to the ONLS which is not reasonable. Thus, the variable ONLS is dropped from model C1b2 to produce the following result:

Model C1b3:

$DT=15.87+0.38*OFFS+0.0015*OFFAS$, $(R^2=0.73)$
(8.16) (3.11) (1.97) (DW=2.44)
[0.00] [0.00] [0.07]

Compared with model A1b, it is clear from model C1b3 that adding the crowding variable OFFAS does improve the explanatory power of the model. The coefficients of variable OFFS and OFFAS are about 0.38 and 0.0015 implying that the marginal dwell time for alighting is about 250 times greater than the crowding variable OFFAS.

Model C2-3:

$DT=15.25+0.24*ONS+0.49*OFFS+0.0064*ONLS$, $(R^2=0.62)$
(13.54) (1.67) (5.24) (3.86) (DW=1.68)
[0.00] [0.10] [0.00] [0.00]

As indicated by the R^2, model C2-3 is a significant improvement over the model A2 when the crowding variable ONLS is added. The coefficients of variables ONS, OFFS, and ONLS are 0.24, 0.49, and 0.0064, which implies that the marginal dwell time for boarding is about half that for alighting, and is about 37 times greater than the crowding variable ONLS. However, compared with model C2-1, the SUMASLS form is preferred since all coefficients are significant at the 0.05 level.

Model C3-3:

```
DT=16.20+0.39*ONS+0.49*OFFS+0.0044*ONLS,  (R2=0.68)
  (9.93)   (2.17)    (4.19)       (2.21)     (DW=1.86)
  [0.00]   [0.04]    [0.00]       [0.03]
```

As indicated by the t-statistics and p-values, all coefficients are significant at the 0.05 level, with an R^2 of 0.68. Compared to model A3, it is clear from model C3-3 that adding variable ONLS to reflect the effect of crowding on board significantly improves the explanatory power of the model. The coefficients of variables ONS, OFFS, and ONLS are 0.39, 0.49, and 0.0044 respectively implying the marginal dwell time for boarding is about 0.8 times that for alighting, and about 87 times greater than the crowding variable ONLS.

3.6.5 Model D: DT = f(ONS,OFFS,AS,LS)

This model assumes that dwell time is affected by crowding on board as well as by the number of passenger boarding and alighting. The crowding effect was examined using two independent variables: AS and LS, with the following results:

Model D1-1:

```
DT=12.36+0.37*ONS+0.39*OFFS+0.05*AS-0.02*LS,  (R2=0.69)
  (7.61)   (2.44)    (2.43)    (0.31)  (-0.13)   (DW=2.12)
  [0.00]   [0.02]    [0.02]    [0.76]  [0.90]
```

Model D1a1:

$$DT=9.79+0.44*ONS+0.58*OFFS+0.043*AS-0.031*LS, \quad (R^2=0.69)$$
$$(3.90) \quad (1.93) \quad (1.85) \quad (0.18) \quad (-0.14) \quad (DW=2.19)$$
$$[0.00] \quad [0.06] \quad [0.08] \quad [0.86] \quad [0.89]$$

Model D1b1:

$$DT=15.18+0.41*OFFS+0.082*AS-0.045*LS, \quad (R^2=0.72)$$
$$(7.96) \quad (3.10) \quad (0.60) \quad (-0.33) \quad (DW=2.32)$$
$$[0.00] \quad [0.01] \quad [0.56] \quad [0.74]$$

Model D2-1:

$$DT=13.26+0.29*ONS+0.61*OFFS-0.18*AS+0.28*LS, \quad (R^2=0.65)$$
$$(13.24) \quad (1.49) \quad (3.11) \quad (-0.90) \quad (1.39) \quad (DW=1.86)$$
$$[0.00] \quad [0.14] \quad [0.01] \quad [0.37] \quad [0.17]$$

Model D3-1:

$$DT=14.40+0.39*ONS+0.58*OFFS-0.14*AS+0.23*LS, \quad (R^2=0.70)$$
$$(10.48) \quad (1.46) \quad (2.22) \quad (-0.52) \quad (0.85) \quad (DW=1.85)$$
$$[0.00] \quad [0.15] \quad [0.03] \quad [0.60] \quad [0.40]$$

All these models show insignificant and/or counter intuitive signs for the crowding coefficients which once again may result from multicollinearity between the variables; as shown in Tables B.4-B.8 the correlation coefficients between the variables AS and LS are between 0.96 and 0.99.

Because the coefficients of the variable LS is insignificant and negative in models D1-1, D1a1, and D1b1, it is dropped from these models, similarly, the variable AS is dropped from the models D2-1 and D3-1 to produce the following results:

Model D1-2:

$$DT=12.37+0.35*ONS+0.41*OFFS+0.027*AS, \quad (R^2=0.69)$$
$$(7.73) \quad (5.20) \quad (4.46) \quad (1.61) \quad (DW=2.11)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.11]$$

As indicated by the t-statistics and p-values, the coefficient of the variable AS is insignificant at 0.10 level implying that the variable AS may contribute little to explaining dwell time.

Compared with the corresponding one car train model, the constant term is about 3 seconds greater, but the coefficient for ONS, OFFS, and AS are smaller, implying that the marginal dwell times for boarding, alighting, and crowding effect on board are smaller in this model.

Model D1a2:

$$DT=9.90+0.41*ONS+0.60*OFFS+0.01*AS, \quad (R^2=0.70)$$
$$(4.21) \quad (3.98) \quad (2.67) \quad (0.36) \quad (DW=2.18)$$
$$[0.00] \quad [0.00] \quad [0.01] \quad [0.73]$$

Similar to model D1-2, the coefficient of the variable AS is insignificant, implying that the marginal dwell time contribution of the crowding term AS is negligible. As indicated by the R^2, it is also clear from the model D1a2 that adding the crowding variable doesn't improve the explanatory power of the model.

Model D1b2:

$$DT=15.00+0.43*OFFS+0.037*AS, \quad (R^2=0.74)$$
$$(8.43) \quad (4.23) \quad (2.11) \quad (DW=2.38)$$
$$[0.00] \quad [0.00] \quad [0.05]$$

As indicated by the t-statistics and p-values, all coefficients are significant at 0.05 level in model D1b2 with an R^2 of 0.74. Compared to model A1b, it is clear that adding the

variable AS to reflect the effect the crowding on board significantly improves the explanatory power of the model. The coefficients of variables OFFS and AS are 0.43 and 0.037 respectively implying that the marginal dwell time for alighting is about 11 times greater than the crowding variable AS.

The constant term is greater than in the D1a2 model. The coefficients for OFFS is similar for these two models, but that for AS is smaller in this two-car model than that of LS in the one-car model, once again, implying that the crowding effect is less significant in the two-car model.

Model D2B:

$$DT = 13.05 + 0.45*ONS + 0.46*OFFS + 0.099*LS, \quad (R^2=0.65)$$
$$(13.44) \quad (5.02) \quad (5.00) \quad (4.77) \quad (DW=1.89)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.00]$$

As indicated by the t-statistics and p-values, all coefficients are strongly significant in model D2-2, with an R^2 of 0.65. Compared with model A2, it is clear that adding the crowding variable LS is a significant improvement. The coefficients of variables ONS, OFFS, and LS are 0.45, 0.46, and 0.099, which implies that the marginal dwell time for boarding is almost the same as that for alighting, and is about 45 times greater than the crowding variable LS.

Model D3-2:

$$DT = 14.29 + 0.52*ONS + 0.46*OFFS + 0.09*LS, \quad (R^2=0.70)$$
$$(10.62) \quad (4.65) \quad (4.02) \quad (3.08) \quad (DW=1.94)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.00]$$

In model D3-2, all coefficients are strongly significant, as indicated by the t-statistics and p-values. As indicated by the R^2 value, adding the crowding variable LS is a significant improvement over model A3. The coefficients of variables ONS, OFFS, and LS

are 0.52, 0.46, and 0.09 respectively, implying that the marginal dwell time for boarding is about 1.1 times that for alighting, as well as about 6 times that for LS.

In conclusion, as indicated by previous analysis by the corrected R^2 value, it is clear that the Two_M1, Two_M1a, and Two_M1b models offer better explanatory power than Two_M2 and Two_M3 models. Thus, in the subsequent sections, further analysis will be concentrate only on these three models.

Finally, the equality of individual parameters from two data sets 1a and 1b (ONS≥OFFS, and OFFS>ONS) was examined by t-statistics again. Table 3.11 presents t-test results for model forms 1a and 1b which shows that the coefficient estimates are insignificantly different at the 0.05 level. Thus, it is concluded that there is not sufficient evidence to reject the hypothesis that the marginal effect of these variables on dwell time are equal for the two sets of data.

3.6.6 Model Estimation with Alternative Forms

The previous model estimations, assumed that the effect on dwell time of crowding on board is linear, however, it may well be nonlinear. To investigate this possibility, nonlinear forms for the variables reflecting crowding is introduced in the following model formulations. As discussed in sections 3.6.3-3.6.5, the crowding effect is examined through six alternative variables: MAXASLS, SUMASLS, ONLS, OFFAS, AS, and LS. Two examples with the MAXASLS and ONS*LS^E forms from the promising models discussed in sections 3.6.3-3.6.5 are presented as follows:

$$DT = b_0 + b_1*ONS + b_2*OFFS + b_3*MAXASLS^E$$

$$DT = b_0 + b_1*ONS + b_2*OFFS + b_3*ONS*LS^E$$

Table 3.11 t test for Equality of Individual Coefficient between Two-Car Train Data Sets (t-statistics shown)

Variable	A1a & A1b models	B1a & B1b models	C1a1 & C1b1 models
Constant	-1.55	-1.22	-1.24
ONS	0.48	-	-
OFFS	0.77	0.46	0.42
MAXASLS	-	-0.41	-
SUMASLS	-	-	-0.41

Models were estimated are done by changing the value of the exponent (E) from 0.0 to 5.0, based on the five approaches discussed earlier in section 3.6; for the Two_M1, Two_M2, Two_M3, Two_M1a, and Two_M1b models.

A. Two_M1 Model:

Figures 3.17 to 3.21 present R^2 as a function of the exponent of the variables MAXASLS, SUMASLS, ONLS, and AS for model estimations based on the approaches discussed earlier. Figure 3.17 shows that R^2 is highest over the range of E from 1.2 to 5.0 in model B1-1 with the MAXASLS form, with the comparable ranges being 1.5-5.0, 2.0-3.5, 1.5-5.0, and 2.0-5.0 for models with the SUMASLS, ONLS, ONS*LS^E, and AS variables respectively (see Figures 3.18-3.21). The wide ranges for E imply that different exponential forms of the variables can be applied in each model. Examples of these models with nonlinear crowding variables are shown below:

Model B1-1:

$$DT = \underset{(7.54)}{14.09} + \underset{(3.06)}{0.27}*ONS + \underset{(4.01)}{0.37}*OFFS + \underset{(2.16)}{1.2*10^{-4}}*MAXASLS^{1.2}, \quad (R^2=0.71)$$
$$[0.00] \quad [0.01] \quad [0.00] \quad [0.04] \quad (DW=2.16)$$

Model C1-1:

$$DT = \underset{(7.71)}{13.88} + \underset{(3.73)}{0.29}*ONS + \underset{(4.48)}{0.39}*OFFS + \underset{(2.21)}{7.2*10^{-6}}*SUMASLS^{1.5}, \quad (R^2=0.71)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.03] \quad (DW=2.17)$$

Model C1-3(A):

$$DT = \underset{(8.14)}{13.79} + \underset{(3.49)}{0.27}*ONS + \underset{(5.92)}{0.45}*OFFS + \underset{(2.57)}{1.6*10^{-7}}*ONLS^2, \quad (R^2=0.72)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.02] \quad (DW=2.16)$$

Figure 3.17 R-SQUARE vs. E of MAXASLS

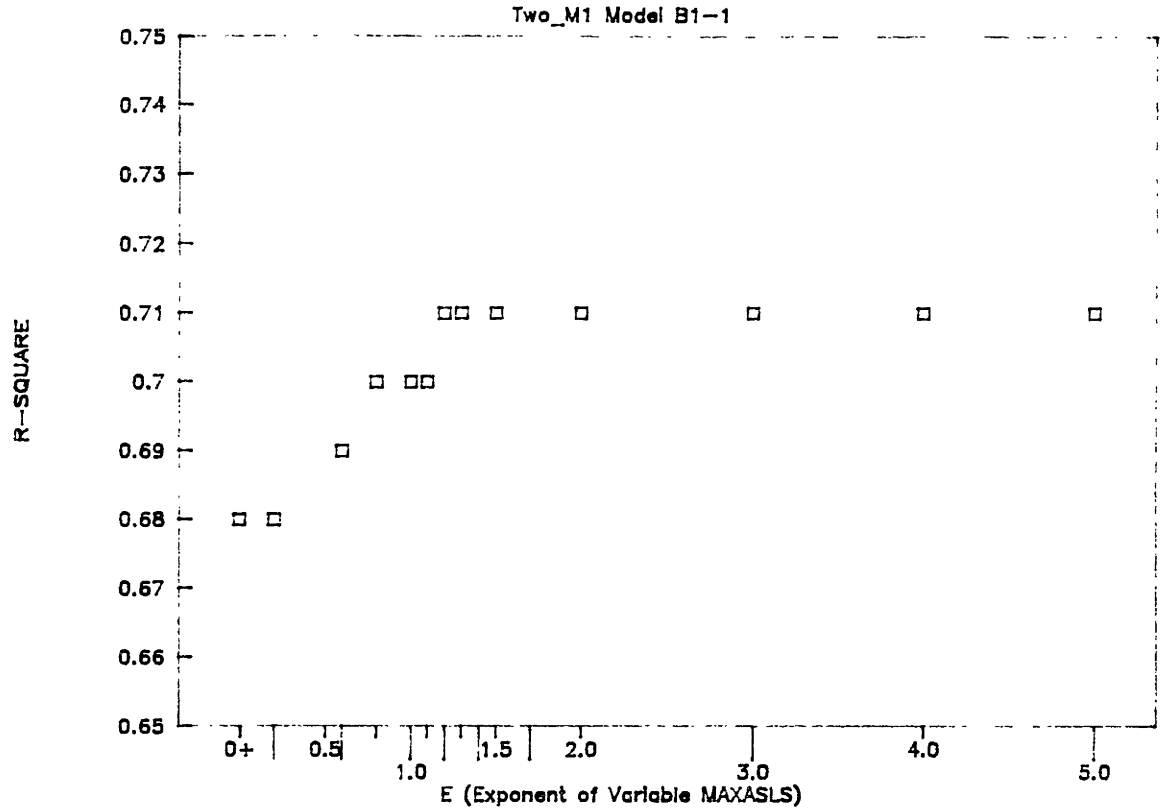


Figure 3.18 R-SQUARE vs. E of SUMASLS

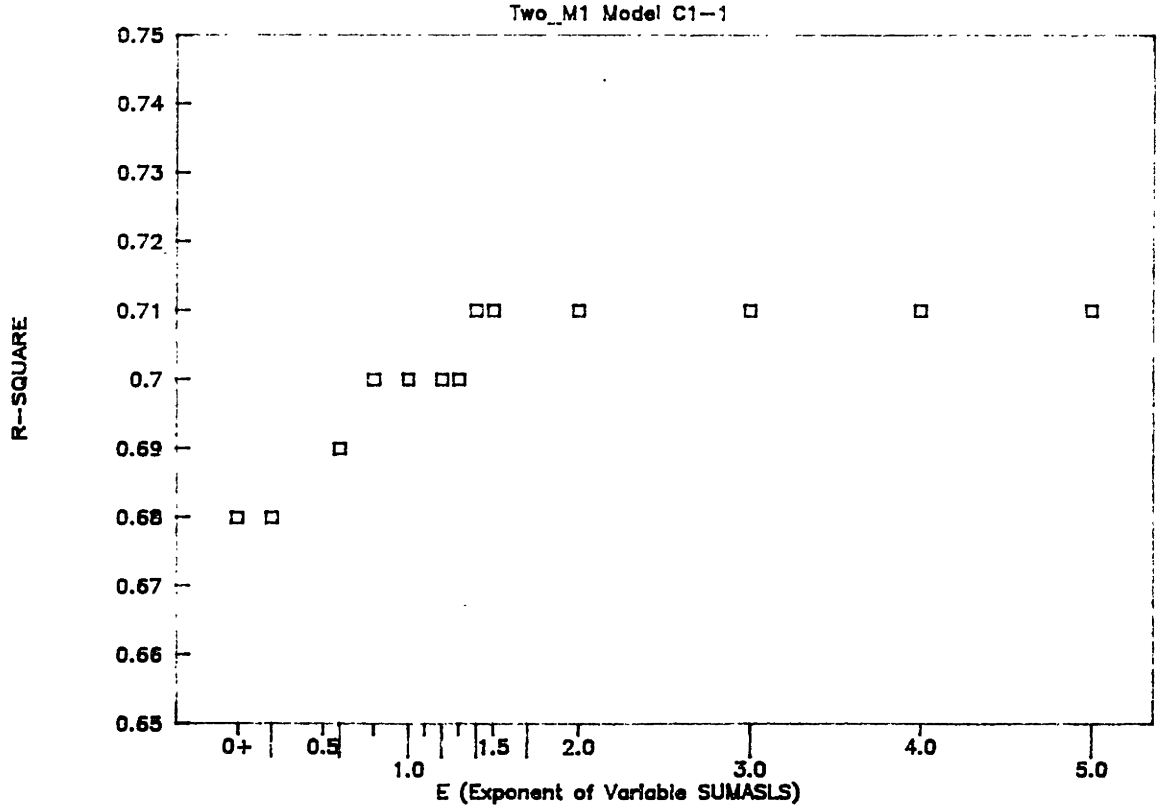


Figure 3.19 R-SQUARE vs. E of ONLS

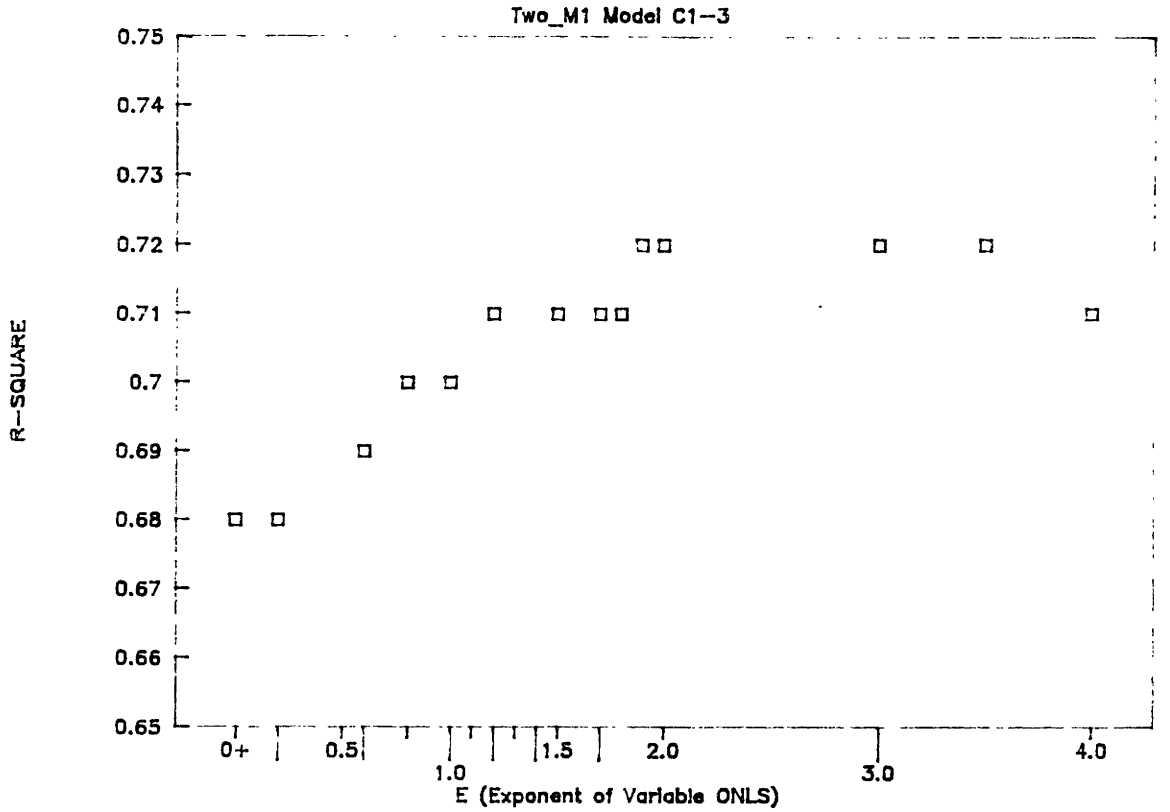


Figure 3.20 R-SQUARE vs. E of LS

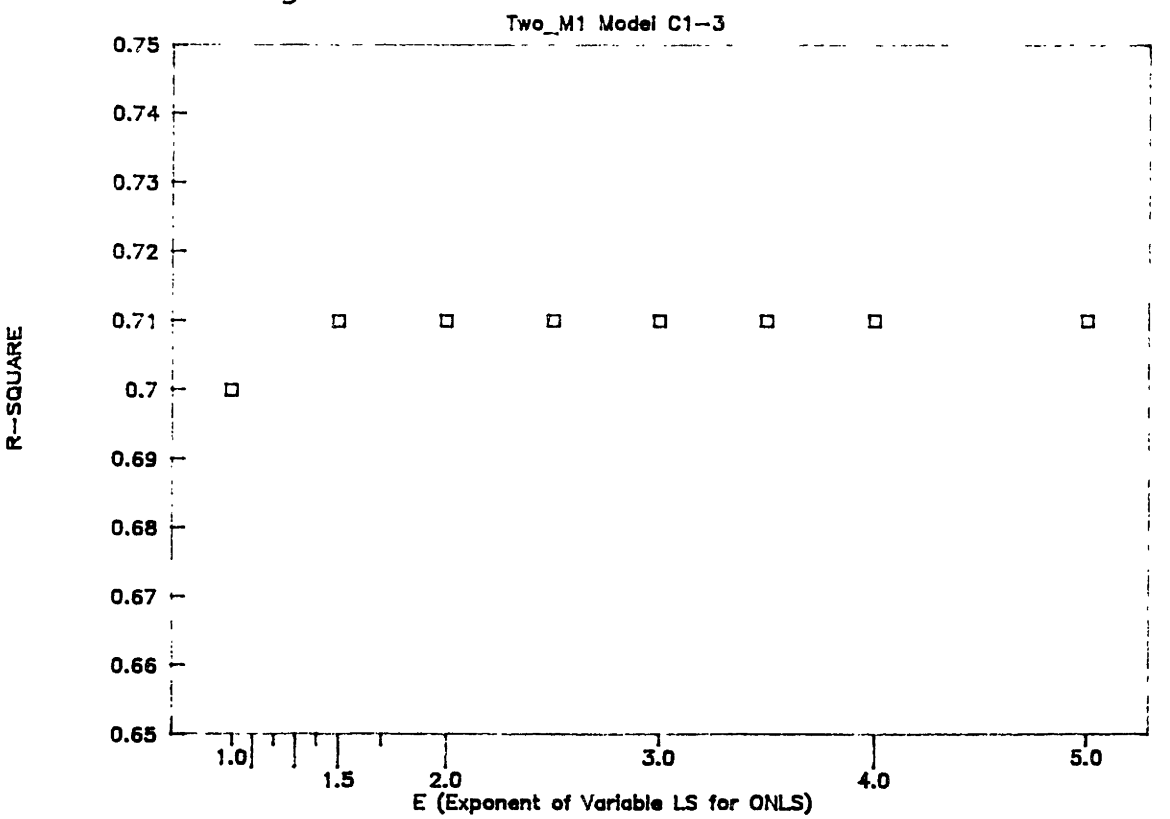


Figure 3.21 R-SQUARE vs. E of AS

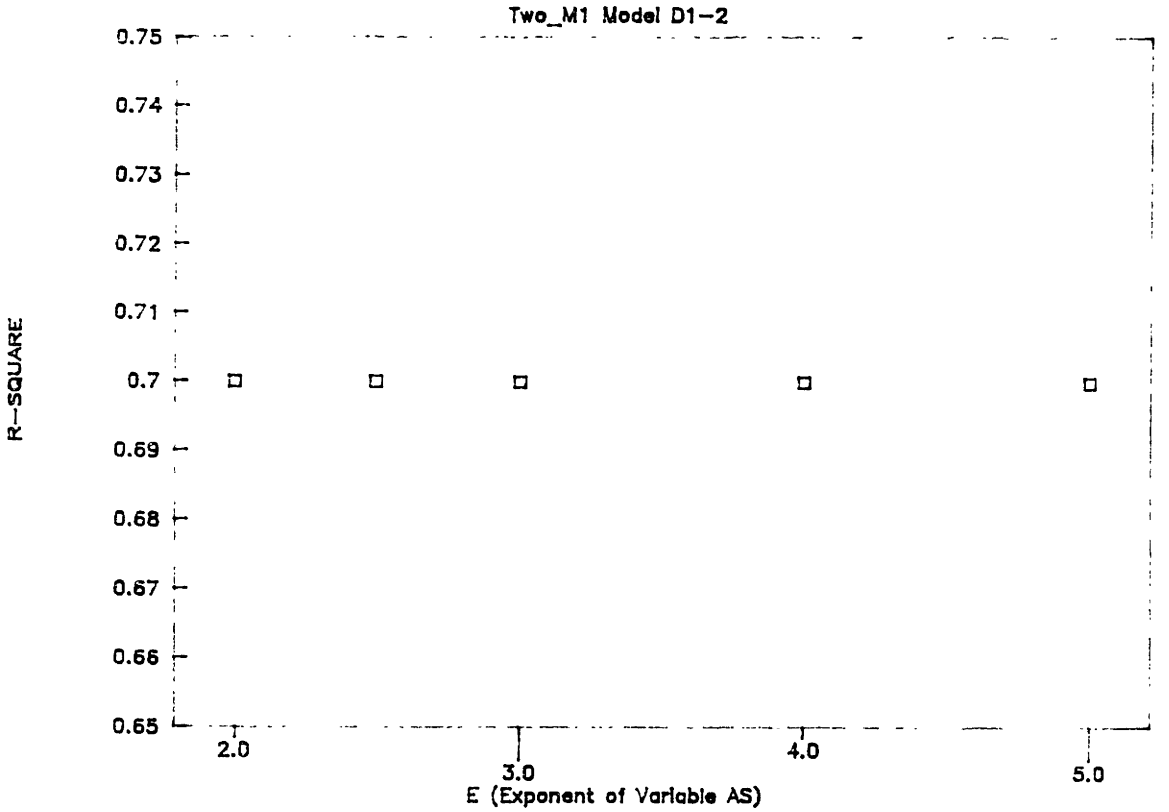
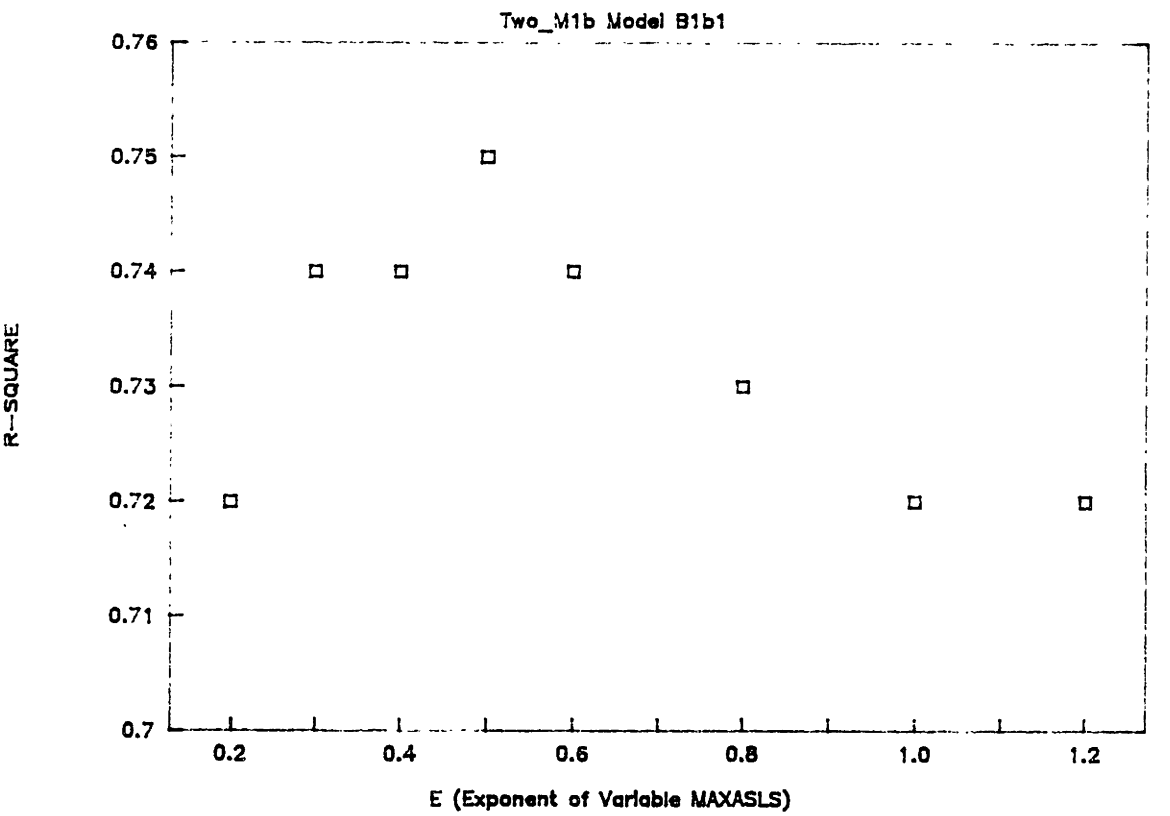


Figure 3.22 R-SQUARE vs. E of MAXASLS



Model C1-3(B):

$$DT=13.54+0.28*ONS+0.44*OFFS+6.0*10^{-6}*ONS*LS^{2}, \quad (R^2=0.71)$$
$$(8.06) \quad (3.70) \quad (5.65) \quad (2.41) \quad (DW=2.14)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.02]$$

Model D1-2:

$$DT=12.72+0.36*ONS+0.42*OFFS+1.3*10^{-6}*AS^{2.5}, \quad (R^2=0.70)$$
$$(7.94) \quad (6.08) \quad (5.01) \quad (2.03) \quad (DW=2.11)$$
$$[0.00] \quad [0.00] \quad [0.00] \quad [0.05]$$

Compared with the linear form with the variables MAXASLS, SUMASLS, ONLS, and AS reflecting crowding effect in the models B1-1, C1-1, C1-3, and D1-2, it is clear that the coefficients of the constant, and variables ONS and OFFS are very similar, while those for the crowding variables are different. In model D1-2, as indicated by the t-statistics and p-values, all coefficients are significant at the 0.05 level with a nonlinear form of AS with a range of E from 2.0 to 5.0, while they aren't at the same significance level with a linear AS form.

As indicated by the above results, it appears that the nonlinear form to reflect the effect of crowding is a slight improvement over the linear model. There also appears to be little to choose between these two model forms since the linear model is more easily interpreted.

B. Two_M1a Model:

The nonlinear form for four alternative variables, MAXASLS, SUMASLS, ONLS, and AS, are introduced in the model formulations with estimations carried out as before. The results from these models with a nonlinear form for the crowding variables are similar to those with the linear forms for the crowding variables, that the coefficient of the crowding variable insignificant at 0.05 level in models B1a1, C1a1, C1a3, and D1a2 implying that these crowding variables contribute little to the dwell time.

C. Two_M1b Model:

Figures 3.22 to 3.26 show R^2 as a function of the exponent of the variables MAXASLS, SUMASLS, OFFAS, and AS for model estimations based on approach 1b. As presented in Figure 3.22, R^2 is highest with E of 0.5 in model B1b1 with the MAXASLS form, similarly, compared with ranges of 0.4-0.5, 0.5, 0.6-0.7, and 0.6 in models C1b1, C1b3 and D1b2 with the SUMASLS, OFFAS, OFFS*AS^E, and AS variables respectively (see Figures 3.23 to 3.26). The wide ranges for E again imply that different exponential forms of the variables can be used in each model. Examples of these models with nonlinear crowding effects are shown below:

Model B1b1:

$$DT = 15.79 + 0.34*OFFS + 0.089*MAXASLS^{0.5}, \quad (R^2=0.75)$$
(8.60) (2.63) (2.28) (DW=2.46)
[0.00] [0.01] [0.04]

Model C1b1:

$$DT = 15.76 + 0.34*OFFS + 0.089*SUMASLS^{0.5}, \quad (R^2=0.75)$$
(8.61) (2.70) (2.29) (DW=2.47)
[0.00] [0.02] [0.04]

Figure 3.23 R-SQUARE vs. E of SUMASLS

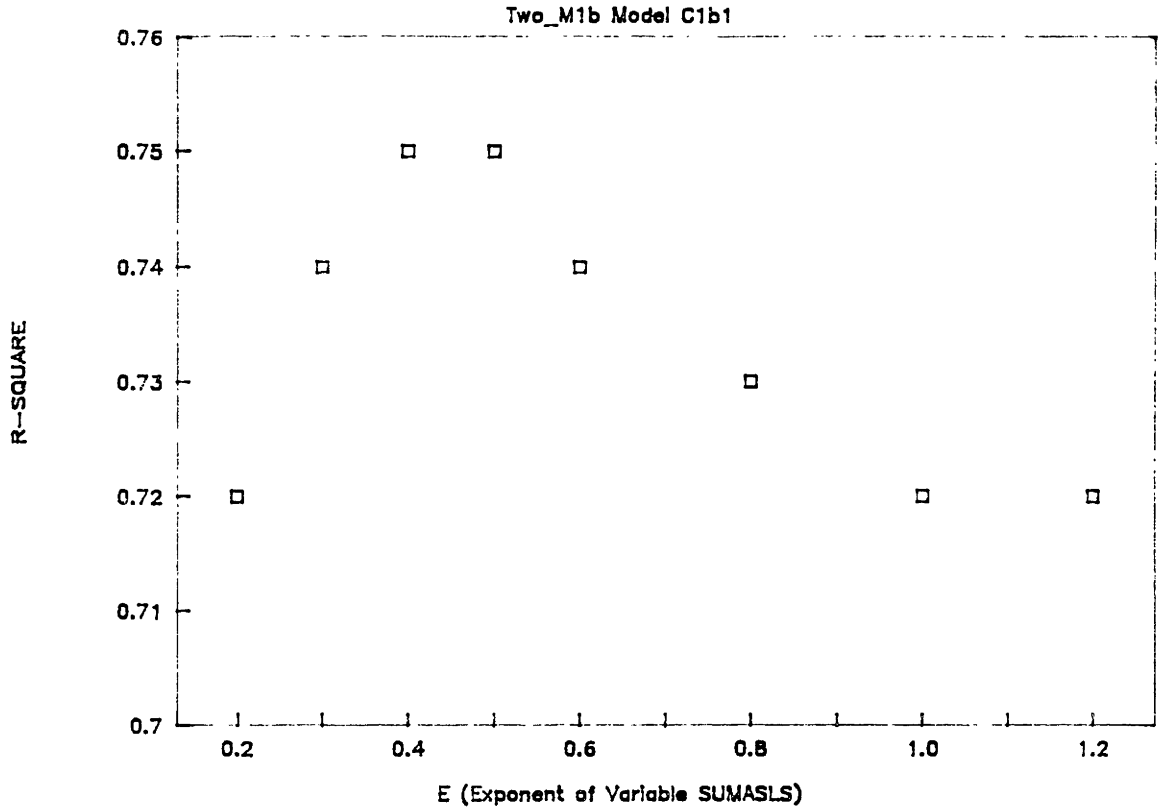


Figure 3.24 R-SQUARE vs. E of OFFAS

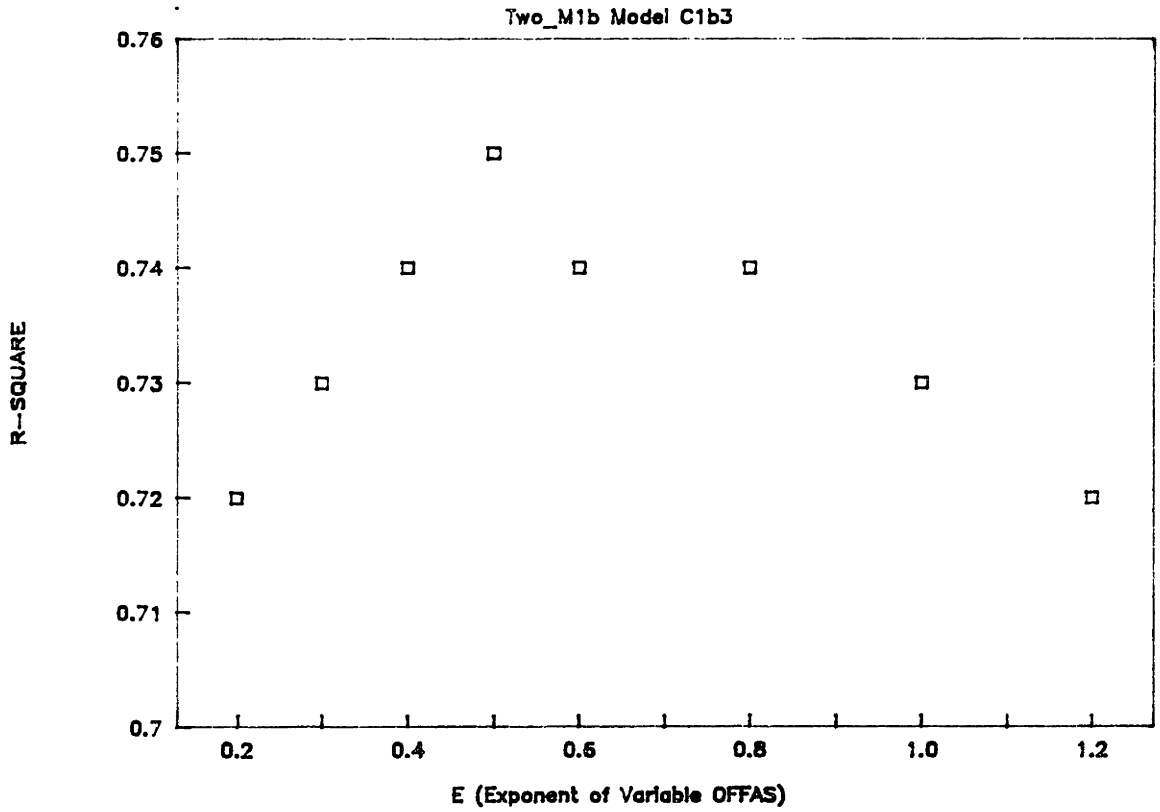


Figure 3.25 R-SQUARE vs. E of AS

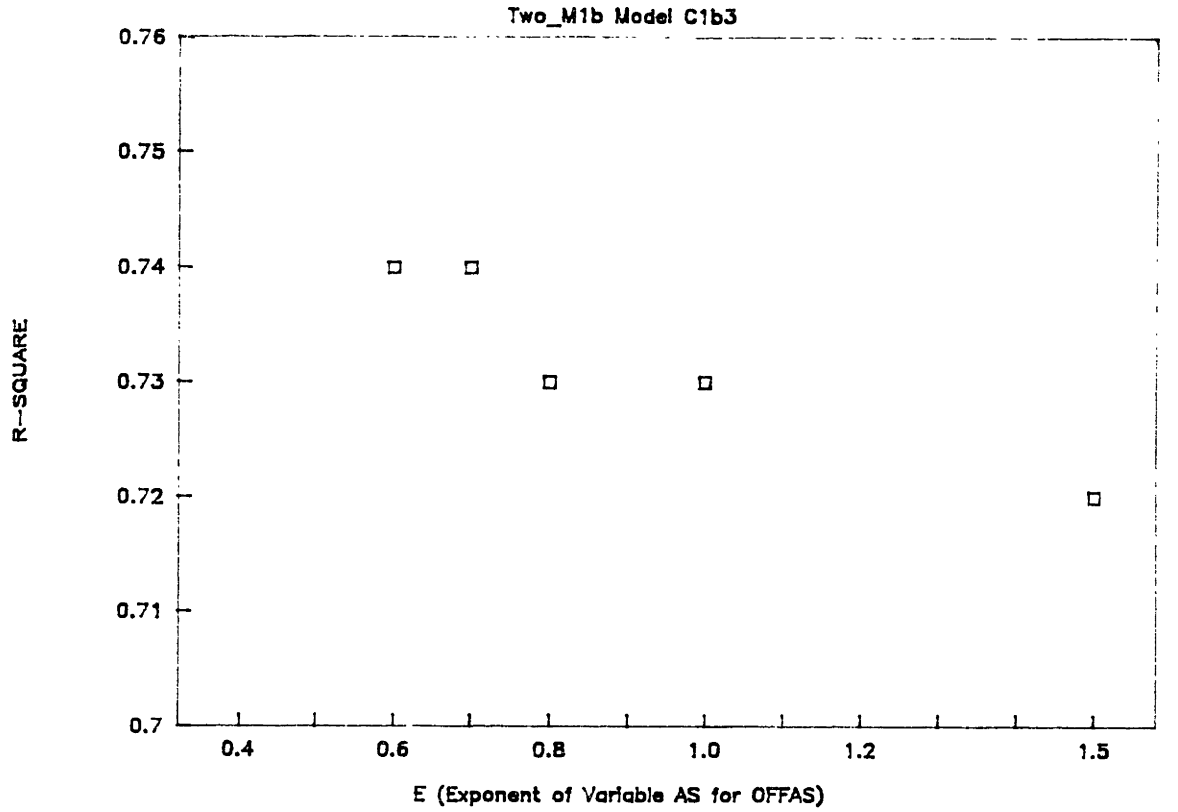
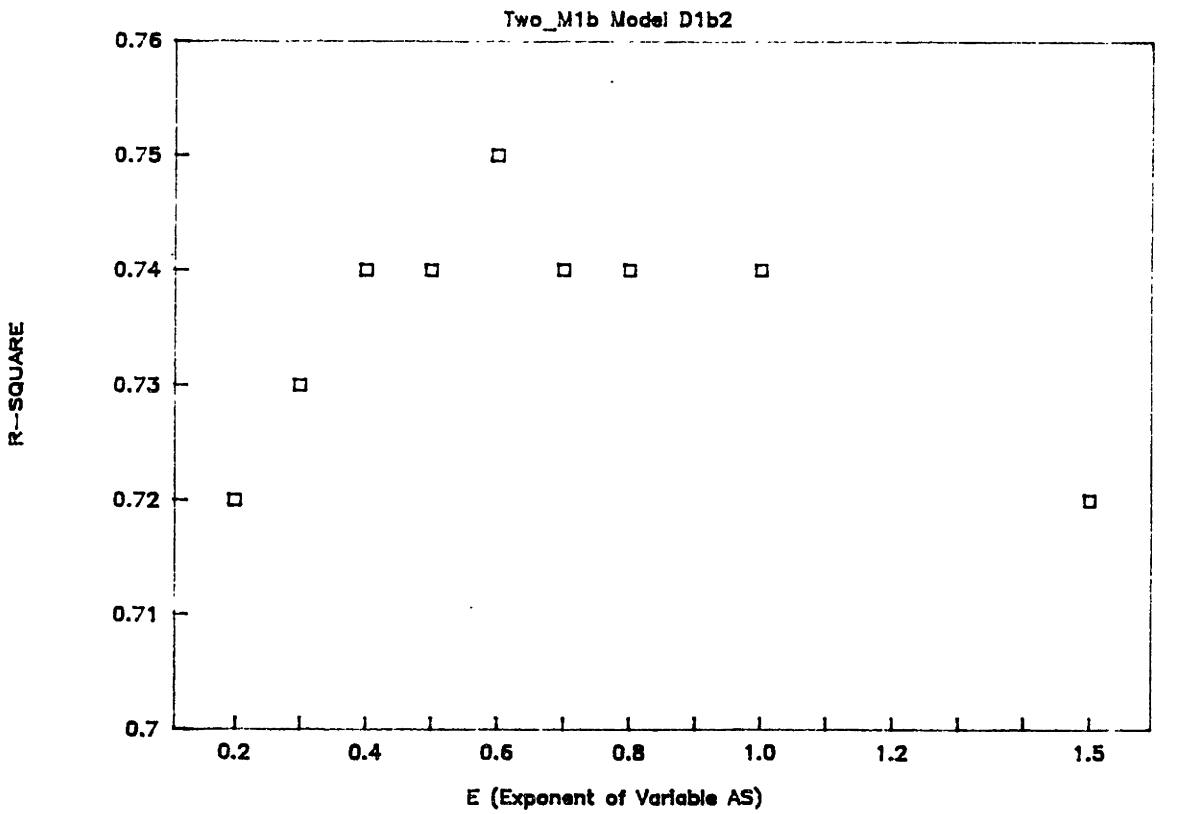


Figure 3.26 R-SQUARE vs. E of AS



Model C1b3(A):

$$DT=15.81+0.33*OFFS+0.12*OFFAS^{0.5}, \quad (R^2=0.75)$$
(8.59) (2.50) (2.28) (DW=2.39)
[0.00] [0.03] [0.04]

Model C1b3(B):

$$DT=16.64+0.29*OFFS+0.014*OFFS*AS^{0.6}, \quad (R^2=0.74)$$
(8.18) (1.94) (2.20) (DW=2.43)
[0.00] [0.07] [0.04]

Model D1b2:

$$DT=14.90+0.40*OFFS+0.30*AS^{0.6}, \quad (R^2=0.74)$$
(8.58) (3.72) (2.27) (DW=2.38)
[0.00] [0.00] [0.03]

As indicated by the R^2 values, it is clear that models B1b1, C1b1, C1b3, and D1bB with nonlinear forms of the crowding variables MAXASLS, SUMASLS, OFFAS, and AS are slight improvements over the linear model. The ranges of exponents for variables reflecting crowding effect which produce highest explanatory power are summarized in Table 3.12.

3.6.7 Checking the Assumptions for Linear Regression

Similar to the one-car models, before firm conclusions cab be drawn from the analysis, the four key assumptions (stated in section 3.5.7) for those most promising two-car models should be checked to make sure that none are violated. Assumptions 1 and 2 are checked by plotting histograms of residuals, while assumption 3 is checked by plotting standardized residuals against predicted dwell times, and assumption 4 is checked using the Durbin-Watson statistic (DW).

Table 3.12 Comparison of Exponents for Variables Reflecting Crowding Effect

Model Form	Two_M1 model	Two_M1b model
MAXASLS	1.2-5.0	0.5
SUMASLS	1.5-5.0	0.4-0.5
ONLS	2.0-3.5	-
ONS*LS(EXP)	1.5-5.0	-
LS	-	-
OFFAS	-	0.5
OFFS*AS(EXP)	-	0.6-0.7
AS	2.0-5.0	0.6

Appendix C presents example of the results for model C1-3 with linear crowding variable form (Figure C.7), and for model C1-1 with nonlinear crowding variable form (Figure C.8). It is clear from the Figures C.7 and C.8, the error component was not spread significantly from the mode at zero and the variance of the error component is virtually the same for all predicted values. Further, the DW statistics are about 2.15 for these desirable models with linear form of the crowding variables MAXASLS, SUMASLS, and LS, implying that no significant autocorrelation exists. Thus, these desirable models look promising as dwell time functions.

As discussed earlier, in Two_M1a model, two variables ONS and OFFS are used to explain the dwell time. Figure C.9 shows the histogram of error component based on this model form. It is clear from Figure C.9 that the error component is not spread significantly from the mode at zero and the variance of the error component is similar for all predicted values. The DW statistics is 2.17 implying that no significant autocorrelation exists. Thus, the model form with only ONS and OFFS variables appears promising as dwell time function in Two_M1a model.

Similarly, Figure C.10 presents the results for model Two_M1b with a linear crowding variable form, while Figure C.11 presents that with a nonlinear form. It is clear from the Figures C.10 and C.11 that the error component is approximately bell-shaped, with the mode at zero, but the variance of the error component does not appear to be the same across all predicted values, although only 19 sample points are available in estimating the Two_M1b model. Furthermore, the DW statistics are about 2.40 which implies that negative autocorrelation may exist. However, this test result is within the inconclusive region implying that additional information is needed before any conclusion can be drawn concerning possible error autocorrelation.

3.6.8 Summary for Two-Car Models

From the preceding analysis, the following conclusions can be drawn:

(1) It appears that linear regression models can accurately describe dwell time using several variables, moreover, the dwell time model should consider the effects of passenger crowding (except Two_M1a Model). There also appears to be little to choose between these five models, although the Two_M1, Two_M1a and Two_M1b are slightly preferred since these models offer better explanation for all observations. As indicated by the previous analysis, it is clear that Two_M1a and Two_M1b models, estimated by using two data sets based on relative magnitude of ONS and OFFS, are a slight improvement over Two_M1, since they better explain the observations and reflect the underlying boarding and alighting process. However, Two_M1 has the advantage that it can be applied to all predictions, while either Two_M1a or Two_M1b should be used depending on the specific situation. The promising Two_M1, Two_M1a and Two_M1b models are summarized below:

Model A1a:

$$DT = 9.69 + 0.42 \cdot ONS + 0.66 \cdot OFFS, \quad (R^2 = 0.71)$$
$$(4.32)\ (4.49)\ (3.99) \quad (DW = 2.17)$$
$$[0.00]\ [0.00]\ [0.00]$$

Model B1-1:

$$DT = 14.04 + 0.26 \cdot ONS + 0.36 \cdot OFFS + 0.0008 \cdot MAXASLS, \quad (R^2 = 0.70)$$
$$(7.43)\ (2.78)\ (3.71)\ (2.07) \quad (DW = 2.15)$$
$$[0.00]\ [0.01]\ [0.00]\ [0.04]$$

Model B1b:

$$DT=15.75+0.40*OFFS+0.0008*MAXASLS, \quad (R^2=0.72)$$
(8.11) (3.39) (1.90) (DW=2.46)
[0.00] [0.00] [0.08]

Model C1-1:

$$DT=13.93+0.27*ONS+0.36*OFFS+0.0008*SUMASLS, \quad (R^2=0.70)$$
(7.43) (2.92) (3.79) (2.03) (DW=2.15)
[0.00] [0.01] [0.00] [0.05]

Model C1b1:

$$DT=15.69+0.41*OFFS+0.0008*SUMASLS, \quad (R^2=0.72)$$
(8.10) (3.50) (1.88) (DW=2.46)
[0.00] [0.00] [0.08]

Model C1-3:

$$DT=13.92+0.22*ONS+0.43*OFFS+0.0013*ONLS, \quad (R^2=0.70)$$
(7.51) (2.00) (5.37) (2.08) (DW=2.12)
[0.00] [0.05] [0.00] [0.04]

Model C1b3:

$$DT=15.87+0.38*OFFS+0.0015*OFFAS, \quad (R^2=0.73)$$
(8.16) (3.11) (1.97) (DW=2.44)
[0.00] [0.00] [0.07]

Model D1b2:

$$DT=15.00+0.43*OFFS+0.037*AS, \quad (R^2=0.74)$$
(8.43) (4.23) (2.11) (DW=2.38)
[0.00] [0.00] [0.05]

(2) A variable reflecting the product of passenger movements and standees, such as MAXASLS, SUMASLS, or ONLS may make more sense than using the standee variable alone, (LS or AS), because if there were standees, but no passengers boarding or alighting,

the number of passengers standing should have no significant impact on dwell time. Based on this viewpoint, models B1-1, C1-3, and C1-3 may be preferred over model D1-2. Similarly, models B1b1, C1b1, and C1b3 are preferred over model D1b2. Furthermore, because the model C1-3 with the SUMASLS form assumes possible effect of the AS on the OFFS and of the LS on the ONS, it appears better able to interpret all possible crowding effects. Based on this viewpoint, the SUMASLS form is most preferred among these model forms.

(3) The constant in all promising Two_M1, Two_M1a, and Two_M1b models (about 13, 11, and 15 seconds respectively) was reasonable, since dwell time always include some time for the doors of the train to open and close, and some time for passengers who may want to alight, even if there is no one waiting to board the train.

(4) It appears that nonlinear forms for variables reflecting crowding effects in models Two_M1 and Two_M1b better describe the observations. Results of these models with nonlinear crowding variables are summarized below:

Model B1-1:

$$DT = 14.09 + 0.27*ONS + 0.37*OFFS + 1.2*10^{-4}*MAXASLS^{1.2}, \quad (R^2=0.71)$$
(7.54) (3.06) (4.01) (2.16) (DW=2.16)
[0.00] [0.01] [0.00] [0.04]

Model B1b1:

$$DT = 15.79 + 0.34*OFFS + 0.089*MAXASLS^{0.5}, \quad (R^2=0.75)$$
(7.94) (6.08) (2.28) (DW=2.46)
[0.00] [0.00] [0.04]

Model C1-1:

$$DT = 13.88 + 0.29*ONS + 0.39*OFFS + 7.2*10^{-6}*SUMASLS^{1.5}, \quad (R^2=0.71)$$

(7.71)	(3.73)	(4.48)	(2.21)	(DW=2.17)
[0.00]	[0.00]	[0.00]	[0.03]	

Model C1b1:

$$DT = 15.76 + 0.34*OFFS + 0.0089*SUMASLS^{0.5}, \quad (R^2=0.75)$$

(8.61)	(2.70)	(2.29)	(DW=2.47)
[0.00]	[0.01]	[0.04]	

Model C1-3(A):

$$DT = 13.79 + 0.27*ONS + 0.45*OFFS + 1.6*10^{-7}*ONLS^{2}, \quad (R^2=0.72)$$

(8.14)	(3.49)	(5.92)	(2.57)	(DW=2.16)
[0.00]	[0.00]	[0.00]	[0.02]	

Model C1b3(A):

$$DT = 15.81 + 0.33*OFFS + 0.12*OFFAS^{0.5}, \quad (R^2=0.75)$$

(8.59)	(2.50)	(2.28)	(DW=2.39)
[0.00]	[0.02]	[0.04]	

Model C1-3(B):

$$DT = 13.54 + 0.28*ONS + 0.44*OFFS + 6.0*10^{-7}*ONS*LS^{2}, \quad (R^2=0.71)$$

(8.06)	(3.70)	(5.65)	(2.41)	(DW=2.14)
[0.00]	[0.00]	[0.00]	[0.02]	

Model C1b3(B):

$$DT = 16.64 + 0.29*OFFS + 0.014*OFFS*AS^{0.6}, \quad (R^2=0.74)$$

(8.18)	(1.94)	(2.20)	(DW=2.43)
[0.00]	[0.07]	[0.03]	

Model D1-2:

$$DT = 12.72 + 0.36*ONS + 0.42*OFFS + 1.3*10^{-6}*AS^{2.5}, \quad (R^2=0.70)$$

(7.94)	(6.08)	(5.01)	(2.03)	(DW=2.11)
[0.00]	[0.00]	[0.00]	[0.05]	

Model D1b2:

$$DT = 14.90 + 0.40*OFFS + 0.30*AS^{0.6}, \quad (R^2=0.74)$$

(8.58)	(3.72)	(2.27)	(DW=2.38)
[0.00]	[0.00]	[0.03]	

(5) The recommended models explain about 70% of the variation in dwell times, implying that while some factors affecting the dwell time were not included, most importantly operator behavior and passenger characteristics, the most significant factors have been captured in these model forms.

3.7 Comparisons between One and Two-Car Train Models

Tables 3.13 to 3.15 compare the parameter estimates for One_M1, One_M1a, and One_M1b models, while Tables 3.16 to 3.18 present those for Two_M1, Two_M1a, and Two_M1b models, and Table 3.19 compares the exponents for the variables reflecting crowding, for both one and two-car models. From Tables 3.13 to 3.19, the following conclusions can be drawn:

(1) Tables 3.13 and 3.16 show that the constant terms in the Two_M1 dwell time models are greater than those in the corresponding One_M1 models, but the marginal dwell time for boarding is significantly smaller. The coefficients of ONS differ by about a factor of 2 in the two car case, that is, there are about half as many ONS per door as in a one car case with the same total ONS. The marginal dwell time for alighting varies between one and two-car models, depending on what model form is chosen, and it depends mainly on

load distribution between cars. It is also clear that the coefficients of the variables reflecting the crowding effect in the one-car models are greater than those in the two-car models implying that the marginal dwell time effect of crowding is greater in one-car trains than in two-car trains. This is due to the six doors available for passengers alighting and boarding in any two-car trains, and passengers would like to board a less crowded car, so that the impacts of AS on OFFS and LS on ONS are less significant.

To test these conclusions, the equality of individual parameter estimates from one and two-car train data sets was examined by t-statistics. It is clear from Table 3.20 that estimates of the variable ONS in A1 model from one and two-car data sets are significantly different. Similarly, estimates of the variable ONS in the other model forms (B1-1, C1-1, and C1-3) from one and two-car data sets are also marginally significantly different. T-tests are also applied to test the equality of parameter estimates of ONS from one car train data set being twice of that from the two car train data set. The t-statistics are 3.32, 0.17, 0.06, and 0.05 respectively for A1, B1-1, C1-1, and C1-3 models, implying that the null hypothesis should be rejected in A1 model, but there are no sufficient evidences to reject the null hypotheses in B1-1, C1-1, and C1-3 models.

As indicated by the corrected R^2 shown in Tables 3.13 and 3.16, it is clear that adding any proposed variable reflecting the crowding effect in One_M1 model significantly improves the explanatory power of the model. These results suggest that the on board crowding effect is more important for the one car train data set.

(2) Tables 3.14 and 3.17 show, that the constant terms in the one car and two car models are similar, but the Two_M1a models have smaller marginal dwell time for boarding and significant marginal dwell time for alighting. The coefficients of ONS differ by about a factor of 2 in the two car case since there are about half as many ONS per door as in a one car case with the same total ONS. The parameter estimates of the variables reflecting crowding are insignificant in all Two_M1a models and those for OFFS are

insignificant in most One_M1a models. The equality of individual parameter estimates from two data sets Two_M1a and One_M1a was examined only for A1a model by t-statistics. Table 3.21 shows that the estimates for the variables ONS and OFFS from two data sets are significantly different which implies that the marginal dwell times for boarding and alighting are significantly different between one and two-car train models with the data sets with ONS≥OFFS. T-test is also applied to test the equality of parameter estimate of ONS from one car train data set being twice of that from the two car train data set. The t-statistic is about 0.20, implying that there is no sufficient evidence to reject the null hypothesis that the marginal dwell time for boarding on one car train is twice that on two car trains while passengers boarding process dominate alighting.

As indicated by the corrected R^2 shown in Tables 3.14 and 3.17, it is clear that adding crowding variables in One_M1a model offer better explanation of dwell times, implying that the on board crowding effect is most important when the boarding process dominates alighting in one car train data set.

(3) Tables 3.15 and 3.18 compare parameter estimates for one and two-car train dwell time models with the data sets OFFS>ONS. It is clear that the marginal dwell time effect for boarding is only marginally significant for both the one car and two car models. This may be because the alighting process is not so significantly affected by passenger crowding. As to two-car train models, the insignificant marginal dwell time for boarding may be due to the six doors available for passengers alighting and boarding in any two-car trains so that the boarding and alighting process can occur simultaneously, resulting in the marginal effect on dwell times from boarding being negligible. As indicated by the R^2 values, using the variables ONS and OFFS in One_M1b, and OFFS in Two_M1b can well explain the dwell times. However, in some cases as Tables 3.15 and 3.18 show, adding a crowding variable does slightly improve the explanatory power. Finally, t tests were used to compare the individual parameter estimates from the two sets of models. As indicated by the t-

statistics shown in Table 3.22, the estimates of the variable ONS for model A1b are significantly different between one and two-car train data sets. Again, t test is also applied to test the equality of parameter estimate of ONS from one car train data set being twice of that from the two car train data set. The t-statistic is 2.42, implying that the null hypothesis should be rejected.

(4) Table 3.19 shows that models B1b1 and C1b1 of Two_M1b model with nonlinear MAXASLS and SUMASLS forms with an exponent of 0.5 are a slight improvement over the models with linear MAXASLS and SUMASLS forms, implying that the on board crowding term is more related to MAXASLS and SUMASLS with an exponent of 0.5. While in models C1-3 and D1-2 of Two_M1 model, it is clear that the models with nonlinear forms of $ONS \times LS^{2.0}$ and $AS^{2.0}$ are significant improvements over the linear models implying that the on board crowding term is more related to LS with an exponent about 2.0. Similarly, models C1-3, D1-2, and D1a2 with nonlinear forms of $ONS*LS^{2.5}$ and $LS^{2.5}$ perform better, implying that the passengers crowding effect is more related to LS with an exponent of 2.5.

(5) In conclusion, the passengers boarding and alighting process appears significantly different. Furthermore, using two sets of data, ONS $\geq$ OFFS and OFFS>ONS, to do analysis will make the results more clear since the aggregation models obscure the particular characteristics related to each data set.

Table 3.13 Comparison of Parameter Estimates of One-Car Train Dwell Time Models (t-statistics in parentheses)

	A1 model	B1-1 model	B1-2 model	B1-3 model	C1-1 model	C1-3 model	D1-3 model
Constant	9.07 (8.87)	12.59 (8.87)	12.34 (8.64)	12.28 (8.89)	12.50 (8.94)	11.94 (8.70)	9.24 (7.19)
ONS	1.55 (8.46)	0.55 (3.66)	0.67 (4.76)	0.53 (3.57)	0.55 (3.76)	0.45 (2.88)	0.71 (5.40)
OFFS	0.63 (5.58)	0.22 (1.89)	0.20 (1.65)	0.33 (3.10)	0.23 (2.03)	0.47 (4.66)	0.52 (5.35)
MAXASLS		0.0076 (6.49)					
ABAS			0.0072 (6.20)				
ABLS				0.0075 (6.76)			
SUMASLS					0.0078 (6.70)		
ONLS						0.013 (6.64)	
LS							0.16 (6.98)
Corrected R-Square	0.48	0.62	0.61	0.62	0.62	0.62	0.63

Table 3.14 Comparison of Parameter Estimates of One-Car Train Dwell Time Models (t-statistics in parentheses)

	A1a model	B1a model	C1a1 model	C1a3 model	D1a3 model
Constant	8.67 (3.91)	12.43 (6.33)	12.32 (6.33)	11.43 (6.15)	8.22 (4.37)
ONS	0.90 (4.03)	0.54 (2.65)	0.56 (2.78)	0.80 (2.12)	0.69 (3.55)
OFFS	1.41 (5.28)				0.73 (2.83)
MAXASLS		0.01 (8.12)			
SUMASLS			0.01 (8.25)		
OFFAS				0.022 (8.84)	
LS					0.18 (5.67)
Corrected R-Square	0.52	0.64	0.65	0.67	0.65

Table 3.15 Comparison of Parameter Estimates of One-Car Train Dwell Time Models (t-statistics in parentheses)

	A1b model	B1b model	C1b1 model	C1b3 model	D1b3 model
Constant	11.98	12.47	12.46	12.19	11.46
	(8.51)	(8.47)	(8.60)	(8.86)	(8.37)
ONS	0.88	0.67	0.65	0.48	0.60
	(4.61)	(2.51)	(2.43)	(1.61)	(2.64)
OFFS	0.43	0.39	0.39	0.47	0.48
	(3.82)	(3.34)	(3.43)	(4.23)	(4.38)
MAXASLS		0.002			
		(1.11)			
SUMASLS			0.0022		
			(1.25)		
ONLS				0.0074	
				(1.74)	
LS					0.066
					(2.09)
Corrected R-Square	0.64	0.64	0.65	0.66	0.67

Table 3.16 Comparison of Parameter Estimates of Two-Car Train Dwell Time Models (t-statistics in parentheses)

	A1 model	B1-1 model	B1-2 model	B1-3 model	C1-1 model	C1-2 model	D1-2 model
Constant	11.73 (7.44)	14.04 (2.78)	13.65 (7.40)	13.89 (7.44)	13.93 (7.43)	13.92 (7.51)	12.37 (7.73)
ONS	0.42 (7.59)	0.26 (2.78)	0.30 (3.67)	0.26 (2.77)	0.27 (2.92)	0.22 (2.00)	0.35 (5.20)
OFFS	0.49 (6.22)	0.36 (3.71)	0.36 (3.61)	0.38 (4.15)	0.36 (3.79)	0.43 (5.37)	0.41 (4.46)
MAXASLS		0.0008 (2.07)					
ABAS			0.0007 (1.88)				
ABLS				0.0008 (2.02)			
SUMASLS					0.0008 (2.03)		
ONLS						0.0013 (2.08)	
AS							0.027 (1.61)
Corrected R-Square	0.68	0.70	0.70	0.70	0.70	0.70	0.69

Table 3.17 Comparison of Parameter Estimates of Two-Car Train Dwell Time Models (t-statistics in parentheses)

	A1a model	B1a model	C1a1 model	C1a3 model	D1a2 model
Constant	9.69 (4.32)	11.38 (3.79)	11.31 (3.83)	11.24 (3.72)	9.90 (4.21)
ONS	0.42 (4.49)	0.34 (2.52)	0.34 (2.62)	0.33 (2.12)	0.41 (3.98)
OFFS	0.66 (3.99)	0.52 (2.23)	0.52 (2.23)	0.58 (2.91)	0.60 (2.67)
MAXASLS		0.0005 (0.85)			
SUMASLS			0.0005 (0.85)		
ONLS				0.00067 (0.78)	
AS					0.01 (0.36)
Corrected R-Square	0.71	0.70	0.70	0.70	0.70

Table 3.18 Comparison of Parameter Estimates of Two-Car Train Dwell Time Models (t-statistics in parentheses)

	A1b model	B1b model	C1b1 model	C1b3 model	D1b2 model
Constant	14.39 (7.46)	15.75 (8.11)	15.69 (8.10)	15.87 (8.16)	15.00 (8.43)
ONS					
OFFS	0.56 (6.29)	0.40 (3.39)	0.41 (3.50)	0.38 (3.11)	0.43 (4.23)
MAXASLS		0.0008 (1.90)			
SUMASLS			0.0008 (1.88)		
OFFAS				0.0015 (1.97)	
AS					0.037 (2.11)
Corrected R-Square	0.68	0.72	0.72	0.73	0.74

Table 3.19 Comparison of Exponents for Variables Reflecting Crowding Effect

Model Form	One_M1 model	One_M1a model	One_M1b model	Two_M1 model	Two_M1b model
MAXASLS	1.0-1.3	0.8-1.4	-	1.2-5.0	0.5
SUMASLS	0.8-2.0	1.0-1.2	-	1.5-5.0	0.4-0.5
ONLS	0.8-1.3	-	-	2.0-3.5	-
ONS*LS(EXP)	2.5	-	1.4-2.8	1.5-5.0	-
LS	2.5-3.0	2.5-3.0	1.2-2.7	-	-
OFFAS	-	1.2-1.8	-	-	0.5
OFFS*AS(EXP)	-	1.6-2.7	-	-	0.6-0.7
AS	-	-	-	2.0-5.0	0.6

Table 3.20 t test for Equality of Individual Coefficient between One and Two-Car Train Data Sets (t-statistics shown)

Variable	A1 model	B1-1 model	C1-1 model	C1-3 model
Constant	-1.42	-0.28	-0.61	-0.86
ONS	5.90*	1.64	1.62	1.20
OFFS	1.02	-0.93	-0.88	0.31
MAXASLS	-	5.52*	-	-
SUMASLS	-	-	5.71*	-
ONLS	-	-	-	5.69*

* significant at 0.05 level.

Table 3.21 t test for Equality of Individual Coefficient between One_M1a and Two_M1a Data Sets (for A1a Models Only)

Variable	DF	t	Critical t-value
Constant	112	-0.32	1.97
ONS	112	1.99*	1.97
OFFS	112	2.39*	1.97

* significant at 0.05 level.

Table 3.22 t test for Equality of Individual Coefficient between One_M1b and Two_M1b Data Sets (for A1b Models Only)

Variable	DF	t	Critical t-value
Constant	55	-0.97	2.00
ONS	55	2.93*	2.00
OFFS	55	-0.41	2.00

* significant at 0.05 level.

3.8 Comparison with Koffman's Dwell Time Models

Compared with the results on the Boston College and Riverside Lines shown in Koffman (1984), the mean dwell time of 23.3 seconds for the one-car trains in this study is greater than those of 10.0 seconds for Boston outbound and 15.4 seconds for Boston inbound (see Table 1.2) because of different definition and measurement of dwell times between two studies. Furthermore, the average passengers boarding and alighting, and the average arrival load (AL) and leaving load (LL) of 8.7, 6.7, 75, and 76 respectively for the one car train data set in this study are greater than those on Koffman's: 3.1, 6.3, and 46.7 for passengers boarding, passengers alighting, and passengers on-board (POB) respectively in Boston outbound; and 4.0, 0.8, and 41.9 for passengers boarding, passengers alighting, and passengers on-board respectively in Boston inbound. The average passengers boarding and alighting, and AL and LL are 19.1, 14.2, 163, and 168 respectively for the two car train data set in this study.

The preliminary analysis reflects the relationships between the dwell time and the key explanatory variables. To explain or predict the dwell time using these variables, multiple linear regression method is applied. The models were estimated in several forms using three explanatory variables: number of passengers boarding (ONS), number of passengers alighting (OFFS), and an alternative variable reflecting crowding effect on board. The constant term in all one-car models is about 10 seconds while that in two-car models is about 13 seconds. The marginal dwell time for boarding, alighting, and the explanatory variable reflecting crowding effect varies, depending on what model form is chosen. For example, the marginal effect on dwell time for boarding, alighting, and crowding effect is 0.55, 0.23, and 0.0078 respectively for the one car train model with the SUMASLS form.

Koffman's models for Boston and San Diego outbound are shown below:

Boston outbound (Free Fare)

Riverside

$DT=3.04+0.65*ONS+0.61*OFFS+0.040*POB, \quad (R^2=0.68)$
(2.50) (7.70) (6.20) (1.40)
[0.02] [0.00] [0.00] [0.18]

Boston College

$DT=2.96+0.84*ONS+0.52*OFFS+0.029*POB, \quad (R^2=0.84)$
(6.00) (13.8) (13.2) (3.00)
[0.00] [0.00] [0.00] [0.00]

San Diego (SSFC)

$DT=8.14+0.67*ONS+0.59*OFFS+0.034*POB, \quad (R^2=0.43)$
(15.0) (17.4) (14.9) (4.80)
[0.00] [0.00] [0.00] [0.00]

Compared with Koffman's results, the constant terms in this study are greater than those in Boston outbound because of different definition and measurement of the dwell times, but the constant terms are close between this study and San Diego case. The definition and measurement of the dwell times are identical in both San Diego case and this study.

The marginal dwell time for boarding and alighting in both Koffman's models and the one car train model D1-3 of this study are close, both have the similar model form. The only difference between these two models is D1-3 model using leaving standees (LS) reflecting crowding effect on board while Koffman's model using POB reflecting it. Compared with the Koffman's results, the marginal effect on dwell time for boarding, alighting, and crowding effect of the alternative model forms in this study may vary.

Chapter 4

Summary and Conclusions

4.1 General

The formulation of dwell time models for rail systems has been addressed in this thesis. Whereas in practice this topic is seldom dealt with in the transit industry, this study demonstrates that the dwell time is positively related to the number of alighting and boarding passengers, and significantly related to the passenger crowding level on board.

The mean dwell time are 23.3 and 26.6 seconds, with standard deviations of 11.41 and 8.40 seconds for the one and two-car train observations. The observations were classfied into four groups for the preliminary load analysis first, based on the following ranges for leaving passenger load (LPL): LPL<53, 53≤LPL<81, 81≤LPL<109, and LPL≥109. The mean dwell times are 16.8, 20.6, 24.0, and 36.0 seconds for the groups with standard deviations of 5.65, 8.35, 6.68, and 13.31 seconds respectively for the one-car train data set. F test shows that the mean dwell time appears significantly related to LPL. Secondly, the observations were classified into four groups for the boarding and alighting (ONOFFS) analysis, based on the following ranges for ONOFFS: ONOFFS≤9, 9<ONOFFS≤17, 17<ONOFFS≤25, and ONOFFS>25. The mean dwell times are 15.8, 20.0, 27.1, and 41.0 seconds, with standard deviations of 6.65, 6.32, 5.90, and 14.96 seconds for the four groups. F test indicates that the mean dwell time is positively related to the ONOFFS variable.

Similarly, a preliminary analysis was conducted for the set of two-car train dwell times based on the LPL and ONOFFS variables. The mean dwell times are 20.4, 23.2, 27.5, and 35.5 seconds with standard deviations of 5.68, 7.39, 6.81, and 6.31 seconds for the four groups based on load analysis. F test suggests the mean dwell time is positively

related to the LPL for the two-car train data set. The analysis based on the ONOFFS variable presents the mean dwell times of 19.3, 22.8, 28.7, and 34.9 seconds with standard deviations of 5.69, 4.87, 6.81, and 6.56 seconds respectively for the four groups. Once again, F test demonstrates that the mean dwell time is positively related to ONOFFS.

The preliminary analysis reflects the relationships between the dwell time and the key explanatory variables. For example, the marginal dwell time for boarding, alighting, and crowding effect on board are 0.55, 0.23, and 0.0078 respectively with a constant of 12.5 seconds for one car train model with SUMASLS form, similarly, those are 0.27, 0.36, and 0.0008 with a constant of 13.9 seconds for the corresponding two car train model. Suppose the number of passengers boarding and alighting are 20 respectively at a station, and AS and LS are assumed 100 respectively (that is a heavy loaded condition for one car train), then the predicted dwell times for both one and two car trains are :

One-car train:

DT=12.5+0.55*20+0.23*20+0.0078*(20*100+20*100)=59.3 (seconds)

Two-car train:

A. the same ONS, OFFS, AS, and LS as one-car train

DT=13.9+0.27*20+0.36*20+0.0008*(20*100+20*100)=29.7 (seconds)

B. Double ONS, OFFS, AS, and LS

DT=13.9+0.27*40+0.36*40+0.0008*(40*200+40*200)=51.9(seconds)

It is clear that the crowding effect on dwell time is much greater in a heavy loaded one-car train than in a two-car train. In real operation, these dwell time functions can be applied in the scheduling process and in operations, thereby achieving the objectives of promoting service quality, line capacity, operations, and reliability.

The major issues and findings relating to the models presented in the case study of

this thesis are summarized in the following sections. Suggestions are also made for further research on this topic.

4.2 Data Collection Issues

The one and two-car train data sets were collected from two stations (Copley and Arlington), and one station (Arlington) respectively on the MBTA (Massachusetts Bay Transportation Authority) Green Line as described in Chapter 3. Because of the unusual level of detail of data required, it was necessary to have a two person team per car, or a four person team for a two car trains, to collect data. Although the use of video-recording may appear to be an alternative way to collect data, this technique cannot be used to capture all the data required simultaneously, thus it appears infeasible for this type of study. Data collection was a labor intensive and difficult process for this study, which limited the data available for comprehensive analysis in this thesis. However, in the transit industry, for further study and building dwell time function purposes, the data collection method stated in Chapter 3 can be adopted to collect additional data, covering several stations, times of day, and different rail system and fare collection types.

4.3 Modelling Issues

Multiple linear regression models were applied to explain the dwell time with several explanatory variables: the number of boarding passengers (ONS), the number of alighting passengers (OFFS), and passenger crowding level on board (such as: MAXASLS, ABAS, ABLS, SUMASLS, ONLS, OFFAS, AS, and LS). A crucial issue in applying linear regression is that these explanatory variables should not be highly correlated with each

other, that is, explanatory variables with high correlation should not be included together in any model formulation. Furthermore, the four key assumptions stated in Chapters 2 and 3 should be met for any linear regression model. Otherwise, remedial measures should be taken, such as applying the generalized least squares (GLS) or weighted least squared (WLS) methods.

4.4 Case Study

Dwell time functions were estimated for one and two-car trains on the Green Line light rail operation of the Massachusetts Bay Transportation Authority (MBTA). A number of findings were reached as a result of the case study:

(1) One and two-car train observations were classified into four groups for the preliminary analysis, based on leaving passenger load (LPL) and the number of passengers boarding and alighting (ONOFFS). The results show that the mean dwell time is positively related to both leaving passenger load and the number of passengers boarding and alighting for one and two-car train observations at the 0.05 significance level.

(2) There is no significant difference in the mean dwell times between the one and two-car data sets for groups with the same ranges of the LPL and ONOFFS at the 0.05 significance level.

(3) Based on the preliminary analysis, three major factors, the number of passengers boarding (ONS), passenger alighting (OFFS), and the effect of crowding on board, were expected to enter into the dwell time function. The models were then estimated for one-car trains based on three approaches: all data together, the data set with ONS being equal to, or greater than, OFFS (ONS≥OFFS) and that with OFFS being greater than ONS (OFFS>ONS). The results show that adding any alternative form of the variable (such as:

MAXASLS, defined as the maximum of the product of ONOFFS and AS (arriving standees), and ONOFFS and LS (leaving standees)); and SUMASLS, defined as the sum of the product of OFFS and AS, and ONS and LS) to reflect the effects of passenger crowding on board in the ONS≥OFFS data set does significantly improve the explanatory power of the model, implying that on board crowding is most important when the boarding process dominates alighting in the one-car train data set.

The model estimations by using two data sets based on relative magnitude of ONS and OFFS are a slight improvement over that by aggregating data set. However, the aggregated model has the advantage that it can be applied to all predictions while the appropriate disaggregate model should be used in the specific situation.

(4) The model specifications with a variable reflecting the product of passenger movements and the standees, such as the variable MAXASLS, SUMASLS, or ONLS (the product of ONS and LS) may make more sense than using the standee variable alone, (LS and AS), because if there were standees, but no passengers boarding or alighting, the number of passengers standing should have no significant impact on dwell time. Furthermore, because the model estimation with the SUMASLS form assumes possible effect of AS on OFFS and LS on ONS, it appears better able to reflect all possible crowding effects. Based on this viewpoint, the SUMASLS form is preferred among these model forms. These models are summarized below, with R^2 value, t-statistics in parentheses, and p-values:

Model B1-1:

$$DT=12.59+0.55*ONS+0.22*OFFS+0.0076*MAXASLS, \quad (R^2=0.62)$$

(8.87)	(3.66)	(1.89)	(6.49)
[0.00]	[0.00]	[0.06]	[0.00]

Model C1-1:

$$DT=12.50+0.55*ONS+0.23*OFFS+0.0078*SUMASLS, \quad (R^2=0.62)$$
(8.94) (3.76) (2.03) (6.70)
[0.00] [0.00] [0.04] [0.00]

Model C1-3:

$$DT=11.94+0.45*ONS+0.47*OFFS+0.013*ONLS, \quad (R^2=0.62)$$
(8.70) (2.88) (4.66) (6.64)
[0.00] [0.01] [0.00] [0.00]

Model D1-3:

$$DT=9.24+0.71*ONS+0.52*OFFS+0.16*LS, \quad (R^2=0.63)$$
(7.19) (5.40) (5.35) (6.98)
[0.00] [0.00] [0.00] [0.00]

(5) Models with nonlinear forms of crowding variables were estimated and in many cases, these models with form such as $ONS \times LS^{2.5}$, are a significant improvement over those with linear form, implying that the on board crowding effect is important for one car train data set, which are summarized below:

Model C1-3:

$$DT=11.43+0.69*ONS+0.48*OFFS+1.35*10^{-5}*ONS*LS^{2.5}, \quad (R^2=0.65)$$
(8.78) (5.38) (4.99) (7.41)
[0.00] [0.00] [0.00] [0.00]

Model D1-3:

$$DT=10.05+0.78*ONS+0.50*OFFS+2.0*10^{-4}*LS^{2.5}, \quad (R^2=0.67)$$
(8.32) (6.70) (5.51) (8.50)
[0.00] [0.00] [0.00] [0.00]

(6) Models were estimated for two-car trains based on five approaches: to treat the train as an entity (including: all data together, the data set with ONS being equal to, or greater than, OFFS ($ONS \geq OFFS$), and that with OFFS>ONS); to deal with a two-car train

as two single cars, and found the relationships between the dwell time and total boarding passengers, total alighting passengers, and crowding effect, for each car separately; and to treat the train based on the theory that the dwell time for a two-car train is the longer dwell time (LDT) for car 1 and car 2 of that train, and found the relationships between the dwell time and the number of passengers boarding and alighting and level of that car with LDT.

The results indicate that treating the train as an entity, the dwell times of trains may be well explained by the explanatory variables ONS, OFFS, and those reflecting passenger crowding effect on board (such as MAXASLS, or SUMASLS). Moreover, models estimated by using two data sets based on relative magnitude of ONS and OFFS, are a slight improvement over an aggregated data set, since the former better reflect the particular boarding and alighting process characteristics in effect. The promising two car train models are shown below:

Model B1-1:

$$DT=14.04+0.26*ONS+0.36*OFFS+0.0008*MAXASLS, \quad (R^2=0.70)$$

(7.43)	(2.78)	(3.71)	(2.07)
[0.00]	[0.01]	[0.00]	[0.04]

Model C1-1:

$$DT=13.93+0.27*ONS+0.36*OFFS+0.0008*SUMASLS, \quad (R^2=0.70)$$

(7.43)	(2.92)	(3.79)	(2.03)
[0.00]	[0.01]	[0.00]	[0.05]

Model C1-3:

$$DT=13.92+0.22*ONS+0.43*OFFS+0.0013*ONLS, \quad (R^2=0.70)$$

(7.51)	(2.00)	(5.37)	(2.08)
[0.00]	[0.05]	[0.00]	[0.04]

(7) Model specifications with a variable reflecting the product of passenger movements and the standees, such as the variable MAXASLS, SUMASLS, ONLS make

more sense than using the standee variable alone, (LS or AS), as stated in one-car model. Furthermore, the model estimation with the SUMASLS form allows possible effects of AS on OFFS and LS on ONS, and so it appears better able to reflect all possible crowding effects, therefore the SUMASLS form is preferred among all two-car model forms.

(8) As discussed in Section 3.6 of Chapter 3, several model specifications with nonlinear forms of crowding variables do offer better fit to observations over those with linear forms of crowding variables, such as model estimation with the MAXASLS form, with the comparable ranges of exponent being 1.2-5.0, and that with SUMASLS form, with the exponent of 1.5-5.0. These model with nonlinear crowding variables are summarized below:

Model B1-1:

$$DT = 14.09 + 0.27*ONS + 0.37*OFFS + 1.2*10^{-4}*MAXASLS^{1.2}, \quad (R^2=0.71)$$

	(7.54)	(3.06)	(4.01)	(2.16)
	[0.00]	[0.01]	[0.00]	[0.04]

Model C1-1:

$$DT = 13.88 + 0.29*ONS + 0.39*OFFS + 7.2*10^{-6}*SUMASLS^{1.5}, \quad (R^2=0.71)$$

	(7.71)	(3.73)	(4.48)	(2.21)
	[0.00]	[0.00]	[0.00]	[0.03]

Model C1-3(A):

$$DT = 13.79 + 0.27*ONS + 0.45*OFFS + 1.6*10^{-7}*ONLS^{2}, \quad (R^2=0.71)$$

	(8.14)	(3.49)	(5.92)	(2.57)
	[0.00]	[0.00]	[0.00]	[0.02]

Model C1-3(B):

$$DT = 13.54 + 0.28*ONS + 0.44*OFFS + 6.0*10^{-7}*ONS*LS^{2}, \quad (R^2=0.71)$$

	(8.06)	(3.70)	(5.65)	(2.41)
	[0.00]	[0.00]	[0.00]	[0.02]

Model D1-2:

$$DT=12.72+0.36*ONS+0.42*OFFS+1.3*10^{-6}*AS^{2.5}, \quad (R^2=0.70)$$

(7.94)	(6.08)	(5.01)		(2.03)
[0.00]	[0.00]	[0.00]		[0.05]

(9) It appears that the desirable one and two-car train models do not violate the key assumptions of the multiple linear regression model, therefore these models look promising as dwell time functions. The recommended models explain about 70% of the variation in dwell times, implying that some factors affecting the dwell time were not included, most importantly operator behavior and passenger characteristics; however, the most significant factors have been captured in these model forms.

(10) The constant in all one-car models is about 10 seconds while that in two-car models is about 13 seconds. These results were reasonable, since dwell time always includes some time for doors of the train to open and close, and some time for passengers who may want to alight, even if there is no one waiting to boarding the train.

(11) The t tests are applied to test the equality of individual coefficient between one and two-car train data sets: the full data set; the data set with ONS≥OFFS ; and the data set with OFFS>ONS. As indicated by the t-statistics, the estimates of ONS in A1 model from one and two-car train full data sets are significantly different while those in the other model firms have a lower level of significance, and the variables reflecting crowding effect in B1-1, C1-1, and C1-3 models are significantly different between the two data sets.

The t-statistics show that the estimates for the variables ONS and OFFS from two data sets Two_M1a and One_M1a (ONS≥OFFS) are significantly different, implying that the marginal dwell times for boarding and alighting are significantly different between one and two-car train models with the data sets ONS≥OFFS. But, t test also shows there is no sufficient evidence to reject the null hypothesis that the marginal dwell time for boarding on one car train is twice that on two car trains while passengers boarding process dominates.

Finally, t tests were used to compare the individual parameter estimates from the two data sets Two_M1b and One_M1b (OFFS>ONS). The t-statistics show that the estimates of the variable ONS for model A1b are significantly different between two data sets while that of OFFS are not significantly different.

4.5 Directions for Further Research

Even though this study made a number of important contributions to the model formulations of dwell time relationships for urban rail systems, numerous areas remain candidates for possible future research.

In term of data inputs and requirements of the formulation, a number of areas remain to be addressed. For example: to incorporate the operator behavior and passenger characteristics into the dwell time model; to reflect the station congestion level in the model; to formulate more than two-car train (say three-car, four-car, and six-car train) dwell time model; and then to test the equality of the estimate of marginal boarding dwell time for alighting and boarding between these models.

In term of modelling specifications, it is expected to include operator behavior, station platform congestion level, and passenger characteristics in the model formulations provided these data can be obtained in further study.

Bibliography

Boardman, T.J. and Kraft, W.H. (1970). "Predicting Bus Passenger Service Time - Part II", Traffic Engineering, Vol. 40.

Cundill, M.A. and Watts, P.F. (1973). "Bus Boarding and Alighting Times", TRRL Report LR 521, Department of the Environment, U.K.

Dubin, J.A. and Rivers, R.D. (1986). "Statistical Software Tools", Durbin/Rivers Research.

Fritz, M.S. (1981). "Effects of Crowding on Light Rail Passenger Boarding and Alighting Rate", M.S. Thesis, Northwestern University.

Fritz, M.S. (1983). "Effect of Crowding on Light Rail Passenger Boarding Times", Transportation Research Record 908.

Greene, W.H. (1990). "Econometric Analysis", MacMillan Publishing Company.

Guenthner, R.P. and Sinha, K.C. (1983). "Modelling Bus Delays Due to Passenger Boardings and Alightings", Transportation Research Record 915.

HCM (1985). "Highway Capacity Manual", Transportation Research Board Special Report 209.

Johnston, J. (1984). "Econometric Method, Third Edition", Mc Graw-Hill Publishing Company.

Koffman, D., Rhyner, G., and Trexler, R. (1984). "Self-Service Fare Collection on the San Diego Trolley", U.S. Department of Transportation.

Kraft, W.H. and Bergen, T.F. (1974). "Evaluation of Passenger Service Times for Street Transit Systems", Transportation Research Record 505.

Kraft, W.H. (1975). "An analysis of the Passenger Vehicle Interface of Street Transit Systems with Applications to Desigh Optimization", Ph.D. Dissertation, New Jersey Institute of Technology.

Kraft, W.H. and Deutschman, H. (1977). "Bus Passenger Service-Time Distributions", Transportation Research Record 625.

Levinson, H.S. (1983). "Analyzing Transit Travel Time Performance", Transportation Research Record 915.

Minitab (1989). "Minitab Reference Manual, Release 7", Minitab Inc..

Never, J., Wasserman, W., and Kutner, M.H. (1983). "Applied Linear Regression Models", IRWIN Inc..

Zografos, K.G. and Levinson, H.S. (1986), "Passenger Service Times for A No-Fare Bus System", Transportation Record 1051.

Appendices

Appendix A

Dwell Time vs. Explanatory Variables

Figure A.1 DWELL TIME vs. ONS

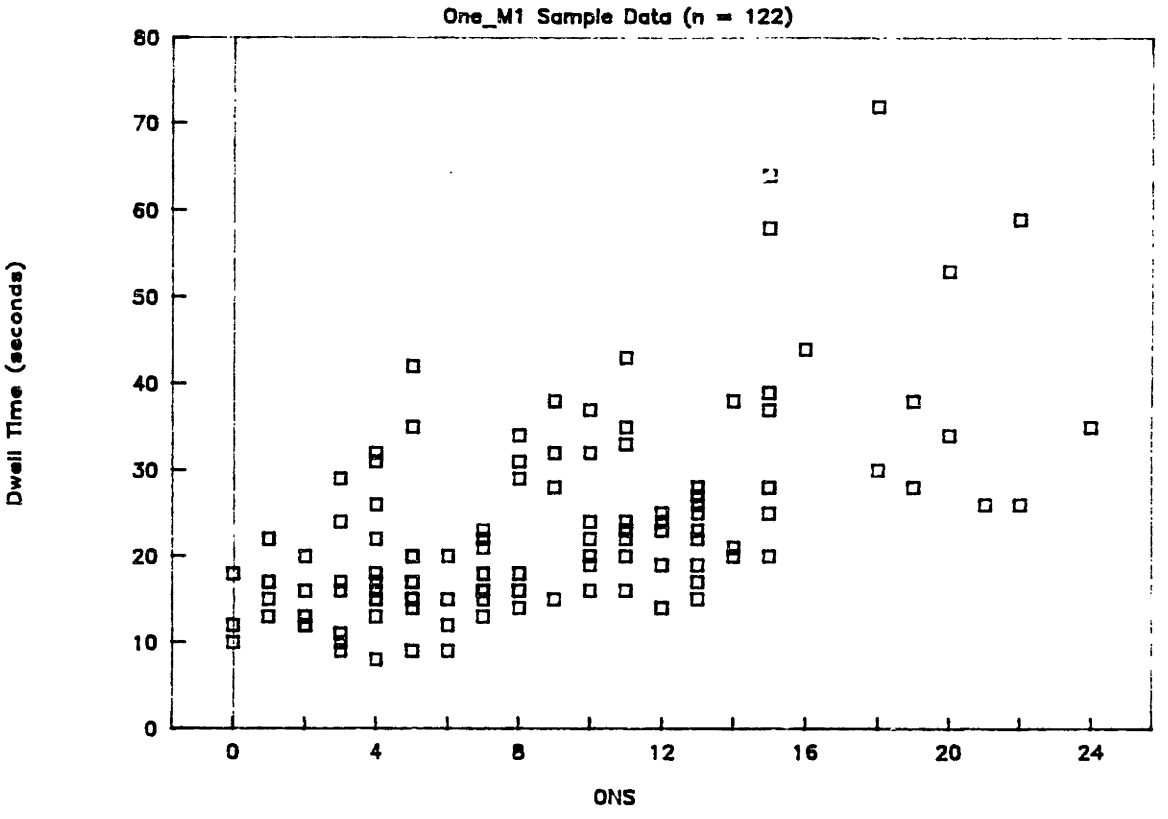


Figure A.2 DWELL TIME vs. OFFS

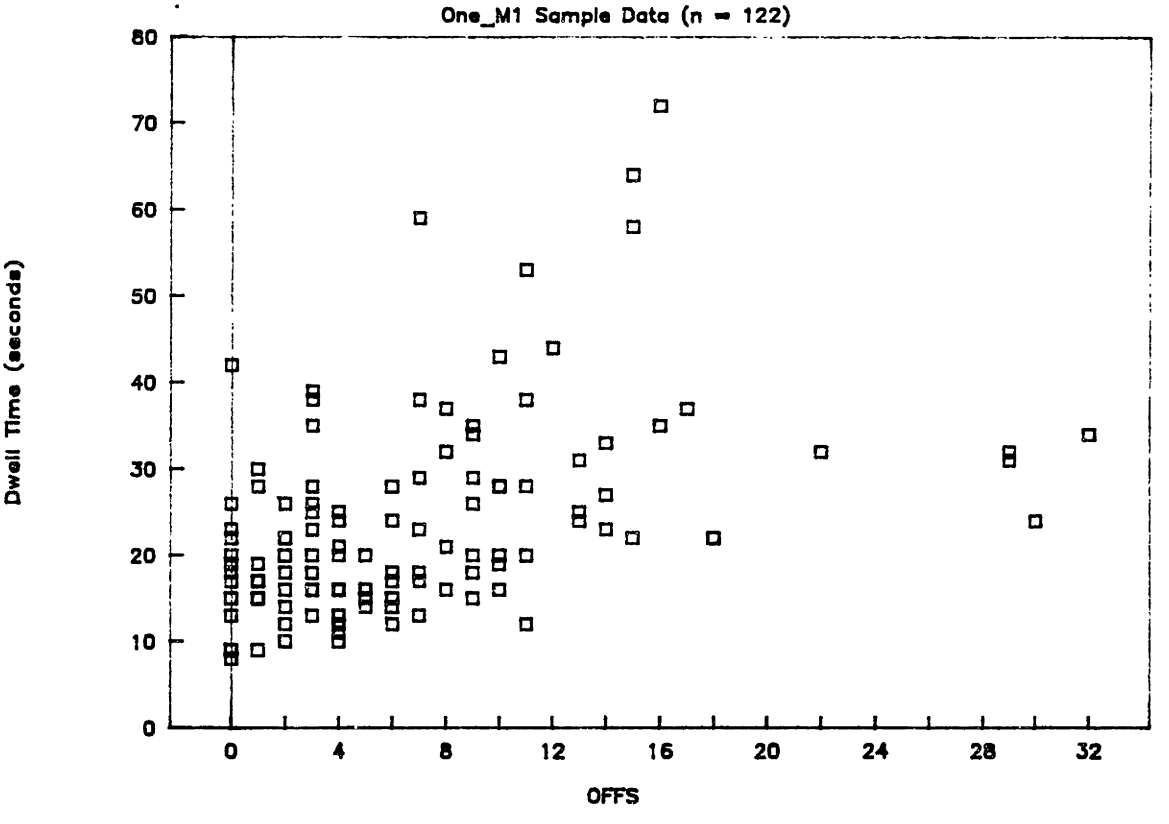


Figure A.3 DWELL TIME vs. ABAS

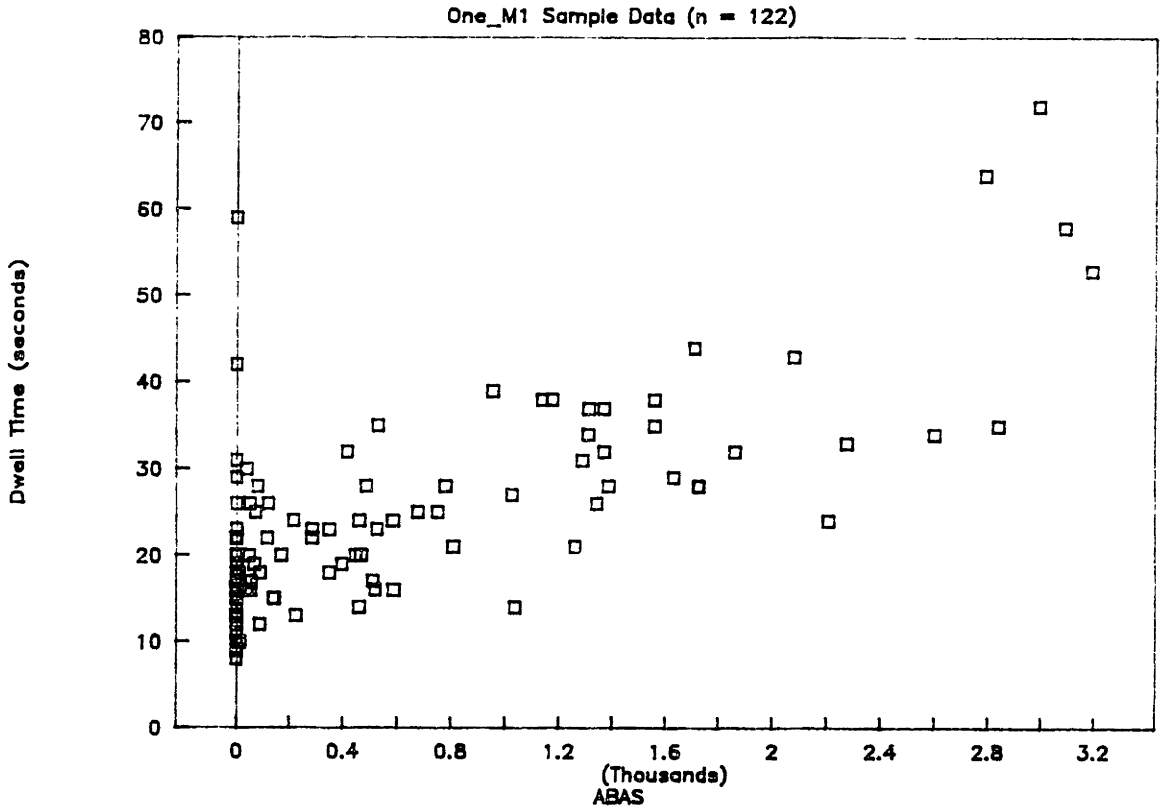


Figure A.4 DWELL TIME vs. ABLS

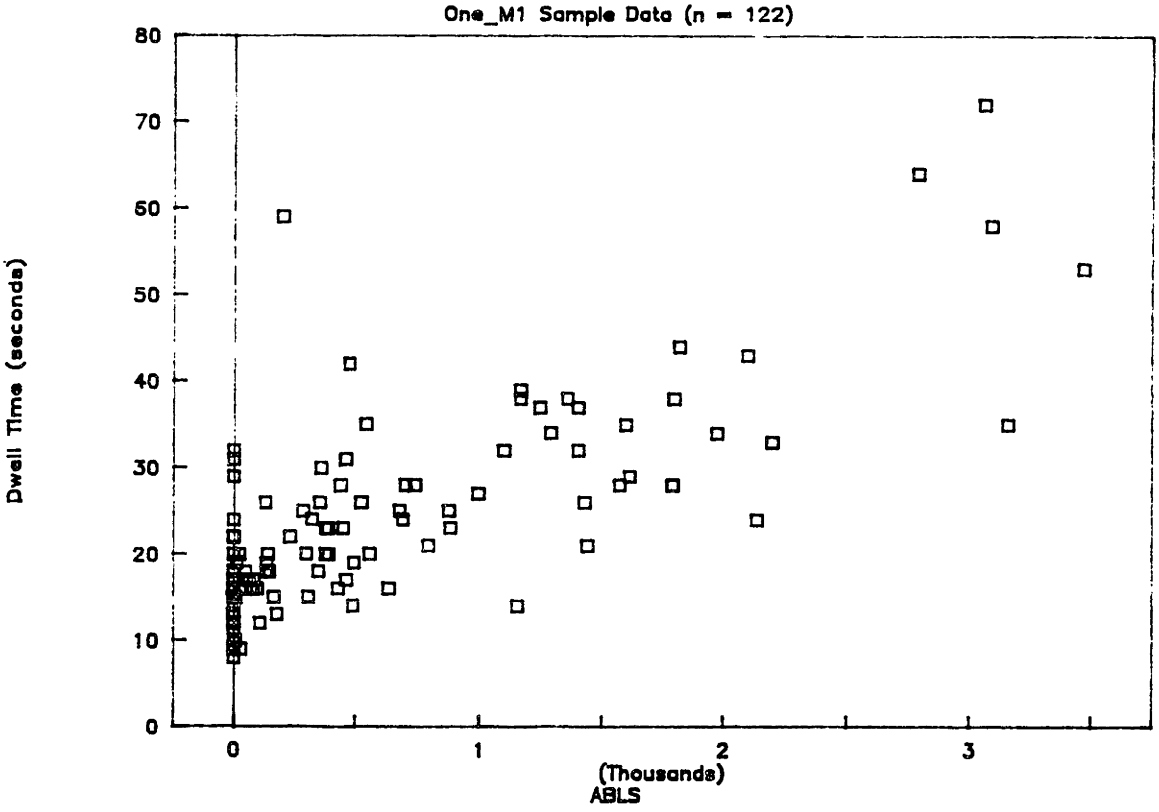


Figure A.5 DWELL TIME vs. MAXASLS

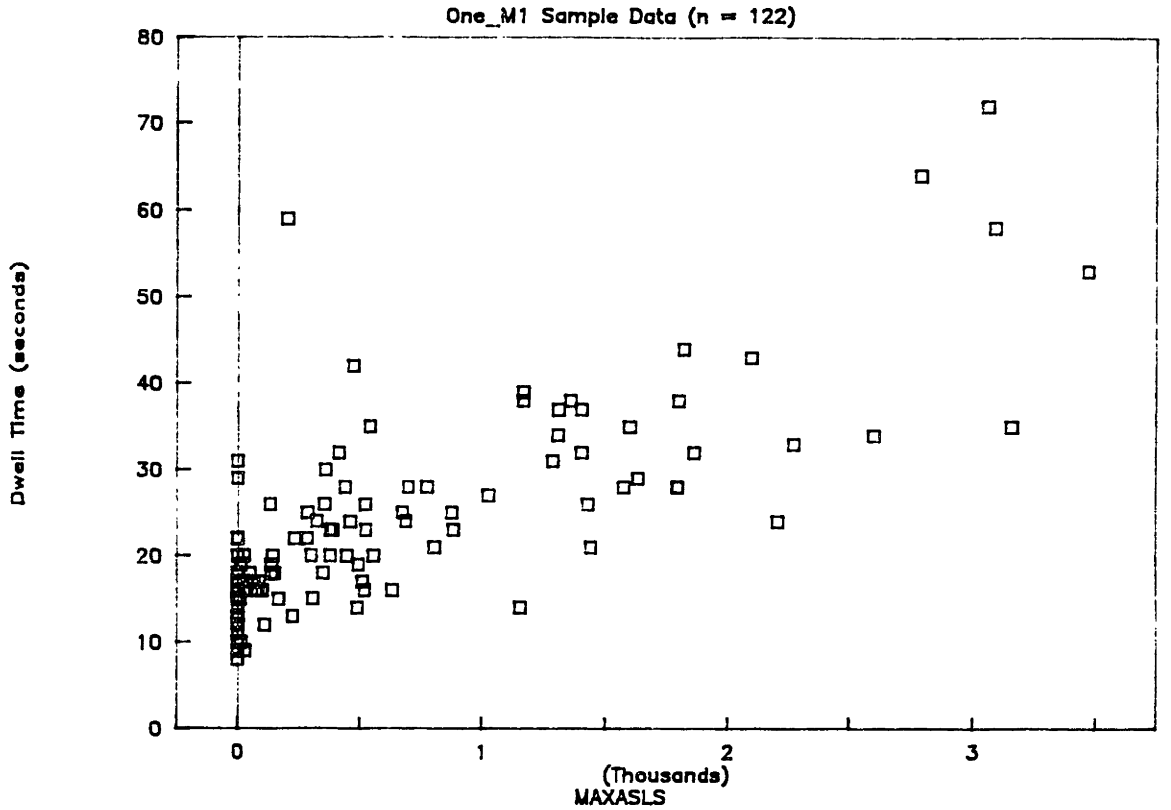


Figure A.6 DWELL TIME vs. SUMASLS

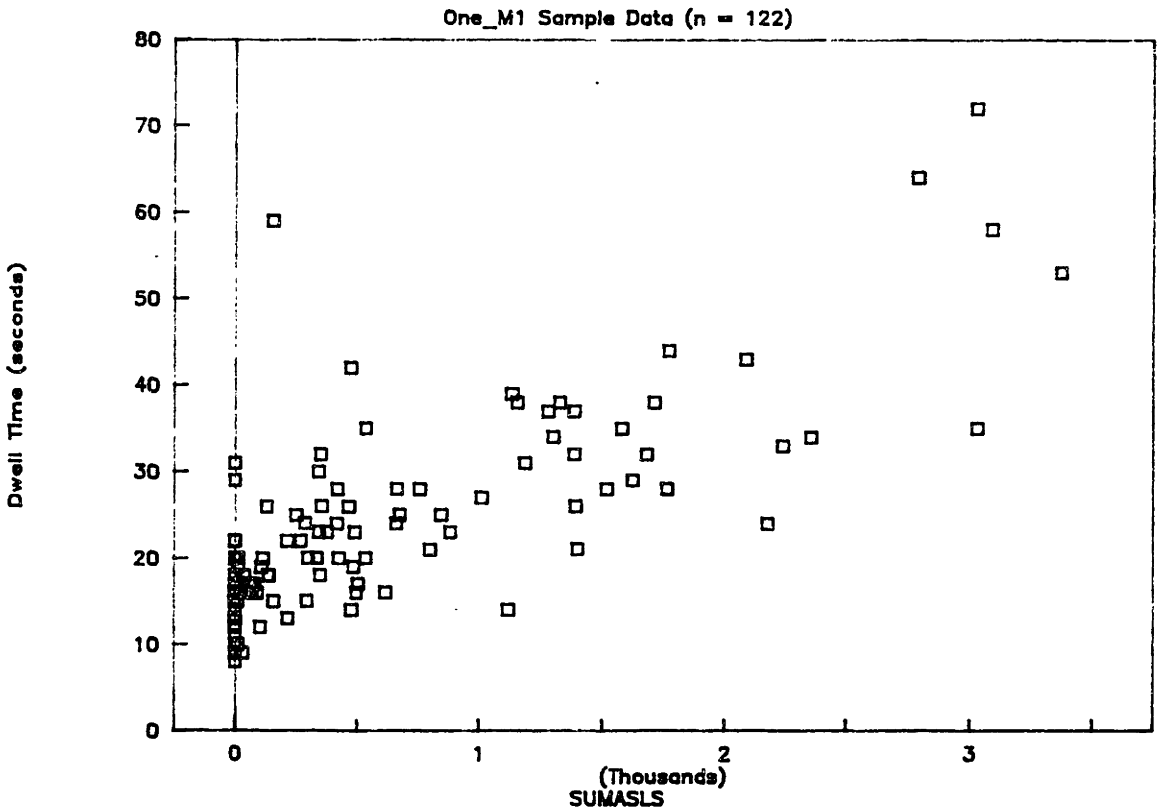


Figure A.7 DWELL TIME vs. ONLS

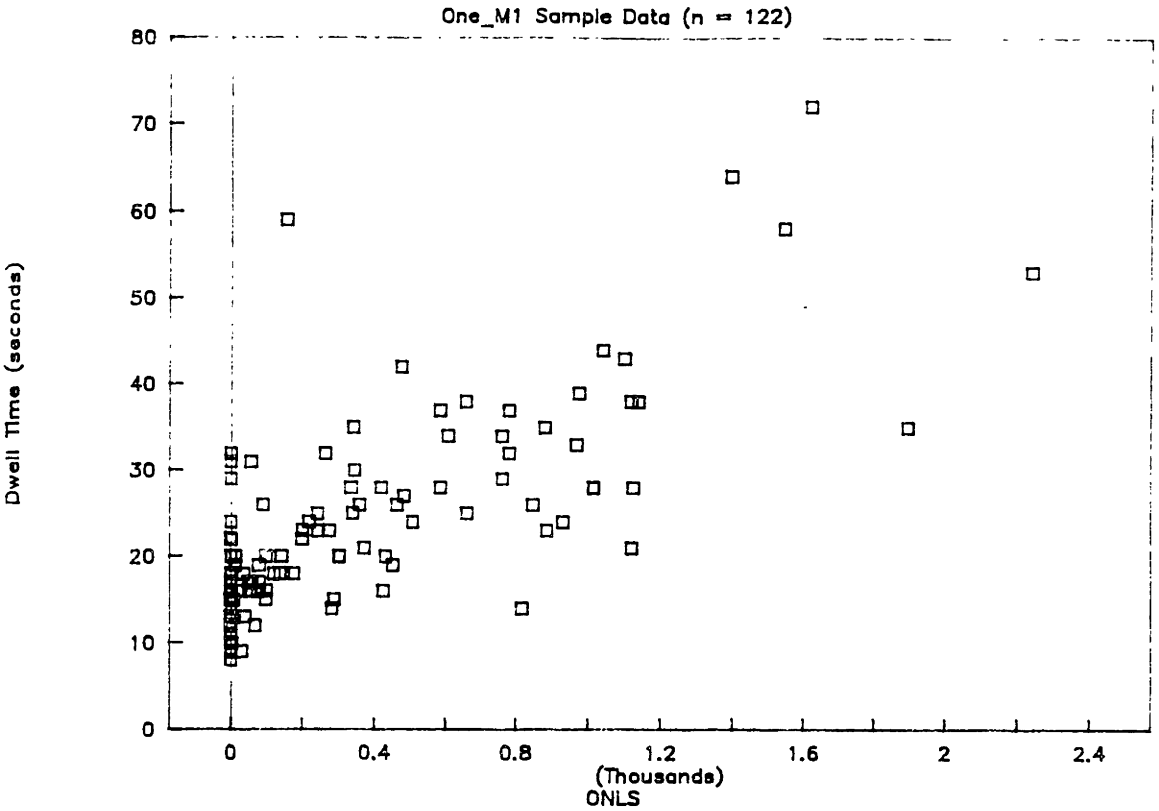


Figure A.8 DWELL TIME vs. OFFAS

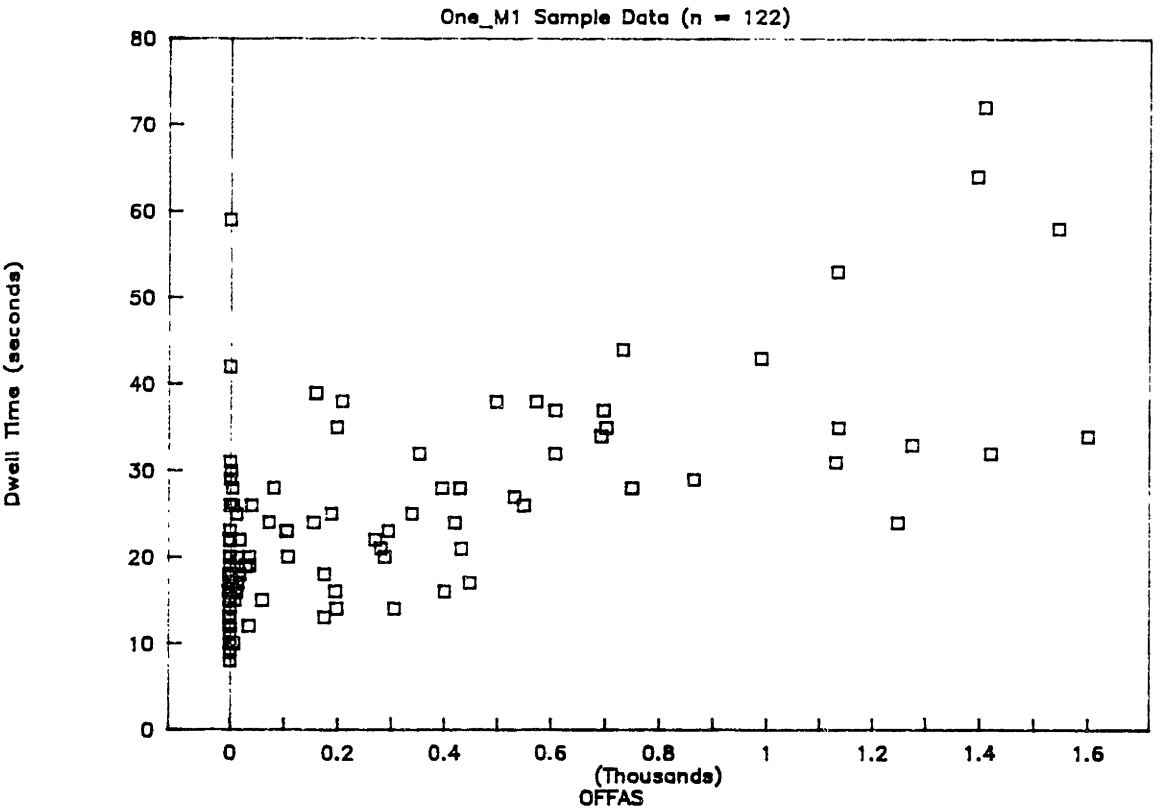


Figure A.9 DWELL TIME vs. LS

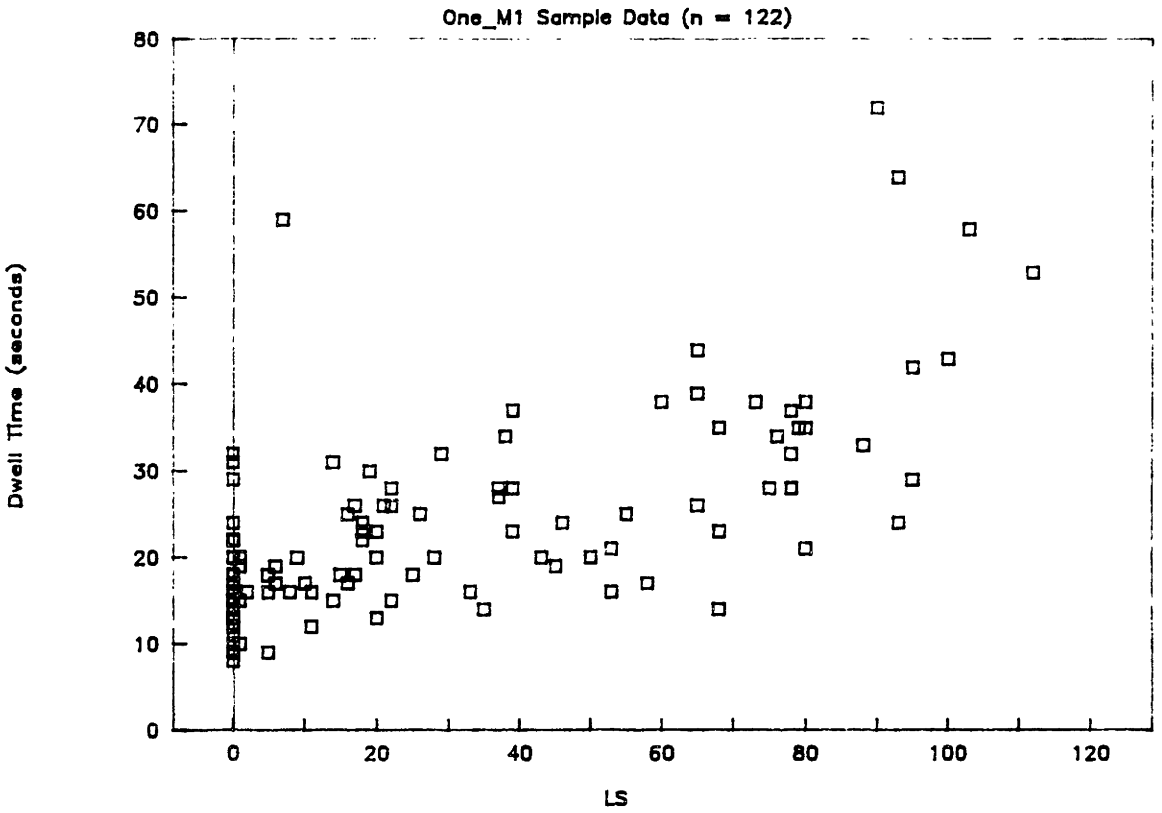


Figure A.10 DWELL TIME vs. AS

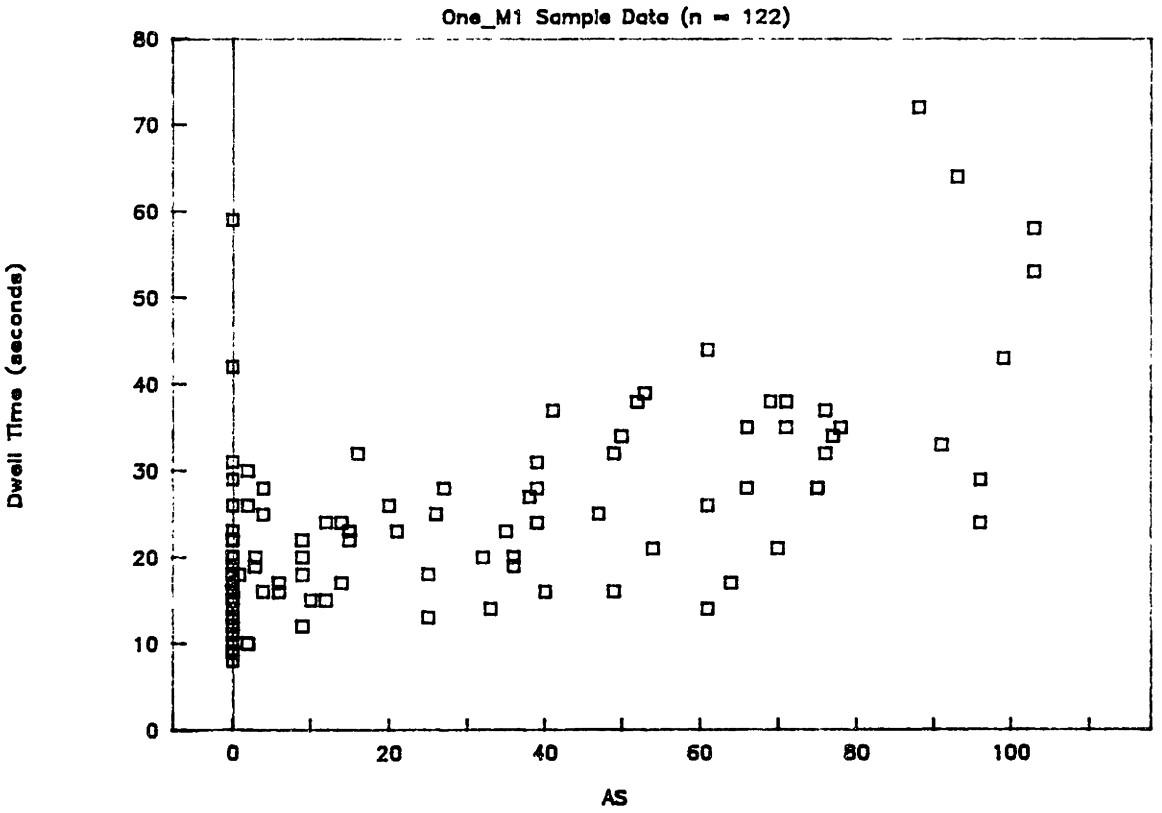


Figure A.11 DWELL TIME vs. ONS

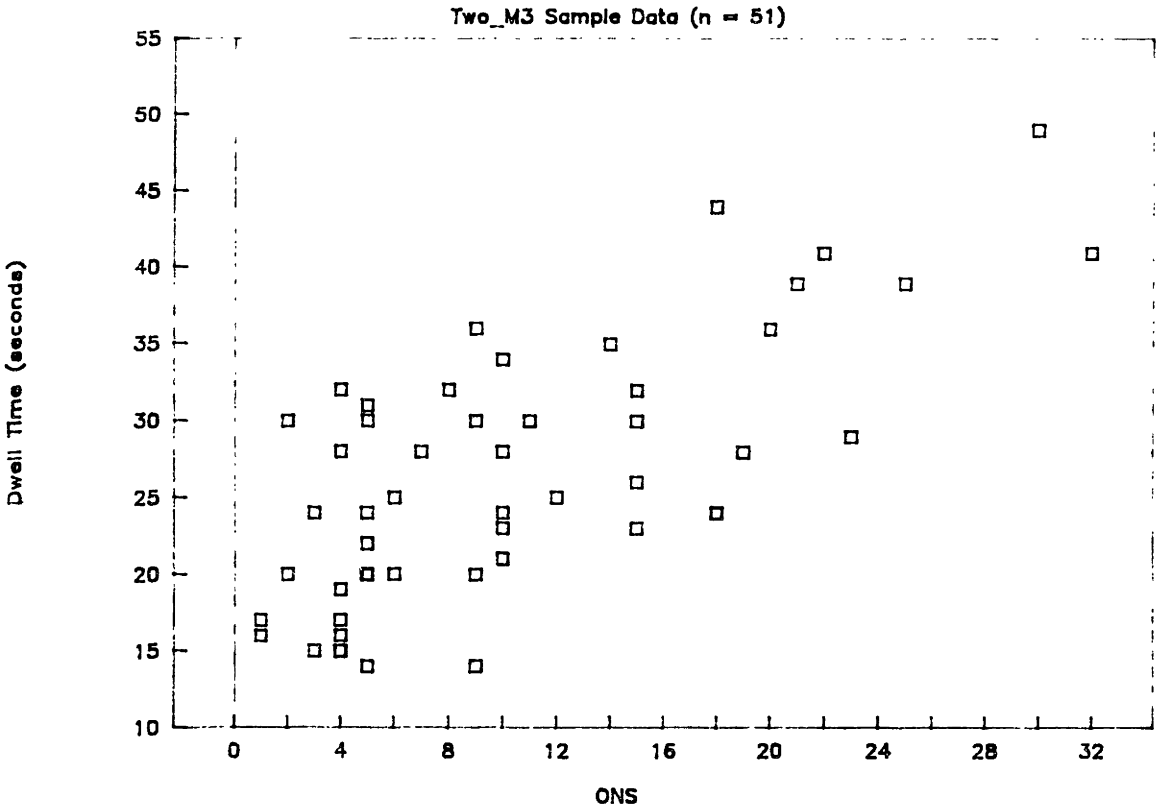


Figure A.12 DWELL TIME vs. OFFS

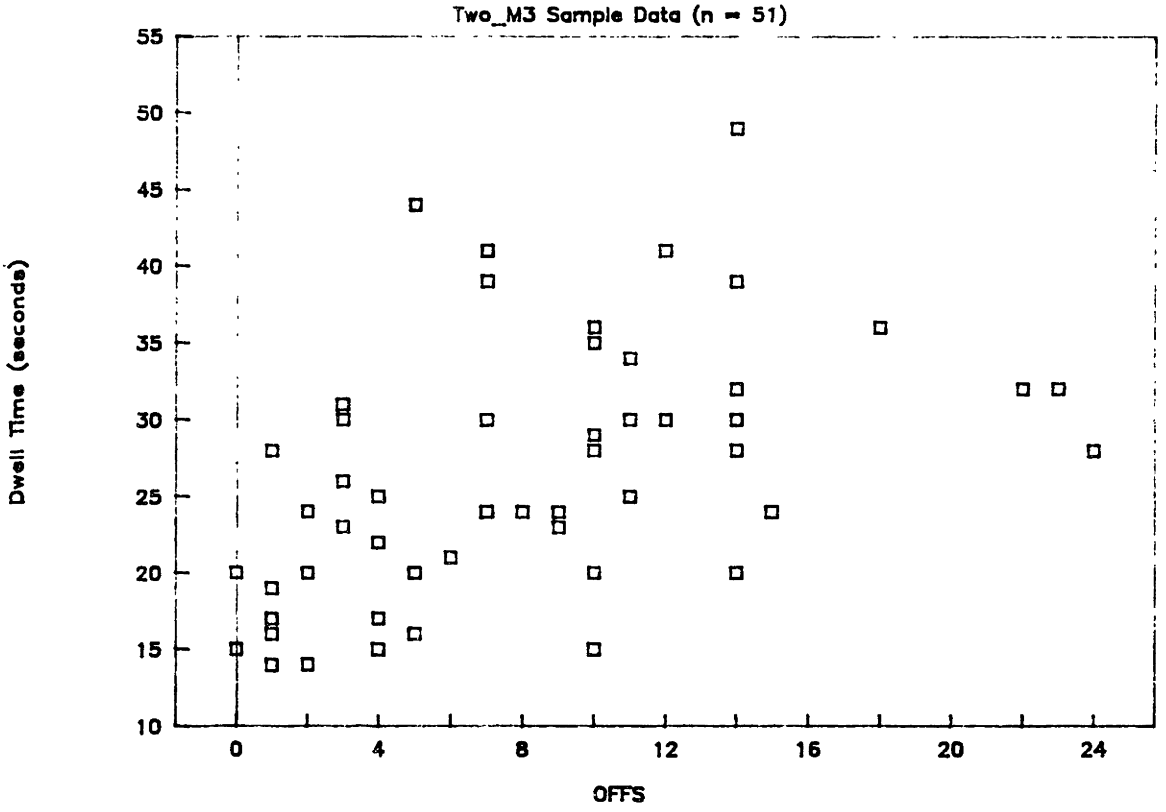


Figure A.13 DWELL TIME vs. ABAS

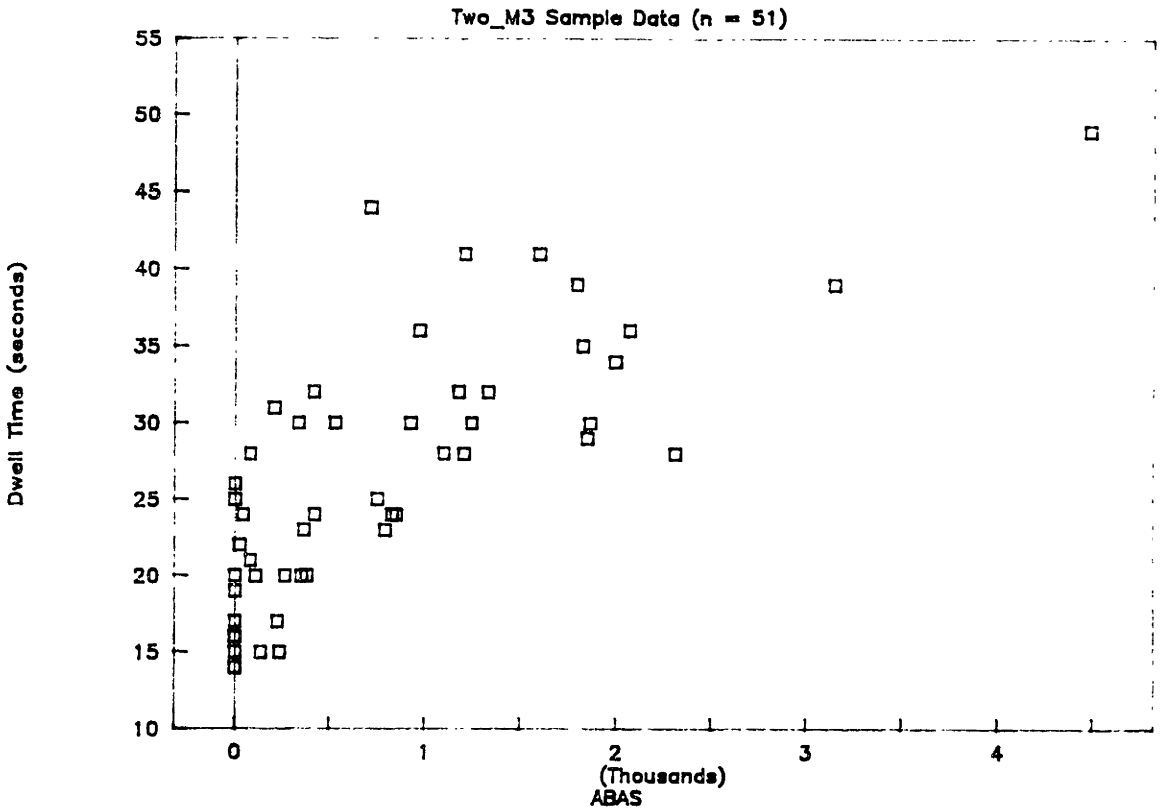


Figure A.14 DWELL TIME vs. ABLS

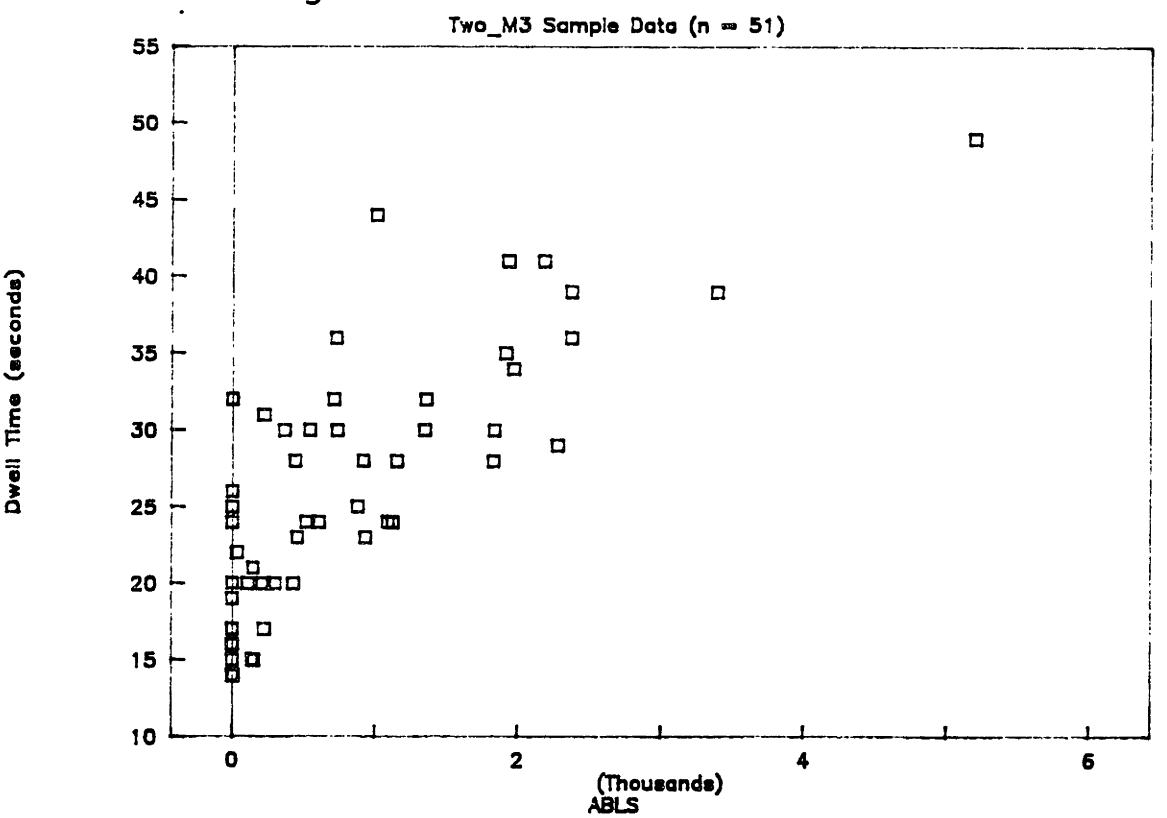


Figure A.15 DWELL TIME vs. MAXASLS

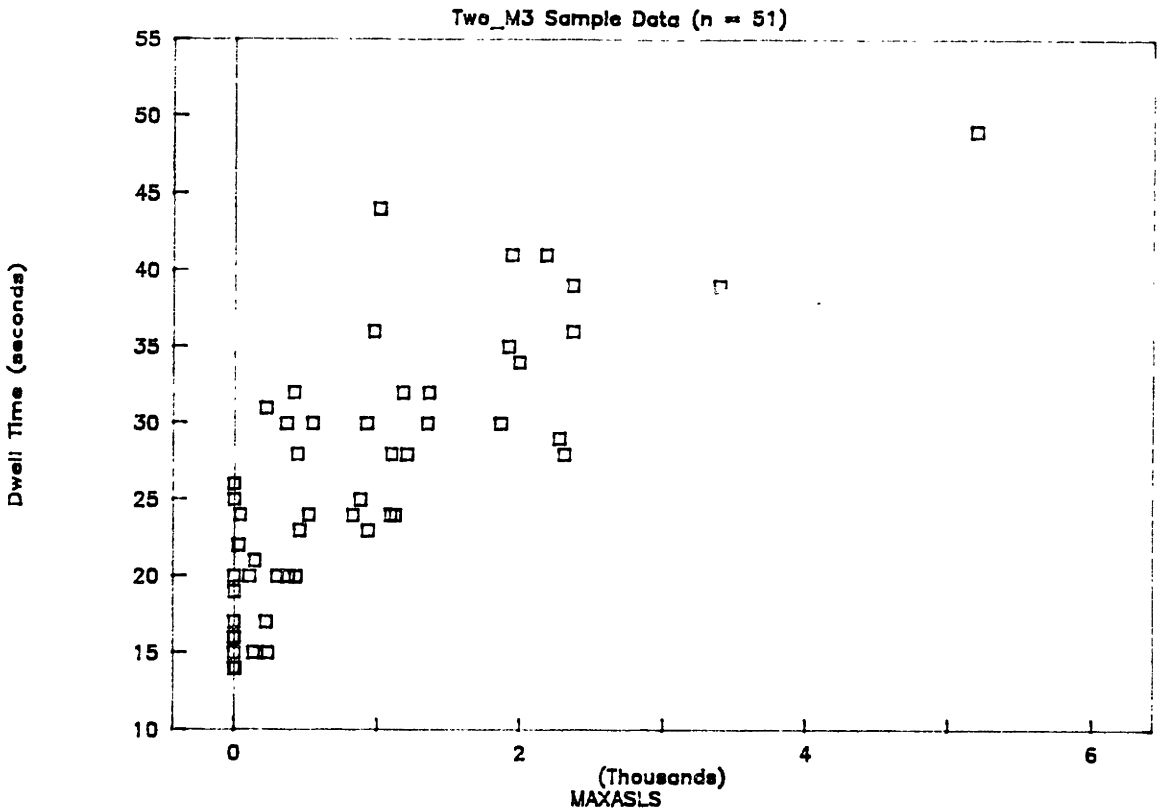


Figure A.16 DWELL TIME vs. SUMASLS

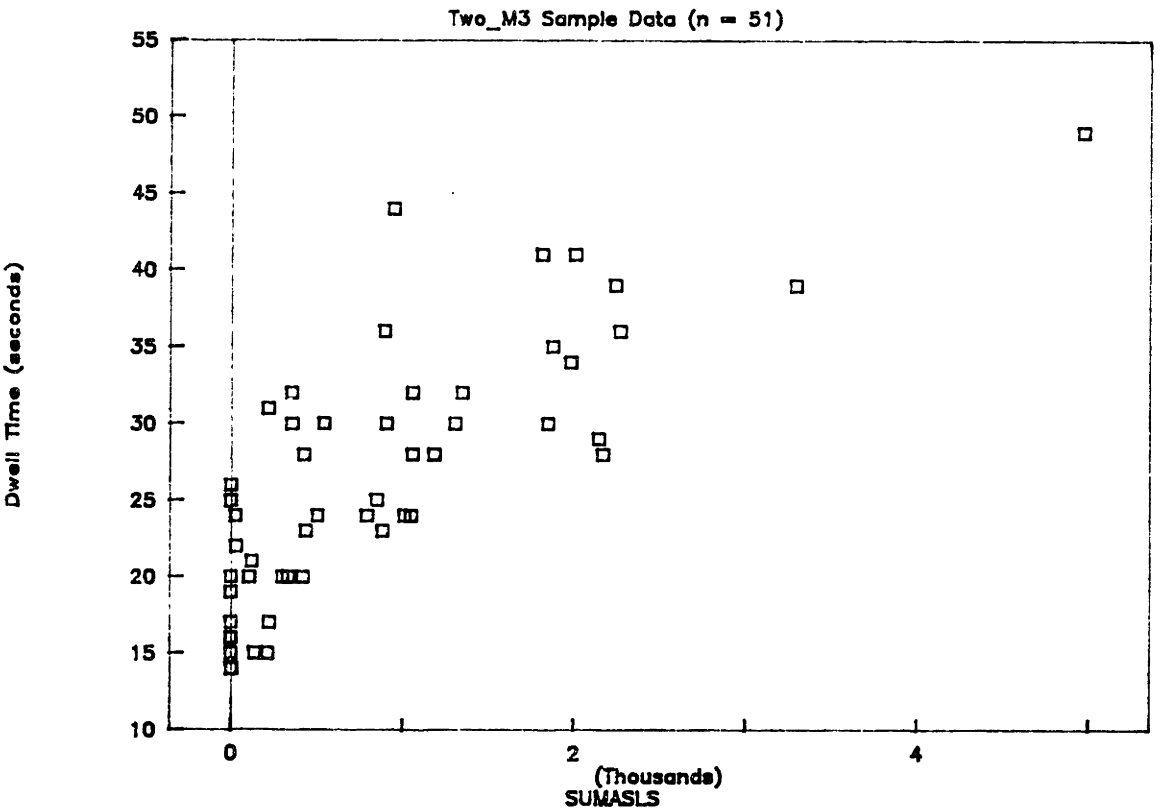


Figure A.17 DWELL TIME vs. ONLS

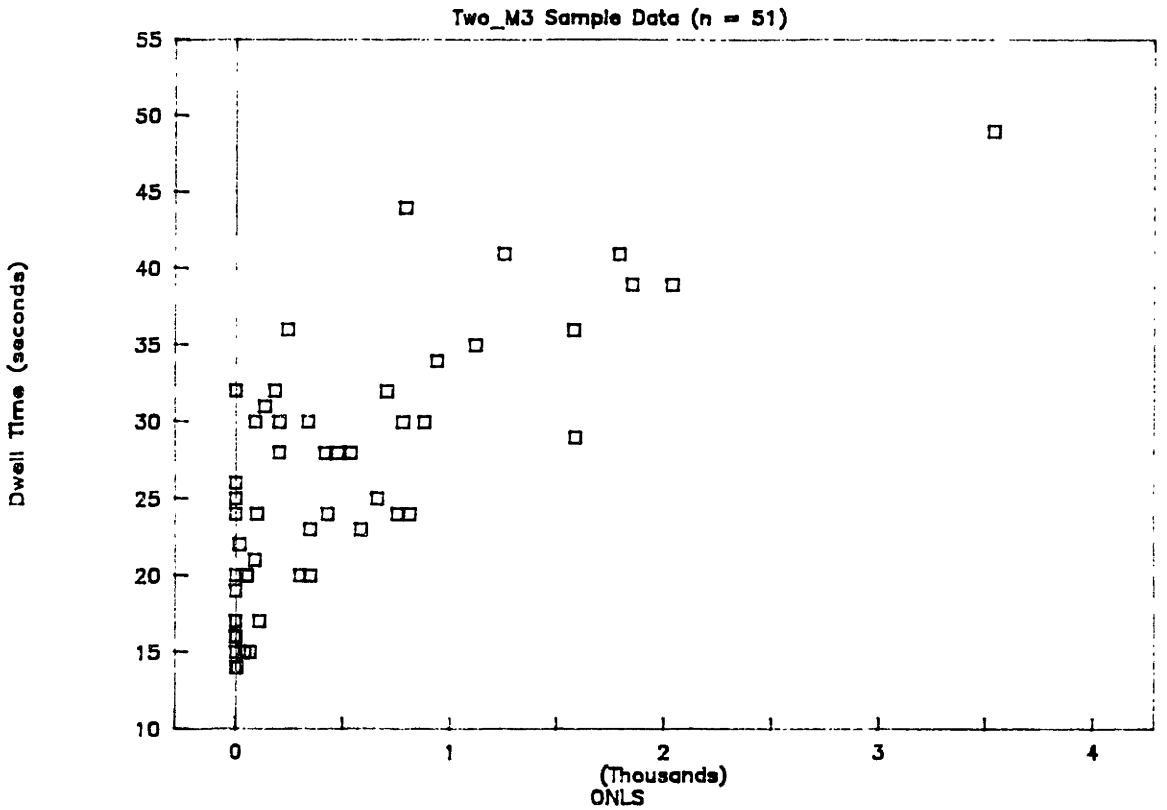


Figure A.18 DWELL TIME vs. OFFAS

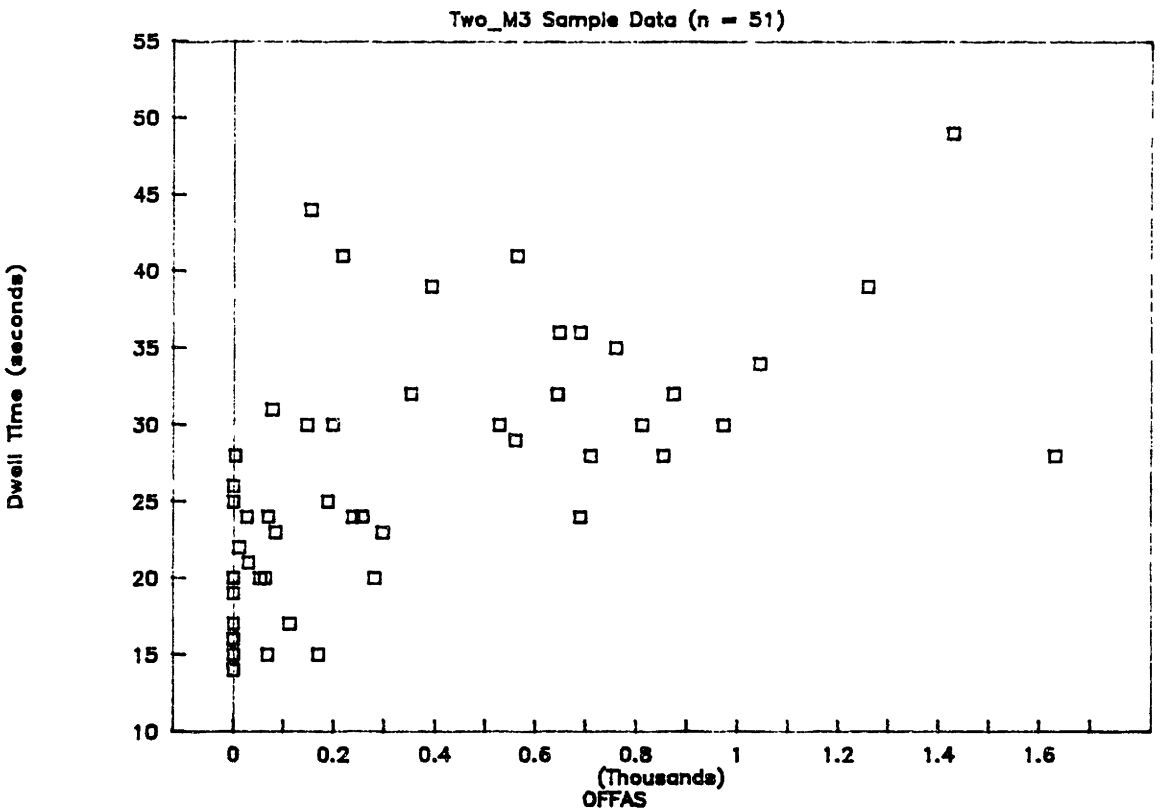


Figure A.19 DWELL TIME vs. LS

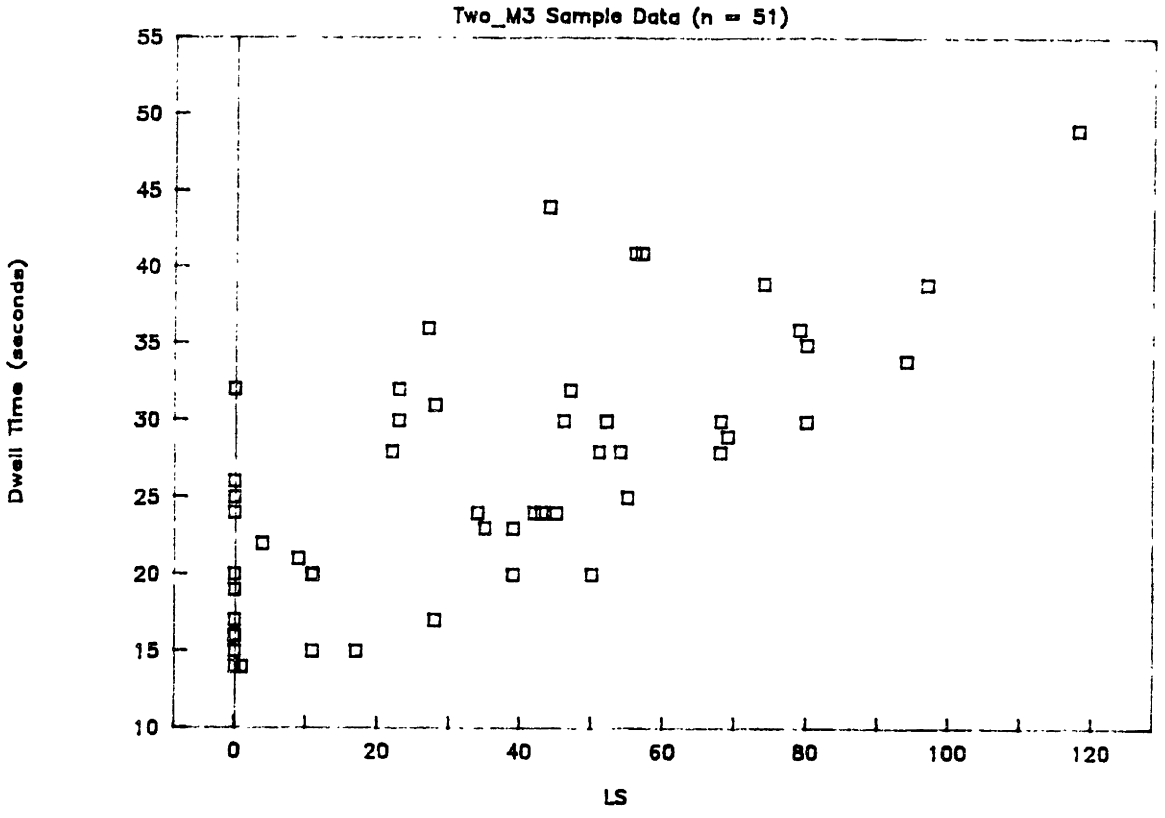
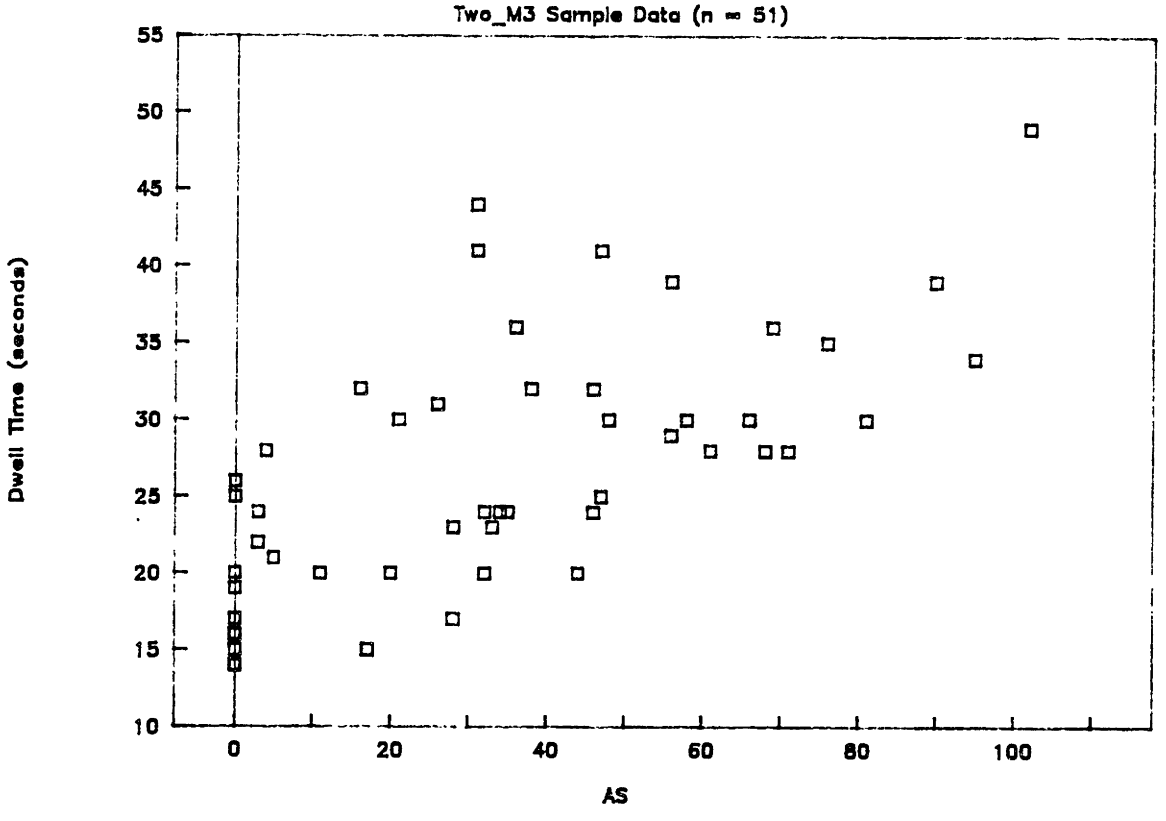


Figure A.20 DWELL TIME vs. AS



Appendix B

Correlation Matrix

Table B.1 Correlation Matrix for One_M1 Model

	DT	ONS	OFFS	ONOFFS	AS	LS	OFFAS	ONLS
ONS	0.599							
OFFS	0.433	0.117						
ONOFFS	0.678	0.692	0.798					
AS	0.654	0.410	0.394	0.535				
LS	0.679	0.493	0.202	0.446	0.918			
OFFAS	0.685	0.400	0.684	0.740	0.841	0.714		
ONLS	0.737	0.673	0.261	0.598	0.845	0.899	0.737	
SUMASLS	0.764	0.585	0.493	0.713	0.905	0.871	0.923	0.941
ABAS	0.737	0.516	0.552	0.714	0.914	0.828	0.954	0.892
ABLS	0.758	0.613	0.402	0.664	0.902	0.891	0.875	0.967
MAXASLS	0.762	0.592	0.506	0.727	0.899	0.863	0.926	0.937
DUMMY	-0.087	-0.252	0.192	-0.014	-0.046	-0.087	-0.046	-0.118

	SUMASLS	ABAS	ABLS	MAXASLS
ABAS	0.988			
ABLS	0.991	0.971		
MAXASLS	0.999	0.988	0.987	
DUMMY	-0.090	-0.074	-0.123	-0.088

MTB >

Table B.2 Correlation Matrix for One_M1a Model

	DT	ONS	OFFS	ONOFFS	AS	LS	OFFAS	ONLS
ONS	0.601							
OFFS	0.656	0.510						
ONOFFS	0.720	0.893	0.842					
AS	0.692	0.395	0.707	0.617				
LS	0.727	0.434	0.596	0.578	0.911			
OFFAS	0.771	0.447	0.834	0.716	0.882	0.795		
ONLS	0.757	0.633	0.686	0.755	0.881	0.910	0.869	
SUMASLS	0.788	0.573	0.773	0.763	0.911	0.890	0.956	0.976
ABAS	0.763	0.503	0.806	0.737	0.931	0.852	0.976	0.939
ABLS	0.787	0.583	0.775	0.770	0.909	0.888	0.952	0.979
MAXASLS	0.787	0.583	0.775	0.770	0.909	0.888	0.952	0.979
DUMMY	-0.058	-0.138	-0.026	-0.100	0.154	0.153	0.013	0.048

	SUMASLS	ABAS	ABLS	MAXASLS
ABAS	0.987			
ABLS	1.000	0.986		
MAXASLS	1.000	0.986	1.000	
DUMMY	0.034	0.034	0.030	0.030

MTB >

Table B.3 Correlation Matrix for One_M1b Model

	DT	ONS	OFFS	ONOFFS	AS	LS	OFFAS	ONLS
ONS	0.723							
OFFS	0.678	0.459						
ONOFFS	0.795	0.776	0.930					
AS	0.564	0.578	0.263	0.434				
LS	0.491	0.566	0.105	0.314	0.978			
OFFAS	0.691	0.714	0.639	0.763	0.829	0.731		
ONLS	0.610	0.782	0.355	0.514	0.878	0.898	0.804	
SUMASLS	0.693	0.777	0.516	0.701	0.891	0.836	0.971	0.923
ABAS	0.697	0.781	0.543	0.722	0.874	0.812	0.979	0.907
ABLS	0.628	0.786	0.340	0.577	0.908	0.901	0.888	0.982
MAXASLS	0.697	0.781	0.543	0.722	0.874	0.812	0.979	0.907
DUMMY	-0.091	-0.271	0.248	0.065	-0.497	-0.560	-0.247	-0.524

	SUMASLS	ABAS	ABLS	MAXASLS
ABAS	0.999			
ABLS	0.971	0.961		
MAXASLS	0.999	1.000	0.961	
DUMMY	-0.371	-0.345	-0.469	-0.345

MTB >

Table B.4 Correlation Matrix for Two_M1 Model

	DT	ONS	OFFS	ONOFFS	LS	AS	ABAS	ABLS
ONS	0.670							
OFFS	0.573	0.121						
ONOFFS	0.832	0.841	0.638					
LS	0.696	0.703	0.374	0.749				
AS	0.691	0.574	0.508	0.722	0.977			
ABAS	0.777	0.693	0.551	0.837	0.936	0.938		
ABLS	0.780	0.787	0.441	0.851	0.942	0.907	0.984	
MAXASLS	0.794	0.769	0.491	0.864	0.940	0.917	0.990	0.998
OFFAS	0.697	0.418	0.747	0.731	0.810	0.892	0.908	0.829
ONLS	0.739	0.867	0.270	0.820	0.897	0.815	0.910	0.966
SUMASLS	0.789	0.762	0.490	0.858	0.943	0.921	0.992	0.998

	MAXASLS	OFFAS	ONLS
OFFAS	0.854		
ONLS	0.954	0.660	
SUMASLS	1.000	0.859	0.952

MTB >
MTB >

Table B.5 Correlation Matrix for Two_M1a Model

	DT	ONS	OFFS	ONOFFS	LS	AS	ABAS	ABLS
ONS	0.759							
OFFS	0.732	0.532						
ONOFFS	0.846	0.938	0.793					
LS	0.717	0.701	0.736	0.807				
AS	0.689	0.609	0.775	0.756	0.988			
ABAS	0.780	0.713	0.815	0.848	0.943	0.944		
ABLS	0.803	0.792	0.784	0.891	0.942	0.922	0.990	
MAXASLS	0.803	0.792	0.784	0.891	0.942	0.922	0.990	1.000
OFFAS	0.730	0.556	0.883	0.762	0.874	0.912	0.960	0.918
ONLS	0.789	0.851	0.693	0.697	0.928	0.883	0.953	0.985
SUMASLS	0.799	0.782	0.785	0.885	0.945	0.927	0.992	1.000

	MAXASLS	OFFAS	ONLS
OFFAS	0.918		
ONLS	0.985	0.835	
SUMASLS	1.000	0.924	0.982

```
MTB >
MTB >
MTB >
```

Table B.6 Correlation Matrix for Two_M1b Model

	DT	ONS	OFFS	ONOFFS	LS	AS	ABAS	ABLS
ONS	0.652							
OFFS	0.936	0.691						
ONOFFS	0.822	0.894	0.941					
LS	0.656	0.578	0.533	0.600				
AS	0.709	0.552	0.604	0.631	0.991			
ABAS	0.761	0.774	0.712	0.802	0.936	0.934		
ABLS	0.707	0.785	0.642	0.764	0.938	0.920	0.993	
MAXASLS	0.761	0.774	0.712	0.802	0.936	0.934	1.000	0.993
OFFAS	0.783	0.694	0.741	0.783	0.943	0.958	0.986	0.965
ONLS	0.660	0.844	0.596	0.763	0.874	0.841	0.960	0.979
SUMASLS	0.752	0.773	0.700	0.794	0.938	0.933	1.000	0.995

	MAXASLS	OFFAS	ONLS
OFFAS	0.986		
ONLS	0.960	0.900	
SUMASLS	1.000	0.984	0.964

```
MTB >
MTB >
MTB >
```

Table B.7 Correlation Matrix for Two_M2 Model

	DT	ONS	OFFS	ONOFFS	LS	AS	ABAS	ABLS
ONS	0.662							
OFFS	0.495	0.191						
ONOFFS	0.759	0.825	0.712					
LS	0.695	0.616	0.307	0.617				
AS	0.668	0.468	0.450	0.594	0.970			
ABAS	0.754	0.672	0.568	0.807	0.875	0.876		
ABLS	0.762	0.781	0.436	0.810	0.883	0.833	0.976	
MAXASLS	0.774	0.757	0.515	0.837	0.872	0.842	0.988	0.994
OFFAS	0.641	0.375	0.765	0.708	0.719	0.819	0.889	0.782
ONLS	0.726	0.857	0.266	0.766	0.835	0.731	0.890	0.963
SUMASLS	0.772	0.750	0.507	0.828	0.890	0.851	0.990	0.995

	MAXASLS	OFFAS	ONLS
OFFAS	0.828		
ONLS	0.941	0.592	
SUMASLS	1.000	0.831	0.940

```
MTB >
MTB >
MTB >
```

Table B.8 Correlation Matrix for Two_M3 Model

	DT	ONS	OFFS	ONOFFS	LS	AS	ABAS	ABLS
ONS	0.712							
OFFS	0.479	0.120						
ONOFFS	0.810	0.813	0.676					
LS	0.718	0.637	0.298	0.648				
AS	0.672	0.455	0.471	0.614	0.959			
ABAS	0.753	0.660	0.548	0.812	0.880	0.879		
ABLS	0.773	0.786	0.398	0.817	0.889	0.824	0.973	
MAXASLS	0.786	0.757	0.486	0.847	0.878	0.837	0.986	0.994
OFFAS	0.616	0.348	0.763	0.706	0.733	0.848	0.886	0.767
ONLS	0.753	0.865	0.231	0.778	0.831	0.706	0.888	0.967
SUMASLS	0.782	0.749	0.479	0.837	0.885	0.846	0.988	0.994

	MAXASLS	OFFAS	ONLS
OFFAS	0.820		
ONLS	0.943	0.583	
SUMASLS	1.000	0.823	0.941

```
MTB >
MTB >
MTB >
```

Appendix C

Histograms of Residuals &

Standarized Residuals vs. Predicted Dwell Times

Figure C.1 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model C1-1, Linear)

```
Histogram of C16   N = 122

Midpoint   Count
  -25.00       0
  -20.00       1  *
  -15.00       3  ***
  -10.00       4  ****
   -5.00      39  ***************************************
    0.00      48  ************************************************
    5.00      15  ***************
   10.00       6  ******
   15.00       2  **
   20.00       2  **
   25.00       1  *
   30.00       1  *

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >


MTB > PLOT C14 C15
```

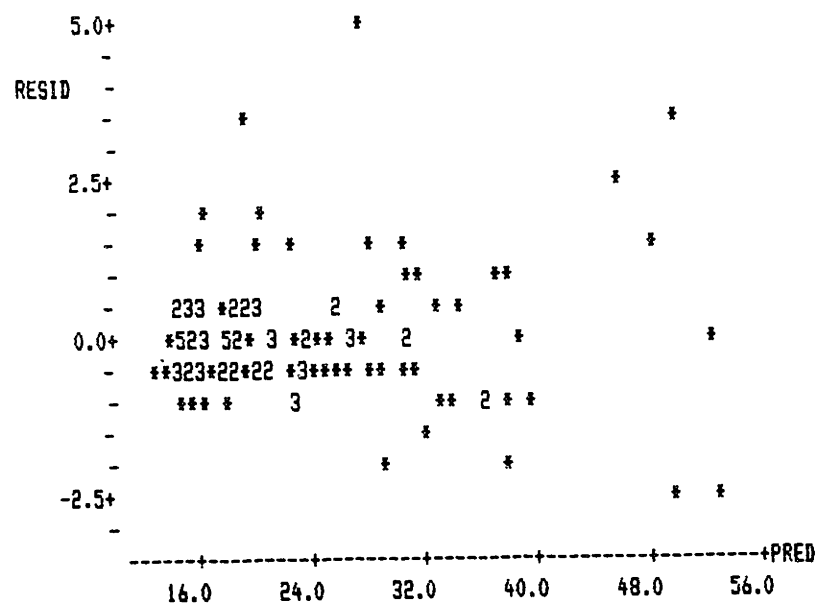


Figure C.2 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model D1-3, Nonlinear, E = 2.5)

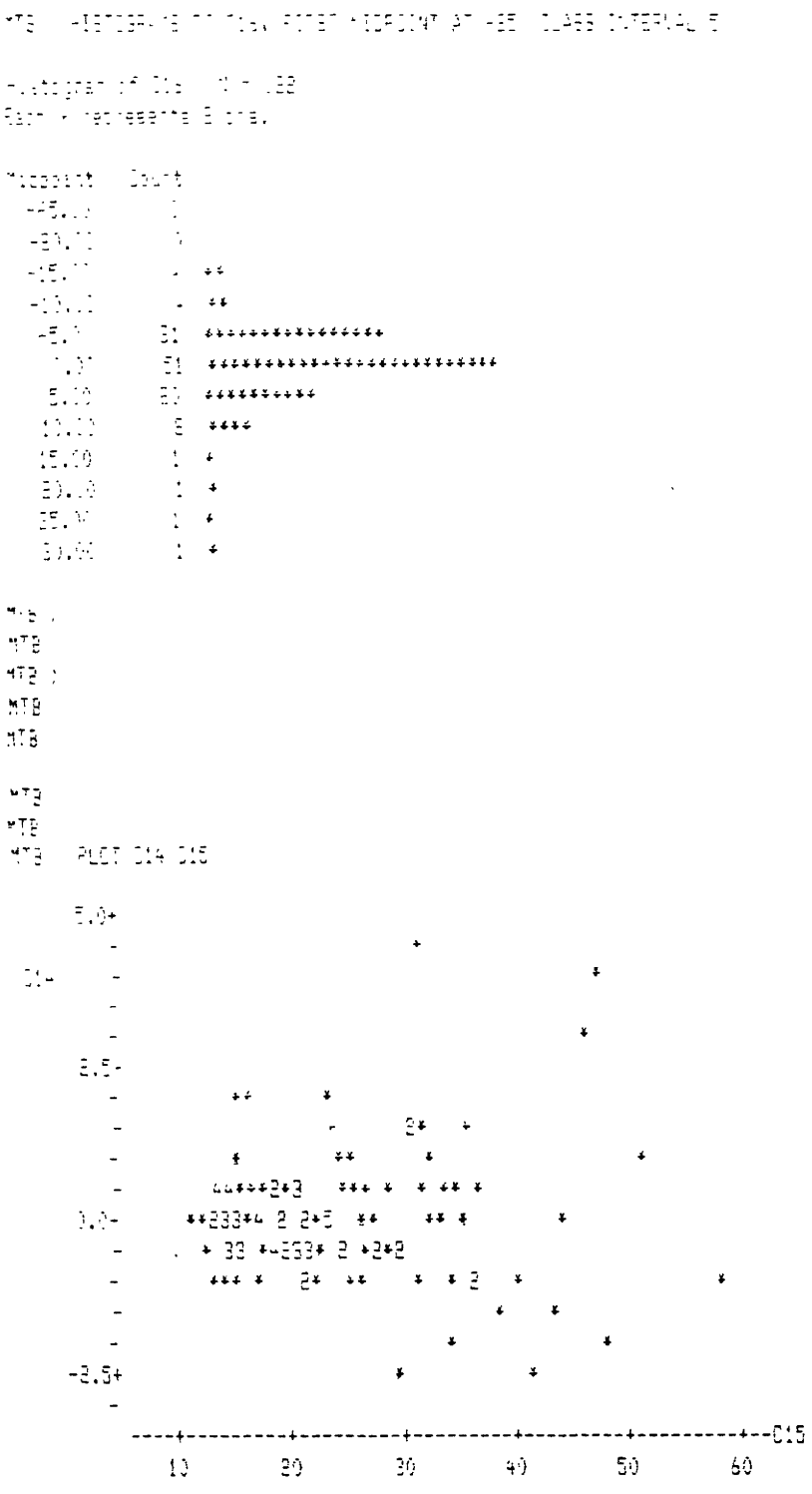


Figure C.3 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model B1a1, Linear)

```
MTB > HISTOGRAMS OF C15, FIRST MIDPOINT AT -25, CLASS INTERVAL 5

Histogram of RESIDUAL   N = 83

Midpoint   Count
  -25.00       0
  -20.00       1  *
  -15.00       2  **
  -10.00       5  *****
   -5.00      18  ******************
    0.00      39  ****************************************
    5.00      11  ***********
   10.00       2  **
   15.00       2  **
   20.00       2  **
   25.00       0
   30.00       0
   35.00       1  *

MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15
```

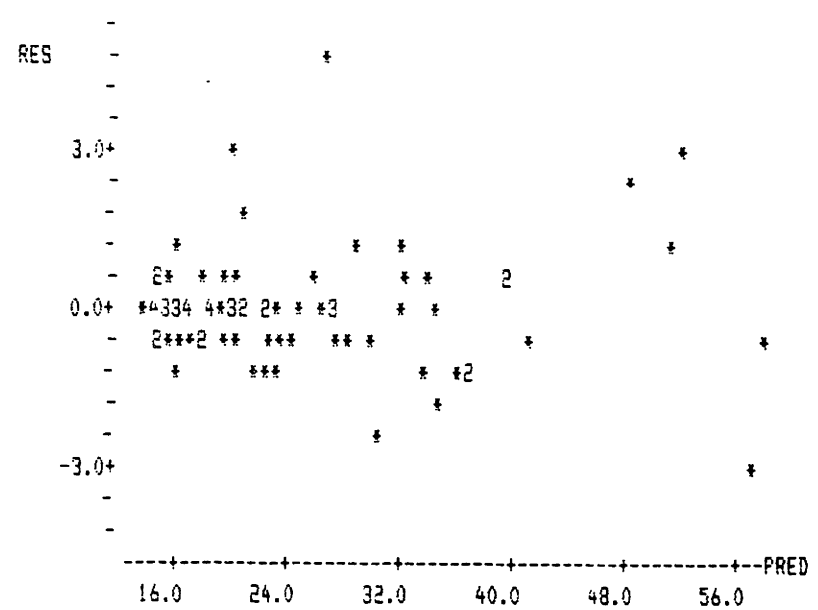


Figure C.4 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model D1-3, Nonlinear, E = 2.7)

```
Histogram of RESIDUAL   N = 33

Midpoint   Count
  -25.00       0
  -20.00       0
  -15.00       2  **
  -10.00       6  ******
   -5.00      21  *********************
    0.00      32  ********************************
    5.00      15  ***************
   10.00       4  ****
   15.00       0
   20.00       1  *
   25.00       1  *
   30.00       1  *

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15

     5.0+
        -                    *
 RES    -
        -                                *
        -
     2.5+                                *
        -
        -    *         *        *
        -    *              * *                    *
        -   32   ** * *  **     **   *
     0.0+  *32*54 *2**** * *    * *  *
        -   * ** *2**4**     2*              *
        -   ***      2        *    3
        -                                   *
        -                                                  *
    -2.5+                    *       *
        -
          +---------+---------+---------+---------+---------+------PRED
         10        20        30        40        50        60

MTB >
```

Figure C.5 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model C1b3, Linear)

```
Histogram of RESIDUAL   N = 39

Midpoint   Count
  -25.00       0
  -20.00       0
  -15.00       0
  -10.00       1  *
   -5.00      12  ************
    0.00      18  ******************
    5.00       4  ****
   10.00       4  ****

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15
```

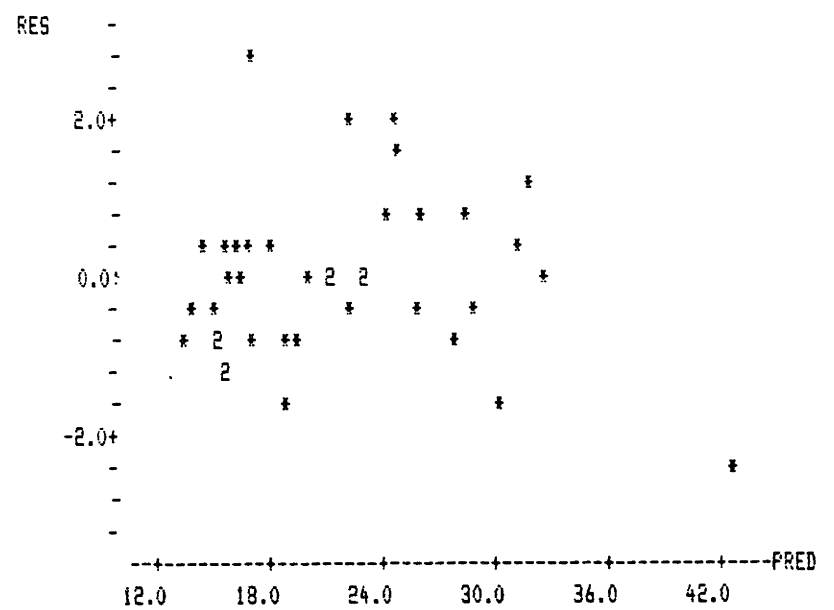


Figure C.6 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model C1b3(B), Nonlinear, E = 2.0)

```
Histogram of RESIDUAL   N = 39

Midpoint   Count
  -25.00       0
  -20.00       0
  -15.00       0
  -10.00       1  *
   -5.00      12  ************
    0.00      19  *******************
    5.00       4  ****
   10.00       3  ***

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15
```

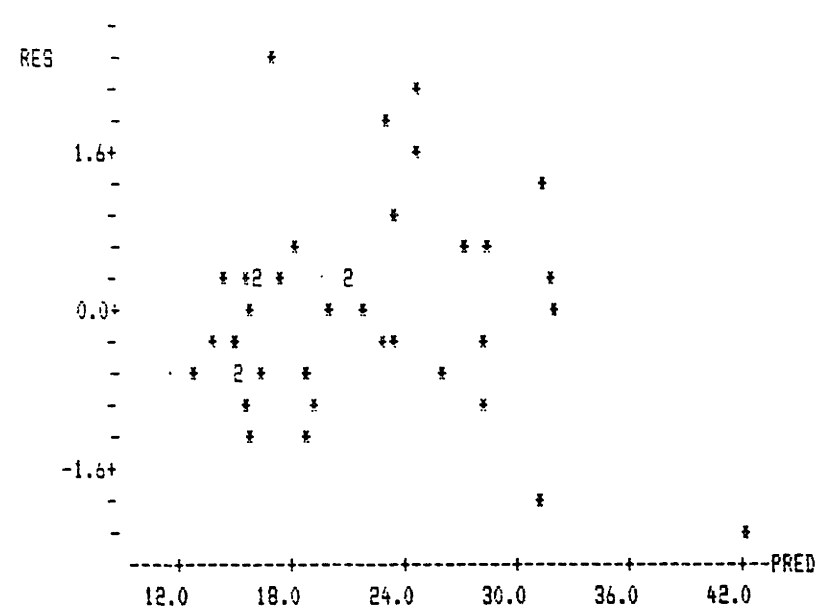


Figure C.7 (a) Histograms of Residuals
(b) Standardized Residuals vs. Predicted Dwell Times
(Model C1-3, Linear)

```
Histogram of PMAX   N = 51

Midpoint   Count
  -25.00       0
  -20.00       0
  -15.00       0
  -10.00       1  *
   -5.00      17  *****************
    0.00      20  ********************
    5.00      10  **********
   10.00       2  **
   15.00       1  *

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15
```

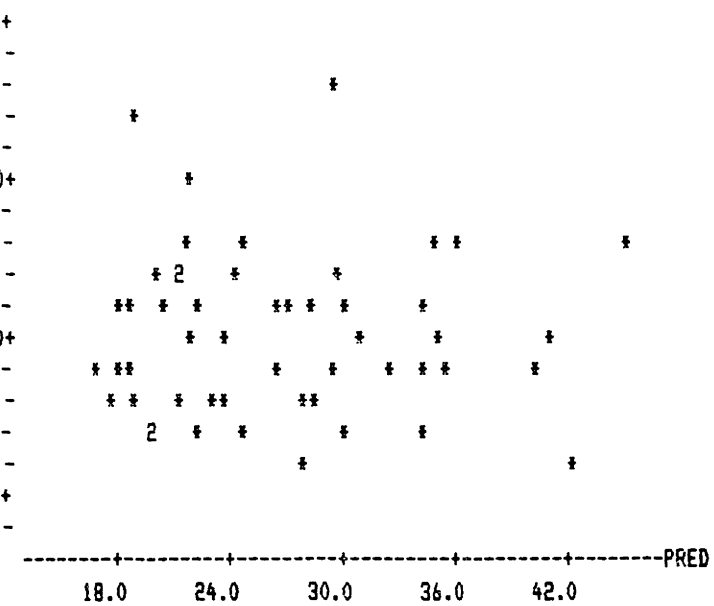


Figure C.8 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model C1-1, Nonlinear, E = 1.5)

```
MTB : HISTOGRAM OF C13, FIRST MIDPOINT AT -25, CLASS INTERVAL 5

Histogram of RESIDUAL   N = 51

Midpoint   Count
  -25.00       0
  -20.00       0
  -15.00       0
  -10.00       0
   -5.00      18  ******************
    0.00      22  **********************
    5.00       8  ********
   10.00       2  **
   15.00       1  *

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15
```

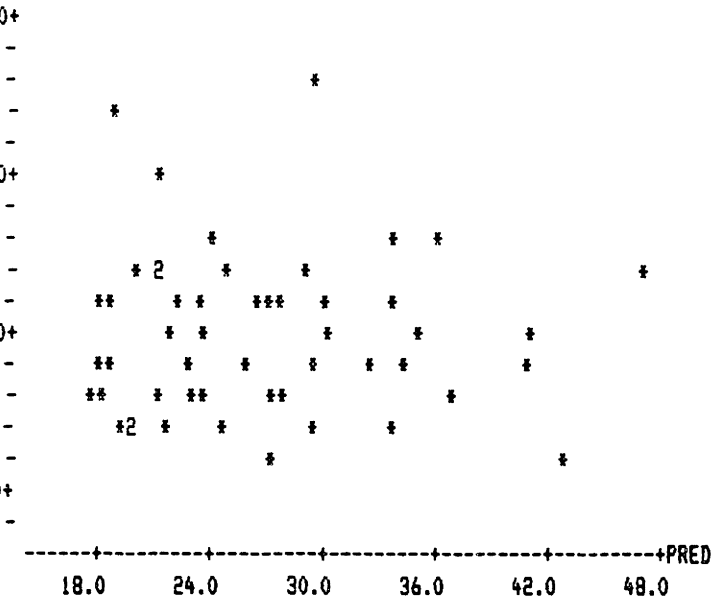


Figure C.9 (a) Histograms of Residuals
(b) Standardized Residuals vs. Predicted Dwell Times
(Model A1a, Linear)

```
Histogram of C16   N = 32

Midpoint   Count
  -25.00       0
  -20.00       0
  -15.00       0
  -10.00       0
   -5.00      11  ***********
    0.00      12  ************
    5.00       7  *******
   10.00       1  *
   15.00       1  *

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15
```

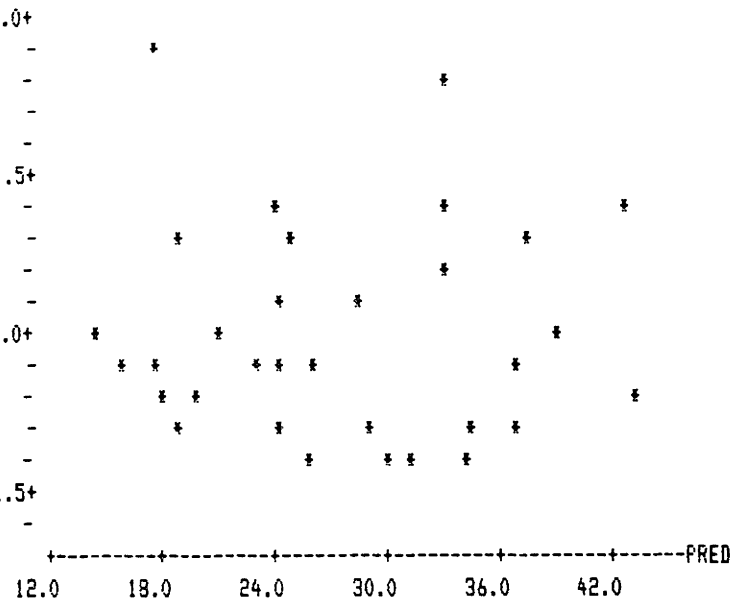


Figure C.10 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model B1b1, Linear)

```
Histogram of RESIDUAL   N = 19

Midpoint   Count
  -25.00       0
  -20.00       0
  -15.00       0
  -10.00       0
   -5.00       4  ****
    0.00      12  ************
    5.00       2  **
   10.00       1  *

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15
```

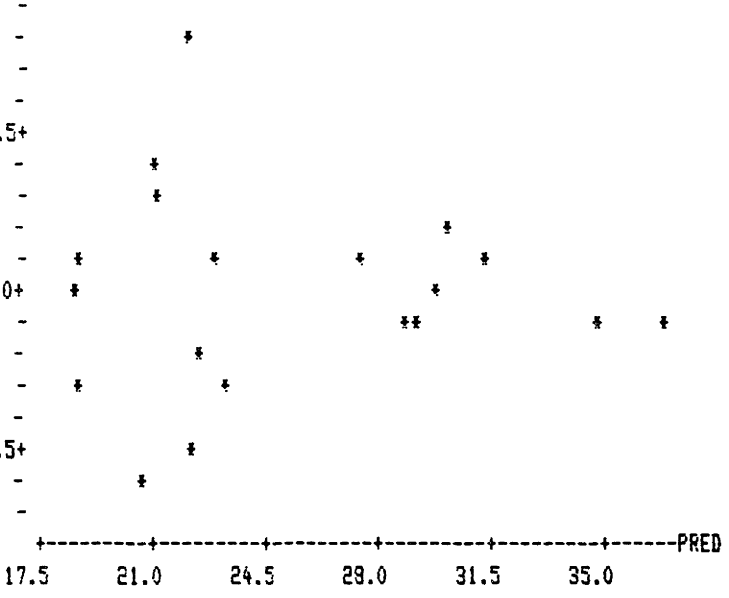


Figure C.11 **(a) Histograms of Residuals**
(b) Standardized Residuals vs. Predicted Dwell Times
(Model D1b3, Nonlinear, E = 0.6)

```
MTB > HISTOGRAMS OF C13, FIRST MIDPOINT AT -25, CLASS INTERVAL 5

Histogram of RESIDUAL   N = 19

Midpoint   Count
  -25.00       0
  -20.00       0
  -15.00       0
  -10.00       0
   -5.00       3  ***
    0.00      13  *************
    5.00       3  ***

MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >
MTB >

MTB >
MTB >
MTB > PLOT C14 C15
```