INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

DETERMINANTS OF SUBWAY TRAVEL IN THE NEW YORK CITY METROPOLITAN AREA: AN EMPIRICAL RESEARCH AND ECONOMETRIC APPLICATION OF DISCRETE CHOICE AND TIME SERIES MODELS TO URBAN TRAVEL DEMAND.

By

MICHEL EMANUEL HANTAR

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2000

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

[signature]

JAN 25, 2000

Date

Chair of Examining Committee

[signature]

JAN. 25, 2000

Date

Executive Officer

Professor Theodore Joyce

Professor Michael Grossman

Professor Linda Edwards

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

DETERMINANTS OF SUBWAY TRAVEL IN THE NEW YORK CITY METROPOLITAN AREA: AN EMPIRICAL RESEARCH AND ECONOMETRIC APPLICATION OF DISCRETE CHOICE AND TIME SERIES MODELS TO URBAN TRAVEL DEMAND.

By

MICHEL EMANUEL HANTAR

Adviser: Professor Theodore Joyce

This dissertation deals with the modeling of transportation decisions in the New York City Area. Three different data sets are used: cross sectional data for subway stations, micro sample data and time series data. The main model estimated is a logistic model developed earlier by McFadden and the independent variables are socio-economic variables, mode characteristic variables, and dummies representing high traffic stations, seasonal indexes, and the 1980 strike effect. Each data set addresses a particular question. By using the first data set we find out that the estimation of the subway ridership is sensitive to spatial area. By using the second data set we estimate the behavior of the riders and show that their optimal decision based on optimizing their utility depends on the mode characteristics. We also derive the value of walking time and the value of auto in-vehicle time and we estimate aggregate elasticities for auto and bus. By using the third

data set we estimate an aggregate elasticity for the demand for subway trips and estimate the impact of public policies such as an increase in fare.

In addition, in this dissertation we attempt to develop an econometric approach that unifies time series and prediction from static models such as the logistic regression. By using one canonical model and varying the underlying assumptions and the distribution of the dependent variable we show that forecasting results can be obtained in ways not so different; only the optimizing algorithm is different. It appears that the two approaches underlined above are useful because they can be used to explain the dynamics of human behavior. This is one of the main contributions to the field of transportation economics.

To Virgil, my father, who is the first economist that I admired.

Table of Contents

List of Tables

Lists of Charts

I) INTRODUCTION

The problem this dissertation addresses is the construction and estimation of an urban travel demand model for the subway. This model is estimated using data for the New York City Metropolitan Area. The results of the model can be applied to other subway systems with very few changes in the underlying assumptions.

This research contributes to the development of transportation economics, and provides the theoretician and the transportation planner with the understanding of what variables affect the demand for subway travel and with a complete model. The model can be used for modeling the impact of changes in the socioeconomic variables and the transportation attributes on the subway travel.

First, a model estimated at the station level will provide a tool to evaluate and predict the impact of capital projects, i.e. opening of new subway stations or new lines. The main contribution is in predicting the sensitivity of the ridership with respect to changes in employment, population, and other economic variables.

Second, a discrete choice approach of the demand for ridership enables the transportation planner to estimate the impact of public policies such as an increase in subsidies. This can be achieved by calculating aggregate elasticities. In addition, modal split modeling will give us a precise idea of the behavioral patterns of transit riders in New York City. Another contribution of this approach is to make predictions about mode choice, i.e. public or private transportation,

given the characteristics of a new individual, Westin [1], Kennedy [2]; this is done by classifying the individuals based on the higher estimated probability.

Third, the estimation of the model using time series will help strategic planners to simulate the impact of macroeconomic variables such as employment, energy price index, on the demand for subway travel. This is also the best way to estimate an aggregate elasticity for the demand for subway travel and to forecast demand levels.

To do the three estimations discussed above we develop a classical demand model for the subway travel. This theoretical model will follow the specification developed by McFadden [3] for the behavioral approach. For the non-discrete approach we will use an approach similar to the one developed by Bollinger and Ihlanfeldt [4]. The theoretical model will be developed and estimated using three data sets, each one having its own limitations. Since both an aggregate demand model for ridership (demand at the station level or at the city level) and a disaggregate model (demand by the individual) are estimated we must use different data sets. The advantage of having this approach was underlined above. The first data set is a cross-section data set from the 1990 Census Population Survey and the M.T.A., and the unit of analysis is the station. The second data set is a cross-section data from the 1990 Subway Intercept Survey and the unit of analysis is the individual. The third data set is a monthly time-series data set from 1978 to 1996, which includes macroeconomic variables from the Metropolitan New York Area, quality service variables such as crime rates in the subway, and subway travel data.

The organization of the dissertation is the following: chapter I is the introduction; in chapter II, the general theory for the demand model for ridership is laid out; in chapter III, we describe each data set and the type of question each one addresses, the variables used, a presentation of descriptive statistics, the appropriate econometric model and estimation technique, and a discussion of the results; in chapter IV, we compare the results of the three data sets, and conclude what possible similarities can be drawn between the econometric approach and the time series approach in terms of predictive power.

II) Theory of Urban Travel Demand.

II.1) Review of the literature.

The measurement of urban travel demand forecasting has been for a long time the area of research of transportation engineers, (McFadden [5]). With the development of modern economic behavior theories and discrete choice theories, and the progress in computational technology, it started to be analyzed by economists.

Wohl and Martin [6] and Kanafani [7] have done the main developments and extensions of the transportation forecasting. Kain [34] has done work on the relationship between transit ridership and employment and on the measurement of transit ridership. Cervero et al [35] showed using the BART system that a significant relationship exists between population, employment and ridership. As we will show in part III.1 they also control for stations and non-station areas by using a radius of 1 to 2 ½ miles around the stations. We chose a radius of ½ mile around the subway stations. But the main innovation in the analysis of transportation behavior has been the development of disaggregate travel demand models, as in Ben-Akiva [8]. These models use discrete choice analysis methods. The models use microlevel data on the behavior of an individual, household or firm. For instance, the choice of travel between bus and subway for a given trip implies a binary independent variable. In addition, if the spectrum of choice the

consumer faces is broader, then the dependent variable can have a natural order, like number of buses ridden, or can be qualitative and have no natural order, like the choice between three alternatives of travel (i.e. car, subway, bus). A review of econometric applications of discrete choice models can be found in Maddala [9], McFadden [31], Manski [10], Amemiya [11].

My contribution to the topic is to derive a forecasting model for the demand for travel that can be used particularly to implement fare policies and to predict the impact changes in the system attributes or sample characteristics on subway travel, (Westin [1]). Following the early work by Becker [12], and the early transportation applications of discrete choice models developed by McFadden [5] we built a model for transportation demand, and give an estimation of the "hedonic" price of travel time. We calculate the elasticities within a given travel mode with respect to different variables. Hensher [36] has shown that fare elasticities and cross elasticities are an important tool for the implementation of fare policies. We will also calculate the aggregate elasticity for the demand for subway travel using time series. After estimating the model with time series we calibrate it to the New York City Metropolitan Area and quantify the effect of different potential public policy changes-variations in fare price, change in employment in the area of study, and change in the energy price index.

II.2) Theory and Model Construction.

II.2.1) Introduction

The main difficulty in the analysis of travel demand behavior is the existence of uncertainty in the choice of the utility function. Indeed, at the individual level, there are large and random variations in the amount of travel undertaken in a day. At the aggregate level, i.e. station or census tract, or borough level, this problem does not exist.

We will first develop the general consumer model and we will use the assumptions developed in Ben-Akiva [8] and in Manski [13] for the discrete choice model. We will follow McFadden [14] [15] to elaborate on the theory of travel time and we will use the conditional logit model to analyze the demand for travel, (McFadden [3]). We will estimate empirically its aggregate elasticity and cross-elasticity for different variables. Since cross section data is not the best tool to derive elasticity we use time series on the subway travel and on the economic indicators for the 1978-1996 period.

Travel is a derived demand. Travel as a service consumed does not supply the household or the rider with utility, but is a necessary complement to the performance of activities at different places and at different times such as work, household production or leisure. To understand this we will show first how the household faces the decisions of travel and how these decisions are integrated in the utility function.

II.2.2) Framework for Choice Behavior.

The choice made by a household can be analyzed as a decision set within different possible choices, and made under defined time and income constraints, (Ben-Akiva [8]). These decisions comprise but are not limited to the following steps:

a) definition of the choice set.

b) description of the alternatives.

c) evaluation of attributes of alternatives.

d) decision rule.

An example of a mode choice problem is shown in Table 1. A commuter going to work faces three different modes of travel: walk, auto, and subway. The commuter faces different costs and derives indirectly different utilities by using these modes.

Table 1

Alternatives	Direct Attributes		Indirect Attributes
	Travel Time	Travel Cost	Utility/Disutility
Walk	P_{t1}	P_{c1}	U_1
Auto	P_{t2}	P_{c2}	U_2
Subway	P_{t3}	P_{c3}	U_3

From Table 1 one can establish a choice between the three alternative based on the direct attributes and indirect attributes of the utility function. The model is called modal split, (Ben-Akiva [8]), and can be written as follows:

$$T_{ji\ auto}/\ T_{ji\ subway}=\psi_4(N_{ij}, N_{ij}, E_i) \qquad (0)$$

Here T_{ji} represents the number of trips allocated by a given rider to mode i or mode j. The number of trips allocated between auto and subway is a function of differences in travel times, monetary costs between modes, and economic variables. The main socioeconomic variables are the employment in the origin zone, population, and income. This model is behavioral and can be used for the implementation of public policies. By definition, these policy variables can only affect the split between the trip distribution. These models usually try to model the probabilistic response of making a transportation decision. After establishing the decisions that the rider faces, the rider

needs to have a utility function that he or she will maximize. This is developed in the following section.

II.3) The consumer theory relevant to transportation choice.

II.3.1) The behavioral approach.

We use the classical theory of consumer choice as a premise to derive a general model of transportation demand. We assume that the rider has a limited horizon of life. Within this time horizon the rider is making decisions regarding different activities that ultimately maximize utility[1]. The rider is making both consumption and production decisions under a defined set of resource constraints. Trips are considered as a joint commodity that involves time, money price, and different socioeconomic characteristics as input. The riders derive an indirect utility by "consuming" the commodity trip. Following Oi and Shuldiner [16], we will consider the trip as an intermediate good that is jointly demanded with other economic goods. The trip can be production oriented-home based work trips- or consumption oriented-leisure trips. Following Becker [12] and Oi and Shuldiner [16], the demand for the joint commodity –the leisure activity and its related trip- depends on four factors:

1) the consumer's preference patterns or tastes: T_i

2) the full income or wealth: $I_i = W_i + V_i$

[1] For computational reason this utility is assumed to be additively separable.

3) the hedonic price of the joint commodity-trip and leisure: H_i

4) the price of complimentary and competing goods: p_i

For statistical estimation, tastes will be assumed to be identical within the population. To estimate this variable, proxies can be derived from observed demographic variables such as age, race, sex, occupational choice, education and unobserved variables such as intelligence, experience, and childhood training. We want also to briefly look at the importance of full income on the demand for subway travel. If we exclude social visits, leisure-related trips can be treated as luxury goods. An increase in income will increase the demand for consumption-oriented travel[2].

We can write a general formalized model of demand for travel as follows:

$$\text{Demand for Travel} = g(I, p_s, p^*, f, TC, S, A, T, D) \qquad (1)$$

Where I is the full income, p_s is a vector of relative price of substitutes for travel mode, f is a vector of fare price for the mode of transportation chosen, p^* is equal to $f\text{-}p_s$, TC is a vector of time cost, S is a vector of socioeconomic variables, A is a vector of characteristics of the mode of transportation chosen, T is a proxy for taste, and D is a vector of dummy variables accounting for station typology and seasonal factors. The full income constraint is represented as follows:

---

[2] In reality the magnitude of these effects depends on the share of Income on transport related consumption. Usually this share is very low.

$$I = R*f+TC*w \quad (2)$$

Where R is the number of rides and w is money wage. Even if we hold the fare price (f) constant, the financial constraint would be different for each consumer since each consumer has a different valuation of the time cost of travel (w). Replacing (2) in (1) the model to estimate becomes:

$$\text{Demand for Travel} = h(R*f+TC*w, p_s, f, TC, S, R, T)+\varepsilon, \quad (3)$$

Since we use three different data sets to estimate the theoretical model in (3), the dependent variable, the explanatory variables, the h (·) function, and the probability distribution function of the residuals ε will differ. We will have a family of three empirical models derived from (3). In doing so we use the approach from Ramsey [16]. Ramsey defines a class of models, say C, and he assumes h (Y, X, θ) is a probability distribution function with Y being a vector of random variables, X a vector of non-stochastic variables, i.e. a vector of explanatory variables, and θ a vector of parameters. The specification of h (·), and the sets Y, X, θ are necessary. The function h (·) defined on the subset of the Cartesian product Y, X, θ defines the set *H*; this set is known in economics as a model. In our case *H*, is given by:

$$Y^* = \theta\Gamma' +\varepsilon \quad (4),$$

Where Y^* will take the forms shown in (5), (6), (7), Γ is the matrix [I, X, Z, P, D]', I is the identity matrix, θ is [δ, α, β, γ, 1], and h (·) is the link function. This h (·) function is

linear (chapter III.1), exponential (chapter III.2), and logarithm (chapter III.3). Using this approach we have three equivalent classes of models:

$Y^* \equiv \text{Ridership}_{\text{station } I}$, if the Y^* is a (Nx1) continuous vector of the variable ridership (5)

$h(\cdot)$ is a linear function f(x).

$Y^* \equiv \text{Log}\left(\frac{\pi}{1-\pi}\right)$, if the $Y^* = \{0,1\}$, i.e. is a (Nx1) vector of discrete choice for mode of travel. (6)

$h(\cdot)$ is $\frac{e(x)}{1+e(x)}$

$Y^* \equiv \text{Log (Aggregate Ridership}^t)$, if $Y_t^* = (Y_1^*, Y_2^*, ... Y_{t-1}^*, Y_t^*)$ is a (NxT) vector of time series for ridership (7)

$h(\cdot)$ is Log(x),

Where Y^* is a (Nx1) vector of the response variable, X is a (NxK) non-stochastic vector of mode characteristics, Z is a (NxK) non-stochastic vector of individual-specific variables, P is a (NxK) non-stochastic vector of relative prices, D is (Nx1) vector of dummy variables, i.e. [0,1]. We define π as the probability associated with the successful

event. We assume that ε is an unobserved (Nx1) vector of random terms identically and independently distributed as N $(0,\sigma^2)$. In order to address the different questions developed in the introduction we estimate (5), (6), and (7). Each model has its own limitations and advantages. The objective here is to address the questions with the appropriate estimation and not to choose which model is the best.

We will analyze now the effects of a change in fare on the number of trips, holding the demand for other commodities and their prices constant. Following equation (3), an increase in the price of the substitute transportation mode, say auto, p_s, decreases the demand for consumption oriented trips. Conversely, a decrease in time cost, TC, or money cost, f, will increase the demand for trips. But the sensitivity of demand to changes in the price of a trip is small if:

a) The trip costs comprise a small cost of the total cost of the joint activity.

b) There are no close substitutes for related leisure activities.

The production-oriented travel is defined by the work trip. The work trip can be a home-to-job trip or on-the-job work trips. Therefore we can assume that the volume of work trip must correspond closely to the level of employment in the destination area. Another common observation is that the price of the work trip itself is small compared to the income earned. Therefore the price of the trip has a small impact on the number of work trips demanded. In some cases we could even say that the elasticity of the demand for work trips with respect to fare prices is inelastic; we will calculate these elasticities in the next chapters.

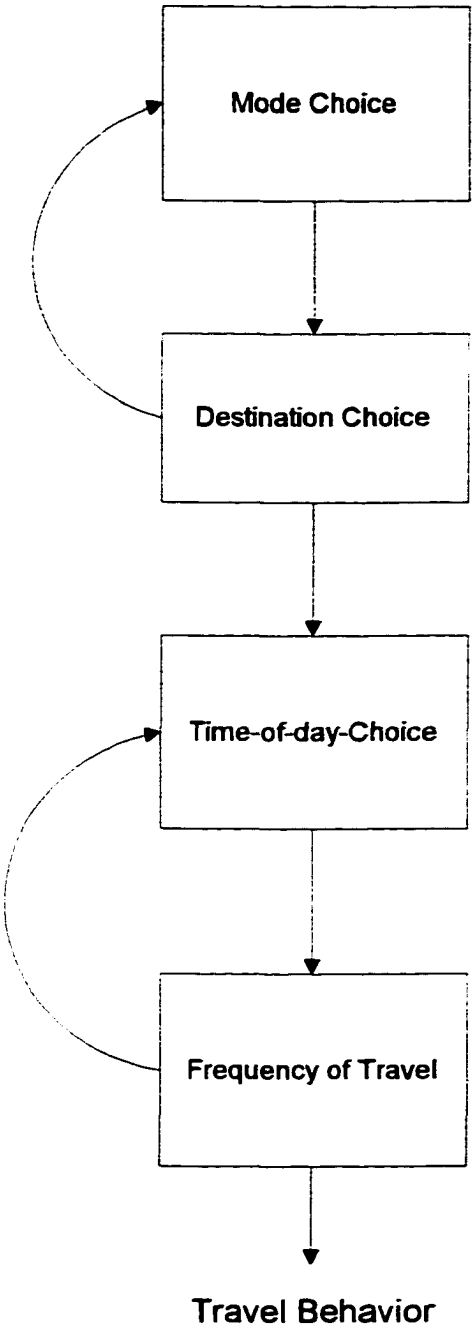
Now we will show how the classical consumer theory and the rider's behavioral choice theory are related. In the classical consumer theory, the consumer, can maximize its utility function under a set of time and budget constraints and an explicit demand function for commodities can be determined. More, and this constitutes the main axiom of the consumer theory, the consumer can maximize this utility by having marginal choices. This implies mathematically that the utility function is continuous and totally differentiable. Another difficult assumption to satisfy is the perfect knowledge of the prices of all other commodities.

These consumption activities have a vector of attributes describing the commodities purchased, work performed, trips taken, leisure time. The rider will choose an activity that will maximize his or her derived utility. The observed attributes will define his observed demand. Finally, this vector of attributes will define the transport demand behavior of the riders.

After showing how the transportation commodity enters the utility function of the rider we need to use this utility with a discrete choice framework. So, after defining a set of choices, showing how the transportation commodity enters the utility function, we will use the probabilistic approach developed by Manski [13] and the conditional logit analysis developed in McFadden [3] to construct the theoretical model. In order to simplify the estimation of the demand function for ridership one can factorize it into its decision components, McFadden [3]: mode choice, time of day travel, trip purpose, and frequency. For our purpose, we will restrict the estimation only to the mode choice for work related trips, given that the individuals leave in the same area. One can now derive an observed travel behavior, like the one shown in Figure 1:

Figure 1

Travel Decisions



Since most of the time the exact form of the utility function for each of the decision shown in Figure 1 is unknown to the rider, we will assume that the utilities are like in real life random, (Manski [13]).

So, in the random utility model the utility to a rider of an alternative is specified as a function of the characteristics of the rider and the attributes of the alternative and an error term. We can express this as follows:

$$P\{\text{choice I}|\ \Omega\}=P\{U\ (\text{choice I})>U(\text{choice K}),\ \ I\ ,K\in\Omega\} \qquad (8),$$

Here Ω is the total number of possible choices. This represents a very realistic approach of the utility theory and Manski [10] identifies four main sources of randomness:

a) Unobserved attributes,

b) Unobserved taste variations,

c) Measurement errors and imperfect information,

d) Instrumental variables.

The following explains the effect of each random source on the utility function.

Unobserved attributes- if the vector of independent variables that affects the decision is incomplete the utility function is a random variable:

$$U_{ti}=U\ (x_{ti}, T_n, {}^{r}x_{ti}) \qquad (a)$$

Where x_{ti} is the vector of values of the attributes of alternative i as perceived by the t^{th} individual, and T_n is the vector of characteristics of the decision-maker, and ${}^{r}x_{ti}$ is a random variable.

Unobserved tastes variations-the utility function is as follows:

$$U_{ti}=U(x_{ti}, T_n, {}^{r}T_n) \quad (b)$$

Where ${}^{r}T_n$ is a taste variable that varies among the decision-makers.

Measurement errors-the utility function is:

$$U_{ti}=U({}^{p}x_{ti}, T_n) \quad (c)$$

Where ${}^{p}x_{ti}$ is not the true attribute observed but is equal:

$${}^{p}x_{ti}=x_{ti}+\varepsilon_{ti} \quad (d)$$

By replacing the attribute vector in (f) with (g) the utility function becomes random:

$$U_{ti}=U(x_{ti}+\varepsilon_{ti}, T_n) \quad (e)$$

Instrumental variables-the utility function is expressed as follows:

$$U_{ti}=U\ (^{I}x_{ti}, T_n) \qquad (f)$$

Some variables of the $^{I}x_{ti}$ vector are not observed. Then an instrumental variable is used instead of $^{I}x_{ti}$ that relates the instruments and attributes:

$$^{I}x_{ti}=z_{ti}+ \varepsilon_{ti} \qquad (g)$$

Where z_{ti} is a vector containing the instrumental variables and the observed variables, x_{ti}.

By replacing in (f) the expression of $^{I}x_{ti}$ with (g) we obtain a non-stochastic expression of the utility function for a given choice:

$$U_{ti}=U\ (z_{ti}+ \varepsilon_{ti},\ T_n) \qquad (h)$$

Then, we can generalize by separating the stochastic utility function in two components: a non-stochastic component that reflects the "representative" tastes of the observed sample, $V\ (x_{ti}, T_n)$, and a stochastic component with mean independent of x_{ti} that reflects the effect of idiosyncrasies in tastes for individual t, $\varepsilon_{ti}(x_{ti}, T_n)$, (McFadden [5]).

The utility is rewritten as: $U_{ti}= V\ (x_{ti}, T_n)+\varepsilon(x_{ti}, T_n)= V_{ti}+\varepsilon_{ti}$ (9)

We can associate with each t^{th} individual making the choice i the level of indirect utility Y_n^* defined in (4). If the rider is faced with m choices then the observed variables are defined as follows:

$$Y_{ti}=1 \quad \text{if} \quad Y_n^* = \text{Max}\,(Y_{t1}^*, Y_{t2}^*, \ldots Y_{tm}^*) \tag{10}$$

$$Y_{ti}=0 \quad \text{otherwise.}$$

And it follows:

$$Y_n^* = V_{ti}+\varepsilon_{ti}= \beta_n' X_{ti}+ \alpha_i' Z_t+\varepsilon_{ik} \tag{11}$$

If the residuals ε_{ti} are IID with type I extreme-value distribution (Gumbel distribution) with the following cumulative distribution function (CDF):

$$F_{EV}(\varepsilon_i<\varepsilon)=\exp[-\exp(-\varepsilon)], \qquad -\infty<\varepsilon<+\infty \tag{12}$$

And with the following probability distribution function (PDF):

$$\int_{EV}(\varepsilon_i)=F_{EV}\exp(-\varepsilon_i), \qquad -\infty<\varepsilon_i<+\infty \tag{13}$$

Then we can show that the probability that any given rider *t* will choose an alternative *i* is given by the probability that the utility of that alternative to the rider is greater than the utility to that rider of all available alternatives:

$$P_{ti}\{\text{choice I}|\ \Omega\}=P\{\ V_{ti}+\varepsilon_{ti} > V_{tk}+\varepsilon_{tk}\ \},\ k, i \in \Omega, t=1\ldots N\} \quad \text{for all } i\neq k \qquad (14)$$

$$P_{ti}\{\text{choice I}|\ \Omega\}=P\{\varepsilon_{tk} < V_{ti}-V_{tk}+\varepsilon_{ti}\ \},\ k, i \in \Omega, t=1\ldots N\ \} \quad \text{for all } i\neq k \qquad (15)$$

Since we know that the residuals are IID with the CDF shown in (12), then

$$P_{ti}=\text{Prob}(Y_{ti}=1)=\text{Prob}(\varepsilon_{tk} < V_{ti}-V_{tk}+\varepsilon_{ti}\ ) \quad \text{for all } i\neq k \qquad (16)$$

$$=\int_{-\infty}^{+\infty}\prod_{i\neq k} F_{EV}(\varepsilon_{ti}+V_{ti}-V_{tk})\cdot f(\varepsilon_{ti})d\varepsilon_{ti}$$

With

$$\prod_{i\neq k} F_{EV}(\varepsilon_{ti}+V_{ti}-V_{tk})\cdot f(\varepsilon_{ti})d\varepsilon_{ti}=\prod_{i\neq k}\exp(-e^{-\varepsilon_{ti}-V_{ti}+V_{tk}})exp(-\varepsilon_{ti}-e^{-\varepsilon_{ti}})$$

$$=\exp\left[\varepsilon_{t}-e^{-\varepsilon_{t}}\left(1+\sum_{i\neq k}\frac{e^{V_{ti}}}{e^{V_{t}}}\right)\right] \qquad (17)$$

If we write

$$\Delta_{ti}=\log\left(1+\sum_{i\neq k}\frac{e^{V_{ti}}}{e^{V_{ti}}}\right)=\log\left(\sum_{k=1}^{m}\frac{e^{V_{tk}}}{e^{V_{ti}}}\right) \qquad (18)$$

then (16) can be transformed as

$$\int_{-\infty}^{+\infty}\exp\left(-\varepsilon_{ti}-e^{-(\varepsilon_{ti}-\Delta_{ti})}\right)d\varepsilon_{ti}=\exp\left(-\Delta_{ti}\right)\int_{-\infty}^{+\infty}\exp\left(-(\varepsilon_{ti}-\Delta_{ti})-e^{-(\varepsilon_{ti}-\Delta_{ti})}\right)d(\varepsilon_{ti}-\Delta_{ti})$$

$$=\exp(-\Delta_{ti})=\frac{e^{V_{ti}}}{\sum_{k=1}^{m}e^{V_{tk}}} \qquad (19)$$

The rider chooses the alternative that maximizes his or her utility. We can simplify (4) by rewriting it as follows for a binary logit response model:

$$P\{\text{choice } I=1 | \Omega\}=h(\theta\Gamma') \qquad (20)$$

Where I is (Nx1) vector of choices that may be different for each rider, Ω is the entire universe of choices, Γ is a (NxK) vector of known constants, θ is a (NxK) vector of unknown parameters, and h $(\cdot)$ the function defined in (5), (6), (7). The estimation of the model depends on the assumptions made about the distribution of the disturbances, ε. In general three distributions are assumed about the error term ε for each observation: ε_k-ε_i distributed uniform defines a linear probability model, ε_k and ε_i distributed normally defines a binary probit model, ε_k-ε_i distributed logistic defines a binary logit model. In the case of polychotomous dependent variables, if the random utility error terms are assumed to be independently and identically distributed with a log Weibull probability distribution function (PDF), then we have a multinomial logit model. The drawback of this model is that it is characterized by the independence of irrelevant alternative assumption, usually denoted IIA. This property implies that the odds of choosing option k over option i are not affected by what other options are available. If the random utility error terms are assumed to be distributed multivariate normally, then we have the multinomial probit model. The model allows the error terms to be correlated across alternatives, in so relaxing the independence of irrelevant alternatives assumption, (Kennedy [2]) . We will assume that these disturbances are independently and identically distributed (IID) with

type I extreme-value distribution. In the binary choice case we estimate equation (6) with a common family of joint probability distribution functions. The choice of F is not critical as long as it is a probability distribution function. The h (·) transformation is also referred to as the link function by some authors, (Ben-Akiva [8]). The well-known forms of the h function are:

Linear Probability Model: $h_a(X)=X$ (21)

Probit Model: $h_b(X)=\Phi(X)=\int_{-\infty}^{X}\frac{1}{\sqrt{2\pi}}*e^{\frac{t^2}{2}}dt$ (22)

Logit Model: $h_c(X)=\Lambda(X)=\frac{e^{(x)}}{1+e^{(x)}}$ (23)

The linear probability model presents one drawback: h (·) is not a proper PDF since it is not constrained to lie between 0 and 1. The probit model uses the normal distribution and can therefore be justified by appealing to the central limit theorem but it becomes computationally expensive with more than four alternatives. The logit model main justification is that the logistic distribution is similar to a normal distribution and has a simpler form. To empirically estimate either one of these models a log likelihood function is constructed and maximized. The likelihood function is the product of the probability of "success" and the probability of "failure". Efficient and consistent estimators are derived only in the case of the probit or logit models.

We will use the specification described in (23) and the model described in McFadden [3]. He starts by analyzing the choice of mode for shopping trips. He accounts for both the attributes of each alternative, Z, transit walk time, transit wait plus transfer time, auto in vehicle time, and individual characteristics, X, occupation, race, income, geographic location. He calls this model the "conditional logit model". The mode selection probability satisfies:

$$P_{ti}=\text{Prob}(Y^{*}_{ti}=1)=\frac{e^{V_{it}}}{\sum_{k=1}^{m} e^{V_{it}}} \tag{24}$$

$$P_{ti}=\text{Prob}(Y^{*}_{ti}=1)=\frac{\exp(\beta' X_{ti}+\alpha_i' Z_t)}{\sum_{k=1}^{m}\exp(\beta'' X_{tk}+\alpha_k' Z_t)} \tag{25}$$

In this case the utility of alternative i to individual t is a function of the m attributes of each alternative, i.e. Z_t, as seen by the rider, and the individual characteristics of each rider, i.e. X_{ti}. Here, we must estimate m coefficients identical for all individuals. If we want to estimate inherent differences between the alternatives that are the same for all individuals, dummy variables for all but one alternative will be included, Kennedy (1998). From equation (25) we can see that there is no intercept term in the conditional logit model. If we include an intercept in the terms of the numerator and denominator then each individual term can be rewritten as $\exp(\beta' X_{it}+\alpha_i' Z_t+\gamma_0)$. This can also be written, using the properties of the exponential function, as $\exp(\beta' X_{it}+\alpha_i' Z_t)\cdot\exp(\gamma_0)$. Because $\exp(\gamma_0)$ appears in every term, this term cancels

out of the fraction. Equation (25) implies that the logit for comparing two travel options i and k is given by

$$\log\left(\frac{\Pr ob(Y_{ti}^{*}=1)}{\Pr ob(Y_{tk}^{*}=1)}\right)=\beta(x_{ti}-x_{tk}) \tag{26}$$

By taking the exponential of equation (26) we obtain the odds that person t will choose option i over k as

$$\exp\{\beta(x_{ti}-x_{tk})\} \tag{27}$$

We estimate this model by maximizing the log-likelihood function which is the product of n factors, each equal to equation (25) for the option chosen.

First, we define the (k×m) matrix X_{tm}=(X_{t1}, X_{t2} ...X_{tm}). We also define the following vectors where P_{ti} and Y_{ti} are defined in (25):

$$Y_{tm}=\begin{bmatrix}Y_{t1}\\Y_{t2}\\\\Y_{tm}\end{bmatrix} \qquad P_{tm}=\begin{bmatrix}P_{t1}\\P_{t2}\\\\P_{tm}\end{bmatrix}$$

Then the log-likelihood function becomes

$$\log L = \sum_t \sum_i Y_{it} \log P_{it} = \sum_t \sum_i Y_{it} (\beta' X_{it}) - \sum_t \log(\sum_k e^{\beta' X_{it}}) \tag{28}$$

We assume that at any point in time the individual will choose only one alternative, i.e., $\sum Y_{ti}=1$. We can write the first order condition as

$$\frac{\partial \log L}{\partial \beta} = \sum_t X_t (Y_t - P_t) = 0 \tag{29}$$

And the second-order condition as

$$-E\left(\frac{\partial^2 \log L}{\partial \beta \partial \beta'}\right) = E\left(\frac{\partial \log L}{\partial \beta} \bullet \frac{\partial \log L}{\partial \beta'}\right) = \sum_t X_t A_t X_t' \tag{30}$$

With $A_t=E(Y_t-P_t)(Y_t-P_t)'=D(P_t)-P_tP'_t$ (31)

And $D(P_t)$ is a diagonal matrix with the j^{th} diagonal element=P_{ti}. The covariance matrix of β_{ML} is $(\sum_t X_t A_t X_t')^{-1}$. The ML estimates are obtained by using an iterative technique such as Newton-Raphson, or the method of scoring. For the last one the iterative function is:

$$\beta_{p+1}=\beta_p+[I(\beta_p)]^{-1}S(\beta_p) \tag{32}$$

where $I(\beta_p)$ is the information matrix defined in (30) and $S(\beta_p)$ is defined in (29). Both these two expressions are evaluated at $\beta=\beta_p$ and β_p is the value of β at the p^{th} iteration. We estimate the coefficients α_i in (25) by creating a set of dummy variables. We define $\alpha'=(\alpha'_1, \alpha'_2, \ldots \alpha_{m-1})$ and we normalize $\alpha_m=0$.

We also define Z_{ti} as follows:

$$Z_{tj}=\begin{bmatrix}0\\0\\.\\Z_{tj}\end{bmatrix} \tag{33}$$

Then if we define

$$\theta = \begin{bmatrix}\beta\\\alpha\end{bmatrix}, W_{tj} = \begin{bmatrix}X_{tj}\\Z_{tj}\end{bmatrix} \tag{34}$$

We can estimate the model as shown above with θ replacing β and W_{ti} replacing X_{ti}. We will estimate the model described in McFadden [3].

There are two modes of transportation: transit or car. We will consider in our case transit is bus. The full price of a trip by bus, f_b, is

$$f_b=w_1t_b + w_2t_b + c_b \tag{35}$$

where t_w is walking time from home to the bus station, w_1 is the value of walking time, t_b is the time spent in the bus, w_2 is the value of in-vehicle time, and c_b is the bus fare. The full price of a trip by car is

$$f_a = w_2 t_a + c_a \tag{36}$$

where t_a is time spent in the car and c_a is auto operating cost. McFadden assumes that w_1 and w_2 are the same for each person but differ from each other. Then he specifies the log odds for comparing the auto and bus modes as:

$$\ln\left[\frac{\pi}{1-\pi}\right] = \alpha(f_a - f_s) + \beta X \tag{37}$$

where X is a vector of socioeconomic variables. By replacing (35) and (36) in (37) we obtain

$$\ln\left[\frac{\pi}{1-\pi}\right] = \alpha w_1 t_w + \alpha w_2 (t_a - t_s) + \alpha (c_a - c_s) + \beta X \tag{38}$$

The relevant regressors are the difference between t_a and t_s, the corresponding difference between c_a and c_s, and the characteristics of the individual. We can rewrite equation (38) as

$$\ln\left[\frac{\pi}{1-\pi}\right]=\delta t_w+\gamma(t_a - t_s)+ \alpha(c_a-c_s)+ \beta X \qquad (39)$$

Where $\delta=\alpha\ w_1$ and $\gamma=\alpha\ w_2$. McFadden uses the estimates of α, δ, and γ to compute w_1 and w_2.

Next, McFadden examines shopping choice of destination for 63 auto mode trips. The number of alternative trips varies between three and five. He uses the inclusive price of a trip from the estimate of equation (39) as a regressor. For a trip by a given person to a given destination, that price (z) is

$$z=\alpha c_a+\gamma t_a=\alpha(\ c_a+w_2 t_a) \qquad (40)$$

And c_a and t_a differ by trip and by person. Note that $c_a+w_2 t_a$ is the full price of a trip. Multiplication by α converts the money price into a utility price but is not really necessary since α is a constant. One issue that McFadden does not consider is that the value of time may differ among people. If so, let y be income. Suppose that $w_1=ky$ and $w_2=ry$, where k and r are constants. Then equation (39) becomes:

$$\ln\left[\frac{\pi}{1-\pi}\right]=\alpha k y t_w +\alpha r y(t_a- t_b)+ \alpha(c_a-c_b)+ \beta X \qquad (41)$$

This is a logit in which y is interacted with t_w and t_a- t_b. Similarly equation (41) becomes

$$\ln\left[\frac{\pi}{1-\pi}\right]=\delta y t_w+\gamma y(t_a-t_b)+\alpha(c_a-c_b)+\beta X \qquad (42)$$

where $\delta=\alpha k$ and $\gamma=\alpha r$.

In order to address the complete set of questions underlined in introduction, we will estimate the theoretical model (4) by using three different data sets: (1) a cross-section data set where the unit of analysis is the station, (2) a cross-section data set where the unit of analysis is the individual, and (3) a time series data set.

The next three parts will each give a description of the data set, an explanation of the type of question addressed, the particular variables used and how they relate to the theoretical model (4), a presentation of the descriptive statistics for the data set, a description of the econometric model and estimation technique to use with the data set, and a discussion of the results.

The final chapter will compare the results obtained with the three data sets, and draw overall conclusions.

III) Empirical Results.

III.1) The station level results.

III.1.1) Introduction.

How sensitive is the subway travel to variations in economic factors? This issue has attracted the attention of managers and economists at the Metropolitan Transportation Authority since the recent changes in fare. This chapter is concerned with the relationship between subway travel and economic factors like employment, income, level of education and the size of the area around the station. It develops a framework for identifying and quantifying which economic, demographic, and service variables affect weekday subway registrations at the station level. This link is important in identifying the impact of changes in population and employment on the ridership at the station level. This micro quantification of the variables should provide us with an estimation of the subway travel at the station level. We note that a model of census tract and employment has been developed by C.R.Bollinger and K.R.Ihlanfeldt [4] to study the economic impacts of Atlanta's M.A.R.T.A rail transit system on employment and population at the station level.

There are three distinct parts in this empirical estimation. In the first part, a comprehensive research and analysis of the data available is undertaken. In the second part, we design manually areas around each station using a radius of around 0.5 miles.

Each area is created manually in order to avoid overlapping of areas around stations less than 1 mile apart. This intricate layer is created in Maptitude (a version of the TRANSCAD software) and is used to retrieve all information for all defined areas (one area per station). In order to account for the variation in the shape and size of the area around the station we control for the area in the model. In the third part, we estimate the model. In order to account for the different type of stations we decide to use a residential/non-residential typology.

The methodology proposed in this part is applied to data obtained from the following sources: Intercept Survey on Bus Riders conducted in 1990, the 1990 Weekday Subway Travel from the Revenue Division, the 1990 Fare Evaders Study and the available jobs tabulated from the C.P.1990. The results obtained will allow forecasting ridership at the station level. A simulation is done for the year 1995 for the entire system and separately station by station. For the entire system the forecasting error is 7.5% in absolute value. For the sample of 24 stations chosen the error varies. Next, we will describe the New York City subway and its history.

III.1.2) Description of the data set.

In this part we describe the brief history of the New York City Subway, the organizational structure for the empirical study and the process by which the research is designed. It also shows how we designed the areas of observation and the nature of the data used.

The first New York City subway began operating on October 27, 1904 in Manhattan. A private company, the Interborough Rapid Transit Company (IRT) operated the original 9.1 miles of subway track. This line had 28 stations from City Hall to 145th and Broadway. The IRT began service to the Bronx in 1905 to Brooklyn in 1908 and to Queens in 1915. Another private company, The Brooklyn Rapid Transit Company (later the Brooklyn-Manhattan Transit Corporation), began service between Brooklyn and Manhattan in 1915. In 1932, the City of New York finished the construction of the Eight Avenue line, creating the Independent Rapid Transit Railroad. This line was owned and operated by the City. By 1940, the City purchased the IRT and BMT. The City would solely own and operate all subway service in NYC until 1953, when, the New York State legislature created the New York City Transit Authority (now MTA) as a public corporation to operate City's owned subway lines. In March 1968 the New York State legislature created the Metropolitan Transportation Authority. Today the subway system is over 230 miles long, has 406 stations and the daily ridership is 3.5 million; this was the entire population of New York City in 1904. The New York subway system is one of the most complex in the world and the most important mass transportation system in United States.

The research starts with the following question: What variables have a demonstrated impact on weekday ridership at the station level? To address this question we make assumptions about how the data is collected. The assumptions on one hand, are made in order to eliminate double counting, and on the other hand are imposed by the technical limitations of the different computer packages. We use mainly Maptitude, which is a Geographical Information System, and SAS, which is a statistical package.

The independent variables are constructed by using a digitized map. Each observation is limited in space by a defined zone around every station. In general, a zone is a circle with a 0.5-mile radius, drawn around any subway station. This captures a proportion of all variables with respect to the area drawn. The zones are manually drawn to avoid overlapping. This is an important part of the data collection, because it prevents two different areas from using the same information. An area is composed of a number of tracts-these are the smallest units of observation made up of the aggregation of observations at the zip code level. These areas are designed so that a person will be assigned to station A in zone A if the person is within the radius (R_A). The person will be assigned to station B in zone B if it is inside radius (R_A) and radius (R_B) but closer to station B. Because the zones are tailored to the geography of each station, the area of the zones is not fixed (the maximum area of a circle of 0.5 miles radius is $0.5^2 x\pi=0.78$ squared miles). It varies due to the fact areas are not circular for all zones; the average radius is 0.33 miles. A third mile radius is chosen for two reasons. First, walking distance is commonly defined as being within a quarter mile of a station as in C.R.Bollinger and K.R.Ihlanfeldt [4]. Second, a third-mile radius results in a minimum of ring overlap for the stations located in Manhattan. These stations average 0.5 miles in contrast to 0.7 miles separating stations in the other boroughs. The layer is designed around all the stations, which represent a sample of 308 areas. We exclude the Manhattan CBD area which corresponds to the area below the 59th street. In the first phase we do not create a layer for this area because we assume that ridership in that case is driven mainly by jobs. Also it seems that the causality would be from ridership to jobs and not from jobs to ridership. This problem of reverse causality would imply biased coefficients.

After we design the layer around the stations, we link this layer with the data tabulated from the Census of Population (C.P.1990) by census tract, with the data from the Intercept Survey (I.S.), with the data on jobs, and with the ridership data by subway station. The I.S. provides us with the number of people entering the subway station from all local buses. The job data contains the number of jobs in the areas surrounding each station. The ridership data is tabulated from turnstile counts. The smallest unit of aggregation is the census tract. A tract corresponds to almost a block in Manhattan, but for the other boroughs it may include more than one block. We choose this unit of aggregation in order to maximize the number of observations. Another reason to choose a small unit (tract) is the fact that data is retrieved proportionately by area of tract inside the area around the stations. The small unit allows for a better allocation of data to each area around the station. In the next section we will show what variables we use.

III.1.3) Description of variables.

In this part we explain how the variables are constructed and how they relate to the theoretical model. We describe in Table 2 the variables used to estimate the ridership at the station level, their expected impact on ridership, and the advantages and limitations of these variables.

The median household income is used as it appears in the C.P.1990. In general, ridership tends to be more sensitive to economic changes in lower income areas than in middle and upper income areas. The middle and upper income areas appear to be more inelastic to changes in economic conditions in the City. The estimation of the demand

equation for ridership will determine if the fare increase has a stronger impact on registrations in lower income neighborhoods. The per capita income is also tabulated in the C.P.1990, but corresponds to only a census tract. While retrieving this data we weigh the average household income variable by the number of households living in the area and the per capita variable by the population living in the area. For instance, here is what would happen if a given zone included two tracts exactly:

Table 3

Calculation of weighted Income.

	Tract 1	Tract 2	Zone of 0.5 miles
Area	0.5 Sq. Mi.	0.2 Sq. Mi.	0.7 Sq. Mi.
Households	2,000	3,000	5,000
Household Income	$35,000	$40,000	75,000
Weighted Income	$14,000	$24,000	$38,000

Source: C.P.1990.

In addition, we calculate two variables: the first one includes all people with high school diploma and the second one includes all people with more than high school education.

Because we could not have reliable counts for immigration, we used a proxy for this variable: we defined as immigrants the persons that declared, at the C.P.1990, that their first language spoken at home was not English. These have the shortcoming of not counting people from India, the Commonwealth Community, some Caribbean countries

and Great-Britain. Immigration has been shown to be an important factor in explaining ridership trends, (MTA Reports [32]). Within this category, and holding the income effect fixed, ridership patterns depend on the ethnicity of the immigrants. For instance, Dominican (and other Hispanic groups) and Chinese (and other Asian groups excluding Japanese) immigrants use the subway more than Jamaican and Italian immigrants. Employment is constructed using the number of civilians in the labor force (we exclude people being in the labor force in the armed forces). The variable "evaders" is constructed using clerk counts and is correlated negatively with the ridership. The number of jobs in the area comes also from the C.P.1990. This figure was calculated by the department of City Planning which estimated the number of employers in the area: proxies were used such as hospitals, firms, and other big employers. Riders entering the subway station from the bus were extracted from the I.S.

In addition, the socioeconomic variables are interacted with a set of three dummy variables representing station type. A station typology is developed by careful visual inspection. The three types of stations are:

Express stations (TYPE 1)

Commuter stations (TYPE 2)

High-intensity stations (TYPE 3)

The TYPE 1 stations are stations where express trains stop, the TYPE 2 stations are feeder stations, and the TYPE 3 stations are high traffic stations. In reality the subway travel varies across stations types. Interacting the station type dummy variables with the

other variables allow the ridership to vary across station types. The population variable is included in the empirical model. Population is in principle correlated with employment, which is a subset of the population in the area. An index will be constructed as the ratio between the number of jobs in the station area and the population to account for the employment gravity. Some stations are "job attractors" and therefore the ridership is overestimated. This correlation will certainly create biased estimations for these variables. The correlation is shown in the Table 4:

Table 4

Changes in Ridership, Population, and Trips per Resident by Borough

Borough	**Subway Trips**	**Population**	**Trips/Resident**
Year	1970/1990	1970/1990	1970/1990
Manhattan	-12.4%	-1.5%	-11%
Bronx	-37.4%	-21.2%	-20.6%
Queens	-5.8%	-3.1%	-2.8%
Brooklyn	-20.8%	-11%	-11%

Source: Metropolitan Transportation Authority Reports 1970-1990.

The conclusion derived from Table 4 is that the use of the subway, as measured by the annual number of trips per resident of each Borough, has decreased within a period of 20 years. Why? Before answering this question by estimating the coefficients of the

explanatory variables, there is evidence that some citywide factors usually affect ridership:

a) Automobile registration. The number of cars owned by City residents declined in the 1970 and increased in the 1980. This latter rise in auto may have depressed subway travel and especially discretionary travel.

b) Employment. The number of non-farm jobs declined steadily in the City. These shifts in employment rates have been procyclical with ridership, which declined in the 1970 and increased in the 1980. Since 1990, the correlation between employment and ridership does not appear to hold although there are evidence showing that internal factors such as increased police presence at turnstiles, improved service and better quality, may be responsible for the change in the correlation between employment and ridership.

c) The real fare. Nominal fares rose steadily from 30 cents in the 1970 to the 1.50 fare hike in 1995.But the inflation rate increased on the average at a higher rate and as a consequence real fare declined.

d) Non-City Resident Ridership. It is likely that non-residents (people that live in suburbs, tourists) contributed positively to the use of the subway. This is more significant in the CBDs[3].

[3] Community Business District-Wall Street.

It shows that other variables have to be explored completely in order to estimate the ridership, i.e. fare, auto ownership, price of substitute goods. These variables will be empirically estimated in the next chapters using different data sets.

The main objective of this section is to determine which variables explain more accurately average weekday ridership at the turnstile, i.e. at the station level. Although using small groups of stations could provide more neighborhood-specific trends the ridership data is disaggregated into stations. The estimation using time series for the Districts or Boroughs will be done in the last chapter. Also some districts will be omitted in the first stage. The four districts in Lower and Midtown Manhattan are excluded from this analysis because ridership in the CBD is driven more by employment than by residential characteristics. One way to account for this, is to use a proxy for proximity to a high employment area, i.e. distance to the CBD. Since the CBD produces discretionary subway travel we decided to exclude this area but kept the number of jobs as a regressor.

III.1.4) Econometric model and estimation techniques.

In this section we estimate the ridership using the variables previously selected, and we identify which of them are statistically significant in explaining subway travel. To determine if these variables are sensitive to the indexation of the stations, we estimate 18 equations and we test the models to see if they are different using an F-test. For instance, we consider a model for the job-related stations and a model for the residential stations. We show how the cutoff point is chosen in order to classify the stations as residential or non-residential. Finally, after correcting these models for heteroscedasticity, we simulate

the forecasts on a sample of stations and on the entire subway system. We estimate the model using the cross-section data set described in 1.3. This model is estimated using weighted least squares. The econometric model is:

$$\text{Ridership}^{4}_{\text{station}I}{}^{\text{index}J} = a_{0\text{station}I}{}^{\text{index}J} + a_1\text{LOCALBUS}_{\text{station}I}{}^{\text{index}J} + a_2\text{JOBS}_{\text{station}I}{}^{\text{index}J} + a_3\text{EMPLOYMENT}_{\text{station}I}{}^{\text{index}J} + a_4\text{POPULATION}_{\text{station}I}{}^{\text{index}J} + a_5\text{IMIGRATION}_{\text{station}I}{}^{\text{index}J} + a_6\text{EDUCATIONLHS}_{\text{station}I}{}^{\text{index}J} + a_7\text{EDUCMHS}_{\text{station}I}{}^{\text{index}J} + a_8\text{TYPE1}_{\text{station}I}{}^{\text{index}J} + a_9\text{TYPE2}_{\text{station}I}{}^{\text{index}J} + a_{10}\text{TYPE3}_{\text{station}I}{}^{\text{index}J} + a_{11}\text{AREA}_{\text{station}I}{}^{\text{index}J} + a_{12}\text{INCOME}_{\text{station}I}{}^{\text{index}J} + e \quad (43)$$

The TYPE1, TYPE2, and TYPE 3 variables are dummy variables used in order to take into account the stations which are defined as "express stations", "commuter stations", and "high-intensity stations", where the ridership will be greater than the population leaving in that area[5]. These variables take the value 1 for the stations defined above and 0 otherwise. The exogenous variables will be observed proportionately to the area defined around the station. However, for the income variable, we will use the average values. Before presenting the regression results, we calculate the mean values of these variables for the entire subway system, and for job-related stations and residential-related stations in Table 5.

We use the index to differentiate between "job" stations and "residential" stations in the same way Bollinger and Ihlanfeldt [4] use different equations for employment and population areas. For all the stations, the mean number of riders entering from bus is 373,

[4] This excludes all passes-high-school student, disabled persons, and all weekends and night ridership.

[5] As a rule the dependent variable ridership should equal to the population of the area.

the mean number of jobs is 3,306, the mean number of people employed is 4,412, the mean population is 3,700, and the mean immigration is 4,565, the highest mean compared to the other variables. On average more than fifty percent of people riding the subway have less than a high-school education level as compared to a high-school diploma or more. The median household income is $25,111. The mean area around all stations is a third of a mile.

The mean number of employed people in the area neighboring the station, the mean population, and the mean immigration, for job and residential stations are similar to the mean for the entire system. The mean of riders entering from bus and the mean number of jobs are lower overall for job stations as compared to residential stations and all stations.

The gap between subway riders with less than a high-school diploma and riders with more than a high-school diploma is maintained for job stations and residential stations. The mean household income is the lowest for stations where the index of jobs over population is greater than 60 percent (it corresponds to the job 60 classification). For the residential stations the mean household income is almost the same as the one for the entire system.

The above comparison of means between job stations, residential stations, and all stations suggests that there are differences for the number of jobs and riders entering the subway from bus. We will run the pooled regressions and test for the significance of the models using an F-test. In addition, the comparison between the independent variables shows that immigration has the highest mean overall and has an impact on the ridership.

We will now run the regression models with ridership as the dependent variable, and determine which model we will use to forecast the ridership.

The model for ridership is estimated by applying weighted least squares to equation (37) to correct for heteroscedasticity using the statistical package SAS[6]. Mainly all the variables have the expected sign. We have three different sets of coefficients: one for the entire subway system, one for job stations and the other for residential stations.

We determine that in most cases there is heteroscedasticity by using the Bartlett's M Specification Error Test (BAMSET) described in Ramsey [17]. Heteroscedasticity is present when all the specifications made for the model in (43) are correct except that $\sigma^2\Omega$ is assumed to be a diagonal matrix with unequal elements on the diagonal. This is a Type II error that leads to a change in the covariance matrix of the estimators. The Bartlett M test is designed to test the null hypothesis, H_0, against the alternative hypothesis, H_1. We define these hypotheses as in Ramsey, H_0 being expressed in terms of the BLUE residuals as

$$H_0: \ u \sim N(\varnothing, \sigma^2 I_{N-K}). \tag{44}$$

The alternative H_1 is also expressed in terms of the BLUE residuals as

$$H_1: \ u \sim N(\varnothing, \Theta), \text{ where } \Theta \text{ is a diagonal positive-definite matrix.} \tag{45}$$

[6] The procedure used is PROC REG.

The BAMSET test involves the calculation[7] of the M statistics:

$$M = -2\ln\{\sum_{i=1}^{k} \frac{[(s_i^2)v_i/2]}{s^2}\} \qquad (46)$$

Where k is the number of subgroups of squared residuals, $s^2_i=\frac{1}{v_i}\sum v_i u^2_j$, each v_i is an integer equal to (N-k)/k, $\sum_{i=1}^{k} v_i = v = (N-k)$, and $s^2=(1/v)\sum_{j=1}^{nk} u^2_i$. Ramsey recommends setting k equal to 3. Under the null hypothesis the M statistic in (46) is asymptotically distributed as central chi-square with (k-1) degrees of freedom. Since we choose k=3 then M=2.71 at a 90% confidence interval. We show in Table 6 that for all model specifications except two, job 90 and residential 10, we reject the null hypothesis of homoscedasticity. Another way to dampen the effect of heteroscedasticity is to use an isomorphism of the initial model, i.e. the log transform. The results were not so satisfying statistically so we kept the linear model.

The results are shown in Table 7a and Table 7b; Table 7a shows the regression results with all the variables of the model, and Table 7b shows only the statistically significant coefficients of the model. For estimation purpose we use the models estimated in Table 7b. If we compare the coefficients for the model including "all stations" some of the variables change in magnitude and sign after elimination of the not significant coefficients. The coefficient for employed people is 60 percent lower, the coefficient for

[7] This calculation is done with a SAS user written program. The alternative is to use the MEANS statement in the ANOVA or GLM procedures with HOVTEST=BARTLETT.

people with more than a high school diploma is 21 percent higher, the coefficient of type 2 station is 4 percent lower, the coefficient of type 3 station is 1.75 percent higher, the coefficient of area is 17 percent higher, and the coefficient for the median household is 50 percent higher. The sign for those variables is unchanged.

Since the population and the employment is not distributed uniformly in the four boroughs we decide to split the total sample of observations into two different zones: residential zones and non-residential, i.e. job areas. This is achieved by constructing an index variable. The index is equal to the number of jobs in the area divided by the population living in the area. We construct different data sets of observations for different cutoff points and we test for the sensitivity of the variables to the index constructed. This index can be considered as a proxy for the job situation in the area (e.g., number of jobs created). For instance, JOB 40 means that the per capita number of jobs is bigger than 40 percent for all stations in the sample. If the index is less or equal to 40 percent, the areas meeting this condition is regarded as residential. Therefore, if we add these two numbers, we should obtain the total number of observations for the entire data set. We use 10 percent increments in order to have a greater choice of models. These regressions are shown in Table 7a and Table 7b.

The errors are normally distributed which satisfies one of the ordinary least square regression's assumption. The explanatory power, (R^2), of the models varies between 0.58 and 0.76 for residential observations and between 0.7 and 0.93 for job observations, Table 7a.

To further assess the impact of the total or the mix of independent variables on ridership and to test whether the stratification job/residential station is statistically

significant we conduct joint and sum significance tests for each index value. For instance, for JOB 10 the joint ridership test is:

$$H_0: \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \\ \\ a_{12} \end{bmatrix} = 0, \tag{47}$$

And the ridership sum test is:

$$H_0: [a_1 + a_2 + \ldots + a_{12}] = 0, \tag{48}$$

The joint significance test is a χ^2 test distributed with 12 degrees of freedom. The sum significance test is a standard normal test. The results are reported in Table 8. For all job and residential indexes, the null hypothesis is rejected for the joint test at the 5 percent level except for Job 40, Job 60, Job 70, Job 80, Job 90. For the sum test the null hypothesis can not be rejected for Job 40, Job 70, Job 80, and Job 90, and Residential 10 at the 10 percent significance level. These results confirm that the variables impacting ridership depend on the index level.

We choose the model that has a significant number of explanatory variables and has a high R^2. We determine that the best model is the one with the cutoff point of .20. Therefore we use this model to estimate the ridership for the entire system. Other models based on different cutoff variables have been tested: choosing zones with same brackets

of income, same area and the same number of people employed. The results are not significant since these indexes reduce even more the size of the sample; it varies between 20 and 30 observations. Other suggestions to test would be to use information coming from zones with constant area. This means that the bands drawn around the stations should be of fixed radius. But this would create an overlapping problem and would lead to double counting the observations.

We predict the ridership in two ways: for the entire system and station by station. The forecast for the entire non-CBD stations fits well the actual ridership. The error is less than 10 percent. In the second case, some stations are predicted very well but others are far from the actual numbers. The predicted standard error using the Residence 20 model is 2,483. The constant is positive for the residential stations and negative for the job stations. The explanation is that other variables than those selected have an impact on the ridership: the quality of service, the price of the fare, tourists visiting. For instance in the summer there is an important number of visitors, which use either the subway or the bus. The Local bus variable, which is similar to the feeder variable, (stations where the number of riders is higher than the population living in the area defined by the 0.5-mile radius), has a positive impact on the subway travel for all models. The higher the Number of Riders entering from Local Bus, the higher the number of subway riders. Employment and the number of jobs created in the area have also a positive contribution to subway travel but the first variable has a bigger weight. The impact of the education variable on the subway travel depends on the level of education: in almost all models (except Job 50) people with an education level at most equal to high school diploma tend to ride the subway less. This fact is explained by the type of jobs that are generated in New York

City, such as professional jobs and high skill jobs. These jobs usually hire people with at least a college degree. In addition, it is known from the economic literature that the more skilled the labor force is, the easier it is to commute to different work places. This is partially reflected in the coefficient of level of education with more than a high school diploma. This may also be explained indirectly by the high correlation between education and income. Population has as expected a positive impact on subway travel. Immigration has a negative impact on the subway travel; as we explained it in the previous chapter certain groups of immigrants usually tend to use other means of transportation than the subway such as private vans or buses. The express station variable is not statistically significant at all. The feeder station dummy has a negative sign and we would expect a positive sign. But the exchange station has the correct sign; we expect that at high volume stations, i.e. exchange stations, the number of subway riders is higher, the number ranges from 2710 to 6970 additional riders for these stations. The area variable, which controls for the size of the area drawn around the station, has a positive impact on subway travel; the greater the area the higher the population and potentially the higher the number of subway riders. The median household income coefficient has mixed signs and varies with the living area, i.e. job area or residential area. In the case of Residential 40 and Residential 50, the median household income has a negative sign. This means that riders with higher income living in residential areas use less the subway. For these riders the subway travel is regarded as an inferior good; an increase of 1 percent in income decrease the number of rides by exactly 1 percent. The number of evaders is significant only in one model but has the expected sign and magnitude, for every person not paying the fare

there is 1 less rider; this can be explained by the fact that most of those evaders are caught so they do not ride the subway.

We show in Table 9 the actual versus the estimated ridership for 1990 for a sample of 233 stations. The estimated model, residential 20 is underestimating ridership for all non-CBD stations by 3.6 percent. We show in Table 10 a random sample of stations and a comparison between actual and estimated ridership. It is clear that on a station by station basis, the model can be off considerably.

III.1.5) Discussion of results.

First, we determined that variables as employment and population have an important impact on the subway travel. We estimated these variables at the station level and we determined that they are sensitive to the area drawn around the station.

Second, we showed that the model applied to the entire system provides a good estimate with an error of less than 4 percent. This model can be used to predict the ridership at new stations or for new segments of subway. But this model has some limitations. One of the limitations is that the data set does not give us an accurate view of who exactly rides the subway and why people living in the same area and given they own a car make different transportation choices for their activities, i.e. using a car versus a bus to go to the subway station and then to use the subway. We will develop this approach using a discrete choice variable in the following section.

III.2) The individual level results.

III.2.1) Description of the data set.

The data set used here is derived from the 1990 New York City Transit Authority Subway Intercept Survey. This data set provides a reliable and geographically defined picture of individual subway travel behavior during an average week. The survey started in October 1st, 1990 and ended in December 1990. It surveyed a defined group of stations per day and it proceeded borough by borough-the first month was spent in Brooklyn, the second month in Manhattan and Queens, and the third month in Bronx. In total, around 2.25 million survey forms were distributed at almost all New York City subway stations. Around 385,000 forms were returned and were manually encoded into subway survey records. These records were validated using two main criteria:

a) identifying any missing data and data values which are not within the boundaries or expected range.

b) verifying the logical consistency of the data for each record- i.e. a train can be boarded at a station.

The data can be devised in two categories:

1) data necessary to define the subway system and the survey forms.

2) data necessary to define the rider's characteristics.

These records include three "critical" flags. If one of these flags were set to true the record would be rejected. These flags are:

1) Bad origin line/station.
2) Bad destination line/station.
3) Bad code for number of trips.

Charles River Associates in cooperation with Urbitran Associates and the MTA did calculate the necessary weights. We use this weighting system for the entire data set. The weight is computed as follows:

$$w_st= e_st/ r_st \quad (49)$$

Where

w_st- is the weight for respondents entering station s in time t;

e_st- is the total number of entries (gate and turnstile) for station s in time t;

r_st- is the total number of qualifying records for station s in time t.

III.2.2) Variables used and descriptive statistics.

In this part a descriptive analysis of the data is done to better understand how those variables impact the choice of mode for riders. In Part III.1 we determined from the cross-sectional analysis what variables have a statistically significant impact on the subway travel. In this part we use this micro data set to better understand mode choice decisions. In particular we address the following question: " If two people live approximately the same distance from the subway and they work in the same area, why does one take a car to the subway station and the other a bus given they both own a car? "

In addition to the main variables used with the first data set we derive additional time and cost variables, (Wohl and Martin [6]). Travel time and auto operating cost were estimated as generic within the data set. As Hensher [36] suggests there is no microeconomic theoretical reason for treating them as data set specific. We follow D.McFadden [5] and define the full transportation cost as the sum of a "pure" monetary cost and travel time cost. This is expressed as follows:

$$P^{*} = M(I) * t + f, \qquad (50)$$

Where M (I) is a linear function of annual income, t is the travel time and f is the fare. In addition, t does vary with usage levels. In general, subways run every five minutes during peak hour, every 15 minutes during midday, and every 20 minutes in

evening[8]. In the transportation theory it is agreed that travel time can be determined as follows:

Travel time = Vehicle Time + Line-Haul Time + Waiting Time + Transfer Time (51),

where

Waiting Time = Transfer Time = ½ Line-Haul Time. (52)

Then we can rewrite (51) as follows:

Travel time = Vehicle Time + 2*Line-Haul Time (53)

Travel time by bus or car to subway station assumes only one of five values, one for each borough. Travel time is expressed in minutes and income is expressed in dollars. We assume that the number of hours worked annually is 2,000 hours and 5 percent of personal disposable income is allocated to transportation. After defining travel time, the next step is to describe the variables from the survey and the one constructed. The variables include the number of rides in a weekday, mode of travel to the subway station, purpose of the trip, fare paid, time of day trip, auto operating cost, auto in-vehicle time, walk time, bus time, socioeconomic variables such as ethnic group of riders, income of riders, gender of riders, age of riders. The auto operating cost is constructed as follows:

Auto Operating Cost = Number of Miles Traveled x Average Cost per Mile (54)

[8] This line-haul time comes from the subway line schedules provided by the MTA.

Where the number of miles traveled to the subway station is known and the average cost per mile is a fixed cost of 33 cents.

The percentage and frequency of these variables are shown in Table 11 through Table 21. Using the weight developed by Charles River Associates, we estimate the total ridership on any given weekday at 3,612,516. From this total, 43 % are originating in the Manhattan CBD, 21.5 % in Brooklyn, 21 % in Queens, 13 % in Upper Manhattan, and 9.3 % in Bronx, Table 15. From these totals the majority 83.4 % walks to subway station, 5.7 % use a local or express bus, 3.4 % are railroad commuters, and 3.1 % drive to the subway station, Table 20. The high percentage of people walking to the station is explained by the fact that almost half of these riders are originating from the Manhattan CBD where the subway stations are within walking distance-in our case within a radius of 0.5 miles from the subway station. The next question to address is where do these riders go when they use the subway? From the total trips, 45 % of trips are work related and 20 % are school, shopping, recreation or other, Table 12. How do these riders pay for these trips? Most of them pay a regular fee, 96%, 2.2 % pay half fare (seniors), and 1.2 % are students, Table 14. How many trips do these riders make per week? 38 % make between 6-10 trips a week, 37 % make between 11-15 trips per week, 11 % make between 16-20 trips per week, 10 % make between 0-5 trips per week, 5 % make more than 21 trips per week, Table 13. The majority of the trips are peak-trips, AM Peak and PM Peak for a total of 70.3 %. The remaining trips are midday trips 22 %, and evening trips 8 %, Table 21. The age of the riders can be broken into three distinct groups- the main group composed of riders whose age is between 25-39 represents 47 %, followed by

the 40-45 group which represents 25 %, and the 18-24 group which represents 15 %, Table 11. The riders are almost evenly divided between female and male, 53 % and 47 % respectively, Table 16. The distribution of income characterizes the means of transportation- 75.3 % of riders earn below $50,000, 10.3 % earn between $50,000 and $75,000 and 9.5 % earn over $75,000, Table 19. The ethnicity of the respondents can be classified in four groups- 48 % White, 22 % Afro-American, 16.4 % Hispanic and 8 % Asian, Table 17. The destination of the riders is primarily to the Manhattan CBD 53 %. The remaining is spread evenly to Brooklyn 15.3 %, to Upper Manhattan 14 %, to Queens 11 %, and to Bronx 7 %, Table 18.

III.2.3) Econometric model and estimation technique.

The econometric model we estimate here was described in Part II and is similar to the one described in McFadden [3]:

$$\ln\left[\frac{\pi}{1-\pi}\right] = \alpha\, w_1 t_w + \alpha w_2 (t_a - t_s) + \alpha (c_a - c_s) + \beta X \qquad (55)$$

In this model π can be formalized as follows:

$\pi = P_n(f,d,m|a)$ = probability of rider n making a trip(f=1) to destination d by mode m, conditional on rider auto ownership, a. The mode options are auto or bus.

The model predicts the probability that a rider will choose a particular mode of travel from home to the subway station, given the mode of travel characteristics, the household characteristics and given the riders live in the same area and they all own a car.

This model estimates the log-odds ratio of choosing the car over the bus to go to the subway station. We estimate this conditional logit model by using the PHREG procedure from SAS. Originally, PHREG was designed to do Cox regression analysis of continuous-time survival data. The method used to estimate a proportional hazards model is the partial likelihood. We can use this procedure because the partial likelihood function is identical to the likelihood function for the conditional model. The sample is composed of 189 individual trips. For each individual trip a time of travel has been calculated, time spent walking, the time spent in the bus, and auto in-vehicle time are available, besides the individual characteristics. In addition to the bus fare, the auto operating cost is calculated based on an operating cost of $.33 per mile. The event modeled here is the odds of choosing auto over bus, i.e. $\pi=1$ if the individual chooses the auto to go to the subway station. We estimate the model by maximizing a log likelihood function having the form of equation (28).

$$\text{Log}[P(\text{Auto})/P(\text{Bus})]=\alpha\text{BUSWALKTIME}+\beta\text{DIFAUTOBUSTIME}+\gamma\text{DIFAUTOBUSCOST}+ +\delta\text{AGE}+\varepsilon\text{INCOME}+\varphi\text{GENDER}+\mu\text{ETHNIC} \quad (56)$$

The detailed results are shown in Table 22 and the mean values by mode are shown in Table 23. Although this model may appear to be a multinomial logit[9] it differs in two ways:

1. The explanatory variables can include characteristics of the choice options and also variables that describe the relationship between the individual and the option.

2. The set of available options may vary among individuals.

The independent variables in this model fall into two classes:

1. Socioeconomic characteristics of the household.

2. Mode-specific variables.

The explanatory variables are: bus walk time (BUSWALKTIME), the difference between auto in-vehicle time and time spent in the local or express bus (DIFAUTOBUSTIME), the difference between auto operating cost and the local or express bus fare (DIFAUTOBUSCOST), income (INCOME), age of riders (AGE), gender of riders (GENDER), and ethnicity of rider (ETHNIC). The list of explanatory variables is shown in Table 24.

The time variables are expressed in minutes, the cost variables are expressed in 1990 dollars, income is expressed in 1990 thousands of dollars, gender is a dummy variable, and ethnic is a categorical variable.

[9] In fact a multinomial logit is a particular case of the conditional logit; one can go from one to another under certain conditions, McFadden [15].

Not all the parameters estimated quite reach statistical significance but all have the expected signs and magnitude; the walking time variable has a p-value of 0.09, the age variable has a p-value of less than 0.01, and the difference in costs for auto and bus has a p-value of 0.21. The results in Table 22 show negative effects of auto in-vehicle time less bus travel time, auto operating cost less fare cost, and age variables. Each additional minute increase in the difference between auto in-vehicle time and time spent in the bus reduces the odds of choosing auto by 1.5% (100x(1-0.985)). An increase in the difference between auto operating cost and bus fare by one unit reduces the odds of choosing the car by 63.8%. An increase in the age by one unit decreases the odds of choosing the car to go to the subway station by 48.1%. Income has a positive impact on the odds of choosing a car versus taking the bus to travel to the subway station. An increase in gender or ethnic bracket by one unit increases the odds of choosing the car to travel to the subway station by 54.6% and respectively by 2.8%. Each additional minute of walking to the subway station increases the odds of choosing the car by 10.8%.

We use the coefficients estimated in equation (56) to derive the value of time for each mode. Since the time variable is in minutes, the hourly value of walk time is equal to (0.10/1.01)*60=$6 and the value of auto in-vehicle time is (0.015/1.01)*60=$0.9. We conclude that the value of walk time is about 6 times higher than the value of in-vehicle time. The same magnitude was produced by McFadden's estimates.

We also estimate the aggregate cross elasticity for auto mode. Aggregate cross elasticities measure the effect of an incremental change in a variable on the expected share of the group choosing alternative i, (Ben-Akiva [8], Hensher [36]). We define $\overline{P}(i)$ as the expected share of the group choosing alternative i as following:

$$\overline{P}(i)=\frac{\sum_n^N \overline{P}_n(i)}{N}, \tag{57}$$

Where N is the number of decision makers in the group. Now, suppose we change the value of some variable x_{jnk} for each individual by a small increment so that

$$\frac{\partial x_{jnk}}{x_{jnk}}=\frac{\partial x_{jn'k}}{x_{jn'k}}=\frac{\partial x_{jk}}{x_{jk}}, \quad \text{for all n, n'=1,2,....,N,} \tag{58}$$

Where

$$x_{jk}=\frac{1}{N}\sum_{n=1}^{N} x_{jnk} \tag{59}$$

Equation (58) shows the percentage change in x_{jnk} is uniform across all members of the group. Then, the aggregate elasticity can be expressed as follows:

$$E_{x_{jk}}^{\overline{P}(i)}=\frac{\sum_{n=1}^{N} P_n(i)E_{x_{jnk}}^{Pn(i)}}{\sum_{n=1}^{N} P_n(i)} \tag{60}$$

The formula in (60) is a weighted average of the individual level elasticities (disaggregate elasticities) using the choice probabilities as weights. These disaggregate elasticities can be written as follows:

$$E_{x_{jnk}}^{P_n(i)} = [\delta_{ij} - P_n(j)]x_{jnk}\beta_k \qquad (61),$$

where δ_{ij} is the Kroenecker delta function, which equals 1 for i = j and 0 for i ≠ j. Since we are only calculating direct elasticities, i.e. $\delta_{ij}=1$, by substituting expression (61) in (60) we obtain

$$E_{x_{ink}}^{\bar{P}(i)} = \frac{\beta_k}{N \cdot \bar{P}(i)} \cdot \sum_{n=1}^{N} P_n(i)[1 - P_n(i)]\bar{x}_k \qquad (62)$$

or $$E_{x_{ink}}^{\bar{P}(i)} = \frac{1}{\sum_{n=1}^{N} P_n(i)} \cdot \beta_k \sum_{n=1}^{N} P_n(i)[1 - P_n(i)]\bar{x}_k \qquad (63)$$

The aggregate elasticities for the group of individuals choosing auto to go to the subway station are shown in Table 25. These elasticities are evaluated at the sample mean value[10] of each variable. For almost all predictors, each percentage increase in the independent variables increases the expected share of the group choosing auto. The age variable has a negative effect on the expected share of the group choosing auto as a mode to go to the subway station.

The measure of goodness of fit for a discrete model, i.e. the dependent variable is qualitative, can be determined either in terms of fit between the calculated probabilities

[10] The average values are shown in Table 23.

and observed frequencies or in terms of the model to forecast observed responses. This applies to individual data and grouped data. We show here three measures of fit:

a) the Hosmer and Lemeshow goodness-of-fit test.

b) the generalized coefficient of determination (R-square).

c) the adjusted generalized coefficient of determination (R-square).

The Hosmer and Lemeshow [18] test involves dividing the data into ten groups of equal size based on the percentiles of the estimated probabilities. The observations are sorted ascending by their estimated probability of the event outcome. The differences between the expected number of observations and the observed are summarized by the Pearson chi-square statistic. This statistic is compared to a chi-square distribution with t degrees of freedom, where t is the number of groups minus 2. The statistic can be written as follows:

$$X^2_{HW} = \sum_{i=1}^{t} \frac{(O_i - N_i \bar{\pi}_i)^2}{N_i \bar{\pi}_i (1 - \bar{\pi}_i)} \tag{64}$$

Where

N_i is the number of observations in the i^{th} group.

O_i is the number of event outcomes in the i^{th} group.

$\bar{\pi}_i$ is the average estimated probability of an event outcome for the i^{th} group.

The Hosmer and Lemeshow statistic is compared to a chi-square distribution with (t-2) degrees of freedom.

For the individual observations we have different generalized coefficients of determination. Cox and Snell [19], Maddala [9], and Magee [20], proposed a generalization of the coefficient of determination to a general linear model:

$$R^2 = 1 - \left[\frac{L(0)}{L(\hat{\beta})}\right]^{\frac{2}{n}} \tag{65}$$

Where *L(0)* is the likelihood of the intercepts only model (restricted model when all parameters are 0), $L(\hat{\beta})$ is the likelihood of the specified model, and *n* is the sample size. Nagelkerke [21] derived an adjusted coefficient, which can achieve a maximum value of 1:

$$R^2_{adj} = \frac{R^2}{R^2_{\max}} \tag{66}$$

In our model we will use the definition of (65):

$$R^2 = 1 - \left[\frac{L(0)}{L(\hat{\beta})}\right]^{\frac{2}{n}} = 0.3 \tag{67}$$

The model predicts the correct choices in 30% of cases. The ability of the model to predict a choice of mode for work trips is acceptable. In this case we have only two choices, so one would expect on a random basis to select the correct choice for 50 % of the sample.

III.2.4) Discussion of results.

We derived the results at the individual level, and we estimated the value of time spent in auto and the one spent on bus while traveling to the subway station. We derived aggregate cross-elasticities for choosing auto with respect to the independent variables. The model can be used in applying fare policies. One can now predict how a given sample of the population having similar characteristics, i.e. living in the same area, taking the subway to work, having a car, will react to such fare policies.

But using cross-sectional data has some limitations. First, we can not determine meaningful fare elasticities, since we do not account for the variation of fare in time. Even with few changes in fare along the years one would derive a good elasticity estimate. Second, by using a discrete approach with micro level transportation data we can only predict a mode choice for a given sample. So, to forecast the subway travel we need to use time series data. We will develop this estimation in the next chapter.

III.3) The aggregate level results.

III.3.1) Introduction.

In the previous two parts we showed the variables having an impact on the subway travel, and we built a discrete choice model for the travel to the subway station. We also estimated the mode elasticity if auto was the choice of travel to the subway station. But in order to build a complete model for transportation choice and to implement public policies as well as to forecast revenue losses or gains due to the implementation of these policies we need to estimate a time series model. We give a detailed explanation of the notations used in this chapter in Annexe 1. In order to build a transportation model that we will use to forecast subway travel and revenue we need to proceed in logical steps.

First, we explore the data graphically to understand the different time trends, seasonal factors, and other outliers.

Second, we decide on the best model estimations to use, i.e. exponential smoothing, dynamic regression. Each model addresses a different question; the exponential smoothing model is used to forecast the demand for subway travel, and the fare. The dynamic regression is used to capture the behavioral response of the subway ridership to the exogenous variables and to estimate the subway travel elasticity. The exponential smoothing model is widely used when one can not use Box-Jenkins or dynamic regressions; this happens when the historical data is too short or it is not stable, i.e. non-stationary. On the other hand, dynamic regression combines time-series dynamic

features and explanatory variables. We use here both techniques. We then forecast the demand for subway travel based on the future values of the exogenous variables.

Third, we validate the model by diagnostic examination of the forecast errors generated by the model over the historical data. If the model passes the tests then we run it on the historical data and we extrapolate from the last historical data point. We need also mention the main difference between forecasting and cross-sectional data models: what matters with time series is the external validity of the model, not the internal validity such as goodness of fit.

So, in this chapter we show the usefulness of the time series data set. We also show the main contribution of using time series, i.e. forecasting. The time series approach extrapolates information from the past and assumes that the past events will continue to propagate into the future. In addition, time series models have the advantage of providing us with a direct relationship between the main macroeconomic variables and the subway travel variable.

We showed that the use of cross-section data as well as the use of microeconomic data at the disaggregate level had some limitations as to the implementation of public policies or new fare schedules. We build in this chapter time series models that allow us to estimate and forecast the following: a complete demand-price model for the entire subway system, the aggregate subway elasticity, a forecast for the demand for subway travel, the fare, and the revenue for two years. In the next part we describe the time series data set.

III.3.2) Description of the data set.

The data set used in this chapter is composed of monthly time series from January 1978 to December 1996. The data set was compiled using data from the Office of Management and Budget at the Metropolitan Transportation Authority, and data from the Office of Management and Budget at the City Planning of New York City. Some of the variables were transformed by dividing by population or by taking the natural logarithm. The geographical area is New York City, and the five Boroughs.

This data set allows us to estimate the subway fare elasticity at the aggregate level, income elasticity, and cross elasticities such as price of energy elasticity. We can also estimate what is the effect of aggregate shocks to the economy such as change in the unemployment rate, or change in the energy price on the subway travel. In addition, we use this data set to forecast the demand for subway trips, the real fare, and other macroeconomic variables. These results can be used as a tool to support project investments such as new subway cars, new subway stations, and new lines. In the next part we describe the variables included in the data set.

III.3.3) Variables used and descriptive statistics.

In this part we describe the variables used and we do a descriptive analysis of the data to better understand how those variables are used in the time series frame. In the previous chapter we estimated the model at the individual level and we derived aggregate elasticities for the demand for auto travel to the subway station. The demand elasticity for

subway travel is derived using time series and after transforming the original variables. We use here five types of transformations. We express the price variables in real terms by dividing them by the consumer price index, we divide the original variables by the population variable, we take the log of those variables and we estimate a log-log model, we multiply the fare variable with dummy variables for fare increases, and we construct a 2 month lagged variable for the real fare. So, some of the variables are expressed in per capita units. In addition, to dampen the effect of differences between months due to different number of days we normalize the subway riders per capita by multiplying this variable by a coefficient, $\kappa = \frac{30}{Days/month}$.

We use three sets of dummy variables, one set to control for the seasonal factors, a second set to control for the strike in 1980, and another set to control for the fare hikes. We construct 12 dummy variables, each one taking value 1 for a given month, i.e. JANUARY=1 if the month is January and 0 otherwise. We construct a dummy variable, HIKE, that takes the value 1 for the month where the fare hike occurs and 0 otherwise. We also construct a dummy variable, STRIKE, that takes the value 1 for the month of April 1980 and 0 otherwise. All these variables are contributing to the theoretical model by allowing us to estimate aggregate elasticities and by accounting for real shocks from the New York City economy through the employment or price index variables. In addition, this data set allows us to test the hypothesis that subway travel in New York City is negatively correlated with subway crime rates. The variables from this data set are: the indexes for Consumer Price Index, the Energy Price Index, Total Employment, Private Employment, the Rate of Unemployment, the Number of Subway Riders, the Subway Fare, the Number of Subway Felonies, the Population, and Aggregate Income.

The constructed variables are the following: the Real Fare, the Real Energy Price, the Real Per Capita Income, monthly dummy variables, i.e. JANUARY, FEBRUARY and so on, monthly dummy variables for fare hikes, i.e. HIKE, and the dummy variable STRIKE. Next we plot the main variables.

As shown in Graph 3, subway ridership per capita is cyclical and has been influenced by economic factors; in 1980, after the second oil shock the subway ridership per capita is at its lowest level as suggested by the observed trough. From 1988 to 1992 the subway ridership per capita is declining. As the New York City region emerges from the 1990 recession the trend is reversed in 1993, and the subway ridership is increasing steadily. Although the nominal fare has increased continuously since 1978, the real fare has been declining during the periods preceding the two economic crises of 1981 and 1990. Partly this explains why even when the nominal fare increased in 1995, the subway ridership continued to grow the following years. As we show in Graph 2, the real fare increases from 1978 to 1986, when it reaches a peak. Then the real fare is slightly increasing and is stable throughout the following years. The relationship between the fare and the subway ridership shown on Graph 1 suggests that ridership and real fare are counter cyclical. From 1978 until 1992, each real fare increase is followed by a small decrease in subway ridership. This period corresponds also to a fragile economic situation. Following the last increase in fare to $1.50 in November 1995, the subway ridership is increasing steadily. Another economic aggregate variable is closely related to the subway ridership, the private employment. We showed in part III.2 that 80% of home-based trips are done for the purpose of work. We show on Graph 5 that private employment is procyclical with subway ridership per capita for the period 1984-1991; as

the number of people employed in the private sector starts to grow, beginning the post recession period of 1992, so does the subway ridership per capita. The relationship between the real income and the subway ridership is more ambiguous and it depends on how the subway trips are perceived. We showed in the main model, in chapter II, if the subway trip is perceived as a "normal[11]" good then as real income increases the number of subway trips increases since the good becomes relatively cheaper. But this is shown to work also in the opposite direction. As real income per capita increases, as we show on Graph 7, the share of income for auto transportation becomes lower than the share for subway transportation, and other transportation means become more affordable. Then subway ridership is redirected towards those new transportation choices such as autos, cabs, private cars, and vans. Another important variable that has an indirect impact on the subway ridership per capita is the real energy price. Particularly, the real energy price is expected to be negatively correlated with any means of transportation that use energy, such as autos, buses, vans, and to some extent the subway. By comparing Graph 3 and Graph 4, we can see that the real energy price index is procyclical with the subway ridership per capita except for the period 1992-1996. The base year used for the energy price and the CPI is 1983. We can distinguish four periods. The first time period, 1978-1982, shows a spike in the real energy price and a continuous decrease in the subway ridership. The second time period, 1982-1987, is followed by a modest variation in subway ridership and a decrease in the real energy price with the level of 1986 corresponding to 1978. In the third time period, 1988-1993, the real energy price continue to decrease slowly and the subway ridership reaches a second through in 1992 with 995

[11] This means the subway trip is neither a luxury good nor a necessity good.

million passenger trips. During the fourth time period, 1993-1996, the real price index shows a slight negative trend whereas the subway ridership is increasing. So, we can conclude that as the real price index rose in the late 1980s, the number of subway riders increased since subway became cheaper than auto, and since auto and subway trips are substitutes. In Graph 5 we plot both the subway ridership per capita and the private employment series. We can easily see that private employment is procyclical with subway ridership. For the 1984-1991 time period, an increase in private employment corresponds to an increase in the subway ridership. This result validates both the cross-sectional results from part III.1, where we show that the number of subway trips is sensitive to employment, and also the micro sample results from part III.2 where we show that most of the trips are job related. Indeed, this relationship is the best shown for the 1991-1996 time period; it appears that each trough from the subway ridership series coincides with a trough from the private employment series. Finally, we looked at the number of subway felonies as a proxy for the quality of subway ridership, Graph 6. We see that between the 1990-1996 the quality of subway ridership increased, i.e. the number of felonies per capita decreased by 38% due to an increase in the number of the subway police officers. For the same period the number of subway riders increased by 18% and we can safely assume that at least a small number of those additional riders is due to the increase in the quality of rides. After showing that a relationship exists between subway ridership per capita and the exogenous variables described above we will estimate that relationship. We use two models; to forecast the subway ridership and other economic variables we estimate an exponential model; to determine the subway elasticity we estimate a log model. We construct these different models in the next part.

III.3.4) Econometric model and estimation technique.

III.3.4.1) Exponential smoothing.

In this section we estimate the demand for subway trips and the fare, and we forecast these two series to the year 2002. We estimate the series by using the Winters technique. Since we suspect the subway travel series to be seasonal we use the Winters technique which estimates three smoothing parameters: level, trend, and seasonal. By using the Winters model, we assume that each observation is the product of a deseasonalized value and a seasonal index for the month. We can write the forecasting equations for the multiplicative Winters model as follows:

$$\hat{Y}_t^j(n) = (l_t + nk_t)\hat{i}_t(n) \tag{68}$$

With $\hat{Y}_t^1 \equiv$(SUBWAY RIDERS/POPULATION) (69)

And $\hat{Y}_t^2 \equiv$(REAL FARE) (70)

And the smoothing equations are:

$$l_t = \alpha \frac{Y_t}{i_{n-p}} + (1-\alpha)(l_{t-1} + k_{t-1}) \tag{71}$$

$$k_t = \beta(l_t - l_{t-1}) + (1-\beta)k_{t-1} \tag{72}$$

$$i_t = \gamma \frac{Y_t}{l_t} + (1-\gamma)i_{t-p} \tag{73}$$

The results and the forecast for each year up to the year 2002 are presented in Table 26 and Table 27 and in Graph 8 and Graph 9. By examining the Graph 9, we can see the justification of the dummy variable, STRIKE, for the year 1980 where we have a trough due to a subway strike. The subway ridership is forecasted to increase within the next years due to the introduction of the Metro Card; from 1999 to 2002 we forecast a 5 percent yearly increase with a high of \$1.47 billion passenger trips in 2002. The average fare per ride is forecasted to decrease 14 percent in 1999, and 10 percent thereafter, from \$0.96 in 1999 to \$0.70 in 2002. The average fare per ride is equal to the total revenue collected from fares divided by the total number of rides in a year. This average fare declined steadily since the introduction of discount fares in 1997, and since the introduction of the Metro Card. Subway riders paid an average of \$1.12 per trip in 1998 compared to \$1.44 in 1996, a 28.5 percent drop in the average fare per trip. This decline in average fare per ride caused a decline in revenue but it was compensated partially by an increase in ridership, (New York City Independent Budget Office [32]). The subway's

revenue is derived from fares, direct subsidies from government, tax-supported subsidies from state and city taxes, and transfers from the TBTA[12]. According to the IBO[13], in 1998, around 56 percent of revenues were generated from fares, 35 percent were from direct subsidies, 4 percent from indirect subsidies, and about 5 percent from other sources such as advertising income. The subway revenue is affected by the number of riders and by the level of fares as shown in Table 27. We forecast the revenue by multiplying the average fare per trip and the total number of trips made during a year. For 1997 and 1998, the NYC Transit subway passenger revenue was 1,483 millions and respectively 1,361, an 8 percent decrease. Based on the forecast for the subway, and the forecast on the average fare per trip, we derive the fare revenue for the years 1999-2002. If we assume that subway trips will increase 5 percent each year and the average fare will decrease in 1999 by 14 percent and in the following years by 10 percent then the fare revenue would decrease each year by 5.5 percent starting in year 2000. We showed here that with an exponential smoothing model we estimate the revenue and the impact of future fare changes on the revenue. But this model does not account for other explanatory variables and is not a dynamic model. In addition, since we did not make any particular assumption about the probability distribution of the series we do not need to do any error diagnostic. We will develop a dynamic-causal model in the next part and we will test for ARCH presence and for co-integration between the predictors.

[12] Triborough Bridge and Tunnel Authority.

III.3.4.2) Dynamic model.

We showed in the previous part the forecast for the demand for the subway travel, real fare, and the subway revenue using exponential smoothing. In this part we develop a dynamic model that includes the explanatory variables. We start with a simple specification and we improve the model by taking into account seasonality and lags in the real fare, i.e. we use 11 dummy variables and a 2 month lag of the real fare, and by accounting for special events such as fare hikes or strikes. We test these models for cointegration between the independent variables and for the presence of ARCH[14] residuals. We use these elasticities to estimate the impact of public policies.

The elasticities are derived directly from the coefficients; we use a log model to dampen the effect of autocorrelation. We estimate the model (7) using the time series data set described in III.3.2. The model is estimated using the Yule-Walker method, Hamilton [22]. We correct for autocorrelation and we also estimate the AR (1) coefficient of the disturbance. The econometric models are shown below:

$$\text{Log}[(\text{AggregateRidership/Population})*k]=a_{0t}+a_{1t}\text{Log}(\text{FARE/CPI})+a_{2t}\text{Log}(\text{EnergyPrice/CPI})+a_{3t}\text{Log}(\text{PrivateEmployment/Population})+a_{4t}\text{Log}(\text{Fellony/Population})+a_{5t}\text{Log}(\text{Income/Population/CPI})+a_{6t}\text{STRIKE}+e_t \quad (74)$$

---

[13] Independent Budget Office.
[14] Autoregressive Conditional Heteroskedasticity is abbreviated to ARCH.

$$\text{Log}[(\text{AggregateRidership/Population})*k]=a_{0t}+a_{1t}\text{Log}(\text{FARE/CPI})+a_{2t}\text{Log}(\text{EnergyPrice/CPI})+a_{3t}\text{Log}(\text{PrivateEmployment/Population})+a_{4t}\text{Log}(\text{Fellony/Population})+a_{5t}\text{Log}(\text{Income/Population/CPI})+\text{JANUARY}+\text{FEBRUARY}+\text{MARCH}+\text{APRIL}+\text{MAY}+\text{JUNE}+\text{JULY}+\text{AUGUST}+\text{OCTOBER}+\text{NOVEMBER}+\text{DECEMBER}+e_t \quad (75)$$

$$\text{Log}[(\text{AggregateRidership/Population})*k]=a_{0t}+a_{1t}\text{Log}(\text{FARE/CPI})+a_{2t}\text{Log}(\text{EnergyPrice/CPI})+a_{3t}\text{Log}(\text{PrivateEmployment/Population})+a_{4t}\text{Log}(\text{Fellony/Population})+a_{5t}\text{Log}(\text{Income/Population/CPI})+\text{JANUARY}+\text{FEBRUARY}+\text{MARCH}+\text{APRIL}+\text{MAY}+\text{JUNE}+\text{JULY}+\text{AUGUST}+\text{OCTOBER}+\text{NOVEMBER}+\text{DECEMBER}+\text{STRIKE}+e_t \quad (76)$$

$$\text{Log}[(\text{AggregateRidership/Population})*k]=a_{0t}+a_{1t}\text{Log}(\text{PrivateEmployment/Population})+a_{2t}\text{Log}(\text{Fellony/Population})+a_{3t}\text{Log}(\text{Income/Population/CPI})+a_{4t}\text{Log}(\text{FARE/CPI})_{t-2}+\text{JANUARY}+\text{FEBRUARY}+\text{MARCH}+\text{APRIL}+\text{MAY}+\text{JUNE}+\text{JULY}+\text{AUGUST}+\text{OCTOBER}+\text{NOVEMBER}+\text{DECEMBER}+e_t \quad (77)$$

$$\text{Log}[(\text{AggregateRidership/Population})*k]=a_{0t}+a_{1t}\text{Log}(\text{FARE/CPI})(a_{2t}\text{HIKE1}+a_{3t}\text{HIKE2}+a_{4t}\text{HIKE3}+a_{5t}\text{HIKE4}+a_{6t}\text{HIKE5}+a_{7t}\text{HIKE6}+a_{8t}\text{HIKE7})+a_{9t}\text{Log}(\text{EnergyPrice/CPI})+a_{10t}\text{Log}(\text{PrivateEmployment/Population})+a_{11t}\text{Log}(\text{Fellony/Population})+a_{12t}\text{Log}(\text{Income/Population/CPI})+e_t \quad (78)$$

The estimated coefficients for the five models shown above are shown in Table 28, Table 29a, Table 29b, Table 30, and Table 31. The coefficients have the expected

magnitude and sign. The log of subway riders per capita is adjusted for the variation of the number of days in the month by multiplying the variable by a coefficient k.

The first model includes the main variables and the variables are all statistically significant. The explanatory power of the model is R^2 =0.54, and the DW=2.16 after autocorrelation correction. The constant is negative and statistically significant which means that if the main predictors are zero no trips are generated. The real price variables have a negative impact on the demand for subway travel. The real price elasticity for subway trips is -0.29 which means that a 10 percent increase in the fare will decrease the number of subway riders by 2.9 percent or on average by 3,086,083 a month based on the forecast of the 1999 annual demand. If we use an average fare of $0.96 per trip, the revenue loss generated by this demand loss would be on average $2,962,640 a month. This real price elasticity for subway trips is higher in absolute value than the real price elasticity for energy which is –0.14; a 10 percent increase in the price for energy would decrease the number of riders by 1.4 percent or the equivalent of 1,489,833 riders per month. The number of subway felonies, a proxy for the quality of the subway travel, has the expected impact on the aggregate number of subway riders. A decrease in the number of felonies by 16 percent, the annual average decrease in felonies for the last seven years, would increase the number of subway trips by 1.5 percent or by 1,596,250 trips on average per month, everything else being constant. At an average fare of $0.96 per trip this would be equivalent to a revenue gain of $1,532,400 a month. Private employment is contributing positively to the subway travel, which corroborates the findings in part III.1. Indeed around 80 percent of trips are job related so employment is an important explanatory variable. An increase in private employment by 1 percent in a given month

would increase the number of subway trips by 712,991 a month. Finally, the real income per capita has a negative impact on the number of trips. As real income increases, the household's share for transportation services becomes smaller and this may imply a change in preferences; higher income households may prefer driving their own car, while lower-income households may have no choice but riding the subway.

The second model accounts for seasonal variations in the demand for subway travel. The explanatory power of the model is R^2=0.5, close to model 1, and DW=1.94. We constructed monthly dummy variables to isolate the effect of seasonality. We know that during summer the number of trips tend to increase due to visitors that come to New York City. We show in Table 29a all coefficients of the model. All the main predictors from model 1 changed except real income per capita. But the changes in the coefficients are very small. The policy implications estimated above remain quasi the same. The seasonal factors have a mixed effect on the number of trips; the months of January, March, April, May, June, October, November, and December have a positive effect on the number of trips while the months of February, July, and August have a negative effect. The magnitude of the seasonal factors is very small which is expected since the demand for subway trips is driven mainly by job trips and not leisure trips. Job trips are not seasonal. After we drop the seasonal factors that are not statistically significant, all seasonal coefficients contribute positively to the number of trips except the coefficient for July.

In the third model we also account for the subway strike in the early 1980 by using a dummy variable. The explanatory power of the model increased, R^2=0.80 and the DW=2.04. The most significant result of this model is the dummy variable used for the

strike; it has a negative effect on the subway ridership per capita and it is very significant. Also by using the strike variable, some of the seasonal coefficients are changing: January, February have the opposite sign as compared to the model above. The coefficients are shown in Table 29b.

The fourth model includes lag of the real fare. The explanatory power is R^2=0.46 and DW=1.82. Since 1976 until 1996 the MTA increased the subway fare at 7 occasions. We want to know if the subway ridership is sensitive to those fare hikes. Since we know that there is a delay between the increase in price and the effect on ridership we include a 2-month lag of the real fare. We showed in the previous models that the real fare has a negative impact on the demand for subway trips but the magnitude of this impact is relatively small. We show that increases in the real fare have a negative impact on ridership and this effect is delayed.

The fifth model includes an interaction between the real fare and the hike periods. The explanatory power is R^2=0.44 and DW=1.70. The predictors we used in model 3 remained unchanged and the interacted real fare with the hike dummies is not statistically significant. Other variables such as real fare, the number of felonies, real energy price, real income per capita, and private employment have a significant impact on subway ridership.

In this part we developed a time series model for the subway ridership. We showed that even though subway ridership is seasonal the effect is small based on the estimated coefficients. We also showed that this effect is changing if we normalize the subway ridership per capita. We incorporated outside shocks such as strikes by including a dummy variable and show the effect was significant. We showed that when the fare was

increased it did have a lagged effect on the subway ridership. The predictors that are significant are real fare, real energy price, real income per capita, and private employment. We estimated the aggregate elasticities for these variables and we simulated the impact of different policies or real economic shocks on the subway revenue, Table 32. We simulated five effects: a ten percent decrease of the real fare, a ten percent increase of the real energy price, a one percent increase in private employment, a sixteen percent decrease in the number of felonies, and a five percent increase in real income per capita. From our estimate we can see that the first important effect is due to the real fare policy; it ranges from $21 million to $36 millions increase annually in the subway revenue. The second effect is the real shock to the economy due to increase in the real price of energy such as crude oil; the decrease of the subway revenue ranges from $14 million to $14 million annually. The impact is lesser when we account for changes in fare. The third effect is related to the quality of subway trips and we estimate the impact of an increase in the quality of subway trips on the subway revenue. If we assume the same rate of decrease in the number of felonies as for the past seven years, sixteen percent, the subway revenue would increase annually between $10 million and $18 million. The fourth effect is a real shock due to the increase in private employment; the subway revenue would increase annually between $3 million and $8 million. The fifth effect is a 5 percent increase in the real income per capita; it would decrease the subway revenue by an annual average of $0.2 to 7 millions.

In the next part we test these models for the variance of the subway ridership using an ARCH test and also for the relation between the predictors using the Phillips-Ouliaris-Hansen test for cointegration.

III.3.5) ARCH and Cointegration Tests.

III.3.5.1) Testing for ARCH.

When we estimate the coefficients of the models described above we assume that the residuals are N (0, σ^2) and that they are uncorrelated and homoscedastic, i.e. their variance is constant over time. It follows from this assumption that the estimators obtained by OLS regression are BLUE[15]. In reality, time series such as subway ridership are very volatile and hard to predict during certain periods where volatility is higher than usually such as oil crisis, recession periods. During those periods of high fluctuations, large and small residuals tend to come in clusters. This implies that the variance of the error depends on the size of the preceding error and varies with time; this is the problem of autocorrelation. In a more formal matter we can say that the variance u_t conditional on u_{t-1} depends linearly on the square of the u_{t-1}. The unconditional variance is constant, so the OLS estimators are BLUE. But because the conditional variance is heteroskedastic we can find a nonlinear estimator by MLE that is more efficient. Green [25] shows how to find an estimate that is asymptotically equivalent to the MLE.

Engle's Autoregressive Conditional Heteroscedastic (ARCH) model deals with the problem described above. We can write it first as a simple autoregressive model:

$$u_t = \lambda u_{t-1} + \varepsilon_t \qquad t=1\ldots,T \qquad \varepsilon_t \approx IN(0, \sigma^2) \qquad (79)$$

[15] Best Linear Unbiased Estimator.

where the conditional mean $E(u_t|u_{t-1})= \lambda u_{t-1}$ depends on t, and the conditional variance $var(u_t|u_{t-1})= \sigma^2$ is constant. The unconditional mean of u_t is 0 and the unconditional variance $var(u_t|u_{t-1})= \sigma^2/(1-\lambda^2)$. The ARCH model is a generalization of the AR(1) model since its conditional variance is also a function of the past observations. A general expression of the ARCH model is:

$$u_t|z_{t-1} \approx N(g(z_{t-1}),h(z_{t-1})) \qquad (80)$$

Engle specifies the conditional mean, $g(z_{t-1})$, as a linear function of the variables z_{t-1} and the conditional variance h as:

$$h_t=\alpha_0+\alpha_1 e^2_{t-1}+\alpha_2 e^2_{t-2}+\alpha_3 e^2_{t-3}+\ldots+\alpha_p e^2_{t-p} \qquad \varepsilon_t \approx IN(0,h_t) \qquad (81)$$

with $\varepsilon_t=u_t-g_t$. In a simple case we estimate the following model:

$$Y_t=\beta X_t+\varepsilon_t \qquad \varepsilon_t \approx IN(0,h_t) \qquad (82)$$

$$h_t = var(\varepsilon_t) = \alpha_0+\alpha_1 e^2_{t-1} \qquad (83)$$

In this case we test for ARCH by using the following null hypothesis:

$$H_0: \alpha_1=0 \qquad \text{versus } H_A: \alpha_1 \neq 0 \qquad (84)$$

Under the null hypothesis we test that the error variance is not a conditional process, i.e. that there is no ARCH process. The most common way to test for ARCH is a LM[16] test. In this test the square of the OLS disturbances is regressed on a constant and its lagged values and the $M=TxR^2$ statistic are calculated. It is distributed under the null hypothesis as a χ^2 distribution with degrees of freedom equal to the number of lags. If the values calculated are greater than the critical value then we reject the null hypothesis, i.e. we have evidence of the presence of ARCH effects. This will allow us to say if the significant DW[17] is due to correlation in the ε_t or due to the ARCH effect. In our example we calculate the M statistic for each model and we show the results in Table 33. Excepting the first model all other models are showing evidence of ARCH process. Therefore we adjusted those models and we estimated the AR(1) error term for each one. In the next part we test the models to determine if the predictors are cointegrated.

III.3.5.2) Testing for Cointegration.

Many economic series (employment, income) are integrated processes, i.e. they are stationary only after differencing. But there are economic forces that keep pairs of

[16] LM stands for Lagrange Multiplier test; it is asymptotically close to the Wald Test or the Likelihood Ratio Test.
[17] DW is the Durbin-Watson statistic.

economic series together, even though individually those economic series may drift. In our case the regression that involves multiple I(d)[18] variables can be written as:

$$Y_t=BX_t+u_t \tag{85}$$

The regression above produces autocorrelated but stationary error terms u_t. Then we can say the Y_t and X_t are cointegrated as in Granger [26]. We interpret the above regression as an equilibrium condition among the variables and u_t represents the deviation from this equilibrium. Then we can use this forecast to predict long run trends even though the errors are autocorrelated because in the long run the errors dampen since they are stationary. But in the short run the forecast can be off.

We can interpret the cointegration relation by looking individually at the graphs of the following series: subway riders per capita, real fare, real energy price, private employment, number of felonies, income per capita. We see that each series has no defined trend but when used in relation to the subway per capita riders series they would have a defined trend. This is what we want to test here, if the predictors are cointegrated. We use the Phillips-Ouliaris-Hansen procedure described in Hamilton [22] to test for cointegration, i.e. we test the following null hypothesis:

$$H_0\text{: no cointegration.} \quad \text{versus} \quad H_A\text{: cointegration.} \tag{86}$$

The estimated cointegrating regressions are the equations shown in (74), (75), (76), (77), and (78) and can be written as:

[18] I(d) means integrated of order d.

$$Y_t=BX_t+u_t \tag{87}$$

The true process for $Z_t(Yt,Xt)$ under the null hypothesis is:

$$\Delta Y_t=\delta+\Sigma\Psi_s\varepsilon_{t-s} \tag{88}$$

Where δ is a vector of constant numbers with at least one nonzero, Ψ_s is a finite matrix of coefficients. Then we construct the following statistic (this is calculated directly in SAS):

$$Z_\rho\equiv(T-1)(\rho_t^{\cdot}-1)-(1/2)\{(T-1)^2x(\sigma_{\rho T}^{\cdot})^2/s_T^2\}((\lambda'_T)^2-c'_{0,T} \tag{89}$$

Where $\rho_t^{\cdot}$ is the estimate of ρ based on OLS estimation of $u'_t=\rho\, u'_{t-1}+e_t$ and the other parameters are calculated as follows, (Hamilton [22]):

$$s^2=(T-2)^{-1}\Sigma(e'_t)^2 \tag{90}$$

$$c'_j=(T-1)^{-1}\Sigma\, e'_t\, e'_{t-j} \tag{91}$$

$$\lambda'^2= c'_0+2\Sigma^{12}{}_{j=1}[1-(j/13)]\, c'_j \tag{92}$$

If the Z_ρ statistic is below the critical value, i.e. Z_ρ is negative and large in absolute value then we reject the null hypothesis of no cointegration. The calculated Z_ρ is shown in Table 34. For each of the four models the Z_ρ is greater in absolute value than the critical value which is equal at the 5% level to –42.5. Therefore we conclude that the predictors and the subway per capita riders are cointegrated. In the long run there exists an equilibrium relation between those variables.

III.3.6) Discussion of results.

In this chapter we built a complete time series model for the demand for subway travel. We started with a simple model and we included stochastic trends, i.e. seasonal factors, and we also looked at the periods where real shocks happened such as fare increase or subway strike (models 2,3,4,5). We derived from the estimated coefficients the price elasticity for the demand for subway travel. Based on this elasticity we estimated the impact of different policies such as real fare change, real energy price change, improvement in the quality of ride, increase in private employment, and increase in real income per capita on the subway travel and the corresponding revenue change. We showed that there is a long-run relationship between those variables and the subway ridership thereby justifying the time series analysis.

IV) Comparison of the results and conclusion.

The results derived in this dissertation show that there is no unique method to estimate a travel demand model. On the contrary we need more than one data set to completely address all the issues that arise from building a transportation model. The main model was developed by McFadden [3], and it addresses discrete decisions of travel. We estimated this model using three different data sets.

The first data set is a cross-section data set where the unit of analysis is the station. The results derived allow the transportation planner to evaluate and to predict the impact of capital projects, such as opening of new subway stations or new lines. Only a study at the station level would give us a thorough estimation of the ridership at this detail level. Another study done by Bollinger and Ihlanfeldt [4] relates variables such as employment and population to the tract and station level. This is necessary to estimate the direct impact of socioeconomic variables of a defined sample of the population on the ridership. We showed that the demand for subway travel at the station level is statistically sensitive to the characteristics of the area: work related area versus non-work related area. We also showed that on an aggregate level the forecast for subway travel is good. But using only a cross-sectional approach and data from the C.P.1990 does not give us an understanding of consumer behavior; only microlevel data with the unit of analysis being the individual rider would address this issue.

The second data set is compiled from a micro sample survey realized in 1990. Additional variables describing the auto and bus attributes have been added to the main

survey. The results we derived from this data set allow us to better understand modal choice decisions. We estimated the model for people living in the same area and having a car. We also estimated the elasticity of demand for auto travel to the subway station. The results found can be used to make predictions about the choice of mode when the characteristics of households change, (Westin [1]). But this approach has one shortcoming; it does not give us an estimate of the elasticity of subway travel. This can only be accomplished by using time series on the subway travel, the fare, and other variables.

Because most of the data available to transportation planners has been cross-sectional, the use of time-series has not been emphasized in travel demand modeling. Most of the time-series methods used in travel demand analysis have been applications of methods developed for the analysis of aggregate economic models, in which the dependent variable is continuous. The results we derived from this approach are showing us that time series is a good approach when it comes to price elasticity estimations. We quantified the impact of changes in various prices such as real fare, real energy prices, as well as the impact of changes in employment and in the quality of rides, i.e. number of felonies. This estimation will help strategic planners to simulate the impact of macroeconomic variables such as employment on the demand for subway travel.

This last approach could be developed more provided the data set necessary for it would be available to use sequences of discrete choices. Indeed, new contributions in modeling time series of discrete decisions have been developed by Heckman [23] in the area of labor economics. In the transportation field Daganzo and Sheffi [24] have developed an algorithm for applying analysis to a time series of discrete decisions.

Annexe 1

$Y_1, Y_2 \ldots Y_T$	-	Univariate or Multivariate historical time series, subscripts being time indexes and T being the sample size.
X_t	-	Vector of explanatory variables or leading indicators.
Z_t	-	Vector of seasonal variables.
ε_t	-	Serially uncorrelated random shock at time t; it is replaced by the measured forecast error e_t.
$b\ Y_t$	-	backwards time operator, i.e. $b\ Y_t = Y_{t-1}$, $b^n\ Y_t = Y_{t-n}$.
Δ	-	differencing operator, i.e. $\Delta = 1-b$, or $\Delta\ Y_t = Y_t - Y_{t-1}$.
$L(\beta)$	-	a time series model, where β is a n-vector of parameters.
ARIMA(p,d,q)-		Autoregressive Integrated Moving Average process, where p,q,d are small integers, i.e. p is the autoregressive order, q is the moving average order, and d is the degree of differencing.
T	-	Number of historical data points.
n	-	forecast horizon.
p	-	number of periods in a year.
l_t	-	smoothed level at end of time t.
k_t	-	smoothed trend at end of time t.
i_t	-	smoothed seasonal index at end of time t.
α	-	smoothing parameter for level.
β	-	smoothing parameter for trend.
γ	-	smoothing parameter for seasonal indexes.
$\hat{Y}_t(n)$	-	forecast for time t+n from origin t.
j	-	index for the series used, i.e. x-series[1], y-series[2].

TABLE 2

Description of variables used in the station level estimation.

Description of variables	Expected Impact on Ridership	Advantage or Limitation of the variable
Median Household Income	Mixed :Could increase or decrease the ridership	Can derive income elasticity
Per capita Income	Zones with higher PCI would have lower impact on ridership	Isolate the effect of population
Individuals with more than high school diploma, Individuals with less than high school diploma	The higher the level of education the lower the number of trips	Good indicator of modal choice
Immigration	High rates of immigration increase or decrease ridership	Not in Census, Variable is constructed
Employment	Employed people would ride more the subway	Is procyclical with ridership
Population	Correlation with ridership	Highly correlated with ridership and other variables
Evaders	Should correlate negatively with ridership	The counts are done randomly
Jobs	Positively correlated with the ridership	Could be used for building new stations
Riders entering from local bus	Should have big impact at feeder statio	Depends on the fare policy used
Area	Mixed impact	Control for variances in the area
Feeder Stations	Mixed impact	Controls for stations with ridership > population
Exchange Stations	Mixed impact	Controls for high volume stations
Express Stations	Should help registration of express stations	Constructed with MAPTITUDE software.

Source : Census Population 1990, Intercept Survey M.T.A. 1990.

TABLE 5

Mean values for the independent variables.

Name of Variable	Index type All stations	JOB 10	JOB 20	JOB 30	JOB 40	JOB 50	JOB 60	JOB 70	JOB 80	JOB 90	RESD 10	RESD 20	RESD 30	RESD 40	RESD 50	RESD 60	RESD 70	RESD 80	RESD 90
Riders entering from local bus	373	364	423	443	423	237	119	109	149	216	667	210	278	337	427	429	414	392	380
Number of jobs in area i	3,306	3,288	3,384	3,276	3,131	3,106	3,448	3,291	2,810	3,156	3,879	3,053	3,347	3,432	3,385	3,275	3,308	3,347	3,312
Number of employed people in area i	4,412	4,430	4,533	4,564	4,222	4,758	5,382	4,649	5,343	4,480	3,810	4,017	4,204	4,549	4,276	4,200	4,376	4,334	4,409
Population	3,700	3,713	3,768	3,734	3,569	3,582	3,413	3,317	3,413	2,922	3,292	3,482	3,655	3,795	3,747	3,763	3,759	3,724	3,732
Imigration	4,565	4,603	4,835	5,008	4,609	5,106	5,763	4,547	5,371	4,537	3,324	3,685	3,961	4,513	4,352	4,303	4,567	4,498	4,566
Level of education < high school	3,230	3,243	3,289	3,301	3,289	3,681	4,057	3,615	3,994	3,165	2,823	3,037	3,133	3,188	3,053	3,049	3,171	3,166	3,233
Level of education > high school	1,955	1,961	2,027	2,062	1,728	1,983	2,310	1,892	2,264	1,977	1,760	1,719	1,809	2,119	1,944	1,877	1,965	1,929	1,954
TYPE 1 station*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TYPE 2 station																			
TYPE 3 station	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Area in square miles	0 3321	0 3308	0 3293	0 3329	0 3246	0 32	0 3102	0 3012	0 2894	0 3	0 3728	0 3412	0 331	0 3375	0 3369	0 3369	0 3368	0 3356	0 3334
Median household income	$ 25,111 00	$ 25,051 00	$ 25,047 00	$ 25,248 00	$ 24,063 00	$ 23,589 00	$ 23,172 00	$ 23,312 00	$ 24,326 00	$ 21,996 00	$ 27,072 00	$ 25,322 00	$ 24,924 00	$ 25,866 00	$ 25,709 00	$ 25,536 00	$ 25,386 00	$ 25,177 00	$ 25,236 00
Number of evaders																			

*Station types 1, Express station, 2. Feeder station, 3. Exchange station
Source : Census Population 1990 Intercept Survey M.T.A. 1990

TABLE 6

BAMSET Test for Homoscedasticity

	Bartlett M Statistic		
Job stations		Residential stations	
JOB 10	152.18	RESID 10	0*
JOB 20	117.27	RESID 20	27.47
JOB 30	88.41	RESID 30	37.90
JOB 40	64.41	RESID 40	86.81
JOB 50	31.27	RESID 50	125.60
JOB 60	10.29	RESID 60	142.19
JOB 70	3.43	RESID 70	146.40
JOB 80	3.08	RESID 80	150.11
JOB 90	2.2*	RESID 90	153.33

* Not significant at the 10% level.

TABLE 7a

Estimation Results for Subway Ridership by station index

Name of Variable	Estimation Results by Index																	
	All stations	Job 10	Job 20	Job 30	Job 40	Job 50	Job 60	Job 70	Job 80	Job 90	Residential 10	Residential 20	Residential 30	Residential 40	Residential 50	Residential 60	Residential 70	Residential 80	Residential 90
Riders entering from local bus	1 54	1 53	1 77	1 79	1 82	-0 88	1 34	4 69	4 12	.	.	0 11	0 19	1 77	1 65	1 51	1 55	1 6	1 54
	(0 21)	(0 21)	(0 23)	(0 26)	(0 29)	(1 09)	(3 27)	(2 87)	(4 89)			(0 89)	(0 59)	(0 45)	(0 21)	(0 22)	(0 22)	(0 21)	(0 21)
Number of jobs in area i	0 15	0 14	0 11	0 27	0 43	1 07	0 63	0 21	-1 1	.	.	0 68	0 15	-0 001	0 07	0 12	0 14	0 13	0 14
	(0 07)	(0 08)	(0 08)	(0 13)	(0 17)	(0 27)	(0 31)	(0 35)	(1 03)			(0 20)	(0 09)	(0 09)	(0 08)	(0 08)	(0 08)	(0 08)	(0 08)
Number of employed people in area i	0 82	0 75	-0 28	0 04	-0 93	-1 26	-0 88	2 02	-3 52	8 85	.	3 97	3 72	2 61	2 7	2 07	1 16	1 15	1 34
	(0 63)	(0 67)	(0 78)	(0 89)	(1 07)	(0 95)	(0 97)	(1 27)	(5 48)	(6 99)		(0 90)	(0 84)	(0 87)	(0 82)	(0 82)	(0 79)	(0 74)	(0 71)
Population	-0 01	-0 01	-0 01	0 1	0 17	0 12	0 37	1	-3 09	-0 2	.	0 53	0 04	-0 12	-0 16	-0 16	-0 13	-0 1	-0 11
	(0 14)	(0 15)	(0 17)	(0 21)	(0 23)	(0 29)	(0 44)	(0 42)	(3 05)	(1 25)		(0 32)	(0 20)	(0 21)	(0 16)	(0 16)	(0 16)	(0 15)	(0 15)
Immigration	-0 17	-0 37	-0 17	-0 13	-0 01	0 09	-0 1	-0 11	0 69	.	.	-0 59	-0 81	-0 51	-0 42	-0 38	-0 4	-0 38	-0 38
	(0 13)	(0 13)	(0 16)	(0 19)	(0 24)	(0 24)	(0 24)	(0 34)	(1 09)			(0 22)	(0 18)	(0 16)	(0 15)	(0 15)	(0 14)	(0 13)	(0 13)
Level of education < high school	-0 31	-0 24	0 37	0 22	0 84	1 22	0 74	-1 29	4 14	-3 58	0 81	-3	-2 61	-1 75	-1 84	-1 27	-0 56	-0 59	-0 76
	(0 51)	(0 55)	(0 63)	(0 73)	(0 84)	(0 72)	(0 72)	(0 86)	(4 54)	(3 20)	(1 53)	(0 85)	(0 69)	(0 72)	(0 68)	(0 68)	(0 66)	(0 61)	(0 58)
Level of education > high school	1 09	1 2	2 01	1 54	1 92	1 94	1 47	-2 73	7 04	-12 06	-0 01	-2 99	-1 8	-0 12	-0 32	0 2	0 96	0 9	0 72
	(0 61)	(0 65)	(0 75)	(0 90)	(1 11)	(1 08)	(1 21)	(1 86)	(8 43)	(9 25)	(2 04)	(1 08)	(0 85)	(0 83)	(0 76)	(0 77)	(0 74)	(0 69)	(0 66)
TYPE 1 station*	3039	692	5268	46593	35649	20567	532	-56028	36687 9	.	.	-15493 2	4127	4381	-3189	21297	2519	4355	11757
	(489 08)	(506 50)	(603 06)	(685 52)	(806 67)	(896 58)	(1155 15)	(1254 88)	(3347 58)			(853 95)	(636 66)	(649 18)	(581 09)	(553 65)	(530 40)	(507 04)	(489 10)
TYPE 2 station	-1939 13	-1774 46	-3122 14	-27057 5	-26835 9	2634 38	42 75	-10289 9	5321 89	.	.	172 88	8793 5	-993 1	-14663 8	-24022 1	-23948	-23209 9	-22476 5
	(751 71)	(788 53)	(978 07)	(1235 83)	(1356 01)	(2118 45)	(3611 71)	(3285 45)	(6098 73)			(1124 13)	(923 11)	(921 77)	(803 29)	(795 58)	(802 92)	(778 35)	(757 27)
TYPE 3 station	5374 08	5308 15	5082 31	58898 3	54542 1	32915 2	-212 38	31853 7	-5367 89	.	.	50976 8	11985 6	27279 8	36518 9	69225 5	62350 5	59571	59271 7
	(885 72)	(905 43)	(975 95)	(1109 44)	(1348 54)	(1616 87)	(1733 91)	(2434 59)	(7678 84)			(2847 73)	(1711 19)	(1718 06)	(1211 36)	(1061 16)	(1015 26)	(924 31)	(906 15)
Area in square miles	23286	25490 8	25268 1	15043 7	13450 3	32658 6	11517	55765	12168	-14765	.	38764 3	25960 7	35567	38773 7	2634	24522 7	21825 7	22348 8
	(2087 83)	(2174 13)	(2642 98)	(3135 75)	(4127 34)	(4010 71)	(4964 33)	(5109 15)	(12363 13)	(34535 78)		(3208 79)	(2611 19)	(2586 43)	(2371 27)	(2126 42)	(2311 53)	(2162 08)	(2096 15)
Median household income	0 04	0 04	0 06	0 08	0 08	0 24	0 07	0 11	-0 33	0 4	0 04	-0 11	-0 08	-0 05	-0 04	0 01	0 02	0 03	0 03
	(0 03)	(0 03)	(0 03)	(0 04)	(0 04)	(0 05)	(0 07)	(0 06)	(0 44)	(0 34)	(0 20)	(0 08)	(0 05)	(0 04)	(0 04)	(0 03)	(0 04)	(0 03)	(0 03)
Number of evaders	0 26	0 26	0 32	-0 19	-0 51	-0 38	-1 06			0 96	0 05	-1 62	-0 07	0 41	0 16	0 14	0 05	0 11	0 11
	(0 56)	(0 57)	(0 64)	(0 73)	(0 78)	(0 99)	(1 15)			(4 71)	(2 01)	(1 00)	(0 80)	(0 82)	(0 64)	(0 63)	(0 63)	(0 67)	(0 66)
Constant	696 56	650 44	456	-440 25	-1084 41	-5547 87	-4089 81	-4986 43	21145	-4863 72	1254 41	2006 3	2918 10	2831 82	2763 64	1521 51	1633 78	1410 61	1386 14
	(1058 86)	(1085 09)	(1204 00)	(1425 15)	(1679 55)	(1728 01)	(2036 82)	(2187 09)	(25540 90)	(8313 05)	(5460 14)	(2254 05)	(1809 14)	(1388 16)	(1230 04)	(1197 99)	(1188 94)	(1108 54)	(1081 20)
N	234	227	179	135	98	66	42	31	18	9	6	55	98	135	167	191	202	215	224
R^2	0 87	0 87	0 88	0 9	0 7	0 83	0 81	0 77	0 93	0 84	0 76	0 74	0 58	0 79	0 76	0 75	0 74	0 74	0 74

*Station types 1. Express station 2 Feeder station 3 Exchange station
Note standard errors in parentheses

TABLE 7b

Estimation Results for Subway Ridership by station index
Statistically significant variables

Name of Variable	Estimation Results by Index																	
	All stations	Job 10	Job 20	Job 30	Job 40	Job 50	Job 60	Job 70	Job 80	Job 90	Residential 10	Residential 20	Residential 30	Residential 40	Residential 50	Residential 60	Residential 70	Residential 80	Residential 90
Riders entering from local bus	1 54 (0 21)	1 52 (0 21)	1 82 (0 23)	1 8 (0 24)	1 81 (0 26)	·	·	4 41 (1 45)	4 72 (1 32)	·	·	·	·	1 54 (0 38)	1 66 (0 21)	1 52 (0 21)	1 59 (0 22)	1 6 (0 21)	1 68 (0 21)
Number of jobs in area i	0 16 (0 07)	0 14 (0 08)	0 17 (0 07)	0 24 (0 10)	0 4 (0 14)	1 11 (0 22)	0 82 (0 12)	·	·	·	0 65 (0 09)	0 67 (0 18)	0 17 (0 08)	·	·	0 11 (0 08)	0 13 (0 08)	0 14 (0 08)	0 39 (0 06)
Number of employed people in area i	0 49 (0 15)	·	·	·	·	·	·	1 12 (0 89)	1 86 (0 85)	1 76 (0 79)	0 43 (0 09)	3 23 (0 75)	2 59 (0 65)	2 32 (0 24)	2 26 (0 21)	2 08 (0 22)	0 45 (0 18)	0 43 (0 17)	1 7 (0 20)
Population	·	·	·	·	·	·	·	0 84 (0 23)	·	·	·	0 34 (0 18)	·	·	·	·	·	·	·
Imigration	-0 34 (0 12)	·	·	·	·	·	·	·	·	·	·	-0 5 (0 20)	-0 72 (0 17)	-0 48 (0 15)	-0 42 (0 15)	-0 38 (0 14)	-0 4 (0 13)	-0 34 (0 13)	-0 36 (0 13)
Level of education < high school	·	·	·	·	·	0 26 (0 16)	·	-0 72 (0 60)	-0 64 (0 59)	·	·	-2 13 (0 62)	-1 61 (0 50)	-1 51 (0 20)	-1 52 (0 18)	-1 33 (0 19)	·	·	-0 77 (0 13)
Level of education > high school	1 32 (0 20)	1 51 (0 17)	1 35 (0 15)	1 45 (0 17)	1 23 (0 27)	0 75 (0 29)	·	-1 36 (1 11)	-2 33 (1 09)	-2 72 (1 32)	·	-2 14 (0 66)	-0 89 (0 60)	·	·	·	1 54 (0 22)	1 36 (0 21)	·
TYPE 1 station[1]	·	·	·	·	·	·	·	·	·	·	·	-1842 5 (764 49)	·	·	·	·	·	·	·
TYPE 2 station	-2015 (735 92)	-1704 66 (784 37)	-3163 69 (962 98)	-2741 64 (1203 44)	-2588 48 (1295 26)	·	·	·	·	·	·	·	·	·	-1380 85 (784 18)	-2309 3 (779 82)	-2392 01 (789 95)	-2263 38 (755 33)	-1567 95 (772 78)
TYPE 3 station	5468 1 (809 20)	5116 81 (832 01)	5514 61 (852 97)	6318 06 (960 41)	5889 2 (1135 86)	2710 23 (1361 49)	·	4023 96 (1254 91)	·	·	·	4200 28 (2300 13)	·	2861 36 (1648 46)	3439 72 (1143 07)	6970 44 (980 70)	6177 81 (978 54)	5820 4 (858 80)	6149 81 (883 14)
Area in square miles	2730 38 (1807 16)	3841 39 (1651 13)	·	·	·	·	3750 03 (2087 23)	·	·	·	·	3699 67 (2186 00)	·	2983 67 (2368 83)	3843 11 (2232 11)	2860 99 (2184 71)	2637 38 (2164 62)	·	·
Median household income	0 06 (0 02)	0 06 (0 02)	0 1 (0 02)	0 11 (0 02)	0 08 (0 03)	0 27 (0 06)	·	0 09 (0 06)	0 03 (0 03)	0 07 (0 06)	·	·	·	-0 05 (0 04)	-0 04 (0 04)	·	·	0 05 (0 03)	·
Number of evaders	·	·	·	·	·	·	·	·	·	·	·	-1 4 (0 91)	·	·	·	·	·	·	·
Constant	·	·	·	·	·	-5087 19 (1436 11)	·	-1790 53 (1543 52)	·	·	·	·	2529 27 (544 81)	2714 03 (1143 91)	2532 86 (1057 33)	1394 9 (818 62)	1788 11 (819 32)	1377 32 (858 90)	·
N	234	227	179	135	88	65	42	30	18	9	7	56	88	135	187	191	202	215	225
R^2	0 87	0 87	0 86	0 9	0 84	0 81	0 87	0 65	0 9	0 82	0 99	0 92	0 84	0 78	0 78	0 75	0 73	0 74	0 86

[1] Station types 1 Express station 2 Feeder station 3 Exchange station
Note standard errors in parentheses

TABLE 8

Joint and Sum Significance Tests

	Joint Test (X^2)	Sum Test (Standard Normal)
All stations	41.75	6.24
Job stations		
JOB 10	40.50	6.38
JOB 20	38.01	3.03
JOB 30	35.06	2.27
JOB 40	13.76*	1.02*
JOB 50	17.49	2.96
JOB 60	8.34*	2.88
JOB 70	3.84*	1.06*
JOB 80	2.75*	0.91*
JOB 90	0.75*	0.19*
Residential stations		
RESID 10	1.58*	0.16*
RESID 20	8.29*	2.68
RESID 30	8.4*	2.08
RESID 40	32.26	2.39
RESID 50	35.37	4.91
RESID 60	38.40	7.66
RESID 70	37.98	5.74
RESID 80	40.71	5.96
RESID 90	42.66	6.68

* Not significant at the 10% level.

TABLE 9

Actual and Estimated coefficients for the NYCT subway system, 1990

	1990	
System	Actual	Predicted
(233 stations)	1,059,607	1,022,891
Error		3.6%
Standard deviation		2,483

Note: The system is comprised of 233 stations and not 409.

Table 10

Estimation of the subway ridership for a random sample of stations, 1990.

Stations	WKD 90	Forecast 9(	Difference
Livonia Ave.	761	1953	-1192
Cypress Hills	784	4118	-3334
80 St.	2175	2804	-629
Van Siclen Ave.	1673	2499	-826
Beach 44 St.	322	3537	-3215
Beach 90 St.	668	3701	-3033
69 St.	3432	4067	-635
Beach 67 St.	1106	3231	-2125
225 St.	2434	3180	-746
155 St.	1821	1814	7
Seneca Ave.	1405	3380	-1975
Elderts Lane	1689	2854	-1165
Beach 98 St.	280	2834	-2554
Lefferts Blvd.	7192	5399	1793
111 St.	1712	2338	-626
104-102 St.	1361	2677	-1316
Pennsylvania Av	2908	2762	146
Shepherd Ave.	1608	2961	-1353
Beach 36 St.	393	3461	-3068
Forest Pkwy.	1667	5565	-3898
219 St.	1998	3927	-1929
Bay Pkwy.	551	1926	-1375

TABLE 11

Age of Respondent

Age of Respondent	Frequency	Percent
< 17	66,826	1.8%
18-24	553,461	15.3%
25-39	1,710,065	47.3%
40-54	893,321	24.7%
55-64	286,336	7.9%
65 +	102,410	2.8%

TABLE 12

Destination Choice

Destination	Frequency	Percent
Home	1,266,901	35.1%
Work	1,631,645	45.2%
School	203,242	5.6%
Shopping	123,781	3.4%
Recreation	90,461	2.5%
Other	296,413	8.2%

TABLE 13

Number of Trips

Number of Subway Trips	Frequency	Percent
0-5	348,905	9.7%
6-10	1,368,316	37.9%
11-15	1,337,807	37.0%
16-20	392,572	10.9%
21+	164,916	4.6%

TABLE 14

Type of Fare

Type of fare	Frequency	Percent
Regular	3,459,996	95.8%
Senior	77,855	2.2%
Disabled	10,607	0.3%
Student Pass	43,962	1.2%
NYCTA Pass	6,100	0.2%
Free Bus Pass	1,188	0.0%
Other	8,940	0.2%

TABLE 15

Originating Trip

Origin Station	Frequency	Percent
Manhattan CBD	1,540,759	42.7%
Upper Manhattan	470,272	13.0%
Bronx	334,754	9.3%
Queens	514,859	14.3%
Brooklyn	751,872	20.8%

TABLE 16

Gender

Gender of Respondent	Frequency	Percent
Female	1,945,953	53.9%
Male	1,666,485	46.1%

TABLE 17

Ethnicity

Ethnicity	Frequency	Percent
Afro-American	813,795	22.5%
Asian	290,954	8.1%
Hispanic	600,881	16.6%
White	1,768,957	49.0%
Other	137,834	3.8%

TABLE 18

Destination Zone

Destination Zone	Frequency	Percent
Manhattan CBD	1,925,062	53.3%
Upper Manhattan	502,781	13.9%
Bronx	249,267	6.9%
Queens	385,125	10.7%
Brooklyn	550,279	15.2%

TABLE 19

Income

Income	Frequency	Percent
< $ 15,000	744,100	20.6%
$15,000-$25,000	866,262	24.0%
$25,001-$35,000	634,341	17.6%
$35,001-$50,000	548,610	15.2%
$50,001-$75,000	376,045	10.4%
>$75,000	346,565	9.6%

TABLE 20

Mode of Travel to Station

Mode of Travel to Station	Frequency	Percent
Walk	3,015,091	83.5%
Automobile	111,900	3.1%
Ferry	19,747	0.5%
Commuter Railroad	126,917	3.5%
PATH	26,069	0.7%
Taxi/Car Service/Van	41,861	1.2%
Local Bus	178,967	5.0%
Express Bus	23,904	0.7%
Other	61,700	1.7%

TABLE 21

Time of Trip

Time of Trip	Frequency	Percent
AM Peak (5-10 AM)	1,314,329	36.4%
Midday (10 AM- 3 PM)	783,515	21.7%
PM Peak (3-7 PM)	1,232,551	34.1%
Evening (7-10 PM)	282,121	7.8%

TABLE 22

CONDITIONAL LOGIT MODEL OF MODE CHOICE;
Dependent Variable Equals The Log Odds of Choice of Auto Mode;
Binary Logit Maximum Likelihood Estimates:Standards Errors in Parantheses

Independent Variables	Coefficients
Walk time (minutes)	0.1
	(0.06)
Auto in-vehicle time less bus time (minutes)	-0.01
	(0.04)
Auto operating cost less bus fare (dollars)	-1.01
	(0.81)
Income of rider (dollars)	1.02E-05
	-9.60E-06
Age of rider	-0.66
	(0.24)
Gender of rider	0.43
	(0.42)
Ethnicity of rider	0.03
	(0.24)

Summary Statistics

Number of observations=189
Number of cases=39
-2LOG L(0)=187.16
-2LOG L(b)=171.71
R^2=0.3

TABLE 23

Mean values of explanatory variables for auto.

Variables	**Mean for Mode Auto**
Household Income	$47,820
Gender	1
Ethnicity	4
Age	3
Auto in-vehicle time minus bus time[1]	3
Auto operating cost minus fare[2]	7
Walking time	$1
	13

Note 1 : figures are in absolute values and rounded.

TABLE 24

List of explanatory variables included in the conditional logit estimation.

Socioeconomic Variables

Automobile ownership (can also be endogenously predicted)
Household Income
Gender
Ethnicity
Age

Level-of-service variables

Auto in-vehicle time minus bus time
Auto operating cost minus fare
Walking time

Source : Subway Intercept Survey, M.T.A., 1990

TABLE 25

Aggregate elasticities from the maximum likelihood estimates of the conditional logit.

	Conditional Logit
Mode share elasticity with respect to:	**Auto**
Income	0.471
Gender	0.556
Ethnicity	0.096
Age	-2.08
Auto in-vehicle time minus bus time	0.103
Auto operating cost minus fare	0.992
Walking time	1.275

TABLE 26

TIME SERIES MODEL;
Estimation of Subway Ridership and the Fare;
Multiplicative Winters Exponential Smoothing.

Components	Subway Ridership	Fare
Level	92376	1.50930
Trend	33.466	0.00407
Smoothing weight for level	0.19538	0.96824
Smoothing weight for trend	0.00567	0.00492
Smoothing weight for seasonal indexe	0.10201	0.59281
Number of observations	228	228
DW	2.15	1.95
R^2	0.63	0.99

TABLE 27

TIME SERIES MODEL;

Annual Forecast for the Subway Ridership, the Fare, and the Revenue;

Year	Subway Riders[1] (in millions)	Average Fare per ride (dollars)	Subway Revenue (in millions of dollars	Revenue Growth
1999	1,277	$ 0.96	$ 1,230	-
2000	1,341	$ 0.87	$ 1,162	-5.50%
2001	1,408	$ 0.78	$ 1,098	-5.50%
2002	1,478	$ 0.70	$ 1,038	-5.50%

1- The forecast is based on data up to December 1996.

TABLE 28

TIME SERIES MODEL;

Dependent Variable equals The Normalized Log of Subway Riders per Population; Yule-Walker estimates with correction for autocorrelation. Standards Errors in Parantheses

Independent Variables[3]	Coefficients and elasticity	t-statistics
Constant	-3.19	-4.37
	(0.73)	
Log(Real Fare)	-0.29	-6.85
	(0.04)	
Log(Real Energy Price)	-0.14	-4.42
	(0.03)	
Log(Real Income per Capita)	-0.12	-2.07
	(0.06)	
Log(Private Employment)	0.67	7.59
	(0.09)	
Log(Felony)	-0.09	-9.81
	(0.010)	
AR(1) coefficient (RHO)	0.085	1.26
	(0.067)	

Summary Statistics

Number of observations=228
DW=2.16
R^2=0.54

1-Before autocorrelation correction
2-After autocorrelation correction
3-The model included dummy variable for the April 1980 strike.

TABLE 29a

TIME SERIES MODEL;

Dependent Variable equals The Normalized Log of Subway Riders per Population;

Model with seasonal factors;

Yule-Walker estimates.

Standard Errors in Parantheses.

	All Variables		Significant Variables.	
Independent Variables	Coefficients	t-statistic	Coefficients	t-statistic
Constant	-3.72	-4.15	-3.61	-4.29
	(0.90)		(0.85)	
Log(Real Fare)	-0.16	-3.3	-0.17	-3.47
	(0.05)		(0.05)	
Log(Real Energy Price)	-0.13	-3.51	-0.13	-3.72
	(0.04)		(0.04)	
Log(Real Income per Capita)	-0.11	-1.63	-0.12	-1.82
	(0.07)		(0.07)	
Log(Private Employment)	0.46	4.19	0.49	4.82
	(0.11)		(0.10)	
Log(Felony)	-0.06	-5.47	-0.07	-6.23
	(0.11)		(0.01)	
January	0.02	1.71	-	
	(0.01)			
February	-0.05	-3.7	-	
	(0.01)			
March	0.08	5.72	0.05	4.54
	(0.01)		(0.01)	
April	0.002	0.11	-	
	(0.01)			
May	0.06	4.3	0.03	2.6
	(0.01)		(0.01)	
June	0.06	4.14	0.06	5.31
	(0.01)		(0.01)	
July	-0.01	-0.77	-0.04	-4.04
	(0.01)		(0.01)	
August	-0.007	-0.49	-	
	(0.01)			
October	0.09	6.05	0.08	7.71
	(0.01)		(0.01)	
November	0.009	0.64	-	
	(0.01)			
December	0.04	2.5	-	
	(0.01)			
AR(1) coefficient (RHO)	-0.04	-0.55	-0.03	-0.41
	(0.07)		(0.07)	
Number of observations=228	DW^1 =1.92	R^2=0.50	DW^1 =1.94	0.5

1-Before autocorrelation correction
2-After autocorrelation correction

TABLE 29b

TIME SERIES MODEL;

Dependent Variable equals The Normalized Log of Subway Riders per Population;

Model with strike;

Yule-Walker estimates with correction for autocorrelation.

Standards Errors in Parantheses

	All Variables	
Independent Variables	Coefficients	t-statistic
Constant	-3.63	-6.81
	(0.53)	
Log(Real Fare)	-0.26	-8.75
	(0.03)	
Log(Real Energy Price)	-0.108	-4.9
	(0.02)	
Log(Real Income per Capita)	-0.089	-2.1
	(0.04)	
Log(Private Employment)	0.54	8.39
	(0.06)	
Log(Felony)	-0.09	-12.2
	(0.007)	
January	-0.005	-0.53
	(0.009)	
February	0.01	1.18
	(0.009)	
March	0.05	5.55
	(0.009)	
April	0.03	2.91
	(0.009)	
May	0.03	3.21
	(0.009)	
June	0.06	6.58
	(0.009)	
July	-0.04	-4.47
	(0.009)	
August	-0.004	-0.46
	(0.009)	
October	0.09	9.35
	(0.009)	
November	0.007	0.71
	(0.009)	
December	0.002	0.19
	(0.009)	
STRIKE 80	-0.55	-18.02
	(0.03)	
AR(1) coefficient (RHO)	0.02	0.31
	(0.07)	
Number of observations=228	DW^1 =2.04	R^2=0.80

1-Before autocorrelation correction
2-After autocorrelation correction

TABLE 30

TIME SERIES MODEL;

Dependent Variable equals The Normalized Log of Subway Riders per Population;

Model includes 2 months lag of Real Fare;

Yule-Walker estimates.

Standards Errors in Parantheses

	All Variables		Significant Variables.	
Independent Variables	Coefficients	t-statistic	Coefficients	t-statistic
Constant	-2.7	-2.83	-2.5	-2.76
	(0.96)		(0.9)	
Log(Real Income per Capita)	0.09	2.11	0.09	2.15
	(0.04)		(0.04)	
Log(Private Employment)	0.35	3	0.39	3.63
	(0.11)		(0.10)	
Log(Felony)	-0.06	-4.44	-0.06	-5.28
	(0.01)		(0.01)	
January	-0.01	-0.78	-	
	(0.01)			
February	0.010	0.71	-	
	(0.01)			
March	0.06	3.7	0.05	4.86
	(0.01)		(0.01)	
April	0.006	0.4	-	
	(0.01)			
May	0.03	2.19	0.03	2.82
	(0.01)		(0.01)	
June	0.06	3.86	0.06	5.03
	(0.01)		(0.01)	
July	-0.05	-3.32	-0.05	-4.3
	(0.01)		(0.01)	
August	-0.01	-0.77	-	
	(0.01)			
October	0.09	6.18	0.09	7.83
	(0.01)		(0.01)	
November	0.008	0.58	-	
	(0.01)			
December	0.002	0.17	-	
	(0.01)			
2 months Lag of Log of Real Fare	-0.18	-3.44	-0.19	-3.62
	(0.05)		(0.05)	
AR(1) coefficient (RHO)	0.103	-1.5	-0.09	-1.31
	(0.07)		-0.07	
Number of observations=228	DW^1 =1.79	R^2=0.46	DW^1 =1.82	R^2=0.46

1-Before autocorrelation correction
2-After autocorrelation correction

TABLE 31

TIME SERIES MODEL;

Dependent Variable equals The Normalized Log of Subway Riders per Population;

Subway fare interacted with "hikes";

Yule-Walker estimates.

Standards Errors in Parantheses

All Variables

Independent Variables	Coefficients	t-statistic
Constant	-3.55	-3.59
	(0.99)	
Log(Real Income per Capita)	0.003	0.08
	(0.04)	
Log(Private Employment)	0.22	1.82
	(0.11)	
Log(Felony)	-0.05	-3.78
	(0.01)	
January	-0.009	-0.55
	(0.01)	
February	0.01	0.8
	(0.01)	
March	0.05	3.44
	(0.01)	0.34
April	0.005	
	(0.01)	
May	0.03	2.15
	(0.01)	
June	0.06	3.71
	(0.01)	
July	-0.05	-3.2
	(0 01)	
August	-0.01	-0.76
	(0.01)	
October	0.09	6.27
	(0.01)	
November	0.01	0.73
	(0.01)	
December	0.008	0.54
	(0.01)	
Fare*Hike1	-0.005	-0.53
	(0.009)	
Fare*Hike2	-0.003	-0.35
	(0.01)	
Fare*Hike3	0.006	0.63
	(0.01)	
Fare*Hike6	0.003	0.32
	(0.01)	
Fare*Hike7	-0.007	-0.72
	(0.01)	
AR(1) coefficient (RHO)	-0.15	-2.2
	(0.07)	
Number of observations=228	DW^1 =1.70	R^2=0.44

1-Before autocorrelation correction
2-After autocorrelation correction

Table 32
Estimation of policy impact on the yearly subway revenue.

Models used and Policy Implemented	Subway Demand Elasticity with respect to : Real Fare 10% decrease	Real Energy Price 10% increase	Private Employment 1% increase	Number of Felonies 16% decrease	Real Income Per Capita 5% increase
Subway Model	-0.29	-0.14	0.67	-0.09	-0.12
Revenue Impact ($ millions)	$ 36	$ (17)	$ 8	$ 18	$ (7)
Subway Model with Seasonal Factors	-0.17	-0.13	0.49	-0.07	-0.12
Revenue Impact ($ millions)	$ 21	$ (16)	$ 6	$ 14	$ (7)
Subway Model with Seasonal Factors and Strike Effect	-0.26	-0.11	0.54	-0.09	-0.09
Revenue Impact ($ millions)	$ 32	$ (13)	$ 7	$ 18	$ (6)
Subway Model with 2 months lagged Real Fare	-	-	0.35	-0.06	0.09
Revenue Impact($ millions)	NA	NA	$ 4	$ 12	$ (5)
Subway Model with Fare Change interacted with Fare	-	-	0.22	-0.05	0.003
Revenue Impact ($ millions)	NA	NA	$ 3	$ 10	$(0.2)

Assumptions used to calculate the impact on subway revenue:
(a) Subway demand in 1999 is forecasted at 1,277 millions trips.
(b) The average subway fare used is $0.96.

Table 33
Tests for ARCH (1) for each model.

Model	LM Statistic for ARCH(1)	ARCH effects
Subway Model	0.3779	yes
Subway Model with Seasonal Factors	0.0260	yes
Subway Model with Seasonal Factors and dummy for St	0.0258	yes
Subway Model with 2 months lagged Real Fare	0.0300	yes
Subway Model with Fare Change interacted with Fare	0.0134	yes

X^2 is ditributed with 1 df and CI 95% =0.004

Table 34
Phillips-Ouliaris Cointegration Test for each model.

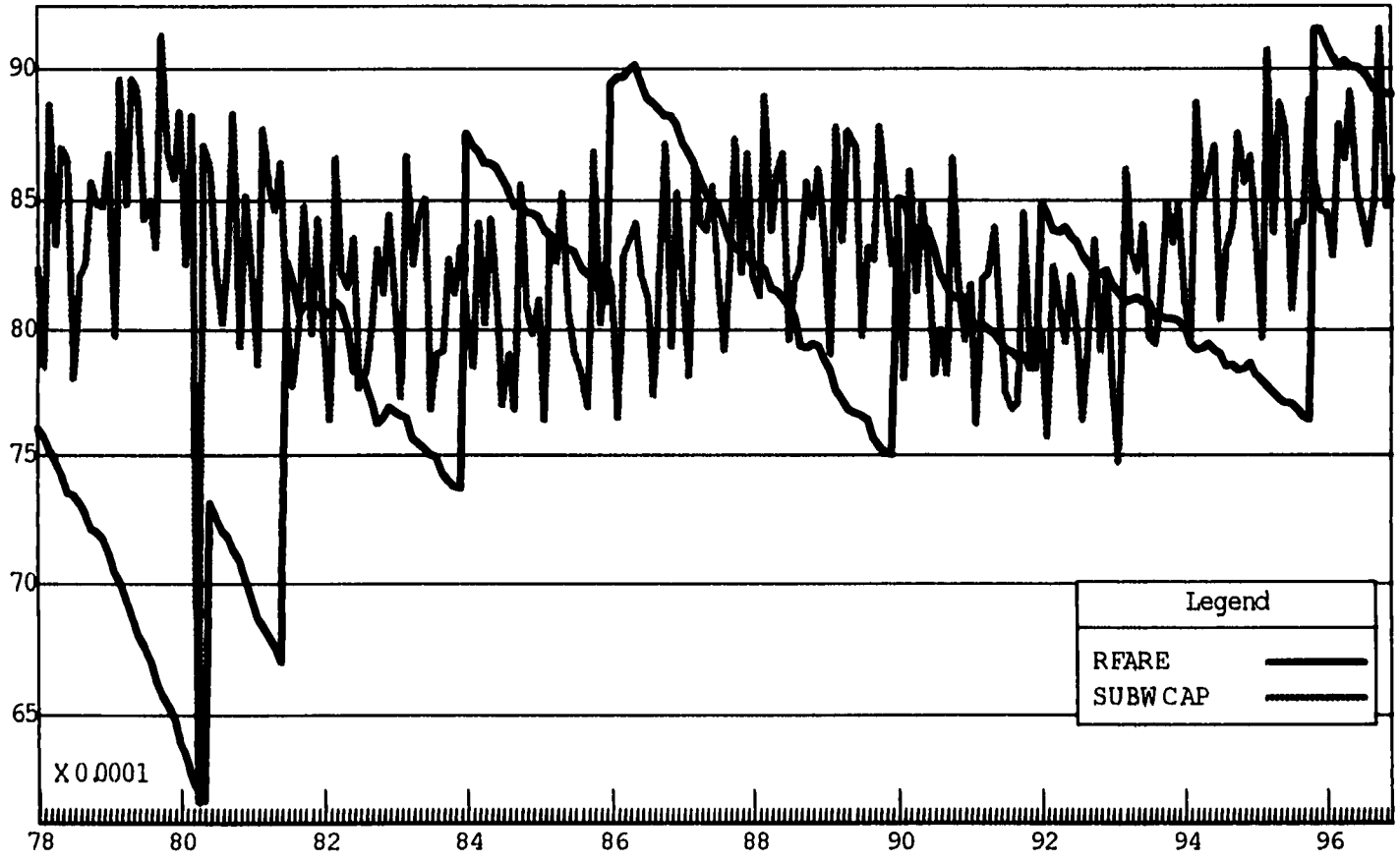
Model	RHO, 1 Lag	Predictors Cointegrated
Critical Value of $Z_{rho\ (1)}$	-42.5	
Subway Model	-245.33	yes
Subway Model with Seasonal Factors	-218.40	yes
Subway Model with Seasonal Factors and dummy for Strike	-232.01	yes
Subway Model with 2 months lagged Real Fare	-203.76	yes
Subway Model with Fare Change interacted with Fare	-189.84	yes

(1) Z_{rho} from table B8 in Hamilton [1994]

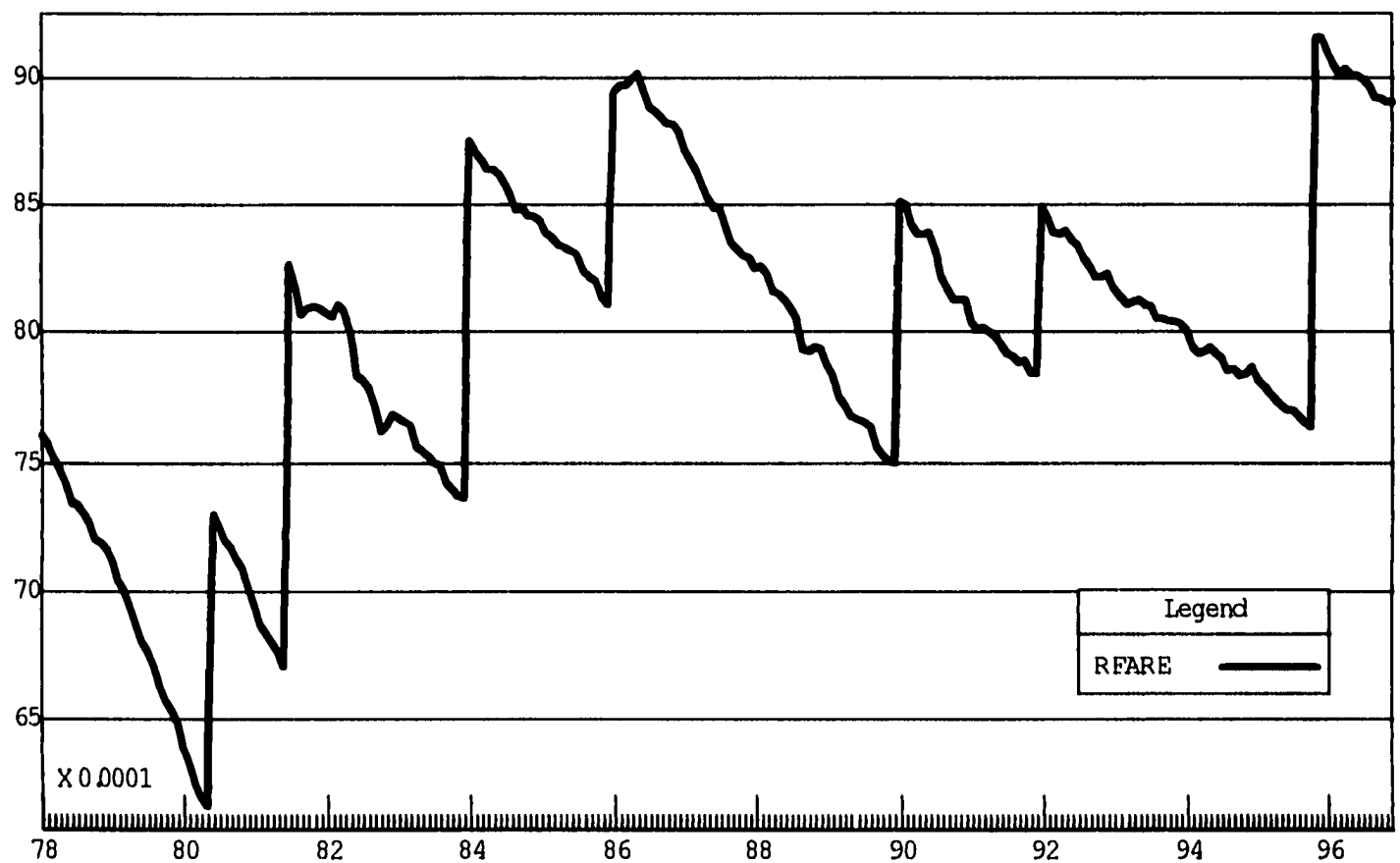
Graph 1

Plots of Real Fare and Number of Subway Riders, 1978-1996.

Units: Real Fare x 0.0001, Subway Riders x 10,000.

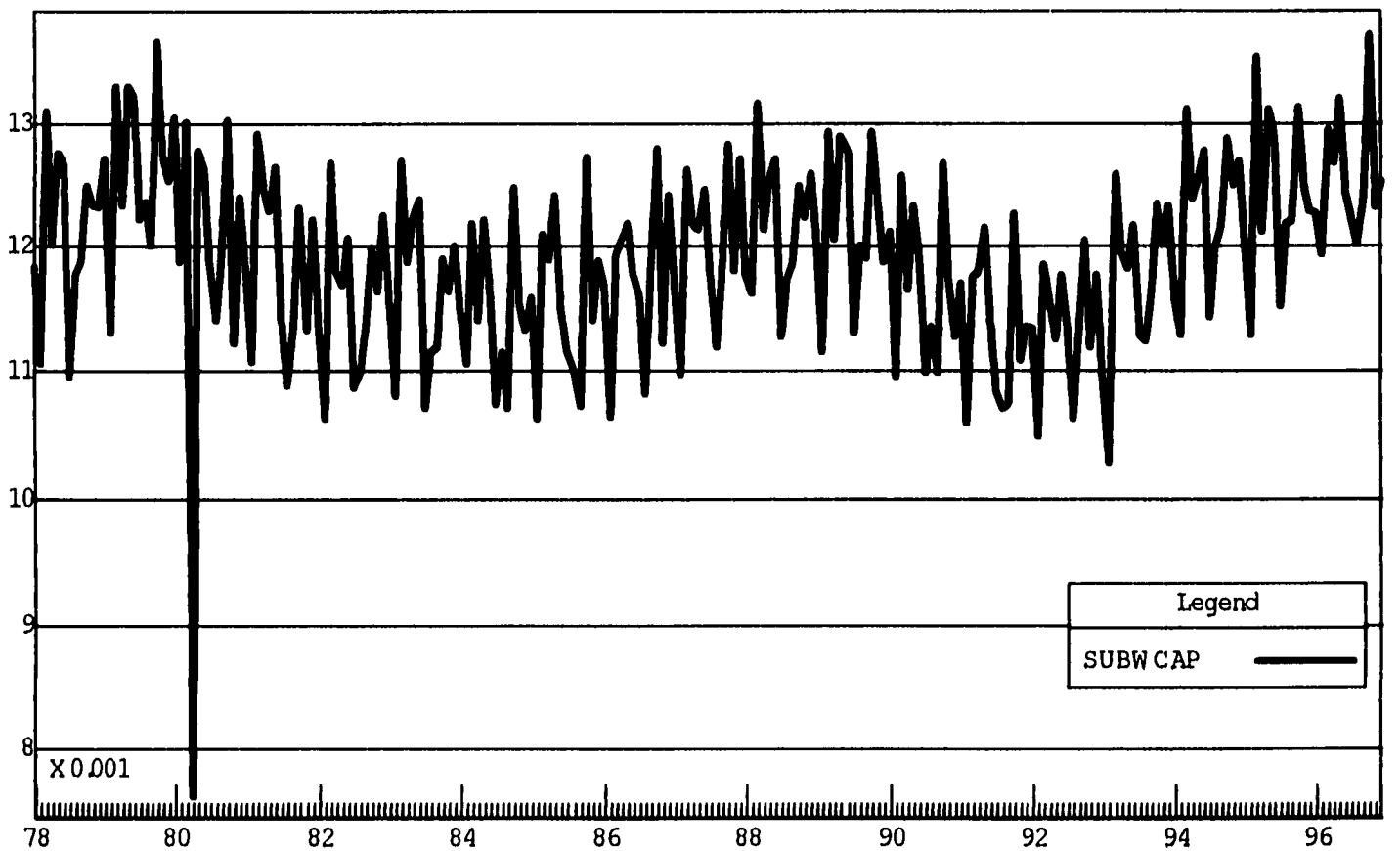


Graph 2
Plot of Real Fare, 1978-1996.

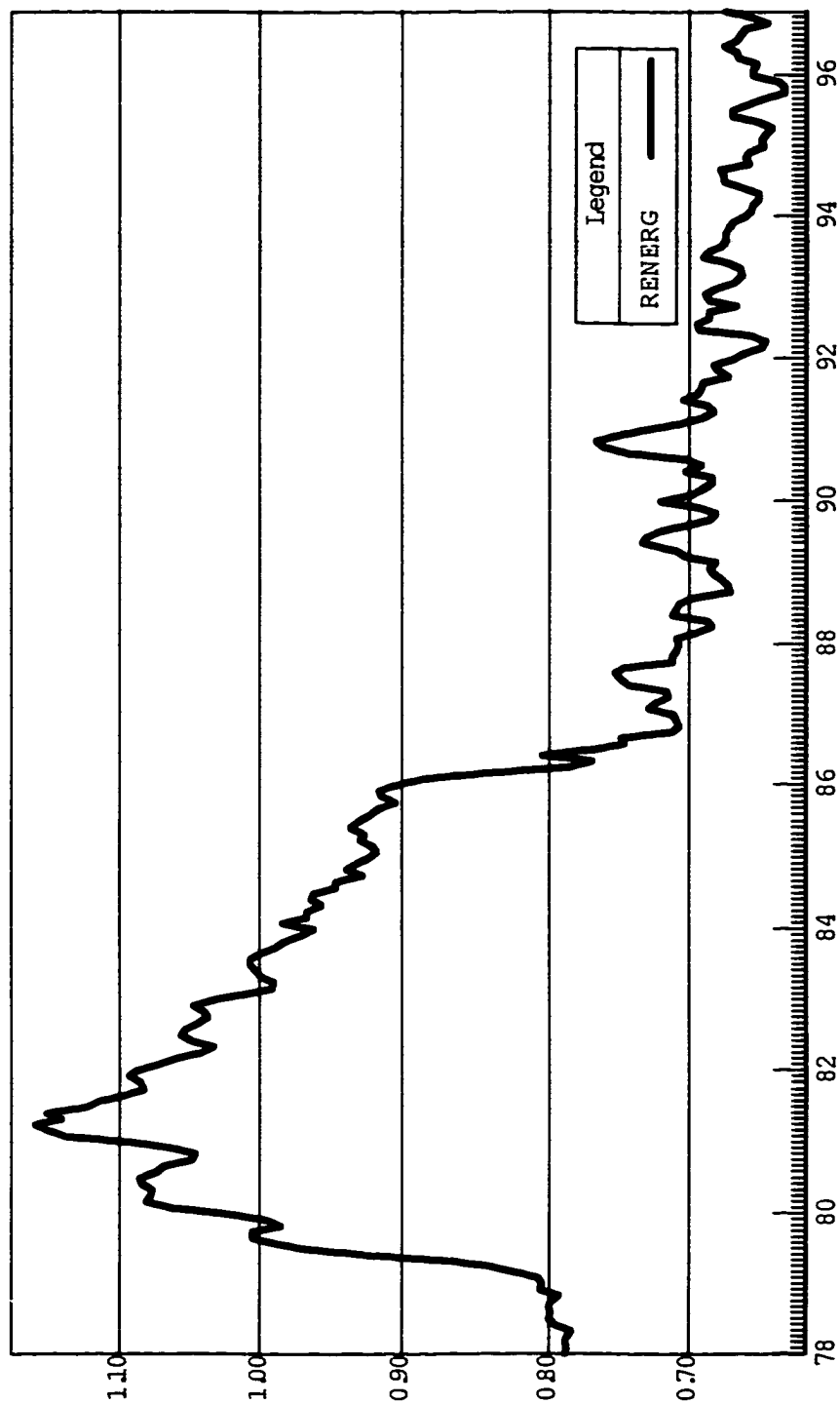


Graph 3

Plot of Number of Subway Riders per Capita, 1978-1996.



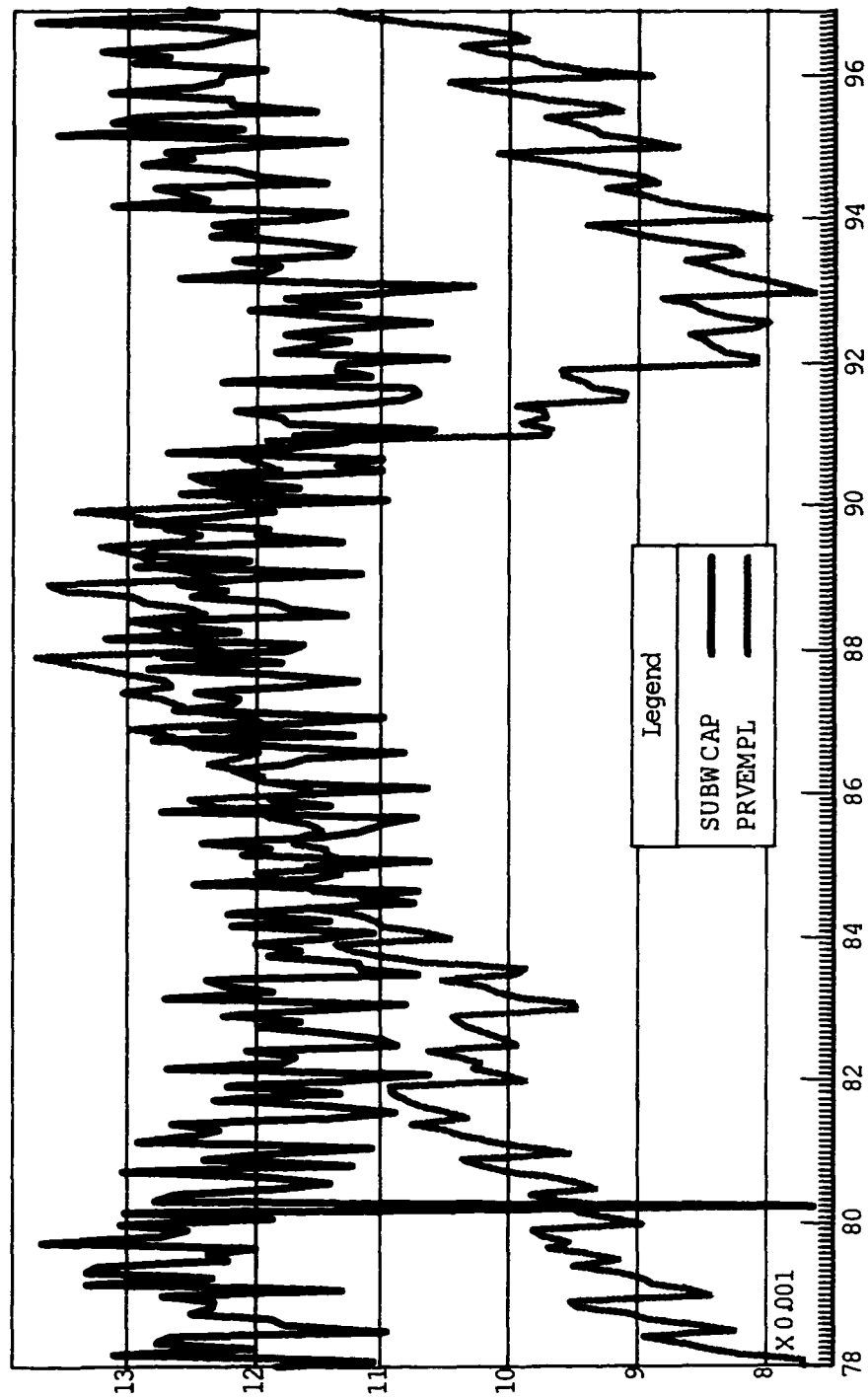
Graph 4
Plot of Real Energy Index, 1978-1996.



Graph 5

Plots of Number of Subway Riders per Capita and Private Employment, 1978-1996.

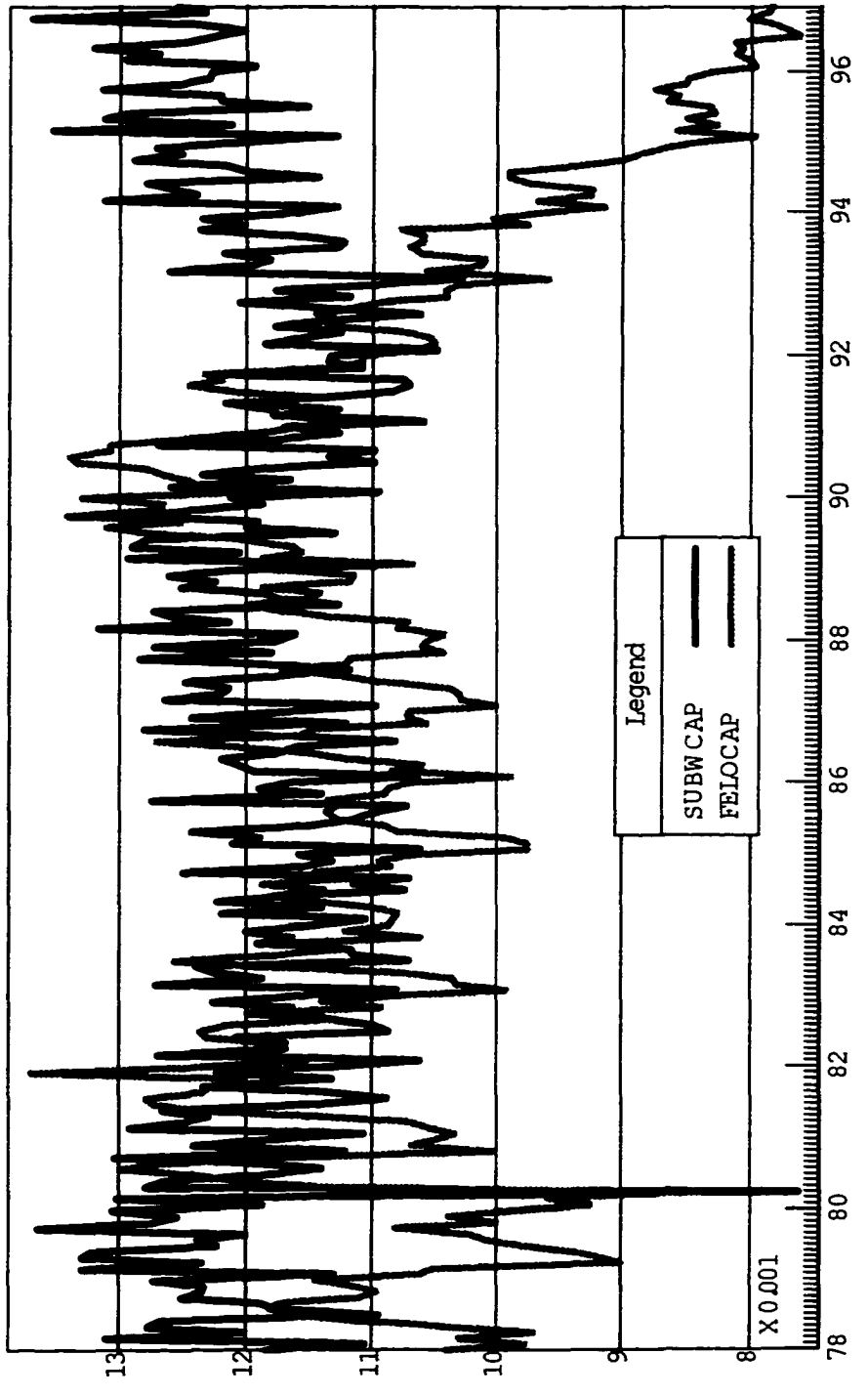
Units: Subway Riders x 0.001, Private Employment January 1978=2649.9



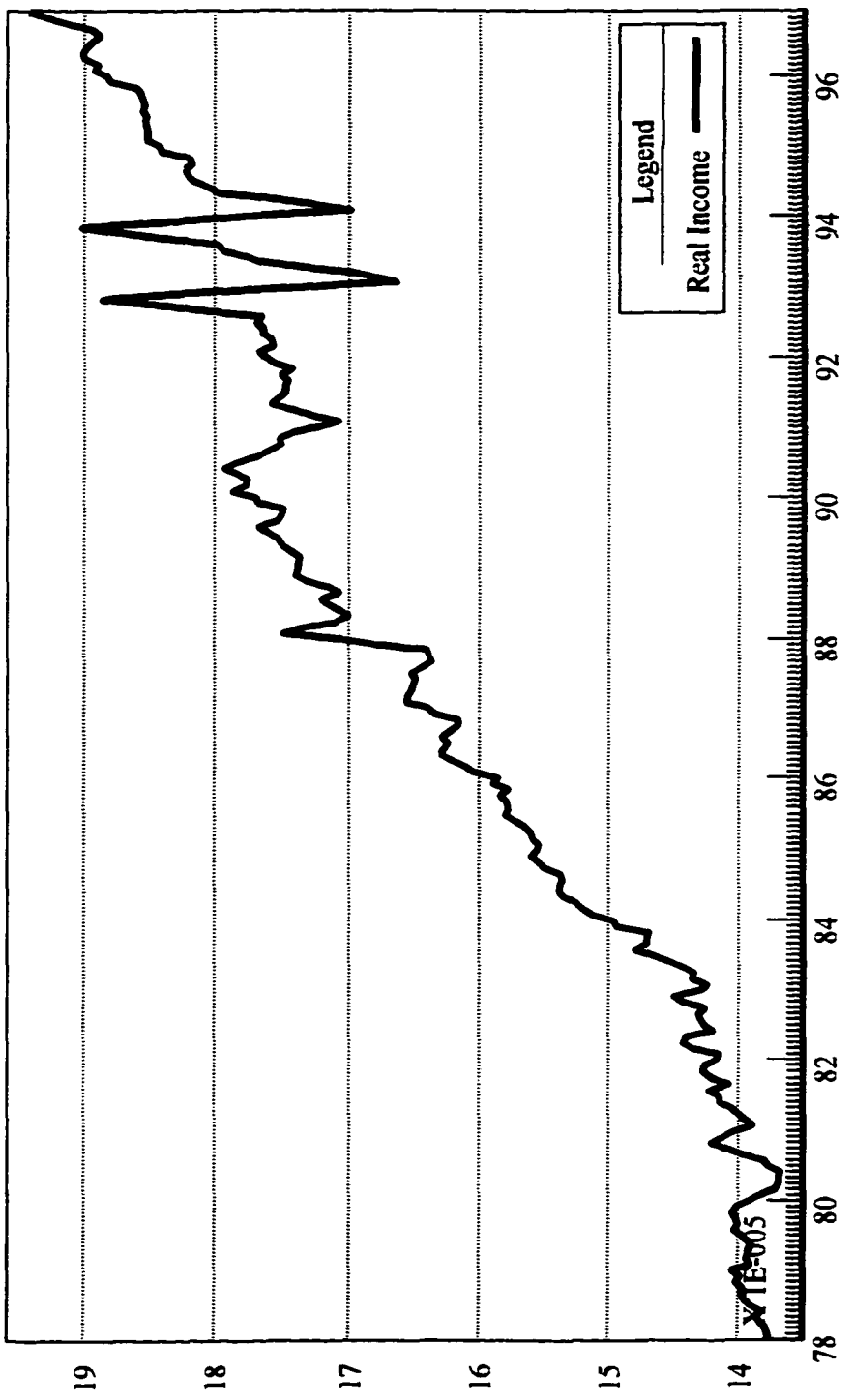
Graph 6

Plots of Number of Subway Riders per capita and Felonies per Capita, 1978-1996.

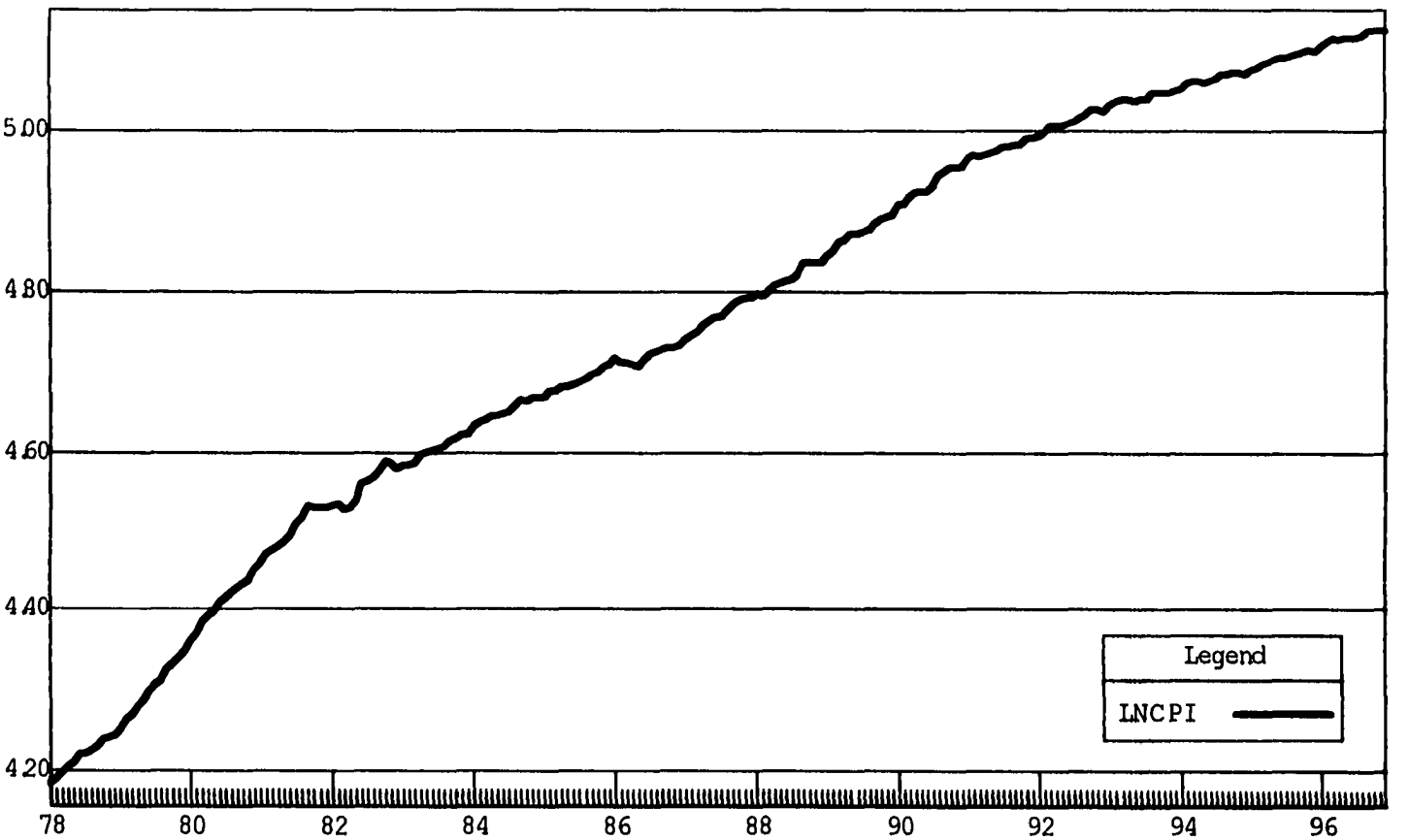
Units: Subway Riders x 0.001, Felonies x 10E-8.



Graph 7
Plots of Real Income per capita, 1978-1996.

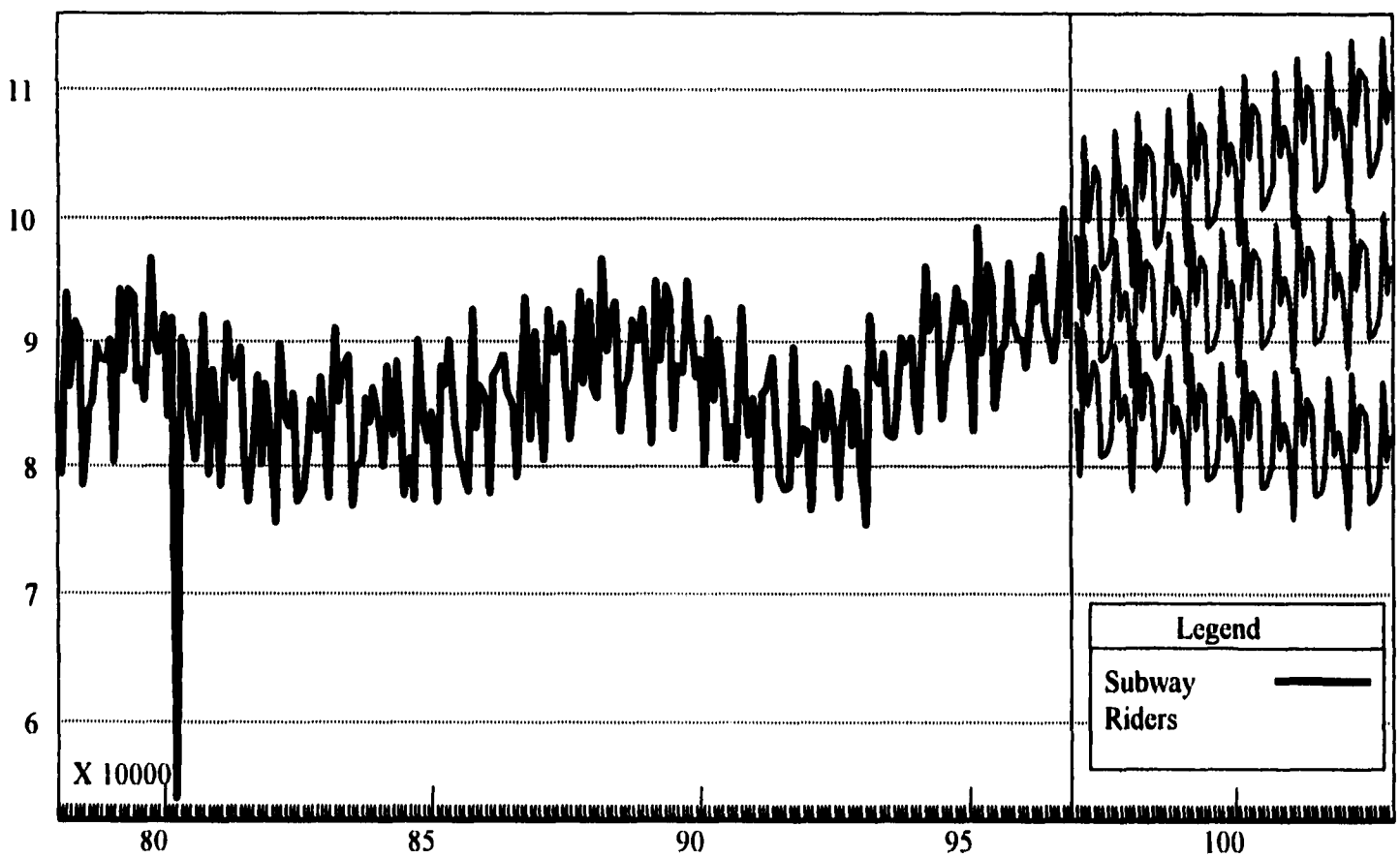


Graph 8
Plots of Log of Consumer Price Index, 1978-1996.

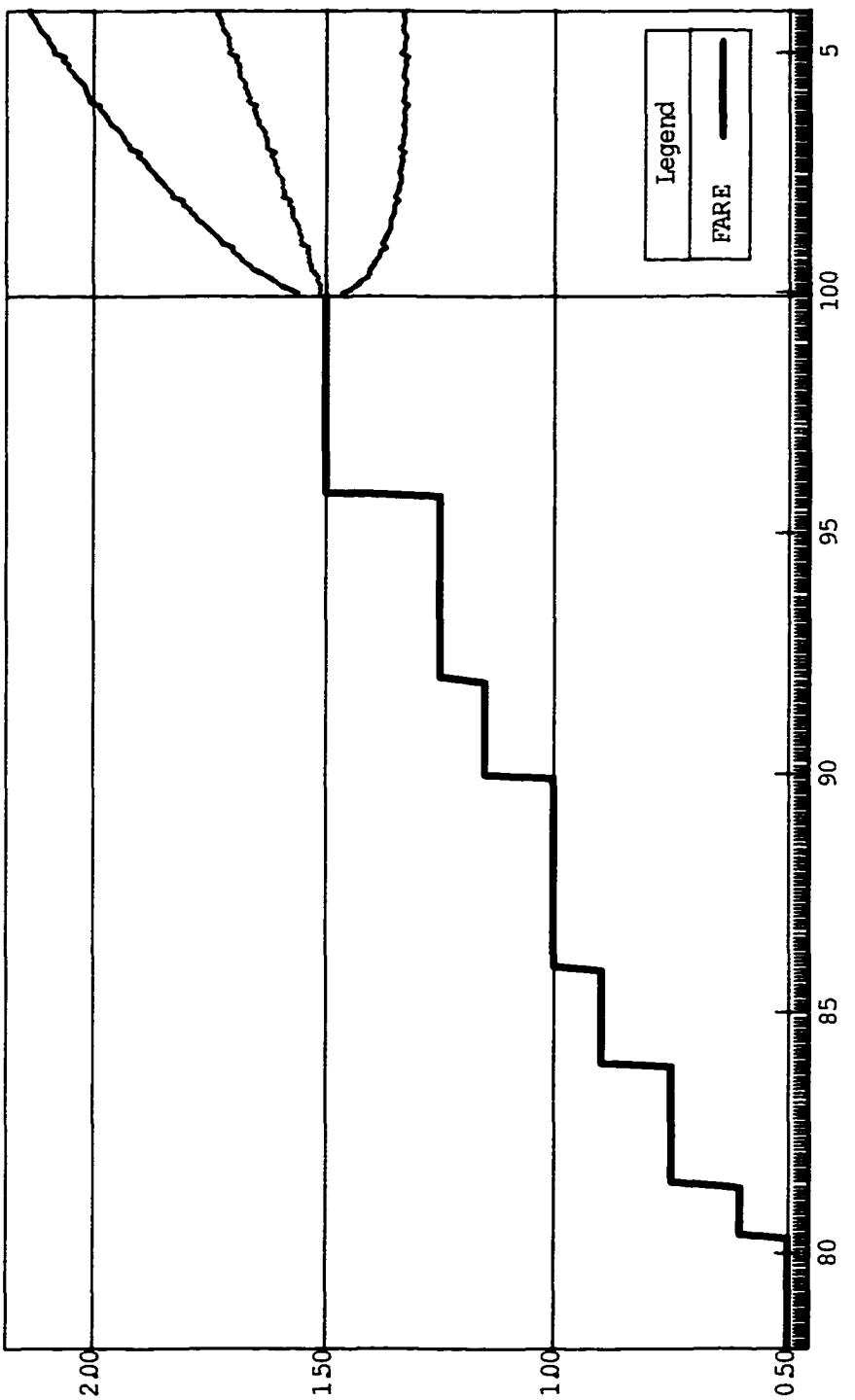


Graph 9

Plot of Number of Subway Riders, Forecast to 2002.



Graph 10
Plot of Fare, Forecast to 2002.



Bibliography

[11] Amemiya, T, 1981, Qualitative Response Models: A survey, Journal of Economic Litterature 19 (Dec), 1483-1536.

[27] Amemiya, T. Advanced Econometrics, Cambridge, Mass: Harvard University Press 1985.

[12] Becker, Gary, A Theory of the allocation of time, The Economic Journal 299, Vol. 75, 1965.

[8] Ben-Akiva, Discrete choice analysis: theory and application to travel demand, MIT Press, 1985.

[28] Ben-Akiva, M., 1972, Structure of travel demand models (Transportation Systems Division, Department of Civil Engineering, MIT, Cambridge, Mass.) unpublished.

[4] Bollinger C. R. and K.R.Ihlanfeldt, 1996, The Impact of Rapid Rail Transit on Economic Development: The case of Atlanta's MARTA.

[35] Cervero R., Landis R., and Landis J., 1995 BART at 20: Land use impacts, Paper presented at 74th Annual Meeting of the Transportation Research Board, Washington, DC.

[19] Cox,D.R. and Snell,E.J. (1989), The Analysis of Binary Data, Second Edition, London: Chapman and Hall.

[29] Daganzo, C, 1979. Multinomial Probit: The Theory and its applications to Demand Forecasting, Academic Press, NY.

[24] Daganzo, C and Sheffi Y, 1982. Multinomial Probit Models with Time Series Data: Unifying State Dependence and Serial Correlation Models, Environnment and Planning A14: 1377-1388.

[37] Greene, W., 1996. Heteroskedastic Extreme Value Model for Discrete Choice. New York University.

[25] Greene, W. H., 1997. Econometric Analysis, 3rd edn. Macmillan, NY.

[26] Granger, C. W. J. and P. Newbold 1986. Forecasting Economic Time Series, 2nd edn. Academic Press, London.

[22] Hamilton,J.D, 1994. Time Series Analysis, Princeton University Press.

[23] Heckman, J., 1981, Statistical Analysis of Discrete Panel Data. In Structural Analysis of Discrete Data with Econometric Applications. C. Manski and D. McFadden, edn. MIT Press, Cambridge, Mass.

[36] Hensher D.A. and King J., 1998, Establishing fare elasticity regimes for urban passenger transport: time-based fares for concession and non-concession markets segmented by trip length, Journal of Transportation and Statistics, volume 1, 43-61.

[18] Hosmer,D.W. and Lemeshow,S., 1989, Applied Logistic Regression, New York: John Wiley and Sons, Inc.

[34] Kain J., 1997, Cost-effective alternatives to Atlanta's costly rail rapid transit system, Journal of Transportation Economics and Policy, 31, 25-49.

[7] Kanafani, A. K., 1983. Transportation Demand Analysis.

[2] Kennedy, P. A guide to econometrics.Cambridge, The MIT Press 1998.

[9] Maddala, G. S. Limited Dependent Variables and Qualitative variables in Econometrics, Cambridge University Press 1994.

[20] Magee,L., 1990, R^2 Measures Based on Wald and Likelihood Ratio Joint Significance Tests, American Statistician,44 250-253.

[10] Manski, C, 1975, Maximum Score Estimation of the Stochastic Utility Model of Choice, Journal of Econometrics 3: 205-228.

[13] Manski, C, and Lerman, S, 1977, The Estimation of Choice Probabilities from Choice-Based Samples, Econometrica 45: 1977-1988.

[3] McFadden, D, 1974, Conditional Logit Analysis of Qualitative Choice Behavior. Frontiers in Econometrics, P Zarembka, Ed. Academic Press, NY, p105-142.

[14] McFadden, D., 1977, Quantitative methods for analyzing travel behavior of individuals: Some recent developments, Cowles Foundation Discussion Paper No.474.

[15] McFadden, D., 1973b, Travel demand forecasting study, BART Impact Study Final Report Series (Institute of Urban and Regional development, University of California, Berkeley, Calif.) unpublished.

[5] McFadden, D., 1974, The measurement of urban travel demand, (Department of Economics, University of California, Berkeley, Calif.) unpublished.

[30] McFadden, D. and F. Reid, 1974, Aggregate travel demand forecasting from disaggregated behavioral models (Department of Economics, University of California, Berkley, Calif.) unpublished.

[31] McFadden, D. 1982. Econometric Analysis of Qualitative Response Models, Working Paper, Department of Economics, MIT, Cambridge, Mass.

[32] Metropolitan Transportation Authority, Office of Management and Budget, Division Revenue, Revenue Reports 1990-1998.

[33] New York City Independent Budget Office, New York City Transit's Fiscal Condition, August 1999.

[21] Nagelkerke,N.J.D. (1991), "A Note on a General Definition of the Coefficient of Determination" Biometrika, 78, 691-692.

[16] Oi, W. and P. Shuldiner, 1962, An analysis of urban travel demands, (Northwestern University Press, Evanston, Ill.).

[17] Ramsey, J.B., 1969, Classical mode selection through specification error tests, Journal of the Royal Statistical Society, Series B.

[1] Westin, R.B., 1973, Predictions from binary choice models, The Journal of Econometrics:1-16.

[6] Wohl, M., and B.V. Martin. 1967. Traffic Systems Analysis for Engineers and Planners, McGraw-Hill, New York.

Presented to the Office of Management and Budget, 1996, Metropolitan Transportation Authority, New York.

Presented to the University Transportation Research Center, 1996, Region II, The City College, New York.