

A Utility-Theory-Consistent System-of-Demand-Equations Approach
to Household Travel Choice

by

Kara Maria Kockelman

B.S. (University of California, Berkeley) 1991
M.S. (University of California, Berkeley) 1996
M.C.P. (University of California, Berkeley) 1996

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy
in

Engineering – Civil and Environmental Engineering

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA, BERKELEY

Committee in charge:

Professor Mark Hansen, Chair
Professor Daniel McFadden
Professor Martin Wachs

Fall 1998

The dissertation of Kara Maria Kockelman is approved:

Chair

Date

Date

Date

University of California, Berkeley

Fall 1998

A Utility-Theory-Consistent System-of-Demand-Equations Approach
to Household Travel Choice

Copyright 1998

by

Kara Maria Kockelman

Abstract

A Utility-Theory-Consistent System-of-Demand-Equations Approach
to Household Travel Choice

by

Kara Maria Kockelman

Doctor of Philosophy in Civil and Environmental Engineering

University of California, Berkeley

Professor Mark Hansen, Chair

Modeling personal travel behavior is complex, particularly when one tries to adhere closely to actual causal mechanisms while predicting human response to changes in the transport environment. There has long been a need for explicitly modeling the underlying determinant of travel – the demand for participation in out-of-home activities; and progress is being made in this area, primarily through discrete-choice models coupled with continuous-duration choices. However, these models tend to be restricted in size and conditional on a wide variety of other choices that could be modeled more endogenously.

This dissertation derives a system of demands for activity participation and other travel-related goods that is rigorously linked to theories of utility maximization. Two difficulties inherent in the modeling of travel – the discrete nature of many travel-related demands and the formal recognition of a time budget, not just a financial one – are dealt

with explicitly. The dissertation then empirically evaluates several such demand systems, based on flexible specifications of indirect utility. The results provide estimates of activity generation and distribution and of economic parameters such as demand elasticities. Several hypotheses regarding travel behavior are tested, and estimates are made of welfare effects generated by changes in the travel environment.

The models presented here can be extended to encompass more disaggregate consumption bundles and stronger linkages between consumption of out-of-home activities and other goods. The flexibility and strong behavioral basis of the approach make it a promising new direction for travel demand modeling.

Dedication

This work is dedicated to my family and to my best friend, Steven Glenn Rosen.

Their constant support has made an extraordinary difference.

TABLE OF CONTENTS

List of Tables	vii
List of Symbols	ix
Acknowledgments	xiv
CHAPTER ONE: INTRODUCTION	1
CHAPTER TWO: REVIEW OF RELATED LITERATURE	6
General	6
Models of Trip Generation	6
Systems of Equations	7
Hybrid/Simulation Models	9
Discrete- and Discrete + Continuous-Choice Models	10
Value-of-Time Models	13
Summary of Related Literature	15
CHAPTER THREE: RESEARCH METHODOLOGY	18
Microeconomic Foundations	18
Roy's Identity in a Two-Budget Framework	22
Theory-Implied Constraints	27
Non-Negativity	28
Concavity of the Expenditure Function	30
Homogeneity	32
Summability	33
Separability	33
Symmetry	36
Validity of Utility Maximization Hypothesis	37
Estimating Benefits and Costs	38
Functional Specification	40
Model Specifications	42
Type 1 Model Specification: Modified Linear Expenditure System	43
Type 2 Model Specification: Modified Translog	44

Type 3 Model Specification: Modified Translog with Constants	45
Type 4 Model Specification: Modified Translog with Constants, using Wage and Total Time Data	46
Statistical Specification	47
Integer Demand Observations and the Poisson Assumption	47
Generalizing the Poisson Assumption through Use of a Negative Binomial	49
Implication of the Assumption of Multiplicative Error Component for Indirect Utility	55
Data Set	57
Definition of the Consumption Space	58
Trip Chaining	63
CHAPTER FOUR: EMPIRICAL ESTIMATION AND MODEL VALIDATION	71
Estimation Techniques	71
Likelihood Maximization	71
Acquiring Starting Parameter Values	71
Variance-Covariance Estimation	72
Results to be Estimated	73
Results of Type One Model: Modified Linear Expenditure System	73
Results of Type Two Model: Modified Translog	76
Economic Implications of the Type Two Model Results	79
Results of Type 3 Model: Modified Translog with Constants	81
I. Discretionary Activity Participation	81
Model Comparisons: Case Example	85
II. Modeling Tours Explicitly	88
Results of Type 4 Model: Modified Translog with Constants, Using Wage and Total Time Data	91
Comparison of All Model's Elasticity Estimates	94
Discussion of Value of Time Estimates	96
A Functional Conflict between Behavioral Indicators	97
Identification of All Income/Wage Terms in Indirect Utility	98
Other Reasons for Incorrect Marginal Utility of Income Estimates	102
Further Qualifications	103
CHAPTER FIVE: HYPOTHESIS TESTING AND COST-BENEFIT ESTIMATION	110
Hypothesis Testing using the Type Two Model	110
Hypothesis 1	111
Hypothesis 2	114
Hypothesis 3	115
Hypothesis 4	117

	vi
Hypothesis 5	118
Cost-Benefit Analysis using the Type Two Model	119
CHAPTER SIX: LIMITATIONS AND EXTENSIONS	127
Data Deficiencies	127
Modeling Expanded Choice Sets	128
Inclusion of Automobile Ownership in the Model	129
Location Choice Decision	130
Modeling Activity-Participation Times	132
Incorporating Different Preference Structures	132
Recognition of Other Constraints on Behavior	133
CHAPTER SEVEN: CONCLUSIONS	137
REFERENCES	140
APPENDIX	149
A-1: List of Possible Functional Form Specifications	150
A-2: Derivation of Roy's Identity	155
A-3: Description of Data Set Used	157
A-4: Description of Negative Binomial Distribution	160
A-5: Examples of the Estimation Algorithms, as programmed in SPlus3.3	161
A-6: Example Algorithm for Travel-Time Cost Calculations, as programmed in Matlab	165
A-7: Estimating the Variability in the Results	166
INDEX	165

List of Tables

Table 3-1: Regressions of Travel Time and Distance on Measures of Urban Form

Table 3-2: Description of Variables Used

Table 4-1a: Parameter Estimates of Modified Linear Expenditure System, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours

Table 4-1b: Economic Implications of Modified Linear Expenditure System, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours

Table 4-2a: Parameter Estimates of Modified Translog Model, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours

Table 4-2b: Economic Implications of Modified Translog Model, as applied to Discretionary Activity Participation across Four All-Jobs Iso-Opportunity Contours

Table 4-3a: Parameter Estimates of Modified Translog Model with Intercept Terms, as applied to Discretionary Activity Participation across Four All-Jobs Iso-Opportunity Contours

Table 4-3b: Economic Implications of Modified Translog Model with Intercept Terms, as applied to Discretionary Activity Participation across Four All-Jobs Iso-Opportunity Contours

Table 4-3c: Comparison of Log-Likelihood Values across Different Models, based on the Modified Translog Model with Intercepts

Table 4-4a: Parameter Estimates of Modified Translog Model with Intercept Terms, as applied to Trip Tours to Four All-Jobs Iso-Opportunity Contours

Table 4-4b: Economic Implications of Modified Translog Model with Intercept Terms, as applied to Trip Tours to Four All-Jobs Iso-Opportunity Contours

Table 4-5a: Parameter Estimates of Modified Translog Model with Intercept Terms and Wage and Total-Time Information, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours

Table 4-5b: Economic Implications of Modified Translog Model with Intercept Terms and Wage and Total-Time Information, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours

Table 4-6: Summary of Elasticity Estimates: Median Values across Households

Table 4-7: Summary of Value-of-Time Estimates: Quartiles across Households

Table 5-1a: Quartiles of the Estimates of the Derivative of Total Travel Time for Participation in Discretionary Activities with respect to Activity Access/Travel Times

Table 5-1b: Quartiles of the Standard Error Estimates of the Derivative of Total Travel Time for Participation in Discretionary Activities with respect to Activity Access/Travel Times

- Table 5-1c: Quartiles of the T-Statistics of the Derivative of Total Travel Time for Participation in Discretionary Activities with respect to Activity Access/Travel Times
- Table 5-2a: Quartiles of the Estimates of the Derivative of Total Discretionary-Activity Participation with respect to Income
- Table 5-2b: Quartiles of the T-Statistics of the Derivative of Total Discretionary-Activity Participation with respect to Income
- Table 5-3a: Quartiles of the Estimates of the Derivative of Total Activity Participation in Discretionary Activities with respect to Activity Access/Travel Times
- Table 5-3b: Quartiles of the T-Statistics of the Derivative of Total Activity Participation in Discretionary Activities with respect to Activity Access/Travel Times
- Table 5-4a: Quartiles of the Estimates of the Derivative of Indirect Utility with respect to Discretionary Time
- Table 5-4b: Quartiles of the T-Statistics of the Derivative of Indirect Utility with respect to Discretionary Time
- Table 5-5a: Quartiles of the Estimates of the Derivative of Indirect Utility with respect to Income
- Table 5-5b: Quartiles of the T-Statistics of the Derivative of Indirect Utility with respect to Income
- Table 5-6a: Quartiles of Indirect Utility Estimates
- Table 5-6b: Quartiles of Indirect Utility Changes Following 50% Increase in Travel Times
- Table 5-6c: Quartiles of Equivalent Variation, in Money and Time Units
- Table 5-6d: Median Travel Times and Low, Median, and High Budget Levels for Sample
- Table 5-6e: Comparison of Equivalent Variation Changes for Specific Household Types
- Table 5-6f: Comparison of Time Valuation for Specific Household Types

List of Symbols

The notation for variables used in this dissertation is provided here, in order of appearance:

Chapter 2:

v	= Indirect utility
Y	= Income or total monetary expenditures
VMT_t	= Vehicle miles traveled in period t
P_{VMT}	= Price per mile of vehicle miles traveled in period t

Chapter 3:

\bar{A}	= Vector of number of different activities participated in per period
\bar{T}	= Vector of time spent participating in different activities per period
\bar{t}	= Vector of access/travel times to different activities
\bar{Z}	= Vector of number of non-activity goods consumed
\bar{P}_A	= Vector of prices for participation in different activities
\bar{P}_{trvl}	= Vector of travel costs to different activities
\bar{P}_Z	= Vector of prices for non-activity goods
w	= Wage, in money per unit time
Y_{un}	= Unearned income per period
T_i	= Time spent participating in activity i per period
T_w	= Time spent participating in the work activity per period
H	= Total time budget (Amount of time available for expenditure in different activities and in travel, per period)
Y	= Income available, per period

- $v(\cdot)$ = Indirect utility function (*i.e.*, maximized utility)
- A_i^* = Optimal consumption of activity i per period
- T = Time budget for set of activities considered, per period
- L = Lagrangian function
- L^{opt} = Optimized Lagrangian function
- u = Direct utility function
- λ_{Time} = Lagrange multiplier on time constraint, in units of utility per unit time
- λ_{Money} = Lagrange multiplier on money-budget constraint, in units of utility per unit money
- \bar{A}^* = Vector of optimal consumption of activities per period
- \bar{T}^* = Vector of optimal activity participation times per period
- \bar{Z}^* = Vector of optimal consumption of other goods per period
- T_w^* = Optimal work-activity participation time per period
- $e_H(\bar{t}, Y_{un}, w, u)$ = Minimum time budget needed to achieve utility u at travel times \bar{t} ,
 unearned income Y_{un} , and wage w
- T_d = Discretionary time available
- \bar{X}_{Food} = Vector of number of foods consumed
- \bar{X}_{Other} = Vector of number of non-food goods consumed
- \bar{P}_{Food} = Vector of prices of foods
- \bar{P}_{Other} = Vector of prices of non-food goods
- Y_{Food} = Total monetary expenditures of foods
- $e_S(\bar{P}, u)$ = Minimum monetary expenditure to achieve utility u at prices \bar{P}
- $h_{i,S}(\bar{P}, u)$ = Money-compensated Hicksian demand for good i per period
- $e_T(\bar{t}, \bar{P}, Y, u)$ = Minimum time budget needed to achieve utility u at travel times \bar{t} ,

prices \bar{P} , and income Y

- $T_{i,T}(\cdot)$ = Time-compensated demand for participation time in activity i per period
- $\bar{A}_T(\cdot)$ = Time-compensated vector of demands for number of activities per period
- EV_S = Equivalent variation, as measured in money units
- EV_T = Equivalent variation, as measured in time units
- $e_S(\bar{t}, T, u')$ = Minimum money budget needed to achieve utility u' at travel times \bar{t}
and available time T
- $e_T(\bar{t}, Y, u')$ = Minimum time budget needed to achieve utility u' at travel times \bar{t}
and available money budget Y
- $h_{i,S}(\bar{t}, T, u')$ = Money-compensated Hicksian demand for good i per period
- $h_{i,T}(\bar{t}, Y, u')$ = Time-compensated Hicksian demand for good i per period
- X_i^* = Optimal consumption of activity i per period
- $\alpha, \alpha_i, \beta_{ij}, \gamma_{iY}, \gamma_{iT}, \gamma_{TY}, \gamma_{iH}, \mu_o, \mu_i$
= Unknown parameters of the indirect utility function, to be estimated
- X_i = Observed integer level of demand of good type i
- \bar{X} = Vector of observed integer demand levels
- λ_i = Mean optimal rate of demand for good type i per time period
- $\bar{\epsilon}$ = Vector of unobserved variation, characterizing an observation's vector of
optimal demands, relative to the population mean
- X_T = Observed total of integer demands
- λ_T = Mean total optimal rate of demand across good types, per period
- m = Size parameter characterizing the negative binomial distribution
- p^* = Probability parameter characterizing the negative binomial distribution

\bar{p}	= Vector of probabilities characterizing the multinomial distribution
$\Gamma(\cdot)$	= Gamma function (defined in Appendix, section A-4)
ε	= Gamma error term characterizing individual observation's optimal demands, relative to population mean
P	= Odds-ratio parameter characterizing the negative binomial distribution
$E(\cdot)$	= Mean or expectation operator
$V(\cdot)$	= Variance operator
$COV(\cdot)$	= Covariance operator
$X_{Autos} = X_A$	= Observed automobile ownership
P_A	= Odds-ratio parameter characterizing the negative binomial distribution for automobile
P_T	= Odds-ratio parameter characterizing the negative binomial distribution for total activity participation
p_A	= Probability parameter characterizing demand for automobiles, relative to total activity plus automobile demand
X_A^*	= Optimal automobile ownership level
X_T^*	= Optimal rate of total activity participation
\bar{L}	= Vector of zeros for non-chosen locations and a one for chosen location
\bar{H}	= Vector of levels of housing attributes consumed
\bar{P}_L	= Vector of prices of housing locations
\bar{P}_H	= Vector of prices of different housing attributes
$\bar{t}_{obs'd}$	= Vector of travel times observed for a given neighborhood, to different activities
\bar{t}_{felt}	= Vector of travel times perceived/"felt" by an individual/household, to different activities
α	= Overdispersion parameter characterizing a negative binomial distribution (equals $1/m$)
μ	= Mean parameter characterizing a negative binomial distribution (equals mP)

Chapter 4:

- $\hat{\theta}$ = Maximum likelihood estimator of unknown parameters
- w_n = Estimate of gradient of likelihood function for the n th household, using $\hat{\theta}$
- W = Estimate of gradients of likelihood function across all sampled households, using $\hat{\theta}$
- $X_{i,n}^*$ = Optimal rate of consumption of activity i for household n
- v_{t_i} = Marginal utility of travel time to activity i
- v_{T_d} = Marginal utility of discretionary time
- $\delta_{iP}, \delta_{iPY}, \delta_Y, \delta_{YY}$ = Unknown parameters characterizing expanded indirect utility function
- $\delta'_{iP}, \delta'_{iPY}, \delta'_{PYT}$ = Identifiable functions of unknown parameters and prices which
characterize the expanded indirect utility function

Chapter 5:

- $E_{\varepsilon}(\cdot)$ = Expectation operator, over unobserved/error term ε

Acknowledgments

The creation of a dissertation can involve many, many people, and this one certainly did. I am indebted to the National Science Foundation and University of California at Berkeley, for their multi-year fellowships, and to the University of California Transportation Center for its generous dissertation grant. These sponsorships afforded me the security to focus on my intellectual interests throughout graduate-school.

I also had great fortune in securing Professors Mark Hansen, Daniel McFadden, and Martin Wachs to advise me in this project. All have provided significant insight and substantially facilitated this work. Michael Mauch, Phil Spector, and Simon Cawley were very generous with their time and programming aid. Chuck Purvis and others in the San Francisco Bay Area's Metropolitan Transportation Commission's modeling group provided hundreds of megabytes of critical data. Finally, I am grateful to Professors Elizabeth Deakin, Carlos Daganzo, Tom Golob, David Gillen, Steve Goldman, Ken Train, and Browyn Hall – for their advice and expertise on specific topics.

Chapter One: Introduction

This research examines a methodology for modeling household travel demand, as tied to out-of-home activity participation. The investigation adheres to the microeconomic theories of rational behavior and utility maximization¹ by the household and incorporates constraints on time, in addition to the common constraint on monetary expenditures. The methodology is tested empirically for several model specifications, using data from the San Francisco Bay Area. The results provide estimates of optimal trip generation and distribution (that is, destination choice) by households together with multiple economic variables, including cross-travel time elasticities, values of time, and welfare changes.

Little prior travel-behavior research has taken into account a time constraint or explicitly recognized travel demand as driven by demand for activities at physically separate destinations. Much of the research regarding time constraints has been theoretical, with little empirical support (*e.g.*, Becker 1965, Johnson 1966, DeSerpa 1971). A primary reason for the absence of empirical method is the difficulty satisfying utility maximization theory while permitting estimation. Other methods of analysis have resorted to substantial simplification of behavior based on strong assumptions such as bindingness of a single constraint (either the money or the time constraint is binding, but not both) and/or strongly additive preferences (*e.g.*, Zahavi 1979a, Zahavi *et al.* 1981, Gronau 1970). Accordingly, these methods have lost many relations of interest.

Discrete-choice models can accommodate the simultaneous (rather than sequential) nature of a variety of decision types and can be consistent with utility-maximizing behavior (McFadden 1974, Ben-Akiva and Lerman 1985, Train *et al.* 1987). However,

many choice variables are *ordered* or *continuous* (for example, the number of dining trips per month, square footage of home parcel). Ordered logit and probit models have been estimated for a single choice and for error-correlated simultaneous choices (*e.g.*, Yen *et al.* 1998), but not for a set of simultaneous choices where the parameters are constrained across equations or where the outcomes are cardinal (such as the number of person-trips by a household to different activity types over a day or more). If ordered choices were modeled simultaneously as non-ordered choices, the independence of irrelevant alternatives (IIA) property of the logit model would not be tenable; and the probit suffers from intractability for large numbers of (non-ordered) choices.

Thus, this dissertation takes a different approach and seeks to illuminate the interactions and trade-offs among demands for out-of-home activities and, therefore, travel. The methodology is sufficiently flexible that other consumption can be incorporated as well. The approach employs models consistent with utility theory so that the basic model structure and resulting predictions yield behaviors that are economically rational under a wide range of circumstances. Moreover, utility theory provides numerous extensions, supplying, for example, estimates of welfare changes, cross-time demand elasticities, and values of time.

In this research, systems of demand functions are derived from flexible functional forms of the indirect utility function through parallels to Roy's Identity. Continuous (though latent) demand levels underlie the system of interdependent equations, and these equations are simultaneously estimated so that cross-equation parameter constraints and correlated error structures are accommodated. The system is estimated as a set of negative binomial regressions, produced from mixing independent Poissons with

stochastic gamma terms and thereby providing for unexplained heterogeneity in behavior. These gamma terms are correlated, recognizing the correlation of unobserved information across multiple responses for a single observational unit.

The methods developed here are intended to further the state of the art in travel-demand modeling. The behavioral foundations of the investigated models are stronger than those of many existing models, lending greater credibility to the results and predictions. And the incorporation of relevant market “prices” (in the form of travel times) as well as two distinct budget constraints makes the models applicable to a variety of policy scenarios.² Moreover, the requisite data are commonly available to metropolitan and local planning organizations, so the methods advanced and applied here can be implemented in the short term.

Additionally, the resulting models allow for various tests of hypotheses concerning travel- and activity-related consumption, such as the existence of constant travel-time budgets. Application of microeconomic theory using the model’s estimated (scaled) indirect utility functions also permits evaluation of “welfare” changes due to policy changes (*e.g.*, Hausman *et al.* 1995, Burt and Brewer 1971, Cicchetti *et al.* 1976). For example, through inversion of the indirect utility function with respect to either one of the constraint levels, measures of a project’s social “cost” or “benefit” can be estimated in units of time and money by using differences in the constraints’ respective expenditure functions across households.

This model’s recognition of simultaneity in decision-making, time constraints on choice, and the discrete nature of travel data, along with its rigorous microeconomic

foundation, offer significant advantages in travel modeling. The ensuing chapters detail the model's specification and illustrate its application.

ENDNOTES:

-
- ¹ The assumption of utility-maximizing behavior, implicit in many models and their constraints (*e.g.*, cross-equation constraints on parameters), can often be tested using empirical results. For examples of such tests, see Christensen *et al.*, 1975, and Deaton and Muellbauer, 1980a.
- ² For example, if sales and service opportunities were to re-locate, the travel-time environment would change. These changes are incorporated directly in the proposed model, permitting immediate estimation of a household's response, via substitution and time-constraint effects.

Chapter Two: Review of Related Literature

General

Over the years, travel behavior has been modeled in a number of ways. Many of the earliest models were developed primarily for prediction; their virtue is that they are easy to apply. Later models are theoretically sounder, based on hypotheses concerning human behavior and focusing on causation. Some of the most plausible travel models acknowledge simultaneity in decision-making by avoiding strictly sequential estimation, hypothesize distinct behavioral mechanisms, and/or suggest new ways of adhering to microeconomic theory. However, shortcomings in existing models persist, and this research seeks to overcome the deficiencies. The purpose of this chapter is to summarize the strengths and weaknesses of existing models.

Models of Trip Generation

In the standard Urban Transportation Planning Model (UTPM), the first step is trip generation – estimation of the number of trips made for different purposes by households. The sequential, rather than simultaneous, estimation of such models and their lack of transportation-supply variables have long been recognized as inherent weaknesses in this mainstay of planning practice (*e.g.*, Dickey 1978, Gur 1971), yet these practices continue in the present day (*e.g.*, MTC 1996, Purvis *et al.* 1996, ITE Journal 1994). In their comprehensive book *Modelling Transport*, Ortúzar and Willumsen (1994) point out that while one's access to opportunities affects trip generation and “offers a way to make trip generation elastic (*i.e.*, responsive) to changes in the transport system”, it “has rarely been used....” (1994, p. 117) For example, in a two-stage “recursive” model of trip and trip-chain generation, Goulias and Kitamura's (1991) explanatory variables are almost

exclusively demographic in nature; for non-demographic data, they use a rural-versus-large city dummy variable and segment their trip-chain model by three city sizes.¹

Few recent methodologies consider total trip demand before addressing other aspects of behavior, such as trip chaining, distribution, timing, and duration. The absence of interest may be due to the apparent inelasticity of total demand with respect to access costs. Following an extensive review of past literature on trip frequency as a function of several rather simple measures of location type and accessibility (such as local-area densities and distance to central business districts) and a correlation-based analysis of their own, Hanson and Schwab (1987) conclude that “accessibility level has a greater impact on mode use and travel distance than it does on discretionary trip frequency” – an “unexpected” result given “the strong trip frequency-accessibility relationship posited frequently in the literature” (1987, p. 735). And Ortúzar and Willumsen observe that the incorporation of typical measures of access “has not produced the expected results, at least in the case of aggregate modeling applications, because the estimated parameters of the accessibility variable have either been non-significant or with the wrong sign.” (1994, p. 147) These results may be questioned, however, since the models and measures used to examine this relationship generally are unrefined. In order to estimate the elasticity of travel demand with respect to access, more sophisticated, behaviorally based models should be used.

Systems of Equations

A set of model equations is estimated as a system when a correlated error structure is hypothesized, there exist endogenous explanatory variables, and/or cross-equation parameter constraints exist. Researchers have applied the technique of structural equation

modeling to predict multiple travel choices in a manner similar to the modeling methodology developed here, but without cross-equation parameter constraints or a strict behavioral basis. For example, Golob and McNally (1997), Golob and Meurs (1987), Golob and van Wissen (1989 and 1990), and Lu and Pas (1997) regress variables such as vehicle-miles traveled (VMT), time spent per day in different activities, mode share, and auto-ownership on exogenous socioeconomic variables as well as on several endogenous variables. Much of the software used by these researchers allows for latent-variable techniques, such as the Tobit and ordered probit. However, the foundation for such systems in a utility-maximizing framework is missing. In a recent paper, Kitamura writes that existing structural equations models “offer no explicit treatment of the decision mechanisms underlying activity engagements.” (Kitamura 1996) One finds that “prices” are absent from these models, and measures of benefit cannot be constructed from their results.

Outside of transportation, there are many simultaneous-equations models of demand for goods and services. Optimal shares of monetary expenditures are typically estimated after applying rigorous microeconomic theory (*e.g.*, symmetry in compensated substitution, homogeneity in prices and income, summability and concavity of expenditures); however, time constraints are not considered. Abundant experience with these models has resulted in an understanding of the limitations of different functional forms and the need for specific cross-equation parameter restrictions for conformance with neoclassical economic theory (such as demands’ zero-degree homogeneity in prices and income). For detailed examples, see Lau 1986, Deaton 1987, Deaton and Muellbauer 1980b, Stone 1954, and/or Pollack and Wales 1978 and 1980.

Hybrid/Simulation Models

Recent, so-called “hybrid” models hypothesize traveler decision mechanisms which require less information than utility maximization yet satisfy spatial and temporal constraints. For example, Recker’s (1995) Household Activity Pattern Problem (HAPP) algorithm minimizes a generalized time cost function (which he calls “disutility”) subject to linear coupling, connectivity, temporal, and budget constraints. However, his method takes demand for participation in activities (as well as their duration and location) as given and neglects actual behavior for calibration of the model or its objective function. While the model is detailed and able to accommodate a variety of constraints, it avoids consideration of the basis for travel demand and is effectively a scheduling problem.

STARCHILD (Recker, *et al.* 1986a, 1986b) and SMASH (Ettema *et al.* 1993, 1995a) are similar to Recker’s HAPP model in that an activity program is provided exogenously, decision rules to choose among alternatives are relatively simplistic, and the models determine scheduling. Another model, AMOS (RDC 1995), can be classified similarly, but it requires more inputs and is tailored for response prediction in a limited policy setting. In comparing these models to econometric models, Bowman and Ben-Akiva (1996) observe that with the “hybrid” models the sample of considered alternatives is often inadequate, the response or decision process is probably too simplistic, and many significant, related decisions must be determined exogenously (*e.g.*, activity type, location, and travel mode).

Discrete- and Discrete + Continuous-Choice Models

Following McFadden’s seminal linkage of the logit model specification to microeconomic theory (1974), many discrete-choice models have been developed for the

purpose of travel demand modeling. The strict application of these models requires a complete specification of the feasible choice set, restricting the simultaneous and flexible estimation of total demand. Notwithstanding this limitation, many of these models remain microeconomically rigorous by assuming and applying the principles of utility maximization, though some of the strongest applications are not in the area of transportation. For example, Cameron (1982 & 1985) tests flexible² indirect-utility specifications in nested logit models for her analysis of home-weatherization choices. While rigorous, the size of her problem is limited; she evaluates two choices – the installation of energy-conserving appliances and, when applicable, the appliance package chosen.

In an early study of travel behavior, Adler (1976) relies exclusively on a multinomial logit across “all” possible non-work trip patterns for households, but the independence of irrelevant alternatives (IIA) assumption implicit in the logit formulation is unlikely to hold in his model. Fortunately, the nested logit structure has provided a useful way to avoid imposing the IIA property. Domencich and McFadden (1975) detail a four-level nested logit specification by modeling shop-trip mode split, time-of-day choice (peak vs. off-peak), destination choice, and “frequency.” Still, their model’s permitted shopping-trip frequency allows just one or no shop trips per household per day, which may be too limiting for many applications.

Incorporating the choice of trip purpose, but assuming fixed total demand, Kitamura and Kermanshah (1984) sequentially estimate a nested logit for trip-purpose and trip-destination choices. In the destination-choice model, a negative and statistically significant coefficient on the time-of-day-times-distance variable, after controlling for

distance by itself, causes them to conclude that longer trips are less likely toward the end of a day, as time constraints become more binding. Their recognition of a possible time-budget effect is important; however, their assumption of the time-of-day variable's exogeneity is questionable, and the time constraint is accommodated obliquely.

Damm and Lerman (1981) recognize travel as a derived demand and combine discrete-choice models of activity participation with the continuous choice of activity duration. This model offers the advantage of providing information on the time-of-day for an individual's travel and a system of simultaneous equations for estimation of the five periods' activity durations. However, while the authors discuss the indirect incorporation of a discretionary-time constraint via an individual's socio-economic characteristics, this constraint is not made explicit. Moreover, the research considers only the choices of workers on a workday given five distinct periods³ during which to choose participation in a non-work activity, and the authors specify linear utility functions with additive separability across each of the five choices.

Kitamura's work in this area (1984) is similar to that of Damm and Lerman (1981), except in the functional form of the time-allocation equation and in the discussion of model set-up. Kitamura's models are more fundamentally linked to economic theory and avoid selectivity bias in parameter estimates (by weighting responses in the duration model according to observations' likelihoods in the discrete-choice model). Nevertheless, due to the substantial complexities of the model, Kitamura relies on very specific functional forms for indirect utility and error structure in order to readily derive activity-participation-time demands. He also considers only two classes of time use: mandatory and discretionary.

In the context of auto ownership and use, Mannering and Winston (1985) also combine a discrete with a continuous choice. Making use of Dubin and McFadden's (1984) appliance-purchase-and-consumption model specification, they specify a linear functional form for demand of a single good, vehicle miles traveled in period "t" (VMT_t), and, using Roy's Identity ($dv/dY \times VMT_t + dv/dP_{VMT_t} = 0$ [where variables here and throughout the paper are as defined in the List of Symbols, immediately following the Table of Contents], Roy 1943), determine the implied functional form for indirect utility (v). They then use this indirect utility in a nested logit model for the number of cars owned – and the type or "class" of vehicle, given the number owned. After estimating the logit – and thus an indirect utility function, the estimated levels of VMT are easily obtained. This modeling method provides another example of a semi-simultaneous mixed discrete-with-continuous model of travel, and it incorporates some basic economic theory for a behavioral basis. Unfortunately, for a case of multiple goods, working "backwards" to derive indirect utility can be very difficult unless one begins with highly constrained demand equations; the connection is more clear if one moves from a functional form for indirect utility to a form for demands. Moreover, Mannering and Winston's necessarily specific choice of functional form for VMT demand leads to a rather limiting indirect utility function, one that is not, for example, homogeneous of degree zero in income and prices (which is a theoretically required condition discussed in Chapter Three). And, unlike this research, their focus is not on activity participation, the influence of time constraints, or the accommodation of multiple, integer demands.

Harvey and Deakin's STEP analysis package (1996) does not simultaneously combine discrete and continuous choices, but it does apply discrete-choice estimation to a

wide array of travel-related decisions by individuals, including location choice and time of travel. Notably, STEP incorporates an entire region's travel "prices" (*i.e.*, interzonal travel times – peak and off-peak, and intrazonal parking prices) into its models of trip distribution. However, STEP is not fully simultaneous and pays little attention to the implications of microeconomic theory for model form.

Value-of-Time Models

For a long time microeconomics and utility theory focused on the money budget and monetary expenditures. In the 1960's and 1970's time valuation, the labor-leisure trade-off, and activity participation choices began to be studied in a variety of ways, using microeconomic principles. Becker (1965), Johnson (1966), DeSerpa (1971), Oort (1969), and Bruzelius (1979) provide theoretical derivations of time's valuation across different activities. However, their hypothesized models typically treat travel as a single activity and/or emphasize the time spent *participating in* (rather than accessing) the other activities. Moreover, their focus is on the theoretical value of time, rather than a working system of demand equations for participation in out-of-home activities.

Becker (1965) argues that time use is a highly relevant aspect of household decision-making and that "total-income losses" due to non-income-producing uses of time are very significant. Thus, he advocates the incorporation of time in economic models of the household. He suggests that a household's "full income is substantially above money income" (1965, p. 517) and acknowledges people's pursuit of "productive consumption," such as eating and sleeping (activities which Golob and McNally [1997] and others have termed "maintenance"). Becker also comments on the intra-household allocation of consumption and production activities, arguing that members offering relative

efficiencies in different areas (*e.g.*, high-wage earners) will contribute relatively more time in those pursuits. More recently, Jara-Díaz (1994) extends time-valuation models into a setting which relies on travel times and can illustrate modal trade-offs. However, his results remain similar to those just mentioned: largely theoretical, based on direct-utility functions, and rarely tested empirically – except when employing random-utility discrete-choice models.

Train and McFadden (1978) look specifically at the labor-leisure trade-off using a discrete mode-choice model. Their work demonstrates how wages might reasonably enter the conditional-utility specification, as well as how workers optimize their time use. But the model only considers the choice of workers and employs a restrictive, two-good, Cobb-Douglas direct-utility specification.

Golob, Beckman, and Zahavi (1981) acknowledge the imposition of both time and income constraints in a setting that uses microeconomic theory, but they either consider only one at a time to be binding or assume travel expenditures are negligible relative to time and/or money budgets. Such assumptions may rarely hold: one can reasonably expect that *both* constraints are binding, as long as people value time and do not experience satiation in consumption. For example, in order to maximize utility, a person can sell his/her time to increase income (while reducing discretionary time available), spend more time in enjoyable activities (*e.g.*, leisure) and/or buy time-saving goods (such as prepared meals). This assumption of the bindingness of constraints is testable in the proposed research⁴. Additionally, Golob, Beckman, and Zahavi neglect activity participation as the underlying basis for travel demands, rely on additive utility functions,

and model total distance traveled, rather than distinguishing trip types or estimating the number of trips made.

Summary of Related Literature

There is a well-documented interest in the modeling of travel-related behaviors. Moreover, substantial progress has been made in the topics of time constraints, simultaneity of travel-related choices, the modeling of both continuous and discrete behaviors, and the implications of microeconomic theory.

Still, deficiencies exist. Most prominently, the existing literature does not consider integer consumption of multiple goods based on a continuous and cardinal latent response in a microeconomically rigorous framework; behaviorally-based time-use research remains largely theoretical; models of simultaneous choices which are consistent with utility maximization tend to be of discrete choices; and supply-side variables have been lacking in models of trip demand.

In contrast, the present research prominently incorporates supply-side variables (in the form of travel times to iso-opportunity contours) while allowing simultaneous estimation of trip generation and trip distribution, based on continuous, underlying demands derived from rigorously applied microeconomic theory⁵. This research provides estimates of numerous behavioral descriptors, such as demand elasticities; and it allows for a variety of extensions, such as estimation of access times' effects on a household's *total* travel time and on its welfare. The methods and model specifications used are considerably different from those found in previous work, and they are described in the following chapters.

ENDNOTES:

-
- ¹ One, somewhat typical exception is Safwat and Magnanti's (1988) Simultaneous Transportation Equilibrium Model (STEM), where total trip generation is estimated as a function of a log-sum accessibility measure (derived via the calibration of their trip-distribution logit model).
 - ² Cameron (1982) investigates household preferences using the rather flexible translog and Leontief functional forms to describe indirect utility. Note that these functional forms are summarized in the Appendix, section A-1, and are discussed in Chapter Three.
 - ³ The periods for non-work activity participation that Damm and Lerman model are: prior to the home-to-work trip, during the home-to-work trip, during work, during the work-to-home trip, and following the work-to-home trip.
 - ⁴ The bindingness of constraints is tested in Chapter Five by calculating the T-statistics for the derivatives of the estimated indirect utility function with respect to the constraint levels for each household; these derivatives theoretically represent the shadow prices of these constraints, which come out of the utility maximization.
 - ⁵ In his ground-breaking time-valuation work, DeSerpa suggests that "(d)espite (its) difficulties" a *system-of-demands approach* to the problem, where travel times effectively represent the minimum amount of time required for participation in/consumption of an activity, "has considerable merit" because "'non-economic' factors, such as comfort and convenience are ... implicitly considered", aggregation of demands "does not depend on any arbitrary assumptions about the individuals comprising the group", and, "most importantly, the measure (of time's value) is compatible with the hypothesis of utility maximisation. No other (time-value) measure can make that claim." (1971, pg. 842) It appears that DeSerpa would strongly support an approach fundamentally very similar to the one proposed here.

Chapter Three: Research Methodology

Microeconomic Foundations

In this research household activity and other, related consumption trade-offs are posited to adhere to microeconomic theories of utility maximization. Estimates of travel-related behaviors, such as trip-making rates and trip distribution, are derived from empirical analyses of statistical models based on this theory. In a rather general formulation of the utility-maximization problem, a household may be assumed to derive its welfare (*i.e.*, utility) from consumption of/participation in a vector of distinct, out-of-home activities \bar{A} (which are location specific, include the household's work activities, and are indexed by i), the time spent participating in each of these activities T_i (and, in particular for the work activity, T_w), the total time spent accessing all of these activities $\bar{t}\bar{A}$ (where \bar{t} is the vector of fixed travel times to access the activities¹), and consumption of all other goods \bar{Z} . It is helpful to think of the consumption/decision variables in this problem as rates; for example, one activity might be the number of shopping trips in the local neighborhood per day. Under the general model, households are subject to *unearned* income (Y_{un}) and available-time (H) constraints which are also rates (*e.g.*, dollars per day, hours per day), and these constraints lead to trade-offs between consumption of the different goods. In equation form, the problem can be written as the following:

$$\begin{aligned}
 & \underset{\bar{A}, \bar{T}, \bar{Z}}{\text{Max Utility}}(\bar{A}, \bar{T}, \bar{t}\bar{A}, \bar{Z}) \\
 & \text{s.t. } \bar{P}_A \bar{A} + \bar{P}_{trvl} \bar{A} + \bar{P}_Z \bar{Z} \leq Y_{un} + wT_w, \sum_i T_i + \bar{t}\bar{A} = H, \text{ and } \bar{A}, \bar{T} \text{ \& } \bar{Z} \geq 0.
 \end{aligned} \tag{3-1}$$

Note that time spent for activity participation is of two types: travel to non-home sites (t_i) and during participation itself (T_i); both of these enter explicitly in the direct utility function, though only the participation time, T_i , is an endogenous variable. The work activity contributes to the income budget level via the wage earned, w ; but participation in most other activities is likely to cost money (with $P_{trvl,i} + P_{A_i}$ representing the monetary price per unit of participation in activity i , due to travel costs and direct participation costs). There is an equality in the time constraint since all time not spent in accessing and participating in activities outside of the home counts as time spent in at-home activities².

The general model is subject to various modifications. For example, if one wishes to focus on discretionary activity choices and assume work and income exogeneity in such decisions, one would not explicitly model work as an activity and would substitute total income, Y , for unearned income, Y_{un} , and discretionary time, T_d (total time minus, for example, work and school time), for total time, H . Also, there are many other constraint possibilities; for example, minimum participation-time constraints may exist for certain activities (such as working, dining out, or seeing a movie in a theater) and only fixed levels of consumption may be permitted (such as working or going to school five times per week).

FIGURE 3-1 here *****

Figure 3-1 is an illustration of what the utility maximization looks like in a simplified, two-activity case; in this illustration a single, per-unit price and time-expenditure exist for each of the two activities and income and time budgets are exogenous/given so that $t_1 A_1 + t_2 A_2 = H$ & $P_1 A_1 + P_2 A_2 = Y$. In a more-realistic, N-good case, the intersection of the two budget constraints is an N-2 dimensional hyperplane; so the optimal choice “bundle” of activities will not appear as a single point of intersection, as it does in the illustrated case of Figure 3-1. Furthermore, choice of activity participation *times* (over a given period), rather than just optimal rates, expands the decision space substantially, yielding a hyperplane of dimension 2N-2.

In practice, a closed-form/analytic solution to constrained maximization of direct utility functions of more than a couple goods is rare, because solution of the Lagrangian equation’s set of first-order conditions is often intractable. In order to derive a system of (optimal) demand equations, it has been found significantly more convenient to work with the *indirect* utility function, as defined in Equation 3-2 (with arguments defined as for Equation 3-1 and in the List of Symbols, which follows the Table of Contents).

$$\begin{aligned} \text{Indirect Utility} &= \{ \text{Max Utility} \mid \text{Budget \& Time Constraints} \} \\ &= v(\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y_{im}, w, H) \end{aligned} \quad (\text{Chapter$$

Three:-2)

By beginning from a specification of indirect utility, one can then rely on a relation called Roy’s Identity (Roy 1943) to provide individual demand equations. The derivation of the entire system from a single indirect utility specification imposes many cross-equation parameter constraints automatically (because many parameters are likely to show up in two or more of the demand equations). However, there are a variety of other

constraints implied by long-held microeconomic theories for the typical, money-based applications of these methods, and these constraints tend to be more subtle; they are discussed shortly, in a section titled Theory-Implied Constraints.

Roy's Identity in a Two-Budget Framework

Roy's Identity is the method for deriving demand functions, whose dependent variables (consumption) can be observed, from indirect utility, which is unobservable and ordinal – rather than cardinal – in nature. Fortunately, Roy's Identity continues to hold in a two-budget framework, although more restrictively than in the typical, single-budget framework. Given a functional specification for indirect utility, v , as well as exogenously determined available time (T) and income (Y) constraints, the relations one can use to identify optimal demand, A_i^* , are shown in Equation 3-3. Details of this equation's derivation are provided in section A-2 of the Appendix.

$$\text{Roy's Identity: } A_i^* = -\frac{\frac{dv}{dt_i}}{\frac{dv}{dY}} = -\frac{\frac{dv}{d(P_{trvl,i} + P_{A_i})}}{\frac{dv}{dY}}, \forall i,$$

where A_i^* = Optimal, long – run rate of consumption per period, (Chapter

v = Indirect utility, t_i = Travel time to Activity i ,

T = Time available per period,

$P_{trvl,i} + P_{A_i}$ = Unit Price to participate in Activity i

(due to travel & participation costs), & Y = Income available per period.

Three:-3)

When income and time budget levels are exogenous and observed, the derivation of optimal demand levels is reasonably straightforward. However, income and discretionary time are likely to be endogenous to the decisions to participate in non-work/discretionary activities; in other words, households probably make choices of how much time to spend

working – earning income while giving up discretionary time – when determining the amounts of other activities they might engage in. In such a situation, the identities allowing one to identify demands do not look so similar to the common form of Roy's Identity, and the estimations of value of time and compensated demand are complicated.

Imagine a situation where total time available to a household's members (*e.g.*, 24 hours each day a member is surveyed), marginal hourly wage of the household, unearned income, travel times, and activity-participation prices are observed. The Lagrangian equation and several of its first-order conditions for utility maximization would look like the following:

$$L(\bar{A}, \bar{T}, \bar{Z}, \lambda_{Time}, \lambda_{Money}) = U(\bar{A}, \bar{T}, \bar{t}\bar{A}, \bar{Z}) + \lambda_{Time} \left(H - \sum_k T_k - \bar{t}\bar{A} \right) \dots$$

$$+ \lambda_{Money} (Y_{un} + wT_w - \bar{P}_A \bar{A} - \bar{P}_{trvl} \bar{A} - \bar{P}_z \bar{Z})$$

$$L^{opt} = v(\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y_{un}, H, w) =$$

$$L \left\{ \begin{array}{l} \bar{A}^* [\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y_{un}, H, w], \bar{T}_A^* [\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y_{un}, H, w], \\ \dots \bar{Z}^* [\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y_{un}, H, w] \end{array} \right\},$$

$$\frac{dL^{opt}}{dt_i} = \frac{dv}{dt_i} = \frac{dL}{dt_i} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dt_i} = -\lambda_{Time} A_i^* + 0,$$

$$\frac{dL^{opt}}{dH} = \frac{dv}{dH} = \frac{dL}{dH} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dH} = \lambda_{Time} + 0,$$

$$\frac{dL^{opt}}{dw} = \frac{dv}{dw} = \frac{dL}{dw} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dw} = \lambda_{Money} T_w^* + 0, \text{ and}$$

(Chapter

$$\frac{dL^{opt}}{dY_{un}} = \frac{dv}{dY_{un}} = \frac{dL}{dY_{un}} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dY_{un}} = \lambda_{Money} + 0.$$

Three:-4)

The endogeneity of discretionary time leads to a form of Roy's Identity which differs from that shown in Equation 3-3. Following some simple manipulation of the first-order conditions found in Equation 3-4, one has the following form:

Roy's Identity with Discretionary – Time Endogeneity:

$$A_i^* = -\frac{\frac{dv}{dt_i}}{\frac{dv}{dH}} = -\frac{\frac{dv}{d(P_{trvl,i} + P_{A_i})}}{\frac{dv}{dY_{un}}}, \forall i,$$

where A_i^* = Optimal, long – run rate of consumption per period, (Chapter
 v = Indirect utility, t_i = Travel time to Activity i ,
 H = Total time available per period,
 $P_{trvl,i} + P_{A_i}$ = Unit Price to participate in Activity i
 (due to travel & participation costs),
 & Y_{un} = Unearned Income available per period.
 Three:-5)

The above identity is not the only one that can be derived from this model specification. Incorporation of the wage variable, w , allows one to identify optimal work time, T_w^* , as the ratio of the derivative of indirect utility with respect to wage and with respect to total time available. And the vector of other goods (\bar{Z}) remains identifiable (as it is under a situation of exogenous income and discretionary time); demands for these goods equal the negative ratio of the derivative of indirect utility with respect to their prices and with respect to unearned income.

Under a situation of endogenously determined budgets, the value-of-time computations change; if unearned income and total time available are observed but discretionary time is endogenous, one can use the following:

$$\text{Value of Time} = \frac{\lambda_{\text{Time}}}{\lambda_{\text{Money}}} = \frac{dv/dH}{dv/dY_{un}} = \frac{de_H(\bar{t}, Y_{un}, w, u)}{dY_{un}} \quad (\text{Chapter$$

Three:-6)

However, if unearned income is *not* observed in the data set (but total time available and wage are, and discretionary time is endogenous), one will need to rely on the following equation:

$$\text{Value of Time} = \frac{\lambda_{\text{Time}}}{\lambda_{\text{Money}}} = \frac{dv/dH}{\left(\frac{dv/dw}{T_w^*}\right)} \approx \frac{dv/dH}{\left(\frac{dv/dw}{T_w}\right)} \quad (\text{Chapter$$

Three:-7)

Note that in this equation one may care to use the *observed* amount of time worked (T_w) to approximate value of time rather than the optimal level of working hours (T_w^*), because unearned-income information may not be available and/or may be measured with significant error. Since unearned-income information is not available in the data set used here for empirical analyses, the approximation in Equation 3-7 is used in those models of Chapter Five that endogenize time expenditures.

Assuming that households are able to optimize their time expenditures and activity participation³, how will models which assume exogenous total expenditures/income and discretionary time compare in their value-of-time computations with those which incorporate these variables endogenously? One way to look at the difference is to manipulate the ratio of derivatives in the income-and-discretionary-time endogenous case; for example:

$$\frac{dv/dT_d}{dv/dY} = \frac{\frac{dv/dH}{\left(\frac{dT_d}{dH}\right)}}{\frac{dv/dY_{un}}{\left(\frac{dY}{dY_{un}}\right)}} = \frac{dv/dH}{dv/dY_{un}} \times \frac{dY/dY_{un}}{dT_d/dH}$$

(Chapter

$$= \text{Value of Time} \times \frac{\left(1 + w \left[\frac{dT_w}{dY_{un}} \right] \right)}{\left(1 - \frac{dT_w}{dH} \right)}$$

Three:-8)

So, if the second term in the last part of the above equation is greater than one, one will over-estimate the value of time. It seems reasonable that, as total time available (H) increases, a household's members will work somewhat more, but not all of the newly available time. Thus, the denominator of the second term is likely to be less than one but not necessarily very close to zero (especially if work restrictions – such as a forty-hour week maximum paid week – are imposed). And, as unearned income increases, one may expect work time to decrease, perhaps so much that wage multiplied by work time exactly cancels unearned income, making the top part of the equation close to zero and causing one's value of time estimate (with the assumption of work-time and income exogeneity) to be much lower than the actual.

If work time is exogenously determined for households, then work time is unresponsive to changes in total time available to a household, H , and unearned income, Y_{un} , and one will be estimating the true value of time, without inflation or deflation. Unfortunately, without observing the variable of unearned income across the sample, it is difficult to analyze how work time depends on total time and unearned income. However, one can crudely estimate work time's response to changes in total time

available by modeling observed work time for this sample as a function of wage, travel times, total time available, and a coarse estimate of unearned income; a simple ordinary least squares model across households with one or more workers produces a derivative value of just 0.0808 hours of work time per hour of total time available to the household (with a T-statistic of 29.5).⁴ The estimate of unearned income on which this crude model relies is a value equal to the household's income if the household has no surveyed workers and zero otherwise. Running this same model specification for *all* sampled households produces a coefficient estimate of just -0.145 hours per \$1,000 of unearned income (with a T-statistic of -27.3). These results suggest that the derivatives of work time with respect to both income and wage are small; in fact the ratio of the derivatives of indirect utility with respect to discretionary time and total income available to the household (as in Equation 3-8) are estimated this way to be about nine percent higher than the true value of time, on average⁵. If this is a good estimate of the bias in this measure, it makes sense to deflate the value-of-time results for models which taken income and discretionary time to be exogenous by five to fifteen percent.

Theory-Implied Constraints

The models estimated here are not as general as the formulation presented in Equation 3-2, due to a lack of data on monetary prices and an inability to microeconomically identify non-work time expenditures in activities; but they are described by a system of equations which determines the optimal number of out-of-home activities accessed per day by household members. In order for a system of demand equations to be consistent with microeconomic theory and common sense, the equations must generally be compatible with several types of constraints; not only do such

restrictions impose consistency with theory, they can be helpful in reducing the dimensionality of the problem (*i.e.*, the size of the parameter space)⁶. Non-negativity of optimal demands is a feasibility limitation, and concavity of total monetary expenditures in prices is a requirement when prices are exogenous and constant; these conditions are generally checked following model estimation. In contrast, zero-degree homogeneity of demands (in prices and expenditures/income) is typically imposed *a priori* and automatically in the functional specification, and summability of expenditures (to equal total budget) and symmetry (of compensated cross-price effects) are often imposed through parameter constraints. If the conditions of summability and symmetry are not needed for parameter identifiability, their viability can generally be tested using differences in the constrained and unconstrained likelihood values. A final constraint on many estimated models is the implicit assumption of separability of preferences from other, non-considered goods. These various constraint types and their usefulness in the models investigated here are discussed below.

Non-Negativity

Generally, people cannot consume negative amounts of a good, unless, for example, they own some and sell or give it to others. In the context of this research, one can argue that people sometimes pay others to participate in out-of-home activities for them (such as food shopping). However, the available data do not provide information on such transactions so all observations are non-negative and this condition is imposed on the estimates. The method of ensuring this condition via the estimation process used here is an assignment of a very low likelihood value every time the iterative maximum-likelihood search mechanism tries parameter sets which produce negative demand

estimates for every demand type of at least one household. If some, but not all, demand types are estimated to be negative for a given household, the parameter set is permitted but optimal demands which are initially predicted to be negative are set to a positive level very close to zero. The optimal demand rates are not set to exactly zero since it is expected that, for the demand types specified, every household will *eventually* have to consume at least one such good. For example, a demand set of four iso-opportunity contours for all types of discretionary trips represents a partitioning of destinations for a type of trip virtually all households eventually make. However, if trip purposes were partitioned quite narrowly, segregating purposes like “education,” “work,” and “child-care”, one would need to incorporate zero-level demands since households without students, workers, and/or children would not reasonably be expected to make such trips⁷.

Before concluding this discussion of non-negativity, one should recognize that the rather *ad hoc* choice of a close-to-zero level of demand to assign to households with a predicted-to-be-negative optimal demand level is not theoretically satisfactory, particularly when the other demand levels are left as initially predicted. In reality, such households find themselves at a corner solution, where Roy’s Identity no longer applies to *all* demand types at once; instead, theory suggests that an optimization over limited choice sets is undertaken and the maximized utilities of distinct scenarios are compared. This added complexity can be accommodated in the models presented here, though it has not been in the estimated models provided in Chapter Four. In fact, Chapter Four’s predicted demands are well above the close-to-zero value for all demand types in almost all the models estimated.

Concavity of the Expenditure Function

Price extremes are preferable to balanced prices; this characteristic is manifest in quasiconvexity of the indirect utility function and concavity of its inverse⁸, the expenditure function. While this characteristic is not immediately intuitive, it is theoretically expected. It is expected because at “average” prices, one can buy no more than one could buy across the combined feasible space of the two price extremes which produced the average, subject to a single budget level; so one cannot be better off at balanced prices than at a combination of the two extremes. Thus, the indirect utility resulting from a weighted average of price vectors can be no higher than that achieved from a weighted average of indirect utilities resulting from the two extreme price sets. Moreover, if one or more prices increase, one is at least as well off if one’s budget increases in an amount equal to the price change (a vector) times the vector of previously optimal quantities; this amount of added income will allow one to consume the old bundle of goods and thereby be just as well-off. But, due to substitution effects, one will likely shift away from consumption of the relatively more expensive goods and be able to be just as well-off, so the amount of expenditures needed to achieve a *given* level of utility is less and thus concave in prices. These relations translate to the matrix of second derivatives of the money-expenditure function in prices being negative semi-definite. (For further discussion of these conditions, see, *e.g.*, Varian 1992.)

How do these conditions apply in the present model, where time characterizes costs? If one were to consider *all* time use, one would expect humans to require more time in a day in order to be just as well-off if travel times increase. However, the amount of additional time required is not necessarily less than the quantity of activities consumed

times their change in travel times. Humans directly experience time use, including travel time, so time expenditures are arguments in the *direct* utility function. This aspect of time use also arises in the following discussion, on homogeneity, and affects the application of many microeconomic theories in a time-expenditure setting.

In reality, more time spent accessing opportunities/activities may require *more* than a full compensation of total time to keep welfare constant; the direct impact on one's welfare may be sufficiently negative⁹. Thus, concavity of time expenditures in travel times is not a required property. And, as one might expect, the sister property of a quasiconvex indirect utility function with respect to travel times does not apply here either. While the time-budget constraint resulting from a weighted averaging of two travel-time vectors leads to a feasible consumption space which is a subset of the union of the two feasible spaces of the original two vectors, one may be better off because the indirect utility function shifts when the time vector changes! There may be a preference for better-balanced travel times because, for example, one can then spend better balanced amounts of time participating in a variety of activities (versus being "stuck" in the few activities which are relatively travel-time inexpensive). Changes in iso-utility contours due to changes in the travel times can bring this about. For these reasons, the conditions of time-budget concavity and indirect utility quasiconvexity are not imposed or expected for the models estimated here.¹⁰

Homogeneity

Since money is just a unit of exchange and does not itself hold value, rational humans are expected to not alter their choices under pure inflation. The theory is that indirect utility and all demands are homogeneous of degree zero in prices and income; so,

if prices and income all change by the same factor, a household's welfare/utility and consumption choices do not change (see, *e.g.*, Deaton and Muellbauer 1980b, Varian 1992). A typical specification of indirect utility and its resulting system of demand equations show prices everywhere divided by total expenditures, so that homogeneity is implicit in the formulation; section A-1 of the Appendix details several such specifications for a money-expenditure setting, but a general description of such a model in a *time-and-money-homogenous* setting is the following:

$$\text{Indirect Utility} = v\left(\frac{\bar{t}}{H}, \frac{\bar{P}}{Y}\right) \quad (\text{Chapter Three:-9})$$

The idea that pure inflation should not change one's consumption patterns is theoretically acceptable in an environment where people pay for goods with money, but this is probably too strong an assumption for consumption which involves time expenditures, since time is not instantly tradable – people directly experience their spending of time¹¹. For this reason, several modifications were made to the typical model specifications, providing greater functional flexibility by not imposing homogeneity with respect to travel times and the time budget; these functional forms are shown in the section titled Model Specifications. Note that if information on monetary prices were available in the data sets, one could include these and impose homogeneity over prices and monetary expenditures.

Summability

The very common assumption of non-satiation¹², that a little more of a good is a positive thing, no matter how much a person already has, implies summability of monetary expenditures when one is considering consumption across all demand

alternatives. Summability is also the condition that the sum of all demands considered in a model times their prices equals total expenditure on the set of goods considered.

In a system of activity-demand equations where one is modeling all uses of time (or all uses of, say, discretionary time), one would probably want to impose summability to ensure that results are consistent with reality (*e.g.*, a 24-hour day). However, when one considers only the *number* of activities accessed, as in this research, rather than also modeling the amount of time spent in each, summability's imposition – in this case across travel-time expenditures – puts the focus on allocating an exogenous total travel time, rather than allocating total time available. Thus, summability would be unnecessarily limiting and is not imposed here.

Separability

The neglect of other goods' price information generally necessitates an assumption of separability and shifts the modeling focus to substitution and trade-offs *within* a subset of consumption over an exogenously determined subset of the budget. Separability exists when direct utility is a function of sub-utility functions having distinct good sets as arguments; if utility is an additive function of these subutility functions, strong separability exists.¹³

As an example, one may collect detailed data on households' consumption of food items but not have any information on their consumption of clothing, lodging, transport, and utilities. To be able to apply the rigorous microeconomic theories associated with utility maximization and estimate a system of demand equations across this limited data set, one would need to argue for separability of preferences and rely on food expenditures

as the exogenous budget constraint, rather than total budget. The utility function and demand functions would then be written as the following:

$$u(\bar{X}) = f(u_1(\bar{X}_{Food}), u_2(\bar{X}_{Other})),$$

$$X_{i,Food} = X_{i,Food}(\bar{P}_{Food}, Y_{Food}).$$

Three:-10)

(Chapter

Separability is a strong assumption; it implies that consumers can order their preferences in each, distinct subset of choices independent of the amounts of other goods consumed. Strong or additive separability is even more restrictive; it rules out the possibilities of inferior subsets of goods and complementarity across subsets while imposing approximate proportionality between own-price and income elasticities. A more detailed discussion of separability can be found in Deaton and Muellbauer (1980a).

A model which assumes separable preferences can be significantly more limiting than a model considering the role of the entire budget available to a consumer. However, if one assumes that prices of all non-considered goods are the same for all households, preference separability is unnecessary. In the case at hand, this condition requires that only the travel-time environments differ across the sample population. The constancy of other goods' prices across the sampled observations means that their effects are not identifiable empirically; so, even if these prices were known, their invariance would effectively conceal their distinct parameters within the set of identifiable effects. One of the limitations this assumption places on model estimates is that the effects of changes in relative prices of the non-considered goods will not be predictable with the results established here.

How valid is the assumption of price invariance across non-considered consumption in the models estimated here? The price of a McDonald's hamburger may be the same regardless of where purchased in a region, but the prices of other goods, such as restaurant dining and food shopping may vary according to land rents, freight delivery costs, and local shoppers' preferences. However, if, for example, demand types are defined sufficiently broadly in a spatial sense (*e.g.*, destination zones are large), average price variability may be rather negligible, with enough opportunities present to match the prices found elsewhere.

If prices of goods not considered in the demand system do vary significantly, one may assume that separability holds and replace the variable of total expenditures with that of the subset's expenditures. Or, if prices move proportionally together, according to one's location (*e.g.*, central-city versus suburban dwellers), one may consider deflating or inflating income measures according to a price index, across sampled consumers. These approaches are not taken here, however, because it is virtually impossible to argue separability of goods consumption and activity participation (since many activities are complements of consumption – for example, recreational activities and entertainment expenses) and because price and monetary-expenditure information is lacking in available data sets.

Symmetry

Symmetry is a condition that arises in the typical system-of-demands frameworks, *i.e.*, in those where only a monetary constraint governs. It refers to the condition of symmetry of *compensated* cross-price effects (Slutsky 1915). Income-compensated or Hicksian demands can be derived simply by taking the first derivatives of the typical,

money-expenditure function; the derivatives of these demands with respect to other goods' prices are the cross-price effects, and these are symmetric thanks to Young's Theorem (which says that the order of differentiation is not important). Thus, the second derivatives of the expenditure function with respect to prices P_i and P_j are symmetric, as illustrated in the following equations:

Expenditure = $e_s(\bar{P}, u)$ = Money needed at prices \bar{P} to achieve utility u ;

Money – Compensated Hicksian Demand = $h_{i,s}(\bar{P}, u) = \frac{de_s(\bar{P}, u)}{dP_i}$; (Chapter

Compensated Cross – Price Effect $_{ij} = \frac{dh_{i,s}(\bar{P}, u)}{dP_j} = \frac{d^2e_s(\bar{P}, u)}{dP_i dP_j} = \frac{dh_{j,s}(\bar{P}, u)}{dP_i}$.

Three:-11)

The matrix of compensated-demand derivatives is called the Slutsky matrix, and theory implies that it is negative semi-definite, since total money expenditures are concave in prices. However, in the model structure investigated here, the first derivative of time expenditures with respect to an activity's travel time is not the compensated demand for that activity. In the common application, expenditures equal prices times amount of goods consumed; but in the decisions considered here, time expenditures are the sum of access costs/travel times multiplied by the number of out-of-home activities consumed *plus* the amount of time spent in each activity (in- and out-of-home). Thus, the derivative of time expenditures with respect to any travel time is no longer equal to the time-compensated demands for activities, so the matrix of second derivatives of the time-compensated expenditure function is no longer the same as the time-based Slutsky matrix and symmetry is not a condition imposed on the demand system estimated here. The following equations illustrate this property:¹⁴

Typically, $e_s(\bar{P}, u) = \bar{P}' \bar{h}_s(\bar{P}, u)$; and,

since $e_s(\bar{P}, u)$ is homogeneous of degree one in prices and

$h_{i,s}(\bar{P}, u)$ is homogeneous of degree zero in prices,

$$\sum_i \frac{de_s(\bar{P}, u)}{dP_i} P_i = e_s(\bar{P}, u) = \sum_i h_{i,s} P_i, \text{ and } \sum_i \frac{de_s(\bar{P}, u)}{dP_i} = \sum_i h_{i,s},$$

$$\text{so } \frac{de_s(\bar{P}, u)}{dP_i} = h_{i,s}(\bar{P}, u).$$

(Chapter

However, $e_T(\bar{t}, \bar{P}, Y, u) = \bar{t}' \bar{A}_T(\bar{t}, \bar{P}, Y, u) + \sum_i T_{i,T}(\bar{t}, \bar{P}, Y, u)$,

and is not homogeneous of degree one in access times and

involves unidentifiable participation – time demands.

Three:-12)

Validity of Utility Maximization Hypothesis

It is important to recognize that many empirical analyses of demand systems, analyzing different consumption sets' shares of monetary expenditures for aggregate, serial data sets and using a variety of common forms (such as the translog and generalized Leontief), have failed to support results satisfying basic economic theories (*e.g.*, Guilkey *et al.* 1983 and Caves and Christensen 1980). For example, imposition of symmetry constraints may reduce the likelihood of the observed sample substantially or the concavity of expenditures in prices may not be satisfied at many observations. (Lau 1986, Deaton and Muellbauer 1980a & 1980b, Pollack and Wales 1978 & 1980) Lack of support for well-accepted economic theory by a model suggests that the model specification is substantially incorrect and/or the households/individuals observed are not economically "rational" (according to a utility-maximization hypothesis of behavior). Any modeler should be conscious of these possible inconsistencies and check for them where practicable. However, as discussed throughout much of this section on theory-implied constraints, very few of the theories which are expected to apply in a money-

expenditure setting are likely to hold here. Without symmetry and summability, Roy's Identity is the origin of virtually all restrictions imposed in the models estimated in Chapter Four; these restrictions are implicit by virtue of the common parameters found throughout the estimated demand equations and are due to the system's derivation from a single indirect utility function. Non-negativity of demands and positivity of the marginal utility of time are the only other conditions imposed here; however, marginal utility estimates are considered for their conformance with theory, and the concavity of estimated expenditure functions and convexity of estimated indirect utility functions are examined briefly.

Estimating Benefits and Costs

“Equivalent” and “compensating variation” are measures of welfare changes following price changes, each using a difference in expenditure functions but at different reference levels of indirect utility. The author knows of no empirical examples where equivalent and/or compensating variation has been quantified with anything other than a money metric. Actual welfare change is not measurable in known units, since it is generally agreed to be the change in *utility* associated with price/cost changes.¹⁵

The distributional effects associated with policy changes are very important. Total benefits exceeding costs/disbenefits only signifies a *potential* for Pareto superiority, *i.e.*, the possibility of a Pareto-preferred redistribution of the benefits so that no one is worse off following a positive-net-benefits change. (Varian 1992) As economist Steven M. Goldman describes it, “Cost-benefit analysis as a welfare measure which is done independently of distributional effects is fundamentally flawed.” (Goldman 1998)

Measurement of welfare changes using a money metric favors projects benefiting those who have the most monetary resources available, rather than those who might experience the most welfare benefit, because those with the most money can place the highest monetary value on a change in conditions¹⁶ (see, *e.g.*, Heap *et al.* 1992, Price 1993). This is of particular concern in the evaluation of projects producing significant time-expenditure differences, such as transportation infrastructure alterations (*e.g.*, Daganzo 1997). A fortunate result of recognizing a time constraint in utility maximization is that the indirect utility function can be inverted with respect to this budget variable and welfare impacts can be assessed with a time metric. Equation 3-13 provides the definitions of equivalent variation which are used here, in terms of money ($EV_{\$}$) and time (EV_T). As illustrated, equivalent variation can be written as the difference in expenditure function values at reference price levels, as well as, under constant budget levels, the integrals of the compensated/Hicksian demand equations.¹⁷

$$\begin{aligned}
 \text{Welfare Change}_{\$} \approx EV_{\$} &= e_{\$}(\bar{t}^o, T, u') - e_{\$}(\bar{t}^o, T, u^o) = \int_{\bar{t}'}^{\bar{t}^o} \bar{h}_{\$}(\bar{t}, T, u') \cdot d\bar{t} \\
 & \hspace{20em} \text{(Chapter} \\
 \&\text{ Welfare Change}_T \approx EV_T &= e_T(\bar{t}^o, Y, u') - e_T(\bar{t}^o, Y, u^o) = \int_{\bar{t}'}^{\bar{t}^o} \bar{h}_T(\bar{t}, Y, u') \cdot d\bar{t} \\
 \text{Three:-13)} &
 \end{aligned}$$

Note that the negative of equivalent variation can be understood to mean the maximum amount of money or time a household would be willing to give up to avoid the change in prices/travel times, if budgets levels are unchanged. Chapter Five's section on cost-benefit analysis uses both the income and time definitions of equivalent variation to estimate the welfare impacts of an increase in travel times.

Functional Specification

Assuming that the demand equations arise from derivatives of the indirect utility function, one may wish to select functional forms for the indirect utility function, v , which are flexible to a second (or greater) order¹⁸. This flexibility permits estimation of cross-price and income elasticities, in contrast to non-interactive functional forms, which produce only non-zero direct elasticities. The transcendental logarithmic's functional form (*i.e.*, the translog) is commonly used in practice (*e.g.*, Cameron 1982, Christensen *et al.* 1973 & 1975, Pollack and Wales 1980) and is quite flexible¹⁹, but it has some drawbacks. Under a situation of no cross-parameter constraints, the number of translog parameters increases with more than two times the square of (rather than linearly with) the number of goods, which may result in statistical insignificance for many parameters and low confidence in estimation – depending on sample size. For empirical estimability over limited sample sizes, one may need to make some *a priori* assumptions as to relationships and assume a relatively parsimonious form for estimation.

Other functional forms for indirect utility are also possible and have been used in money-expenditure systems of demand. A variety of forms are shown and discussed briefly in the Appendix (A-1), but the simplest to estimate impose untenable assumptions implicitly. For example, in a theoretically consistent linear-in-unknowns demand system of monetary expenditures on three or more goods, all income elasticities must equal one. (Lau 1986) And, in the traditional consumption framework where only a monetary budget governs, the Cobb-Douglas and utility-consistent Rotterdam (Barten 1964, Theil 1965) functional forms impose additivity and homotheticity assumptions on preferences – along with a constant, unitary elasticity of substitution²⁰ across all pairs of goods! (Greene

1993, Christensen *et al.* 1975) In reality, the substitutability of consumption goods may vary widely, given different relative combinations (for example, at extremes of goods ratios, less substitutability is expected than at better-balanced levels). The generalized Leontief functional form may be more flexible than the translog when unequal or low elasticities of substitutions exist across the choices (Guilkey *et al.* 1983, Caves and Christensen 1980), but can be rather intractable in its most general form (Diewert 1971 & 1974) and is not expected to perform as well when high and unequal substitution elasticities exist (Caves and Christensen 1980).

There is no particular reason to expect similarity of substitution across different activity types, but there is an expectation of high substitutability across certain choice definitions. For example, in the empirical investigations pursued here, activities are distinguished by the iso-opportunity zones in which they take place, rather than the activity type or purpose; therefore, one may expect very high substitution effects across zones and opt for a translog specification. Substitution is expected to be less when one considers very distinct activity types, such as personal business versus social, so the Generalized Leontief may be most useful in these cases; however, some trip types, such as non-food shopping and recreation, may to a certain degree still act as substitutes. Among the models estimated here, in Chapter Four, one of the specifications resembles Stone's Linear Expenditure System (1954), while the others are based on modifications of the translog specification.

Model Specifications

Four distinct model types are tested empirically in Chapter Four, and their functional forms are specified here, with typical economic notation for demand (X_i^*) replacing the

notation for optimal rate of activity participation, A_i^* . The first of these four models is an attempt at a relatively simple specification using an indirect utility specification similar to that which generates Stone's Linear Expenditure System (1954). The other three are modifications of the translog model (Christensen *et al.* 1973 & 1975), and they are presented here in order of increasing generality. All models are used to estimate long-run, optimal out-of-home activity participation rates (per day) for households, and all but the third are used only once, to model participation in discretionary activities. The third of these specifications is also used to model entire home-based tours of activities, rather than just individual stops, and these tours can include non-discretionary trip-making. All specifications shown, except the fourth, rely on discretionary time (total time minus work and school time) and income as exogenously provided arguments. The empirical results from those analyses are provided in chapters Four and Five.

Type 1 Model Specification: Modified Linear Expenditure System

In an effort to begin with as simple a functional specification as possible, Stone's Linear Expenditure System or "LES" (1954) was examined for use. However, without the ability to impose homogeneity in the time dimension and due to the presence of *two* budget variables, the resulting demand system is not nearly as simple as Stone's. The indirect-utility specification and resulting demand equations used for this modified-LES specification are as follows:

Indirect Utility = $v = \{ \text{Max Utility} | \text{Budget \& Time Constraints} \}$

$$= \frac{\prod_i t_i^{\alpha_i}}{-T_d - \frac{1}{2} \sum_{i,j} \beta_{ij} t_i t_j - \sum_i \beta_{iY} t_i Y}, \quad \text{so...}$$

$$X_i^* = \frac{\alpha_i}{t_i} \left[T_d + \frac{1}{2} \sum_{j,k} \beta_{jk} t_j t_k + \sum_j \beta_{jY} t_j Y \right] - \sum_j \beta_{ij} t_j - \beta_{iY} Y, \quad (\text{Chapter$$

where $t_i = \text{Travel Time to Activity } i, Y = \text{Income},$

& $T_d = \text{Discretionary Time Available},$

and $\beta_{ij} = \beta_{ji} \quad \forall ij$ (for identifiability of parameters).

Three:-14)

Note that Stone's *original* specification produces a system of demand equations whose parameter space increases only linearly with the number of goods consumed, " T ". While Stone's system requires the estimation of $2I-1$ parameters, the modified system used here has a parameter set which grows quadratically with the number of goods considered, requiring the estimation of $2I+I(I+1)/2$ parameters²¹! The assumption of homogeneity saves a modeler many degrees of freedom for estimation purposes; however, there exist many major weaknesses with the LES, as discussed in Section A-1 of the Appendix.

Under the modified LES specification used here, the value of time is independent of the budget levels, depending only on access times, and the time-budget elasticities of demand are independent of all variables but the demand's own access time; such functional inflexibilities pose a serious problem. For example, this model's estimation results, which are provided in Chapter Four, produce *negative* values of time for all households in the 10,834-observation sample! A more flexible model is almost certainly necessary.

Type 2 Model Specification: Modified Translog

Having considered the strengths and weaknesses of various functional forms, many of which were discussed in a previous section, titled “Functional Specification”, a modified version of the Christensen *et al.*'s translog form (1975) was chosen to represent the indirect utility function for the remaining set of models estimated here. The translog was chosen for its second-order functional flexibility as well as for its ability to flexibly model substitutes well.²² The most restrictive form of this general specification that is analyzed here is termed the “Type 2 Model Specification”, and it is as follows:

$$\begin{aligned} \text{Indirect Utility} = v &\approx \text{Translog}(\bar{t}, T_d, Y), \\ v &= \alpha_o + \sum_i \alpha_i \ln(t_i) + \sum_{ij} (1/2)\beta_{ij} \ln(t_i) \ln(t_j) + \\ &\sum_i \gamma_i \ln(T_d) \ln(t_i) + \sum_i \gamma_{iY} \ln(Y) \ln(t_i) + \gamma_{TY} \ln(T_d) \ln(Y) \end{aligned} \quad (\text{Chapter$$

Three:-15)

The optimal demand levels which result from application of Roy's Identity (with respect to time) to the above formulation are the following:

$$\text{So, } X_i^* = \frac{-\left(\frac{1}{t_i}\right) \left(\alpha_i + \sum_j \beta_{ij} \ln(t_j) + \gamma_{iY} \ln(Y) + \gamma_{iT} \ln(T_d) \right)}{\frac{1}{T_d} \left(\sum_j \gamma_{jT} \ln(t_j) + \gamma_{TY} \ln(Y) \right)}, \quad (\text{Chapter$$

where t_i = Travel Time to Activity i , Y = Income,

& T_d = Discretionary Time Available,

and $\beta_{ij} = \beta_{ji} \forall ij$ & $\gamma_{TY} = 1$ (for identifiability of parameters).

Three:-16)

Notice that the number of parameters in this modified translog system increases quadratically with the number of good types considered. The system of equations

requires the estimation of $3I+I(I+1)/2$ parameters, which is $2I$ more than in the LES-based, Type 1 model.

Type 3 Model Specification: Modified Translog with Constants

The Type 2 model specification can not be nested with a no-information model specification (*i.e.*, a model without any explanatory information) since all of its unknown parameters interact with explanatory variables. Therefore, a more flexible model of this form was investigated, adding $I+1$ parameters to the modified-translog specification to effectively function as constant terms; this change produces the following:

Indirect Utility =

$$v = \alpha_o - \sum_i \mu_i t_i + \mu_o T_d + \sum_i \alpha_i \ln(t_i) + \sum_{ij} (1/2) \beta_{ij} \ln(t_i) \ln(t_j) + \sum_i \gamma_i \ln(T_d) \ln(t_i) + \sum_i \gamma_{iY} \ln(Y) \ln(t_i) + \gamma_{TY} \ln(T_d) \ln(Y) \quad (\text{Chapter$$

Three:-17)

The optimal demand levels which result from application of Roy's Identity (with respect to time) to the above formulation are the following:

$$So, X_i^* = \frac{\mu_i - \left(\frac{1}{t_i} \right) \left(\alpha_i + \sum_j \beta_{ij} \ln(t_j) + \gamma_{iY} \ln(Y) + \gamma_{iT} \ln(T_d) \right)}{\mu_o + \left(\frac{1}{T_d} \right) \left(\sum_j \gamma_{jT} \ln(t_j) + \gamma_{TY} \ln(Y) \right)}, \quad (\text{Chapter$$

where $t_i = \text{Travel Time to Activity } i, Y = \text{Income},$

& $T_d = \text{Discretionary Time Available},$

and $\beta_{ij} = \beta_{ji} \forall ij$ & $\gamma_{TY} = 1$ (for identifiability of parameters).

Three:-18)

The expectation is that this more flexible specification will provide more reasonable estimates of behavior, such as demand elasticities and values of time; it also allows the

nesting of the Type 2 specification within the Type 3 and so provides a means of gauging the need for Type 3's added flexibility.

Type 4 Model Specification: Modified Translog with Constants, using Wage and Total Time Data

As discussed in the section on application of Roy's Identity, one's discretionary-time and income budgets may be endogenous to the choice of discretionary-activity participation. Thus, a model that allows for these choices in a simultaneous manner may prove useful. Taking the most flexible of the model specifications suggested, *i.e.* that of the modified translog with constants, a specification based on wage rates and *total* time availability to a household's members is the following:

Indirect Utility =

$$v = \alpha_o - \sum_i \mu_i t_i + \mu_H H + \sum_i \alpha_i \ln t_i + \sum_{ij} (1/2) \beta_{ij} \ln(t_i) \ln(t_j) + \sum_i \gamma_{iH} \ln(H) \ln(t_i) + \sum_i \gamma_{iw} \ln(w+1) \ln(t_i) + \gamma_{wH} \ln(w+1) \ln(H) \quad (\text{Chapter$$

Three:-19)

The optimal demand levels which result from application of Roy's Identity (with respect to time) to the above formulation are the following:

$$So, X_i^* = \frac{\mu_i - \left(\frac{1}{t_i} \right) \left(\alpha_i + \sum_j \beta_{ij} \ln(t_j) + \gamma_{iw} \ln(w+1) + \gamma_{iH} \ln(w+1) \right)}{\mu_H + \frac{1}{H} \left(\sum_j \gamma_{jH} \ln(t_j) + \gamma_{wH} \ln(Y) \right)}, \quad (\text{Chapter$$

where t_i = Travel Time to Activity i , w = Wage Rate,

& H = Total Time Available,

and $\beta_{ij} = \beta_{ji} \forall ij$ & $\alpha_{iH} = 1$ (for identifiability of parameters).

Three:-20)

Since income is not exogenous in this model, values of time are estimated using Equation 3-8's approximation, which requires an estimate of the unobserved variable optimal work time, T_w^* . As described in Table 3-2, a household's work time is assumed to be eight hours per day for full-time workers plus four hours per day for part-time workers. These results of these computations are shown in the following chapter, Chapter Four.

Statistical Specification

Integer Demand Observations and the Poisson Assumption

Observed demands can be visibly discrete in *limited*-period data sets. However, one may expect that continuous and smoothly differentiable preference and demand functions underlie observed behavior, since households are typically free to optimize their choices over relatively long periods of time. This is the assumption made here, so a link to a model of cardinally ordered discrete demand levels is needed for empirical estimation. This link may be best provided via the Poisson distribution, which is defined over the set of non-negative integers.

Given an assumption of Poisson-distributed demands, the various activity types i (*e.g.*, near vs. far, or dining vs. social activities) can be characterized as in Equation 3-21. This set of Poisson random variables is simultaneous in nature, since the derivation of all mean demands from a single indirect utility specification introduces common parameters across the demand specifications.

$$X_i \sim \text{Poisson}(\lambda_i), \forall i,$$

$$\text{where } \lambda_i = X_i^* = f_i(\bar{P}_A, \bar{P}_{trvl,A}, \bar{P}_Z, \bar{t}, Y, T).$$

Three:-21)

(Chapter

The Poisson distribution arises naturally from counts of independent events that occur at a specified rate, so it would be a plausible distributional assumption if household members make trips at randomly and independently selected times throughout their window of discretionary time. In reality, household members are often constrained to temporally and spatially coordinate their trip-making due to limitations on automobile, driver, and transit availability, closures of activity sites (*e.g.*, stores late at night), and the desire to engage in activities together. Moreover, activity participation and travel take time, undermining the assumption that such events occur independently in time.²³

Without independent and identically distributed exponential inter-event times, the Poisson may still characterize the counts of activity participation across households; this may be particularly true over longer periods of time, as the short-term/daily realities of trip chaining and activity coordination take on less importance relative to long-run behavior. Unfortunately, the household travel surveys with sufficient sample size and detail for use in this study tend to be of short duration (*e.g.*, one to two days, typically); so the Poisson remains a significant assumption. However, as described next, the Poisson is mixed here with a gamma distribution, in order to capture unobserved heterogeneity across different households having the same set of observed characteristics.

Generalizing the Poisson Assumption through Use of a Negative Binomial

One limitation of the typical Poisson regression model is that its variance is constrained to equal its mean. Cameron and Trivedi (1998) describe the failure of the Poisson assumption of equidispersion as qualitatively similar to a failure of homoscedasticity in a linear regression model, but with possibly much larger effects on standard errors. However, allowing variation in the Poisson's parameter λ by mixing the

Poisson with another distribution can help one avoid a restrictive equi-dispersion assumption and accommodate the effect of unobserved factors on each household's mean trip-making rates.

Overdispersion is common in behavioral data (Cameron and Trivedi, 1998), and it was found to be present in the trip-making data sets, after controlling for a variety of market characteristics and demographic explanatory variables and then applying statistical tests described in Cameron and Trivedi (1998). Even though a large set of explanatory variables is used, the dispersion coefficients (the α parameters) are highly statistically significant in all models, indicating a decisive rejection of an equidispersion hypothesis.

Factors other than travel times and income and time budget levels play an important role in household activity participation rates. Whether a household is active or inactive, profligate or frugal, may mean significant differences in optimal rates of trip-making. Thanks to these *unobserved* characteristics, one also would expect there to be some variation in trip-making rates across households with the same *observed* characteristics. It is therefore useful to add a "second layer" of stochasticity by mixing a Poisson with a second distribution. Additionally, one may reasonably hypothesize correlation across the unobserved components of the various demands by a single household, since one can expect the deviation in a household's demand for one type of trip to be associated with deviations in its other trip demands. For example, if a household is taking part in a certain out-of-home activity more than one expected (given its time and income constraints and the set of travel times it faces), it may also tend to participate in other out-of-home activities with a higher-than-expected frequency. Information on one set of

demands observed for a specific household, relative to expectations, is likely to help one better predict other consumption by the same household.

The use of a compounded and correlated error structure within a system of Poisson equations is unusual. Few modeling efforts have used a multivariate Poisson form to model demands, particularly in a rigorous micro-economic framework with more than two choice types. Hausman, Hall, and Griliches (1984) use the seminal set-up of Bates and Neyman (1952) in order to model the number of patents received by a panel of firms over time as “fixed-effect Poissons”, which integrate to negative binomials; but there are no “prices” or explicit links to profit maximization in their model. Hausman, Leonard, and McFadden (1995) estimate the choice of recreational sites using a multinomial distribution conditioned on total number of trips, where the total is a fixed-effects Poisson and travel costs are included in the set of explanatory variables. Hausman, Leonard, and McFadden’s model provides measures of welfare/benefits via a logit model’s log-sum maximum-expected-utility. However, their model does not consider other types of trips or related consumption, and the two decision stages (*i.e.*, total number of trips and allocation of these trips across sites) are estimated sequentially, rather than simultaneously.

If significant flexibility of the error terms’ covariance structure (*e.g.*, a multivariate normal distribution across the “ ε_i ’s”) were permitted, the maximum-likelihood equation’s values would almost certainly have to be computed using numerical integration or distribution simulation over the multiple of probabilities. Such an approach is illustrated by Equation 3-22, with $g(\bar{\varepsilon})$ representing an assumed joint-density, such as

the multivariate normal, with fixed mean and a variance-covariance matrix to be determined.

$$\text{Prob}(\bar{X} = \bar{k} | \lambda_1(\bar{\varepsilon}), \dots, \lambda_I(\bar{\varepsilon})) = \int_{\varepsilon_1=-\infty}^{+\infty} \dots \int_{\varepsilon_I=-\infty}^{+\infty} \prod_i \exp(-\lambda_i) \left(\frac{\lambda_i^{k_i}}{k_i!} \right) g(\bar{\varepsilon}) d\bar{\varepsilon} \quad (\text{Chapter$$

Three:-22)

Many have used simulation for estimation of complex specifications; for example, Yen *et al.* (1998) have used it for a set of correlated-in-unobserveds ordered probits, and Train (1996), McFadden and Train (1996), and Mehndiratta (1996) have used simulation successfully for a random-parameters logit model. However, estimation times tend to be long – and a second simulation will invariably lead to a set of different estimates.

Instead, if one can specify the second layer of stochasticity (*i.e.*, the layer within the Poisson's own lambda parameter) so that the random component can be tractably integrated out, the estimation is much simplified. For this reason, an integrable error structure was sought in this research, leading to the mixing of a Poisson with a gamma to produce a negative binomial distribution; the use of the *same* gamma error term across all of a household's demands allows for a cancellation of these terms in the probabilities of a multinomial (which is conditioned on a negative binomial for total demand), as illustrated in the following equations:

If $X_i \sim \text{Poisson}(\lambda_i = X_i^* \varepsilon)$, and

$\sum_{i=1}^I X_i = X_T \sim \text{Poisson} \hat{\wedge} \text{Gamma} = \text{Negative Binomial}(m, p^*)$, where $\frac{m(1-p^*)}{p^*} = \sum_{i=1}^I X_i^*$,

then $\text{Prob}(X_1, X_2, \dots, X_I | p_1, p_2, \dots, p_I) = \text{Multinomial}(\bar{X} | \bar{p}, X_T) \text{Neg. Bin} \left(X_T | \sum_{i=1}^I X_i^* \right)$

$$= \left(\frac{X_T!}{\prod_{i=1}^I X_i!} \prod_{i=1}^I p_i^{X_i} \right) \left(\frac{\Gamma(X_T + m)}{X_T! \Gamma(m)} \right) (1-p^*)^{X_T} (p^*)^m$$

where $p_i = \frac{X_i^* \varepsilon}{\sum_{j=1}^I (X_j^* \varepsilon)} = \frac{X_i^*}{\sum_{j=1}^I X_j^*}$.

(Chapter

Three:-23)

As implied in the above equations, two parameters characterize a negative binomial.

The parameters m and p^* are used in Equation 3-23, and these can be thought of as a size and probability parameter. Appendix section A-4 describes the negative binomial distribution in more detail.

The negative binomial assumption has been used in empirical work for several decades. For example, Chatfield *et al.* (1966) used a single negative binomial regression equation to model household purchases, but Rao *et al.* (1973) were the first to use a *system* of equations and thus a specification similar to (yet much simpler than) the set-up followed here. Rao *et al.* (1973) modeled the number of boys and the number of girls born to a pair of parents as symmetric binomials (*i.e.*, with probability of either equal to 0.5) conditioned on a negative binomial for the total.

As mentioned, a negative binomial distribution (NB) can be thought of as a Poisson whose parameter varies as a gamma (*i.e.*, Poisson $\hat{\wedge}$ Gamma). And the entire system of demand equations can still be considered a system of Poissons, but with variation permitted in the rates (λ_i 's). Knowing total count, a system of *independent* Poissons

becomes a multinomial (MN) distribution; knowing total count under a system of correlated Poissons conditioned on a negative binomial (*i.e.*, $\text{Poissons} \wedge \text{NB}(\lambda_T)$) also implies a MN. However, for one to be able to identify the probabilities of the choices (p_i 's) with a closed-form solution – and avoid simulation or numerical integration, one must make some assumptions and thereby constrain the system's "double stochasticity" to a certain form. Here, the assumption that the system conditioned on total count is a MN with p_i 's equaling λ_i/λ_T implies that the variation in each λ_i is equal to the factor of variation in the λ_T times p_i . Thus, for such a set-up, each *multiplicative* gamma error component is the same value as the gamma random variable that affects total trips.

Since a gamma variable times a constant is also a gamma variable, all marginal distributions of trips (X_i 's) are negative binomial($m, p_i P$), with their mean rate having a gamma distribution; in statistical notation: $X_i \sim \text{Poisson}(\lambda_i) \wedge \text{Gamma}(m, m/X_i^*)$.²⁴ The density function for a gamma distribution is shown in Equation 3-24, helping illustrate why the rates for individual demands are also gamma distributed. The stochastic assumptions of observed demands having Poisson distributions whose rates interact multiplicatively with the same unobserved gamma variable, for a given household, imply that individual rates can be thought of as gamma variables with the same size parameter as total demand (m), but with modified scale parameters (m/X_i^* , rather than m/X_T^*).

$$\lambda_T \sim \text{Gamma}(m, m/X_T^*) \rightarrow \text{pdf}_{\lambda_T}(\lambda_T) = \frac{m e^{-\lambda_T m/X_T^*} (\lambda_T m/X_T^*)^{m-1}}{X_T^* \Gamma(m)},$$

$$\text{where } m > 0, \lambda_T \geq 0, \text{ and } \Gamma(m) = \int_0^{\infty} e^{-x} x^{m-1} dx \quad [= (m-1)! \text{ if } m \text{ is integer}] \quad (\text{Chapter$$

$$\text{So } p_i \lambda_T \sim \text{Gamma}(m, m/p_i X_T^*) = \text{Gamma}(m, m/X_i^*).$$

Three:-24)

Typically, a multinomial's component levels are negatively correlated, because of a fixed sum. However, when the sum or total is allowed to vary as permitted here, the unconditional correlation becomes positive, as shown in the following set of equations. Overdispersion, as previously discussed, is also a property of this distribution and illustrated in the following equations.

$$\text{If } (X_1, X_2, \dots, X_I) \sim \text{Multinomial}(\bar{p}, X_T) \underset{X_T}{\wedge} \text{Negative Binomial}(m, P = \frac{1-p^*}{p^*}),$$

$$\text{then } E(X_T) = mP = \sum_i X_i^*,$$

$$\text{and } V(X_T) = mP(1+P) = E(X_T)(1+P) > E(X_T) \rightarrow \text{Overdispersion.}$$

$$E(X_i) = mPp_i = E(X_T)p_i,$$

(Chapter

$$\text{and } V(X_i) = p_i^2 PE(X_T) + p_i E(X_T) > E(X_i) \rightarrow \text{Overdispersion.}$$

$$\begin{aligned} COV(X_i, X_j) &= E_{X_T} (COV(X_i, X_j | X_T)) + COV_{X_T} (E(X_i, X_j | X_T)) \\ &= -p_i p_j E(X_T) + E_{X_T} (p_i p_j X_T^2) - p_i p_j E(X_T)^2 = p_i p_j PE(X_T) > 0, \text{ as expected.} \end{aligned}$$

Three:-25)

In sum then, the system of demand equations can be termed a multivariate negative binomial, since each of the demands is marginally represented by a negative binomial. Moreover, the special, same-gamma-term assumption allows the system to collapse to a multinomial for the different splits of activity types, conditioned on a negative binomial for total activity participation. As a point of comparison, the third model specification in Chapter Four is run with the same-gamma-term assumption and without it (*i.e.*, as a system of independent Poissons); the correlations of residuals resulting from this later specification are investigated.

Implication of the Assumption of Multiplicative Error Component for Indirect Utility

The assumption that households having the same observed characteristics can have different long-run, optimal rates of activity participation according to a gamma distribution implies something about the variation across these households' indirect utility values. Since the average optimal rates, X_i^* 's, are derived via Roy's Identity, the gamma error component must come out of one or both of the derivatives which are used: the derivative of indirect utility with respect to travel time or that with respect to available time. One stochastically convenient theory is that the households, although well aware of their marginal utility of available time, observe their travel-time environment with some error such that the travel time they perceive is really distributed like the inverse of a gamma random variable around the "true" or observed travel time. Another possibility is that the travel time data observed and used to estimate the models provide the mean travel times within different neighborhoods, but the actual, household-specific travel times within that neighborhood are inversely gamma distributed around that neighborhood's mean. These stochastically equivalent assumptions translate to the following:

Let $t_{i,obs'd} = \varepsilon t_{i,felt}$, where $\varepsilon \sim \text{Gamma}(m, m)$.

Then $v = v(\bar{t}_{felt}, T, Y) = v(\bar{t}_{obs'd}, \varepsilon, T, Y)$, and (Chapter

$$X_i^* = \frac{\left(\frac{dv}{dt_{i,felt \text{ or actual}}} \right)}{\left(\frac{dv}{dT} \right)} = \frac{\left(\frac{dv}{dt_{i,obs'd}} \times \frac{dt_{i,obs'd}}{dt_{i,felt \text{ or actual}}} \right)}{\left(\frac{dv}{dT} \right)} = \frac{\varepsilon \left(\frac{dv}{dt_{i,obs'd}} \right)}{\left(\frac{dv}{dT} \right)}$$

Three:-26)

The above is likely the simplest method of integrating back from the error assumption on the observed demand system to the unobserved indirect utility function;

but other processes also could lead to the multiplicative gamma specification in the demand system. However, no matter what method of accommodating the gamma error specification at the level of the indirect utility function, welfare analysis is likely to be complicated and one should be wary of using the indirect utility functions and their inverted expenditure functions as originally specified, without explicitly acknowledging the stochastic components. The need for averaging a measure like equivalent variation over its unobserved, random components is generally important when the error does not enter additively with a mean of zero, since expectation is a linear operation. McFadden (1996) provides a nice discussion of this situation in discrete-choice models.

Note that a simulation-of-likelihoods estimation method would allow one to estimate a more distributionally complex model, without imposing the same gamma error term on all the Poisson rates at the level of observed demand. For example, one could begin by specifying the unobserved heterogeneity to occur in the indirect utility function and then see what that implies for the demand functions. An error term which arises additively and which is independent of travel times and the total-time-budget variable would *not* show up in the demand equations for number of trips (though it would typically be relevant in average welfare estimates). A more reasonable assumption would be of random parameters, *i.e.* of unobserved differences in household's preference structures; such an assumption would be very similar in nature to Train's (1996) and Mehndiratta's (1996) random-parameters logit models.

Data Set

The 1990 Bay Area Travel Surveys (BATS) was used for the empirical portion of this research. They detail trip-making of over 10,000 households in the San Francisco

Bay Area for periods of one, three, or five workdays. While the BATS are not surveys of activities, *per se*, BATS households' activity participation can be inferred from the trip purposes and the start and end times of consecutive trips. Since survey lengths vary across the BATS households, a time component is included explicitly in the likelihood. For example, if the Poisson rates, X_i^* 's, are for a one-day period, one must multiply them everywhere in the Poisson specification with the variable *days*, where “*days*” is the number of days for which the household was observed participating in activities. In the multivariate negative-binomial likelihood equation used here (shown in Equation 3-23) the multinomial portion of the likelihood remains the same, but the negative binomial's probability for total observed trips (X_T) changes. The expectation of a multi-day survey's total number of trips, X_T , is *days* times the single-day level, but the variance increases more than linearly (unlike a Poisson). The parameter p^* for a multi-day survey must be replaced by $m/(m + (days \times X_T^*))$ everywhere, so the variance remains equal to $\mu + \alpha\mu^2$, but the mean, μ , has become $X^* \times days$. Note that the process remains a negative binomial with the same gamma term, as long as one assumes that the heterogeneity for an household is constant for each of the days the household is surveyed.²⁵ The BATS data set is described in more detail in the Appendix section A-3, and a definition of all variables used is provided in Table 3-2.

Definition of the Consumption Space

Interestingly – but not too surprisingly, investigations for this research indicate that access times for activities distinguished simply by type or purpose (*e.g.*, dining versus recreational) are *endogenous*, given a household's location. In other words, even given

their relatively fixed residential locations, households can, to a significant extent, choose how long they spend accessing different types of activities.

Initially, per-trip travel times and distances for the San Francisco Bay Area were regressed on a wide variety of urban form variables (*e.g.*, accessibility to all jobs by automobile, accessibility to sales and service jobs by walking, entropy across the proportions of half-mile-radius-neighboring land uses, mix of neighborhood land uses, and developed-area densities, as defined in Kockelman 1996 & 1997) in order to instrument for the travel times and costs associated with different locations, after controlling for trip purpose/activity type. The predictive power of these models was minimal; for example, ordinary least squares regressions of per-trip travel times and distance on the large set of detailed urban-form variables produces R-squareds of just 0.002 and 0.016, respectively. The R-squared results of OLS regressions controlling for mode and trip type are shown in Table 3-1, where it appears clear that such models are effectively useless for prediction. This set of access measures does not predict statistically significant reductions in per-trip travel times or distances, even after controlling for mode and/or trip purpose!²⁶ The first of these two general results is in agreement with the combination of Zahavi *et al.*'s constant travel-time-expenditures hypothesis (Zahavi 1979a & 1979b, Zahavi and Talvitie 1980, Zahavi and Ryan 1980) and the travel-time-inelastic nature of trip demands described by Ortúzar and Willumsen (1994) and Hanson and Schwab (1987) (as mentioned in Chapter Two's literature review). As a result of all these indications, the possibility of instrumenting for the travel costs needed for the system-of-demands approach by using characterizations of a household's environment appears very remote.

The evidence suggests that people travel further than they need to; this may very well be because they wish to expand their choice set of activity sites and thereby increase the expected “quality” of the activity they do engage in, at their chosen sites. For example, while one probably will travel only to the closest of a very specific activity type (such as eating out at a McDonald’s), one will not often travel to the closest dining establishment.²⁷ As long as the marginal value of travel time plus the monetary cost of travel remains below the marginal value of increased opportunities brought about by traveling further, people can be expected to lengthen their journeys.

Table 3-1: Regressions of Travel Time and Distance on Measures of Urban Form

TIME Regression's Dependent Variable:	R-Squared
PV-Trip Travel Time for Personal-Business Trips	0.006
PV-Trip Travel Time for Social Visit Trips	0.004
PV-Trip Travel Time for Dining/Eat Trips	0.006
PV-Trip Travel Time for Recreation Trips	0.009
PV-Trip Travel Time for Grocery/Food Shop Trips	0.002
PV-Trip Travel Time for Non-Food Shopping Trips	0.004
<hr/>	
Non-PV-Trip Travel Time for Personal-Business Trips	0.020
Non-PV-Trip Travel Time for Social Visit Trips	0.073
Non-PV-Trip Travel Time for Dining/Eat Trips	0.008
Non-PV-Trip Travel Time for Recreation Trips	0.009
Non-PV-Trip Travel Time for Grocery/Food Shop Trips	0.061
Non-PV-Trip Travel Time for Non-Food Shopping Trips	0.033

DISTANCE Regression's Dependent Variable:	R-Squared
PV-Trip Travel Distance for Personal-Business Trips	0.011
PV-Trip Travel Distance for Social Visit Trips	0.011
PV-Trip Travel Distance for Dining/Eat Trips	0.010
PV-Trip Travel Distance for Recreation Trips	0.016
PV-Trip Travel Distance for Grocery/Food Shop Trips	0.014
PV-Trip Travel Distance for Non-Food Shopping Trips	0.020
<hr/>	
Non-PV-Trip Travel Distance for Personal-Business Trips	0.011
Non-PV-Trip Travel Distance for Social Visit Trips	0.026
Non-PV-Trip Travel Distance for Dining/Eat Trips	0.009
Non-PV-Trip Travel Distance for Recreation Trips	0.006
Non-PV-Trip Travel Distance for Grocery/Food Shop Trips	0.027
Non-PV-Trip Travel Distance for Non-Food Shopping Trips	0.028

Unfortunately, in virtually all existing travel data sets there is no information regarding the *quality* of activities pursued. For example, except for general activity-purpose categories, there are no survey questions regarding the grade or class of establishments visited or the unit prices of activity consumption. To deal with this lack of detail in the data, one may choose to segment activities by some measure of quality, relative to an observation-specific origin (*e.g.*, the household's home location). One such measure is the number of choices a trip-maker has, which increases with time and/or distance traveled²⁸. Thus, the number of *jobs* has been used here to distinguish activity

quality for trip types. Discretionary trips to locations within bands of 60,000, 300,000, 900,000 and two million jobs serve as the four types of trips in the models investigated.

The variables derived from the travel surveys and from travel-time and employment data are described in Table 3-2. The focus is on the household as a unit, rather than intra-household trade-offs and decisions. So the total time available and income budget²⁹ apply to the entire household, and the *sum* of activity engagements over the households' members is the observed demand. Travel times for the four good groups distinguished in the data set represent average travel times to access the four different iso-opportunity contours from a household's *home location*.

In addition to the number of jobs, the amount of *land* area in different uses can be used to measure opportunity levels, particularly for activities like outdoor recreation. One may also wish to include trip-making from non-home trip-making bases, such as work. However, the size of the demand set may increase multiplicatively; for example, trip-making to four iso-opportunity contours from the home and work bases of a one-worker household across all trip types would mean eight different demand types (and eight different travel times upon which to apply Roy's Identity).

Table 3-2: Description of Variables Used**Dependent Variables:**

Number of Person Trips - Number of trips by surveyed household members (*i.e.*, those members aged five and over) in the region on the survey days(s); does not include trips to home location

to:

Discretionary Activities, including:

Medical/Dental Activity

Social Visit

Dining - Eat meal

Recreation

Grocery Shopping

Non-Food Shopping

Non-Discretionary Activities, including:

Work and Work-Related Activities

Personal Business Activity

Education

Other - Child care, serve passenger, change travel mode, other reason

Explanatory Variables:

Income "Y" - Pre-tax household income in 1989

Marginal Wage "w" - Estimate of average wage per hour for household (\$/hour)

= $\text{Income}/50/(40 \times \# \text{full-time workers} + 20 \times \# \text{part-time workers in household})$

Discretionary Time "T_d" - Estimate of non-work-related and non-school time in a day available to a household's members age five and older (hours/day)

= $24 \times \text{Household Size} - \text{Time in work-related \& school activities}$

Total Available Time "H" - Household Size (#members age five & older) \times 24 hours (hours/day)

Travel Times to Iso-Opportunity Contours - Average total travel time by single-occupant vehicle during free-flow conditions to access successively further sets of opportunities, relative to household's home traffic analysis zone (TAZ); computed sequentially to nearest TAZs in turn (and exclusive of travel times to TAZs lying in other iso-opportunity contours).

Contour Levels constructed at: 60,000, 300,000, 900,000 and two million total jobs, cummulatively.

Another way of creating more detailed consumption sets involves segmenting iso-opportunity contours by modes of travel and by trip type. For example, the different modes available would generate different travel times, recreational trips' travel times would come from contours based on entertainment and other recreational employment, and shopping trip travel times would come from those based on sales jobs. Clearly, there will be very high substitutability among these classes, which can be accommodated using

a flexible system of demand equations, as described in the section on functional specification in this chapter.

Trip Chaining

The chaining of trips into “tours” is a common phenomenon which complicates the analysis of activity-participation demands by altering access times. Within the Bay Area Travel Surveys (BATS), 36.6% of home-based trip tours involve more than one non-home stop. 51.6% of the BATS person sample are full-time workers (and six percent are part-time workers); so a large percentage (12.3%) of the BATS trips are between work and some non-home purpose, and 5.74% of sequential trip pairs represent a tour from work and back (*i.e.*, they have “work” as the first trip’s origin and as the second trip’s destination). However, more than half (56.4%) of the chained trips are unrelated to work.

The marginal cost of adding a stop to one’s tour can be relatively small, if that stop is anywhere near the general path between primary activity locations. The nature of home-based tours found in the data set was investigated and it was found that most tours contain a *single* major leg from home, even though the average number of stops per tour is close to three, at 2.71. The mean and median total travel time per tour³⁰ are 21.4 and 11.9 minutes, respectively, across all tours made (which number almost 40,000); and the travel time from home to the furthest destination accessed in each tour (with “furthest” measured by travel time) are 9.6 and 5.2 minutes. Thus, a single leg of the tour accounts for about 45-percent of the tour’s travel time, which can be taken to mean that about 90-percent of the tour time is spent accessing a single destination. These results suggest that a single destination accounts for much of the tour’s travel time, while additional stops are relatively marginal in travel time cost.

Weekday non-work trip-making by workers tends to not be very complicated. For example, in their 1981 data set of Nagoya, Japan, workers, Kwakami and Isobe (1990) found that of the 15% of workers making non-work trips before or after work, only 2.3% made a trip on their way to work, 6.7% made one stop on their way home, and 4.2% made a single trip after arriving home. Kwakami and Isobe's simulation results, which took work time to be exogenous, predict that as the time spent working during the day falls, workers travel further per non-work trip in addition to making more non-work trips; this is likely due to the loosening of the discretionary-time budget constraint and the ability to consume a higher quality of activity by traveling further. In a related study of trip-chaining by workers, Kitamura *et al.* (1990) found that mid-chain stop locations between work and home "tend to cluster along the line segment than connects the home and work bases as commuting distance increases" (1990, p. 153); they also found that intensity "peaks" of stop location form toward the home and work ends of the segment.

These same sorts of tour characteristics were found in the analyzed data sets, leading to a specification which accommodates chaining behavior; further description of this model's definition of demands, along with empirical results, are presented in the following chapter.

ENDNOTES:

-
- ¹ The chaining of trips as well as the linking of activities at a single opportunity site (e.g., shopping and entertainment at a shopping center) complicate the analysis since access times can be reduced and, to a significant extent, endogenized. To accommodate this effect, one can introduce variables for the possibilities of linking trips and/or model endogenously the number of chained trips to better account for the impact such travel behaviors have on a household's choice set and utility. Chapter Four presents and estimates a model, using the Type 3 model specification described later in the current chapter.
- ² In-home activities are included in the vector of activities, A , but they have zero travel time and zero travel costs.
- ³ One should be aware that these formulations assume a two-constraint case. If other constraints apply and lead to corner solutions for variables such as work time, the specified model will be insufficient and equations such as 3-7 and 3-8 will not apply. Moreover, if a household's *perceived* wage or marginal return to an extra hour of work is unobservable, one may need to construct a model which accommodates this fact.
- ⁴ For this regression model, observed work time is the amount of hours spent at work and in work-related activities during the survey day(s) for each household with workers, across its members. Table 3-2 defines the wage variable, w , used in these regressions (which is estimated using income and the number of full- and part-time workers) as well as the total time available to the household (per day), H .
- ⁵ The median wage estimate for the household sample is \$15.58/hour in 1990 pre-tax dollars, and this was substituted for the wage variable, w , in Equation 3-8. Substituting -0.000145 hours/dollar and +0.0808 hours/hour for the derivatives of work time with respect to unearned income and total time, respectively, yields a bias estimate of +9.04 percent.
- ⁶ There is significant debate as to the validity of constraints implied by the theory of demand, such as homogeneity and symmetry (of the substitution matrix). For example, empirical tests of aggregate, serial demand systems by Deaton and Muellbauer (1980a) and Christensen *et al.* (1975) reject these restrictions. Deaton and Muellbauer suggest that their model's "rejection of homogeneity may be due to insufficient attention to the dynamic aspects of consumer behavior." (1980a, p. 312) They suggest adding time-trend variables, lagged values, and stocks as explanatory variables. And Polak and Wales (1978 & 1980) cite the importance of analyzing stocks, rather than flows, of durable goods, which requires a rigorous dynamic treatment of behavior. Thus, one should use care in the analysis of some of the goods of interest here, for example the number of automobiles or size of home; and any results using basic methods of analysis for such goods should be considered with some caution. However, the data likely to be available for the research at hand will not be serial, so there can be little consideration of these effects. Additionally, the current research will not experience the problems of high collinearity in prices and aggregation biases, which aggregate serial data are prone to (e.g., Deaton and Muellbauer 1980a, Barten 1977).
- ⁷ Note that the log-likelihood equation must be re-written to accommodate certain zero-level demands, since one or more multinomial probabilities will equal zero and the logarithm of zero is undefined. If a demand level is optimally zero, the multinomial's choice set collapses, eliminating the zero-level possibilities.
- ⁸ An $\mathbb{R}^n \rightarrow \mathbb{R}^1$ function $f(x)$ is convex/concave if $f(\alpha x_1 + (1-\alpha)x_2)$ is less/greater than or equal to $\alpha f(x_1) + (1-\alpha)f(x_2)$, for $\alpha \in [0,1]$. Therefore, concavity of the money-expenditure function implies that the amount of expenditures needed to achieve a given utility level is no lower at a set of average prices than at two initial sets of unbalanced prices. Very similarly, *quasiconvexity* of the indirect utility function implies that the upper contour set of a such a function is a convex set; so, if two sets of prices, given income, lead to the same level of indirect utility, any average of those prices can be no better for the consumer. In a money-expenditures setting, this condition can be written as the following: the set of prices P such that $k \geq v(P, Y)$ is a convex set.
- ⁹ Since added consumption of some activities (without binding budget constraints) may require so much added travel time and produce a net negative impact on utility, the commonly assumed property of "nonsatiation" or "monotonicity" is not likely to always be viable here. After a certain point, strictly

more of an activity is not necessarily a good thing. This is important to note because even if preferences are complete, reflexive, transitive and continuous (as described in Varian 1992), there may not exist a continuous utility function which represents those preferences.

- ¹⁰ Concavity implies that the symmetric matrix of second derivatives has only non-positive diagonal terms (i.e., $d^2e_T/dt_i^2 \leq 0 \forall$). The matrices of second derivatives in prices of the expenditure function which result from estimation of the Type Two model were computed for all 10,834 households, separately. Rather interestingly, for all 10,834 households, the first three of the four diagonal terms were positive and the fourth was negative; this result may suggest a condition more closely resembling *convexity* in discretionary-time expenditures with respect to travel times to the first through third contours!

Convexity of the indirect utility function is a little easier to examine, since indirect utility is an immediate product of the models' parameter estimation. In the case of the modified translog model specifications with and without constant terms, the diagonal terms of the matrix of second derivatives (with respect to travel times) is β_{ii}/t_i^2 . These should be non-negative if the function is convex, but one finds that the β_{ii} (which determine the sign of this derivative) are estimated with negative signs for two to three of the demand types in the four models of this type (as shown in Tables 4-2a, 4-3a, 4-4a, and 4-5a). Thus, convexity of the indirect utility function is not apparent in travel times.

- ¹¹ With monetary expenditures, people simply "hand over" their money; it is an immediate transaction, not requiring effort at the moment of use and affecting the spender only in how much money he/she has left over.
- ¹² One word of caution as to expectations of non-satiation here: Activity participation can be tiring and eventually undesirable for an individual, so non-satiation may not exist in terms of out-of-home time expenditures alone. Having more goods is easy when compared with experiencing activities, since the former requires storage space (or friends who are willing to cart away your belongings). Thus, the viability on non-satiation in activity participation may not make sense, particularly at the level of the individual. Still, one must experience his/her entire day (in contrast to not having to spend one's entire income); restorative activities such as sleeping help make up for energy, and summability across *all* time expenditures is clearly a valid condition.
- ¹³ Note that strong separability allows a monotonic transformation of the direct utility function to produce an equivalent direct utility function which is *explicitly additive* in the sub-utility functions.
- ¹⁴ Even if one were to make the highly heroic and unreasonable assumptions that discretionary time expenditure is homogeneous of degree one in access times, time-compensated activity demands are homogeneous of degree zero in access times, and the sum of the derivatives of the various time-compensated time-in-activity demands [$I_{i,T}(t, P, Y, u)$'s] with respect to a single access time equal zero, one can still not argue that the time-compensated activity demands are the first derivatives of the time-compensated expenditure function. The following equations make this apparent:

$$\begin{aligned} \text{Even if } \sum_i \frac{de_T(\bar{t}, \bar{P}, Y, u)}{dt_i} &= \sum_i h_{i,T} + \sum_j \left(\sum_i \frac{dh_{j,T}}{dt_i} t_i \right) + \sum_i \left(\sum_j \frac{dT_{j,T}}{dt_i} \right), \\ &= \sum_i h_{i,T} + 0 + 0 \text{ and} \\ \sum_i \frac{de_T(\bar{t}, \bar{P}, Y, u)}{dt_i} t_i &= e_T(\bar{t}, \bar{P}, Y, u) = \sum_i h_{i,T} t_i + \sum_i T_{i,T}, \\ \text{the solution to these two equations is NOT } \frac{de_T}{dt_i} &= h_{i,T}. \end{aligned}$$

If the time-compensated activity demands were in fact the first derivatives of the time-compensated expenditure function, one then could impose Slutsky symmetry on the estimable/identifiable demands, derived from one's indirect utility function. Symmetry of the Slutsky matrix in the common problem formulation (i.e., one with purely a monetary expenditure constraint) is generally very useful because, together with a condition for negative semi-definiteness (i.e., concavity of expenditures in prices), it guarantees integrability (see, e.g., Jorgenson and Lau, 1979); this means that these two conditions guarantee the existence of an indirect utility function that could generate the demand system estimated.

However, as discussed in a prior section, one cannot assume concavity of the time-expenditure function, thanks to travel time's direct effect on one's welfare. And, without the expenditure derivatives producing compensated demands here, one cannot logically impose symmetry on the Slutsky relation, as illustrated by the following relation.

$$\text{Given: } h_{i,T}(\bar{t}, \bar{P}, Y, u) = X_i^*(\bar{t}, \bar{P}, Y, e_T(\bar{t}, \bar{P}, Y, u)).$$

$$\text{Without } \frac{de_T}{dt_i} = h_{i,T},$$

$$\frac{dh_{i,T}}{dt_j} = \frac{dX_i^*}{dt_j} + \frac{dX_i^*}{dY} \frac{de_T}{dt_j} \neq \frac{d^2 e_T}{dt_i dt_j}$$

$$\left(\text{and } \frac{dh_{i,T}}{dt_j} \neq \frac{dX_i^*}{dt_j} + \frac{dX_i^*}{dY} X_j^* \text{ either} \right).$$

- ¹⁵ Changes in what is known as “consumer surplus” are a special case of equivalent and compensating variation; the value of change in consumer surplus lies between these two estimates and is defined as the integral of demand (vectors) over a change in prices. (Varian 1992)
- ¹⁶ The argument for a money measure of benefits/disbenefits is that it best accommodates society's values of benefits/disbenefits to everyone over a variety of impacts experienced. For example, one can argue that the time of high-income persons should be more valued by society than that of other persons, since their elevated incomes are typically due to their higher-valued labor-market activities; in other words, society optimally trades their time at a higher rate.
- ¹⁷ Equivalent variation was originally defined by Hicks (1956) as the difference in expenditures at a reference utility level, rather than at a reference price level. However, when budget levels are held constant, the expenditures under the before and after scenarios are the same, so the definitions provided in Equation 3-13 are then equivalent to those given by Hicks. The definitions used here can be found in Varian (1992); they are the negative of Deaton and Muellbauer's (1980) definitions, when available income levels are unchanged.
- Note also that Equation 3-13 assumes income (Y) and available time (T) are exogenous. With income and work time endogenous instead, one would write the lower set of equalities with unearned income, Y_{un} , in place of total income, Y , and both sets of equalities with an added argument of wage, w .
- ¹⁸ Flexibility to a certain order means that any set of values for that order of derivatives can be achieved (with a single, variable set of parameters).
- ¹⁹ Christensen *et al.* (1975) note that the translog provides a second-order approximation to *any* (typical) direct or indirect utility function; thus, the resulting demand functions provide a *first*-order approximation to any system. The same is true of the Almost Ideal Demand System (Deaton and Muellbauer, 1980a) and the generalized Leontief (Lau 1986).
- ²⁰ An elasticity of substitution is the dimensionless version of the derivative of the ratio of two goods with respect to their marginal rate of substitution (MRS). MRS is effectively a utility-constant measure of substitution between two goods. The following equations illustrate this definition:

$$MRS_{ij} = - \frac{du/dX_j}{du/dX_i} = \text{Rate of substituting } X_i \text{ for } X_j \text{ to keep utility constant.}$$

$$\text{Elasticity of substitution}_{ij} = \frac{d\left(\frac{X_i}{X_j}\right)}{d(MRS_{ij})} \times \frac{MRS_{ij}}{\left(\frac{X_i}{X_j}\right)}.$$

- ²¹ The parameterization of the distributional assumptions, which are described in the following section – Statistical Specification, adds additional parameters requiring estimation to each of the model types discussed in this section.
- ²² Second-order flexibility is not fully realized with the current specification, because it does not include a $\log(Y)^2$ term (which would not be identifiable from the demand system estimated). However, the ability of this functional specification to capture substitution relationships is likely to remain superior to that of the other most popular form for such models, that of the Leontief. As discussed earlier, in the section on Functional Specification, substitution is important in the empirical analysis pursued here because the demand sets modeled in Chapter Four differ by quality of destination, rather than by activity type or purpose. So significant substitutability is anticipated.
- ²³ Some researchers are working with activity-duration models, acknowledging that activities endure separately rather than overlap (*e.g.*, Ettema *et al.* 1995b, Bhat 1996), but micro-economic or other rigorous behavioral linkages are missing from these models. For example, using Weibull-based hazard functions, Kim (1994) models activity duration separately from trip generation but simultaneous with trip travel time.
- ²⁴ Recall that a gamma random variable can be thought of as the sum of m independent exponential random variables, with each exponential sharing the rate λ . Thus, a constant p (which is less than one) times a gamma can be thought of as the sum of m independent exponentials, each with a longer rate of λp ; so the inverse of the rate (which is the average time between events) is shorter and the sum of the exponential times between events is shorter.
- ²⁵ One can think of the sum of “ t ” days worth of a household’s travel data as being the sum of “ t ” Poisson random variables, each with the same mean over the population having this household’s characteristics – and each interacted with the same gamma term, which represents the heterogeneity within a population of similar observed characteristics. As is well known, the sum of “ t ” *independent* Poissons is Poisson; the “ t ” days of Poissons considered here for a single household (indexed by n) are independent when one conditions on knowing the gamma error term and the mean rate (over the population with this household’s characteristics). Thus:

$$\sum_{s=1}^t (X_{i,n,s} | \lambda_{i,n} = X_i^* \varepsilon_n) \sim \text{Poisson}(\lambda_{i,n}(t) = tX_i^* \varepsilon_n), \text{ and}$$

$$\sim \text{Neg. Binomial} \left(m' = m, P' = \frac{tX_i^*}{m} \right).$$

- ²⁶ More accessible environments *do* appear to lower automobile ownership, reduce *total* travel distances, and shift mode of travel to slower modes (such as bus and walking), as described by Kockelman (1996 & 1997).
- ²⁷ The aggregation of trips into the broad categories asked for in surveys (*e.g.*, recreational vs. shopping trips) obscures the subtle but important distinctions across activities and renders travel times endogenous. In theory, if one had a large enough sample of observations, one could model a system of activity demands where essentially *every* destination-and-mode (and time-of-day!) combination was a possible “good” to be consumed by everyone residing in the region. Travel costs would be fairly obvious (given inter-zonal travel times and distances), and with regional data one could ensure that an individual’s responses to a limited survey would not bias his/her vectors of travel costs while implicitly controlling for quality- and price-of-activity differences.
- ²⁸ It merits mention that much of Zahavi’s work (*e.g.*, Zahavi 1979a) measures utility by total distance traveled, essentially asserting that it is access to opportunities that determines one’s welfare – an idea similar to those discussed here. However, distance may be a seriously flawed utility measure; for example, who can say with certainty that several short-distance journeys are preferred to a few longer

journeys, just because the first choice involves less distance? The approach advocated in this dissertation allows the data to interpret preferences far more flexibly than a distance metric.

²⁹ The survey used, like most surveys of significant size, does not provide income-per-worker or hourly wage information, so analysis at the individual level would not have been feasible.

³⁰ The travel times referred to in this section are not as reported by survey respondents (who tend to report times in increments of five minutes); instead, they come from interzonal free-flow automobile-travel times provided by the region's metropolitan planning organization, which is the Metropolitan Transportation Commission in the case of the Bay Area Travel Surveys.

Chapter Four: Empirical Estimation and Model Validation

Estimation Techniques

Likelihood Maximization

The likelihood maximization relies on S-Plus statistical computing software (produced by MathSoft Co.), using a model/trust-region approach described by Gay (1983). Due to the constraint of strict positivity on the Poisson rates (X_i^*) and the complexity of the likelihood equations' first and second derivatives¹, the algorithm employs its own, numerical approximations to the derivatives utilized in a quadratic approximation to the likelihood function over an iterative series of neighborhoods it “trusts”².

Acquiring Starting Parameter Values

When a model's regression equations are not linear in unknown parameters, as is the case for the functional forms considered here, the choice of a method to achieve starting parameter estimates can be quite difficult. This is particularly true when negative estimates of the marginal utility of time and negative demand estimates are effectively disallowed (as discussed in the section on non-negativity in Chapter Three). In the present specification, one could run an iterative maximum-likelihood search procedure on each demand equation individually, using negative binomial stochastic assumptions; however, such individual regressions require their own sets of feasible start values for many of the unknown parameters, along with a likelihood search, without guarantee of unique convergence.

To begin the maximum-likelihood search routines in the models estimated here, several different sets of parameter estimates were attempted for each system of equations as a whole, with theory often guiding the choice of sign (*e.g.*, the derivative of indirect utility with respect to each travel time should be negative, so signs on the starting values of prominent coefficients of travel times were chosen appropriately). Section A-5 of the Appendix provides actual code and starting values used in typical program files. Fortunately, not too many parameter sets were required to find a feasible set from which to begin the iterative search in each model. Additionally, several very distinct but feasible starting parameter sets were used on each model, and the likelihood values of their solutions were compared in an attempt to avoid convergence to a local, rather than global, maximum.

Variance-Covariance Estimation

The model complexities complicate the estimation of errors in estimation, so the specification of the log-likelihoods' gradients and Hessians is not elementary. To facilitate computations, the Berndt *et al.* (1974) method (BHHH) of Hessian estimation is used, requiring only gradient information. The variance-covariance matrix for the parameter set is thus estimated applying the following equation, but employing the parameter estimates generated from the sampled observations (Greene 1993, p. 326):

Asymptotic Variance – Covariance Matrix =

$$AVC(\hat{\theta}) = E \left(\left[\sum_{n=1}^{n=N} w_n w_n' \right]^{-1} \right) = E \left([W'W]^{-1} \right), \quad (\text{Chapter$$

*where $\hat{\theta}$ is the MLE, $w_n = \text{Gradient of Likelihood of } n\text{'th observation,}$
and $W = \text{Matrix whose rows are the } w_n \text{'s.}$*

Four:-1)

Results to be Estimated

To limit the problem size while analyzing a variety of different functional specifications and illustrating use of the method, only four demand types were distinguished here; and they are used with all four model types (Type 1 through Type 4, as specified in Chapter Three). Specifically, the four demand types used in four sets of results are the number of discretionary trips (*i.e.*, non-work, non-education, and non-serve-passenger trips) made to each of four iso-opportunity contours using the Bay Area Travel Surveys. The contours are defined, as discussed in Chapter Three, by (free-flow, automobile) travel times from a household's specific "neighborhood"/traffic analysis zone (TAZ) out to contours of 60,000, 300,000, 900,000, and two million jobs in the region. The fifth set of results presented here also relies on just four distinct demand types across the four all-jobs iso-opportunity contours described, and it uses the Type 3 model specification; however, it explicitly accommodates trip chaining by focusing on activity tours, rather than individual stops/single activities. These five sets of results are described here now.

Results of Type One Model: Modified Linear Expenditure System

The estimates for a Type 1 model of discretionary-activity participation (represented by Equation 3-14) are shown in Table 4-1a; these represent the likelihood-maximizing set over the full sample (N=10,834) after starting from a variety of parameter values. The median levels of first-order estimates of demand elasticities and the value of time for this model are shown in Table 4-1b.

Even though the great majority of the parameter estimates appear to be highly statistically significant in this model, the value-of-time and aggregate mean estimates

differ significantly from expectations. For example, the quartiles of the value-of-time estimates are all negative across the sample; these estimates are presented in Table 4-7, along with value-of-time estimates from other models, in the section titled “Discussion of Value of Time Estimates”, toward the end of this chapter.

There are other clues that the model is far off. For example, the average X_1^* across households is 2.67 trips/day, which is more than twice the observed mean of 1.09 trips/day, and the ratio of the sum of X_2 to X_3 is 2.25 while the model predicts 1.55. Thus, this model does not appear to be sufficiently flexible for our purposes, and its results should not be taken too seriously.

Table 4-1a

Parameter Estimates of Modified Linear Expenditure System, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours, using the BATS data set

Parameter	Final Estimates	Standard Errors	T-Statistics	P-Values
α	0.902	1.94E-01	46	0.000
α_1	0.0929	3.26E-02	28	0.000
α_2	0.160	4.68E-02	34	0.000
α_3	0.0947	2.16E-02	44	0.000
α_4	0.0580	3.15E-02	18	0.000
β_{1Y}	-4.39E-05	2.24E-06	-20	0.000
β_{2Y}	-5.06E-05	1.74E-06	-29	0.000
β_{3Y}	-3.21E-05	3.06E-07	-105	0.000
β_{4Y}	-2.99E-05	6.23E-07	-48	0.000
β_{11}	-2.18E-01	3.79E-02	-5.8	0.000
β_{12}	-6.72E-02	1.52E-02	-4.4	0.000
β_{13}	-8.91E-02	7.19E-02	-1.2	0.215
β_{14}	2.69E-02	5.50E-02	0.5	0.625
β_{22}	5.14E-02	1.87E-02	2.7	0.006
β_{23}	-1.15E-01	1.66E-02	-7.0	0.000
β_{24}	5.49E-02	7.41E-03	7.4	0.000
β_{33}	1.95E-01	2.54E-03	77	0.000
β_{34}	-5.93E-02	2.10E-03	-28	0.000
β_{44}	9.96E-03	3.66E-04	27	0.000

$L = -46,696$

$N = 10,834$

Table 4-1b

Economic Implications of Modified Linear Expenditure System, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours, using the BATS data set

VALUE ESTIMATED:	Median of Sample
Discretionary-Time Elasticity of Demand:	
Immediate Zone Demand	0.389
Near Zone	0.672
Moderate Zone	0.572
Far Zone	0.304
Income Elasticity of Demand:	
Immediate Zone Demand	0.044
Near Zone	0.163
Moderate Zone	0.411
Far Zone	0.696

Cross-Time Demand Elasticities:				
w/r/t Time of:	Immediate	Near	Moderate	Far
Immediate Zone:	-0.054	0.081	0.234	-0.191
Near Zone:	0.016	-0.725	0.575	-0.617
Moderate Zone:	0.284	0.814	-2.711	1.021
Far Zone:	-0.216	-0.658	-0.209	-0.541

Notes:

Demands are in trips per day, Discretionary Time is hours,
Travel Times are minutes, & Income is before-tax dollars.

Results of Type Two Model: Modified Translog

The translog functional form of Equations 3-15 and 3-16 was used, with the expectation that its larger parameter set would provide more flexible estimation and better results than that of the modified linear expenditure system. For purposes of parameter identification, this model's γ_{TY} parameter was fixed to equal positive one and the β_{ij} 's are constrained to equal β_{ji} 's. The parameter estimates are shown in Table 4-2a.

The log-likelihood value for this estimation is -46,431.6, but it cannot be compared with a no-information situation (where each X_i^* is modeled as a constant, independent of time and income information) or even a full-information situation (where *each*

household's $X_{i,n}^*$ is modeled as its *own* constant), because there are no free constants in this model – every unknown parameter is interacted with explanatory variables which vary across households.

Observe that the variance in the data is substantially reduced by using explanatory information. For example, the estimate of the overdispersion parameter, α , falls from 1.638^3 (for total trips) to 1.001 here, signaling a tighter distribution thanks to explanatory information and the model structure itself.

The average X_i^* estimates of this model fall much closer to the sample means than those of the modified linear expenditure system, suggesting much better accuracy in aggregate prediction. The average X_i^* 's are estimated to be 1.14, 0.66, 0.29, and 0.19, respectively, while the observed per-day average demands are 1.09, 0.62, 0.27, and 0.19. Theoretical considerations aside, this model appears to predict aggregate behavior well.

Table 4-2a

Parameter Estimates of Modified Translog Model, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours, using the BATS data set

Parameter	Final Estimates	Standard Error	T-Statistics	P-Values
α	1.00	0.00	123	0.000
α_1	1.53	1.29	1.2	0.238
α_2	-8.35	1.96	-4.3	0.000
α_3	-9.69	1.74	-5.6	0.000
α_4	-3.52	1.55	-2.3	0.023
β_{11}	-7.96	0.91	-8.7	0.000
β_{12}	-0.643	0.24	-2.7	0.007
β_{13}	-1.05	0.24	-4.4	0.000
β_{14}	-0.202	0.19	-1.1	0.286
β_{22}	-1.03	0.51	-2.0	0.041
β_{23}	-3.90	0.60	-6.5	0.000
β_{24}	0.858	0.29	3.0	0.003
β_{33}	8.47	1.02	8.3	0.000
β_{34}	-2.27	0.40	-5.7	0.000
β_{44}	1.95	0.48	4.1	0.000
γ_{1Y}	1.60	0.19	8.6	0.000
γ_{2Y}	1.15	0.16	7.2	0.000
γ_{3Y}	-0.576	0.10	-6.0	0.000
γ_{4Y}	-0.836	0.12	-6.8	0.000
γ_{1T}	-0.178	0.10	-1.9	0.064
γ_{2T}	0.913	0.23	4.0	0.000
γ_{3T}	1.44	0.26	5.6	0.000
γ_{4T}	1.50	0.28	5.4	0.000
γ_{TY} (fixed)	1.00	n/a	n/a	n/a

$L = -46,432$

$N = 10,834$

Economic Implications of the Type Two Model Results

This model's estimates' of elasticities are shown in Table 4-2b. Overall, this model's results appear reasonable, including the value-of-time estimates across the household sample (whose quartiles are provided in Table 4-7).

Discretionary-time elasticities are positive, as one would expect (*i.e.*, more discretionary time available to the household leads to more discretionary activity participation). Income elasticities, on the other hand, are positive for far and moderate zone activities but negative for closer activities; it appears that money is spent on access to consumption of activities further away, rather than near one's home. It is interesting that near trips are not found to be "inferior" with respect to time, but they are with respect to income (albeit to a minor extent). Note that these results are not definitive because part of this income effect is due to the purchase of automobiles, which effectively reduce per-trip marginal costs and travel times, and part is arguably due to the higher-income households having more specialized workers who must travel further on workdays and so undertake more activities at sites remote from home, but near their work locations. The presentation of the fifth set of model results more explicitly considers this question of trip chaining.

Finally, observe that own-travel-time elasticity estimates are generally negative as one would expect of economically "normal" goods, but not for the nearest zone's activity participation rates. And most cross-time elasticities are positive, suggesting the expected substitution effects (rather than complementarity), since the demands are only defined across "quality" here (*i.e.*, level of opportunity choice), not activity type (*e.g.*, social and personal business activities are less likely to be substitutable).

Table 4-2b

Economic Implications of Modified Translog Model, as applied to Discretionary Activity Participation across Four All-Jobs Iso-Opportunity Contours, using the BATS data set

VALUE ESTIMATED:	Median of Sample
Discretionary-Time Elasticity of Demand:	
Immediate Zone Demand	1.028
Near Zone	0.869
Moderate Zone	0.678
Far Zone	0.706
Income Elasticity of Demand:	
Immediate Zone Demand	-0.294
Near Zone	-0.208
Moderate Zone	0.086
Far Zone	0.122

Cross-Time Demand Elasticities:					
	w/r/t Time of:	Immediate	Near	Moderate	Far
Immediate Zone:		0.257	0.061	0.103	-0.033
Near Zone:		0.100	-0.891	0.496	-0.188
Moderate Zone:		0.242	0.832	-2.952	0.442
Far Zone:		0.047	-0.207	0.383	-1.446

Notes:

Demands are in trips per day, Discretionary Time is hours,
Travel Times are minutes, & Income is before-tax dollars.

Results of Type 3 Model: Modified Translog with Constants

The Type 3 model specification, a modified translog which includes constant terms, (as shown in Equation 3-18) has been applied here to two different demand sets. The first covers the discretionary-activity participation demands used in the previous two models; the second looks at home-based tours of all trip types. Both rely on the four iso-opportunity contours used previously, which count all job types as opportunities.

I. Discretionary Activity Participation

The estimated parameters for a Type 3 model across four divisions of discretionary-activity participation are shown in Table 4-3a. The estimate of the overdispersion parameter α has dropped to 0.938, suggesting that estimates are falling closer to observations than in the two previous models; and the demand estimates accurately estimate aggregate behavior. The average X_i^* 's are estimated to be 1.08, 0.62, 0.28, and 0.19, respectively, while the observed per-day average demands are 1.09, 0.62, 0.27, and 0.19.

An advantage of the translog specification with constant terms is that one can nest a no-information case within the specification. Table 4-3c provides a summary of the likelihood values resulting from a variety of specifications linked to this particular specification. The log-likelihood value for the no-information case⁴ is -47,688; in contrast, the log-likelihood of the full model is -46,218. The p-value for the hypothesis that the no-information model is the proper model, given the assumption that this third model specification encompasses the true model as a nested specialization, is 0.000; so one must reject this hypothesis (given the assumption).

Table 4-3a

**Parameter Estimates of Modified Translog Model with Intercept Terms,
as applied to Discretionary Activity Participation across Four All-Jobs Iso-
Opportunity Contours, using the BATS data set**

Parameter	Final Estimates	Standard Errors	T- Statistics	P-Values
α	0.938	0.020	47	0.000
μ_0	0.267	0.049	5.5	0.000
μ_1	0.930	0.234	4.0	0.000
μ_2	0.189	0.062	3.1	0.002
μ_3	-0.075	0.026	-2.9	0.004
μ_4	-0.033	0.030	-1.1	0.266
α_1	-0.831	2.306	-0.4	0.718
α_2	-1.066	3.325	-0.3	0.749
α_3	-1.112	2.676	-0.4	0.678
α_4	-0.944	5.668	-0.2	0.868
β_{11}	-4.191	1.159	-3.6	0.000
β_{12}	0.051	0.329	0.2	0.876
β_{13}	-1.453	0.414	-3.5	0.000
β_{14}	0.714	0.268	2.7	0.008
β_{22}	-0.123	1.287	-0.0	0.924
β_{23}	-4.464	1.007	-4.4	0.000
β_{24}	2.237	0.537	4.2	0.000
β_{33}	7.845	1.513	5.2	0.000
β_{34}	-2.842	0.709	-4.0	0.000
β_{44}	-0.200	1.658	-0.1	0.904
γ_{1Y}	1.491	0.292	5.1	0.000
γ_{2Y}	0.078	0.141	0.6	0.579
γ_{3Y}	-1.126	0.216	-5.2	0.000
γ_{4Y}	-1.187	0.230	-5.2	0.000
γ_{1T}	-0.813	0.146	-5.6	0.000
γ_{2Y}	0.262	0.310	0.8	0.398
γ_{3Y}	1.443	0.435	3.3	0.000
γ_{4Y}	1.888	0.521	3.6	0.000
γ_{TY} (fixed)	1.000	n/a	n/a	n/a

$L = -46,219$

$N = 10,834$

Table 4-3b

Economic Implications of Modified Translog Model with Intercept Terms, as applied to Discretionary Activity Participation across Four All-Jobs Iso-Opportunity Contours, using the BATS data set

VALUE ESTIMATED:	Median of Sample
Discretionary-Time Elasticity of Demand:	
Immediate Zone Demand	0.774
Near Zone	0.652
Moderate Zone	0.440
Far Zone	0.415
Income Elasticity of Demand:	
Immediate Zone Demand	-0.206
Near Zone	-0.039
Moderate Zone	0.144
Far Zone	0.132

Cross-Time Demand Elasticities:					
	w/r/t Time of:	Immediate	Near	Moderate	Far
Immediate Zone:		0.574	-0.014	0.124	-0.144
Near Zone:		0.020	-0.630	0.440	-0.303
Moderate Zone:		0.252	0.688	-2.608	0.384
Far Zone:		-0.073	-0.316	0.347	-1.231

Notes:

Demands are in trips per day, Discretionary Time is hours, Travel Times are minutes, & Income is before-tax dollars.

Economic Implications of the Type Three, Discretionary Activities Model Results

Estimates of various economic implications of this model are shown in Table 4-3b.

While elasticity signs and magnitudes appear to be in general agreement with those estimated for the previous, no-constant-terms model for these data, the value of time estimates differ dramatically. Even though this model is more flexible (and offers a significantly higher log-likelihood value, of -46,218 versus -46,432, for a difference of just five parameters), its value-of-time estimates are highly negative and thus contrary to expectations – in clear contrast to the value-of-time results for the previous model. This

model's value-of-time results are provided and discussed along with the value-of-time results for other models, toward the end of this chapter.

Model Comparisons: Case Example

Ten variations of this Type 3 system of discretionary-activity demands were run for purpose of likelihood comparisons. They differ primarily in their assumptions about the stochastic nature of the unobserved heterogeneity; but, the “no-information” and “full-information” variants make assumptions about the access to explanatory information. A broad comparison of likelihoods like this can be done for any of the modified-translog-with-intercept models since their specification includes intercepts, allowing them to be rigorous nested within a full-information case and over a no-information case. Note, however, not all of the ten variants are specialization's or generalizations of all others. All ten cases are described briefly and compared by means of their log-likelihood values in Table 4-3c.

Several interesting results emerge from these log-likelihood values. One is that the imposition of the same-gamma heterogeneity assumption significantly constrains the model as estimated. Without removing any estimated parameters yet allowing optimal activity-participation rates to vary independently of one another (given their population means, for a given set of household characteristics), the log-likelihood rises dramatically, from -46,218 to -42,592, given a difference of just three identifying restrictions⁵.

Following this change in stochastic specification, the overdispersion factor, α or “a” in the table, rises substantially as well, from 0.938 across the set of trips to 1.5 for “immediate” or very local trips, 2.2 for “near” trips, 2.8 for “moderate” trips, and 7.4 for “far” trips. These results suggest that the imposition of the same-gamma assumption for

characterizing the unobserved heterogeneity in optimal participation rates is too strong – at least for the data set used, where short-duration observations are likely to be more heavily influenced by chaining and intra-household trip coordination. More flexible models, which still provide for some correlation in unobserved information, should prove useful, though these likelihoods will probably require simulation.

Table 4-3c

Comparison of Log-Likelihood Values across Different Models based on the Modified Translog Model with Intercepts (analyzing Discretionary-Activity Participation across Four Iso-Opportunity Contours, using the BATS data set)

Model Description:	α	Log-Likelihood
i. Poisson (no unobserved heterogeneity, $\alpha=0$):	0	-52,052
ii. No Information (no explanatory variables):	1.012	-47,688
iii. MODEL AS ESTIMATED (same-gamma heterogeneity):	0.938	-46,218
iv. Semi-Independent Negative Binomials (with same overdispersion " α "):	2.13	-42,981
v. Totally Independent Negative Binomials (different " α "s):	1.5 to 7.4	-42,592
vi. Full Information (all optimal rates = observed rates, minimized variance):	0	-32,749
vii. Individually Estimated Negative Binomials for each Demand		
(without cross-equation parameter constraints):		
Immediate Trips:	1.51	-16,640
Near Trips:	2.20	-12,538
Moderate Trips:	3.04	-7,635
Far Trips:	6.18	-5,864
	Sum =	-42,677

Another way to look at these results is to compute the fraction of total likelihood difference, between the full- and no-information cases (cases vi and ii), that is “explained” by the specified model (case iii). This ratio is often referred to as a pseudo- R^2 , and it is 9.84% for this model⁶. This is actually higher than one might expect, given the disaggregate and short-term nature of the data. As mentioned in Chapter Two’s literature review, little if any research has found significant elasticities of trip demand with respect to travel times (Ortúzar and Willumsen 1994, Hanson and Schwab 1987). The percentage of explained variation in models of single-day trip-making, other than the

model proposed in this research, tend to be on the order of five percent (*e.g.*, Hanson and Schwab 1987). However, one should not put too much stock in this measure of explained variation; simply a longer survey period with fewer zero-level observations of demand would increase the percentage (or R-squared), without any change in parameter estimates, because the zeros do not diminish the full-information log-likelihood at all⁷.

Another point of interest is that the removal of the cross-equation parameter constraints does not do much for the log-likelihood. In case vii, where each demand is estimated completely independently of the others (but with the same general functional *form* given in Equation 3-18), 18 more parameters are being estimated than in the set of demand equations derived from case v's single indirect utility specification⁸ (of Equation 3-17); yet this only translates to a likelihood increase of $-42,592 - (-42,677)$, or 85. This difference still provides for a highly statistically significant likelihood ratio test of the difference⁹, but the magnitude of the difference appears small when compared with the differences other changes in the model create. For as many observations as there are in the data set (10,834 households times four dependent-value observations per household), it is not surprising that one would get a statistically significant result for most tests; what is surprising is the relatively small size of this difference for this particular test. It suggests that the derivation of a set of demands from a single indirect utility specification is not so presumptuous or limiting! However, the value-of-time results remain unbelievable, so the model structure is imperfect.

II. Modeling Tours Explicitly

Due to the prevalence of trip-chaining or "tour-making" in many observations of activity-participation, the incremental travel time faced by a household to pursue an

added activity can be substantially less than the round-trip travel time from home. Since most tours appear to involve a primary stop or leg with a significant travel time from the home location, the data set of demands constructed for analysis here is based on the number of tours made, with the furthest destination visited during the tour determining the tour type (according to which of the four distinct iso-opportunity contours the tour belongs). Since the number of tours that are exclusively non-work related is rather small (about 15,000 tours in the BATS data set) and many of the tours containing a work purpose also contain discretionary-purpose stops, tours of all types were assembled here for analysis, providing roughly 40,000 tours across the BATS households surveyed. Estimation results are given in Tables 4-4a and 4-4b.

Table 4-4a

**Parameter Estimates of Modified Translog Model with Intercept Terms,
as applied to Trip *Tours* to Four All-Jobs Iso-Opportunity Contours,
using the BATS data set**

Parameter	Final Estimates	Standard Errors	T- Statistics	P-Values
α	0.076	0.004	20.7	0.000
μ_0	-0.079	0.035	-2.3	0.022
μ_1	1.339	0.412	3.3	0.001
μ_2	0.339	0.120	2.8	0.005
μ_3	0.187	0.073	2.6	0.010
μ_4	0.127	0.097	1.3	0.189
α_1	-0.044	2.847	-0.0	0.988
α_2	-0.931	4.114	-0.2	0.821
α_3	-1.101	5.650	-0.2	0.845
α_4	-0.729	14.639	-0.0	0.960
β_{11}	-2.702	1.226	-2.2	0.028
β_{12}	1.231	0.510	2.4	0.016
β_{13}	-0.835	0.522	-1.6	0.110
β_{14}	0.812	0.437	1.9	0.063
β_{22}	-0.000	1.594	-0.0	1.000
β_{23}	-4.450	1.457	-3.1	0.002
β_{24}	3.012	0.920	3.3	0.001
β_{33}	7.866	2.696	2.9	0.004
β_{34}	-1.974	1.006	-2.0	0.050
β_{44}	0.445	4.695	0.0	0.924
γ_{1Y}	-0.032	0.119	-0.3	0.789
γ_{2Y}	-0.625	0.186	-3.4	0.000
γ_{3Y}	-1.506	0.378	-4.0	0.000
γ_{4Y}	-2.833	0.722	-3.9	0.000
γ_{1T}	1.225	0.575	2.1	0.033
γ_{2Y}	1.225	0.575	2.1	0.033
γ_{3Y}	2.343	0.862	2.7	0.007
γ_{4Y}	4.959	1.607	3.1	0.002
γ_{TY} (fixed)	1.000	n/a	n/a	n/a

$L = -48,469$

$N = 10,834$

Table 4-4b

Economic Implications of Modified Translog Model with Intercept Terms, as applied to Trip *Tours* to Four All-Jobs Iso-Opportunity Contours, using the BATS data set

VALUE ESTIMATED:	Median of Sample
Discretionary-Time Elasticity of Demand:	
Immediate Zone Demand	0.970
Near Zone	0.956
Moderate Zone	0.848
Far Zone	0.704
Income Elasticity of Demand:	
Immediate Zone Demand	-0.022
Near Zone	0.039
Moderate Zone	0.123
Far Zone	0.191

Cross-Time Demand Elasticities:				
w/r/t Time of:	Immediate	Near	Moderate	Far
Immediate Zone:	0.356	-0.139	0.015	-0.196
Near Zone:	-0.157	-0.400	0.397	-0.433
Moderate Zone:	0.052	0.407	-1.344	0.070
Far Zone:	-0.092	-0.259	0.093	-0.711

Notes:

Demands are in trips per day, Discretionary Time is hours,
Travel Times are minutes, & Income is before-tax dollars.

Observe that the overdispersion factor α is very close to zero here, suggesting less of a negative binomial and more of a Poisson distribution; this reduced value also suggests more stability in estimation thanks to less unobserved variation (assuming that a Poisson holds). So tour-making may be less variable than individual stop-making, which makes some sense given the fixed cost of getting ready to leave one's home and take care of business and activities outside one's home; the marginal cost of adding stops is relatively small once is already "out and about". Moreover, the same-gamma-error assumption may apply better here because gross estimates of the α terms¹⁰ for the different, individual

demands are much more stable; they are 0.84, 0.85, 0.70, and 0.52, rather than 1.7, 2.6, 3.4, and 7.6, as estimated for individual-activity (non-tour) demands.

This model's estimates of demand are reasonable predictors of aggregate behavior. The average X_i^* 's are estimated to be 1.22, 0.57, 0.37, and 0.28, respectively, while the observed per-day average trip-chain rates to the different contours are 1.14, 0.64, 0.38, and 0.32.

In Table 4-7, the value of times estimated for this model are negative, though they are not as extreme as those implied by the previous translog-with-constants model of discretionary trip-making. The travel-time elasticity matrix (shown in Table 4-4b) resembles earlier estimates of this matrix, but three of the four income elasticities are now positive.

Results of Type 4 Model: Modified Translog with Constants, Using Wage and Total Time Data

As discussed in Chapter 3, the work decision, and thus the income and discretionary-time budgets are likely to be made simultaneously with the discretionary-activity decisions. Thus, a model where these budgets are endogenous may prove useful. The Type 4 model accommodates these decisions by relying on wage and total-time data, rather than income and discretionary-time data, but the demand set analyzed is the same, four-zone discretionary-trip data set analyzed in the first three models discussed here. The results of this analysis are shown in Tables 4-5a and 4-5b.

Table 4-5a

Parameter Estimates of Modified Translog Model with Intercept Terms and Wage and Total-Time Information, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours, using the BATS data set

Parameter	Final Estimates	Standard Errors	T-Statistics	P-Values
α	0.935	0.020	47	0.000
μ_H	0.387	5.529	0.0	0.944
μ_1	3.212	1.537	2.1	0.037
μ_2	0.528	0.270	2.0	0.050
μ_3	-0.361	0.174	-2.1	0.038
μ_4	-0.192	0.117	-1.6	0.100
α_1	3.564	5.456	0.7	0.514
α_2	-3.332	6.137	-0.5	0.587
α_3	-4.253	5.436	-0.8	0.434
α_4	-2.461	11.412	-0.2	0.829
β_{11}	4.211	2.565	1.6	0.101
β_{12}	0.841	0.861	1.0	0.329
β_{13}	-4.060	2.088	-1.9	0.052
β_{14}	1.978	1.231	1.6	0.108
β_{22}	-0.270	2.660	-0.1	0.919
β_{23}	-4.443	2.480	-1.8	0.073
β_{24}	-2.426	1.432	-1.7	0.090
β_{33}	2.300	2.421	1.0	0.342
β_{34}	-6.493	3.401	-1.9	0.056
β_{44}	-2.749	3.779	-0.7	0.467
γ_{1W}	0.527	0.377	1.4	0.162
γ_{2W}	0.586	0.437	1.3	0.179
γ_{3W}	-0.223	0.203	-1.1	0.273
γ_{4W}	-2.267	1.132	-2.0	0.045
γ_{1H} (fixed)	1.000	n/a	n/a	n/a
γ_{2H}	3.948	1.520	2.6	0.009
γ_{3H}	5.084	2.238	2.3	0.023
γ_{4H}	4.935	2.174	2.3	0.023
γ_{WH}	4.580	2.318	2.0	0.048
γ_H	1.617	0.246	6.6	0.000

$L = -46,267$

$N = 10,834$

Table 4-5b

Economic Implications of Modified Translog Model with Intercept Terms and Wage and Total-Time Information, as applied to Discretionary-Activity Participation across Four All-Jobs Iso-Opportunity Contours, using the BATS data set

VALUE ESTIMATED:	Median of Sample
Total-Time Elasticity of Demand:	
Immediate Zone Demand	0.705
Near Zone	0.554
Moderate Zone	0.379
Far Zone	0.435
Wage Elasticity of Demand:	
Immediate Zone Demand	-0.077
Near Zone	-0.079
Moderate Zone	-0.031
Far Zone	0.104

Cross-Time Demand Elasticities:				
w/r/t Time of:	Immediate	Near	Moderate	Far
Immediate Zone:	0.850	-0.100	0.180	-0.181
Near Zone:	-0.058	-0.567	0.169	0.065
Moderate Zone:	0.297	0.291	-1.878	0.434
Far Zone:	-0.146	0.117	0.383	-1.716

Notes:

Demands are in trips per day, Discretionary Time is hours,
Travel Times are minutes, & Income is before-tax dollars.

Predictions of aggregate behavior using this model do not appear to be as strong as those from the Type 3 model estimates, but they are quite reasonable. The average X_i^* 's are estimated to be 1.07, 0.62, 0.28, and 0.19, respectively, while the observed per-day average demands are 1.09, 0.62, 0.27, and 0.19.

The travel-time elasticity matrix corresponds roughly with those estimated previously and total-time elasticities are positive, as expected. However, the wage elasticities are generally negative and negligible, except for the furthest zone. One might expect more significantly negative wage effects on discretionary trip-making as workers choose to

work more and engage in fewer discretionary activities (during weekdays at least).

However, the act of working more often may add to discretionary activity participation because of added purchasing power and because work travel can put workers in contact with many activity sites (along travel routes to and from work) for lower travel-time costs than home-based trips.¹¹

Comparison of All Model's Elasticity Estimates

To facilitate comparisons, Table 4-6 provides a summary of elasticities estimated for all five of the model specifications analyzed. The reported values are the median values for the 10,834-household sample, and only the first three models are strictly comparable in terms of all estimates shown, since their response and explanatory variables sets are the same. As described earlier, the fourth model analyzed relies on the same functional form for demand as the third, but it models trip tours to all activity types, rather than individual, discretionary-activity stops. The fifth model allows for work-time (and, therefore, much of discretionary-time) and income endogeneity, so its reported elasticities are with respect to total time and wage variables.

Table 4-6

Summary of Elasticity Estimates: Median Values across Households

Model Used:	Type 1	Type 2	Type 3	Type 3* (All Tours)	Type 4** (Endogen. Work)
Discretionary/Total** Time Elasticity of Demand:					
Immediate Zone Demand	0.389	1.028	0.774	0.970	0.705
Near Zone	0.672	0.869	0.652	0.956	0.554
Moderate Zone	0.572	0.678	0.440	0.848	0.379
Far Zone	0.304	0.706	0.415	0.704	0.435
Income/Wage** Elasticity of Demand:					
Immediate Zone Demand	0.044	-0.294	-0.206	-0.022	-0.077
Near Zone	0.163	-0.208	-0.039	0.039	-0.079
Moderate Zone	0.411	0.086	0.144	0.123	-0.031
Far Zone	0.696	0.122	0.132	0.191	0.104

Cross-Time Demand Elasticities:	Type 1	Type 2	Type 3	Type 3*	Type 4**
Demand for Activities in:	w/r/t to Travel Time to Immediate Zone:				
Immediate Zone:	-0.05	0.26	0.57	0.36	0.85
Near Zone:	0.02	0.10	0.02	-0.16	-0.06
Moderate Zone:	0.28	0.24	0.25	0.05	0.30
Far Zone:	-0.22	0.05	-0.07	-0.09	-0.15
	w/r/t to Travel Time to Near Zone:				
Immediate Zone:	0.08	0.06	-0.01	-0.14	-0.10
Near Zone:	-0.73	-0.89	-0.63	-0.40	-0.57
Moderate Zone:	0.81	0.83	0.69	0.41	0.29
Far Zone:	-0.66	-0.21	-0.32	-0.26	0.12
	w/r/t to Travel Time to Moderate Zone:				
Immediate Zone:	0.23	0.10	0.12	0.02	0.18
Near Zone:	0.57	0.50	0.44	0.40	0.17
Moderate Zone:	-2.71	-2.95	-2.61	-1.34	-1.88
Far Zone:	-0.21	0.38	0.35	0.09	0.38
	w/r/t to Travel Time to Far Zone:				
Immediate Zone:	-0.19	-0.03	-0.14	-0.20	-0.18
Near Zone:	-0.62	-0.19	-0.30	-0.43	0.06
Moderate Zone:	1.02	0.44	0.38	0.07	0.43
Far Zone:	-0.54	-1.45	-1.23	-0.71	-1.72

Notes:

*The first three models and the fifth model discretionary activity participation;
the fourth models trip *tours* and includes all activities.

**The fifth model allows for discretionary-time and income endogeneity, while the others do not.

Demands are in trips per day, Discretionary Time is hours,
Travel Times are minutes, & Income is before-tax dollars.

Of the first three, comparable models, signs on estimates are strongly consistent between the second and third models; but the first model, which relies on a modified version of the linear expenditure functional form, is not so consistent with these two. As described briefly in Chapter Three and earlier in this chapter, the first model suffers from

several constraints on its predictions – and the second and third specifications have their own inflexibilities.

Despite their differences, all models predict strong elasticities of demand with respect to time budgets; however, only the second model produces an estimate exceeding one. The *own*-travel-time elasticities of demand tend to be significantly negative, with elasticities for demand for activities in the *moderately* distant contour estimated to be the most notably inelastic. Additionally, though generally positive, many *cross*-time effects on demand to the *furthest* iso-opportunity contour are estimated to be negative.

All five models predict rather negligible elasticities of discretionary-activity and tour demands with respect to income and wages. This effect may be due to a lack of identification of all income contributions to indirect utility (as discussed earlier), but, also, it may be that additional money does not lead to additional activity participation, everything else constant. For example, the quality of activities and the amount of money spent on them may be substantially affected by income and/or wages, but rates of activity participation may not change much, given fixed time constraints. More detailed data sets, including expenditure and price information, may resolve this question.

Discussion of Value of Time Estimates

In contrast to the reasonable and rather stable elasticity results, the value-of-time estimates vary substantially across the five models. The quartiles of the sample's value-of-time estimates are provided in Table 4-7.

Table 4-7

Summary of Value-of-Time Estimates¹²: Quartiles across Households

Model Used:	Type 1	Type 2	Type 3	Type 3 (All Tours)	Type 4** (Endogen. Work)
Value of Discretionary/Total** Time:					
Minimum	\$ (10.73)	\$ 0.22	\$(7.67E+6)	\$ (82.11)	\$ 0
First Quartile	\$ (8.92)	\$ 7.94	\$ (187.62)	\$ (13.92)	\$ 11.72
Median	\$ (8.06)	\$ 13.13	\$ (102.80)	\$ (9.09)	\$ 26.17
Third Quartile	\$ (6.22)	\$ 21.47	\$ (58.95)	\$ (5.61)	\$ 42.66
Maximum	\$ (1.98)	\$ 151.80	\$ 2.42E+6	\$ (0.16)	\$ 384.90

All values are before-tax, 1990 dollars per hour.

**The fifth model allows for discretionary-time and income endogeneity, while the others do not.

Thus, the fifth model's value of time elasticities are with respect to *total* time and wage.

While of the expected order of magnitude, the signs are opposite of one's expectations in three of the five models! As mentioned in Chapter Three's specification of the first model, the modified linear expenditure system's structure is so limiting that its value-of-time estimates depend only on travel times – rather than income and discretionary time levels. This reasoning may explain a large part of the unexpected results for the first of the five models. But what about the other negative estimates? Isn't a modified translog form flexible enough to provide a first-order estimate of the value of time?

A Functional Conflict between Behavioral Indicators

One interesting but restrictive consequence of the modified-translog specification is that there may be some conflict between the income elasticity signs and those on the marginal utility of income (dv/dY); both depend on the γ_{iY} and γ_{TY} terms, and for one to be positive the other is likely to wind up negative. Equation 4-2 illustrates the conflicting incorporation of these parameters for the Type 3 model specification. In this set of equations, note how the final term of the income elasticity is implicitly negative and the

first term has a negative denominator; therefore, it is just the γ_{iY} parameter that has flexibility to determine the sign on income elasticity for the i th demand.

$$\begin{aligned} \text{Value of Time}_{\text{Type 3}} &= \frac{\mu_o + \left(\frac{1}{T_d}\right) \left(\sum_j \gamma_{jT} \ln(t_j) + \gamma_{TY} \ln(Y) \right)}{\left(\frac{1}{Y}\right) \left(\sum_j \gamma_{jY} \ln(t_j) + \gamma_{TY} \ln(T_d) \right)} \\ &= \frac{v_{T_d}}{\text{Marginal Utility of Income}}, \end{aligned} \quad (\text{Chapter$$

$$\text{Income Elasticity of Demand}_i = \frac{dX_i^*}{dY} \times \frac{Y}{X_i^*} = \frac{\gamma_{iY}}{t_i v_{t_i}} - \frac{\gamma_{TY}}{T_d v_{T_d}},$$

where $T_d > 0$, $t_i > 0$, $\gamma_{TY} = +1$ (fixed, a priori, for model identifiability),

v_{t_i} = Marginal Utility of Travel Time < 0 ,

and v_{T_d} = Marginal Utility of Discretionary Time > 0 .

Four:-2)

The same conflict holds true for the less flexible, Type 2 Model. However, its value-of-time results happen to be much more in line with expectations here¹³. One should be wary of these ostensibly flexible specifications, and further functional flexibility may be sought where practical¹⁴. But something more fundamental to the structure of the models may be causing the unanticipated results.

Identification of All Income/Wage Terms in Indirect Utility

In fact, the primary reason for a negative marginal-utility-of-time result may be that the indirect utility functions underlying the estimated models are limited in their representation of income (or wage) effects. Everywhere one finds an income term (Y) [or wage term (w)] in the different model specifications, it is interacted with either a travel time or available-time-budget term, allowing immediate estimation of the assumed indirect utility function from the results of the system-of-demands estimation. However,

if there are other, isolated income (or wage) effects, in the form of $g(Y)$ (for example, $\log(Y)^2$), these will impact the marginal-utility-of-income estimates and thus the value-of-time estimates.

If one is relying on a system of demand equations derived from the application of Roy's Identity in a *time* environment, one can only identify the magnitude of the effects that are available from derivatives of indirect utility with respect to time and available time. In order to identify the purely income (or wage) effects (or these effects interacted with the vector of prices, which are assumed not to vary across households and thus show up as constants or concealed within fundamental parameters in the regression equations), one needs observable information based on these effects. Essentially then, one needs a system of demands derived from application of Roy's Identity in a *money/price* environment so that parameters characterizing the derivatives of indirect with respect to income (or wage) are all present.

Assume then that one has the system of demand equations as developed from the negative ratios of the derivatives of indirect utility with respect to (invariable) prices and income. The entire model should be estimated in a simultaneous fashion, so that the estimates of optimal demand levels developed in the time setting equal those developed in the price/money setting. One can impose equality across the two demand systems by substituting rather complicated functions of explanatory variables and parameters for several of the constant terms (*e.g.*, the μ_i 's). Given this imposed equality, one can then maximize the likelihood of the sample observations using this single set of significantly more complicated demand equations and one should have access to all parameters of interest.

There is a different way to assess the magnitude of income effects which do not appear in derivatives of indirect utility with respect to travel times and available time; however, it is not as elegant and may not produce consistent estimators. It requires taking the results of the existing models and regressing these estimates of optimal activity participation on a system of demands developed in a price environment. This method was used with the third model to take a closer look at the marginal utility of income, and it consistently produced *positive* marginal utilities, thanks to the incorporation of Y , $\log(Y)$ and $\log(Y)^2$ effects. The indirect utility specification used is the following:

Indirect Utility =

$$\begin{aligned}
 v = v_1(\bar{t}, T_d, \bar{P}(\text{implicitly}), Y) + v_2(\bar{P}, Y) = \\
 \alpha_o - \sum_i \mu_i t_i + \mu_o T_d + \sum_i \alpha_i \ln(t_i) + \sum_{ij} (1/2) \beta_{ij} \ln(t_i) \ln(t_j) + \quad (\text{Chapter} \\
 \sum_i \gamma_i \ln(T_d) \ln(t_i) + \sum_i \gamma_{iY} \ln(Y) \ln(t_i) + \gamma_{TY} \ln(T_d) \ln(Y) + \\
 \sum_i \delta_{iP} \ln(P_i) + \sum_i \delta_{iPY} \ln(P_i) \ln(Y) + \delta_Y Y + \delta_{YY} \ln(Y)^2
 \end{aligned}$$

Four:-3)

The optimal demand levels which result from application of Roy's Identity in a *price* environment to the above formulation are the following:

$$X_i^* = \frac{-\frac{\delta_{iP}}{P_i} - \frac{\delta_{iPY}}{P_i} \ln(Y)}{\delta_Y + \left(\frac{1}{Y}\right) \left(\sum_i \gamma_{iY} \ln(t_i) + \gamma_{TY} \ln(T_d) + \sum_j \delta_{jPY} \ln(P_j) + 2 \delta_{YY} \ln(Y) \right)}$$

$$= \frac{-\delta'_{iP} - \delta'_{iPY} \ln(Y)}{\delta_Y + \left(\frac{1}{Y}\right) \left(\sum_i \gamma_{iY} \ln(t_i) + \gamma_{TY} \ln(T_d) + \delta_{PYT} + 2 \delta_{YY} \ln(Y) \right)}, \quad (\text{Chapter$$

where t_i = Travel Time to Activity i , P_i = Price for Activity i ,

Y = Income, T_d = Discretionary Time Available,

and $\delta'_{iP} = \frac{\delta_{iP}}{P_i}$, $\delta'_{iPY} = \frac{\delta_{iPY}}{P_i}$, $\delta_{PYT} = \sum_j \delta_{jPY} \ln(P_j)$,

& $\gamma_{TY} = +1$ (for identifiability of parameters).

Four:-4)

Note that assumption of constant price levels leads to non-identification of the fundamental price-interacted parameters but produces a similar, estimable functional form, where the price-interacted parameters are subsumed into identifiable parameters. This implicit incorporation of the price-specific parameters can be thought of as having occurred in all model specifications used here, particularly those with the constant terms (which thereby allow for isolated-price effects).

Using this specification, a solution was sought which minimized the squared difference between earlier estimates of X_i^* 's (derived in a time environment and constructed using Table 4-2a's parameter estimates) and the estimates arising from Equation 4-4's demand specification (with $\gamma_{TY} = +1$ and the already-estimated γ_{iY} 's substituted in directly). This method produced estimates of the marginal utility of income (*i.e.*, the denominator in the demand equation of Equation 4-4) which are dominated by a positive δ_Y term. The negative terms in the marginal utility of income which come from dv_1/dY are negligible in comparison with the highly positive estimate of δ_Y . The new

value-of-time estimates are all positive, but their magnitude is too low by several orders (*e.g.*, the median value is \$.00045/hour). Moreover, the second set of estimates relies so substantially on the constant terms in the demand equations that estimates are predicted to vary little across households. Apparently then, this expanded indirect utility specification and/or the methods used to estimate this system's parameters (including simply minimizing the sum of squared differences over all demands and strong assumptions like price invariability) remain lacking. These modeling complexities are a prime area for additional research.

Other Reasons for Incorrect Marginal Utility of Income Estimates

In addition to full identification of the indirect utility function and flexibility in functional form, there are other issues to consider in the estimation of the value of time. For example, in the fifth model estimated here, which comes from the Type 4 Model specification, the assumption of exogenously determined income and discretionary time budgets is dropped, theoretically removing any unmodeled dependencies across these explanatory variables and demand which may have caused erroneous results. The loss of this implicit and strong exogeneity assumption – which is present in the previously estimated models – may be what gives this final model its reasonable value-of-time estimates.¹⁵

Another reason for a negative marginal utility of income (and thus negative value of time) results may be that high income households are able to live in lower travel-time environments, so the parameters affecting the marginal utility of travel times might pick up an income effect, leaving the final income effect rather ambiguous and ostensibly negative in many models. For example, after normalizing income for total-time

availability, H , one finds that average travel times to each of the four all-job iso-opportunity contours gradually fall as normalized income increases; the mean travel times between the first and fourth quartiles of Y/H fall from 11.5 to 11.0 minutes to reach the first contour, 22.1 to 20.6 to reach the second, 33.5 to 31.3 to reach the third, and 49.4 to 47.2 to reach the fourth. It appears that high-income/low-time households are residing in locations that better fit their constraints, as one would expect; thus, if household location choice were made endogenous to the model, one could avoid some of the biases this dependence may create.

Further Qualifications

While the results of this research are interesting, one should recognize that the data are imperfect and the model assumptions are strong. For example, the travel-time data are measured with some error, thanks to zonal aggregation and reliance on free-flow, automobile travel times (– and due to the chaining of trips, as discussed at the end of Chapter Three). And the income and wage variables either come from survey-bin mid-points, in the case of income-reporting households, or have been estimated, for the non-reporting households. Simply the use of a model with one or more explanatory variables measured with error leads to highly uncertain impacts on estimates. (Greene 1993) Unfortunately, most models of travel behavior are subject to such deficiencies in the data set, since income tends to be reported by ranges and/or travel times come from a time-of-day-independent data base.

There also is the concern that cross-sectional data do not provide the necessary variation to discern heterogeneity from state dependence in the unobserved information which influences decisions. Meurs (1990) recommends use of panel data for estimation

of trip generation due to difficulties arising in cross-sectional models from omitted time-invariant/fixed effects across individuals; for example, Meurs's models using the Dutch Mobility Panel data set indicate that cross-sectional income elasticities of demand tend to be biased high. Kitamura (1988) uses a three-year panel data set to study trip generation rates and finds the serial correlation to be substantial "suggest(ing) that important determinants of trip generation lie outside the set of variables that have traditionally been considered in travel behavior analysis."

Kitamura *et al.* discuss the need for longitudinal calibration to avoid a "longitudinal extrapolation of cross-section variations" (Kitamura *et al.* 1996). In other words, cross-sectional elasticities are observed over different *individuals* yet often "applied as if they represent longitudinal elasticities that capture the change in behavior that follows a change in a contributing factor within each behavioral unit." (Kitamura *et al.* 1996, pg. 274). The use of cross-sectional elasticities for estimation of longitudinal behavior is only rigorously valid under restrictive conditions, such as response being immediate and its magnitude being independent of past behaviors, according to Goodwin *et al.* (1990). In some cases, the greater the amount of time between a change in an independent variable and measurement of behavioral response, the higher the likelihood that cross-sectional estimates apply.

Becker (1965) voices some concern about the interpretation of cross-sectional elasticities for a different reason. His primary thesis is that the true cost of "commodity" consumption involves a time cost, not just a monetary cost for the non-time factors/goods used to produce commodities. Thus, he argues that "traditional cross-sectional estimates of income elasticity (which) do not hold either factor or commodity prices constant..." are

“...biased downward for time-intensive commodities, and give a misleading impression of the effect of income on the quality of commodities consumed.” (1965, pg. 517)

Unfortunately, without adequate longitudinal data sets, Kitamura’s comments can only be used to qualify the results of this research, in the estimates of such elasticities. Accommodation of Becker’s fully general model requires information on the production technology of commodities (*e.g.*, the combinations of time and money that produce a dining-out experience), so Becker’s concerns may only be stated as qualifications here.

The globalness of the likelihood’s maxima used to estimate parameter values and assess covariance also represents an assumption of these results. While the global maximum is a consistent estimate of the true parameter values – assuming the model and its distribution have been correctly specified, there is no guarantee that the search routine has converged upon the function’s global maximum¹⁶. This particular model’s requirement of *positive* activity-participation rate-parameter estimates for a calculable likelihood value¹⁷ often makes the acquisition of feasible starting parameter values a significant chore, particularly for the most functionally flexible models; thus, it is not easy to try starting at a wide variety of highly distinct parameter sets and comparing final points of convergence in an effort to avoid local maxima. However, as long as the results seem reasonable (*e.g.*, as long as estimates of the means and proportions of trip-making correspond well with observed values), one may expect that one’s results are not a local maximum of poor prediction quality. And, as long as the convergent set seems robust to some changes in starting values, one may expect to be at the *neighborhood’s* maximum. Both these details were confirmed for the results presented here.

The assumption of a single gamma variable characterizing the unobserved heterogeneity across all consumption types considered is a very strong one. The assumption of activities taking place as a Poisson process is also a strong one; it implies independent increments, meaning that the number and types of events to show up in a particular time interval, knowing the rates of occurrence, are independent of those appearing in another, non-overlapping interval. In reality, one expects that people will not over-consume any certain type of activity; so, if, for example, one knows that a household engaged in several social activities in the morning, one would expect fewer such activities in the evening, given the household's optimal rate of social-activity participation. However, this issue of behavior conforming to a Poisson process is likely to be less of a concern over longer survey periods.

As a means of comparing different stochastic specifications, consider the nesting of the Type 3 specification for the four-contours discretionary trip making within less stochastically restrictive models. For example, consider a model of *no* gamma-error-term correlation, *i.e.*, a system of independent negative binomials (with related means, thanks to the sharing of parameters across the X_i^* specifications). As shown in Table 4-3c, this difference of models produces a likelihood ratio test statistic of $2(-42592 - (-46218))$ or 7,252, for a three-parameter change (via the addition of three more over-dispersion factors). This result clearly calls for a rejection of the assumption that the unobservable gamma errors are equal across the four demand types¹⁸. And it suggests that a more complex model of correlation in unobserved information is needed – in place of the perfect-correlation/same-gamma assumption being made here. To allow more flexibility and thus complexity, one may wish to consider, for example, the incorporation of a *set* of

unobservable random variables having a multivariate log-normal stochastic distribution – all within the Poisson specification of Equation 3-21. However, it may well be that *no* form of unobserved heterogeneity – other than that of the single gamma term – allows the observation probabilities to collapse to a closed-form solution. And, without a tractable solution, one will need to rely on a technique like the simulation of likelihoods to approximate maximum likelihood estimation.

It is anticipated that the correlations of unobserved information are at least positive for the current specification of demand types, and this belief is supported by correlation results. For example, sample correlations of the ratios of observed trip-making to average, optimal demands (predicted using the independent-negative-binomials-with-shared-parameters model, type v in Table 4-3c) are all positive – though none exceeds 0.155. These sample correlations also are observed to depend on similarity of demand types: they all support a general trend of decrease with dissimilarity (for example falling from 0.154 between the first and second contours, to 0.036 between the first and third contours, to 0.003 between the first and fourth contours). Note, however, that this is a coarse way to estimate the correlation because the observed activity participation is highly discrete (owing to the short sample period of the surveys).

In the short run, correlation may be less apparent because the phenomenon of trip chaining produces a dependence among the number of trips taken to the more distant activities (in addition to allowing the traveler to face a different set of trip costs than the estimated models assume). For example, when one is analyzing stops (rather than tours) and one observes at least one distant/far trip on a specific day for a specific household, one may expect more far trips but fewer short/near trips that same day than the long-run

optimal rates for that household would suggest. However, as the length of a survey period increases (*e.g.*, from one day to a week or more), more balance in destination choices is likely, so the trip-chaining phenomenon is expected to affect parameter estimates less.

ENDNOTES:

-
- ¹ S-Plus's symbolic differentiation capabilities are limited to general operands (*e.g.*, multiplication, logarithms); it will not automatically take derivatives of the gamma functions which exist in the likelihood formulation used here. To facilitate the maximization process, entire derivatives may be specified for the S-Plus minimization routine, `minsum()`; however, such derivatives must be specified within a couple lines of code, which was not possible with the complex specification used here.
- ² Unfortunately, due to reasons of proprietary information protection, the numerical derivatives that the algorithm creates and uses for optimization are not available to the modeler; if available, these would prove useful for speedier estimation of the variance-covariance matrix of parameter estimates as well as for relative ease in checking the global probability of likelihood-maximizing parameter sets using Gan and Jiang's (1997) suggested method.
- ³ The 1.638 value is computed using the difference of the variance and the mean of the total number of discretionary trips across households, divided by the squared mean. This is simply the solution using the negative binomial's variance formula: $\sigma^2 = \mu + \alpha\mu^2$. This no-information dispersion-parameter estimate for each of the X_i 's individually produces alphas of 2.4, 3.4, 5.0, and 10.0; in contrast, a solution involving explanatory variables in a Type 3 model specification (with cross-equation parameter constraints in effect) produces estimates of 1.5, 2.2, 2.8, and 7.4, all of which are statistically significantly lower.
- ⁴ Note that the no-information situation models each $\Lambda_{i,n}$ as a constant across all households, independent of time and budget information. In contrast, a *full*-information model for this particular data set produces a log-likelihood value of -32,749. The full-information's likelihood-maximizing case occurs when each $\Lambda_{i,n}$ is set to equal the observed number $\Lambda_{i,n}$ and variance is minimized (*i.e.*, the over-dispersion factor "a" is zero).
- ⁵ The three additional restrictions on the constrained model are that the correlations are perfect between the unobserved gamma errors terms of the first and the second, third, and fourth demands (implying perfect correlation across *all* six distinct pairings of these demands' unobserved gamma terms).
- ⁶ The pseudo-R²'s for the other, comparable models estimated in this chapter are: 8.41% for the modified translog without intercepts (whose results are provided in Tables 4-2a and 4-2b), and 9.51% for the modified translog with intercepts estimated on the tour data (whose results are shown in Tables 4-4a and 4-4b).
- ⁷ Observations of zero translate to full-information, Poisson rates of zero; and, since the probability one observes zero trips if the rate is zero is one, the log-likelihood is zero (rather than negative) for these observations.
- ⁸ When the four equations are estimated separately, three more μ_o 's are estimated than previously, six more β_{ij} 's are estimated and nine more γ_{it} 's are estimated.
- ⁹ The hypothesis that the cross-equation demand-parameter constraints are valid, assuming that the model allowing total independence is valid, produces a p-value of 9.1e-27, which is certainly less than any meaningful cut-off level for insignificance.
- ¹⁰ The estimates of "α" terms associated with the no-information cases of individual demands (denoted by i) were computed using the following equation: $\hat{\alpha}_i = (\hat{\sigma}_i^2 - \mu_i) / \mu_i^2$.
- ¹¹ The elasticities of demand with respect to wage also may be negligible because the wage data are inferred (and assumed to be zero for households with no workers), so the variation in this variable may not be really capturing wage differences. Moreover, workers may have relatively little choice over total hours worked; their decision may be largely discrete in that they choose to work full- or part-time and their hours are largely fixed from that decision.
- ¹² The values shown in the table of value-of-time quartiles for the second model type considered come from a first-order approximation, where estimates of the marginal utility of time and income are simply substituted into the ratio for a value of time; this is the method used for all of the measures provided, unless otherwise noted. A more refined, second-order estimator of the value of time in this model

accounts for the positive correlation between the marginal utilities (which is estimated to range from +.10 to +.41 for the sample) and produces the following quartiles: \$0.22, \$8.05, \$13.33, \$21.86, \$156.46. One observes that even though there is significant correlation, the first-order value-of-time estimates lie close to these second-order estimates.

- ¹³ Standard errors of these value-of-time estimates were computed (using the methods described in the Appendix, section A-6); and all the value-of-time estimates for this model are highly statistically significantly different from zero, assuming the model specification and data are correct.
- ¹⁴ To add flexibility through additional identifiable terms, a model which added the terms $\log(T_d)^2$, $\log(Y)\log(T_d)^2$, and $\log(Y)^2\log(T_d)$ to Equation 3-17's indirect utility specification was also estimated. It was hoped that this model would produce value-of-time results more consistent with expectations; however, while the log-likelihood value rose significantly (from -46,218 to -46,012), all value of time estimates across the 10,834 BATS households remained negative, with a median value of -\$6.49. Thus, the problem may lie outside the specification of the time-based demand system, which can only identify effects that interact with travel times and available time.
- ¹⁵ While quite reasonable, the estimates for the fifth model are probably too high. The average pre-tax wage for *working* households in the BATS survey (for the 1989 tax year) was estimated to be \$20.17; across all households it was \$16.99.¹⁵ Post-tax values are likely to be from 20 to 35 percent less, and there may be a bias (estimated to be a positive nine percent, in Chapter Three) for not allowing income and discretionary-time budgets to be endogenously determined.
- ¹⁶ Gan and Jiang (1997) have recently devised a relatively simple test for assessing the globalness of one's likelihood maxima; however, it relies on one's estimation of the likelihood's matrix of second derivatives, which is rather tedious to compute for this particular model and thus was not used.
- ¹⁷ Poisson and gamma probabilities are not valid with negative means; for example, if λ were negative, probabilities would be negative for odd values of "k", and the function $I(x)$ requires a positive argument. Besides, there is no such thing as negative counts of activity participation.
- ¹⁸ If one were to add no new parameters by imposing the same over-dispersion term on each of the independent negative binomials, the log-likelihood value for this particular data set is -42,981. This results in a difference of log-likelihoods (from that of the constrained model) of over 3,000, representing a very significant freeing of the model with no added parameters.

Chapter Five: Hypothesis Testing and Cost-Benefit Estimation

One can test a multitude of hypotheses with the model results. For example, given fixed income and discretionary-time levels, is total travel time by a household independent of the travel-time environment? Is the total number of trips independent of income or travel times? And are the time and budget constraints binding? One can also measure the impacts of changes in access to opportunities in both time and money units, by inverting estimated indirect utility functions with respect to both constraint levels.

These hypotheses are tested and disbenefits estimated using the Type-Two model's parameter estimates (shown in Tables 4-2a and 4-2b) and structure (illustrated by Equations 3-15 and 3-16). This particular model was chosen for these tests and computations because of its apparent agreement with theory relative to many of the other models' results (*i.e.*, its positive value of time and marginal utility of income) and its inclusion of the discretionary-time and income variables, which are an explicit part of several of the hypotheses posed and which produce the desired measures of disbenefits.

Hypothesis Testing using the Type Two Model

Stated in equation form, the hypotheses described in this chapter's introduction and tested here are the following:

$$\text{Hypothesis 1: } \frac{d(\text{Total Discretionary Travel Time})}{d(\text{Travel Time}_j)} = \frac{d\left(\sum_i X_i^* t_i\right)}{dt_j} = 0, \forall j$$

$$\text{Hypothesis 2: } \frac{d(\text{Discretionary Activity Participation})}{d(\text{Income})} = \frac{d\left(\sum_i X_i^*\right)}{dY} = 0, \forall j$$

$$\text{Hypothesis 3: } \frac{d(\text{Discretionary Activity Participation})}{d(\text{Travel Time}_j)} = \frac{d\left(\sum_i X_i^*\right)}{dt_j} = 0,$$

$$\text{Hypothesis 4: } \text{Marginal Utility of Discretionary Time} = \frac{dv}{dT_d} = 0,$$

$$\text{Hypothesis 5: } \text{Marginal Utility of Income} = \frac{dv}{dY} = 0.$$

Hypothesis 1

After studying a variety of aggregate data sets of travel-time expenditures across regions and across countries, Zahavi and others (Zahavi 1979a & b, Zahavi and Talvitie 1980, Zahavi and Ryan 1980, Zahavi *et al.* 1981) have proposed that *total* travel time expenditures are inelastic with respect to travel-time costs. However, Zahavi's observations and conclusions (Zahavi 1979a & b, Zahavi and Ryan 1980, Zahavi and Talvitie 1980, Zahavi *et al.* 1981) tend to be based on aggregate data and simple correlations; so this hypothesis is particularly interesting to test here, where the explanatory model is disaggregate, behaviorally based, and quite complex. However, since the Type Two model data consider only discretionary-activity participation and assume round-trip travel from one's home (without the chaining of trips into tours), the test of this first hypothesis is somewhat different from Zahavi *et al.*'s proposition.

The results of the test of Hypothesis 1 are shown below, in Table 5-1a.

Table 5-1a

**Quartiles of the Estimates of the Derivative of Total Travel Time
for Participation in Discretionary Activities
with respect to Activity Access/Travel Times (units are dimensionless)**

Derivative Quartiles:	Based on Travel Time to the following Zone:			
	Immediate	Near	Moderate	Far
Minimum	0.14	0.03	-0.90	-0.330
25%	1.02	0.20	-0.21	-0.088
Median	1.62	0.34	-0.15	-0.065
75%	2.32	0.50	-0.10	-0.046
Maximum	8.87	2.35	-0.023	-0.014

Evidently, total travel time to access discretionary activities increases when the travel times to access the closer opportunities increase, indicating a dependence on these nearby activities. But the derivatives tend to fall when the more distant opportunities become more time-consuming to access, suggesting that people substitute nearer, less travel-time-costly activities for those far activities. The effects are probably strongest for the nearer activities since the data indicate greater rates of activity participation in the closer iso-opportunity contours: the mean rates across the observed set of households are: 1.09, 0.617, 0.234, and 0.189, for trips to the immediate through the far zones, respectively.

There is always uncertainty in estimates. These particular derivatives involve ratios of random variables (the parameter estimates); and, using first-order Taylor series approximations to the variance of the derivative functions (as described in Appendix section A-7), the quartiles of the standard errors for these total-travel-time-derivative estimates are provided in Table 5-1b.

Table 5-1b

Quartiles of the Standard Error Estimates of the Derivative of Total Travel Time for Participation in Discretionary Activities with respect to Activity Access/Travel Times (units are dimensionless)

Standard Errors of Derivative Quartiles:	Based on Travel Time to the following Zone:			
	Immediate	Near	Moderate	Far
Minimum	0.006	0.003	0.002	0.002
25%	0.028	0.018	0.009	0.007
Median	0.044	0.028	0.015	0.012
75%	0.062	0.042	0.024	0.017
Maximum	0.246	0.172	0.116	0.076

These results lead to the following T-statistics for this hypothesis test:

Table 5-1c

Quartiles of the T-Statistics of the Derivative of Total Travel Time for Participation in Discretionary Activities with respect to Activity Access/Travel Times

T-Statistics of Derivative Quartiles:	Based on Travel Time to the following Zone:			
	Immediate	Near	Moderate	Far
Minimum	19.6	6.4	-15.0	-10.3
25%	35.9	11.3	-11.6	-6.6
Median	37.1	11.8	-9.7	-5.6
75%	37.7	12.4	-8.8	-5.0
Maximum	39.0	14.0	-7.4	-3.8

Interestingly, all household observations call for a rejection of Hypothesis 1, so the results are not consistent with the hypothesis. Moreover, an ordinary least squares regression of total travel time per household (using *reported* travel times from the BATS data set) consistently produces statistically significant coefficients on three of the four iso-opportunity access times, after controlling for household income and total time availability¹.

Hypothesis 2

The estimated quartiles across the 10,834-household population for the derivative of total daily discretionary-activity participation (X_T^*) with respect to income are the following:

Table 5-2a

Quartiles of the Estimates of the Derivative of Total Discretionary-Activity Participation with respect to Income (units are daily activity number per dollar)

Derivative	Minimum	0.25	Median	0.75	Maximum
Quartiles:	-5.34E-04	-1.44E-05	-8.52E-06	-5.31E-06	-3.95E-07

While the results are negligible for this model, their consistently negative sign suggests that total discretionary trip-making does not go up when income rises, *ceteris paribus*. Note that Meurs's models (1990) using the Dutch Mobility Panel data set indicate that cross-sectional income elasticities of demand tend to be biased high (in the positive direction), so actual, derivatives may even more negative. Some possible explanations for this result are that high-income households work so hard that they are too tired to take as many discretionary trips as others, they access higher-quality opportunities to accomplish many demands at once, and/or they use their wealth to purchase goods and services that enable them to avoid making trips.

The quartiles for these estimates' T-statistics are the following:

Table 5-2b

Quartiles of the T-Statistics of the Derivative of Total Discretionary-Activity Participation with respect to Income (units are daily activity number per dollar)

Quartiles of Derivative	Minimum	0.25	Median	0.75	Maximum
T-Statistics:	-23	-0.61	-0.36	-0.23	-0.02

Few of the estimates are significant enough for us to reject the hypothesis, suggesting that income plays little role in a household's rate of discretionary trip-making/out-of-home activity participation. This is a remarkable result, given how influential one might assume income is. However, this result does not speak to income's role in the consumption of other, more material goods; it is in this other consumption, not modeled here, that income is likely to substantially influence choice, as has been indicated in results of the more typical system-of-demands analysis (see, *e.g.*, Deaton 1987, Pollack and Wales, 1978 & 1980, and Stone 1954).

Hypothesis 3

The quartiles for the derivatives of total discretionary-activity participation with respect to the four different travel times sets are shown below, in Table 5-3a.

Table 5-3a

Quartiles of the Estimates of the Derivative of Total Activity Participation in Discretionary Activities with respect to Activity Access/Travel Times (units are daily participation rate per minute change in travel time)

Derivative Quartiles:	Based on Travel Time to the following Zone:			
	Immediate	Near	Moderate	Far
Minimum	-0.157	-0.142	-0.096	-0.030
25%	0.005	-0.021	-0.017	-0.010
Median	0.039	-0.015	-0.008	-0.007
75%	0.092	-0.010	-0.004	-0.005
Maximum	0.917	0.048	0.011	-0.001

The T-statistics for these estimates are as follows:

Table 5-3b

Quartiles of the T-Statistics of the Derivative of Total Activity Participation in Discretionary Activities with respect to Activity Access/Travel Times

Quartiles of Derivative T-Statistics:	Based on Travel Time to the following Zone:			
	Immediate	Near	Moderate	Far
Minimum	-70.4	-40.4	-29.7	-27.0
25%	2.0	-11.2	-12.3	-11.1
Median	12.2	-8.4	-8.3	-9.4
75%	19.2	-6.0	-5.2	-7.9
Maximum	34.8	3.3	2.5	-3.0

With the exception of household response with respect to the immediate zone's travel times, it appears that the total discretionary-activity-participation response may be largely negative, in a statistically significant sense. However, the magnitude of net response for most travel times is quite minor when compared with mean activity-participation frequencies of 1.09, 0.62, 0.27, and 0.19 (for the four contours, respectively), suggesting quite stable activity-participation rates. Thus, the results are consistent with previous work, discussed in Chapter Two, in which trip frequency is found to be insensitive to supply-side variables.

Additionally, Golob, Beckmann, and Zahavi (1981) reference works (*e.g.*, Smith & Schoener 1978 and Zahavi 1979b) which cause them to conclude that “when travel speeds increase, travelers prefer to trade-off saved time for longer trips, rather than for more trips,” and “(w)hen incomes increase, travelers tend to purchase higher speeds (such as by transferring from bus to car travel) and travel longer distances, instead of generating more trips.” (p. 378) These characterizations suggest that one should expect to find nearly constant activity participation levels, regardless of travel times and/or costs, which is consistent with the results found here.

In a related test (but one that work-trip travel time as given), Kitamura (1984) applies models of non-work activity choice and time allocation to estimate the effect of work-trip travel time, with data from the 1977 Baltimore Travel Demand Data set. Kitamura finds this variable to not be statistically significant, in contrast to the other explanatory variables considered, such as cars per driver, work duration, and gender. In some contrast, Purvis *et al.* (1996) use work-trip duration in ordinary-least-squares models of home-based shop/other and social/recreational trip *generation*, and their results indicate an inverse relationship between work-trip duration and non-work home-based trip frequencies (using a 1990 data set), suggesting a binding total-time budget.²

Hypothesis 4

The quartile estimates for the derivative of indirect utility with respect to discretionary time budget are the following:

Table 5-4a

**Quartiles of the Estimates of the Derivative of Indirect Utility
with respect to Discretionary Time (units are utility per daily hour)**

Derivative	Minimum	0.25	Median	0.75	Maximum
Quartiles:	0.117	0.439	0.616	0.914	1.558

As expected, these estimates of the shadow price of this constraint are strictly positive. One cannot effectively comment on the magnitude of these derivatives or their range of estimated values, given that utility is an ordinal measure and uniquely identified only up to a monotonic transformation. However, one can estimate their T-statistics, and the quartiles of these are the following:

Table 5-4b

**Quartiles of the T-Statistics of the Derivative of Indirect Utility
with respect to Discretionary Time (units are utility per daily hour)**

Quartiles of Derivative	Minimum	0.25	Median	0.75	Maximum
T-Statistics:	7.1	8.7	9.1	9.3	10.2

Fortunately, the standard errors are sufficiently small for these estimates that one can be confident that the true marginal utilities of time are strictly positive, assuming the model has been specified correctly.

Hypothesis 5

The quartiles for the marginal utility of income estimates are the following:

Table 5-5a

Quartiles of the Estimates of the Derivative of Indirect Utility with respect to Income (units are utility per annual dollar of income)

Derivative	Minimum	0.25	Median	0.75	Maximum
Quartiles:	2.45E-05	8.86E-05	1.25E-04	1.86E-04	3.11E-03

Recall that this particular model was chosen in large part because it is one of the few that has a positive value of time (thanks to its positive marginal utility of income³).

While the marginal utilities of income appear small in magnitude, indirect utility is not scaled to any known dimension, so one cannot assume that the marginal utilities are relatively small. The quartiles of the T-statistics for these estimates are a better way to assess the magnitude of the marginal utility estimates, and they are provided below, in Table 5-5b.

Table 5-5b**Quartiles of the T-Statistics of the Derivative of Indirect Utility with respect to Income (units are utility per annual dollar of income)**

Quartiles of Derivative	Minimum	0.25	Median	0.75	Maximum
T-Statistics:	4.9	6.6	7.2	7.6	9.7

These values call for a rejection of the null hypothesis that the marginal utility of income is zero or negative. Note that the sign of this result is not consistent with the results of two of the other translog-based models, where marginal utilities of income were estimated to be negative; but that discrepancy is likely to have much more to do with an incomplete indirect utility specification than with the true shadow price on this constraint.

Cost-Benefit Analysis using the Type Two Model

Thanks to the microeconomic rigor of the model developed in this research, one also can estimate the benefits and costs associated with a variety of policies. For example, what are the equivalent variations in units of time and money of a policy which causes travel times to all opportunity contours to rise 50 percent? For purposes of illustration, this particular environmental change is considered here for all sampled households – and for a “typical” household facing different time and income budgets. The method is illustrated using the estimation results of the Type 2 model.

In order to apply the method of equivalent variation (as discussed in Chapter Three’s section on Estimating Benefits and Costs), one must invert the indirect utility function with respect to the budget level.⁴ In the model developed in this research, there are *two* budget levels, providing more information to policy-makers than strictly money-budget models. However, one should be aware that the effects of unidentified, isolated income terms which exist in the full, true indirect utility specification are not captured by the

single time-setting system of demand equations estimated in this research. Thus, inversion of the indirect utility function's time-interacted, identified effects with respect to the *income* variable may not tell the whole story. However, the equivalent variation estimates in terms of time units are theoretically sound in this regard.

For the Type 2 model defined in Chapter Three (Equations 3-15 and 3-16), inversion with respect to the time and income budgets produces the following expenditure functions:

$$e_{T_d}(\bar{t}, u, Y) = \exp \left[\frac{u - \left(\sum_j \alpha_i + \sum_{ij} \beta_{ij} \ln(t_i) \ln(t_j) + \sum_i \gamma_{iY} \ln(t_i) \ln(Y) \right)}{\left(\sum_i \gamma_{jT} \ln(t_j) + \gamma_{TY} \ln(Y) \right)} \right], \&$$

$$e_Y(\bar{t}, u, T_d) = \exp \left[\frac{u - \left(\sum_j \alpha_i + \sum_{ij} \beta_{ij} \ln(t_i) \ln(t_j) + \sum_i \gamma_{iT} \ln(t_i) \ln(T_d) \right)}{\left(\sum_i \gamma_{jY} \ln(t_j) + \gamma_{TY} \ln(T_d) \right)} \right]$$

where " t_i " = Travel Time to Iso – Opportunity Contour " i ", " Y " = Income,
& " T_d " = Discretionary Time Available.

Note that the quartiles of the expenditure functions for these sampled households at their *current* indirect utility and travel-time levels are just the observed levels of income and discretionary time (since one inverts the estimated indirect-utility formula with respect to these two variables). Using the Type 2 model parameter estimates (shown in Table 4-2a), the quartiles of the first-order estimates of optimized utility levels for the sample households are as follows:

Table 5-6a**Quartiles of Indirect Utility Estimates**

Quartiles of Indirect Utility Estimates across Household Sample:	Minimum	0.25	Median	0.75	Maximum
	-52.5	17.9	30.9	39.3	76.8

The quartiles of the changes in indirect utility level estimates, following a 50-percent increase in all travel times for all households are as follows:

Table 5-6b**Quartiles of Indirect Utility Changes Following 50% Increase in Travel Times**

Quartiles of Indirect Utility Drop (following 50% travel-times increase):	Minimum	0.25	Median	0.75	Maximum
	-19.0	-8.4	-7.0	-5.8	-1.5

Since utility does not enjoy an understood scale or dimension, it is generally more useful to look at the equivalent variation associated with a change; this is the amount of money or time lost (or gained) that the change would be equivalent to, under existing price/travel-time conditions. The quartiles for equivalent variation in units of money and time are as follows:

Table 5-6c**Quartiles of Equivalent Variation, in Money and Time Units**

Equivalent Variation Quartiles in Dollars per Day:	Minimum	0.25	Median	0.75	Maximum
	\$ (409.18)	\$ (119.81)	\$ (85.30)	\$ (59.03)	\$ (4.23)

Equivalent Variation Quartiles in Hours per Day:	Minimum	0.25	Median	0.75	Maximum
	-51.5	-13.3	-9.8	-7.1	-2.6

All are negative changes, as expected, since everyone has to engage in some level of activity participation outside the home and travel is almost always viewed as a cost; so having to spend more time to participate in the same types of activities is expected to be a disbenefit. The disbenefits are estimated to be substantial; the median equivalent variation in daily income is negative \$85.30, which means that the median amount a

household would be willing to spend per day just to avoid travel-time increases of 50 percent is \$85.30! This translates to over \$31,000 per year or a roughly \$300,000 premium for a home location⁵ which provides such access, over one that faces 50 percent longer travel times.

Moreover, the median amount of discretionary time such a change represents to the sampled households is estimated to be 9.8 hours per day! In contrast to the same policy being imposed in a money-expenditures environment, this level of equivalent variation is much more than one would expect the household's travel time to increase. If monetary prices were to increase by 50 percent, one would need a proportional increase in income to remain at the same level of utility. But, for the sample of 10,834 households, the median of round-trip travel times multiplied by observed activity participation is just 47.3 minutes per day (with a mean of 1.42 hours). So, even if households were to continue making the same number of trips to the same zones, the expectation is for a median travel-time increase of just 24 minutes – which is nowhere close to the 9.8 hours of equivalent variation. From these results it seems clear that *how* one experiences one's time is of great import (*e.g.*, traveling versus leisure). And access to opportunities is highly valued by households; households appear willing to spend a great deal of money and/or time in order to avoid increases in travel times.

If one wishes to consider a specific set of household characteristics, for example a low-income household versus a high-income household facing the same set of travel times, one can get a feeling for the differences in these households' valuations of changes in access to opportunities. As an illustration of this, consider four households which face the median set of travel times for the San Francisco Bay Area region to the four iso-

opportunity contours modeled: two of these households have the same, sample-median amount of discretionary time, but face very significant differences in their income constraint, while the other two face the same, sample-median income constraint but not the same discretionary-time constraint. How do their valuations of a fifty-percent increase in all travel times differ? The set of median travel times for the region along with the different income and discretionary time constraints considered specifically here (*i.e.*, the tenth and 90th percentiles, as well as the median) are shown below.

Table 5-6d

Median Travel Times and Low, Median, and High Budget Levels for Sample

Median Travel Times to Access the Four Iso-Opportunity Contours:				
(in minutes)	Immediate	Near	Moderate	Far
	9.78	17.02	27.37	42.67

Levels of Income and Discretionary Time Used:		
Income: (1989 pre-tax \$/year)	10%	\$ 17,500
	Median	\$ 42,500
	90%	\$ 87,500
Discretionary Time: (hours/day)	10%	17.93
	Median	38.83
	90%	73.63

The equivalent variation estimates which result from a 50-percent increase in travel times are shown below. As expected, lower-income and lower-time households are less able to place a high equivalent value on such a change.

Table 5-6e**Comparison of Equivalent Variation Changes for Specific Household Types**

Comparison of Equivalent Variation Estimates for Low/High-Income and -Time Households: (following a 50% increase in all travel times):	
Low Income Household:	\$ (35.18) per day
High-Income Household:	\$ (165.00) per day
Low-Discretionary-Time Household:	-5.21 hours per day
High-Discretionary-Time Household:	-16.45 hours per day

It is interesting to note that the differences in the measure of indirect utility estimated for these different households are *greater* for the low-income and low-time households (*i.e.*, -7.3 and -8.0 are the utility differences for the low-income and low-time households considered, versus -6.4 and -5.9 for the high-income and -time households). Since any measure of utility is only unique up to an order-preserving/monotonic transformation, we actually cannot tell if the low-income and low-time households “suffer” more from such a change, but it is possible. These results may suggest that, even with similar utility or welfare *differences*, the ability to place a monetary or time value on such changes can be very different. As a point of comparison, the values of time which correspond to these households are shown below; the results imply that time availability, rather than solely money availability, plays a substantial role in time valuation.

Table 5-6f**Comparison of Time Valuation for Specific Household Types**

Comparison of Value-of-Time Estimates for Low/High-Income and -Time Households (\$/hour):	
Low Income Household:	\$ 5.00
High-Income Household:	\$ 26.79
Low-Discretionary-Time Household:	\$ 31.78
High-Discretionary-Time Household:	\$ 5.96

Before leaving this chapter, the stochastic nature of the presented estimates deserves some serious discussion. It is very rare that the first-order estimates of functions of

variables, such as indirect utility and expenditures as functions of estimated parameters, are the “best” estimates of these functions’ mean values. For the benefit measures considered here, this result can be written in the following form:

$$\begin{aligned} \text{Expected Monetary Benefit} &= E_{\varepsilon} \left(EV_s(\bar{t}, T, u | \varepsilon) \right) \neq EV_s(\bar{t}, T, u) \text{ generally,} \\ \& \text{ Expected Time Benefit} &= E_{\varepsilon} \left(EV_T(\bar{t}, Y, u | \varepsilon) \right) \neq EV_T(\bar{t}, Y, u), \text{ generally.} \end{aligned} \quad \text{(Chapter Five:-1)}$$

Expectation is not equal to the first-order estimates here thanks to correlation between variables and non-linear transformations of variables; the biases these relations create in such an estimate arise in important policy variables like equivalent variation and should be accounted for wherever possible. McFadden (1996) discusses this general estimation difficulty, which tends to be overlooked in the literature, and he suggests a bounding of the typical cost/benefit measures generated by logit model estimates.

ENDNOTES:

-
- ¹ The parameter estimates are positive for the first and third contour's times, negative for those of the second and fourth contour. The p-values which result from tests of the hypotheses that the true parameters equal zero are 0.004, 0.632, 0.108, and 0.000 for the coefficients on access times to the first through fourth contours, respectively.
 - ² Purvis *et al.* also point out that Goulias and Kitamura's work (1989) has found "a significant inverse relationship between work trip frequency and shopping trip and social trip frequency" (Purvis *et al.* 1996, pg. 3), a result also supported by Golob and McNally's work (1997).
 - ³ All models were estimated maintaining the marginal utility of discretionary time positive, so it was only this second condition of the sign on marginal utility of income that produces the sign on the value of time estimates.
 - ⁴ If a model is insufficiently specified with respect to one of the constraints, the expenditure function will obviously be incomplete. This is a concern for the inversion of indirect utility with respect to the income variable here, since all income effects may not be identified when applying Roy's Identity exclusively in a time setting. For example, the marginal utility of income was estimated to be negative in the Type 3 model of discretionary activity participation. Thus, those results lead to an estimate of the income expenditure function whose value falls as utility rises; this is a clearly unreasonable result.
 - ⁵ The present value of a stream of 30 payments of \$31,000 per year at a personal discount rate of ten percent per year is just over \$292,000.

Chapter Six: Limitations and Extensions

There are a variety of ways in which the models described here can be enhanced and extended. For example, more specific demand types should prove useful and allowance could be made for differences in preference structures as well as more flexible stochastic specifications. Additionally, use of longitudinal data, recognition of intra-household dynamics in choices, and incorporation of scheduling constraints may prove useful.

These extensions are discussed here now.

Data Deficiencies

In their review of activity-based travel models, Bowman and Ben-Akiva (1996) remark, “The fundamental problem facing the activity based travel modeler is combinatorial.” In other words, the high dimensionality of choice sets – particularly when the dimensions of time and space are involved – can quickly lead to computationally impractical models. The system-of-demand equations approach creates the same problem: every consumption item of interest must be identified distinctly and with a unique price; therefore, adequate incorporation of the time and space dimensions can be difficult.

Since their purpose is primarily one of illustration, the models estimated here are inherently limited in their scope. Of course, more detailed demand sets can be studied, such as different trip types by different modes, different members of the household, and different times of day to more, distinct zones. However, highly detailed demand studies may be limited by the data. For example, existing travel data tend to be general in nature. Many travel surveys divide discretionary trips into only six categories: social, personal business, food shopping, non-food shopping, dining, and recreation. These and other

categories lack information on the quality of activity. Moreover, employment and interzonal travel time data are merely rough approximations – in addition to being highly correlated across different job types, activity types, and mode types. Finally, there is the problem of sample size: more demand types means more than quadratic increases in the number of parameters requiring estimation for second-order-flexible functional forms – creating less confidence in the resulting parameter estimates.

Modeling Expanded Choice Sets

Notwithstanding the described limitations, the research presented here can be extended in several ways. Actual wage and unearned-income data would be very beneficial for the models described here because they allow the work-time and total-income decisions to be endogenously determined without relying on coarse estimates (as done here, in model Type 4). Data on activity-participation prices should also prove useful, because their incorporation may substantially strengthen marginal-utility of income estimates by allowing more model details in a system of demand equations developed within a price context. And the simultaneous estimation of *two systems* of demand equations (developed from the time and price versions of Roy's Identity) should be helpful for estimation of an *entire* indirect utility function and its resulting estimates of income's marginal utility and a monetary expenditure function.

The use of a panel data set, where households are surveyed at various points in time, may also prove useful, particularly if there is variation observed over time in households' travel-time environments. Chapter Five presents this issue, citing arguments by Kitamura (1996), Goodwin *et al.* (1990), and Becker (1965) against the use of cross-sectional data.

While data availability may place practical limitations on analysis of detailed demand systems, the model presented here is highly applicable in ways requiring far less data. For example, an experiment across a set of locations differing only in their access costs to a few specific activity types (*e.g.*, access to different sizes or qualities of shopping centers and/or parks) can, through the use of the methodologies described here, lead to an analysis of the choices of households in a microeconomically rigorous way. The same holds true for incorporating other types of demands, such as expenditures on in-home entertainment equipment, telecommunications, or other personal goods. The system of demand equations are derived as before, but the error-structure assumptions and likelihood specification for these more continuous demand types should require more flexibility.

Inclusion of Automobile Ownership in the Model

A difficulty with the stochastic specification used here (a multinomial distribution conditioned on a negative binomial) is that many consumption choices, like car ownership, are very distinct from trip-making decisions. Consider summing up cars (long-life capital goods¹) with short-run decisions (like the number of trips per week) and assuming that the relative probabilities of optimal choices represent the relative means; intuitively, this distributional assumption may seem highly unlikely.²

However, one can always include automobile ownership and other, non-activity demand types in the system of demand equations; the basic model structure described here is sufficiently flexible to accommodate these, as long as one can identify the demand level (by either observing variation in its “price” across observational units or by being able to assume that no variations in its price occur across observations). It is the desire to

incorporate correlation in unobserved heterogeneity which complicates the specification; however, model estimation using simulated likelihoods should allow for such correlations.

Location Choice Decision

Location choice is also a decision which should be accommodated to achieve a more complete model. In this research's present formulation, location is given so travel times are taken as exogenous in the activity-participation decision. However, households choose locations based partly on accessibility and expected travel expenditures³. In other words, different travel-time and travel-cost environments lead to different residential location decisions and thus different activity participation choices; for long-run predictions, one should consider travel costs' impacts on *both* the location decision and activity participation (given location) in order to consistently estimate full travel-time elasticities and welfare impacts of policies.

One may choose to model location choice in great detail (for example, over the thousand-plus census tracts one typically encounters in a major metro area) or more coarsely, with far fewer zones employing general access/travel cost information. Since travel costs faced are conditional on location choice and virtually all households choose a single location, the entire problem may be described as in the following equation.

$$\begin{aligned}
& \underset{\bar{L}, \bar{H}, \bar{A}, \bar{T}, \bar{i}\bar{A}, \bar{Z}}{\text{Max}} \text{Utility}(\bar{L}, \bar{H}, \bar{A}, \bar{T}, \bar{i}\bar{A}, \bar{Z}) \\
& \text{s.t. } \bar{P}_L \bar{L} + \bar{P}_H \bar{H} + \bar{P}_A \bar{A} + \bar{P}_{trvl,l} \bar{A} + \bar{P}_Z \bar{Z} \leq Y_{un} + wT_w \\
& \& \sum_k T_k + \bar{i}\bar{A} = T \text{ for the chosen, utility - maximizing location, } l, \\
& \sum_l L_l = 1, L_l \in \{0,1\}, \text{ and } \bar{H}, \bar{A}, \bar{T}, \& \bar{Z} \geq 0; \\
& \text{where } \bar{L} = \text{Location Choice, } \bar{H} = \text{Housing Attributes Choice,} \\
& \bar{P}_L = \text{Prices of locations, } \& \bar{P}_H = \text{Prices of housing attributes.}
\end{aligned}$$

(Chapter Six:-1)

Note that the set-up described in the above equation is not suitable for the system-of-demand-equations approach taken in this research since disaggregate parallels to Roy's Identity can only identify optimal levels of continuous choices (where price variation is observed across observations). Instead, one will probably need to model the location choice decision using a random-utility discrete-choice model (*e.g.*, McFadden 1974, Quigley 1976, Lerman 1977), conditioning the current model on this decision and maximizing the likelihood simultaneously. One possible set-up, based on a multinomial logit for location choice, is illustrated by Equation 6-2. The optimal activity choices in this likelihood (\bar{X}) cannot be identified without conditioning on location, but they can be assumed to follow the model described in this dissertation, once location is given.

Given choice of a single location,

$$\text{Prob}(\text{Choose location " } l \text{ " \& activities } \bar{X}) = \text{Prob}(\bar{X} | \text{Location } l) \text{Prob}(\text{Location } l)$$

$$\text{where } \text{Prob}(\text{Location } l) = \frac{e^{v_l}}{\sum_m e^{v_m}} \quad (\text{Chapter Six:-2})$$

$\& v_m = \text{maximized / indirect utility, given location choice } l.$

Six:-2)

Modeling Activity-Participation Times

Duration modeling in a strict system-of-demand-equations context, allowing integration back to an indirect utility specification and all the information that it provides (*e.g.*, expenditure functions and welfare measures), requires identifiability of demands via an exogenous price or constraint. But when duration is a continuous variable facing no binding constraints, the optimal level is not identifiable (using parallels to Roy's Identity). Without such identifying information, this research's modeling paradigm is insufficient for strict estimation of this dimension of activity demand. The modeling of activity durations has been studied using a system of Tobit regressions linked to binary logits (*e.g.*, Damm and Lerman 1981, Kitamura 1984) and is currently being analyzed with hazard models of individual activities (*e.g.*, Bhat 1996, Ettema *et al.* 1995b), but these approaches typically lack flexibility and economic underpinnings and tend to be limited in their predictive scope. Certainly, there is much study to be done in this area. Ideally, a single model can be developed which acknowledges the simultaneity of the various decisions and permits estimation of all such choice variables.

Incorporating Different Preference Structures

Allowance for preference differences across households is another possible area of extension. Many variables, such as day of week surveyed and age distribution of household members, may provide good measures of such differences. Techniques known as demographic scaling and translating (Pollack and Wales 1980) shift or scale parameters according to functions of the demographic variables. These are likely to be useful, even when estimating homogeneity- or summability-constrained models. However, demographic scaling and translating can add substantially to the parameter set;

more flexible techniques of incorporating demographic information are feasible when such theory-imposed constraints do not apply, such as in the time-identified system of demands studied here.

The techniques of permitting *random* variation in the parameters themselves, as used by Train (1996), McFadden and Train (1996), and Mehndiratta (1996), and more flexible correlation in the compounded error structure, as used by Yen *et al.* (1998), may also prove useful by allowing additional unobserved heterogeneity across observations. However, these methods require a simulation-of-likelihoods technique for parameter estimation. Another possibility for consideration is the specification of tractably integrated compounded error structures within ordered-choice models, such as that used by Bhat and Singh (1998) in a full-information maximum likelihood estimation of a logit and two probits – all related through errors in the latent response variables.

Recognition of Other Constraints on Behavior

Intra-household dynamics and activity scheduling constraints were not addressed in this model, though a suggestion was made for incorporating time-of-day effects by further disaggregating the demand types. A household can spread its income among its different members, but time cannot be traded except by making certain members perform specific tasks; the balancing of the competing needs and preferences of a household's distinct members is an interesting problem and has been investigated by Golob and McNally (1997). Since household members often coordinate their day-to-day activity participation, short-period observations of demand will contain many short-term dependencies; a flexible latent error structure may accommodate these effects.

Certain constraints are likely to be critical in the choice and timing of activities. For example, most shops close shortly after the end of the working day, so workers cannot participate in the same range of weekday activities enjoyed by non-workers. The models developed here can provide the input necessary for scheduling models like HAPP (Recker 1995), STARCHILD (Recker *et al.* 1986a, 1986b), and SMASH (Ettema *et al.* 1993, 1995a), where coupling, authority, and capability constraints (Hägerstrand 1970) are accommodated explicitly. Alternatively, explicit incorporation of such constraints in the utility-optimization problem and their characteristics in the resulting indirect utility function is yet another possible extension to these models.

ENDNOTES:

- ¹ High-value, long-life capital goods (*i.e.*, stocks, rather than flows) can be thought of as contemporaneous decisions with other, short-life goods if rental markets for such stocks exist and are perfectly competitive (Dubin and McFadden 1984, p. 347). Thus, where the auto-leasing market is significantly competitive with the auto-purchase market, one can reasonably incorporate auto ownership into the system of demand equations.
- ² If one assumes that the number of vehicles owned is distributed as a Poisson with mean equal to the population mean (given a set of explanatory variables) times the *same* unobserved gamma-distributed error term that trip-making depends on (for unobserved heterogeneity), one ends up with a system that looks just like a multinomial conditioned on a negative binomial. The assumption of the same gamma term does not seem too unreasonable if one believes that more trip-making typically means more dependence on personal vehicles and probably a close-to-proportional increase in vehicle miles traveled. Making this assumption produces the following likelihood:

$$X_A \sim \text{Negative Binomial} \left(m, P_A = \frac{P_T X_A^*}{(X_A^* + X_T^*)} = \frac{X_A^*}{m} \right),$$

$$\text{Prob}(X_1, X_2, \dots, X_I, X_A / p_1, p_2, \dots, p_I, X_T^*, X_A^*) =$$

$$\text{Multinomial}(\bar{X} | \bar{p}, X_T + X_A) \text{Neg. Bin}(X_A + X_T | X_A^* + X_T^*)$$

$$= \left(\frac{(X_T + X_A)!}{X_A! \prod_{i=1}^I X_i!} p_A^{X_A} \prod_{i=1}^I p_i^{X_i} \right) \left(\frac{\Gamma(X_A + X_T + m)}{\Gamma(X_A + X_T + 1) \Gamma(m)} \right) (1 - p^*)^{X_A + X_T} (p^*)^m,$$

$$\text{where } X_A = \# \text{Automobiles}, p_A = \frac{X_A^*}{(X_A^* + X_T^*)}, p_i = \frac{X_i^*}{(X_A^* + X_T^*)},$$

$$X_T = \sum_{i=1}^I X_i \text{ and } X_T^* = \sum_{i=1}^I X_i^*.$$

This model structure was attempted, but it soon became clear that the dispersion property of the auto ownership decision is very different than that found in the discretionary trip-making/activity-participation observations. After controlling for travel time and budget variables using a reasonably flexible model structure (*e.g.*, that of the modified translog), it was estimated that the variance of auto ownership is less than the mean; so the assumption of a negative binomial appears implausible.

A different, but related, assumption for the incorporation of the auto-ownership decision which still permits correlation in unobserved information may be that the number of autos owned is distributed as a binomial and that as the number of trips made deviates from the average, so does the long-run, optimal number of autos to own. The binomial's scale and probability parameters could be defined as X_T and P_{Auto} , conditioned on the negative binomial of total number of trips (X_T), with P_{Auto} specified as $\mathcal{L}_{Auto} / \mathcal{L}_{Total \#trips}$. However, this specification actually permits P_{Auto} to be greater than one, and it requires that the observed number of cars be less than or equal to the observed number of trips. These are unreasonable requirements, unless one is surveying for a sufficiently long period that all households will be making many trips.

- ³ It is of interest to note that following a survey of residents of five San Francisco Bay Area neighborhoods, Kitamura *et al.* (1994) conclude that *attitudinal* characteristics explain most of the variance they observe in respondents' travel behaviors, rather than the demographic and the simple, rather subjective neighborhood characteristics they attempt to control for. One could argue, however, that attitudes are

substantially shaped by one's environment – in addition to the authors' point that people choose their environments according to their preferences for travel and the like.

Chapter Seven: Conclusions

It is axiomatic that there would be no travel absent demand for participation in geographically separate activities. Yet, few existing models of travel behavior explicitly accommodate the derived nature of travel demand. Moreover, there is a need for a simultaneous-equations approach to a household's choice of out-of-home activity participation while maximize household utility, subject to both time and money constraints. In a review of activity-based travel demand research, Kitamura (1988) writes that a full analysis of household travel demands "is an overwhelming problem. In fact no model has been constructed that determines activity patterns on the sole basis of the utility maximization principle." (1988, pp. 20-21) The research presented here offers a highly flexible and systematic approach to these problems, making use of utility theory as a basis for behavior.

This research allows for illumination of travel-related trade-offs by households. The results include estimates of out-of-home-activity generation and distribution; income, time, own- and cross-"price" elasticities; the variability of travel-time budgets and total trip-making; and responses to changes in a variety of transportation-supply, land use, and demographic variables. The research also provides a working statistical methodology for simultaneous, closed-form estimation of cardinally ordered integer behaviors possessing unobserved heterogeneity. These behaviors that are subject to time and income constraints, within a rigorous microeconomic structure, and their estimation readily yields estimates of benefits and costs in units of both time and money.

The empirical results of this dissertation suggest that income has little effect on manifest demand for discretionary activities (after controlling for travel times and a

household's time budget); this particular result does not imply, however, that income does not exert a significant effect on the specific class of activity chosen or on monetary expenditures while engaging in activities. The results also indicate that available time exerts a strong, positive effect on all demands; yet the time-budget effects of travel-time changes are sufficiently strong that cross-travel-time elasticities are often estimated to be negative. The sections describing hypothesis tests and welfare analyses suggest that total travel time expenditures (to access discretionary activities) fall with increasing travel times and a household's time budget, not just its income, is an important determinant of its value of time.

The methodologies developed and the results demonstrated here are not merely of theoretical interest, but are meaningful to practitioners of transportation planning. They theoretically and statistically advance the modeling of travel demand, and are shown here to be empirically practical, relying on data sets typically available to metropolitan planning organizations (MPOs). The methodologies also are applicable to common policymaking situations because their inputs are both the travel times and costs that distinguish opportunities for activity participation and the discretionary-time and money budgets faced by households.

The models can be made more specific or general as desired. For example, the goods considered can be distinguished not only by distance and opportunity type but by travel mode and time of day. The application also can be local or regional. The methodology is quite flexible (though data limitations may require aggregation of some goods where more flexible functional specifications of demand are desired).

At their most elementary level, the models require information about interzonal travel times, zonal opportunity levels, income or wages, household sizes, and workers per household – information common in forecast inputs for metropolitan planning. Additional information about an area's resources and its population's demographic qualities can also be examined, to distinguish across travel modes or to anticipate preference differences based on observed characteristics. Promising extensions of the methods illustrated in this dissertation include likelihood simulation – to allow more flexible patterns of unobserved heterogeneity, inclusion of the residential location choice – so this decision can be endogenously determined, and simultaneous estimation of a *second* system of demand equations, derived using price variation.

In summary, the methods developed in this research are of theoretical interest and practical use; they advance the art and science of travel-demand modeling while providing insight into human preferences and the prediction of household activity-participation and travel choices. The flexibility and behavioral rigor of the methods make them a promising direction for travel demand theory and application to follow.

References

- Adler, T. J. 1976. "Modelling Non-Work Travel Patterns." Ph.D. Thesis, Department of Civil Engineering, MIT, Cambridge, Massachusetts.
- Barten, Anton P. 1964. "Consumer Demand Functions under Conditions of Almost Additive Preferences." *Econometrica* 32 (Jan./April): 1-38.
- Barten, Anton P. 1977. "The Systems of Consumer Demand Functions Approach: A Review." In *Frontiers of Quantitative Economics* Vol. IIIA. Papers invited for presentation at the Econometric Society Third World Congress, Toronto, 1975. North-Holland Publishing, Amsterdam.
- Bates, Grace E., and Neyman, J. 1952. "Contributions to the Theory of Accident Proneness." *University of California, Publications in Statistics* 1: 215-253.
- Becker, Gary S. 1965. "A Theory of the Allocation of Time." *Economic Journal* 75: 493-517.
- Berndt, E., B. Hall, R. Hall, and J. Hausman. 1974. "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement* 3 (4): 653-665.
- Ben-Akiva, Moshe, and Steven R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press; Cambridge, Massachusetts.
- Bhat, Chandra. 1996. "A Hazard-Based Duration Model of Shopping Activity with Nonparametric Baseline Specification and Nonparametric Control for Unobserved Heterogeneity." *Transportation Research* 30B (3): 189-207.
- Bhat, Chandra R., and Sujit K. Singh. 1998. "A Joint Model of Work Mode Choice, Evening Commute Stops, and Post-Home Arrival Stops." Paper presented at the 77th Annual Meeting of the Transportation Research Board. Washington, D.C.
- Bowman, John, and Moshe Ben-Akiva. 1996. "Activity-Based Travel Forecasting." Proceedings of the June 1996 Activity-Based Travel Demand Forecasting Conference. Travel Model Improvement Program, U.S. DOT.
- Bruzelius, Nils. 1979. *The Value of Travel Time*. London: Croom Helm.
- Burt, Oscar R., and Brewer, Durward. 1971. "Estimation of Net Social Benefits from Outdoor Recreation." *Econometrica* 39 (Sept.): 813-827.
- Cameron, Trudy Ann. 1982. "Home Weatherization Choices: Alternative Logit Models and Policy Implications." Discussion Paper No. 82-31, Department of Economics, The University of British Columbia; Vancouver, Canada.

- Cameron, Trudy Ann. 1985. "A Nested Logit Model of Energy Conservation Activity by Owners of Existing Single Family Dwellings." *Review of Economics and Statistics* 67 (May): 205-211.
- Cameron, Colin A., and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data*. Econometric Society Monograph No.30. Cambridge University Press, Cambridge.
- Caves, Douglas W., and Laurits R. Christensen. 1980. "Global Properties of Flexible Functional Forms." *The American Economic Review* 70 (3): 422-432.
- Chatfield, C., A.S.C. Ehrenberg, and G.J. Goodhardt. 1966. "Progress on a Simplified Model of Stationary Purchasing Behaviour." *Journal of the Royal Statistical Society, Series A* 129: 317-367.
- Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau. 1973. "Transcendental Logarithmic Production Frontiers." *Review of Economics and Statistics* 65: 367-383.
- Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau. 1975. "Transcendental Logarithmic Utility Functions." *The American Economic Review* 65 (No. 3, June): 367-383.
- Cicchetti, Charles J., Anthony C. Fisher, & V. Kerry Smith. 1976. "An Econometric Evaluation of a Generalized Consumer Surplus Measure: The Mineral King Controversy." *Econometrica* 44 (Nov.): 1259-1276.
- Daganzo, Carlos. 1997. *Fundamentals of Transportation and Traffic Operations*. Elsevier Science, Oxford.
- Damm, D., and Steve R. Lerman. 1981. "A Theory of Activity-Scheduling Behavior." *Environment and Planning A* (13): 703-718.
- Deaton, Angus. 1974. "A Reconsideration of the Empirical Implications of Additive Preferences." *The Economic Journal* 84 (June): 338-348.
- Deaton, Angus. 1987. "Estimation of Own- and Cross-Price Elasticities from Household Survey Data." *Journal of Econometrics* 36 (Sept./Oct.): 7-30.
- Deaton, Angus, and John Muellbauer. 1980a. "An Almost Ideal Demand System." *The American Economic Review* 70 (No. 3, June): 312-326.
- Deaton, Angus, and John Muellbauer. 1980b. *Economics and Consumer Behavior*. Cambridge University Press, Cambridge.
- DeSerpa, A.C. 1971. "A Theory of the Economics of Time." *Economic Journal* 81: 828-846.

- Dickey, J. 1978. *Metropolitan Transportation Planning*, 1st Edition. Hemisphere, Washington, D.C.
- Diewert, W.E. 1971. "An Application of the Shephard Duality Theorem, A Generalized Leontief Productions Function." *Journal of Political Economy* 79: 481-507.
- Diewert, W.E. 1974. "Intertemporal Consumer Theory and the Demand for Durables." *Econometrica* 42: 497-516.
- Domencich, Thomas A., and Daniel McFadden. 1975. *Urban Travel Demand: A Behavioral Analysis*. North-Holland Publishing, Amsterdam.
- Dubin, Jeffrey A., and Daniel L. McFadden. 1984. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica* 52 (2): 345-362.
- Ettema, Dick, Aloys Borgers, and Harry Timmermans. 1993. "Simulation Model of Activity Scheduling Behavior." *Transportation Research Record No. 1413*: 1-11.
- Ettema, Dick, Aloys Borgers, and Harry Timmermans. 1995a. "SMASH (Simulation Model of Activity Scheduling Heuristics): Empirical Test and Simulation Issues." *Activity Based Approaches: Activity Scheduling and the Analysis of Activity Patterns*. Eindhoven, The Netherlands.
- Ettema, Dick, Aloys Borgers, and Harry Timmermans. 1995b. "Competing Risk Hazard Model of Activity Choice, Timing, Sequencing, and Duration." *Transportation Research Record No. 1493*: 101-109.
- Gan, Li, and Jiming Jiang. 1997. "A Simple Test for Global Maximum." Unpublished.
- Gay, David M. 1983. "ALGORITHM 611: Subroutines for Unconstrained Minimization Using a Model/Trust-Region Approach." *ACM Transactions on Mathematical Software* 9 (4): 503-524.
- Goldman, Steven. 1998. Informal, electronic correspondence. (August 20)
- Golob, Thomas F., Martin J. Beckmann, and Yacov Zahavi. 1981. "A Utility-Theory Travel Demand Model Incorporating Travel Budgets." *Transportation Research* 15B (6): 375-389.
- Golob, Thomas F., and Michael G. McNally. 1997. "A Model of Activity Participation and Travel Interactions between Household Heads." *Transportation Research B*, 31 (3).
- Golob, Thomas F., and Henk Meurs. 1987. "A Structural Model of Temporal Change in Multi-Modal Travel Demand." *Transportation Research* 21A (6): 391-400.
- Golob, Thomas F., and Leo van Wissen. 1989. "A Joint Household Travel Distance Generation and Car Ownership Model." *Transportation Research B*, 23 (6): 471-491.

- Golob, Thomas F., and Leo van Wissen. 1990. "Simultaneous Equation Systems Involving Binary Choice Variables." *Geographical Analysis*, 22: 224-243.
- Goodwin, Phil B., Ryuichi Kitamura, and Hank Meurs. 1990. "Some Principles of Dynamic Analysis of Travel Behaviour." In *Developments in Dynamic and Activity-Based Approaches to Travel Analysis*, Peter Jones (Editor). Brookfield, Vermont: Gower Publishing.
- Goulias, Konstandinos G., and Ryuichi Kitamura. 1991. "Recursive Model System for Trip Generation and Trip Chaining." *Transportation Research Record No. 1236*: 59-66.
- Greene, William H. 1993. *Econometric Analysis*. McMillan, New York.
- Gronau, Reuben. 1970. *The Value of Time in Passenger Transportation: The Demand for Air Travel*. National Bureau of Economic Research, New York, Occasional Paper No. 109.
- Guilkey, David K., C.A. Knox Lovell, and Robin C. Sickles. 1983. "A Comparison of the Performance of Three Flexible Functional Forms." *International Economic Review* 24 (Oct.): 591-616.
- Gur, Yehuda. 1971. "An Accessibility Sensitive Trip Generation Model." Chicago Area Transportation Studies, Chicago.
- Hägerstrand, Torsten. 1970. "What about People in Regional Science?" *Regional Science Association Papers* 24: 7-21.
- Hanson, Susan, & M. Schwab. 1987. "Accessibility and Intraurban Travel." *Environment and Planning A* 19: 735-748.
- Harvey, Greig W., and Elizabeth Deakin. 1996. "Description of the STEP Analysis Package." Unpublished paper. (February 1996)
- Hausman, Jerry A., Bronwyn H. Hall, and Zvi Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R and D Relationship." *Econometrica*, 52: 909-938.
- Hausman, J.A., G.K. Leonard, and Daniel McFadden. 1995. "A Utility-Consistent, Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses due to Natural Resource Demand." *Journal of Public Economics*, 50: 1-30.
- Heap, Shaun Hargreaves, Martin Hollis, Bruce Lyons, Robert Sugden, and Albert Weale. 1992. *The Theory of Choice: A Critical Guide*. Blackwell Publishing, Oxford, U.K.
- Hicks, J.R. 1956. *A Revision of Demand Theory*. Oxford University Press, Oxford, U.K.

Howe, H., Robert A. Pollack, and Terence J. Wales. 1977. "Theory and Time Series Estimation of the Quadratic Expenditure System." Discussion Paper #388, University of Pennsylvania. (Oct.)

Houthakker, Hendrik S. 1960. "Additive Preferences." *Econometrica* 28 (April): 244-257.

ITE Journal. 1994. "Travel Demand Forecasting Processes used by 10 Large MPO's." Feb., pg. 32.

Jara-Díaz, Sergio R. 1994. "A General Micro-Model of Users' Behavior: The Basic Issues." *Proceedings of the Seventh International Conference on Travel Behaviour*, pp. 91-103.

Johnson, M. Bruce. 1966. "Travel Time and the Price of Leisure." *Western Economic Journal* 4: 135-145.

Jorgenson, Dale W., and Lawrence J. Lau. 1979. "The Integrability of Consumer Demand Functions." *European Economic Review* 12: 115-147.

Kawakami, Shogo, and Tomohiko Isobe. 1990. "Development of a One-Day Travel-Activity Scheduling Model for Workers." In *Developments in Dynamic and Activity-Based Approaches to Travel Analysis* (Ed. P. Jones), pp. 184-205. Gower Publishing, Aldershot, England.

Kim, Soon-Gwan. 1994. "Econometric Modeling of All-Day Activity Involvement and Duration with Transportation Panel Data." Unpublished Dissertation, University of Washington, Department of Civil Engineering.

Kitamura, Ryuichi. 1984. "A Model of Daily Time Allocation to Discretionary Out-of-Home Activities and Trips." *Transportation Research B* 18 (3): 255-266.

Kitamura, Ryuichi. 1988. "An Evaluation of Activity-Based Travel Analysis." *Transportation* 15: 9-34.

Kitamura, Ryuichi. 1990. "Panel Analysis in Transportation Planning: An Overview." *Transportation Research* 24A (6).

Kitamura, Ryuichi. 1996. "Applications of Models of Activity Behavior for Activity Based Demand Forecasting." Proceedings of the June 1996 Activity-Based Travel Demand Forecasting Conference. Travel Model Improvement Program, U.S. DOT. (Paper was revised August 1996)

Kitamura, Ryuichi, and Mohammad Kermanshah. 1984. "Sequential Model of Interdependent Activity and Destination Choices." *Transportation Research Record No. 987*: 81-89.

Kitamura, Ryuichi, Kazuo Nishii, and Konstadinos Goulias. 1990. "Trip Chaining Behaviour by Central City Commuters: A Causal Analysis of Time-Space Constraints." In *Developments in Dynamic and Activity-Based Approaches to Travel Analysis*. P. Jones (Ed.). Gower Publishing, Aldershot, England: 145-170.

Kitamura, Ryuichi, Patricia L. Mokhtarian, Laura Laidet. 1994. "A Micro-Analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area." Institute of Transportation Studies, UC Davis (UTD-ITS-RR-94-28).

Kitamura, Ryuichi, Eric I. Pas, Clarisse V. Lula, T. Keith Lawton, and Paul E. Benson. 1996. "The Sequenced Activity Mobility Simulator (SAMS): An Integrated Approach to Modeling Transportation, Land Use, and Air Quality." *Transportation* 23 (3): 267-291.

Kockelman, Kara. 1996. "Travel Behavior as a Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from the San Francisco Bay Area." Unpublished Master's Thesis. Department of City and Regional Planning, University of California at Berkeley.

Kockelman, Kara. 1997. "Travel Behavior as a Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from the San Francisco Bay Area." *Transportation Research Record No. 1607*: 116-125.

Lau, Lawrence J. 1986. "Functional Forms in Econometric Model Building." In *Handbook of Econometrics III*, Eds. Z. Griliches and M.D. Intriligator. Elsevier Science, New York.

Lerman, Steven. 1977. "Location, Housing, Automobile Ownership, and the Mode to Work: A Joint Choice Model." *Transportation Research Record No. 610*: 6-11.

Leser, C.E.V. 1941. "Family Budget Data and Price Elasticities of Demand." *Review of Economic Studies* 9: 40-57.

Lu, Xuedong, and Eric Ivan Pas. 1997. "A Structural Equation Model of the Relationships among Socio-Demographics, Activity Participation and Travel Behavior." Paper Presented at the 76th Annual Transportation Research Board Meeting. Washington, D.C.

Mannering, Fred, and Clifford Winston. 1985. "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization." *Rand Journal of Economics* 16 (2): 215-236.

McFadden, Daniel. 1964. "Existence Conditions for Theil-Type Preferences." Mimeograph, University of California, Berkeley.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*. P. Zarembka (Ed.). Academic Press, New York: 105-142.

- McFadden, Daniel. 1996. "On Computation of Willingness-To-Pay in Travel Demand Models." Paper presented at the *Theoretical Foundations of Travel Choice Modeling Conference*, held in Stockholm, Sweden, August 7-11.
- McFadden, Daniel and Kenneth Train. 1996. "Mixed MNL Models for Discrete Response." Working Paper; Department of Economics, University of California, Berkeley.
- Mehndiratta, Shomik. 1996. "Time-of-Day Effects in Intercity Business Travel." Unpublished Ph.D. Thesis; Department of Civil and Environmental Engineering, University of California, Berkeley.
- MTC. 1996. "San Francisco Bay Area 1990 Travel Model Development Project: Compilation of Technical Memoranda." Vol. 3. Metropolitan Transportation Commission, Oakland, CA.
- Meurs, Henk. 1990. "Trip Generation Models with Permanent Unobserved Effects." *Transportation Research B* 24 (2): 145-158.
- Oort, C.J. 1969. "The Evaluation of Travelling Time." *Journal of Transport Economics and Policy* 3: 279-286.
- Ortúzar, Juan de Dios, & Luís G. Willumsen. 1994. *Modelling Transport* 2nd Edition. Wiley & Sons, England.
- Pollack, Robert A., and Terence J. Wales. 1978. "Estimation of Complete Demand Systems from Household Budget Data: The Linear and Quadratic Expenditure Systems." *The American Economic Review* 68 (No. 3): 348-359.
- Pollack, Robert A., and Terence J. Wales. 1980. "Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects." *Econometrica* 48 (No. 3): 595-612.
- Price, Colin. 1993. *Time, Discounting, and Value*. Blackwell Publishing, Oxford, U.K.
- Purvis, Charles, Miguel Iglesias, and Victoria A. Eisen. 1996. "Incorporating Work Trip Accessibility in Non-Work Trip Generation Models in the San Francisco Bay Area." *Transportation Research Record No. 1556*.
- Quigley, John. 1976. "Housing Demand in the Short Run: An Analysis of Polytomous Choice." *Explorations in Economic Research*, 3: 76-102.
- RAC (Royal Automobile Club). 1995. "Car Dependence: A Report for The RAC Foundation for Motoring and the Environment." RAC Foundation for Motoring and the Environment.
- Rao, B.R., S. Mazumdar, J.M. Waller, and C.C. Li. 1973. "Correlation Between the Numbers of Two Types of Children in a Family." *Biometrics*, 29: 271-279.

- RDC, Inc. 1995. "Activity-Based Modeling System for Travel Demand Forecasting". U.S. Department of Transportation, and the U.S. Environmental Protection Agency. Washington, D.C.
- Recker, W.W. 1995. "The Household Activity Pattern Problem: General Formulation and Solution." *Transportation Research B* 29(1): 61-77.
- Recker, W.W., M.G. McNally, and G.S. Root. 1986a. "A Model of Complex Travel Behavior: Part I – Theoretical Development." *Transportation Research* 20A (4): 307-318.
- Recker, W.W., M.G. McNally, and G.S. Root. 1986b. "A Model of Complex Travel Behavior: Part II – An Operational Model." *Transportation Research* 20A (4): 319-330.
- Rice, John. 1995. *Mathematical Statistics and Data Analysis*, 2nd Edition. Duxbury Press.
- Roy, René. 1943. *De l'Utilité: Contribution à la Théorie des Choix*. Hermann, Paris.
- Safwat, K.N.A., and T. Magnanti. 1988. "A Combined Trip Generation, Trip Distribution, Modal Split, and Trip Assignment Model." *Transportation Science* 22(1).
- Slutsky, E. 1915. "Sulia Teoria del Bilancio des Consomatore." ("On the Theory of the Budget of the Consumer.") *Giornale degli Economisti* 51: 1-26. English translation in 1952: *Readings in Price Theory*, Eds. G.J. Stigler and K.E. Boulding. Chicago University Press.
- Smith, M.E., and C.E. Schoener. 1978. "Testing for Induced Tripmaking and Travel in Providence, Rhode Island." Paper Presented at the 57th Annual Meeting of the Transportation Research Board, Washington, D.C.
- Sommeyer, W.H., G.M. Hilhorst, and J.W.W.A. Wit. 1962. "A Method of Estimating Price and Income Elasticities from Time Series and Its Application to Consumers; Expenditures in the Netherlands, 1949-1959." Statistical Studies No. 13. Central Bureau of Statistics, The Hague.
- Stone, J.R.N. 1954. "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand." *The Economic Journal* 64 (Sept.): 510-527.
- Theil, Henri. 1965. "The Information Approach to Demand Analysis." *Econometrica* 33 (Jan.): 67-87.
- Train, Kenneth. 1996. "Unobserved Taste Variation in Recreation Demand Models." Working Paper; Department of Economics, University of California, Berkeley.
- Train, Kenneth, and Daniel McFadden. 1978. "The Goods/Leisure Tradeoff and Disaggregate Work Trip Mode Choice Models." *Transportation Research* 12: 349-353.

Train, Kenneth, Daniel McFadden, Moshe Ben-Akiva. 1987. "The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices." *Rand Journal of Economics* 18 (1): 109-123.

Varian, Hal. 1992. *Microeconomic Analysis*, Third Edition. W.W. Norton & Company, New York.

Yen, Jin-Ru, Hani S. Mahmassani, and Robert Herman. 1998. "A Model of Employee Participation in Telecommuting Programs Based on Stated Preference Data." In *Travel Behaviour Research: Updating the State of Play*. Eds. Juan de Dios Ortúzar, David A. Hensher, and Sergio R. Jara-Díaz. Pergamon, Oxford. Zahavi, Yacov. 1979a. "The 'UMOT' Project." Prepared for US DOT, Report No. DOT-RSPA-DPB-20-79-3, Washington, D.C.

Zahavi, Yacov. 1979b. "Travel over Time." Prepared for the US DOT, Federal Highway Administration, FHWA PL-79-004.

Zahavi, Yacov, and M. Ryan. 1980. "The Stability of Travel Components over Time." *Transportation Research Record No. 200*.

Zahavi, Yacov, and Antti Talvitie. 1980. "Regularities in Travel Time and Money Expenditures." *Transportation Research Record No. 199*.

Zahavi, Yacov, Martin J. Beckmann, Thomas F. Golob. 1981. "The 'UMOT'/Urban Interactions." Prepared for U.S. DOT, Report No. DOT-RSPA-DPB-10-7, Washington, D.C.

Zellner, A. 1962. "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias." *Journal of the American Statistical Association* 58: 348-368.

APPENDIX

A-1: List of Possible Functional Form Specifications

Cobb-Douglas:

$$u = \prod_i X_i^{\alpha_i}$$

$$v = \prod_i \left(\frac{P_i}{Y} \right)^{\alpha_i}$$

While satisfying regularity conditions globally (*e.g.*, monotonicity/non-satiation of preferences and strict quasiconvexity (implying that the matrix of cross-elasticities is negative semi-definite), this form is highly inflexible for systems of 3 or more goods; it implies homothetic and additive preferences and thus restricts all elasticities of substitution to equal one (Deaton 1974, Christensen *et al.* 1975).

Stone's Linear Expenditure System (LES, 1954):

$$v = - \frac{\prod_i P_i^{\alpha_i}}{Y - \sum_i \beta_i P_i}$$

$$P_i X_i = P_i \beta_i + \alpha_i (Y - \sum_j \beta_j P_j)$$

s.t. Summability: $\sum_i \alpha_i = 1.$

This system implies constant marginal expenditures (and demand) with respect to income. Zero-degree homogeneity of demands and Slutsky symmetry are automatic. Note that this functional form comes from a utility function which can be written in linear logarithmic form, and is thus both additive and homothetic; these properties imply that all resulting expenditure portions are constant, all elasticities of substitution equal one* (Christensen *et al.* 1975), and uncompensated price derivatives of demand are symmetric. The number of parameters requiring estimation is $2I-1$ (where I is the number of distinct good types being modeled).

* Note: An elasticity of substitution is the dimensionless version of the derivative of the ratio of two goods with respect to their marginal rate of substitution (MRS). The MRS is a utility-constant measure of substitution between two goods. The following equations illustrate this definition:

$$MRS_{ij} = - \frac{du/dX_j}{du/dX_i} = \text{Rate of substituting } X_i \text{ for } X_j \text{ to keep utility constant.}$$

$$\text{Elasticity of substitution}_{ij} = \frac{d\left(\frac{X_i}{X_j}\right)}{d(MRS_{ij})} \times \frac{MRS_{ij}}{\left(\frac{X_i}{X_j}\right)}.$$

Howe, Pollack, and Wale's Quadratic Expenditure System (QES, 1977):

$$v = -\frac{\prod_i P_i^{a_i}}{Y - \sum_i b_i P_i} + \lambda \frac{\prod_i P_i^{a_i}}{\prod_i P_i^{c_i}}$$

$$P_i X_i = b_i P_i + a_i (Y - \sum_j b_j P_j) + (c_i - a_i) \lambda \prod_j P_j^{-c_j} (Y - \sum_k b_k P_k)^2$$

s.t. *Summability (& Homogeneity)*: $\sum_i a_i = 1$ & $\sum_i c_i = 1$.

Note that the LES is a special case of the QES and the QES does not impose constant marginal budget shares. In both the LES and QES the number of parameters is a linear function of the number of good-classes considered (*i.e.*, $2I-1$ & $3I-1$).

Barten (1964) and Theil's (1965) Rotterdam Model:

Starting from: $\log(X_i) = \alpha_i + \eta_{Y,i} \log(Y) + \sum_j \eta_{i,j} \log(P_j)$,

one can arrive at:

$$w_i d \log(X_i) = \beta_i d \log(\bar{Y}) - \sum_j \gamma_{ij} d \log(P_j),$$

where $d \log(\bar{Y}) = \sum_j w_j d \log(X_j)$,

$$\beta_i = w_i \eta_{Y,i}, \text{ \& } \gamma_{ij} = w_i \eta_{ij}^*.$$

Note that $\eta_{Y,i}$ = Income Elasticity of *i*'th good, &

η_{ij} & η_{ij}^* = Uncompensated & Compensated Cross – Price Elasticities.

Summability: $\sum_j \beta_j = 1$ & $\sum_i \gamma_{ij} = 0$,

Homogeneity: $\sum_i \gamma_{ij} = 0$, & Symmetry: $\gamma_{ij} = \gamma_{ji}$.

The primary equation is estimated after replacing the differentials with finite approximations and treating the parameters as though they are constants. (Deaton & Muellbauer, 1980a) As noted by McFadden (1964), the Rotterdam model with constant parameters is consistent with utility maximization only if the utility function can be written in linear logarithmic form, which then, like the LES, is both additive and homothetic. As with the LES, these properties imply that all resulting expenditure portions are constant and all elasticities of substitution equal one (Christensen *et al.* 1975), cross-price elasticities equal one, and own-price elasticities equal negative one (Deaton & Muellbauer 1980a).

Johansen's (Very) General Additive Utility Function:

$$v = \sum_i \frac{\beta_i}{\alpha_i} \left(\frac{P_i}{Y - Y_o} - \rho_i \right)^{\frac{\alpha_i}{\alpha_i - 1}},$$

$$\text{where } \alpha_i < 1, \beta_i > 0, \rho_i < \frac{P_i}{Y - Y_o}, \& Y_o = \left(\sum_k \gamma_k P_k^v \right)^{1/v} < Y;$$

$$X_i = \gamma_i \left(\frac{Y_o}{P_i} \right)^{1/v} + (Y - Y_o) \left(\frac{\beta_i \left(\frac{P_i}{Y - Y_o} - \rho_i \right)^{\frac{1}{\alpha_i - 1}}}{\sum_k \beta_k P_k \left(\frac{P_k}{Y - Y_o} - \rho_k \right)^{\frac{1}{\alpha_k - 1}}} \right), \forall i = 1, \dots, n - 1.$$

Number of parameters requiring estimating is on the order of $4I$. This specification is almost never used in practice, although specializations of this function (e.g., Direct Addilog & LES) are used.

Leser (1941), Somermeyer et al. (1962), & Houthakker's Direct Addilog (1960):

$$v = \sum_i \frac{\beta_i (\alpha_i - 1)}{\alpha_i} \left(\frac{P_i}{M} \right)^{\frac{\alpha_i}{\alpha_i - 1}}$$

$$X_i = \beta_i \left(\frac{P_i}{M} \right)^{\frac{1}{\alpha_i - 1}} \bigg/ \sum_k \beta_k \left(\frac{P_k}{M} \right)^{\frac{\alpha_k}{\alpha_k - 1}}$$

s.t. β scaling (e.g., $\sum \beta = 1$).

If $\alpha_i = \alpha \forall i$, this is CES system.

If $\alpha_i = 0 \forall i$, this is Cobb – Douglas system.

Christensen et al.'s Indirect Transcendental Logarithmic Utility Function:

$$v = \kappa + \sum_i \beta_i \ln \left(\frac{P_i}{Y} \right) + \sum_{ij} (1/2) \beta_{ij} \ln \left(\frac{P_i}{Y} \right) \ln \left(\frac{P_j}{Y} \right),$$

$$X_i = \left(\frac{Y}{P_i} \right) \left(\alpha_i + \sum_j \beta_{ij} \ln(P_j/Y) \right) \bigg/ \left(\sum_j \alpha_j + \sum_{ij} \beta_{ij} \ln(P_j/Y) \right),$$

s.t. Symmetry: $\beta_{ij} = \beta_{ji}$,

and Identifiability: $\sum_i \alpha_i + \sum_{ij} \beta_{ij} = 1, \forall ij$.

Summability and homogeneity are automatic. The number of independent parameters for this specification (when symmetry is imposed) is $\frac{I^2 + 3I - 2}{2}$.

Diewert's Generalized Leontief:

$$v = \left[2 \sum_i \alpha_i (P_i/Y)^{\frac{1}{2}} + \sum_{ij} \beta_{ij} \left(\frac{P_i}{M} \right)^{1/2} \left(\frac{P_j}{M} \right)^{1/2} \right]^{-1}$$

$$X_i = \frac{Y}{P_i} \left[K_i / \left(\sum_j K_j \right) \right],$$

where $K_j = \alpha_j (P_j/Y)^{\frac{1}{2}} + \sum_l \beta_{jl} (P_j/Y)^{\frac{1}{2}} (P_l/Y)^{\frac{1}{2}}$

Note: This formulation is very general and comes from Caves and Christensen (1980); Diewert generally appears to have used less complex forms (e.g., Diewert 1974). Note that the constraint $\alpha_i = 0 \forall i$ implies homothetic preferences here.

Deaton and Muellbauer's Almost Ideal Demand System:

$$e(u, \bar{p}) = \alpha_o + \sum_k \alpha_k \ln(P_k) + \frac{1}{2} \sum_{kj} \gamma_{kj}^* \ln(P_k) \ln(P_j) + u \beta_o \prod_k P_k^{\beta_k},$$

$v = e^{-1}(u, \bar{p})$ with respect to u ,

$$X_i = \frac{Y}{P_i} \left[\alpha_i + \sum_j \gamma_{ij} \ln(P_j) + \beta_i \ln(Y/P^*) \right],$$

$$\gamma_{ij}^* = \frac{1}{2} (\gamma_{ij} + \gamma_{ji}) \text{ \& } \ln(P^*) = \ln(\text{Price Index})$$

$$= \alpha_o + \sum_k \alpha_k \ln(P_k) + \sum_{jk} \gamma_{jk} \ln(P_j) \ln(P_k),$$

$$\text{summability: } \sum_i \alpha_i = 1, \sum_i \gamma_{ij} = 0, \text{ \& } \sum_i \beta_i = 0,$$

$$\text{homogeneity: } \sum_j \gamma_{ij} = 0, \text{ \& symmetry: } \gamma_{ij} = \gamma_{ji}.$$

Note: In the typical set-up, AIDS aggregates “perfectly” (without requiring parallel expenditure expansion paths of different consumers/households) and expenditure *shares* [$w_i = (X_i P_i/Y)$] can be estimated in a linear fashion (except for the price index, P^*), subject to linear constraints. In practice, “econometricians typically use an arbitrary price index to calculate the (Y/P^*) terms” (Varian 1992, pg. 213) and estimate the remaining parameters via a linear system.

A-2: Derivation of Roy's Identity

Roy's Identity (1943) derives from constrained maximization of a direct utility function and can be generated for a general two-exogenous-constraints situation by beginning with the following:

$$\text{Max}_{\bar{A}, \bar{T}, \bar{Z}} \text{Utility}(\bar{A}, \bar{T}, \bar{t}\bar{A}, \bar{Z})$$

$$\text{s.t. } \bar{P}_A \bar{A} + \bar{P}_{trvl} \bar{A} + \bar{P}_Z \bar{Z} \leq Y, \sum_k T_k + \bar{t}\bar{A} = T, \text{ and } \bar{A}, \bar{T} \text{ \& } \bar{Z} \geq 0.$$

This problem formulation results in the following Lagrangian equation and first-order conditions for maximization:

$$L(\bar{A}, \bar{T}, \bar{Z}, \lambda_{Time}, \lambda_{Money}) =$$

$$U(\bar{A}, \bar{T}, \bar{t}\bar{A}, \bar{Z}) + \lambda_{Money} (Y - \bar{P}_A \bar{A} - \bar{P}_{trvl} \bar{A} - \bar{P}_Z \bar{Z}) + \lambda_{Time} \left(T - \sum_k T_k - \bar{t}\bar{A} \right)$$

$$L^{opt} = v(\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y, T) =$$

$$L(\bar{A}^*[\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y, T], \bar{T}^*[\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y, T], \bar{Z}^*[\bar{P}_A, \bar{P}_{trvl}, \bar{P}_Z, \bar{t}, Y, T]),$$

$$\frac{dL^{opt}}{dP_{A_i}} = \frac{dv}{dP_{A_i}} = \frac{dL}{dP_{A_i}} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dP_i} = -\lambda_{Money} A_i^* + 0,$$

$$\equiv \frac{dL^{opt}}{dP_{trvl,i}} = \frac{dv}{dP_{trvl,i}} = \frac{dL}{dP_{trvl,i}} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dP_{trvl,i}} = -\lambda_{Money} A_i^* + 0,$$

$$\frac{dL^{opt}}{dt_i} = \frac{dv}{dt_i} = \frac{dL}{dt_i} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dt_i} = -\lambda_{Time} A_i^* + 0,$$

$$\frac{dL^{opt}}{dY} = \frac{dv}{dY} = \frac{dL}{dY} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dY} = \lambda_{Money} + 0, \text{ and}$$

$$\frac{dL^{opt}}{dT} = \frac{dv}{dT} = \frac{dL}{dT} + \frac{dL}{dA_i} \Bigg|_{\substack{\bar{A}=\bar{A}^*, \bar{T}=\bar{T}^* \\ \bar{Z}=\bar{Z}^*}} \times \frac{dA_i^*}{dT} = \lambda_{Time} + 0.$$

With some minor manipulations of the above, one has the following relation as the optimal number of times to participate in activity i :

$$A_i^* = -\frac{\frac{dv}{dt_i}}{\frac{dv}{dT}} = -\frac{\frac{dv}{dP_{A_i}}}{\frac{dv}{dY}} \equiv -\frac{dP_{trvl,i}}{dY}, \forall i.$$

Note that $dv/dP_{A_i} = dv/dP_{trvl,i}$, so there are only two distinct ratios in the above relation; nevertheless, this relationship imposes many more constraints across parameter

sets than would a single equality for each demand in a traditional, money-based system-of-equations set-up. However, the above relationship may not be very rigorously applied in its entirety because the purchase prices of activities (P_{A_i}) and the travel prices ($P_{trvl,i}$) are not known/provided in most data sets, so many cross-equation parameter constraints are concealed by unknown price levels and one may end up having to rely on many constant terms, rather than the more interesting interactions of variables for explanatory information.

Furthermore, the use of Roy's Identity in identifying the optimal amount of *time* to be spent in each activity i (T_i^*) is not feasible when there are not clear "prices" attached to each time expenditure or clear minimum-time constraints, such as those DeSerpa invokes (1971). For this reason, the question of optimal time expenditures was not considered in this research.

A-3: Description of Data Set Used***Bay Area Travel Surveys (BATS)***

Usable Sample Size: 10,834 households, ~21,300 individuals

Types of Information Asked:

Demographic:

Age, gender, household income, education, employment, driver's license, ...

Location:

Census tract of residence and work

Tenure of dwelling unit (own vs. rent)

Travel Diaries:

1, 3, and 5-day activity diary on all household members age five and over;
~9,400 households surveyed for a single day, and ~1,400 households
surveyed for three or five days

16 trip purposes, trip durations, travel modes and times and fares, parking at
destination)

Vehicle ownership

Other Data for Use:

Interzonal travel times & zonal land-use and employment characteristics.

Other Data Set Possibilities: 1994/1995 Portland Area Activity Surveys

Sample Size: 4,451 households, 10,048 individuals

Types of Information Asked:

Demographic:

Age, gender, income, education, employment, driver's license, ...

Location:

Census tract of residence and work

Tenure of dwelling unit (own vs. rent)

Activity Participation:

2-day activity diary on all household members (24 activity types, durations, inter-activities travel modes and times and fares, parking at destination)

All out-of-home activity durations and all in-home activities of duration \geq 30 minutes

Vehicle ownership

Other Data for Use:

Interzonal travel times & zonal land-use and employment characteristics.

Advantages:

Provides some weekend data (approximately 2,000 of the household-days surveyed were a Saturday or Sunday, while close to 7,000 were weekdays).

Availability of land-use and travel-time data for the region.

Weaknesses:

29% of the reported trips lack either origin- or destination-zone information, so over 50% of the surveyed households have incomplete activity-location information, essentially rendering them unusable.

Other Data Set Possibilities: 1990 U.K. RAC Data (RAC 1995)

Sample Size: 392 Adults, 280 Households, 13 Different Towns/Areas

Types of Information Asked:

Demographic:

Age, Gender, Income, Education, Employment, Driver's License, ...

Location:

How long living at current location; where lived previously & why moved

#Dwelling units in structure; Tenure & mortgage/rent/...

#Bedrooms, bathrooms, kitchens, ...; Nearby friends, relatives, ...

Activity Participation:

7-day activity diary on all household members (9 activity types, durations, inter-activities travel modes and times and fares, parking at destination)

Hours per week in work & school

Household chores, types & hours/day

Child care, hours/day and trip needs for children

Travel Modes & Expenditures:

Vehicle ownership, parking availability at home location, ...

Expenditures on each vehicle (for insurance, maint., parking, & road taxes)

VMT per vehicle for one week

Primary travel modes to work, school, shop, ... & travel times

Availability of alternative modes (other than primary mode used)

Employer provision of parking & parking costs

Amount spent per week on transit

Distance, usual mode, & travel time to local newsagent, food store, doctor's office, rail station, & bus stop

Bus frequency to main shopping center

Non-Travel Expenditures:

Costs of groceries, rent/mortgage, & utilities per month

Information:

Knowledge of transit supply locally

Attitudes:

Toward driving, traffic, transit, & the environment

Toward personal-vehicle attributes (for car purchases)

Toward possible public travel-related policies

Advantages:

Covers a week's worth of activity participation and trip-making for individuals.

Provides information that is likely to indicate taste differences among the households, as well as information on dwelling-unit choice and other significant consumption.

Weaknesses:

Very limited sample size.

Lack of local land-use and interzonal travel-time data.

A-4: Description of Negative Binomial Distribution

In addition to its usefulness as the result of mixing a Poisson with a gamma distribution, a negative binomial represents the number of *failures* (“ N ”) before “ m ” successes are achieved in a series of independent Bernoulli trials where the probability of success is “ p^* ”. Thus, the distribution has the following properties:

$$Prob(N = k) = \frac{\Gamma(k + m)}{\Gamma(m)!(k!)} (p^*)^m (1 - p^*)^k,$$

$$E(N) = \frac{m(1 - p^*)}{p^*} = mP,$$

$$\text{and } V(N) = mP(1 + P).$$

Note that the mean and variance equations of this distribution are quite similar to those for a (positive) binomial, except that “ Q ”=1+ P (vs. $q=1-p$). Recognize that negative binomials do not require an integer “ m ”, so the likelihood functions specified here incorporate gamma functions, rather than factorials, for their combinatorials. Thus, “ $k+m-1$ choose k ” can be written as the following:

$$Combo(k + m - 1, k) = \frac{\Gamma(k + m)}{\Gamma(m)\Gamma(k + 1)},$$

$$\text{where } \Gamma(s) = \int_0^{\infty} e^{-x} x^{s-1} dx \quad [= (s - 1)! \text{ if } s \text{ is integer}].$$

To show that the mean and the variance of the result of mixing a Poisson with a gamma are the same as those for the negative binomials used in this research, the following formulae are helpful:

$$\text{If } \lambda_i \sim \text{Gamma}\left(m, \frac{m}{X_i^*}\right), \text{ then}$$

$$\mu_{\lambda_i} = X_i^* \text{ \& } \sigma_{\lambda_i}^2 = \frac{(X_i^*)^2}{m};$$

$$\text{and if } X_i \sim \text{Poisson}(\lambda_i), \text{ then}$$

$$\mu_{X_i|\lambda_i} = \sigma_{X_i|\lambda_i}^2 = \lambda_i.$$

$$\text{So, } \mu_{X_i} = \mu_{\lambda_i} = X_i^*, \text{ and}$$

$$\sigma_{X_i}^2 = E_{\lambda_i}[\sigma_{X_i|\lambda_i}^2] + V_{\lambda_i}[\mu_{X_i|\lambda_i}]$$

$$= E_{\lambda_i}(\lambda_i) + V_{\lambda_i}(\lambda_i) = X_i^* + \frac{(X_i^*)^2}{m} = X_i^* + \alpha(X_i^*)^2.$$

A-5: Examples of the Estimation Algorithms, as programmed in SPlus3.3

COMMAND FILE FOR MAXIMUM LIKELIHOOD ESTIMATION:

```

source("readData.s")
source("likelihood.s")

# starting values for estimation:

startPars <- list(
  a = 1,
  a1 = -1,
  a2 = -1,
  a3 = -1,
  a4 = -1,

  b11 = 0,
  b12 = 0,
  b13 = 0,
  b14 = 0,
  b22 = 0,
  b23 = 0,
  b24 = 0,
  b33 = 0,
  b34 = 0,
  b44 = 0,

  g1 = 1,
  g2 = 1,
  g3 = 1,
  g4 = 1,

  g1T = 1,
  g2T = 1,
  g3T = 1,
  g4T = 1) # gTY is fixed to equal +1 (for identifiability of other parameters)

cat("dim data is ",dim(dataMU),"\n")

# calculate the MLE's
origFitorigMU <- ms(
~negLogLikelihoodMU(x1,x2,x3,x4,t1,t2,t3,t4,dta,Y,days,a,
  a1,a2,a3,a4,b11,b12,b13,b14,b22,b23,b24,b33,b34,b44,g1,g2,g3,g4,
  g1T,g2T,g3T,g4T,gTY),
  data = dataMU,
  start = startPars,
  control = list(maxiter = 400, scale=c(0.1,rep(100,22)), maxfcalls = 800, tol=1e-4, rel.tol=1e-5),
  trace=T)

#report the parameter estimates and derivatives
start <- unlist(startPars)
final <- origFitorigMU$par
source("dLikelihoodMU.s")
vcov <- dLogLikelihoodMU(origFitorigMU$par,dataMU)

```

```
pvals <- 2*(1-pnorm(abs(origFitorigMU$par/sqrt(diag(vcov))))))
round(cbind(start,final,sqrt(diag(vcov)),pvals),digit=5)
cat("TOTAL final function value:",origFitorigMU$value,"\n")
```

NEGATIVE LOG-LIKELIHOOD FUNCTION SPECIFICATION

(this file sourced from Command file as "likelihood.s", for the Type 2 Model specification):

```
BIG <- 1e10
SMALL <- 1e-10
pos <- function(x) {
  x[x<SMALL] <- SMALL
  x }

negLogLikelihood <- function(x1,x2,x3,x4,t1,t2,t3,t4,dta,Y,days,a,a1,a2,a3,a4,
  b11,b12,b13,b14,b22,b23,b24,b33,b34,b44,g1,g2,g3,g4,g1T,g2T,g3T,g4T,gTY,type=2) {

  n1 <- pos((-1/t1)*(a1+(b11*log(t1)+b12*log(t2)+b13*log(t3)+b14*log(t4))+g1*log(Y)+g1T*log(dta)))
  n2 <- pos((-1/t2)*(a2+(b12*log(t1)+b22*log(t2)+b23*log(t3)+b24*log(t4))+g2*log(Y)+g2T*log(dta)))
  n3 <- pos((-1/t3)*(a3+(b13*log(t1)+b23*log(t2)+b33*log(t3)+b34*log(t4))+g3*log(Y)+g3T*log(dta)))
  n4 <- pos((-1/t4)*(a4+(b14*log(t1)+b24*log(t2)+b34*log(t3)+b44*log(t4))+g4*log(Y)+g4T*log(dta)))

  v <- n1 + n2 + n3 + n4
  p1 <- n1/v
  p2 <- n2/v
  p3 <- n3/v
  p4 <- n4/v

  d <- (1/dta)*(g1T*log(t1)+g2T*log(t2)+g3T*log(t3)+g4T*log(t4)+gTY*log(Y))

  mask1 <- (n1<=SMALL)>0
  mask2 <- (n2<=SMALL)>0
  mask3 <- (n3<=SMALL)>0
  mask4 <- (n4<=SMALL)>0
  nFail <- (sum(nMask <- (mask1 & mask2 & mask3 & mask4)) > 0)
  dFail <- (sum(mask5 <- (d<=SMALL)>0))
  aFail <- (sum(mask6 <- (a<0)>0))
  if(nFail | dFail | aFail) {
    cat("hitting a forbidden value\n")
    if(nFail) {
      cat(sum(nMask),"individuals had all n's <= 0\n")
      badList <- c(badList,list(c(1,counter,(1:length(nMask))[nMask])))
    }
    if(dFail) {
      cat(sum(mask5),"of d's were <= 0\n")
      badList <- c(badList,list(c(5,counter,(1:length(mask5))[mask5])))
    }
    if(aFail) {
      cat("a was <= 0\n")
    }
    if (sum(bigMask <- (nMask|mask5|mask6)) <= 10) {
      index <- (1:length(bigMask))[bigMask]
      cat("Maybe you should drop rows ",index,"\n")
    }
  }
}
```

```

.value <- BIG
} else {
  xTstar <- v/d
  xT <- x1+x2+x3+x4

# Negative Binomial's Specification
  m <- 1/a
  pStar <- 1/(days*xTstar*a+1)
  logpStar <- log(pStar)

  A0 <- lgamma(x1+1)+lgamma(x2+1)+lgamma(x3+1)+lgamma(x4+1)
  A <- -x1*log(p1) - x2*log(p2) - x3*log(p3) - x4*log(p4)
  B <- lgamma(m) - lgamma(xT+m)
  C <- -xT*log(1-pStar) - m*logpStar
  .value <- A0+A+B+C
}
.value
}

```

ESTIMATION OF VARIANCE-COVARIANCE MATRIX OF PARAMETER ESTIMATES

(this file sourced from Command file as "dLikelihood.s", for the Type 1 Model specification):

```

source("psi.s")
SMALL <- 1e-10
pos <- function(x) {
  x[x<SMALL] <- SMALL
  x }

# estimate is a vector of the parameter estimates,
dLogLikelihoodMU <- function(estimate,myData) {
  N <- dim(myData)[1]
  P <- length(estimate)
  deriv <- data.frame(matrix(0,N,P))
  names(deriv) <- names(estimate)

  attach(myData)
  xT <- x1+x2+x3+x4

  estimate <- data.frame(t(estimate))
  attach(estimate)
  n1 <- pos((-1/t1)*(a1+(b11*log(t1)+b12*log(t2)+b13*log(t3)+b14*log(t4))+g1*log(Y)+g1T*log(dta)))
  n2 <- pos((-1/t2)*(a2+(b12*log(t1)+b22*log(t2)+b23*log(t3)+b24*log(t4))+g2*log(Y)+g2T*log(dta)))
  n3 <- pos((-1/t3)*(a3+(b13*log(t1)+b23*log(t2)+b33*log(t3)+b34*log(t4))+g3*log(Y)+g3T*log(dta)))
  n4 <- pos((-1/t4)*(a4+(b14*log(t1)+b24*log(t2)+b34*log(t3)+b44*log(t4))+g4*log(Y)+g4T*log(dta)))

  d <- (1/dta)*(g1T*log(t1)+g2T*log(t2)+g3T*log(t3)+g4T*log(t4)+gTY*log(Y))

  x1Star <- n1/d
  x2Star <- n2/d
  x3Star <- n3/d
  x4Star <- n4/d
  xTstar <- x1Star + x2Star + x3Star + x4Star

```

```

pStar <- 1/(days*a*xTstar+1)
m <- 1/a

temp <- a*days*(xT/(1-pStar) - m/pStar)/(days*a*xTstar+1)^2 - xT/xTstar
M1 <- (x1/x1Star + temp)
M2 <- (x2/x2Star + temp)
M3 <- (x3/x3Star + temp)
M4 <- (x4/x4Star + temp)

termA <- (psi(m) - psi(xT+m) - log(pStar))/a^2
termB <- (days*xT/(1-pStar)-m/pStar)*xTstar*days*pStar^2
deriv["a"] <- (termA + termB)
deriv["a1"] <- -(M1/(t1*d))
deriv["a2"] <- -(M2/(t2*d))
deriv["a3"] <- -(M3/(t3*d))
deriv["a4"] <- -(M4/(t4*d))
deriv["b11"] <- -(M1*log(t1)/(t1*d))
deriv["b12"] <- -(M1*log(t2)/(t1*d))-(M2*log(t1)/(t2*d))
deriv["b13"] <- -(M1*log(t3)/(t1*d))-(M3*log(t1)/(t3*d))
deriv["b14"] <- -(M1*log(t4)/(t1*d))-(M4*log(t1)/(t4*d))
deriv["b22"] <- -(M2*log(t2)/(t2*d))
deriv["b23"] <- -(M2*log(t3)/(t2*d))-(M3*log(t2)/(t3*d))
deriv["b24"] <- -(M2*log(t4)/(t2*d))-(M4*log(t2)/(t4*d))
deriv["b33"] <- -(M3*log(t3)/(t3*d))
deriv["b34"] <- -(M3*log(t4)/(t3*d))-(M4*log(t3)/(t4*d))
deriv["b44"] <- -(M4*log(t4)/(t4*d))
deriv["g1"] <- -(M1*log(Y)/(t1*d))
deriv["g2"] <- -(M2*log(Y)/(t2*d))
deriv["g3"] <- -(M3*log(Y)/(t3*d))
deriv["g4"] <- -(M4*log(Y)/(t4*d))

termA <- (M1*x1Star + M2*x2Star + M3*x3Star + M4*x4Star)/(d*dta)
termB <- log(dta)/d
deriv["g1T"] <- -(termA*log(t1) + M1*termB/t1)
deriv["g2T"] <- -(termA*log(t2) + M2*termB/t2)
deriv["g3T"] <- -(termA*log(t3) + M3*termB/t3)
deriv["g4T"] <- -(termA*log(t4) + M4*termB/t4)

detach("estimate")
detach("myData")
deriv <- as.matrix(deriv)
scale <- apply(deriv,2,sum)
deriv <- deriv / matrix(rep(scale,N),N,P,byrow=T)
vInv <- (t(deriv) %*% deriv)
diag(1/scale) %*% solve(vInv) %*% diag(1/scale)
}

```

A-6: Example Algorithm for Travel-Time Cost Calculations, as programmed in Matlab

```

ms = numzones;           %%% Numer of Zones (TAZs) in Region
empl = alljobs;         %%% Vector of Opportunities per Zone
time = freeflowtimematrix; %%% Interzonal Travel Times

totTime = zeros(ms,1);  %%% Total Time Vector from Origin Zones to Furthest
                        %%% Contour in Iso-Opportunity Contour
totEmp = zeros(ms,1);   %%% Total Employment Vector
avgTime = zeros(ms,1);  %%% Average Time to Access Contour Vector

maxTotEmp=200000;

for i=1:ms
    while totEmp(i) < maxTotEmp
while totEmp(i) < maxTotEmp
    mnIndx = min(find(dist(i,:)==min(time(i,:)))));
    if (totEmp(i)+empl(i,mnIndx)) < maxTotEmp
        totTime(i) = time(i,mnIndx);
        totEmp(i) = totEmp(i) + empl(i,mnIndx);
        avgTime(i) = avgTime(i) + empl(i,mnIndx).*dist(i,mnIndx);
        empl(i,mnIndx) = 0;
        time(i,mnIndx)=inf;
    else
        pctZone = (maxTotEmp - totEmp(i))/empl(i,mnIndx);
        totTime(i) = totTime(i) + pctZone*(time(i,mnIndx)-totTime(i));
        avgTime(i) = avgTime(i) + pctZone.*empl(i,mnIndx).*time(i,mnIndx);
        empl(i,mnIndx) = empl(i,mnIndx)+totEmp(i)-maxTotEmp;
        totEmp(i)=maxTotEmp;
    end
end
end
avgDst = avgDst./maxTotEmp;

```

A-7: Estimating the Variability in the Results

The results of most interest to researchers and policy-makers are the estimates of elasticities, value of discretionary time, benefits and costs, and other transformations of the underlying model's parameter estimates. Unfortunately, the variation of the output of a non-linear function of variables is generally very difficult to compute exactly, so the results provided in this paper are based on a Taylor Series approximation technique called "Propagation of Error" or the "Delta Method" (Rice 1995) shown here:

Since $Y = f(\theta) = f(\mu_\theta + \delta)$, where $\theta = \text{Random Variable}$,
then...

$$Y \approx f(\mu_\theta) + (\theta - \mu_\theta)f'(\mu_\theta) + \frac{1}{2}(\theta - \mu_\theta)^2 f''(\mu_\theta) + \dots$$

$$\text{So } E(Y) = E(f(\theta)) \approx f(\mu_\theta) + \frac{1}{2}\sigma_\theta^2 f''(\mu_\theta)$$

and $V(\theta) \approx \sigma_\theta^2 [f'(\mu_\theta)]^2$ when θ is scalar.

$$\text{When } \theta \text{ is a vector, } E(Y) \approx f(\bar{\mu}_\theta) + \frac{1}{2} \sum_{i,j} \sigma_{i,j} f_{i,j}(\bar{\mu}_\theta),$$

$$\& V(Y) \approx \sum_{i,j} \sigma_{i,j} f_i(\bar{\mu}_\theta) f_j(\bar{\mu}_\theta),$$

$$\text{where } \sigma_{i,j} = \text{Cov}(\theta_i, \theta_j), f_i = \frac{df}{d\theta_i}, \& f_{i,j} = \frac{df^2}{d\theta_i d\theta_j}.$$

In the research undertaken here, the θ_i random variables represent various parameter estimates of the models (*i.e.*, the α 's, β 's, γ 's, and μ 's). The estimates of means provided in the research are based on a first-order Taylor series expansion around the mean (rather than the more complex, second-order formula shown above), but the estimates of variances as are shown above.

As an example of the use of the Delta Method, the expectation and variance of the ratio of two variables can be estimated using the following:

$$\text{If } Y = f(\bar{X}) = \frac{X_1}{X_2}, \text{ then...}$$

$$E(Y) \approx \frac{\hat{\mu}_{X_1}}{\hat{\mu}_{X_2}} - \frac{\hat{\sigma}_{1,2}}{\hat{\mu}_{X_2}^2} + \frac{\hat{\sigma}_2^2 \hat{\mu}_{X_1}}{\hat{\mu}_{X_2}^3} \text{ (second - order approximation)}$$

$$\approx \frac{\hat{\mu}_{X_1}}{\hat{\mu}_{X_2}} \text{ (first - order approximation), \&}$$

$$V(Y) \approx \frac{1}{\hat{\mu}_{X_2}^2} \left(\hat{\sigma}_{X_1}^2 - \frac{2\hat{\mu}_{X_1} \hat{\sigma}_{1,2}}{\hat{\mu}_{X_2}} + \frac{\hat{\mu}_{X_1}^2 \hat{\sigma}_{X_2}^2}{\hat{\mu}_{X_2}^2} \right).$$

Note that the variance and expectation of *linear* combinations of variables can be estimated with more exact formulae, where an approximation arises only because one is relying on one's estimates of parameters and estimates of their asymptotic variances, rather than their true values.

And the covariance of two linear combinations of variables can be computed using the variance-covariance matrix of their combined vector. The formulae for all these relations are the following:

$$\begin{aligned}
 E(\bar{\beta}' \bar{X}) &= \bar{\beta}' \bar{\mu}_{\bar{X}} \approx \bar{\beta}' \hat{\mu}_{\bar{X}}, & \\
 V(\bar{\beta}' \bar{X}) &= \bar{\beta}' \Sigma_{\bar{X}} \bar{\beta} \approx \bar{\beta}' \hat{\Sigma}_{\bar{X}} \bar{\beta}, \\
 &\& Cov(\bar{\beta}'_1 \bar{X}, \bar{\beta}'_2 \bar{X}) = \bar{\beta}'_1 \Sigma_{\bar{X}} \bar{\beta}_2 \approx \bar{\beta}'_1 \hat{\Sigma}_{\bar{X}} \bar{\beta}_2, \\
 &where \hat{\Sigma}_{\bar{X}} = Estimate\ of\ VarCov(\bar{X}).
 \end{aligned}$$

Index

—C—

concavity, 8, 29, **31**, 32, 39

—D—

demand elasticities, 2, 7, 35, 43, 48, 77, 82, 87, 97,
101, 108, 109, 116, 119, 135, 142, 156, 158
discrete-choice models, 1, **10**, 11, 12, 13, 15, 16, 136
discretionary activities, 19, 23, 24, 28, 34, 76, 84, 88,
95, 115, 117, 119, 120, 126, 133

—E—

expenditure function, 4, 31, 32, 37, 38, **40**, 41, 59,
116, 124, 125, 137

—G—

gamma heterogeneity, 51, **54**, 55, 56, 57, 58, 59, 89,
94, 110, 111, 138, 140, **166**

—H—

homogeneity, 8, 29, 32, **33**, 34, 45, 69, 137, 156, 159

—I—

identifiability, 25, 29, 36, 103, 137

—N—

negative binomial, **51**, 54, 56, 57, 60, 94, 111, 134,
140, **166**

—O—

opportunity contours, **61**, 64, 65, 66, 76, 84, 107, 117,
124, 127
overdispersion, 51, 57, 80, 84, 89, 94, 140

—P—

Poisson distribution, **49**, 50, 51, 52, 53, 54, 55, 59, 60,
74, 94, 110, 166

—R—

Roy's Identity, 21, **22**, 40, 46, 49, 58, 103, 133, 136,
161, 162

—S—

separability, **35**, 36, 37
symmetry, 8, 29, **37**, 38, 39, 156
systems of demand equations, 2, **21**, 33, 39, 42, 45,
47, 52, 55, 57, 66, 90, 103, 132, 133, 134, 135,
156

—T—

translog functional form, **42**, 43, 44, 46, 47, 48, 79,
84, 88, 115, 123, 140, **159**
trip chaining, **66**, 67, 82, 91
trip generation, 1, 6, 16, 17, 72, 108, 115, 120, 121

—V—

value of time, 14, 15, 23, 26, 27, 46, 49, 62, 77, 87,
91, 100, 123

—W—

welfare analysis, 1, 2, 3, 18, 32, 33, **40**, 41, 42, 53, 59,
71, 115, 124, 126, 127, 129, 137