# A 24–HOUR HOUSEHOLD–LEVEL ACTIVITY–BASED TRAVEL DEMAND MODEL FOR THE GTA

by

Len K. Eberhard

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Civil Engineering
University of Toronto

**A 24–Hour Household–Level Activity–Based**

**Travel Demand Model for the GTA**

Master of Applied Science

Len K. Eberhard

Graduate Department of Civil Engineering

University of Toronto

2002

The art and science of travel behaviour forecasting is in transition. Having relied on the traditional four–stage urban transportation modelling system for decades, modellers and planners are beginning to consider new methods for predicting travel behaviour. This has been fostered by the realisation that travel is a demand derived from the need to participate in activities that occur in varying space and time, rather than simply making trips for the sake of travelling. Activity–based modelling represents the next generation of travel demand forecasting techniques. This work documents the initial theory, development and testing of a weekday twenty–four hour, person/household–level, activity–based travel behaviour/demand model developed for the Greater Toronto Area (GTA). The model is applied to a 1996 base year population and transportation network. Aggregate trip estimations compare well with travel survey data and disaggregate simulated activity schedules are realistic.

## ACKNOWLEDGEMENTS

I would like to thank the other members of the study team involved in the development of this model, namely Eric Miller, Matthew Roorda, Jeffrey Guan and Marek Litwin. I would also like to thank Murtaza Haider for his assistance with the statistical analysis as well as for providing some additional data.

# TABLE OF CONTENTS

Acknowledgements ..... iii
Table of Contents ..... iv
List of Tables ..... vi
List of Figures ..... vii
1 Introduction ..... 1
2 Literature Review ..... 3
3 Data Preparation & Descriptive Analysis ..... 9
3.1 Data Preparation ..... 9
3.2 Preliminary Investigation ..... 10
3.3 Duration, Start Time & Frequency ..... 12
3.4 Joint Duration, Start Time & Frequency ..... 26
3.5 Mode Choice & Trip Chaining ..... 29
3.6 Location Choice & Trip Chaining ..... 34
3.7 Serving Dependants ..... 38
4 Model Development ..... 44
4.1 Generation of Activity Episodes ..... 48
4.2 Location Choice ..... 54
4.3 Travel Episodes, Mode Choice & Trip Chains ..... 57
4.4 Modelling Process ..... 65
5 Model Results & Validation ..... 66
5.1 Simulated Activity Behaviour ..... 66
5.2 Simulated Travel Behaviour ..... 68
6 Conclusion ..... 73
References ..... 77

Appendix A: Glossary of Terms ..... 81
Appendix B: Data Preparation & Record Deletion ..... 83
Appendix C: Code Used for Data Preparation ..... 88
Appendix D: Duration, Start Time & Frequency T–Test Results ..... 102

Appendix E: Definition of Probability Distribution Categories ................................................ 107
Appendix F: Thematic Maps of Activity Levels Within the GTA ............................................ 113
Appendix G: Location Choice Model Parameter Estimation .................................................... 117
Appendix H: Mode Choice Model Parameter Estimation ........................................................ 124

# List of Tables

Table 3–1 Comparison of Independent & Joint Statistics ........................................................... 27

Table 4–1 Summary of Variables Used for Episode Attribute Distribution Categorisation ........ 53

Table 4–2 Comparison of Predicted and Observed (Input) Total Trips and Mode Split.............. 64

Table 5–1 Comparison of Observed (TTS) and Predicted Mode Split (%) by Time Period........ 70

Table 5–2 Comparison of Observed and Predicted Number of Trips (1000's) by Mode and Trip Purpose ………………………………………………………………………………… 71

Table 5–3 Comparison of Total Number of Trips by Mode and Mode Split (%) ........................ 72

# List of Figures


Figure 3–1 Number of Trips by Destination Purpose & Start Time ........ 11
Figure 3–2 Percentage of Trips by Destination Purpose & Start Time ........ 11
Figure 3–3 Percentage of Trips by Destination Purpose & Start Time (Ages 0 to 15) ........ 13
Figure 3–4 Percentage of Trips by Destination Purpose & Start Time (Ages 26 to 40) ........ 13
Figure 3–5 Percentage of Trips by Destination Purpose & Start Time (Professional/Management/Technical) ........ 14
Figure 3–6 Percentage of Trips by Purpose & Start Time (Retail Sales & Service) ........ 14
Figure 3–7 Percentage of Trips by Destination Purpose & Start Time (Management/Construction/Trades) ........ 15
Figure 3–8 Percentage of Trips by Destination Purpose & Start Time (General Office/Clerical) 15
Figure 3–9 Percentage of Trips by Destination Purpose & Start Time (Unemployed) ........ 16
Figure 3–10 Percentage of Trips by Destination Purpose & Start Time (Auto–Passenger) ........ 16
Figure 3–11 Percentage of Trips by Destination Purpose & Start Time (Public Transit) ........ 17
Figure 3–12 Percentage of Trips by Destination Purpose & Start Time (Auto–Driver) ........ 17
Figure 3–13 Percentage of Trips by Destination Purpose & Start Time (No Driver's Licence).. 18
Figure 3–14 Percentage of Trips by Destination Purpose & Start Time (Driver's Licence) ........ 18
Figure 3–15 Histogram of *Work* Activity Duration ........ 19
Figure 3–16 Histogram of *School* Activity Duration ........ 19
Figure 3–17 Histogram of *Market* Episode Duration ........ 20
Figure 3–18 Histogram of *Other* Episode Duration ........ 20
Figure 3–19 PDF of *Work* Activity Duration by Occupation ........ 22
Figure 3–20 PDF of *Work* Activity Duration by Employment Status ........ 22
Figure 3–21 PDF of *School* Activity Duration by Student Status ........ 23
Figure 3–22 PDF of *Market* Episode Duration by Driver's Licence Possession ........ 23
Figure 3–23 PDF of *Work* Activity Start Time by Occupation ........ 24
Figure 3–24 PDF of *Market* Episode Start Time by Employment Status ........ 24
Figure 3–25 Percentage of *Work* Episode Frequency by Gender ........ 25
Figure 3–26 Percentage of *Market* Episode Frequency by Licence Possession ........ 25
Figure 3–27 *Market* Time–Share (Household with No Children) ........ 28
Figure 3–28 *Market* Time–Share (Household with Children) ........ 28

Figure 3–29 *Other* Time–Share (Household with 1 Vehicle)...................................................... 30
Figure 3–30 *Other* Time–Share (Household with 2+ Vehicles)................................................. 30
Figure 3–31 Percentage of First Mode Used by Number of Modes in Trip Chain ...................... 32
Figure 3–32 Percentage of Number of Modes by First Mode Used in *Work* Trip Chain............. 32
Figure 3–33 Percentage of Number of Modes by First Mode Used in *School* Trip Chain........... 33
Figure 3–34 Percentage of Number of Modes by First Mode Used in Neither *Work* or *School* Trip Chain ............................................................................................................... 33
Figure 3–35 Percentage of First Mode by Second Mode for Two Mode Chains ......................... 34
Figure 3–36 Percentage of Activity Type by Mode Used for Trip (No Driver's Licence)........... 35
Figure 3–37 Percentage of Activity Type by Mode Used for Trip (Driver's Licence, No Vehicle) ................................................................................................................................ 35
Figure 3–38 Percentage of Activity Type by Mode Used for Trip (Driver's Licence, Vehicle Not Available)............................................................................................................... 36
Figure 3–39 Percentage of Activity Type by Mode Used for Trip (Driver's Licence, Vehicle Available)............................................................................................................... 36
Figure 3–40 Percentage of Trips per Chain by Mode (Single Mode Chains Only) ..................... 37
Figure 3–41 PDF of Trip Distance by Mode (*Market* Activities) ............................................... 40
Figure 3–42 PDF of Trip Distance by Mode (*Other* Activities)................................................. 40
Figure 3–43 PDF of Activity Location Distance from Home by Mode (*Market* Activities)....... 41
Figure 3–44 PDF of Activity Location Distance from Home by Mode (*Other* Activities).......... 41
Figure 3–45 PDF of Activity Location Distance from Home by Activity Type .......................... 42
Figure 3–46 PDF of Activity Location Distance from Work by Activity Type........................... 42
Figure 3–47 Number of School-based Facilitate Passenger Trips by Trip Start Time & Mode for Households With Children Aged 4 to 10 ............................................................. 43
Figure 3–48 Number of School-based Daycare Trips by Trip Start Time & Mode..................... 43
Figure 4–1 Travel Behaviour Decision Triangle ......................................................................... 47
Figure 4–2 Household Activity Structure.................................................................................... 50
Figure 4–3 Comparison of *Work* Episode & Activity Duration PDFs. ....................................... 51
Figure 4–4 Comparison of Observed and Predicted Trip Distances (*Other* Activity Locations). 58
Figure 4–5 Comparison of Observed and Predicted Trip Distances (*Work–Business,* No Usual Place of Work Activity Locations) ......................................................................... 58

Figure 4–6 Chain-Level Mode Choice Decision Structure ..... 60
Figure 4–7 Observed and Predicted Mode Split by Time Period ..... 64
Figure 5–1 Default Simulated Person Activity Schedule ..... 67
Figure 5–2 Simple Simulated Person Activity Schedule ..... 67
Figure 5–4 Complex Simulated Person Activity Schedule ..... 68
Figure 5–5 Example of Complex Simulated Trip Pattern ..... 69

# 1 INTRODUCTION

The art and science of travel behaviour forecasting is in transition. Having relied on the traditional four–stage urban transportation modelling system for decades, modellers and planners are beginning to consider new methods for predicting travel behaviour. This has been fostered by the realisation that travel is a demand derived from the need to participate in activities that occur in varying space and time, rather than simply making trips for the sake of travelling. The traditional methods function (as well as can be expected) with aggregate data, predicting traffic flows from zone to zone. Activity–based methods, as they are commonly referred to, however, only make sense if they are applied at a disaggregate level, such as the person or household (i.e. a zone does not exhibit a singular travel/activity behaviour). Fortunately, this transition is occurring at a time when computing power is becoming sufficiently large to be able to handle the greatly increased demands that disaggregate, or microsimulation, methods require relative to traditional, aggregate methods.

The majority of work in activity–based, microsimulation research has focussed on the individual as the decision–making unit (Scott, 2001). Individuals rarely live alone, however, but rather are usually part of a household (the composition of which varies, but is typically comprised of one male head, one female head, and one or more children). It seems unlikely, then, that in the household decision–making process trips are planned independently or that an individual's activity participation and travel behaviour decisions are made in isolation. Rather, the needs of the household are probably translated into a number of activity stops to be made during the day, which are then organised as best as possible (Bhat and Koppelman, 1999). Some of these needs/stops will be fulfilled by individuals and some will be fulfilled jointly by members of the household. This organisation is likely based on factors such as the number of vehicles available to the household, the locations of all the activities in relation to each other, the time of day at which the activities must occur, with whom, and the frequency (i.e. daily, weekly, monthly, *et cetera*). This organisation must also incorporate a travel mode choice (e.g. automobile, public transit, walking) for all given household trip linkages, or trip chains. In order to understand how this organisation occurs, one must understand the intra–household interactions and decision–making processes that take place within the family. Bhat and Koppelman (1999) acknowledge that:

> "An area that has received limited attention thus far in the activity analysis literature is the interactions among individuals in a household and the effect of such interactions on individual activity episode patterns. Interactions among individuals might take the form of joint participation in certain activities (such as shopping together or engaging in recreational/social activities together), 'serve–passenger' and 'escort' activities where one individual facilitates and oversees the participation of another in activities (for example, the 'soccer mom' phenomenon), and allocation of autos and activities among individuals (especially in multi–adult, one–car households). Such interactions can lead to constraints that may be very important in individual activity/travel responses to changes in the transportation or land–use environment" (p. 54).

Others have argued that it is the household, as an entity itself, that is the decision–making unit. Thus, it seems that there is some connection, or integration, of the person–level and household–level utility gained from travel and activity participation and that this connection should be considered when developing an activity–based travel demand model. As such, it is the goal of this research to identify, define, and incorporate these interactions into an activity–based travel demand model in the hope of explicitly capturing this household decision structure and improving current travel demand forecasting methods. This work will document the initial theory, development and testing of a weekday twenty–four hour, person/household–level, activity–based travel behaviour/demand model developed for the Greater Toronto Area (GTA). Chapter 2 will review the literature that has been written on the topic to date. Chapter 3 provides a brief discussion of the process that was used to prepare the trip–oriented data for use in an activity–based model. This is followed by a full descriptive analysis of the data from an activity–based perspective. Chapter 4 discusses the development of the model. Chapter 5 presents the results of a base year model run and investigates the validity of these using both standard as well as unconventional techniques. Finally, Chapter 6 concludes with a critique of the model and proposes some possible improvements to its structure.

# 2 LITERATURE REVIEW

Previous research into intra–household interactions has been undertaken by several disciplines, namely those of economics, psychology and transportation engineering. As is evident from the above, however, such research is usually a sub–component of the broader field of activity–based analysis. The result of this research has been a varied array of approaches to the activity problem. There is even variation in the classification of the various approaches; for example, Arentze and Timmermans (2000) divide them into single–facet, constraints–based, utility–maximizing, computational process, and microsimulation approaches, whereas Wen and Koppelman (1999) categorise them as either tour–based, structural equation, or activity scheduling models. In reviewing the literature, it appears that the computational process and, in particular, the utility–maximising approaches are the only ones to consider intra–household interactions to any degree, whether implicitly or explicitly. Nevertheless, the following is a summary of the relevant work on the topic without consideration of classification.

Transportation modelling has often depended on economic theories to accomplish its goals, most notably in its use of random utility choice theory to determine travel mode choice. As microeconomic theory has attempted to predict consumer behaviour at the household level, it is not surprising that there has been substantial work in the field of intra–household interactions and decision–making. Beginning with the work of Becker (1965), the New Home Economics attempted to explain household production and consumption by maximising a single household utility function. This assumed, for one, that an altruistic household head distributed resources equally within the family and conveniently avoided the problem of aggregating the individual utility functions of family members. Given the obvious weaknesses inherent to such an approach, economists turned next to cooperative game theory and employed the solution developed by Nash (1950). In the Nash solution, the product of the individual gains to male and female household heads from marriage is maximised (Ott, 1992):

$$\text{Max } N = [U^m(X^m) - D^m] * [U^f(X^m) - D^f] \qquad [2\text{–}1]$$

such that $(X^m + X^f)(p) = Y$ (budget constraint)

and $U^i \geq D^i$ $(i = m, f)$

| | | |
|---|---|---|
| where: | $U^i$ | = utility level of individual i if an agreement is reached |
| | $X^i$ | = vector of goods of individual i |
| | $D^i$ | = conflict payoff of individual i, or the 'threat point' |
| | p | = price vector |
| | Y | = total household income |

The two terms in the objective function are the 'gains to cooperation' of each household member, and are determined in part by the threat point, $D^i$. This point usually represents the achievable utility level possible for the individual outside of the marriage (i.e. under divorce). Thus, it is also a measure of each individual's 'bargaining power', as the person who has the larger threat point will receive the greater utility in the Nash solution. This approach was first employed by Manser and Brown (1980) as well as McElroy and Horney (1981). It can also be extended to include time constraints, which, of course, would be critical to an activity–based travel model. Recent endeavours to improve these models have focussed on defining a more realistic threat point by using a non–cooperative solution in place of divorce in the event of disagreement (for example, Chen and Woolley, 1999).

The transportation discipline has produced many activity/travel models at the household level. The majority, however, provide no truly explicit mechanism for intra–household interaction. This is true of STARCHILD, for example, developed by Recker *et al.* (1986a, 1986b), despite having considered household member interdependencies:

> "…through the use of several household interactions such as for example activities performed jointly by several household members, the temporal availability of household automobiles, and decision objectives such as the amount of time spent at home with other household members" (Arentze and Timmermans, 2000, p. 26).

SCHEDULER, developed by Gärling *et al.* (1989), "…entails an interrelated set of decisions made by the individual, interactively with other (household) individuals, concerning who will participate in the activities…" (Arentze and Timmermans, 2000, p. 47). Furthermore:

> "The interaction within a household, or a wider social group, of which the individual is a member, is made dependent upon the kind of activity. Some activities are assumed to require the participation of at least two household members, some can be performed by either household member, whereas still others must be performed by the individual. The scheduling of the first two types of activities requires information of the other person's schedule. If constraints are

> such that the joint activity is impossible, the activities are rescheduled or replaced" (Arentze and Timmermans, 2000, p. 49).

The CHASE (Computerised Household Activity SchEduling) software, first developed by Doherty and Miller (2000), is the first survey to capture the week–long scheduling behaviour of the entire household. Presently, it has been used to record the:

> "…household's activity agenda from which all activities are drawn, and to track the *sequence* of steps whereby activities from the agenda are added, deleted, and subsequently modified during their execution to form individual household member's weekly activity schedules" (p. 78).

More important, though, are future plans to meet a third objective of identifying the decision rules that underlie the various individual and joint decisions made during the household activity scheduling process.

Another important activity/travel model is ALBATROSS (Arentze and Timmermans, 2000). While 'travel party' was taken into consideration, the authors admit to a need to "elaborate interactions within the household…" as "coupling constraints have only been partly taken into account" (p. 405). As a result, Arentze and Timmermans (2001) subsequently extended their research by conceptualising a framework for modelling the formation of activity agendas at the household level, integrating both person and joint activities. The framework identifies the activity decisions and arranges them in a hierarchy where 'agents' distinguish the levels. Decisions at one level are used in successive decisions (e.g. the maximum amount of time allotted to leisure activities at one level is subsequently divided into three subcategories of leisure at another level). Outputs of the decision agents would be an agenda specifying (i) time allocation across persons, activities, and days; (ii) the person, start time, duration, and location of institutionalised activities; and (iii) transport mode access. An empirical modelling technique was applied only to illustrate a possible application of the framework, but showed reasonable results.

The previous activity/travel models have attempted to include intra–household interactions as a sub–component of a much larger household activity scheduling framework, so far having done so more implicitly (i.e. capturing them empirically) than explicitly (i.e. in terms of an explicit theoretical mechanism). Given the complexity and importance of the problem, some researchers have focused on it more specifically. Golob and McNally (1997) used a structural equation method to model time allocation to activities and travel by each household head simultaneously, but neglected to include joint activity participation or travel. Goulias and Kim (2001) estimated two logit models for each of both activity and travel patterns, accounting for person and household variation jointly. Both the activity and travel pattern alternatives consist of four broadly defined categories; for example, the travel pattern alternative is composed of non–motorized, car/carpool, public transportation, or immobile. Both studies identified correlation among some patterns at the household level and at the individual level, leading to conclusions about the importance of accounting for this in such models. Despite their importance in showing that an individual's activity participation and travel behaviour decisions are not made in isolation, however, these models still only implicitly account, to some degree or another, for the interactions among the members of a household (for other examples, see Simma and Axhausen (2001), Fujii *et. al.* (1999), Chandraskharan and Goulias (1999), Bhat *et. al.* (1999), Golob (1999), and Recker (1995)). However, some attempts have been more successful at doing this explicitly.

Miller and Rhamey (1987) modelled the mode choice of two–worker households. They tested several nested and joint logit structures against the conventional independent structure, which considers only the individual worker without regard for the impact that the decision might have on the mode choice of other household members. The authors found that a nested structure significantly outperformed an independent one (in two–worker households), and concluded that:

> "…the work trip mode choice decision should be considered as one component within a larger household decision–making framework which includes both longer–run decisions concerning auto holdings (i.e. the number and type(s) of autos owned by the household) and shorter–run decisions concerning daily activity and travel patterns" (p. 1568).

Proposals for future research also called for the "development of improved household interaction variables and decision structures" (p. 1568).

Algers *et al.* (1997) constructed one of the first operational activity–based forecasting models to specifically account for interactions among household members. Nested logit models incorporating automobile and shopping task allocation between family members are developed separately for work, school, business and shopping trips. This was a precursor to the work of Wen and Koppelman (1999), who followed with a two–stage model consisting of a nested logit model of maintenance activity generation and stop/auto allocation to two family members (i.e. a male and female head) as well as a logit model of the number of tours and the assignment of stops to tours for each individual. Suggested improvements to the model include the use of more realistic estimation techniques (e.g. the paired combinatorial or the cross–nested logit models), as well as the inclusion of joint activity episodes, multi–stop tour pattern alternatives, stop location, tour travel mode and time of day choices.

Scott (2001) derived a joint probit model of the number of daily out–of–home male (or worker), female (or non–worker), and joint activity episodes. The model, called the Trivariate Ordered Probit Model, accounts for household interactions by correlating the three random error terms of the model. Correlated models are statistically compared to non–correlated ones to test the hypothesis that these decisions are made jointly. Results from the empirical analysis of three household types (couple non–worker, one–worker, and two–worker) support this hypothesis and show that the correlated models also predict more accurately.

Gliebe and Koppelman (2001) constructed a sequential utility–based model of joint activity participation among two adult household heads. The two–level model is based on the maximisation of the total household utility as defined by the collective utility of its individual members. The choice set consists of six 'time–share' alternatives representing both individual and joint activity categories. Four 'independent' alternatives comprise the lower level, and are represented along with two joint alternatives in the upper level. Joint activities are hypothesised to be motivated by considerations of efficiency, altruism, and/or companionship (after Townsend, 1987), which is the underlying principle of the functional form. The model is constrained by the available time in a day. Empirical results from parameter estimation reinforce the common assumption that non–work activities (i.e. leisure and maintenance) are scheduled around work hours (i.e. subsistence). Other important findings include task specialisation (i.e. a movement from joint to individual activity participation) in households with young children as well as gender–based role differences. Automobile ownership levels had little impact on activity time–share. This model is important in that it is one of the first where "…the utility of multiple decision makers are represented in both an individual and a collective sense for the purpose of explaining joint activity participation and travel" (p. 3).

Given the complexity of the activity scheduling problem, especially in considering the dynamics involved in the process (i.e. altering one's schedule over time), there is a need for some fundamental concepts to provide guidance in the formulation and development of activity scheduling processes and frameworks. One of these concepts, proposed by Axhausen (1998), is the notion of the 'project'. Axhausen defines a project as "a set of acts/activities which belong together to produce a specific object or to stage a specific event" (p. 310). That is, all activities form part of some greater objective or goal, rather than each episode in a person's life occurring as a separate entity unto itself. For example, hosting a dinner party would likely involve several different, yet related, activities, such as calling the guests, purchasing the food, preparing the meal, and cleaning up, to name a few. Again, some of these activities will be performed by individuals and some will be performed jointly, requiring coordination with other members involved in the project. Ultimately, these projects evolve from fundamental human needs, such as sustenance, health and shelter. Each of these, in turn, are composed of several levels of sub–projects; for example, the dinner party sub–project might evolve from the need for companionship and require the execution of the food shopping sub–sub–project. The scheduling of activities would then occur in a decision hierarchy, whereby a person's agenda is determined by 'schedulers' ranging from the sub–project level to the project level to the person level and finally the household level. Upper level schedulers would mediate conflicts between the lower ones, passing them temporal (and possibly monetary) constraints to accommodate. Such an organising principle would be quite powerful, as it "decomposes and encapsulates the extraordinarily complex activity scheduling problem into 'bite size' pieces using the same decision machinery over and over" (Miller, 2001a).

# 3 DATA PREPARATION & DESCRIPTIVE ANALYSIS

The following is a summary of the process used to prepare the data used in the model as well as the descriptive analysis used to guide its development. Note that activity–based related terms used hereinafter (initially in bold print) are defined in Appendix A. Their usage may or may not be original or consistent with that of other activity–based literature and are defined in terms of how they are utilised within this model.

## 3.1 DATA PREPARATION

The data used in developing the model are obtained from the 1996 *Transportation Tomorrow Survey* (TTS). This survey consists of the one–day travel behaviour of approximately 5% of households in the GTA. This includes all trips made by all members of the household that are 11 years of age and above, as well as their relevant household and personal socio–economic attributes (for use as explanatory variables). The challenge in using this data set to develop an activity–based model is that it was initially collected for use in conventional trip–based models. As such, the data are not targeted toward describing activity behaviour. It is possible, however, to elicit such information to a certain degree from the data, at least as is required for our purposes.

In extrapolating this trip–based data set into an activity–based one, it is necessary to look at trip records grouped by person and by household. Given the available information collected for each trip, combined with estimated travel time information for the GTA, it is possible to determine the **duration**, **start time** and **frequency** of all **activity episodes**, as well as **joint** activity participation and trip chain information. **Activity** types are defined based on the trip destination purpose of the original trip–based data[1]. These include *home, work*, *school*, *market*, *other* (consisting mainly of personal business and entertainment/leisure activities), *daycare* and *facilitate passenger*. Activity episode duration for a given trip/activity record is calculated by finding the difference between the start time of the trip and the subsequent trip and then

[1] In making the transition from a trip–based to an activity–based framework, terms such as 'trip purpose' become synonymous with 'activity type'.

subtracting the estimated travel time of the trip.[2] Trip chains are assumed to be **home–based chains**. A home–based chain begins upon leaving home and ends upon returning home and includes all trips and activities completed in between. This might consist of, for example, going from home to work to lunch to a store to work and back to home. Joint activity participation is defined as two or more people from the same household travelling to and participating in the same activity together. This is determined by comparing and matching the trip start time, trip purpose, travel mode and trip destination (i.e. activity location) of each trip made by a person with those of others in the household.[3] This is a simplification of the phenomenon given the many variations of joint participation that are possible (e.g. a joint activity but not a joint trip, a joint activity with overlapping but unequal duration and/or start time, *et cetera*), but is considered acceptable for an initial model.

After completion of the data preparation and the deletion of unwanted records, the data set consists of 434,583 trip records, representing the travel behaviour of 148,299 persons or 73,034 households (for further details of the data preparation process and record deletion criteria, see Appendices B and C).

## 3.2 PRELIMINARY INVESTIGATION

To gain a better understanding of the data issues that would affect model development, a variety of descriptive analyses were performed to determine which factors (or, more specifically, explanatory variables) are significant[4]. First, the trip records themselves were used to investigate trip purpose (i.e. activity type) and trip start time as functions of various third variables. Figures 3–1 and 3–2 show the absolute number and the proportion of trips made for each purpose by start time, respectively, as a base reference. Both reveal expected results (for a typical weekday) as the morning hours depict *work* and *school* activities, moving towards home activities as the day

[2] Estimated travel times are from origin traffic zone centroid to destination traffic zone centroid and were obtained from previous trip–based travel models for the GTA.
[3] This is done only for certain trip purposes as, for example, the activities of each person are different when one is facilitating another, meaning that the activity is not joint. This is explained in further detail later.
[4] This first analysis was actually performed on the entire TTS data set whereas the remainder are based on the prepared one discussed previously. Also, it included use of the 1986 data set to separate out personal business and entertainment activities from the other activity type (for this analysis only).

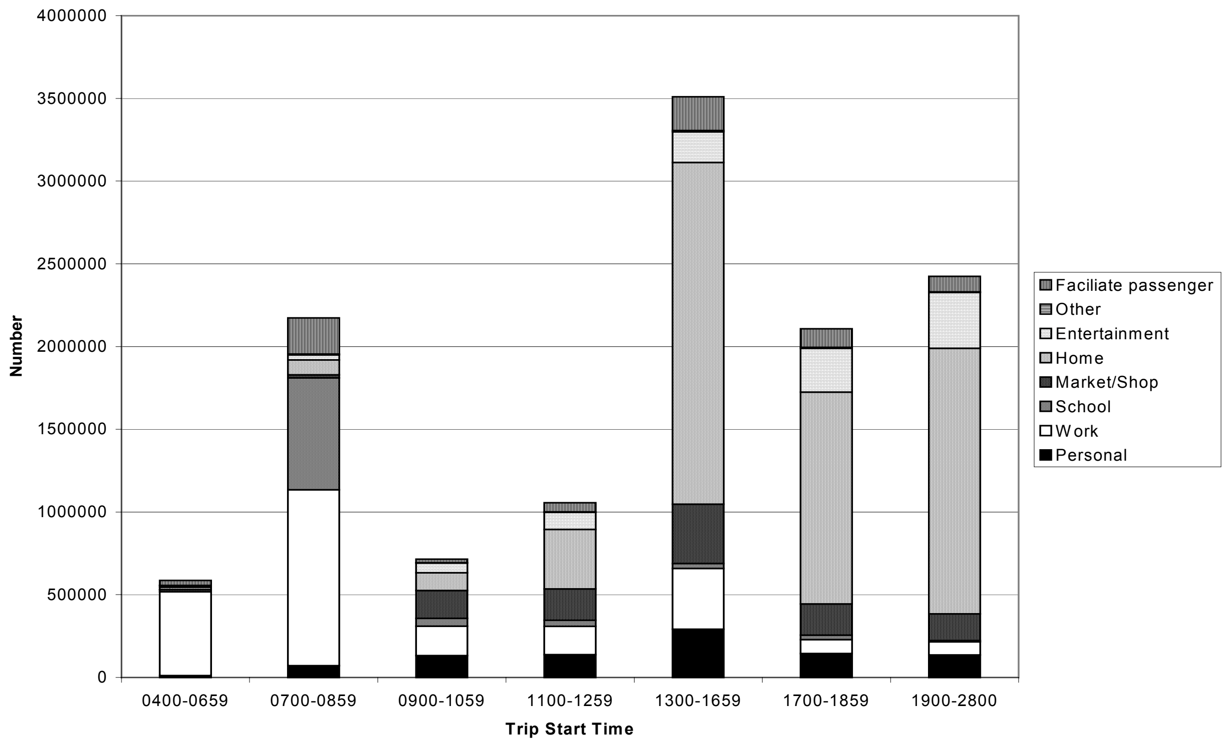



**Figure 3–1 Number of Trips by Destination Purpose & Start Time**

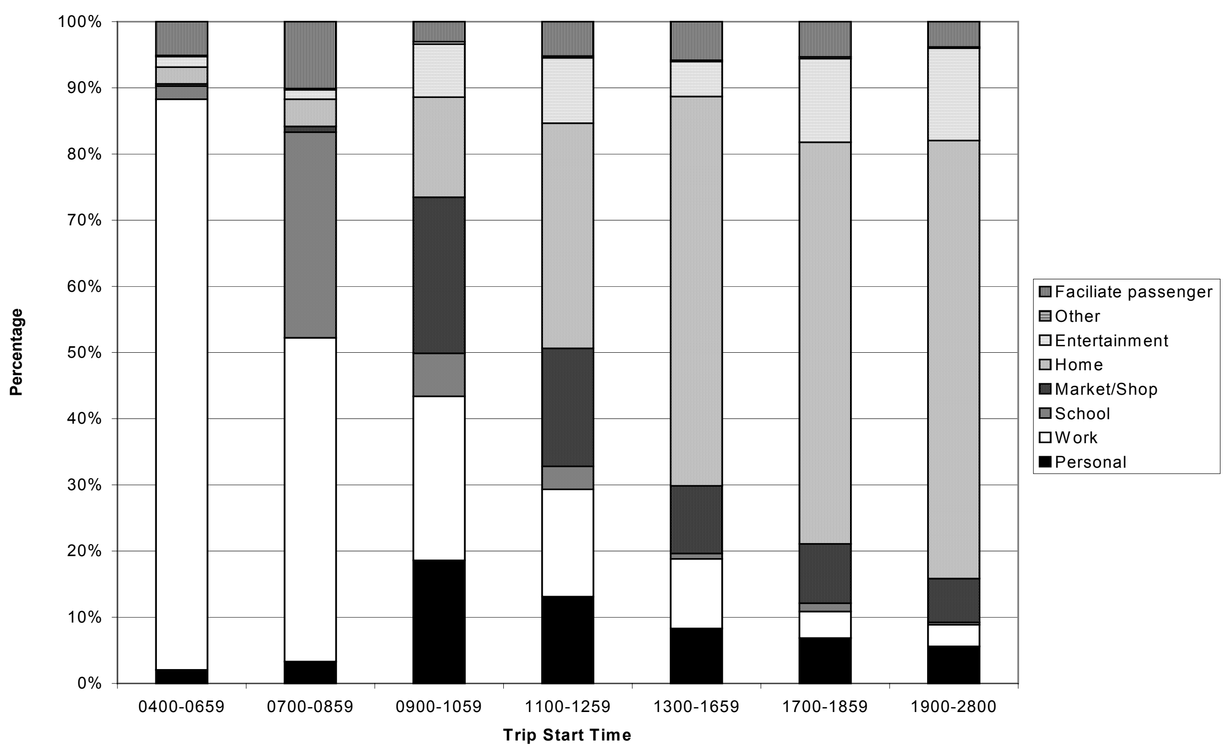



**Figure 3–2 Percentage of Trips by Destination Purpose & Start Time**

progresses. The *market* and *personal business* activities peak at midday, those for *entertainment* in the evening, with *facilitate passenger* fairly constant all day except for a peak in the morning rush hour. Figures 3–3 and 3–4 show the obvious progression from the *school* to *work* activity as age increases through three categories from child to adult. In relation to Figures 3–5, 3–6 and 3–7, one can see in Figure 3–8 that the retail sales and service sector has a larger percentage of young employees who attend school during the day, as is obviously also true for the unemployed sector (Figure 3–9). Finally, the effect of travel mode and driver's licence possession on *work* trips is illustrated in Figures 3–10 through 3–14. That is, younger people going to school, who do not possess a licence, are generally driven or take public transit, followed by the customary shift to the auto–driver mode as people begin to work. As would be expected, there is a high degree of correlation between the various socio–economic variables (such as age, employment status, occupation, *et cetera*) used in these analyses of activity and start time.

## 3.3 DURATION, START TIME & FREQUENCY

The second descriptive analysis examines the distributions of the calculated activity duration, start times and daily frequencies as functions of similar socio–economic variables. Figures 3–15 through 3–18 show the distributions of activity duration for *work*, *school*, *market* and *other*, respectively (as histograms) using 5–minute time intervals. Durations for *work* and *school* are given as total daily activity duration (i.e. the sum of all activity episodes in the day), whereas the others are given as episode duration.[5] The results are very reasonable, as work activity duration averages about 8.5 hours and market episode duration averages 1.3 hours. Figures 3–19 through 3–26 illustrate how activity duration, start time and frequency are affected by selected socio–economic variables (as probability density functions). Occupation and employment status both affect *work* activity duration. The variance in duration for the retail sales and service sector is much higher as compared to that of the general office/clerical sector. The distribution for full–time workers is much more efficient than for either part–time or at–home workers[6], as these workers generally have more flexibility in this activity. Student status has a similar effect on *school* activity duration except that the part–time student distribution is less varied. The

[5] The reasons for this will be discussed in the next chapter.
[6] Note that episodes for at–home workers are for out–of–home episodes, presumably for business meetings.

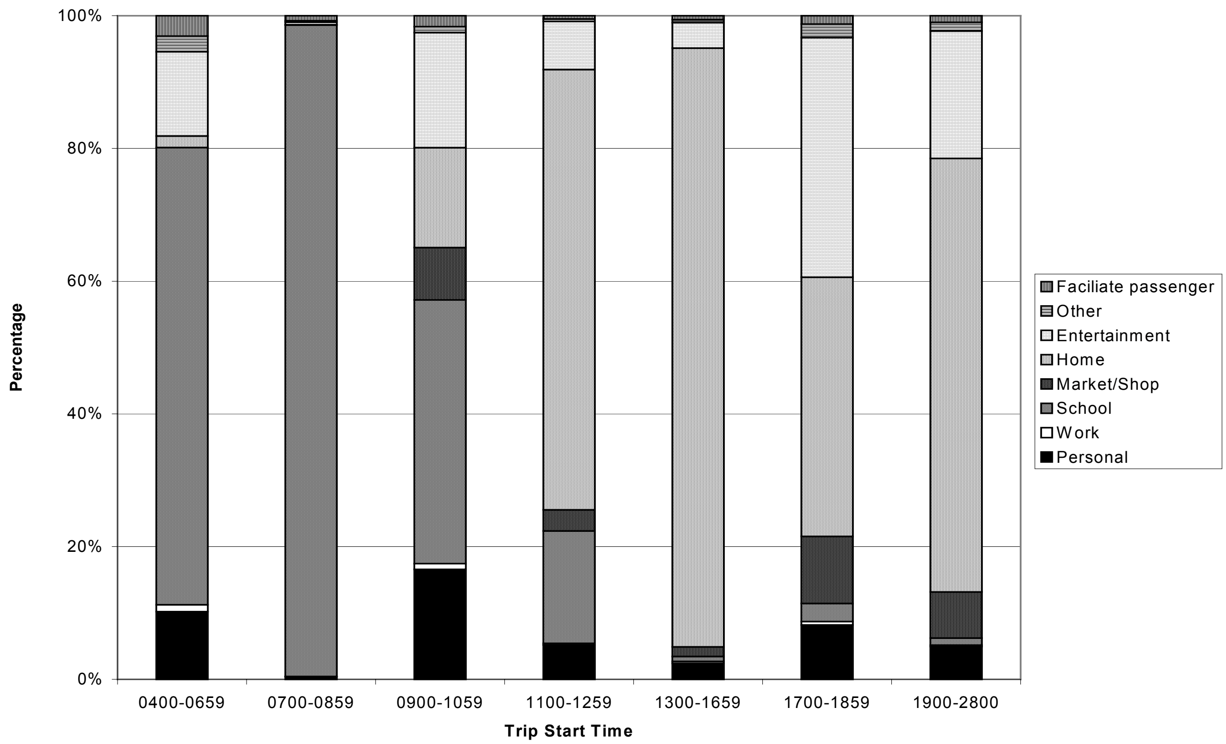



**Figure 3–3 Percentage of Trips by Destination Purpose & Start Time (Ages 0 to 15)**

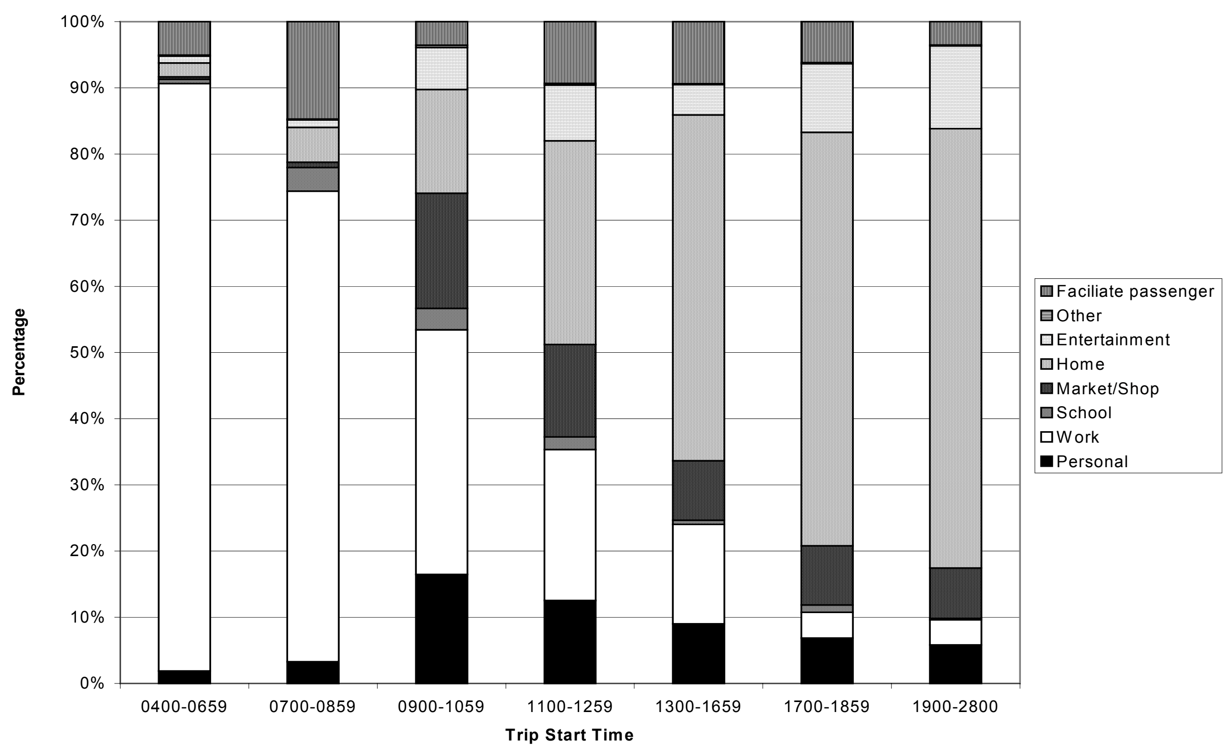



**Figure 3–4 Percentage of Trips by Destination Purpose & Start Time (Ages 26 to 40)**

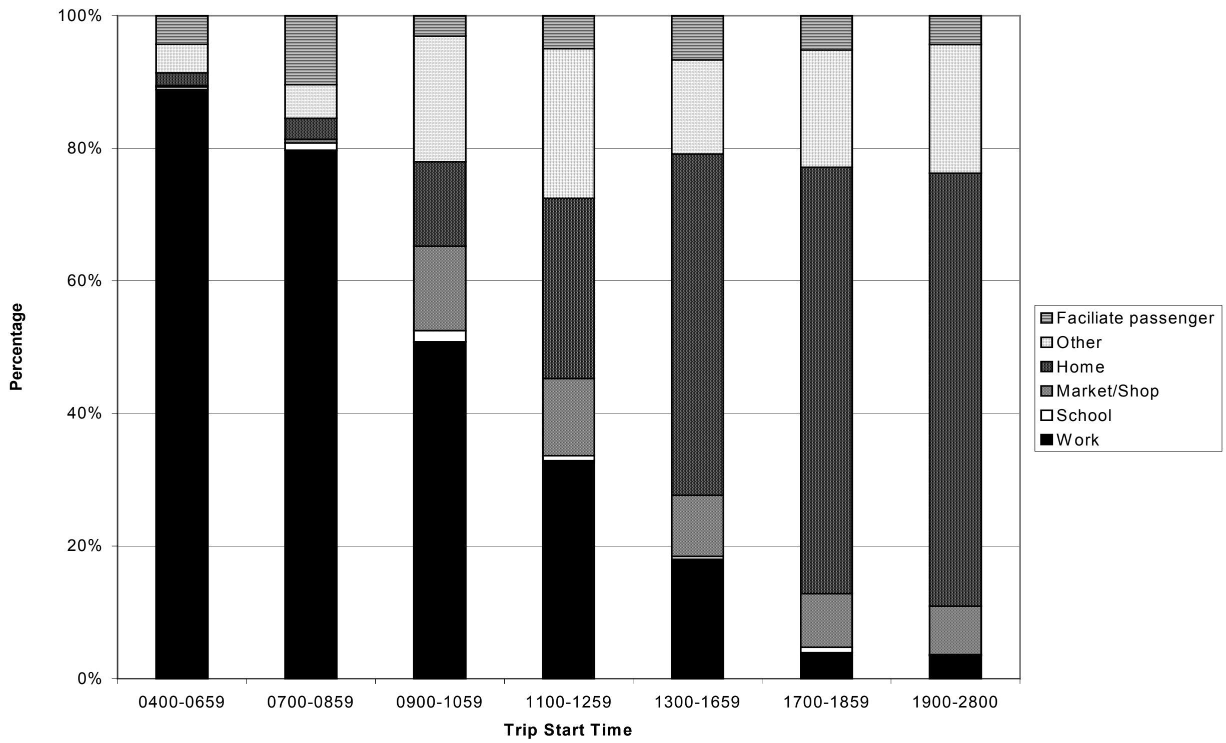



**Figure 3–5 Percentage of Trips by Destination Purpose & Start Time (Professional/Management/Technical)**

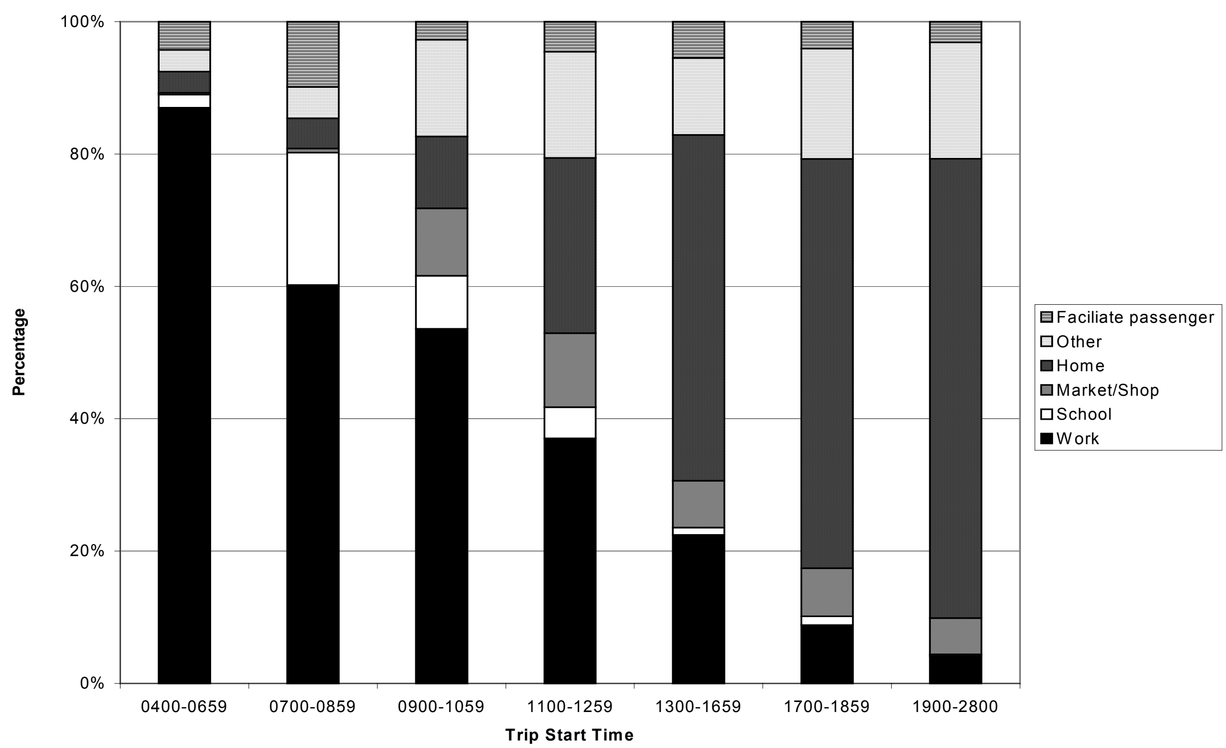



**Figure 3–6 Percentage of Trips by Purpose & Start Time (Retail Sales & Service)**

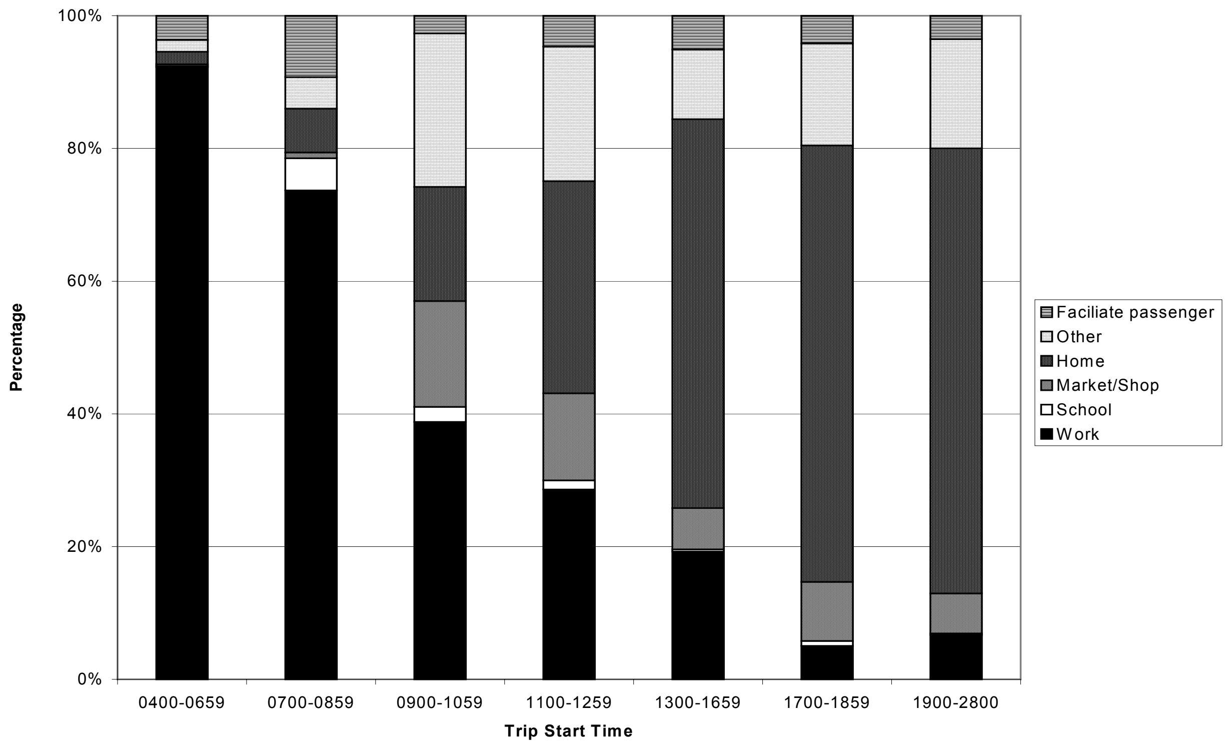



**Figure 3–7 Percentage of Trips by Destination Purpose & Start Time (Management/Construction/Trades)**

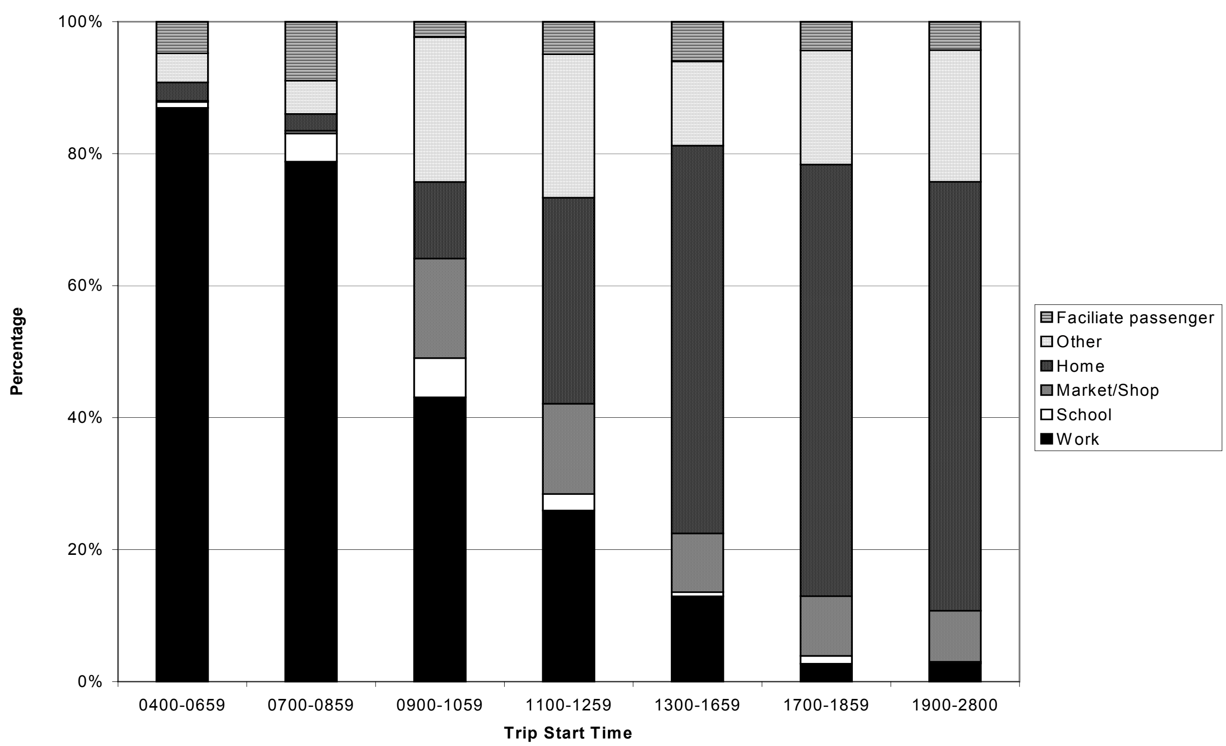



**Figure 3–8 Percentage of Trips by Destination Purpose & Start Time (General Office/Clerical)**

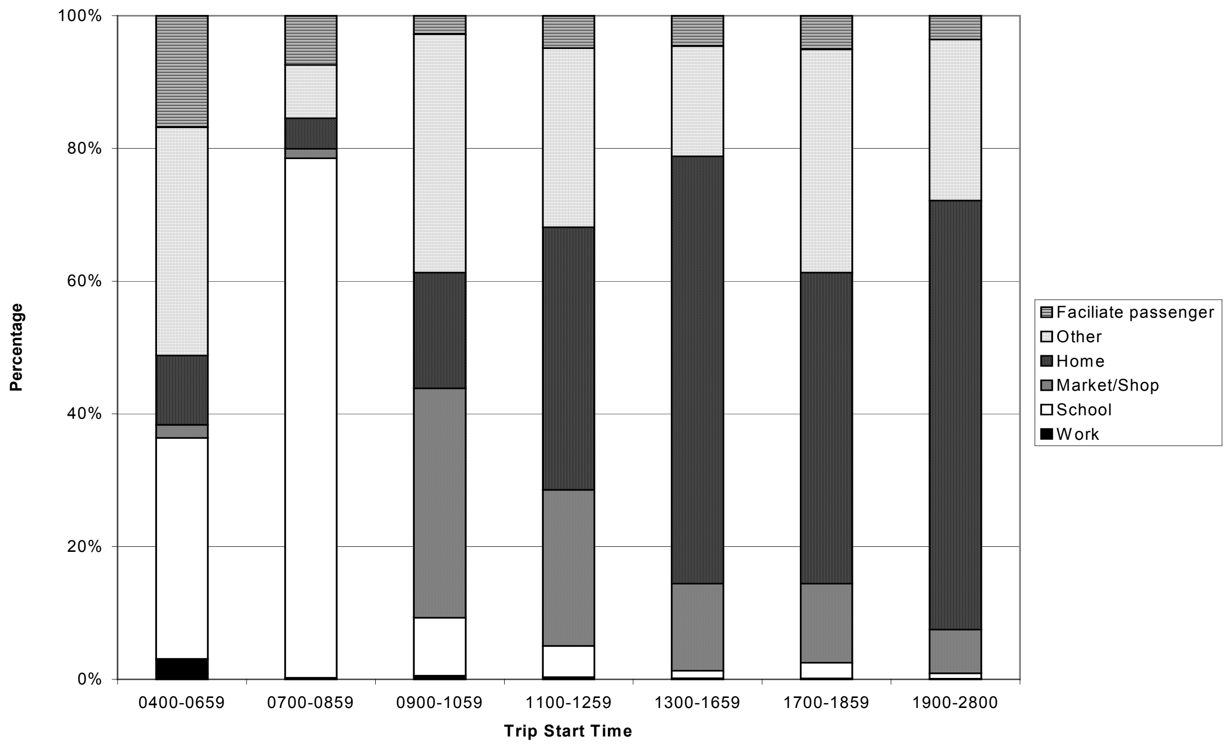



**Figure 3–9 Percentage of Trips by Destination Purpose & Start Time (Unemployed)**

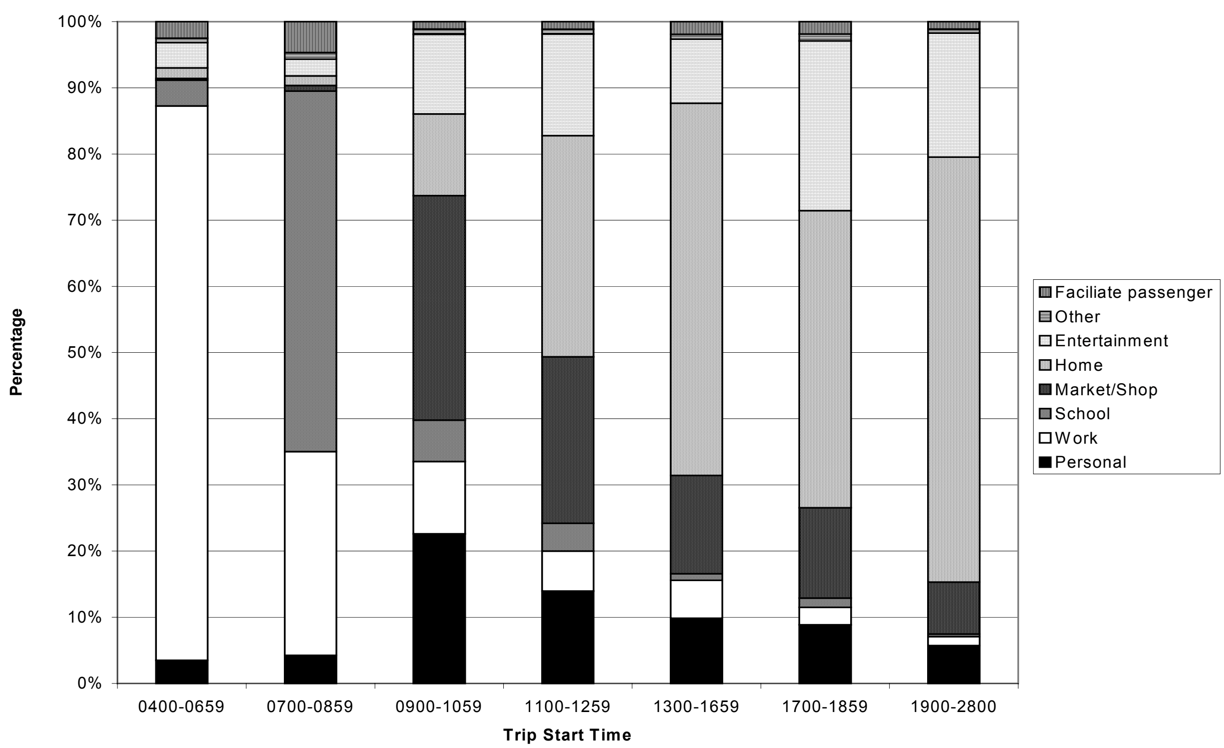



**Figure 3–10 Percentage of Trips by Destination Purpose & Start Time (Auto–Passenger)**

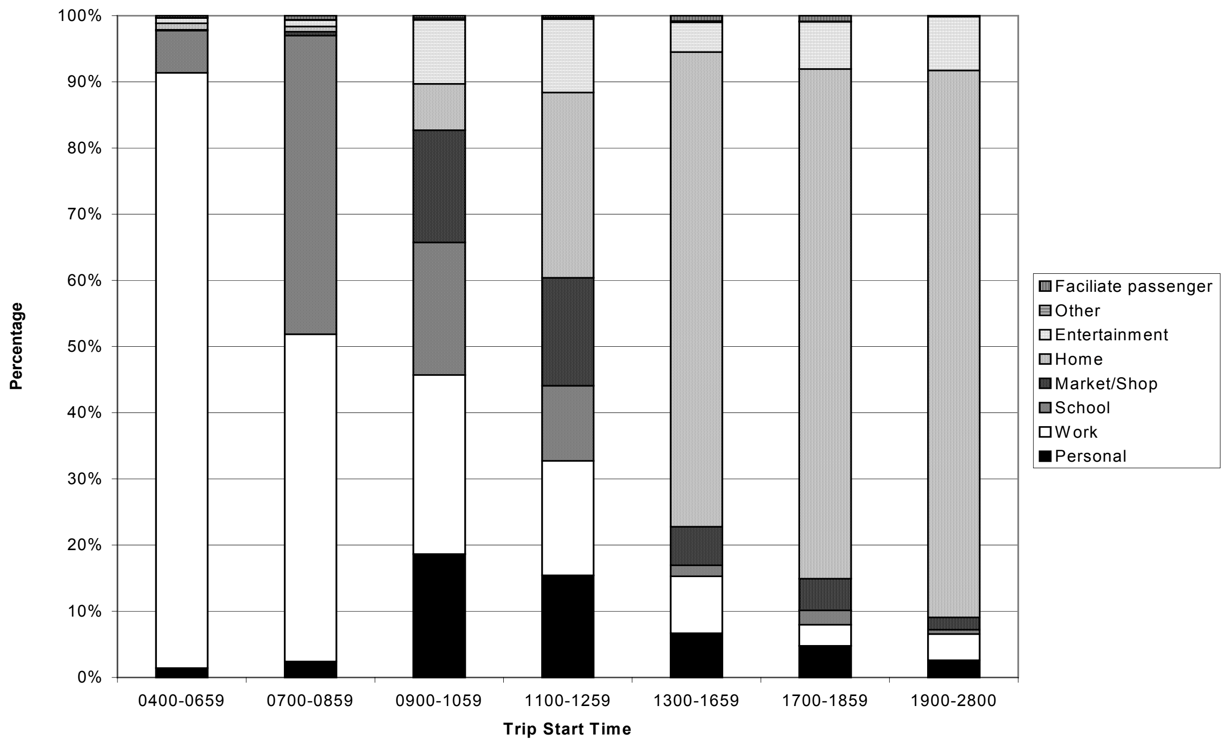



**Figure 3–11 Percentage of Trips by Destination Purpose & Start Time (Public Transit)**

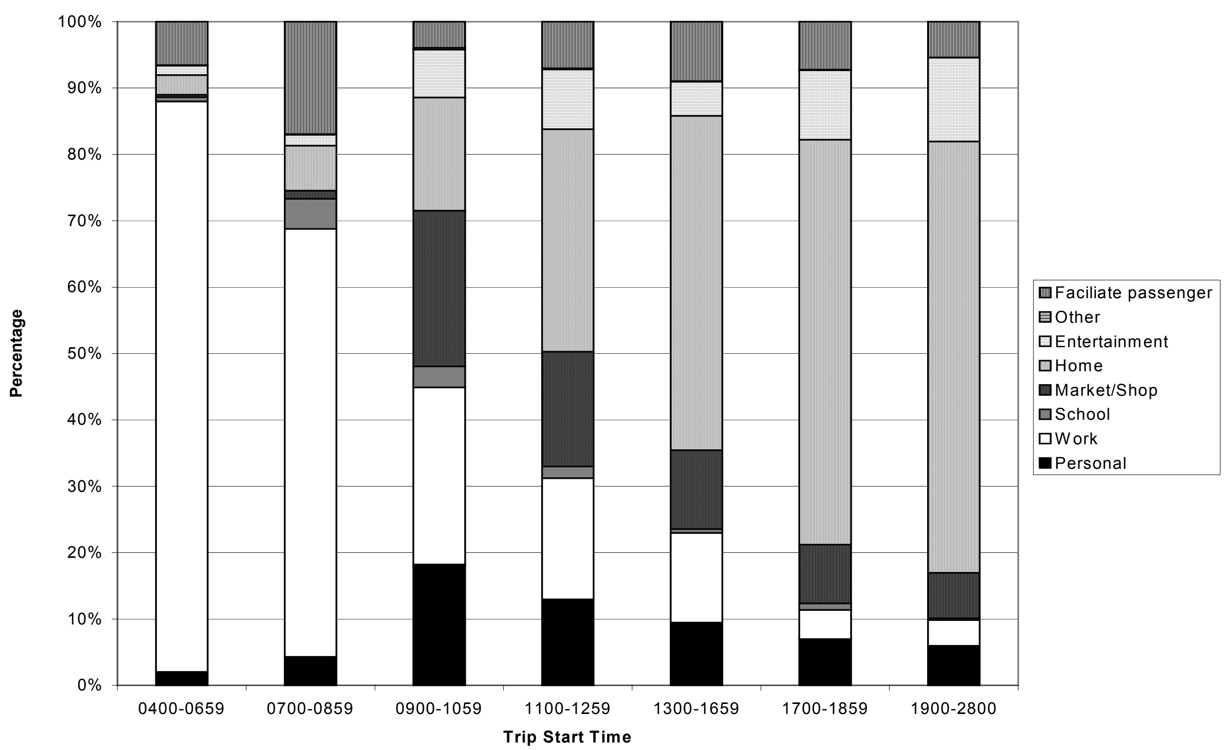



**Figure 3–12 Percentage of Trips by Destination Purpose & Start Time (Auto–Driver)**

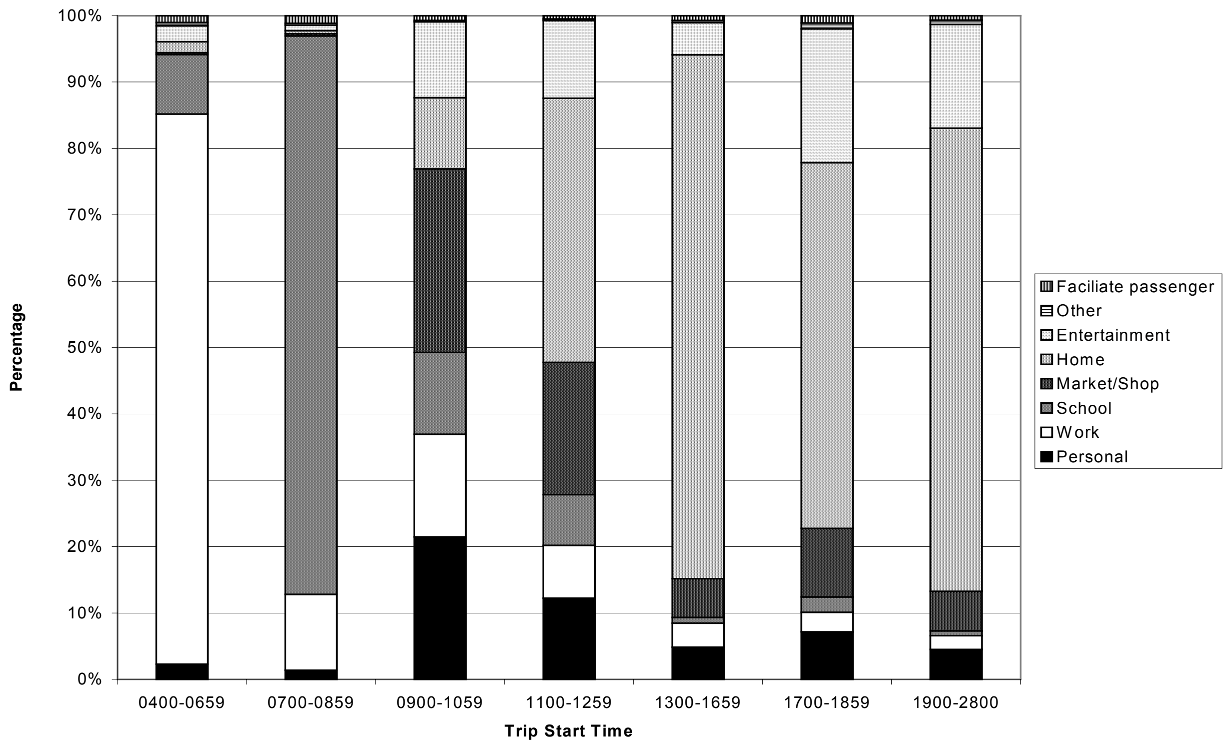



**Figure 3–13 Percentage of Trips by Destination Purpose & Start Time (No Driver's Licence)**

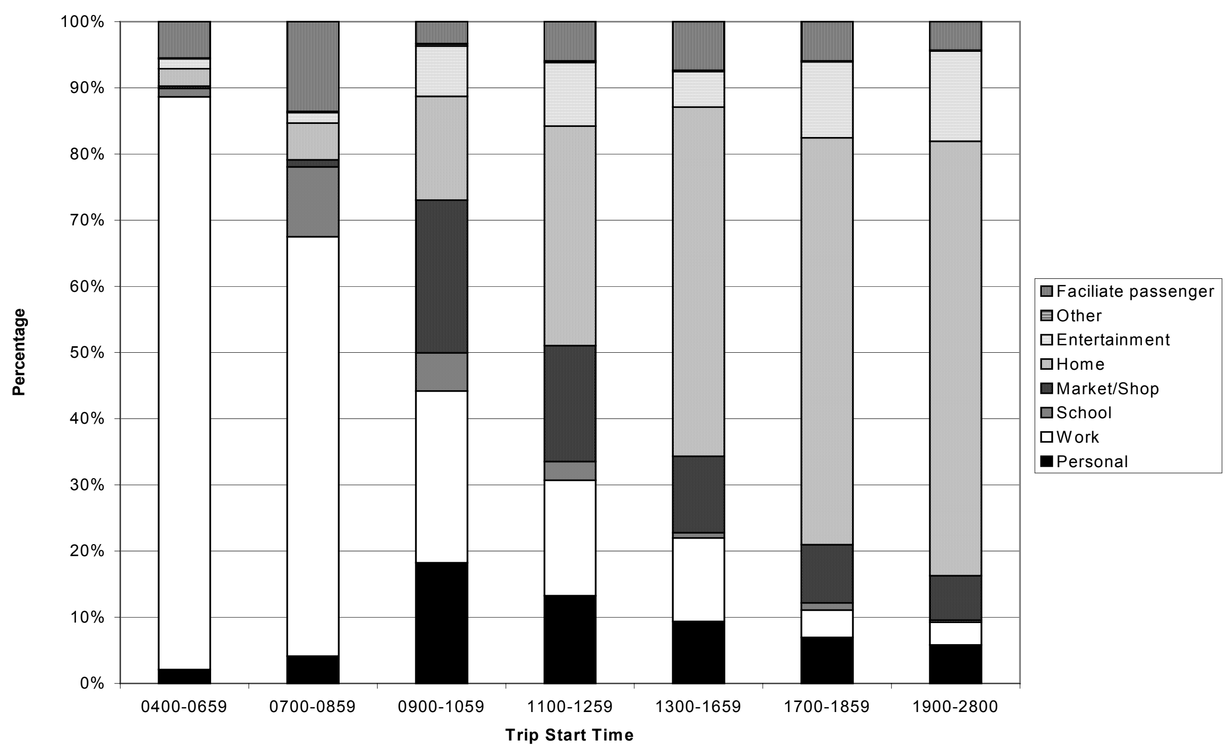



**Figure 3–14 Percentage of Trips by Destination Purpose & Start Time (Driver's Licence)**

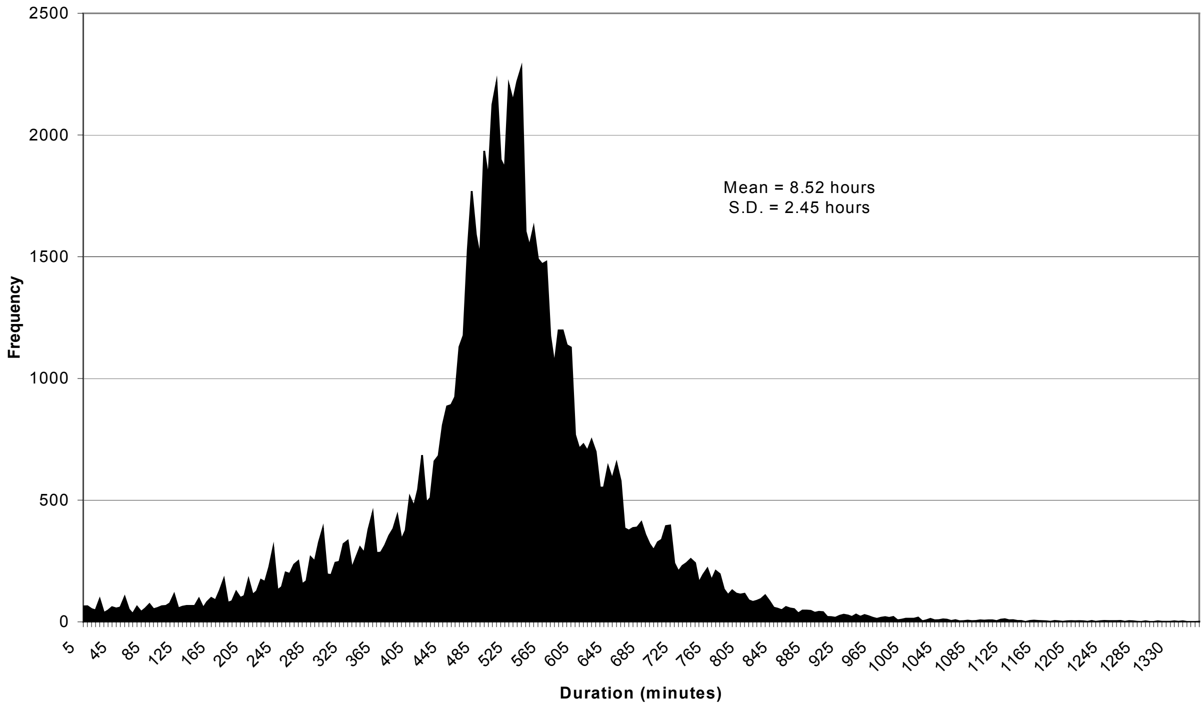



**Figure 3–15 Histogram of *Work* Activity Duration**

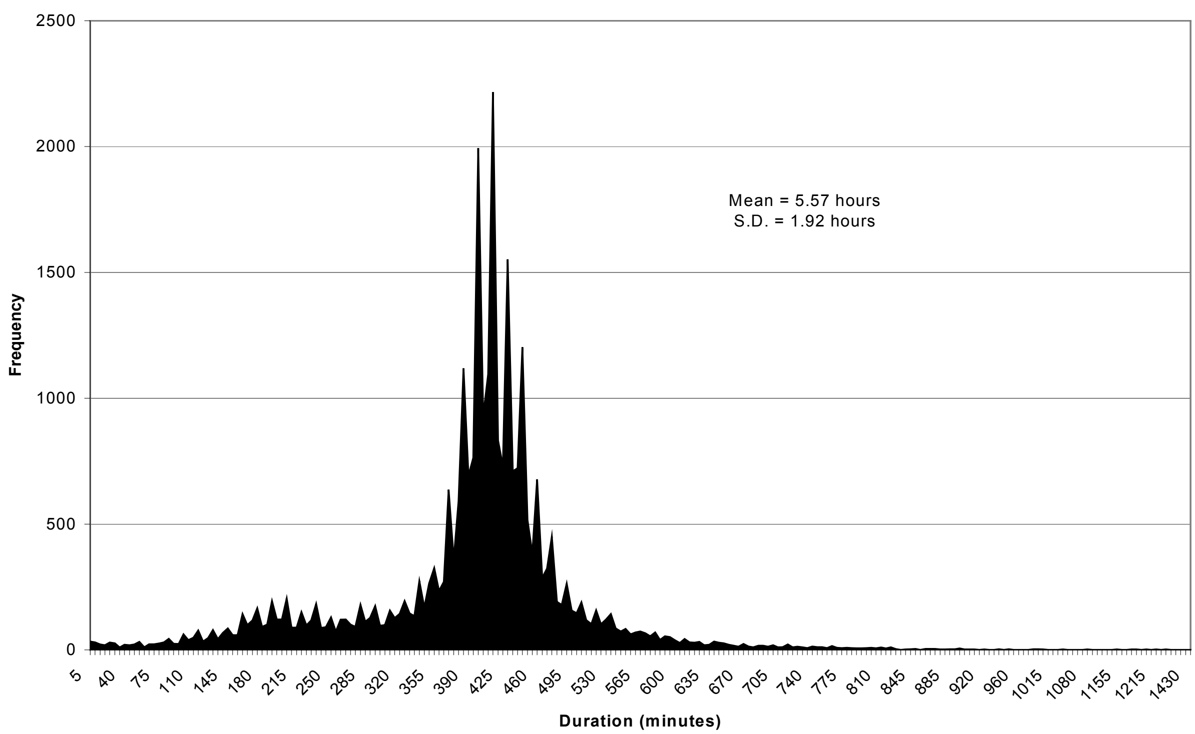



**Figure 3–16 Histogram of *School* Activity Duration**

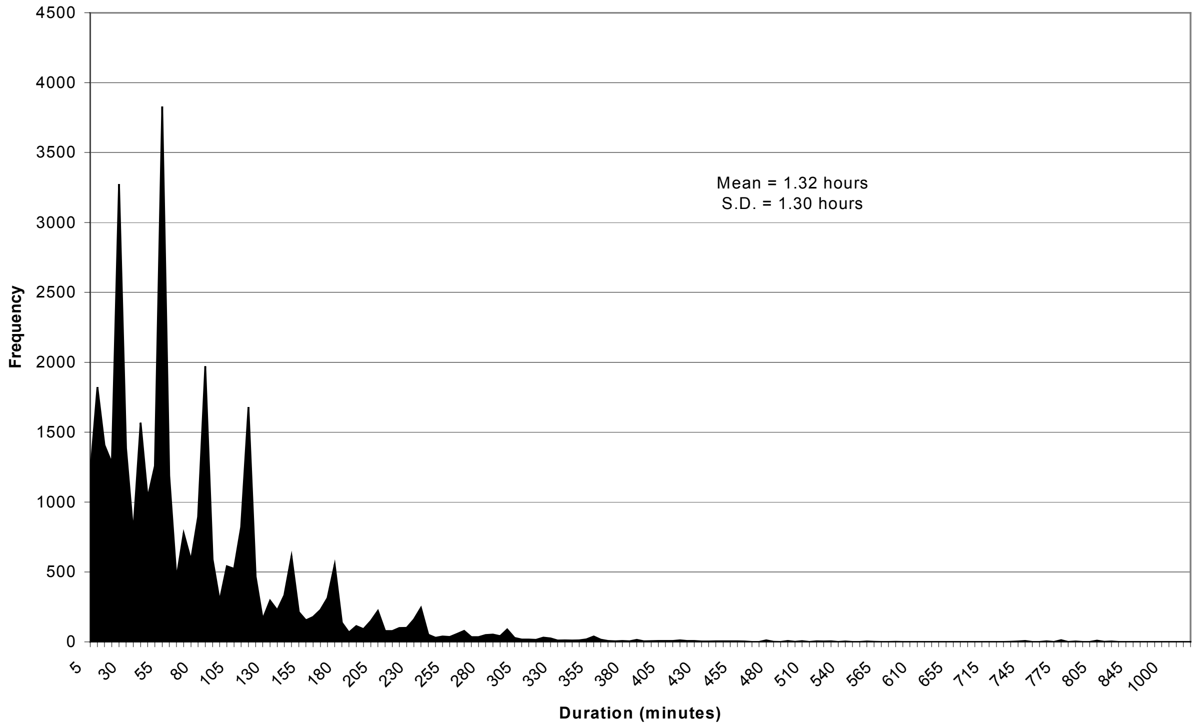



**Figure 3–17 Histogram of *Market* Episode Duration**

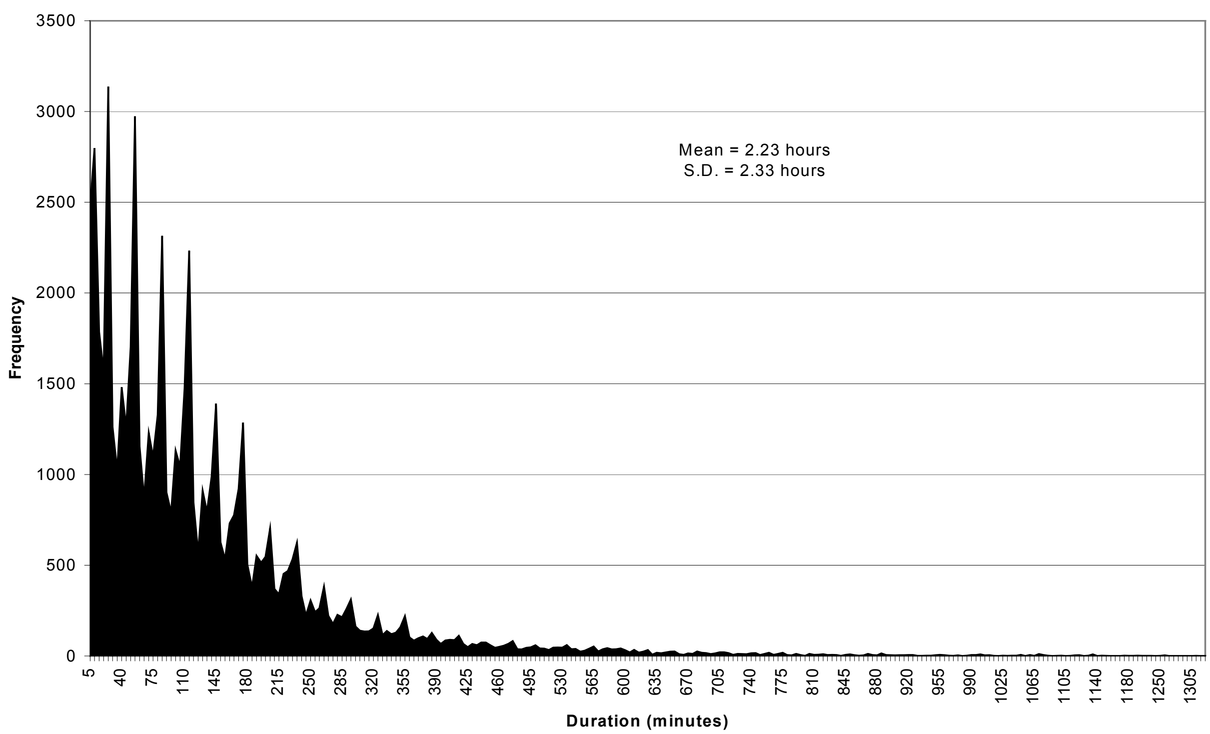



**Figure 3–18 Histogram of *Other* Episode Duration**

effect on average duration for both is obvious. Possessing a driver's licence appears to result in shorter *market* activity duration as compared to not having one. The start time of *work* activities is also affected by occupation. Again, as would be expected, the manufacturing/construction/trades sector starts earliest in the day. Afternoon and evening shifts can also be seen later in the day. As *market* activities are not constrained to begin at common periods of the day, the variance of these distributions is large. It can be seen, however, that full–time workers are forced to engage in this activity after *work*. Gender–based differences are present in considering the number of daily *work* episodes engaged in by males and females. Finally, possession of a driver's licence increases the likelihood of engaging in multiple *market* episodes.

As noted in the preliminary investigation, there is correlation between the duration, start time and frequency of activities. Take, for example, the correlation between the duration and frequency of *market* activities as functions of driver's licence possession (Figures 3–22 and 3–26). In considering both, it can be seen that persons with a licence engage in shorter, more frequent *market* episodes compared to the longer, less frequent episodes by persons without a driver's licence. This correlation is likely a result of the convenience associated with using the auto–driver mode; that is, if one cannot drive, they are more likely to fulfill their *market* requirements in fewer episodes than someone who could drive. Accordingly, these episodes take longer to complete.

Tests were also performed to determine statistical differences between the arithmetic means of the given variable and the overall distributions as well as between the variables themselves. This statistical analysis confirms what is observed graphically. T–tests (assuming equal population variances) were performed on all distributions to test for differences between sample means.[7] The test results for duration, start time and frequency are presented in Appendix D. The calculated t–statistics are tested against a critical value of 1.645 (a one–tailed, 5% significance level). As can be seen, most of the variables prove to be significantly different from

[7] The Kolmogorov–Smirnov goodness–of–fit test was also performed on some distributions, but was discontinued as it revealed almost identical results as those of the T–tests.

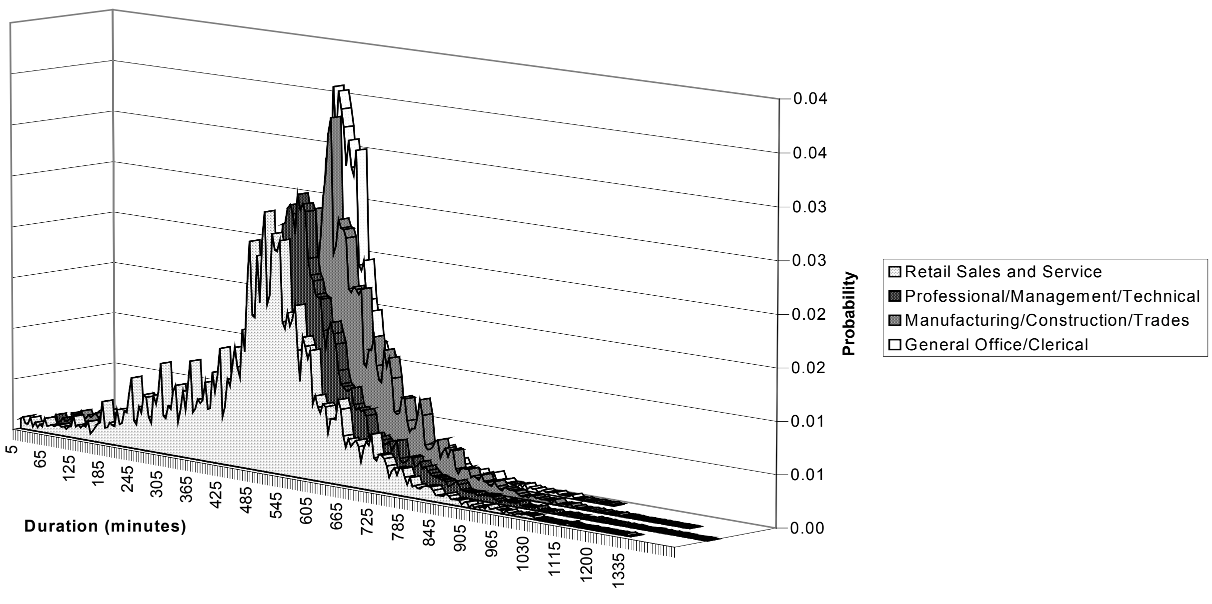



**Figure 3–19 PDF of *Work* Activity Duration by Occupation**

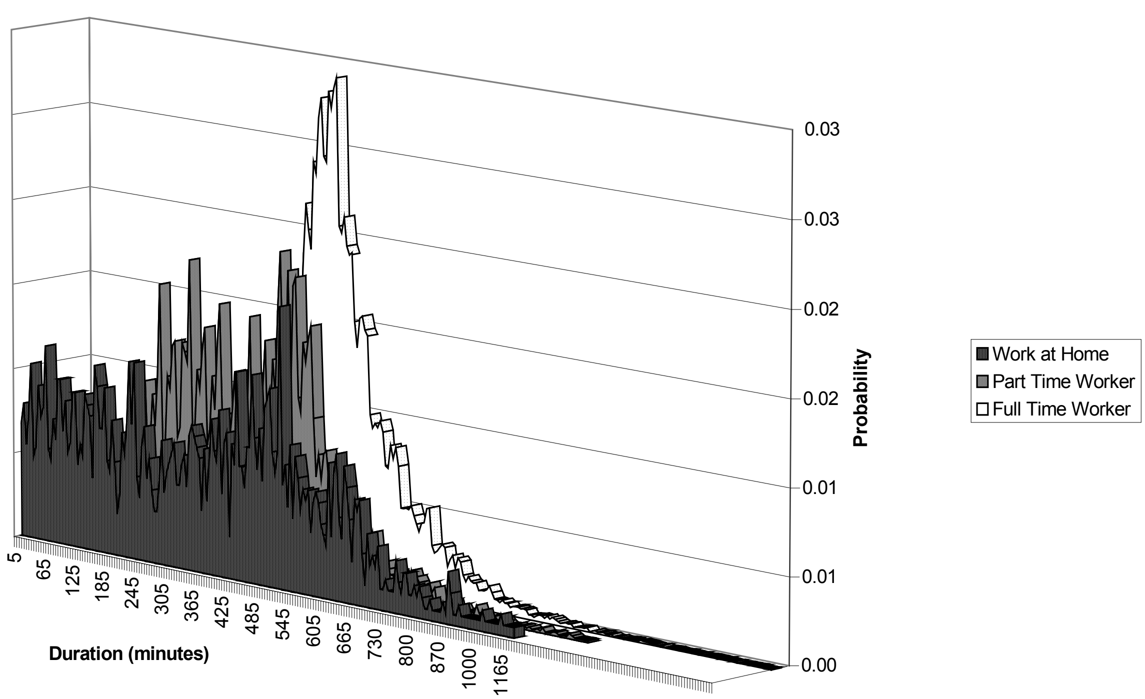



**Figure 3–20 PDF of *Work* Activity Duration by Employment Status**

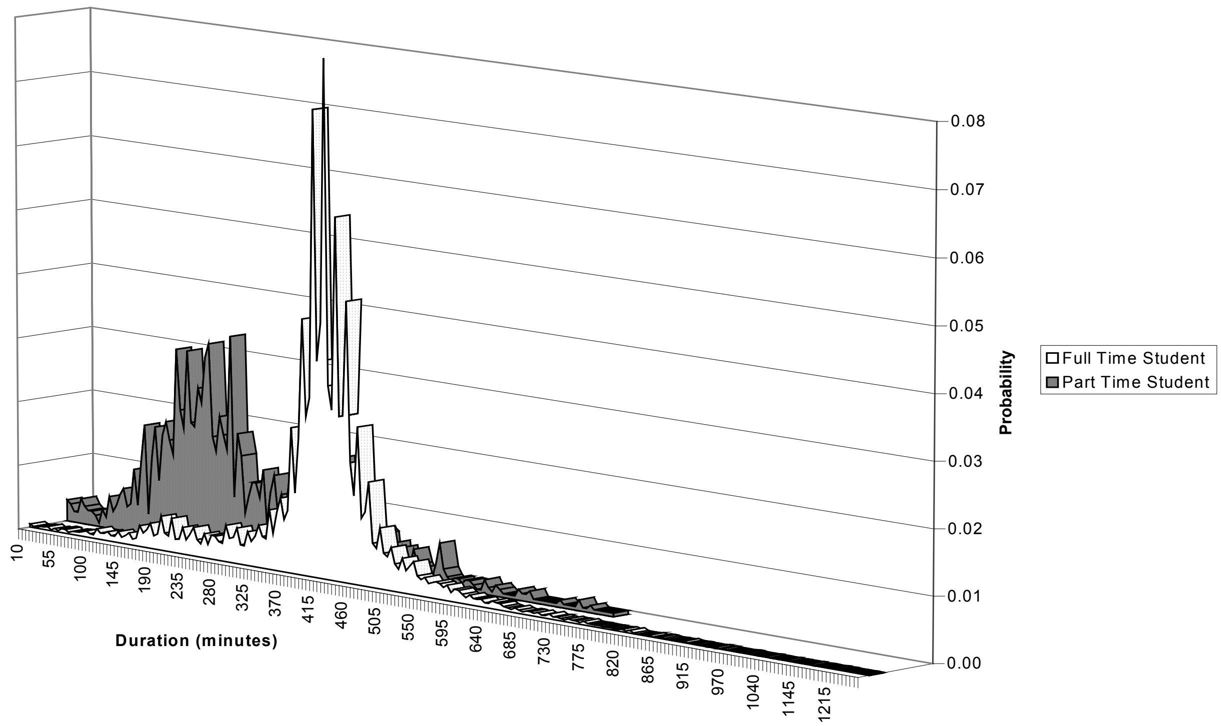



**Figure 3-21 PDF of *School* Activity Duration by Student Status**

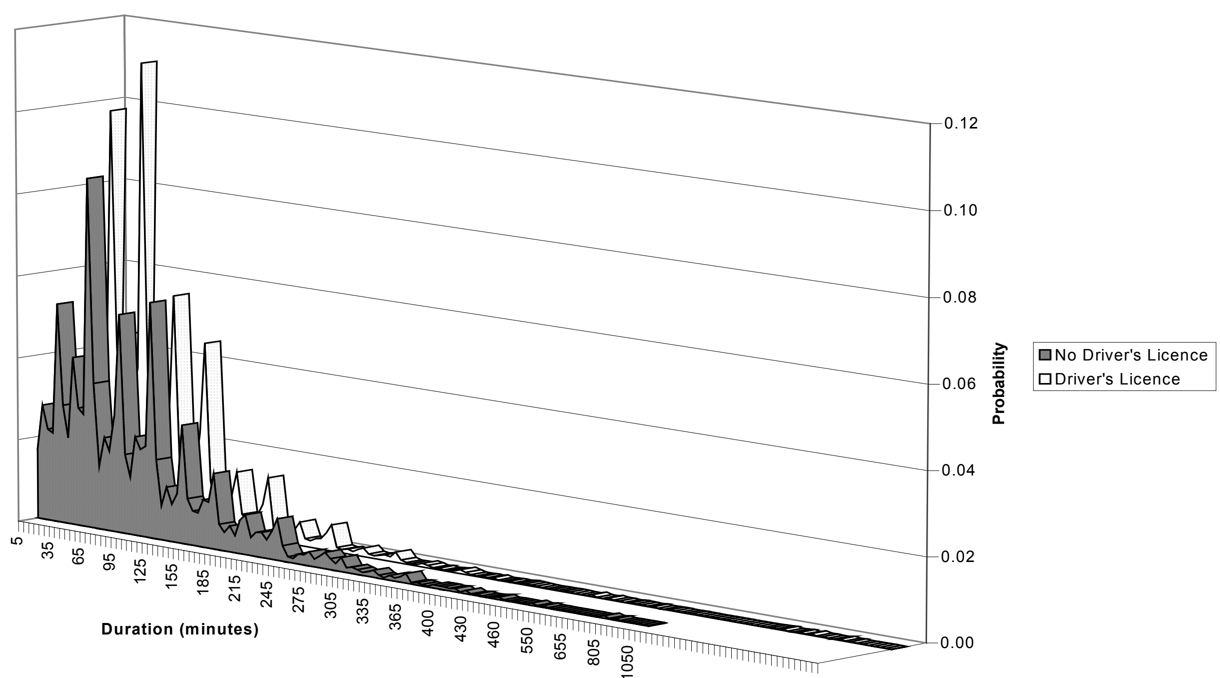



**Figure 3–22 PDF of *Market* Episode Duration by Driver's Licence Possession**

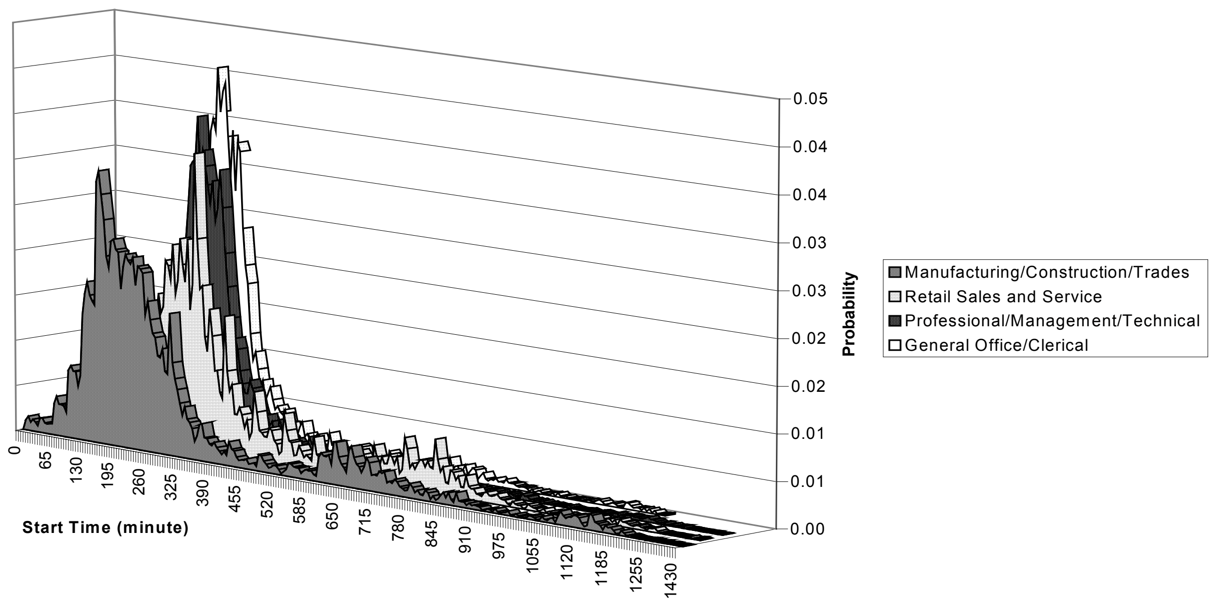



**Figure 3–23 PDF of *Work* Activity Start Time by Occupation**

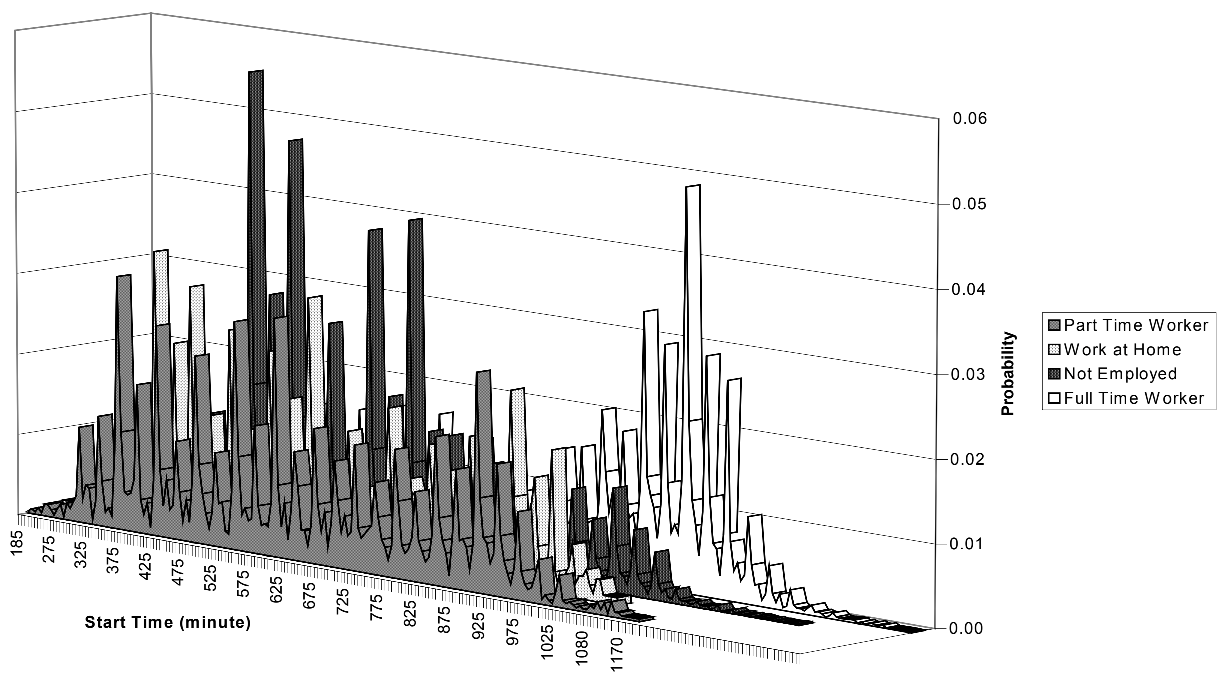



**Figure 3–24 PDF of *Market* Episode Start Time by Employment Status**

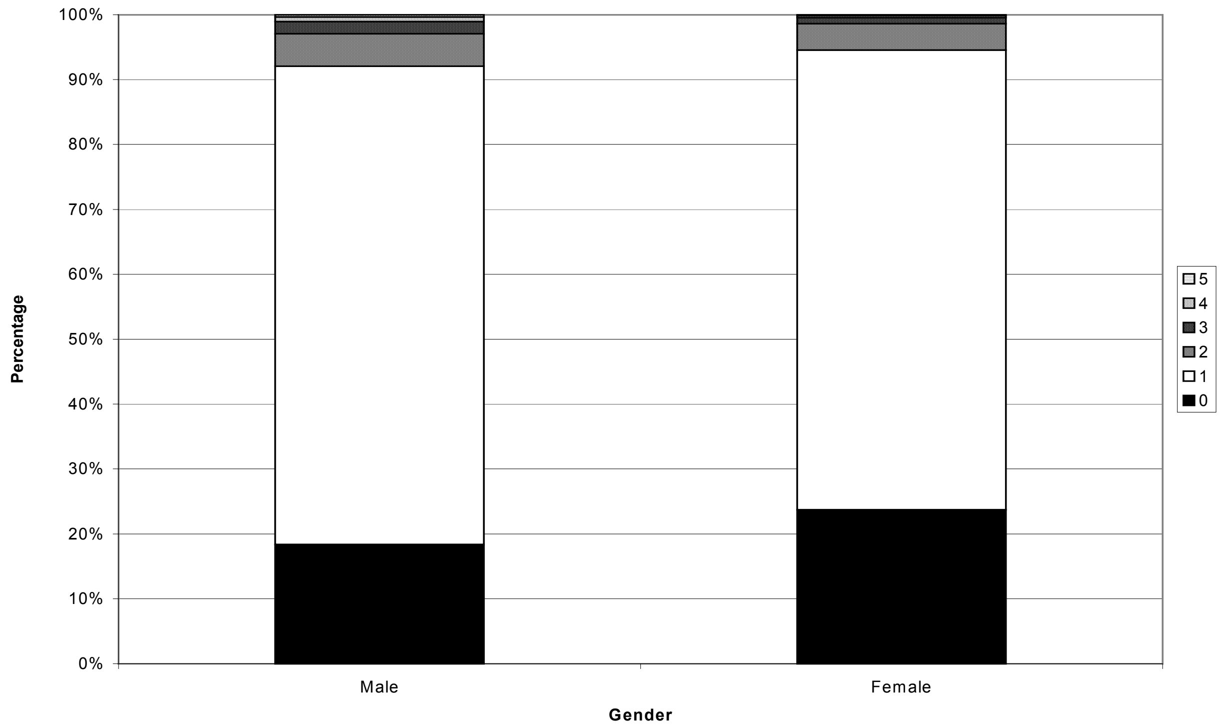



**Figure 3–25 Percentage of *Work* Episode Frequency by Gender**

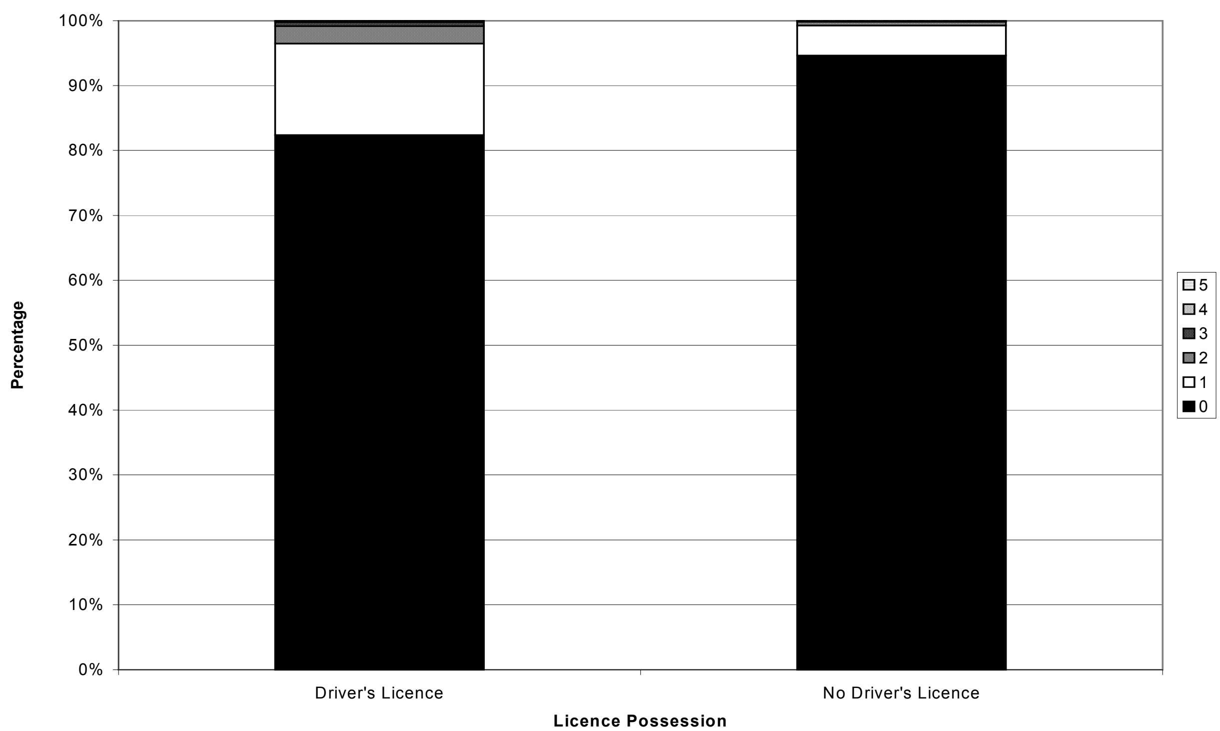



**Figure 3–26 Percentage of *Market* Episode Frequency by Licence Possession**

the overall distribution, as well as from each other. It is acknowledged that this is due largely to the high number of observations as, for example, a difference in means of only 5 minutes often proves to be statistically significant. These tests, however, are still helpful in determining which variables have the most impact on activity duration, start time and frequency.

## 3.4 JOINT DURATION, START TIME & FREQUENCY

The previous analysis looked at the duration, start time and frequency of each activity episode in the database individually. However, as mentioned previously, variables were created that identified when activity episodes occurred between two or more people within a household at the same time and place, or, in other words, jointly. It is hypothesised that these joint activities have somewhat different properties than their individual counterparts, especially in terms of their start time and frequency. After all, these activities are constrained to occur when all parties involved can agree to the schedule and therefore involve what essentially can be called a contract between parties to fulfil certain obligations to each other. For this reason, it is likely that these joint activities have a certain priority over individual ones in terms of their generation and scheduling.

Given that joint activity participation is an integral part of a household–level model, an analysis was performed to investigate these differences. The analysis focuses on joint *market* and *other* activity types only. This is because concurrent *work* or *school* activities are not considered as joint as, for example, two children attending the same school at the same time are not really participating in the activity jointly[8]. Also, joint participation in *home* activities is not observed in the data set and therefore cannot be included.

It was anticipated that joint activity episodes would be similar in duration, would have later start times and would occur less frequently relative to independent activity episodes. These anticipated differences were attributed to the increased constraints inherent to joint activities relative to independent activities, as explained above, which would account for reduced frequency. Later start times would result from the likelihood that people are more able to participate in joint activities outside of the work/school day, which would most often be in the

[8] In other words, they are not gaining utility from altruism, companionship or efficiency, which are the motivators of joint activity engagement as proposed by Townsend (1987).

afternoon and evening. Table 3–1 compares overall average values for *market* and *other* independent and joint activity episodes.[9] The results support *a priori* expectations.

| Attribute | Market | | Other | |
|---|---|---|---|---|
| | Independent | Joint | Independent | Joint |
| Duration | 1.27 hr | 1.40 hr | 2.25 hr | 2.13 hr |
| Start Time | 2:20 p.m. | 2:50 p.m. | 3:10 p.m. | 4:10 p.m. |
| Frequency | 0.28 | 0.16 | 0.55 | 0.21 |

**Table 3–1 Comparison of Independent & Joint Statistics**

Previous activity/time–use studies have investigated the relationship between the independent and joint behaviour of adults using various person and household explanatory variables. Gliebe and Koppelman (2001) found that independent and joint 'maintenance' and 'leisure' time–share utility[10] varied with the number of children and vehicles present in the household. That is, time–shares shifted from joint maintenance and leisure activities to independent maintenance and leisure activities as the number of children in a household increased and the number of vehicles increased from one to two or more. The proposed explanations are that "parents with children specialise in their tasks and find less time for out–of–home joint leisure activities" and that one–vehicle households are forced to participate in activities jointly because of the need for vehicle sharing. Based on this reasoning, joint activities were further investigated in terms of these household–level variables. Figures 3–27 and 3–28 depict independent and joint market time–share without and with children[11], respectively. Similar to the results noted above, one can see that as a household shifts from one without

[9] It should be noted that, here, values are calculated based on aggregation to a joint/household level. For example, average episode frequency is expressed as 'occurrences per household per day', as opposed to the person–level calculation of 'occurrences per person per day' used in the previous analysis. In determining average values, it is important not to let the number of people participating in the joint activity weight the observations. Further complicating matters is the appropriate definition of joint activity. For example, if three household members participate in an activity together, should this be defined as one joint activity episode or three joint activity episodes?

[10] The *maintenance* and *leisure* categories used in Gliebe and Koppelman's analysis are somewhat analogous to the *market* and *other* categories used in this model. Also, time–share is analogous to total daily activity duration.

[11] Children are defined as persons 15 years of age and below.

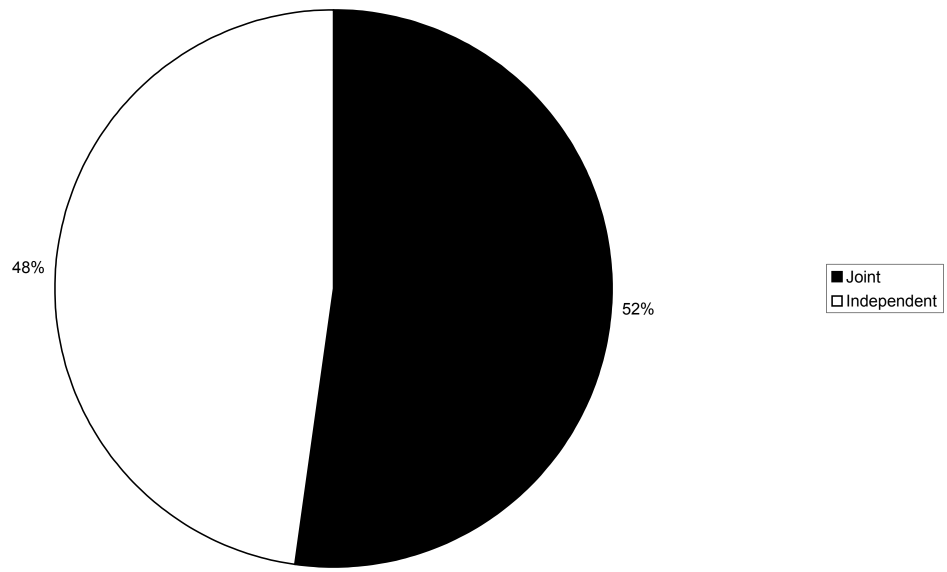



**Figure 3–27 *Market* Time–Share (Household with No Children)**

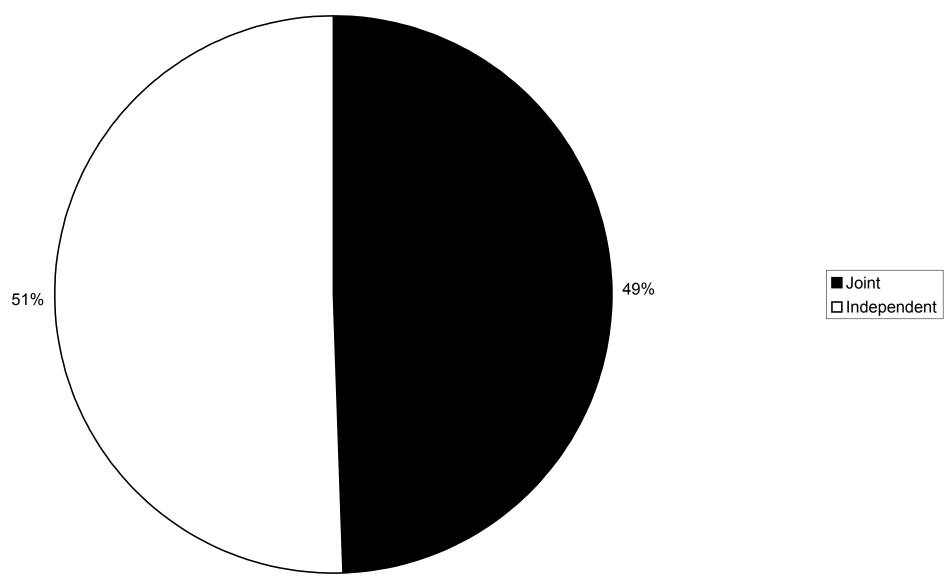



**Figure 3–28 *Market* Time–Share (Household with Children)**

children to one with children, *market* time–share shifts from joint to independent.[12] Figures 3–29 and 3–30 reveal the same trend for *other* activities in shifting from a one–vehicle household to a two– or greater vehicle household.[13]

## 3.5 MODE CHOICE & TRIP CHAINING

All travel demand models obviously must consider the mode of travel. Trip–based models, however, need only consider the single mode used for the single trip made in the analysis period (commonly the morning peak period). A 24–hour, household level model, on the other hand, should account for such phenomena as multi–mode trip chains and the use of one vehicle by several members of the household at different periods of the day. For example, trip–based mode choice models such as that by Miller and Rhamey (1987) could assume that once a household vehicle has been taken by a household member, it was no longer available to other household members, thus making vehicle availability simple to determine. This assumption obviously can not realistically be applied to the daily travel and activity patterns of a household over the period of a day, thus making the availability of a vehicle to household members much more difficult to determine.

To assist with the mode choice component of the model, descriptive analyses were performed to investigate multi–mode trip chaining as well as mode choice variation as functions of vehicle availability and activity episode type. The data revealed that 94.2% of all identified trip chains were made using a single mode and that 5.5% were made using two. Thus, a significant amount exhibited a multi–modal choice behaviour. Figure 3–31 extends this to incorporate travel mode type. As would be expected, the single mode chain is dominated by the auto–driver mode with 99% of the share. Figures 3–32, 3–33, and 3–34 explore the first mode used in each trip chain as a function of the number of modes used, for chains containing a *work*

[12] The changes, however, are not drastic. A likely reason is that, here, joint trips include those made with children aged 11 to 15, whereas Gliebe and Koppelman considered only those with adults only. The comparison is not quite equal, as some joint trips previously made with a spouse would likely be replaced with joint trips made with children.

[13] Again, it is expected that these changes would be larger given a different definition of a joint activity, as there is likely a high positive correlation between household size/number of children and automobile ownership (i.e. joint trips made with a spouse using one vehicle may be replaced by joint trips with children using several vehicles).

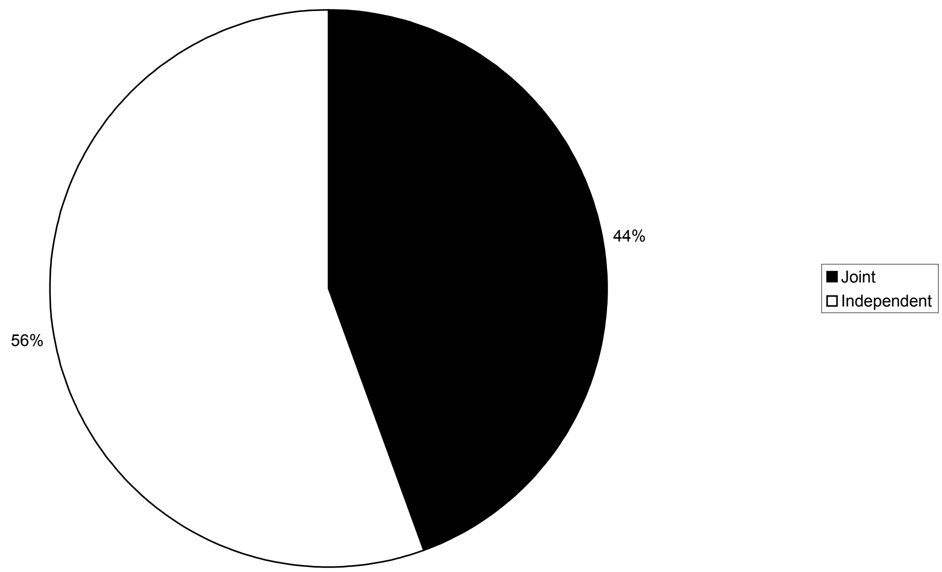



**Figure 3–29 *Other* Time–Share (Household with 1 Vehicle)**

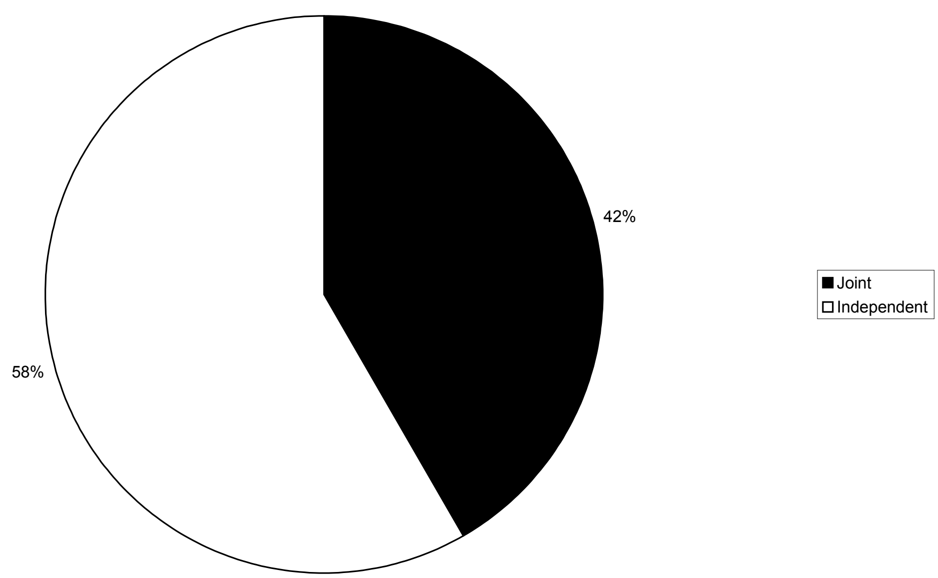



**Figure 3–30 *Other* Time–Share (Household with 2+ Vehicles)**

episode, a *school* episode or neither a *work* nor *school* episode, respectively. Again, one sees the dominance of the auto–driver mode in single mode chains, with the auto passenger, transit and walk modes gaining share in the multi–modal trip chains. Age effects are very apparent in comparing the school chain with the other two, as the auto passenger and walk modes are more dominant. Figure 3–35 shows the second mode used given the first mode used for two mode trip chains. Again, the auto–passenger, transit and walk modes dominate modes such as auto–driver and bicycle, modes which are not logical second modes.

One would expect mode choice to vary with availability of the household vehicle (if any) as well as with activity episode type; that is, the mode choice set is dependent on whether household vehicles are being used by other household members as well as on the suitability of the mode for the given chain. To investigate these factors, the mode choices for varying activity types were explored as functions of four measures of auto availability. The four categories of auto availability used are: (i) no driver's licence; (ii) driver's licence, no household vehicles; (iii) driver's licence, no vehicle available[14]; and (iv) driver's licence, vehicle available.[15] Figures 3–36 through 3–39 show the mode splits for various activity types[16] for each of the auto availability categories[17]. Those with no driver's licence are obviously highly dependent on the auto–passenger, transit and walk modes. In this category, the household can still possess a vehicle, resulting in high proportions of the auto passenger–mode in joint activities. Those with a driver's licence but who live in a household with no vehicle are still dependent on the auto–passenger, transit and walk modes, but the share shifts significantly from auto–passenger to transit. The remaining auto–passenger share is likely the result of carpooling. Notice also the auto–driver share present in the *work–business* activity type[18], representing those using a company vehicle to travel. Those with a licence but no available vehicle are again heavily

[14] This should be distinguished from not possessing a household vehicle.
[15] A binary 'Yes' or 'No' variable was created to determine vehicle availability based on whether a household vehicle was present at home when a person left home. Availability remained the same for the remainder of the trip chain. Details of the rules and assumptions are presented in Appendices B and C.
[16] These are not the TTS activity types used in previous examples. They are adaptations that will be defined in the next chapter. For now, it is the general acitivity that is important.
[17] Note that a significant proportion of trip chains involved more than one activity, the implication of which is that not all modes were chosen specifically for each activity type. Assuming that each trip chain has a single mode–influential activity, it would be impossible to determine which activity of each chain was the representative activity.
[18] This activity will be defined subsequently. It refers to a work activity engaged in a location that is not the person's usual place of work.

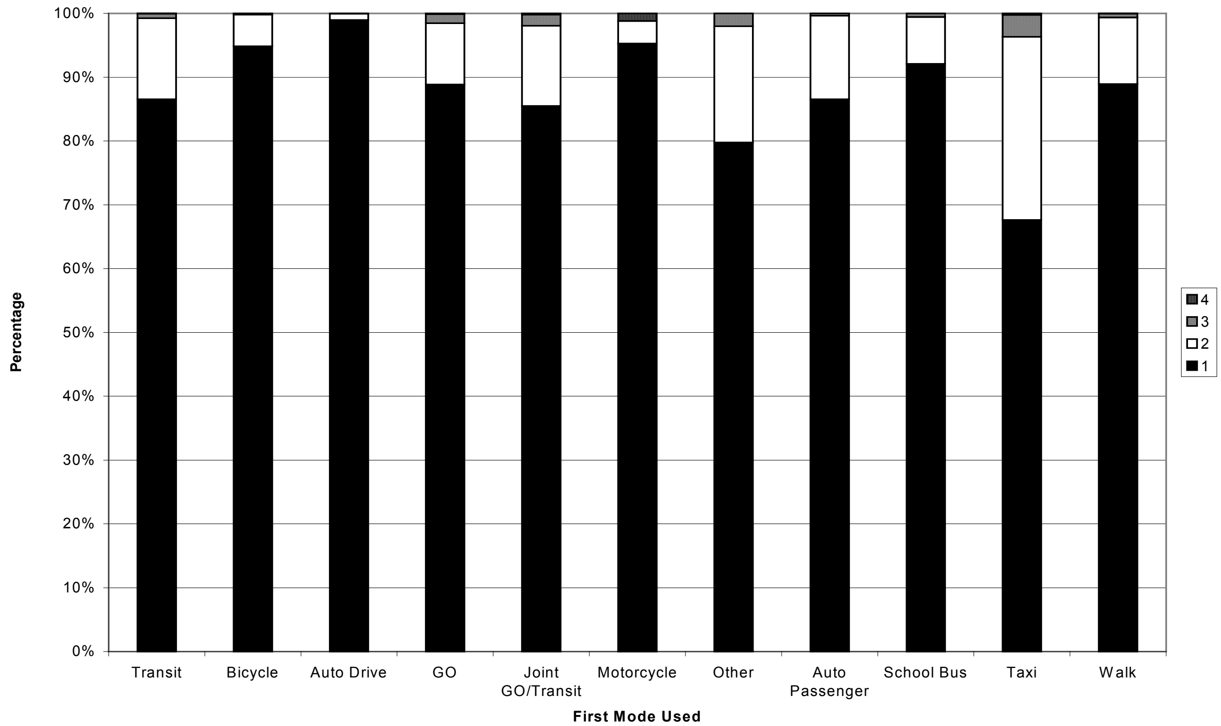



**Figure 3–31 Percentage of First Mode Used by Number of Modes in Trip Chain**

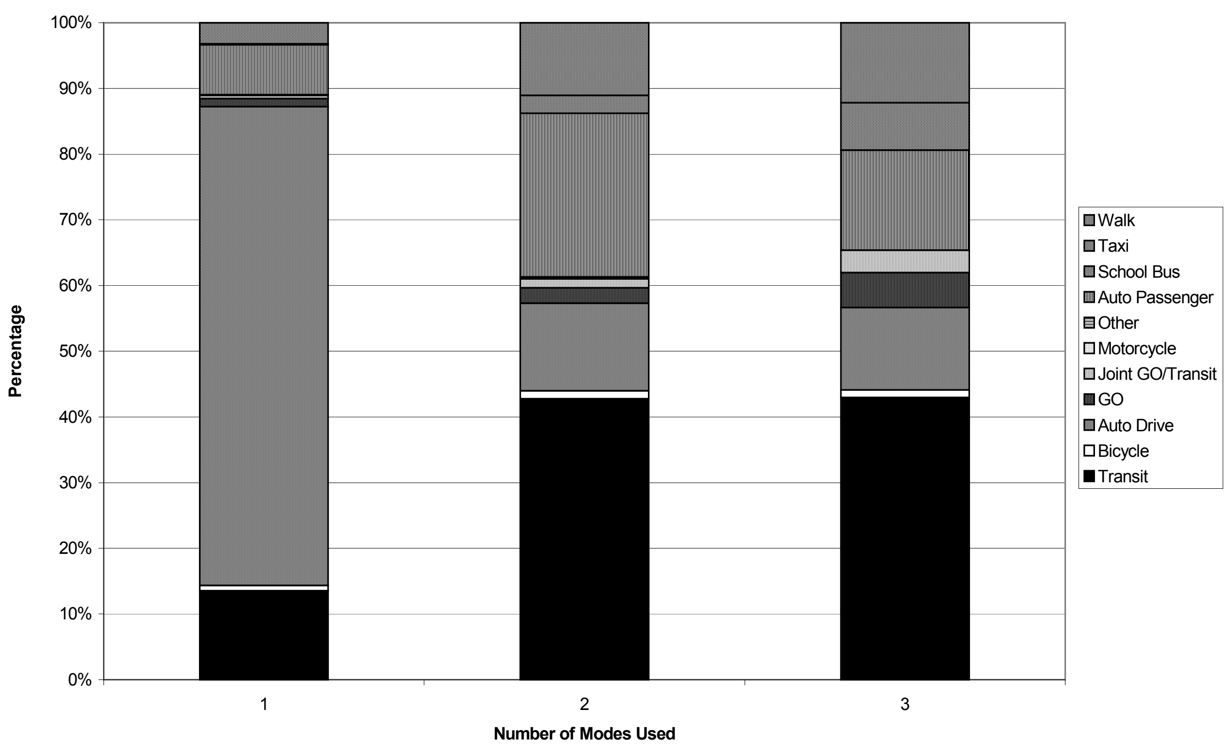



**Figure 3–32 Percentage of Number of Modes by First Mode Used in *Work* Trip Chain**

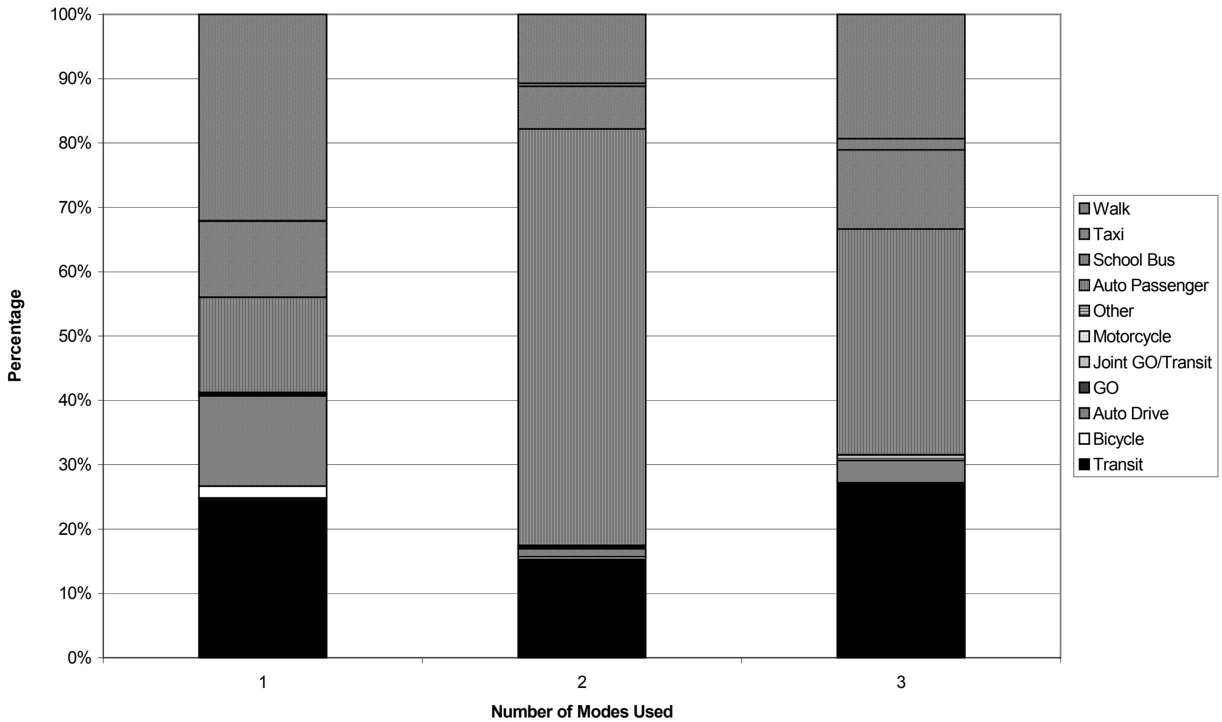



**Figure 3–33 Percentage of Number of Modes by First Mode Used in *School* Trip Chain**

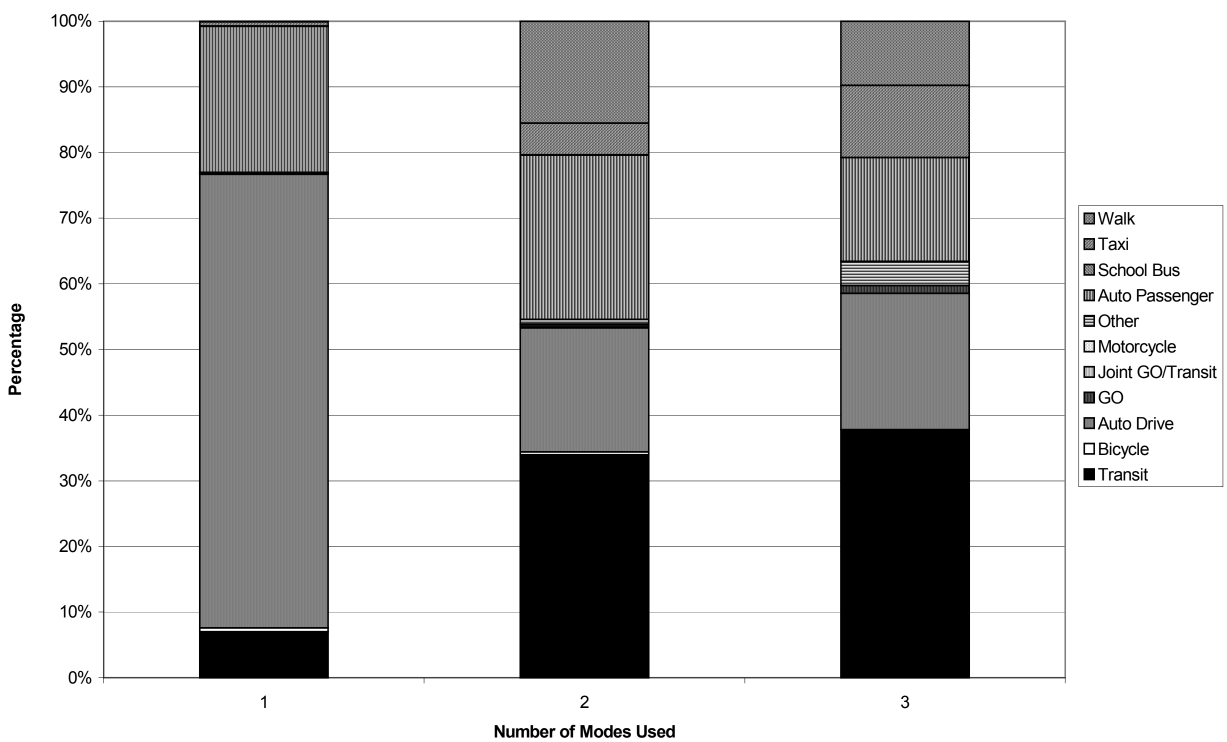



**Figure 3–34 Percentage of Number of Modes by First Mode Used in Neither *Work* or *School* Trip Chain**

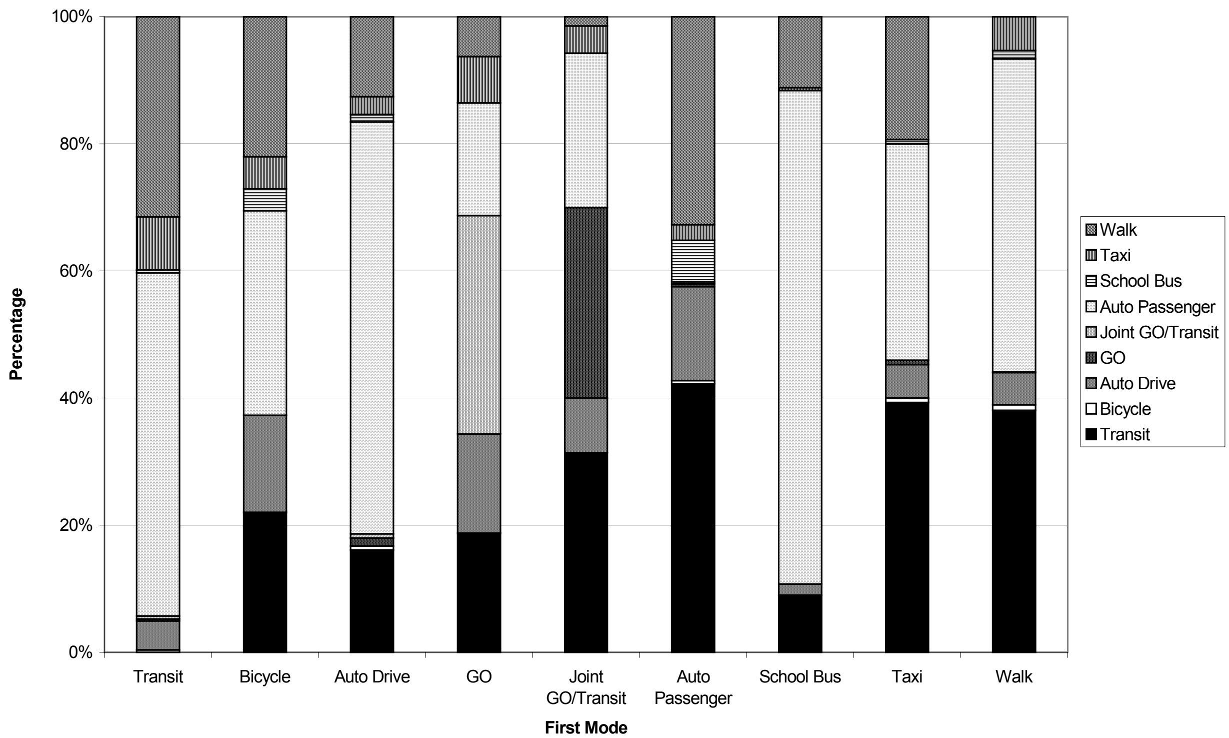



**Figure 3–35 Percentage of First Mode by Second Mode for Two Mode Chains**

dependent on the auto passenger mode, especially for joint activities.[19] The auto–driver shares present in some activity categories are likely the result of borrowing a non–household vehicle (such as from work or a friend).[20] Finally, the dominance of the auto–driver mode for those with a licence and an available vehicle is not surprising.

## 3.6 LOCATION CHOICE & TRIP CHAINING

As the daily activities of a household are distributed in time and space, there must be an important connection between the choice of activity location and the sequence in which they are carried out, which is presumably motivated by efficiency and convenience. It is therefore critical to investigate these phenomena before attempting to model them. Given, however, that there are

[19] The availability variable assumed, given two people leaving home by auto at the same time (one auto driver, one auto passenger), that a vehicle would be subtracted from the household fleet due to the driver, which often resulted in a vehicle not being available for the auto passenger.

[20] As the auto availability variable is based on household vehicle ownership, this behaviour created such illogical results.

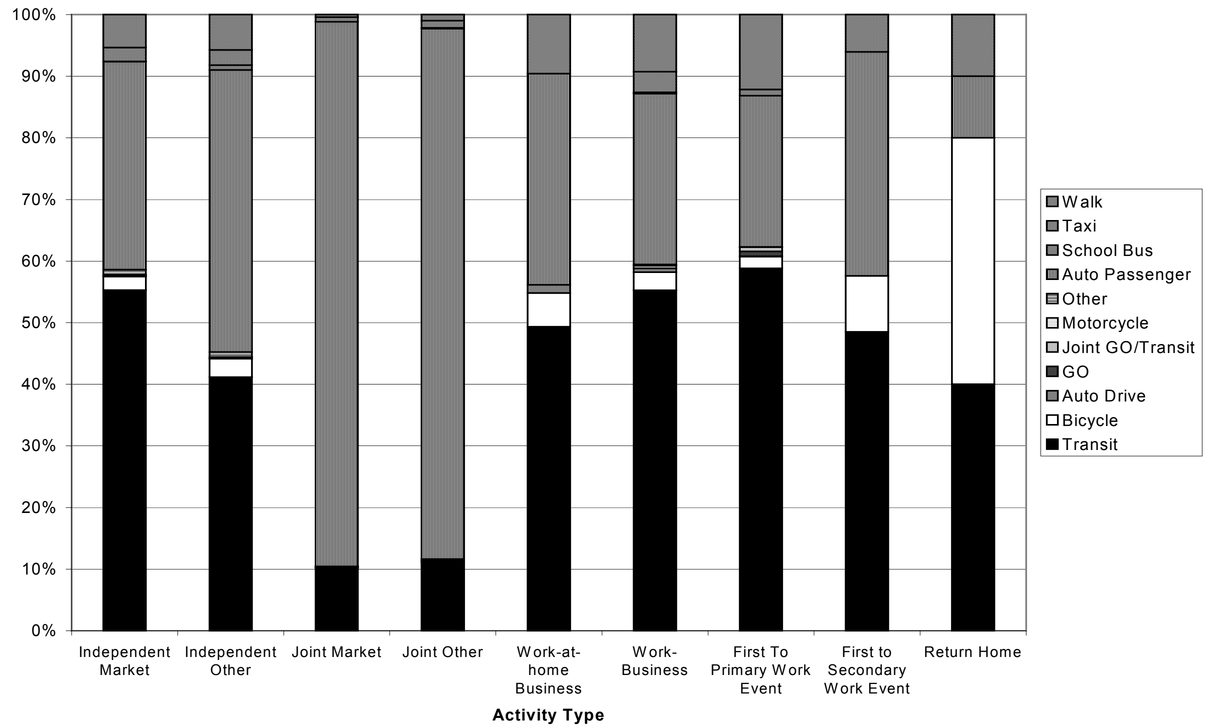



**Figure 3–36 Percentage of Activity Type by Mode Used for Trip (No Driver's Licence)**

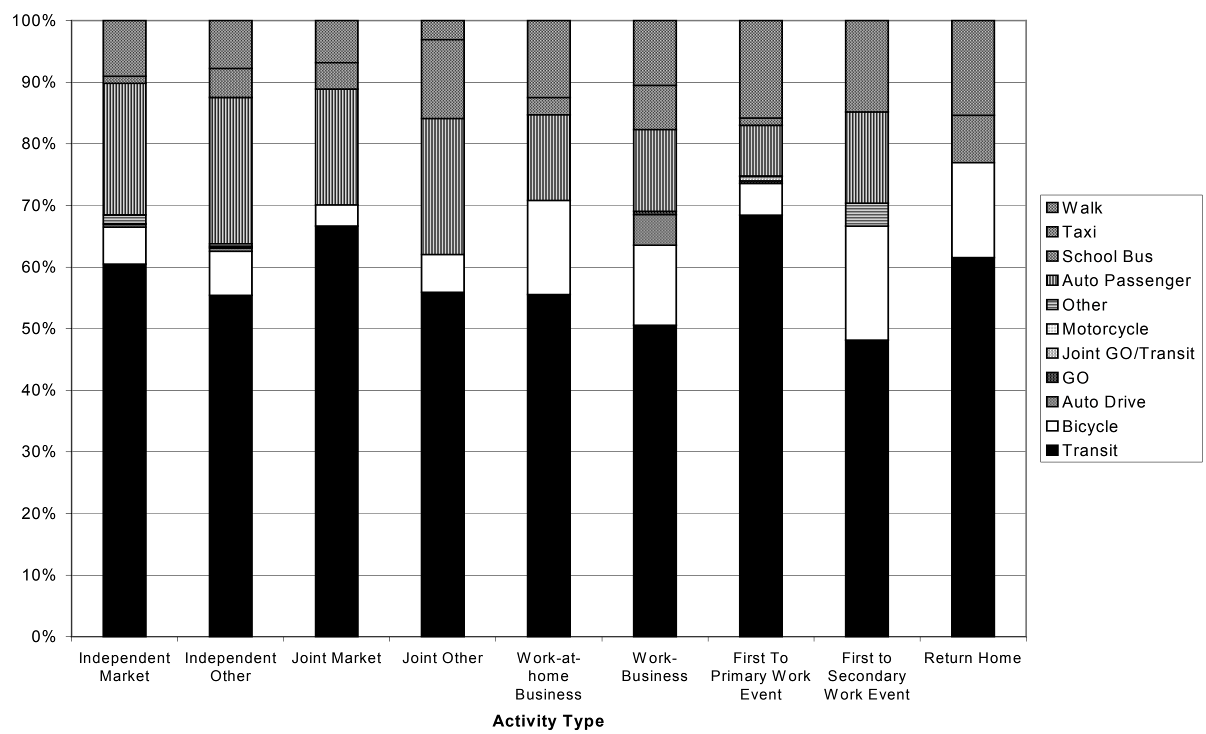



**Figure 3–37 Percentage of Activity Type by Mode Used for Trip (Driver's Licence, No Vehicle)**

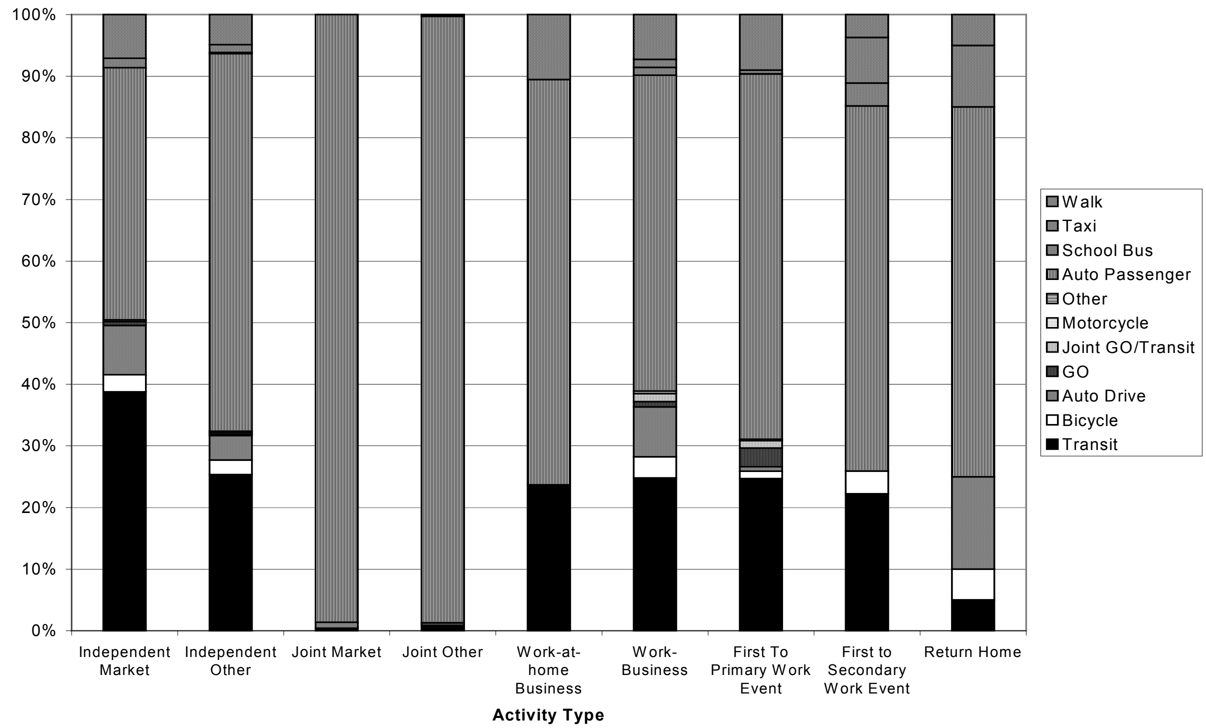



**Figure 3–38 Percentage of Activity Type by Mode Used for Trip (Driver's Licence, Vehicle Not Available)**

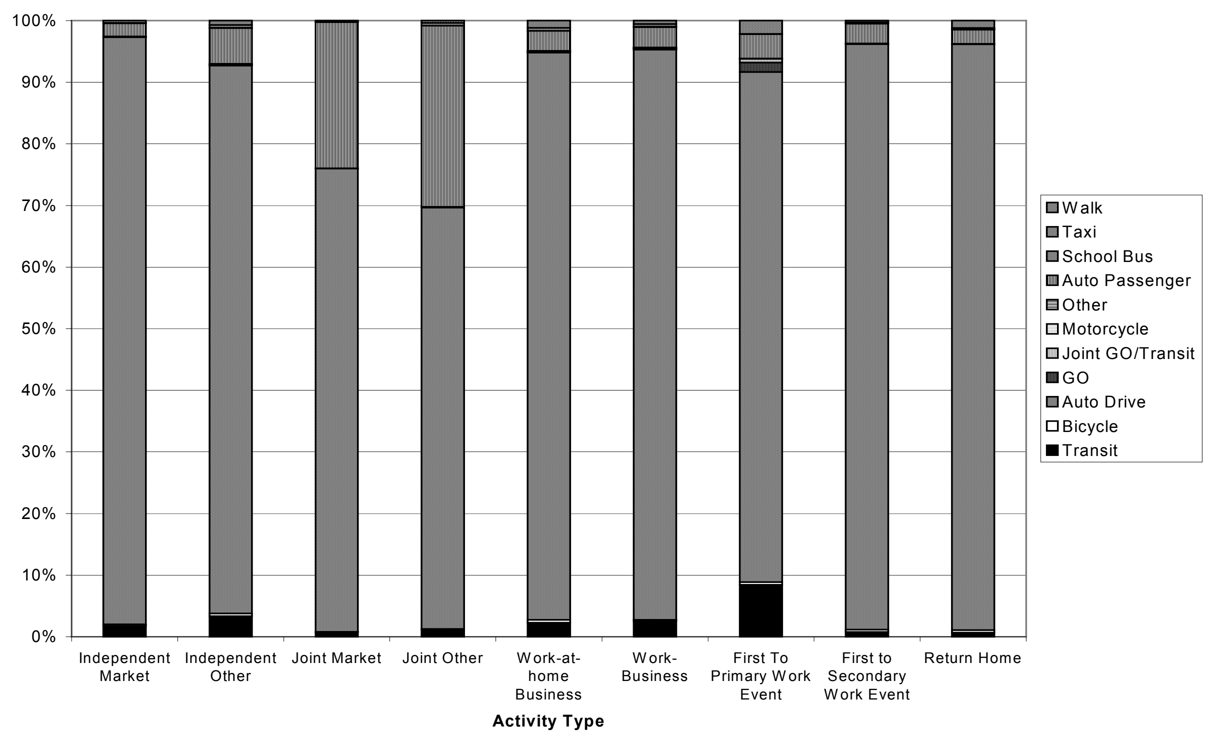



**Figure 3–39 Percentage of Activity Type by Mode Used for Trip (Driver's Licence, Vehicle Available)**

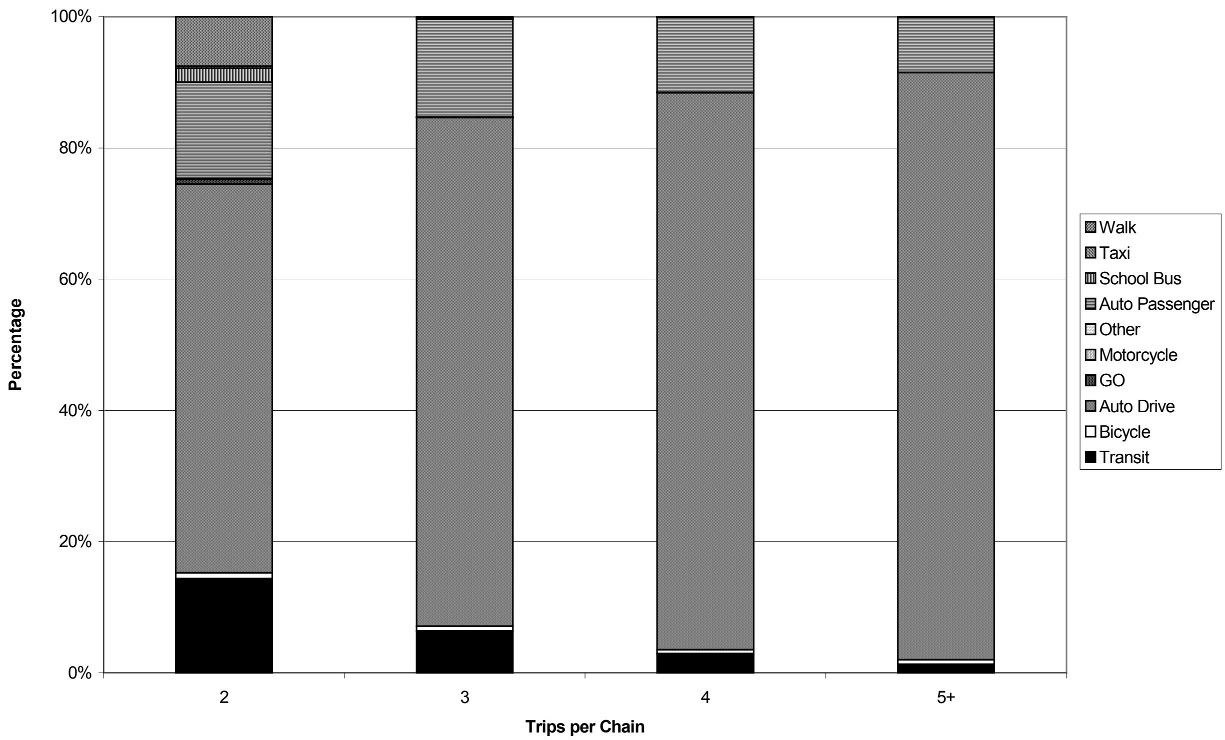



**Figure 3–40 Percentage of Trips per Chain by Mode (Single Mode Chains Only)**

a very large number of combinations of location and sequence, quantification of such is no trivial task and would have no meaning descriptively. As such, the distance travelled between activity locations as well as distance from home and work are used as a measure of location choice. First, however, trip chaining behaviour is explored further. Figure 3–40 shows the mode used as a function of the number of trips in the chain.[21] It can be seen that the auto–driver mode dominates chains consisting of two trips, but that other modes such as walk, school bus, GO rail and public transit are still represented. More specifically, the modes associated with travelling to and from work or school (a two–trip chain) are represented. As the chain becomes increasingly complex (in terms of the number of stops), however, the auto drive mode share becomes increasingly dominant. This is obviously a result of the convenience of the auto mode and the increased difficulty associated with completing multi–stop chains.

Figures 3–41 through 3–44 show the distribution of trip distance from the previous activity location and activity location distance from home for *market* and *other* activity types by

[21] Only the 94.2% of chains identified in the previous section as using a single mode are used here.

mode, respectively. First, note the tighter distributions associated with trip distances from the previous activity as opposed to distances from home. These distances are generally shorter and are a result of the efficiency gained from trip chaining. The distributions for each mode are also as expected. The walk and bicycle modes produce the tightest distributions and involve the shortest distances, followed by taxi and then by the transit and auto modes with more varied and longer distances. Once again, the crucial tie between location and mode choice is illustrated. Figures 3–45 and 3–46 display activity location distance from home and from work by activity type, respectively.[22] These figures show the fixed and constrained nature of work–related location choices in relation to *market* and *other* activity locations. That is, in both cases, *market* and *other* activity locations can be chosen so as to require less travel (although the differences are not as large in the 'distance from work' analysis) whereas work–related locations are generally predetermined and therefore have a fixed distance. There are, of course, exceptions to this generalisation. Work–related activities do not have to occur at a 'work' location; for example, a meeting might be held at a nearby restaurant. Conversely, a trip to the doctor's office (an *other* activity) is fixed in location, which, however, has likely been chosen long before the appointment to be reasonably close to home.

## 3.7 SERVING DEPENDANTS

*Daycare* and *facilitate passenger* activities involve serving the needs of another person; that is, the person engaging in these activities does not gain utility in the conventional way but rather 'gives' it to someone else. If that person is a household member, household utility is gained and the person is said to be 'serving a dependant'. This obviously has major implications for a household–level model.[23] Unfortunately, the data is very limited in terms of what can be determined about this behaviour. This is because:

- No trip records exist for persons under 11 years of age. As the majority of dependants are commonly in this age group, the majority of information is missing; and

[22] Only activity types that require location choice modelling are shown. Work and school locations are predetermined.
[23] This, in particular, has implications for mode choice, in that being an auto–passenger with a household member is actually using a different mode than being an auto–passenger with a non–household member (i.e. 'carpooling').

- These activities are not limited to serving persons of the same household. Thus, whether the person is serving a household member, a non–household member or a combination of the two cannot be determined.

Thus, it is only really possible to accurately determine *facilitate passenger* activities that serve another adult household member, which represents just a portion of this behaviour. Despite this, some information can be extracted about serving dependants. Figure 3–47 shows *facilitate passenger* trips to the school of a child within a household having children aged 4 to 10, by trip start time and mode.[24] Figure 3–48 also shows *daycare* trips by trip start time and mode.[25] Both figures show the morning and afternoon peaks as well as the very dominant auto–driver mode that are expected to be associated with both activities.

[24] Many assumptions are made here, including the activity of the person served being *school* and, more important with respect to serving dependants, that they are a household member.
[25] Here, the activity and age of the person served is known, in that they are going to a daycare and are 4 years old or less. It is still possible, however, that the person is not a household member.

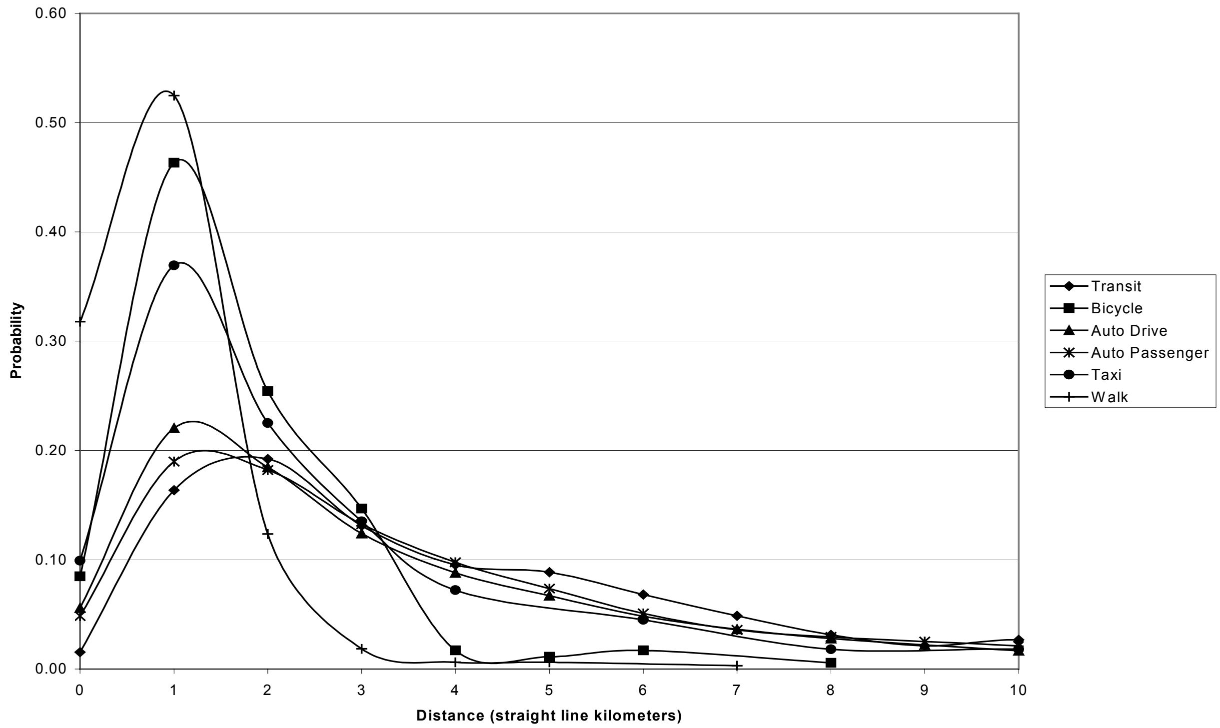



**Figure 3-41 PDF of Trip Distance by Mode (*Market* Activities)**

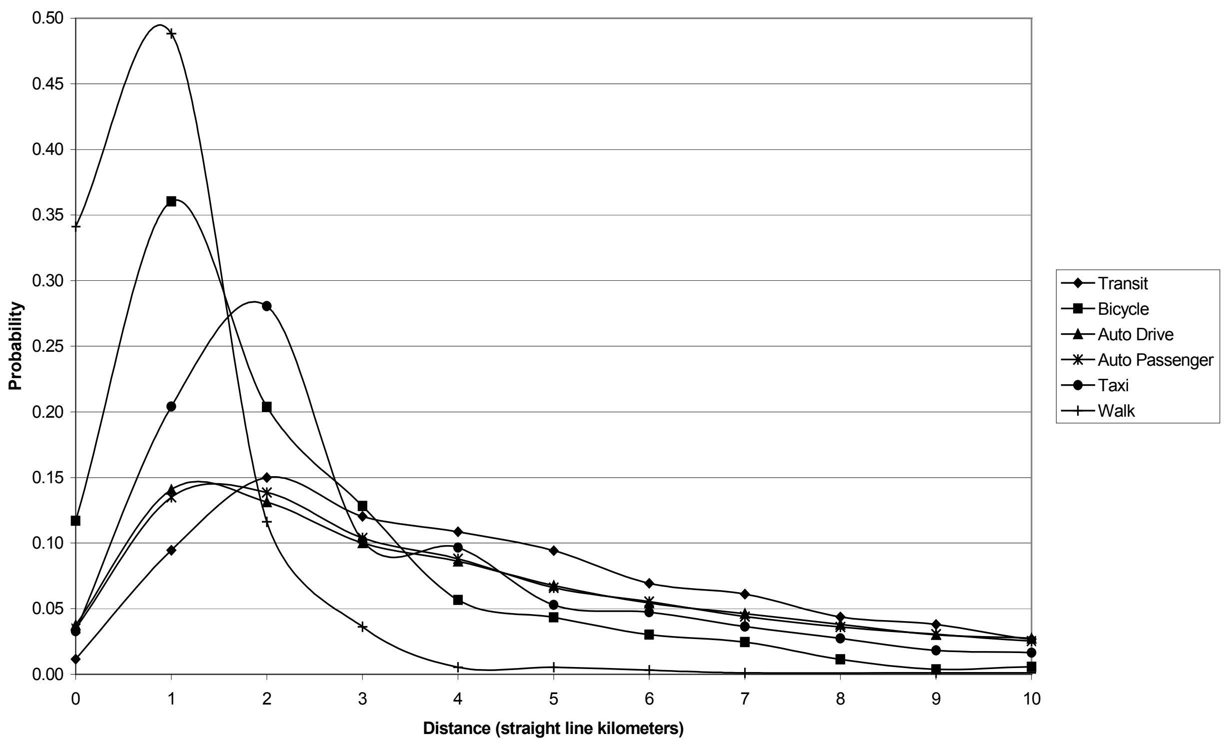



**Figure 3–42 PDF of Trip Distance by Mode (*Other* Activities)**

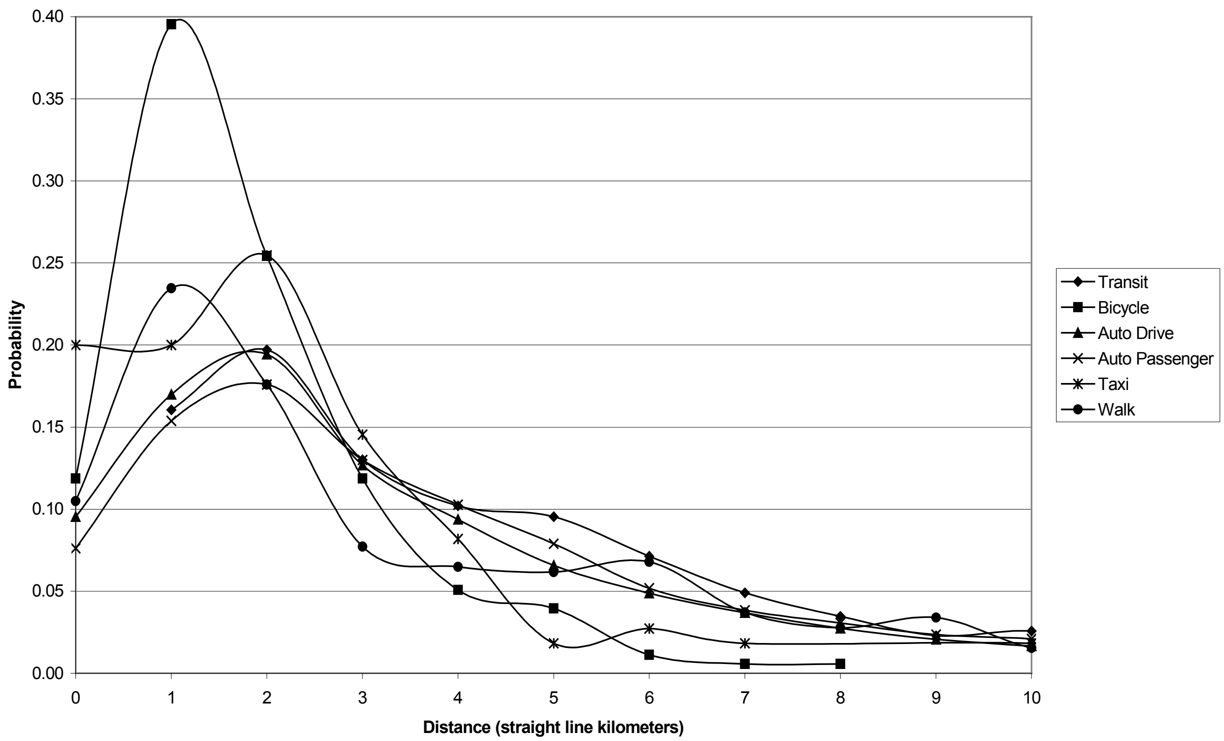



**Figure 3–43 PDF of Activity Location Distance from Home by Mode (*Market* Activities)**

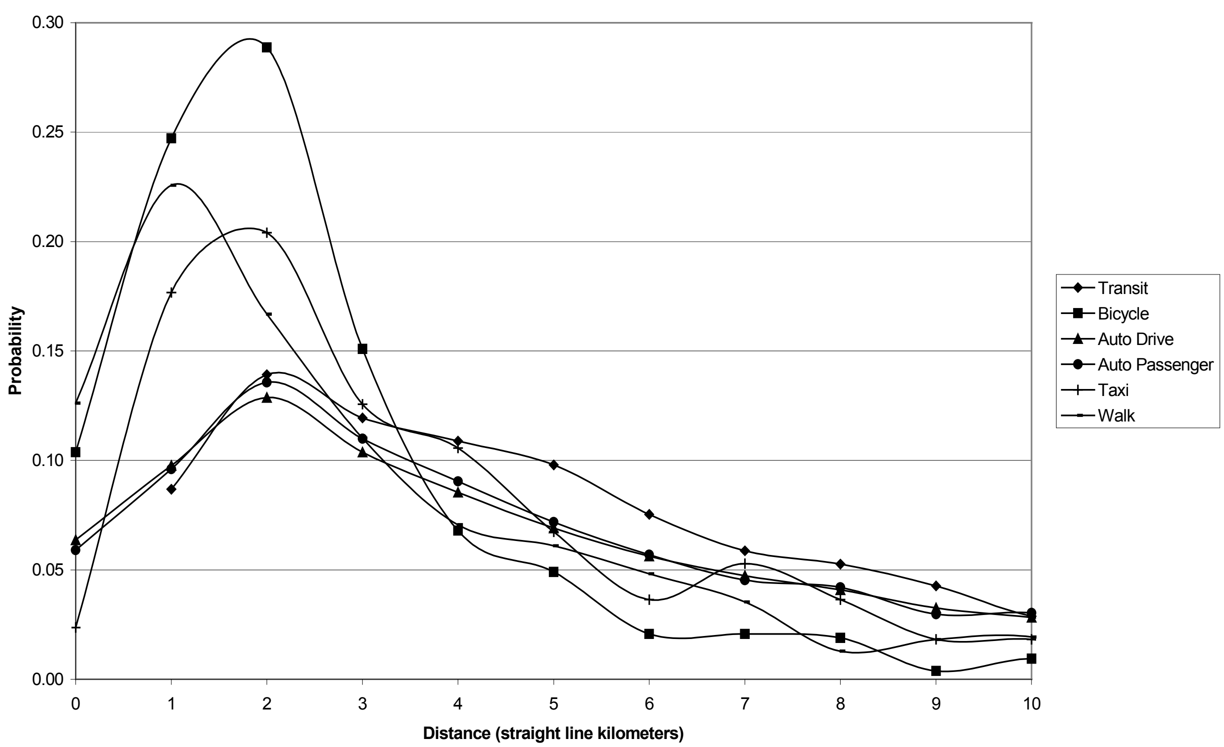



**Figure 3–44 PDF of Activity Location Distance from Home by Mode (*Other* Activities)**

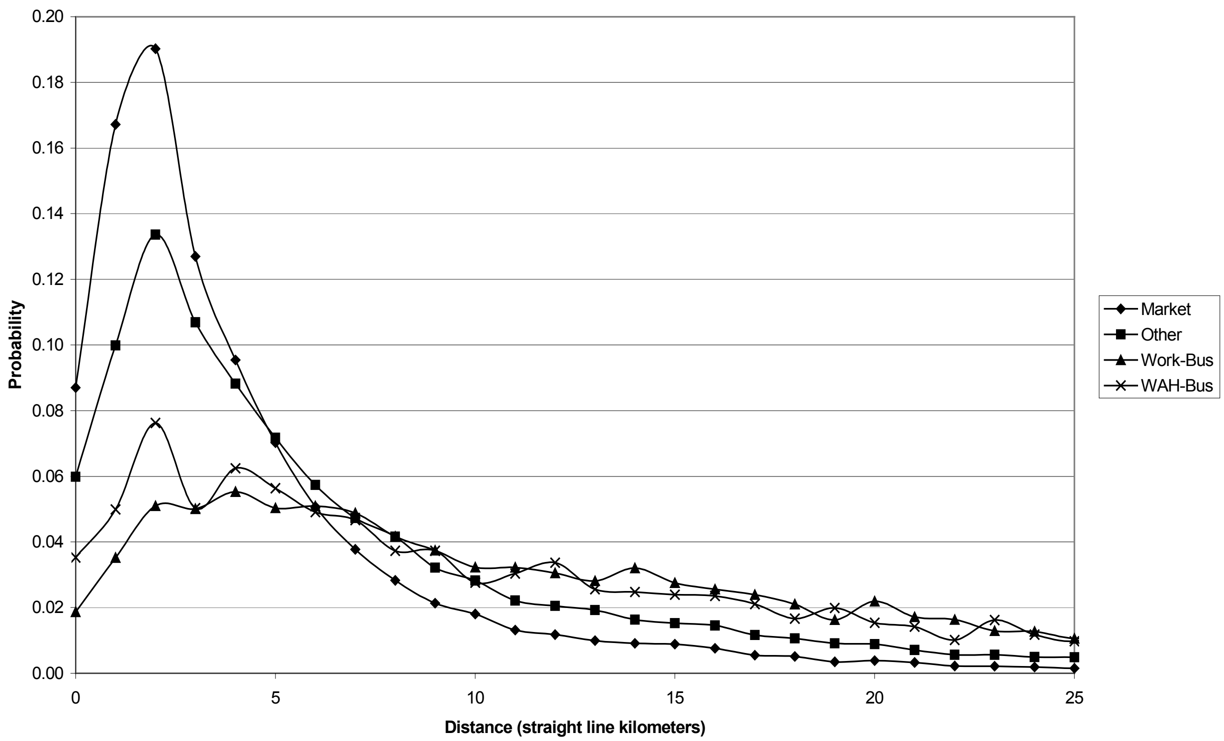



**Figure 3–45 PDF of Activity Location Distance from Home by Activity Type**

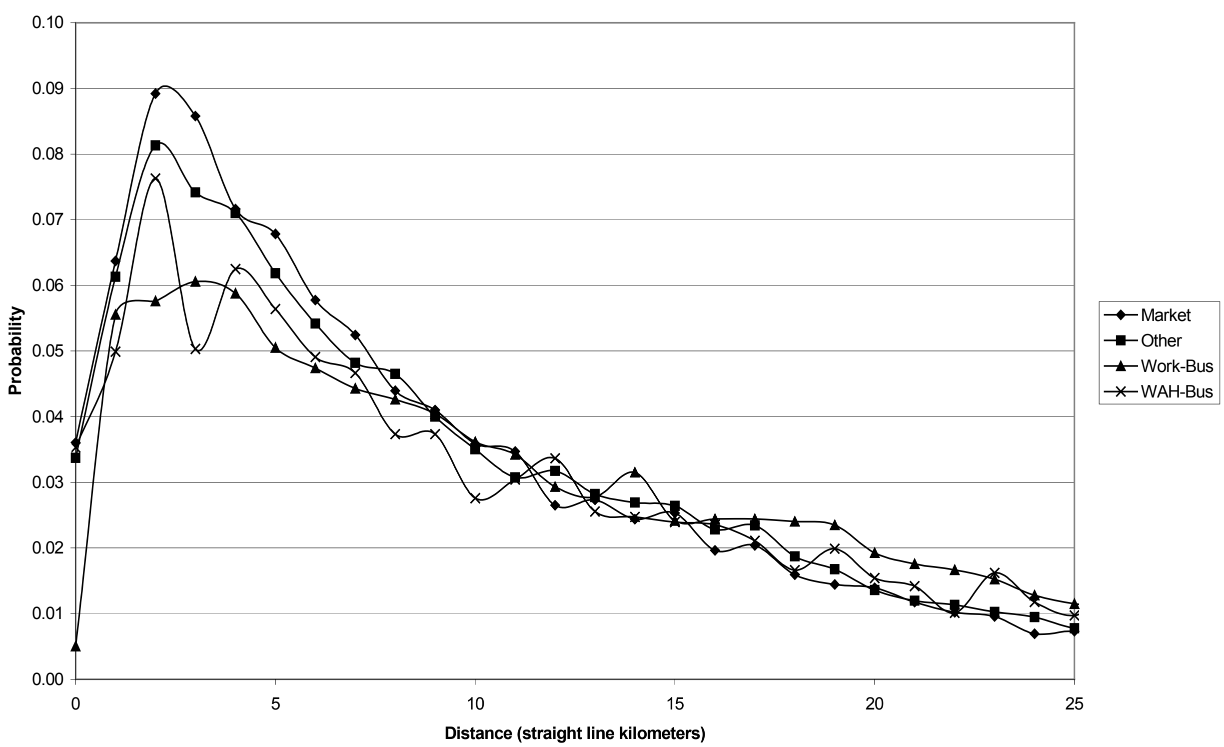



**Figure 3–46 PDF of Activity Location Distance from Work by Activity Type**

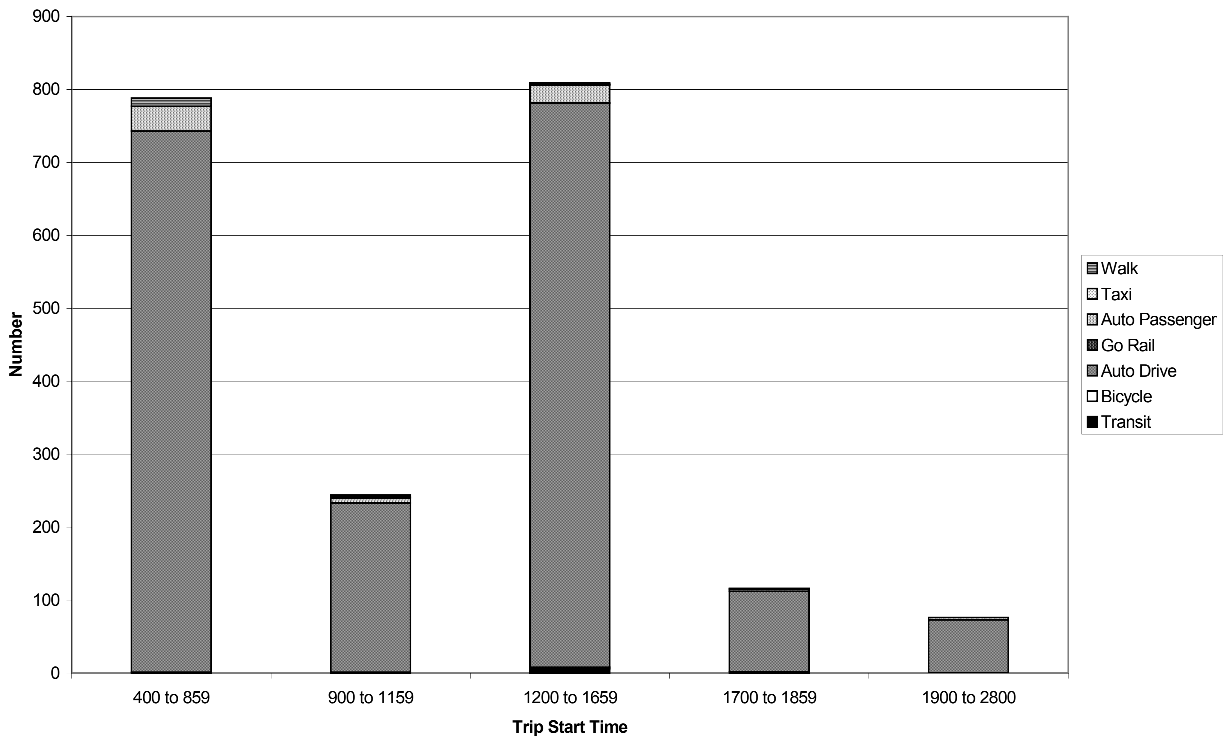



**Figure 3–47 Number of School-based Facilitate Passenger Trips by Trip Start Time & Mode for Households With Children Aged 4 to 10**

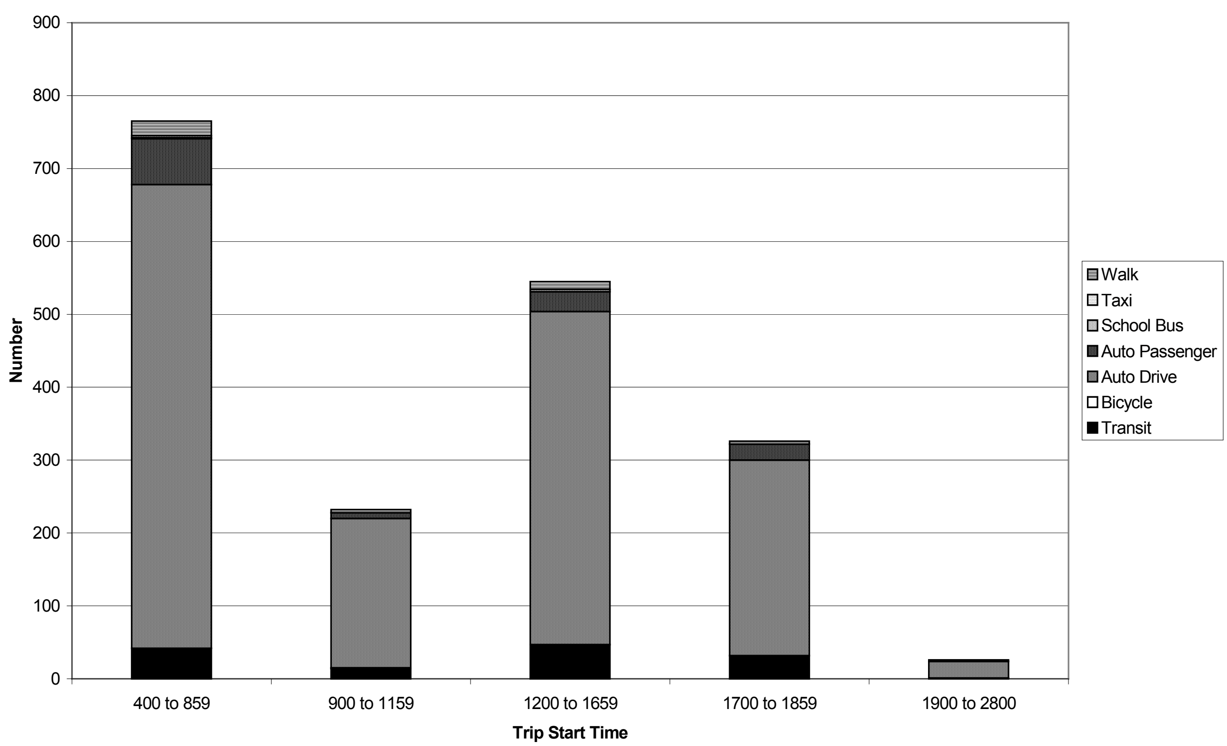



**Figure 3–48 Number of School-based Daycare Trips by Trip Start Time & Mode**

# 4 MODEL DEVELOPMENT

As mentioned in the introduction, the intention is to develop a model that estimates household travel as a direct result of participation in out–of–home activities. This requires that we simulate the **generation** and **scheduling** of various activities as they might occur within a household during a typical weekday. To do this, we combine the observed travel/activity behaviour of a population sample with a set of rules that we assume are employed by rational **decision–makers** to simplify the very complex task of generating various activities (as they are required in order to fulfil certain fundamental needs) and organising them within a constrained amount of time in a realistic and efficient manner (or, perhaps, as efficient as we feel is realistically achievable by a rational decision–maker, given the complexity of the problem). The task we are simulating is complex because it requires the decision–maker to process a large amount of information that is rarely complete and that, even worse, is dynamic in that it alters with time. The combinations and permutations involved in scheduling even a small number of activities are vast. Hence, we rely on the notion that people are problem simplifiers, who quickly reduce the number of choices down to a reasonable number that can be dealt with.

As modellers, we are faced with the challenge of taking this process and applying the mechanics of mathematics to it. We do not, of course, expect to capture perfectly the intricacies of human behaviour. Rather, we hope to capture the basics of this decision–making process so that we can reasonably predict future behaviour. There are, indeed, certain aspects of the process that, for now at least, will elude us due mainly to the limitations of the data, which, again, are geared toward trip–making and not activity participation. These aspects include:

- The dynamic nature of the process. Schedules are constantly being modified and re–modified as available information changes. This model is static, not in the sense that the analysis time period is instantaneous (it encompasses a 24–hour period in five minute intervals rather than only, say, morning peak period work trips), but in that we only observe the final decision, or action, and can therefore only model the same. That is, we do not observe the dynamic process leading up to execution of the final activity, the timeframe of which can span minutes, days, months or years (see next).

- The different scheduling mechanisms involved in fixed and unfixed activities, the degree of fixation being defined as the flexibility in choice of location, start time and possibly duration. The location, start time and mode choices of fixed activities are likely determined by a person or household far in advance of actual participation in the activity, as opposed to the relatively instantaneous choices made for unfixed activities. For example, one does not wake up each day and decide where and when one is going to go to work, as this has likely (in most cases) been determined far in advance. On the other hand, choosing the location of a hardware store is much more flexible, as is when to visit it. In addition, fixed choices will also likely be highly co–dependent with such factors as automobile holdings.[1] Again, for the unfixed activities of *market* and some components of *other*, the data only reveal the final decision in terms of the chosen location and the position in the sequence of activities. For the fixed activities of *work* and *school*, the data also include the usual location, if it exists. The determinants of the choice of usual location are presumably contained in the socio–demographic data of the household and the land–use characteristics of the surrounding urban area. As such, the distinction between the two types of scheduling mechanisms is lost in the data.
- The full utility of the activity. Our data inform us of the 'what, when and where' of the activity, in terms of the general purpose of the activity (e.g. work, shopping, *et cetera*), the mode, the start time and the location. Why the activity episode was performed and to who's benefit (personal or household), however, are unobserved. The 'who' is especially important in determining joint participation in activities, a phenomenon that adds even more complexity to the scheduling process. To illustrate, consider the subtle difference between shopping for eyeglasses and shopping for groceries. The former provides personal utility and the latter household utility, yet both can be performed by the individual or jointly with another household member. The 'why', on the other hand, concerns the real purpose of the activity episode. Individual activity episodes are likely rarely carried out as autonomous entities, but rather as part of a larger overall goal or objective such as, for example, grocery shopping for a dinner party. We will refer to interrelated collections of activity episodes as **projects** (after Axhausen, 1998). Although

[1] Note the difference between automobile holdings, referring to the number of vehicles possessed by a household, and automobile availability, meaning the ability to access a household vehicle at any point in time.

we have no way of tying together related activity episodes in this way with the available data, we will use this concept as a fundamental organising principle in the model, to be elaborated further in the future.

The model will loosely follow the basic logic and sequence of the traditional four–stage modelling system, in that it first generates trips, then determines the travel mode and finally assigns flows based on an equilibrium assignment. There are, however, many important departures from this system that make this model unique and groundbreaking. These include:

- Full future populations are synthesised with corresponding socio–demographic attributes;
- Trips are generated at the disaggregate person/household level as a *result* of participation in multiple out–of–home activity episodes, for a 24–hour period. This microsimulation approach means that all of the origins and destinations, as well as the sequence of these, for each person in a future population are explicitly determined. This enables the trip distribution stage of the traditional system to be skipped, as flows from zone to zone are not balanced, but explicitly determined;
- Activity location choices are modelled as is required by the activity–based, microsimulation framework;
- Given that the sequence of trips (i.e. the 'trip chain') is known, mode choice can be modelled at the chain level rather than at the trip level. This is much more realistic, especially when considering the auto–driver mode as the car must, in general, be used to complete a chain once it has been used to start one; and
- Joint participation in travel/activity episodes between members of a household is modelled.

Three components are critical to determining daily travel behaviour: activity **schedule**, location choice and mode choice. One can argue that persons attempt to balance these components so as to maximise their utility as well as that of the household to which they belong. This is no trivial task as activity schedule, location choice and mode choice are all part of an interrelated, simultaneous decision triangle (see Figure 4–1). That is, there is no linear sequence to the decision process as each component is dependent upon the others. The fixing of and/or omission of one or two of these components may lead to a sequence that is seemingly unique and

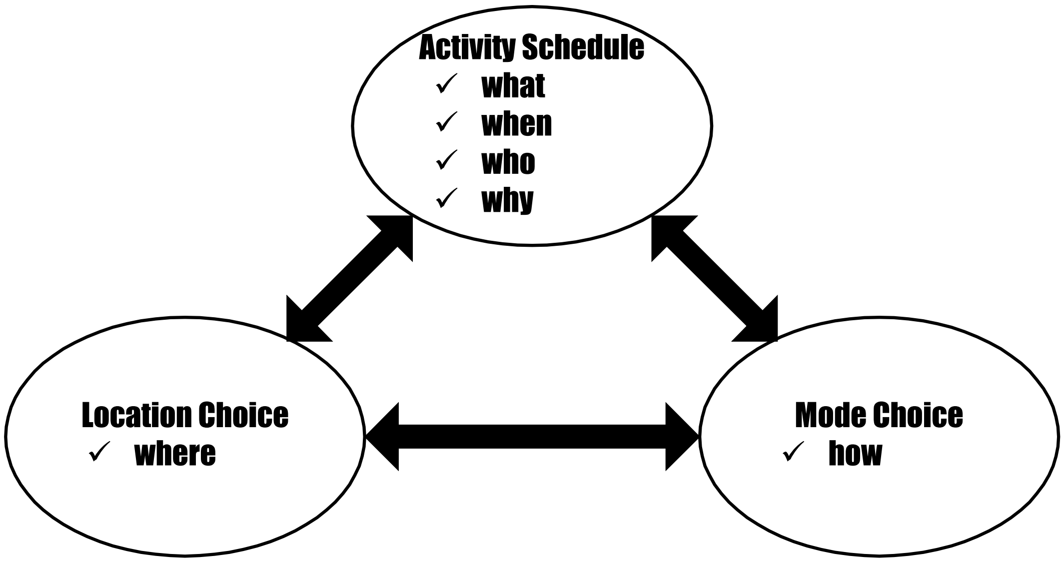



**Figure 4–1 Travel Behaviour Decision Triangle**

optimal based on the attributes of a particular activity episode, as is evidenced by "I need to see my mother who is located close to home, therefore I will walk". Here, the activity is fixed in location and the mode is chosen without consideration of the activity schedule. To assume a sequence based solely on the attributes of an activity episode, however, one would have to make the poor assumption that it is independent of the effects of the component attributes of other activity episodes within the schedule. This assumption is contradicted by "I will pass my mother's house on the way home from the supermarket, to which I will drive, therefore I will drive to my mother's." Here, inclusion of the activity schedule into the decision process has resulted in a mode choice that is different from the one reached by the otherwise reasonably assumed decision sequence as a result of the simultaneous consideration of activity schedule, location and mode. As such, realistically, the decision process involves simultaneous consideration of the component attributes, whether fixed or unfixed, of all activity episodes scheduled for the analysis time period. Ultimately, this process is also continuous throughout the time period, responding to the dynamic nature of the activity schedule.

The realisation of the absence of a linear sequence implies immense complexity in the process, which, of course, extends to modelling it. As such, with the initial model, we do propose a sequence to this process with the idealistic future goal of removing it. The sequence is as follows. First, activity episodes are generated and locations are determined for those activities that require choice. Scheduling of these episodes is then completed assuming auto travel times. The mode choice model is applied after scheduling of all activity and travel episodes is complete, as their sequence is important to this model. Thus, the schedule is already imperfect because of the simultaneous relationship between the activity schedule and mode choice decisions, in that one cannot determine an accurate schedule without knowing of the mode(s) (as differing modes have differing travel times) and *vice–versa*. The ultimate objective of the model is, however, to forecast the correct number, origin/destination, start time and mode of all trips with accuracy that is sufficient for our purposes. Accurate specification of the activity schedule is, therefore, secondary in importance in that it is merely a step towards meeting this objective.

The following describes the general concepts and methods used to generate the potential activity episodes for a household as well as documents the development and estimation of the location and mode choice models.

## 4.1 Generation of Activity Episodes

The objective of episode generation is to produce candidate activity episodes for insertion into an **agenda** and, subsequently, a schedule. A complete agenda contains a set of internally feasible episodes with known start time, duration and location attributes. These attributes will be conditioned by scheduling/time constraints as well as by the attributes of the decision–maker. The former are imposed by the scheduling algorithm and the latter are imposed explicitly through socio–demographic categorisations determined to be significant during the descriptive analysis. The model will use observed frequency distributions of activity duration, start time and frequency to randomly generate potential activity episodes for each activity type for a given population. The creation of realistic synthesised person and household schedules cannot, however, be achieved by arbitrarily placing these episodes into a schedule as they exist in the data set. Examination of the data reveals an underlying logic to the scheduling of these episodes that is lost as a result of separating them out from each other by activity type. It quickly becomes apparent that this logic has to be recaptured and applied in order to simulate behaviour with any accuracy. To do so, we define some concepts to be used to make it possible to produce a realistic and workable synthetic schedule (see Miller, 2001b).

The project represents a powerful concept that can be used to return this logic to an activity–based model. The *work*, *school*, *market*, *other* and *facilitate passenger/daycare* activity types defined in the data (as destination purposes) potentially represent what Miller (2001a) refers to as 'primary projects'. That is, they are the highest level of activity "in which household members engage in attempting to meet their fundamental needs." A complete list of primary projects should be representative of all of these fundamental needs through reasonably broadly defined categories. Projects themselves are composed of an agenda of one or more activity episodes of varying type and, conversely, all activity episodes evolve from projects. While this conceptual difference is important, it is lost at the level of detail used in this model, which is, unfortunately, much too low to truly make use of this concept. For this, the activity categories in the data would have to be much more detailed and each activity episode linked to others within the household based on their contribution to one of the defined primary projects. These could then be aggregated to the primary project level and used for project generation (with an agenda) and then activity episode generation. While we are considering this concept within the model, it is only in the most basic way. As such, the terms activity and project are used synonymously hereinafter, a project being comprised of activity episodes of the same type. The activity types to be implemented in the model are: *work*, *school*, *independent market*, *joint market*, *independent other*, *joint other*, *return–home* and *in–home*, which is the default activity. A flow chart showing each within the household activity structure is presented in Figure 4–2.

In attempting to develop a method to generate and schedule episodes for each activity type, one realises that there are fundamental differences between them, thereby requiring slightly different approaches in their definition and application. *Work* is by far the most complicated. To illustrate, consider the ambiguities involved in conceptualising the workday. Does it consist of one 8–hour episode that includes lunch, or is it one 3–hour (morning), one 1–hour (lunch) and one 4–hour (afternoon) collection of episodes? After all, people rarely work for 8 hours straight (i.e. without lunch), but they likely conceptualise and schedule the whole time as *work* time.[2] This issue first became apparent in a practical sense while performing the descriptive analysis.

[2] Or, again, as belonging to the *work* project. Although eating lunch is not a work activity (it may even be a market activity if one eats at a restaurant), it is done, at least in part, so that one has the energy to continue the *work* activity. Both activities evolve from the fundamental need for sustenance.

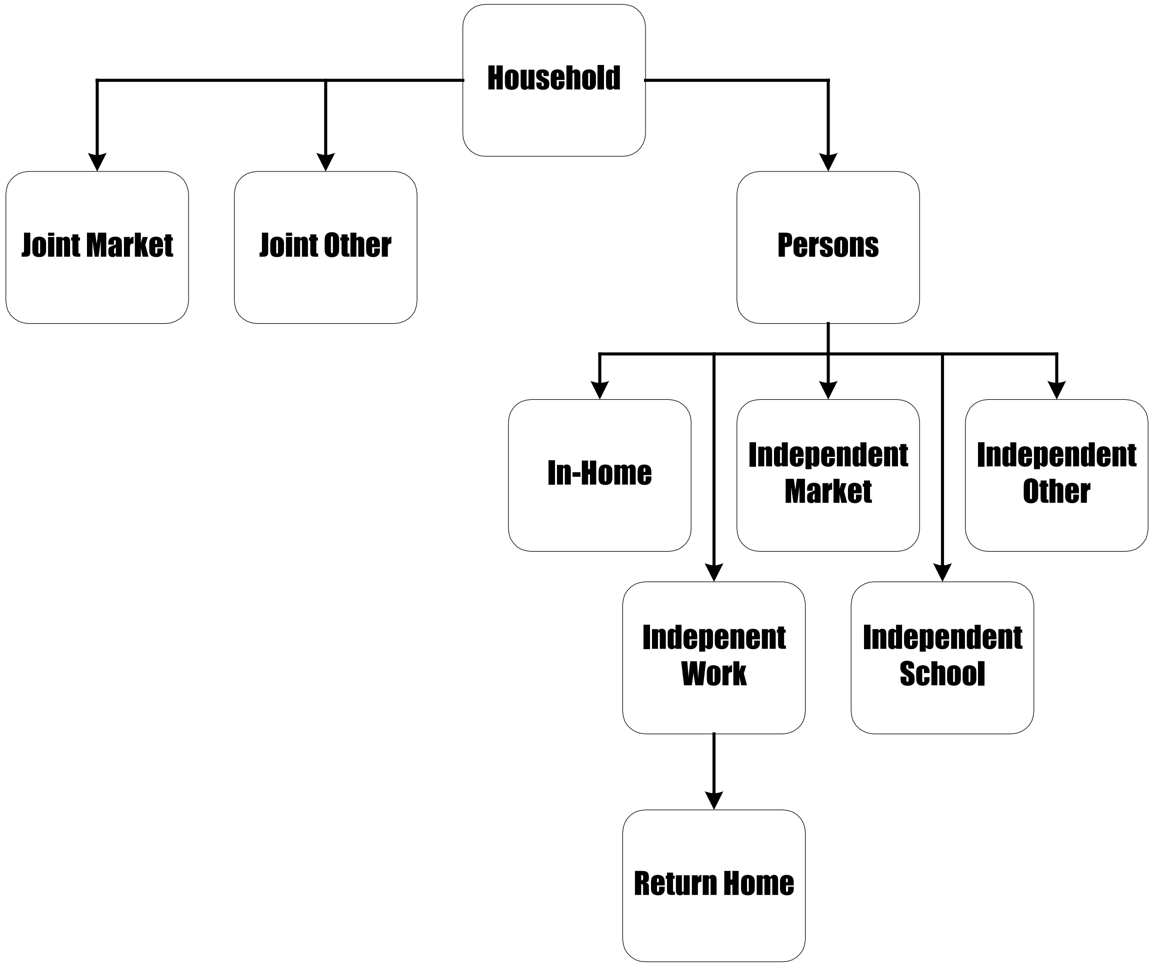



**Figure 4–2 Household Activity Structure**

The majority of workers report single 7.5 hour (the mean) *work* episodes. Some, however, report multiple *work* episodes that are the result of either (i) a change in location of the *work* activity; (ii) participation in another type of non–*work* activity between two *work* episodes(including returning home); or (iii) a combination of the two. The sum of the duration of these episodes, however, represents a full workday. The result is that the variance of the *work* episode duration is considerably greater and the distribution skewed to the left as compared to that of the total daily activity duration due to the presence of multiple short episodes. This is shown in Figure 4–3. The same is true of the distributions of the other activity types but to a lesser degree, as their total daily activity duration is generally considerably less than that of *work*.[3] Therefore, in

[3] Although school activity duration is also relatively long, this activity was generally not divided between several episodes and therefore did not pose the same problem, as might be expected.

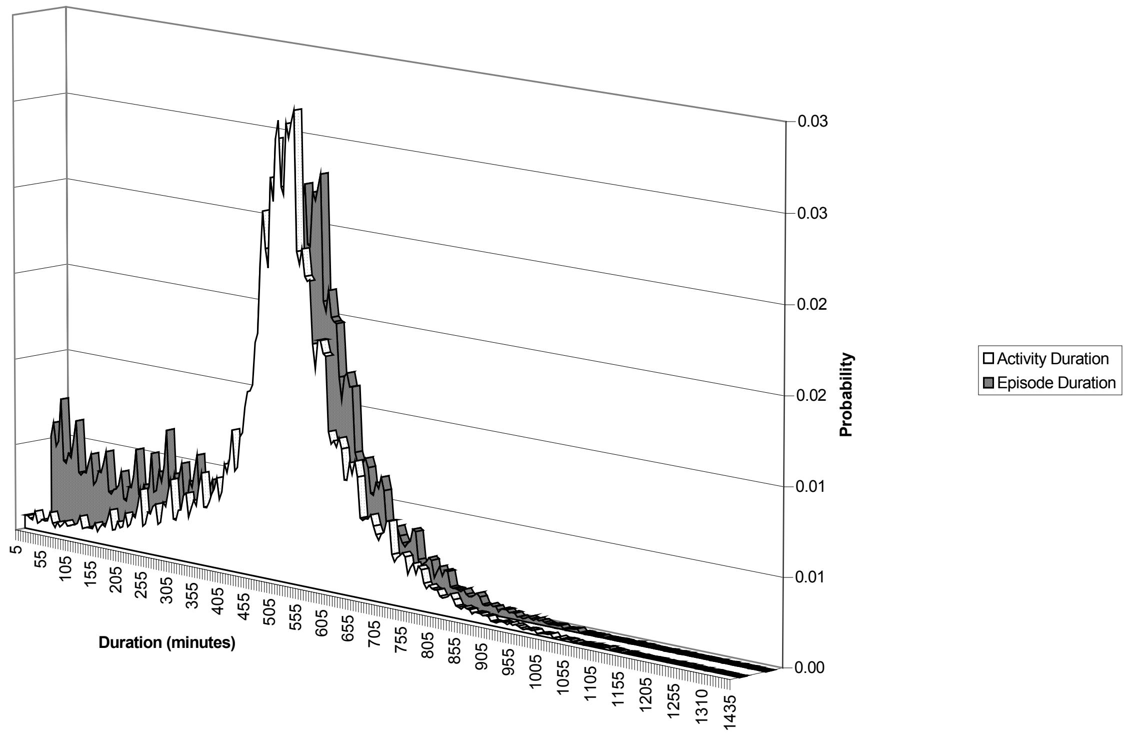



**Figure 4–3 Comparison of *Work* Episode & Activity Duration PDFs**

considering how to generate and schedule the *work* activity, it is apparent that randomly inserting work episodes into the person schedule would be unrealistic and messy, much more so than for other activity types. In addition, it is a commonly accepted assumption that work is a 'cornerstone' activity that has priority and around which all other activities are scheduled. This is because the start time and duration of this activity are generally not negotiable for the decision–maker (i.e. the activity schedule component has fixed attributes), it involves contracts with other parties and provides a relatively high utility to the person/household. Thus, it is especially important that this activity be scheduled in a realistic manner. Therefore, to deal with these complexities, we define the *primary work* **event** as the sequence of *work* episodes beginning at the first home–to–work episode of the day. In generating the distributions upon which these events are based, the start time is that of the first trip, the duration is the sum of all work episodes within and the frequency is either 0 or 1 (i.e. the worker goes to work or does not). In this way, the workday is now a conceptual block of time that can be manipulated, rather than a collection of episodes with distinct properties. In order to handle cases where a worker goes to a second job or works in the evening, we define the *secondary work* event as the sequence of *work* episodes beginning from the second home–to–work episode that starts after 3 p.m., conditional on the *primary work* event having occurred and having the same functional definition as the *primary work* event. As the *secondary work* event is a rare occurrence (in the data) and is generally shorter in duration, we assume, for simplicity, that it consists of a single episode. However, for the reasons mentioned above, the *primary work* event may consist of multiple episodes; that is, through participation in non–*work* activities or *work* activities at other locations. To deal with the latter, we define a *work–business* episode as work at a location that is not the usual place of work. Functionally, this episode has a start time, duration and frequency as determined by observed distributions. The episodes of the *work* activity, therefore, are generated by first determining the attributes of the *primary work* event (including whether there is one) and then, possibly, adding *work–business* episodes, the duration of which are contained within that of the *primary work* event. The base location of the *primary work* event is the usual place of work; thus, it is similar in nature to the *in–home* activity in that it acts as the default location during this period.

Two special types of worker exist which require special consideration, those being people who have no usual place of work and people who work at home. For the first case there is, of course, no base location for a *primary work* event to occur at. Conceptually, then, all of these *work* episodes are *work–business* episodes and therefore are treated as such; that is, they are generated and scheduled as regular activity episodes, without enforcing the organising principles of the *primary work* event. For the case of working at home, the concept of the *primary work* event is also meaningless as it is indistinguishable from the *in–home* project, which is the usual place of work. As such, we define the *work–at–home business* episode as a another unique type of *work* episode, having start time, duration and frequency characteristics as observed in the data. These will also be treated on an episode by episode basis as are the remainder of the non–*work* activities.

Non–*work* activities do not present the same difficulties associated with the *work* activity as, for example, one generally does not interrupt grocery shopping to see a doctor, only to return to grocery shopping later. A possible exception to this assumption is the *school* activity. However, as multiple episodes and location change are much less likely within this activity, it is deemed reasonable to assume a single episode for which the duration distributions consist of the total observed event times. *Market* and *other* activity episodes can be performed either individually or jointly with another household member. As it was shown that individual and joint episodes have differing start time, duration and frequency attributes, separate distributions are used to generate episodes for each type. Furthermore, as it makes no sense to generate joint trips at the person level, *joint market* and *joint other* episodes are generated for the household. The number of adults engaging in the joint activity is then also randomly determined, again based on observed distributions. A *return–home* episode is a special case of the *in–home* episode and is only eligible to exist within the *primary work* event for the purposes of representing workers going home for lunch. Finally, the *in–home* activity is, as mentioned before, the default activity, which initially has a duration of 24 hours and is modified as needed during scheduling.

The complete set of distributions used for event and episode generation consists of 180 categories for each of the duration, start time and frequency attributes. Duration and start time distributions use 5–minute intervals. Each of these are defined and presented in Appendix E. A summary of the explanatory variables used for categorisation of each activity type defined for the model is presented in Table 4–1. Again, each variable has been chosen based on its statistical and logical impact on the activity type. The same variables are used to determine all three

| Activity Type | Explanatory Variables Used for Categorisation |
|---|---|
| Primary Work | Age, Occupation, Employment Status |
| Secondary Work | Occupation, Employment Status |
| Work–Business | Age, Occupation, Employment Status |
| Work–at–Home Business | Age, Occupation, Employment Status |
| School | Age, Student Status |
| Return Home | Occupation, Employment Status |
| Independent Other | Age, Possesion of a Driver's Licence, Employment Status |
| Joint Other | Presence of Children, Number of Adults in Household, Auto Accessibility |
| Independent Market | Age, Employment Status, Gender |
| Joint Market | Presence of Children, Number of Adults in Household, Auto Accessibility |

**Table 4–1 Summary of Variables Used for Episode Attribute Distribution Categorisation**

attributes; that is, the same variables influence the duration, start time and frequency of episodes or each activity type. This may be a poor assumption as, for example, employment status likely has a much larger impact on *work* duration than it does on frequency. The use of multiple variables, however, increases the likelihood that each attribute is properly characterised. As can be seen, age, occupation and employment/student status are predominant. The joint activity attributes are determined by household level variables that were found to have impact by Gliebe and Koppelman (2001).

As mentioned previously, potential activity episodes are generated randomly. The first step in this process involves extracting the cumulative distribution function from the probability distribution function of each distribution. Generation of potential activity episodes involves applying a random number (ranging from 0 to 1) to the specific cumulative distribution of the simulated person (with predefined socio–demographic characteristics) involved. Given that the gap of the distribution is larger for more probable intervals, the overall activity episode attributes of a population will approach that of the observed distributions with increasing application of the process. These potential activity episodes are then forwarded for scheduling. For technical details of the generation and scheduling procedures, see Roorda *et al.* (2002). Note that this process assumes that the duration, start time and frequency of an episode are determined independently. This assumption is obviously a very poor one and will be discussed later.

## 4.2 LOCATION CHOICE

As discussed previously, location choice can be fixed or unfixed in nature.[4] Miller (2001c) differentiates the two as exogenous and endogenous to the scheduling process, respectively. This has important implications as the location choices of exogenous activities are determined independently of the schedule and those of endogenous activities dependent on the schedule, the latter making the scheduling process much more complicated. In the model, location choices will be required for all episodes generated for *individual* and *joint market, individual* and *joint other, work–business* (usual and no usual place of work) and *work–at–home business* activities. *Work–business*, *work–at–home business* and some components of *other* fall into the exogenous

[4] Ultimately, no activity is truly fixed as, for example, people continuously change work and school location throughout their lives. It is a matter of scale and, from the perspective of modelling the workday, these activities can certainly be treated as fixed in location.

category, whereas *market* and the remaining components of *other* are endogenous to the scheduling process. Our data are insufficient to distinguish between the two as, for example, we cannot readily make a distinction between whether a person has shopped at a fixed location (such as a speciality store) or whether the location was chosen to be convenient at the time. As such, the location choice for all of these activity types will be determined exogenously to the schedule. Following is a discussion of the development and specification of the location choice model.

The model is an entropy model that estimates the probability of choosing an activity location, where the choice set is the entire set of 1996 GTA traffic zones. Separate models are developed for each activity type for which location choices are required. Issues related to specification of the models include:

- As the location choice is determined exogenously to scheduling of the activity, modal level–of–service effects are not relevant, including auto availability to the person;
- Spatial effects will, of course, be important. Thus, distance will act as an impedance to location choice;
- Location attractors can consist of population and employment totals as well as dummy variables denoting various levels of 'activity centres';
- It is likely that socio–demographic characteristics, such as occupation or auto accessibility, influence location choice and should be considered; and
- The commonly Gamma–shaped trip distance distribution may require use of 'short distance' dummy variables to improve model goodness–of–fit.

Given the above, the model takes on the typical entropy/logit formulation. Thus, the probability that a person who lives in zone i chooses location zone j is defined as (see Miller, 2001c):

$$P_{j|i} = \frac{\exp\left(\sum_k \delta_{jk}\left[\alpha_k + \beta_k \log\left(E_j\right) + \phi_k \log\left(P_j\right) + \gamma_k d_{ij}\right]\right)}{\sum_{j'} \exp\left(\sum_k \delta_{j'k}\left[\alpha_k + \beta_k \log\left(E_{j'}\right) + \phi_k \log\left(P_{j'}\right) + \gamma_k d_{ij'}\right]\right)} \qquad [4\text{–}1]$$

where:

$\delta_{jk}$ = 1, if zone j belongs to zone activity category k; 0 otherwise

$E_j$ = employment in zone j
$P_j$ = population in zone j
$d_{ij}$ = distance from zone i to zone j
$\alpha_k, \beta_k, \phi_k, \gamma_k$ = parameters to be estimated

Data inputs into the model include observed location choices as well as a matrix of dummy variables defining the activity centres. Observed location choices are for each home zone except for *work–business* (usual place of work), which utilises the workplace location zone as the reference. In other words, it is hypothesised that these location choices are affected by the location of the workplace, whereas the others are affected by the location of the home. Activity centre matrices are defined to indicate attraction for *work*, *market* and *other* activities. There are three levels to which a zone can be assigned (exclusively). All zones in planning district one (i.e. the city core) are defined as level one. Level three is the default level. Level two zones are defined as follows[5]:

- For *work*, any zone with greater than or equal to 3,000 employed per square kilometre;
- For *market*, any zone with greater than or equal to 100,000 square feet of shopping mall floor space; and
- For *other*, any zone with greater than or equal to 50 retail stores per square kilometre.[6]

Thematic maps of the GTA showing the level assigned to each zone based on these criteria are presented in Appendix F.

Model output files for the location choice parameter estimates as well as explanatory variable definitions are contained in Appendix G. Common among all models are distance, employment and short trip variables. The short trip dummy variables significantly improve model performance. All models are also conditioned by a maximum distance allowable, which varies by model, to minimise the effect of outliers. The *work–business* (usual place of work) model is separated out by occupation type, as this proves beneficial and is operational due to the large number of observations. The activity centre specification only proves beneficial for the

[5] Data for shopping mall and retail stores are provided by Murtaza Haider.
[6] The definition of the second level market and other activity centres is somewhat arbitrary in that, for example, movie theatres (an *other* activity) exist in shopping malls and market activities can obviously be located in retail stores. Retail stores, however, encompassed a wide range of services as well as goods that would best describe the personal business and entertainment sub–components of the *other* activity type.

*market* location choice model, which, coincidentally, also has the best goodness–of–fit statistics including a rho–square value of 47%. Otherwise, rho–square values range from 32% for *other* activity locations down to about 20% for the *work* activity models. These values are not surprising as decision–makers probably have the least control over *work–business* locations relative to that of all other activities. While one has control over the location of, for example, his/her usual place of work, *work–business* activities likely involve clients who have made location choices in isolation of the location of the decision–maker, relinquishing his/her control and essentially removing the 'choice'. This is not in contradiction to Central Place Theory as the distance parameters of these models are still negative and significant. It does imply, however, that the location choice process for this activity has much less to do with the needs or characteristics (in terms of home or usual place of work location) of the decision–maker than for the other activities and is therefore much more difficult to determine accurately. It is also interesting to note that model performance was best for the unfixed/endogenous *market* activity as compared to the fixed/exogenous *work* activities, with the mixed *other* model lying in between, which reinforces the previous observation.

Selected trip distance probability distributions for observed input and predicted location choices are presented in Figures 4–4 and 4–5 as another measure of model performance. The fits to the observed distributions indicate good prediction results.

## 4.3 Travel Episodes, Mode Choice & Trip Chains

Conventional mode choice models are trip–based; that is, they predict a chosen mode as a function of its utility for a single origin–destination pair, independently of trips made before and after it. This method is appropriate when considering the objectives of these models, but has obvious limitations in considering the factors that really affect how a person chooses a mode for a trip that is usually contained within a trip chain. It is also a fair approach when the modes involved are all non–drive (e.g. transit, walk, auto–passenger). It is definitely a poor one, however, when also taking into account the drive mode (i.e. auto–driver). This is because, generally, once a person has decided to take the household automobile, they are in possession of it until they return home and, conversely, once they have decided not to take the car, it is not

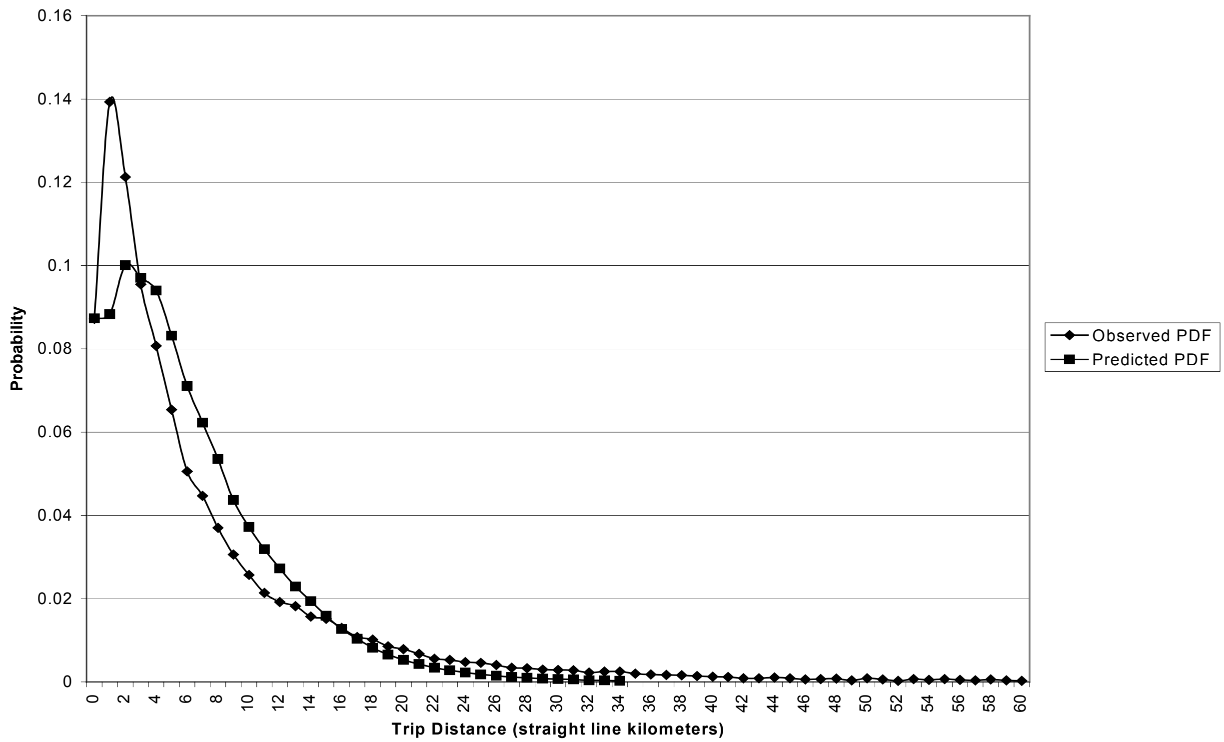



**Figure 4–4 Comparison of Observed and Predicted Trip Distances (*Other* Activity Locations)**

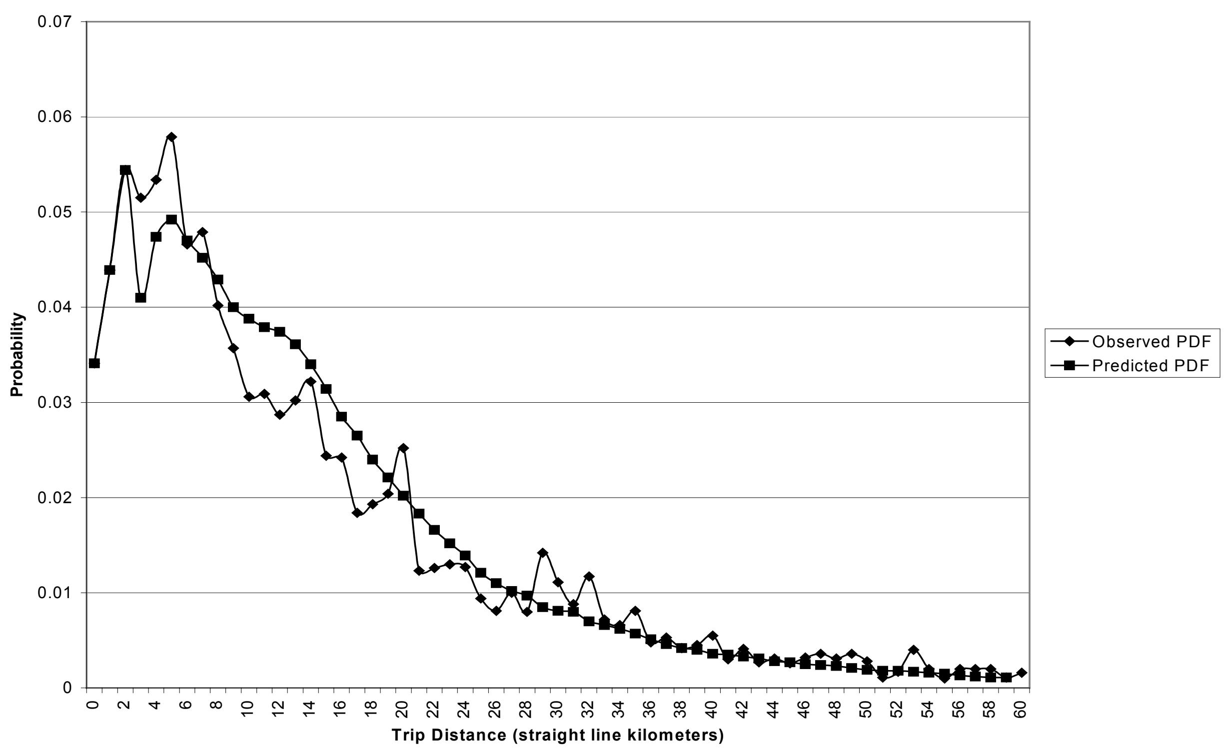



**Figure 4–5 Comparison of Observed and Predicted Trip Distances (*Work–Business,* No Usual Place of Work Activity Locations)**

available to them until they return home.[7] Thus, the mode choices for all of the trips within this home–based chain are dependent on this decision. One of the critical advances of this model is that it determines mode choice at the chain level, rather than the trip level. This is possible because the model generates full–day activity schedules for all persons from which trip chains evolve. Following is a discussion of the development, structure and parameter estimation for the mode choice model.

The generation and scheduling of activity episodes in varying space and time also implies the generation and scheduling of **travel episodes** (i.e. trips), which, indeed, represent the ultimate objective of the model. For each generated activity episode inserted into a schedule, two travel episodes are generated, as either:

- The new episode is inserted into an *in–home* or *primary work* event episode, in which two new travel episodes are required; or
- The new episode is inserted between two regular episodes, in which the one travel episode between the prior and posterior episodes is replaced by two travel episodes, occurring from the prior episode to the new episode and then from the new episode to the posterior episode.

As with activity episodes, travel episodes have a start time and duration. The latter is synonymous with travel time, which is dependent on the mode of travel. The frequency of these, however, is defined directly by the number of activity episodes rather than randomly.

Given a complete person schedule, home–based chains (again, the sequence of all activity and associated travel episodes contained within the departure from home and the subsequent return to home) develop naturally. In practice, any point in space can be used to define a trip chain *anchor point*, the point that defines the beginning and end of a chain. As an extension to this, sub–chains can develop within chains using, for example, the usual place of work as the anchor point. Note that these sub–chain anchor points only have practical implications for chains involving the auto–driver mode as they represent places where, should the car be left behind (i.e. not chosen as the mode for the sub–chain), the person must return in order to retrieve

[7] There are, of course, exceptions to this, such as when two household members meet outside of the home to exchange possession of the automobile.

it. Otherwise, their existence is of no interest. We use home as the most 'natural' definition of an anchor point and the point in space that is most critical to mode choice. It is critical as it is the place to where the car must be returned to at the end of the chain and also where it is remade available to other members of the household, an important consideration in a household–level model. As such, the terms 'chain' and 'home–based chain' are synonymous. We consider the usual place of work and rail parking lots as natural sub–chain anchor points.

Using the preceding concepts, chains are defined as being *drive* or *non–drive*, differentiated by whether the auto–driver mode is used for the first trip of the chain (in which case the person is assumed to be in possession of the car for the remainder of the chain) or not. For each chain, a mode choice decision structure exists as is shown for the chain in Figure 4–6. In choosing the *non–drive* alternative, the mode choice for each trip in the chain is made on a trip–by–trip basis in the conventional way, with all remaining, relevant, *non–drive* modes available to choose from. Relevant modes are defined as a function of the primary purpose of the chain, which is to be discussed later. Choice of the *drive* alternative implies that all trips in the chain will be completed using the auto–driver mode. It is conditional upon the decision–maker possessing a driver's licence and the existence of at least one household vehicle. As an extension, if a sub–chain exists within a *drive* chain with the usual place of work as the anchor point, the *drive*/*non–drive* decision is repeated in the same fashion based on the trips

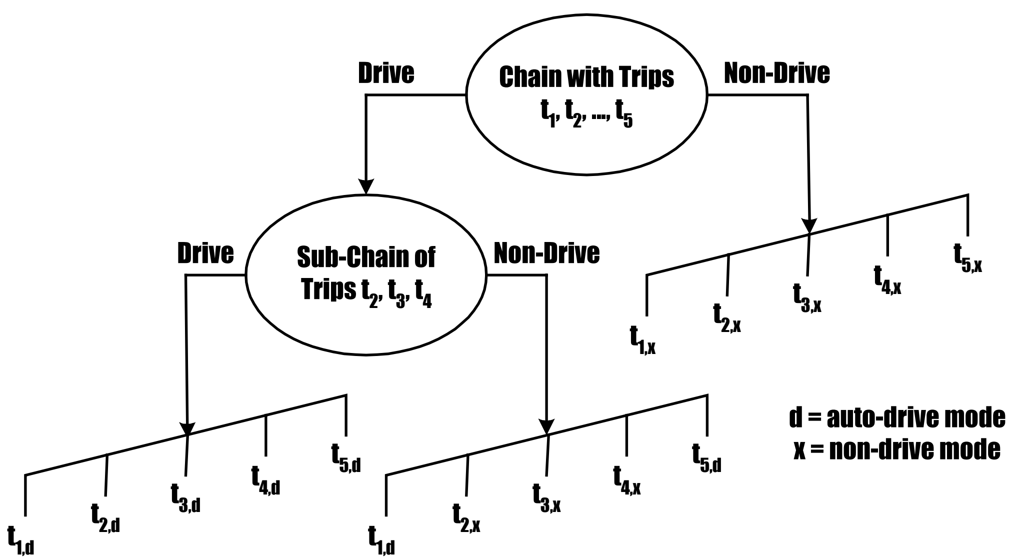



**Figure 4–6 Chain-Level Mode Choice Decision Structure**

contained within the sub–chain. Sub–chains anchored at rail parking lots would require no such decision; that is, the *non–drive* alternative is implied and the *non–drive* mode choices determined accordingly. These sub–chain types, however, are not considered in this model version.

A logit utility maximisation formulation is used. Given a chain, the utility for each trip is defined as (see Miller, 2001d):

$$\begin{aligned} U_{mtn} &= V_{mtn} + \varepsilon_{mtn} \\ V_{mtn} &= \beta'_{tn} X_{mtn} \end{aligned} \qquad [4\text{–}2]$$

where:

$U_{mtn}$ = random (actual) utility of using mode m for trip t for person n
$V_{mtn}$ = systematic (observable) utility of using mode m for trip t for person n
$X_{mtn}$ = vector of explanatory variables for mode m, trip t and person n
$\beta_t$ = vector of parameters

Thus, for *non–drive* chains, a chosen mode, *m**, is determined for each trip in the chain as the one with the maximum utility for the decision–maker, as:

$$U_{m^*tn} \geq U_{mtn} \forall m^*; m, m^* \in M_{tn} \qquad [4\text{–}3]$$

where $M_{tn}$ is the set of *non–drive* modes available for trip t and person n. This represents the right side decision of Figure 4–6. The utility of the chain is then calculated simply as the sum of the utilities of the constituent trips, as:

$$U_{cn} = \sum_{t \in T_c} U_{m^*tn} + \eta_{cn} \qquad [4\text{–}4]$$

where:

$T_c$ = set of trips comprising chain c
$U_{cn}$ = random (actual) utility of chain c for person n
$\eta_{cn}$ = random disturbance term

This represents the input to the upper level decision. Practically, the sum of trip utilities for the *non–drive* alternative (i.e. the utility of performing the *non–drive* trip chain) is actually represented by a 'sum of logsums' term, as:

$$\begin{aligned} E[U_{cn}] &= \sum_t E\left[\underset{m}{Max}(U_{mtn})\right] \\ &= \sum_t Logsum_t \end{aligned} \qquad [4\text{–}5]$$

$U_{cn}$ can be determined for the *drive* and *non–drive* alternatives. The explanatory variables used in the *drive* alternative are expressed at the chain level, as it can only be executed utilising the single auto–driver mode (e.g. cost is represented as the sum of the costs of completing each trip in the chain by car). Thus, the upper level decision rule is to choose the *drive* alternative if:

$$U_{cn}^{drive} \geq U_{cn}^{non-drive} \qquad [4\text{–}6]$$

Otherwise, choose the *non–drive* alternative.

It is assumed that utility parameters vary by the *primary purpose* of the chain. The primary purpose is intended to indicate the *raison d'être* for the chain as the utility of the *drive* chain is likely different when, for example, a *work* activity episode exists in the chain as compared to when one does not. Primary purpose is categorised hierarchically into five groups, defined as:

- Level 1, if at least one *work* episode at the usual place occurs for an out–of–home worker;
- Level 2, if at least one *work–business* episode occurs for an out–of–home worker;
- Level 3, if at least one *school* episode occurs for a student 19 years or older;
- Level 4, if at least one *school* episode occurs for a student 18 years or younger; and
- Level 5, if none of the above instances occur.

Chains are classified as one of these contingent on not having the attribute of an upper level. As such, parameter estimation for ten mode choice models are required in total. The parameters for primary purpose levels 1, 2 and 3 have been estimated for a previous model. Therefore, only parameters for levels 4 and 5 require estimation.

The *non–drive* trip and *drive*/*non–drive* chain mode choice models for primary purpose levels 4 and 5 are presented in Appendix H along with definitions of the explanatory variables used. The lower level *non–drive* trip models are similar in structure to conventional models. They contain a standard mix of level–of–service and socio–economic variables such as in–

vehicle travel time, travel cost, number of household drivers, age, home location, proximity to transit, and trip distance. There are, however, two key departures from convention. First, the modal alternatives do not include the auto–driver mode. This means that they do not have to compete against this dominant mode (at least not yet) and are on a more level playing field. The second difference is the presence of chain–level variables. Here, a variable representing the number of trips in the chain, *ntrips*, is used to measure the complexity of the chain, which is not to be confused with the distance travelled. It is hypothesised that the existence of multiple stops lends utility to the auto–passenger alternative given the convenience associated with it. The variable *njoint*, the number of joint trips in the chain, is also meant to capture this convenience as it is likely that these trips will be made by car. Given these specifications, both of the non–drive trip models exhibit common goodness–of–fit statistics as well as significant, correctly signed parameter estimates.

The upper level *drive*/*non–drive* models are unconventional. As mentioned, the *non–drive* chain utility is represented solely by a 'sum of logsums' term passed up from the *non–drive* trip models. It is obviously expected to have a positively–signed parameter. *Drive* chain utility is represented by chain–level level–of–service as well as socio–demographic variables, including in–vehicle travel time, the number of household vehicles, total chain distance, age, home location, parking cost, and occupation. All parameters are correctly signed and most are significant. Rho–square values are also good. It should be noted that sub–chains would not be incorporated for chain primary purpose levels 4 and 5 given the nature of the chain and/or the attributes of the decision–maker involved; for example, students 18 years of age and under are not observed performing a sub–chain anchored at a rail parking lot in the data and thus would not be offered this alternative.

As another measure of performance, Table 4–2 compares observed inputs and predicted values for travel mode. As can be seen, the results are quite good. Although some of the modes have larger errors than others, these are for less significant modes; that is, the predictions of number of total auto, transit and walk trips, as well as overall mode split, are very good. Figure 4–7 graphically depicts the mode split for each of four time periods. Again, the results are encouraging.

| Mode | Predicted | Observed | Difference | Percent Difference | Predicted Split | Observed Split | Difference |
|---|---|---|---|---|---|---|---|
| Auto-passenger | 1,325,203 | 1,298,573 | 26,630 | 2.1 | 15.7 | 15.4 | 0.3 |
| Transit all way | 1,009,499 | 1,092,584 | -83,085 | -7.6 | 12.0 | 12.9 | -1.0 |
| Subway, auto access | 86,484 | 0 | 86,484 | 0.0 | 1.0 | 0.0 | 1.0 |
| GO-Rail, transit access | 16,250 | 22,574 | -6,324 | -28.0 | 0.2 | 0.3 | -0.1 |
| GO-Rail, auto access | 43,058 | 45,027 | -1,969 | -4.4 | 0.5 | 0.5 | 0.0 |
| Walk all way | 524,598 | 522,851 | 1,747 | 0.3 | 6.2 | 6.2 | 0.0 |
| Auto-driver | 5,220,658 | 5,201,913 | 18,746 | 0.4 | 61.9 | 61.6 | 0.2 |
| Bicycle | 43,399 | 70,153 | -26,754 | -38.1 | 0.5 | 0.8 | -0.3 |
| School-Bus | 143,888 | 135,868 | 8,020 | 5.9 | 1.7 | 1.6 | 0.1 |
| Taxi | 26,425 | 38,643 | -12,218 | -31.6 | 0.3 | 0.5 | -0.1 |
| Other | 0 | 11,313 | -11,313 | -100.0 | 0.0 | 0.1 | -0.1 |

**Table 4–2 Comparison of Predicted and Observed (Input) Total Trips and Mode Split**

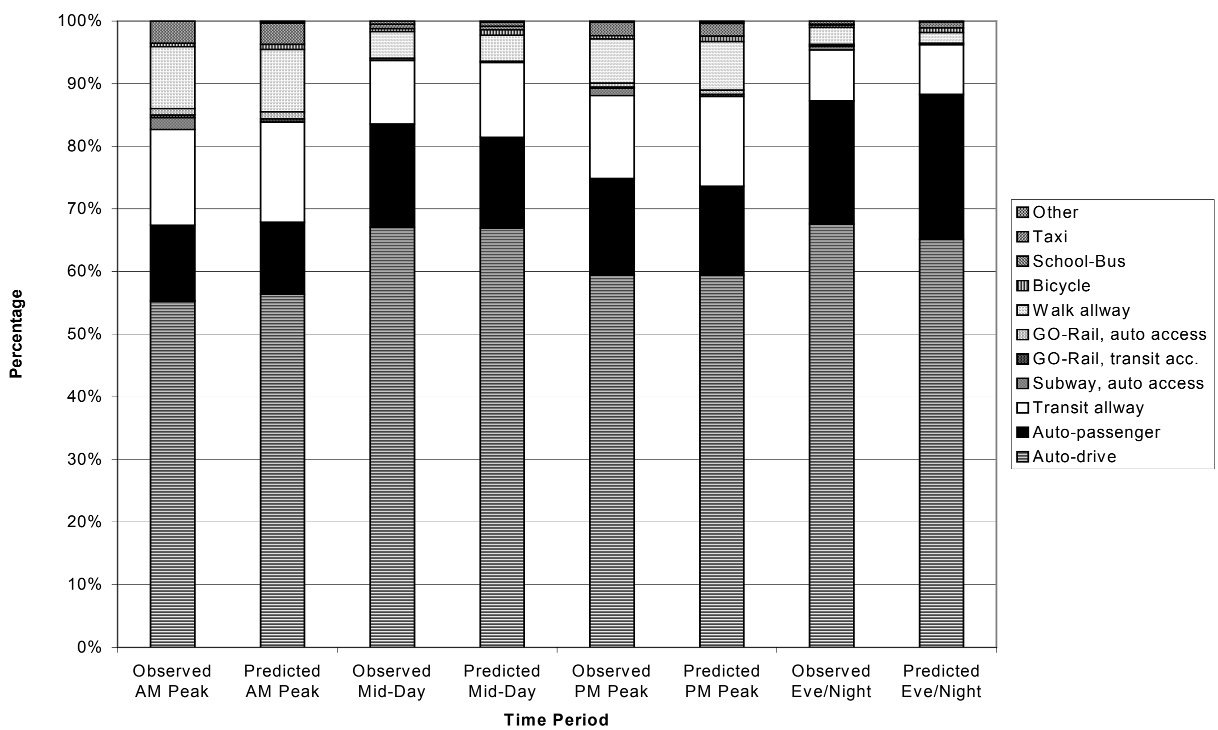



**Figure 4–7 Observed and Predicted Mode Split by Time Period**

## 4.4 MODELLING PROCESS

Given the above, the model can predict the activity and travel behaviour of a population. This process is executed in the following four major steps.

1. For future year forecasting, the base population is 'grown' to the future year population. This is accomplished by adjusting the weight factor assigned to each person in the 1996 TTS so that the sum across all persons in a given zone is that of the forecasted future population. In this way, socio–demographic characteristics of the household are preserved. The exception to this is place of work, which is subsequently altered to meet future year employment forecast totals.
2. This population is used by the activity/travel simulator to determine full–day schedules for each person within each household using 5–minute time steps. Each out–of–home activity episode also has location and start time attributes. The final start times are non–conflicting; that is, any conflicts between generated activity episodes are resolved in a rule–based manner. This resolution may alter the initial duration attributes of an episode or, in the extreme, its feasibility. Locations for *work–business*, *market* and *other* activity episodes are determined using the location choice model. Again, *in–home* is the default activity. Trips times are assumed to be those of the auto mode.
3. Trips to and from these activities are extracted as home–based trip chains and sent to the mode choice model.
4. Trips made by the auto and transit modes are assembled into origin–destination matrices using morning peak, mid–day, afternoon peak and evening/night time periods. These are assigned to the road and transit networks using standard techniques.

Thus, the final output of this model, in meeting with the overall objective, is auto and transit volumes by transportation network link by time period. The output is that of a traditional four–stage model run for each time period. The process leading up to this, however, is radically different.

# 5 MODEL RESULTS & VALIDATION

The model was tested using the 1996 TTS base year population and transportation network. Therefore, the travel and activity behaviour of the simulated population should be similar to that of the base year population as no 'evolution' of the population or the land–use/transportation network was applied. Validation of a travel demand model typically consists of comparing the spatial and temporal distribution of aggregate trips by type and travel mode. However, given that this is an activity–based model, it only seems sensible to also investigate the simulated schedules being generated for the population to ensure that they are reasonably realistic. As such, the following investigates the results of this test by first looking at the simulated activity behaviour of the population and then at the trips that evolve from it.

## 5.1 SIMULATED ACTIVITY BEHAVIOUR

Activity behaviour can be investigated from either a macro–level, aggregate perspective or a micro–level, disaggregate perspective. As the behaviour of the observed population was well illustrated at the aggregate level in the descriptive analysis, this section will validate the model by checking that the simulated schedules of some sample persons are reasonable. That is, the schedules should be realistic in terms of the duration, start time and frequency of the composite episodes for each activity type or, in other words, realistic from a temporal standpoint.

The resulting schedules simulated by the model vary greatly in the degree of complexity of behaviour, just as the real schedules of persons in the TTS do. This is illustrated in Figures 5–1 through 5–3. These represent the simulated daily activity behaviour of three persons whose socio–economic attributes are unknown. The x–axis depicts the sequence of activity episodes for the entire day and the y–axis represents the values of both duration and start time for each episode (in minutes, with zero start time equal to 4 a.m.). The first example depicts the most simple possibility: the person does not engage in any out–of–home activities and remains *in–home* all day. Of course, in this case, no trips are required by the person and they have no impact on the transportation network. The second example illustrates what is likely the most common result. Here, the person executes a simple home–based work chain. The *work* episode duration is slightly more than 8 hours and starts at exactly 8:30 in the morning. The remainder of the day

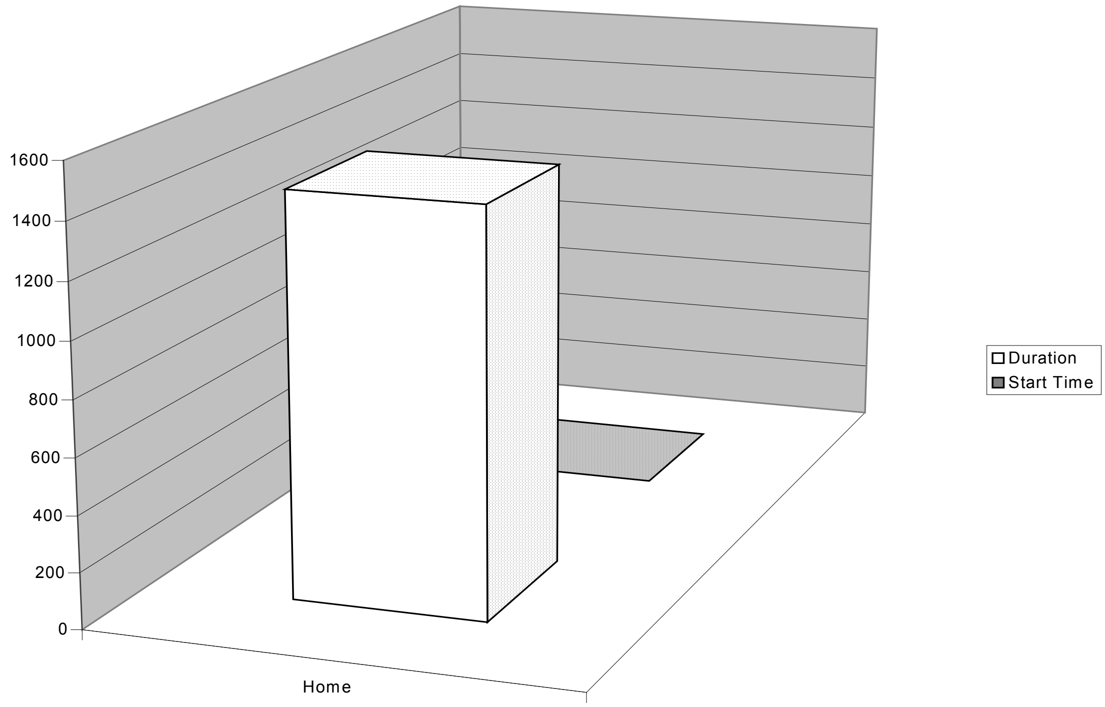



**Figure 5–1 Default Simulated Person Activity Schedule**

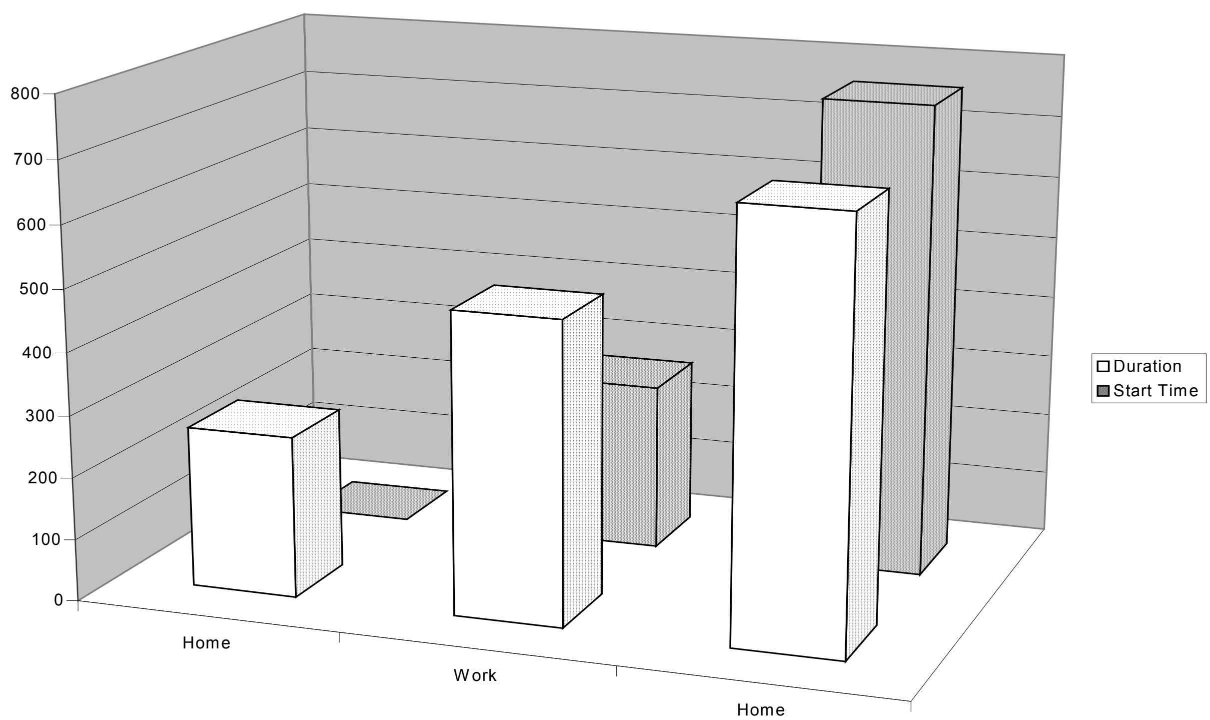



**Figure 5–2 Simple Simulated Person Activity Schedule**

is spent at home. This behaviour is very prevalent in the TTS. Finally, the third example shows a very complicated activity behaviour. This person engages in three home–based chains. The first chain consists of three *work* episodes, each lasting about 5.5, 4.0 and 1.0 hours, respectively. These three episodes are broken up by an *independent other* episode and a *joint market* episode. The other two chains consist of another *joint market* episode and a *joint other* episode. Even with this complexity, the results are not unreasonable. The first *work* episode starts at around 6:30 a.m. and the last *work* episode ends around 6:30 p.m. The *market* and *other* activity episodes are also all of reasonable duration and start time.

## 5.2 SIMULATED TRAVEL BEHAVIOUR

The spatial accuracy of a model is typically verified by comparing the estimated aggregate flow of trips across specific screenlines for which real cordon counts have been completed. The spatial distribution of activity locations and the subsequent trip between each is of little meaning at the disaggregate level; that is, one cannot say whether the simulated travel pattern of one person is right or wrong for the purposes of validation. However, it would be interesting to

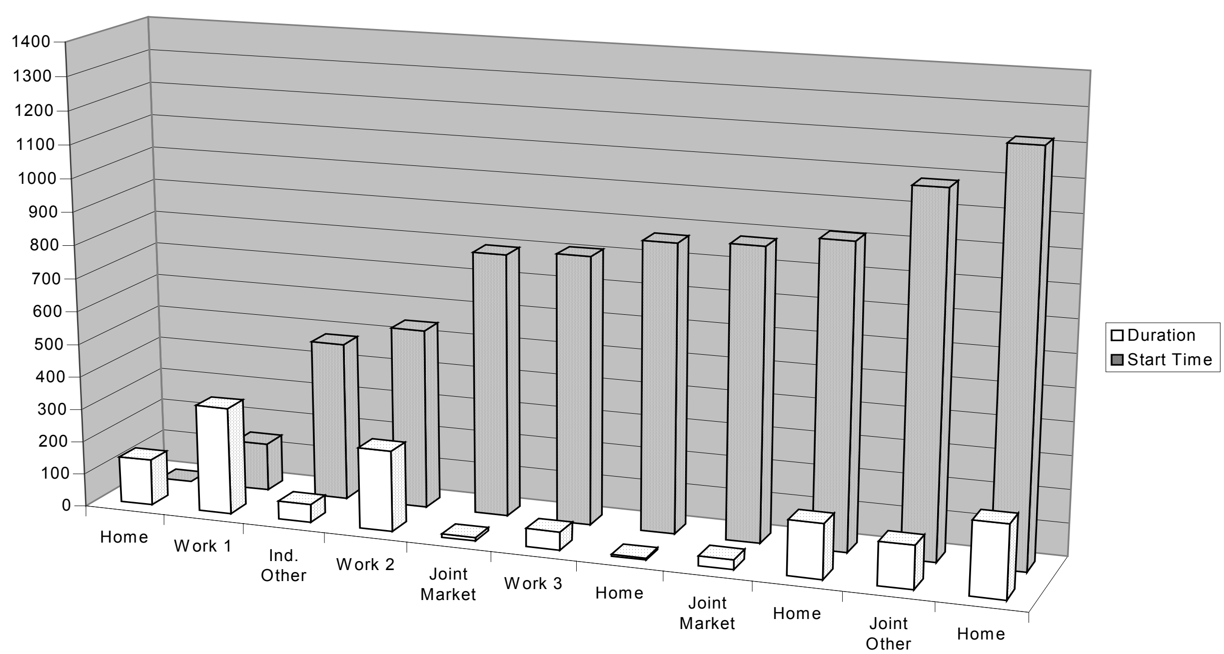



**Figure 5–3 Complex Simulated Person Activity Schedule**

investigate the sequence of trips made to each activity episode location so as to observe how optimal, or efficient, it is. The simulated activity schedule of the person in the last example would require the execution of ten trips to various fixed and unfixed locations. Figure 5–4 displays this sequence of trips. Note that the arrows ends indicate the destination zone centroid only and not any particular point within the zone. The sequence begins in the home zone 331 and proceeds tail–to–tip throughout for all three chains or ten trips. Looking at this pattern without consideration of the context, one might say that it is highly sub–optimal and could be completed much more efficiently in, say, a clockwise direction. This might also lead to the conclusion that the model is producing unrealistic results, as no rational person would behave in such a way. Remember, however, that this represents travel behaviour over an entire day for which activities are spatially distributed with fixed and unfixed start times and which are often subject to dynamic changes to location and start time. Sub–optimal execution of travel patterns is, indeed, a reality. One of the toughest questions to answer is not whether inefficiency should be present in the results in order to be realistic, but rather to what degree. Here, sub–optimality is introduced randomly by the activity scheduling algorithm and the location choice model so the degree of inefficiency is not based in reality. Employing more realistic means, dynamic activity

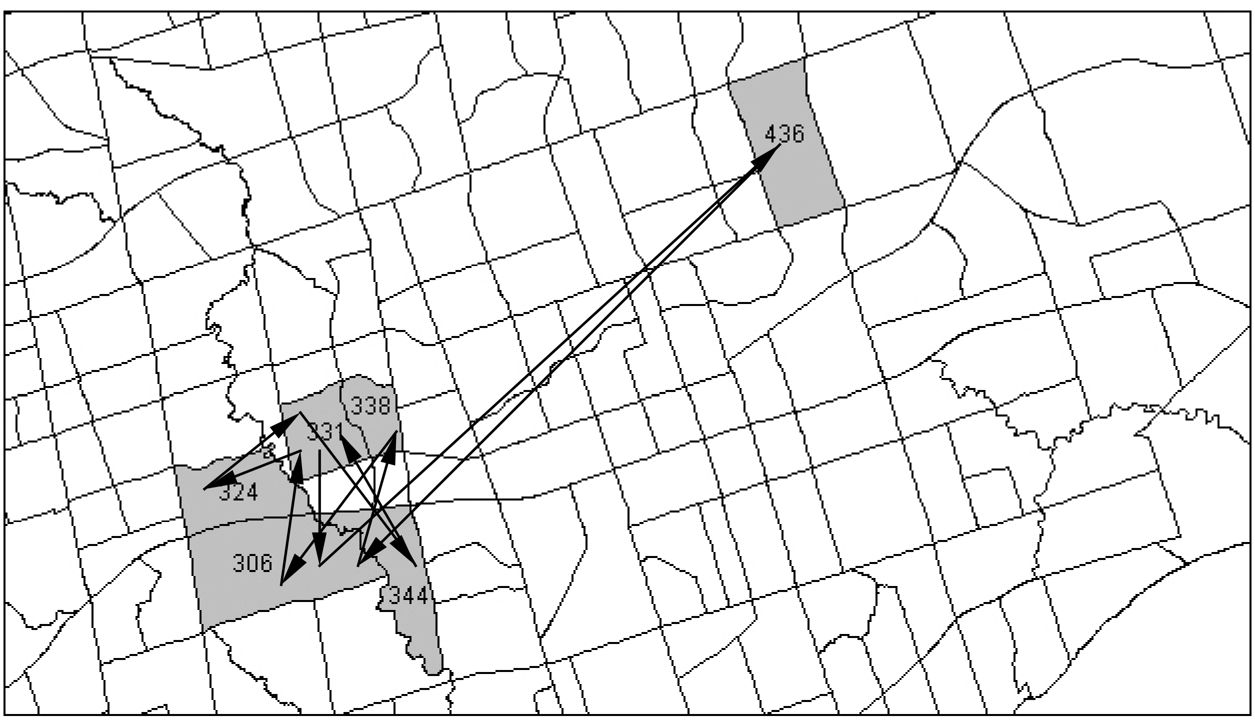



**Figure 5–4 Example of Complex Simulated Trip Pattern**

scheduling represents a possible non–random generator of sub–optimality. This is discussed further in the conclusion.

As a more conventional validation of the model, one can compare the aggregate production of trips predicted by the model to those observed in the TTS. However, a truly direct comparison between the two is not possible for two main reasons. First, although TTS collects personal data for persons under 11 years of age and under, it does not collect trip data for these persons. The model estimates school activity episodes and trips for all children. Second, the model does not estimate trips equivalent to the facilitate passenger, daycare or unknown trip purposes in the TTS. Given these differences, there is still plenty of common ground remaining between the two for comparison.

Overall, it is most important that a travel demand model be able to accurately determine the number of trips by mode and time period. Given that this model is activity–based, the importance of capturing the trip purpose is greater as compared to traditional trip–based models in terms of credibility. Tables 5–1 and 5–2 compare mode split by time period and the number of trips by mode and trip purpose, respectively. Note that the modes and trip purposes are aggregated so as to be common for the purposes of comparison, but are actually defined somewhat differently within the TTS and model.

| Mode | AM Peak | | | Mid–Day | | | PM Peak | | | Evening/Night | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | TTS | Model | Diff. | TTS | Model | Diff. | TTS | Model | Diff. | TTS | Model | Diff. |
| Auto Driver | 53.8 | 46.1 | -7.7 | 66.3 | 63.4 | -2.9 | 58.1 | 52.6 | -5.5 | 64.2 | 66.5 | 2.4 |
| Auto Pass. | 12.9 | 12.8 | -0.1 | 14.7 | 14.9 | 0.2 | 15.3 | 16.1 | 0.8 | 24.1 | 19.5 | -4.6 |
| Transit | 16.6 | 15.7 | -0.9 | 12.2 | 13.8 | 1.6 | 14.4 | 13.1 | -1.2 | 7.8 | 9.4 | 1.6 |
| GO | 2.0 | 1.2 | -0.7 | 0.2 | 0.3 | 0.2 | 1.2 | 0.7 | -0.5 | 0.3 | 0.5 | 0.1 |
| Walk | 10.1 | 18.6 | 8.5 | 4.2 | 5.0 | 0.8 | 7.6 | 13.3 | 5.8 | 1.7 | 3.1 | 1.4 |
| Bicycle | 0.8 | 1.0 | 0.2 | 0.9 | 0.7 | -0.3 | 0.8 | 0.9 | 0.1 | 0.7 | 0.4 | -0.3 |
| School Bus | 3.6 | 4.6 | 1.0 | 0.6 | 1.2 | 0.6 | 2.1 | 2.9 | 0.8 | 0.0 | 0.2 | 0.2 |
| Taxi | 0.2 | 0.0 | -0.2 | 0.6 | 0.6 | 0.1 | 0.3 | 0.3 | 0.0 | 1.1 | 0.5 | -0.6 |
| Other | 0.2 | 0.0 | -0.2 | 0.3 | 0.0 | -0.3 | 0.2 | 0.0 | -0.2 | 0.2 | 0.0 | -0.2 |
| Total | 100.0 | 100.0 | 0.0 | 100.0 | 100.0 | 0.0 | 100.0 | 100.0 | 0.0 | 100.0 | 100.0 | 0.0 |

**Table 5–1 Comparison of Observed (TTS) and Predicted Mode Split (%) by Time Period**

| Mode | Home | | | Work | | | School | | | Market | | | Other | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | TTS | Model | % Diff. | TTS | Model | % Diff. | TTS | Model | % Diff. | TTS | Model | % Diff. | TTS | Model | % Diff. |
| Auto Driver | 2,569 | 2,273 | -12 | 1,604 | 1,681 | 5 | 95 | 86 | -10 | 554 | 444 | -20 | 917 | 863 | -6 |
| Auto Passenger | 680 | 680 | 0 | 189 | 213 | 13 | 148 | 188 | 27 | 179 | 158 | -12 | 361 | 246 | -32 |
| Transit | 566 | 570 | 1 | 313 | 348 | 11 | 167 | 175 | 5 | 60 | 72 | 21 | 137 | 101 | -26 |
| GO | 42 | 31 | -27 | 42 | 32 | -24 | 3 | 2 | -20 | 1 | 1 | 75 | 3 | 2 | -37 |
| Walk | 289 | 503 | 74 | 71 | 107 | 51 | 194 | 391 | 101 | 8 | 12 | 59 | 23 | 18 | -25 |
| Bicycle | 34 | 36 | 5 | 16 | 1 | -93 | 11 | 24 | 117 | 4 | 6 | 60 | 12 | 7 | -46 |
| School Bus | 76 | 114 | 50 | 1 | - | -100 | 78 | 113 | 45 | 0 | 1 | 1045 | 2 | 3 | 41 |
| Taxi | 24 | 16 | -36 | 8 | 2 | -80 | 1 | - | -100 | 2 | 8 | 209 | 12 | 9 | -28 |
| Other | 8 | - | -100 | 4 | - | -100 | 1 | - | -100 | 1 | - | -100 | 6 | - | -100 |
| Total | 4,287 | 4,222 | -2 | 2,247 | 2,384 | 6 | 698 | 980 | 40 | 809 | 702 | -13 | 1,474 | 1,248 | -15 |

**Table 5–2 Comparison of Observed and Predicted Number of Trips (1000's) by Mode and Trip Purpose**

These results show that the model predicts fairly well. In general, transit and auto–passenger trips are predicted quite well while auto–driver trips are under predicted and walk/school bus trips are over–predicted. This is confirmed in Table 5–3, which compares the total, all–day mode splits. This discrepancy is undoubtedly related, at least somewhat, to the two differences mentioned above; that is, the inclusion of trips for children 10 years of age and under (which are likely dominated by the walk and school bus modes) as well as the omission of facilitate passenger and daycare trips (which are likely auto–driver dominated) from the model. Furthermore, TTS reports an additional 590,000 trips for facilitate passenger, daycare and unknown purposes, while the model generates an additional 72,000 trips that occur external to the GTA. Taking these trips into consideration, the model actually under predicts total trips by almost 500,000. This is obviously a non–trivial amount that will be accounted for in future versions of the model. However, it should be noted that these 'chauffeur' types of trips are generally not too far out of the way in terms of the overall route taken by the driver. Thus, they represent more of a route correction than a trip entity. It is the small amount of trips made specifically for this purpose only that truly have an impact. Otherwise, the model should ideally estimate a total number of trips approximately equal to the observed TTS total plus the amount generated for persons under 11 years.

| Mode | TTS | Model | Difference | Percent Difference | TTS Split | Model Split | Difference |
|---|---|---|---|---|---|---|---|
| Auto Driver | 5,738,737 | 5,347,024 | -391,713 | -6.8 | 60.3 | 56.1 | -4.2 |
| Auto Passenger | 1,557,502 | 1,484,761 | -72,741 | -4.7 | 16.4 | 15.6 | -0.8 |
| Transit | 1,242,019 | 1,267,636 | 25,617 | 2.1 | 13.1 | 13.3 | 0.2 |
| GO | 90,460 | 67,954 | -22,506 | -24.9 | 1.0 | 0.7 | -0.2 |
| Walk | 585,436 | 1,030,906 | 445,470 | 76.1 | 6.2 | 10.8 | 4.7 |
| Bicycle | 77,255 | 73,794 | -3,461 | -4.5 | 0.8 | 0.8 | 0.0 |
| School Bus | 156,421 | 230,572 | 74,151 | 47.4 | 1.6 | 2.4 | 0.8 |
| Taxi | 47,958 | 33,839 | -14,119 | -29.4 | 0.5 | 0.4 | -0.1 |
| Other | 19,933 | 0 | -19,933 | -100.0 | 0.2 | 0.0 | -0.2 |
| Total | 9,515,721 | 9,536,488 | 20,767 | 0.2 | 100.0 | 100.0 | 0.0 |

**Table 5–3 Comparison of Total Number of Trips by Mode and Mode Split (%)**

# 6 CONCLUSION

The preceding has documented the initial development and testing of a model that represents the next generation of travel demand forecasting methodologies. The first application of the model will involve estimating carbon dioxide emissions from personal use vehicles over a 24–hour period in the GTA (see Miller and Roorda, 2002). It will eventually be implemented as the main travel demand component of the Integrated Land–Use, Transportation and Environment (ILUTE) modelling system that is currently under development at the University of Toronto. It is envisioned that ILUTE will one day replace the current trip–based travel demand model also developed at the University and currently used by the City of Toronto for transportation planning.

The limited nature of the data has been mentioned numerous times throughout the preceding as a hindrance to the potential of this model. The ultimate objectives and subsequent structure of even this initial specification are meant to be supported by data that is far more detailed than is available through the TTS, at least in the activity–based sense. Given, however, that the model backbone is in place, this potential may be realised once such data becomes available. Some of the major initial model limitations, possible future improvements and/or ensuing data requirements are summarised as follows.

- Account for the correlation between activity duration, start time and frequency. One of the most unrealistic aspects of this model is that these activity attributes are determined independently of one another. For example, activity episode duration is likely to decrease as episode frequency increases. Given that these activity episode attributes are determined probabilistically, some unrealistic results are bound to be generated. This represents possibly the biggest challenge to this next generation of travel models.
- Include ‘serving dependants’ as a project type. Dependants represent a special challenge for two reasons. First, data on the most common dependent, children, are scarce as most surveys concede that respondents may not wish to report the travel behaviour of their children. Second, this project essentially involves ensuring that dependants are supervised and transported for the entire duration of the analysis period according to their schedules. Distinction can be made between the dependant ‘tagging along’ during the

adult's activity (in which case the location is determined by the adult) and the adult specifically accompanying the dependent to an activity (in which case the location is determined by the dependent). In either case, the activity types engaged in by adult and dependent are different unless the activity is a joint one, which is conceptually distinct. Data for this is partially contained in the TTS as facilitate passenger and/or daycare trip purposes. One, however, can only determine that the facilitated person is a household member that is over 10 years of age. Facilitated children and non–household members are not observed.

- As an extension of the previous point, when driving another person as an auto passenger, distinguish between facilitating a household member and a non–household member as 'serving a dependent' and 'carpooling'. The distinction is very important to a household model in terms of the gains to household utility.
- Merge activity schedule, location choice and mode choice into a simultaneous decision structure so that the effects of prior and posterior activity locations as well as present travel mode (as either drive or non–drive) are accounted for.
- Implement vehicle allocation/availability and holdings models. Currently, mode choices are made without consideration of whether there is a car presently available. This unrealistic aspect of the model is due to time constraints only and will be corrected in the future; the mere fact that this is possible is a giant leap forward for travel demand models. The allocation model might utilise a utility maximisation approach that is based on 'with car/without car' schedules to determine which household member(s) is 'most deserving' of the vehicle(s) at that time. Vehicle holdings models have already been developed and are needed for forecasting in a microsimulation framework as it unreasonable to assume that a household will have the same number and type of automobile ten years into the future.
- Incorporate fully the concept of the project as a fundamental organising principle. This would add a more realistic structure to the activity episode generation and scheduling process by tying together different, yet interrelated, activity types as well as giving a reference for determining the utility of each episode relative to other episodes of the same type. This concept is also important in representing a true household decision structure as each project can involve several household members and thus represents a means by

which their schedules can be tied together. An important implication of this is that the *in–home* activity could no longer be used as the generic default activity as some specific activity episodes evolving from the project would be executed at home, requiring the same scheduling process as out–of–home episodes. Data that relates activity episodes both within a person's schedule and to those of other household members would be required. The involves the concept of episode utility discussed earlier, which would have to be quantified.

- Include sub–chains anchored at rail transit parking lots (i.e. Kiss N' Ride) in the mode choice model. The required data is already present in the TTS, it is only a matter of altering the mode choice model.
- Move from a static scheduling process to a dynamic one. Even in synthesising a static schedule, modification of existing episodes is often required in order to add additional ones. This is somewhat indicative of the dynamic nature of the scheduling process as it occurs in reality, where information affecting decision–making arises at varying times. Capturing this process in a model would require data that measures addition and modification/deletion of activity episodes as well as the degree of fixation (see below) of the activity (i.e. long–term versus spontaneous), which is likely highly codependent with its utility. Such data is currently being collected using the CHASE software developed by Doherty and Miller (2000).
- Schedule fixed and unfixed activities according to their inherent differences. This model simulated this difference through priority in scheduling order. While this worked well for the single day analysis period, a more realistic approach would likely involve a week long period. Fixed activities, such as work at the usual place, would be set at the beginning of the analysis period. A dynamic scheduling process would then be responsible for adding the unfixed activities. This would require definition of which activity types are fixed and which are not, which might include varying degrees of such.

These current limitations and potential improvements may seem so numerous and significant as to question the validity of the present model and its results. However, one should consider that each of the specified improvements are only even a possibility as a result of the household–level, activity–based microsimulation framework and could not be implemented into a conventional trip–based model. It is conceivable that the biggest limitation to the potential of this next generation of travel demand models is the quality, variation and enormous amount of data that is required to support them. The goal of these models is to capture the behaviour of the rational decision–maker, hopefully to the point where sub–optimal results arise more often from our simulation of the real–life sub–optimal behaviour of decision–makers observed in the data rather than randomly. Indeed, over–optimization is a potential problem (albeit a modeller's dream), however some of these potential improvements may also provide for this. Consider, for example, that, within a dynamic framework incorporating a vehicle allocation and availability model, a person might initially not be allocated the household car for a particular trip chain due to the execution of a trip chain with higher household utility by another household member. It is possible that a dynamic modification to that person's schedule within the trip chain may cause that chain to have a higher household utility than that of the other household member, meaning that the vehicle should be allocated to them. However, given that the other member had already been assigned the vehicle for their trip chain, the vehicle would not be available and therefore could not be re–allocated, resulting in execution of the trip chain under sub–optimal conditions. A similar result might occur from a dynamic change to the location of an activity episode, altering the optimal location choices made for prior and/or posterior episodes. Finally, given that these modules could be successfully completed, it is unlikely that they would have much meaning within the context of a single day. All of the above, to some degree or another, would realistically require an analysis period spanning a week or even a month. This is still minuscule compared to the yearly time steps likely to be taken by the longer–term modules, such as a vehicle holdings model, that will evolve the population to the given future year, which is typically 20 years from the present.

## REFERENCES


Algers, S., A. Daly and S. Widlert (1997). Modeling Travel Behaviour to Support Policy Making in Stockholm, In *Modeling Travel Behaviour in an Era of Change*, P.R. Stopher and M. Lee–Gosselin (Eds.), Pergamon, Guildford, pp. 547–569.

Arentze, T., F. Hofman, H. Van Mourik and H. Timmermans (2000). ALBATROSS: A Multi–Agent Rule–Based Model of Activity Pattern Decisions, forthcoming, *Transportation Research Record*.

Arentze, T. and H. Timmermans (2000). *ALBATROSS: A Learning Based Transportation Oriented Simulation System*. European Institute of Retailing and Services Studies.

Arentze, T. and H. Timmermans (2001). Modeling the Formation of Activity Agendas Using Reactive Agents, forthcoming, *Transportation Research Record*.

Axhausen, K.W. and T. Gärling (1992). Activity–Based Approaches to Travel Analysis: Conceptual Frameworks, Models and Research Problems, *Transport Reviews*, (12), pp. 324–341.

Axhausen K.W. (1998). Can We Ever Obtain the Data We Would Like to Have?, In *Theoretical Foundations of Travel Choice Modeling*, T. Gärling, T. Laitila, and K. Westin (Eds.), Elsevier, The Netherlands, pp. 305–323.

Becker, G.S. (1965). A Theory of the Allocation of Time, *The Economic Journal*, vol. 75(299), pp. 493–517.

Berkowitz, M.K., N.T. Gallini, E.J. Miller and R.A. Wolfe (1990). Disaggregate Analysis of the Demand for Gasoline, *Canadian Journal of Economics*, pp. 253–275.

Bhat, C. and F.S. Koppelman (1999). Activity–Based Modeling of Travel Demand, In *The Handbook of Transportation Science*, R. Hall (Ed.), Kluwer Academic Publishers, Norwell, Massachusetts, pp. 35–61.

Bhat, C., J. Carini and R. Misra (1999). Modeling the Generation and Organization of Household Activity Stops, *Transportation Research Record*, (1676), pp. 153–161.

Chandrasekharan, B. and K.G. Goulias (1999). Exploratory Longitudinal Analysis of Solo and Joint Trip Making in the Puget Sound Transportation Panel, *Transportation Research Record*, (1676), pp. 77–85.

Chen, Z. and F. Woolley (1999). A Cournot–Nash Model of Family Decision Making, Ph.D. dissertation, Carleton University, Ottawa, Ontario, pp. 1–45.

Damm, D. (1980). Interdependencies in Activity Behavior, *Transportation Research Record*, (750), pp. 33–40.

DMG (1996). *The Transportation Tomorrow Survey: Data Guide (Version 2.1)*. Joint Program in Transportation, University of Toronto, Toronto, Ontario.

DMG (1997). *1996 Transportation Tomorrow Survey: Interview Manual*. Joint Program in Transportation, University of Toronto, Toronto, Ontario.

Doherty, S.T. and E.J. Miller (2000). A Computerized Household Activity Scheduling Survey, *Transportation*, vol. 27(1), pp. 75–97.

Ettema, D.F., A. Daly, G. de Jong and E. Kroes (1997). Towards an Applied Activity–Based Travel Demand Model, Paper presented at the IATBR Conference, Austin, Texas.

Ettema, D.F. and H. Timmermans, Eds. (1997). *Activity–Based Approaches to Travel Analysis*. Pergamon, Amsterdam.

Fujii, S., R. Kitamura and K. Kishizawa (1999). An Analysis of Individuals' Joint Activity Engagement Using a Model System of Activity-Travel Behavior and Time Use, *Transportation Research Record*, (1676), pp. 11-19.

Gärling, T., K. Brannas, R. Garvill, R.G. Golledge, S. Gopal, E. Holm and E. Lindberg (1989). Household Activity Scheduling, *Transport Policy*, vol. 4 (Management and Technology: Selected Procedings of the Fifth World Conference of Transport Research), pp. 235–248.

Gärling, T., R.G. Golledge and M.–P. Kwan (1994). Computational Process Modeling of Household Travel Activity Scheduling, *Transportation Research*, vol. 28(B), pp. 355–364.

Gärling, T., T. Kalen, J. Romanus, M. Selart and B. Vilhelmson (1998). Computer Simulation of Household Activity Scheduling, *Environment and Planning*, vol. 30(A), pp. 665–679.

Gärling, T., T. Laitila and K. Westin, Eds. (1998). Theoretical Foundations of Travel Choice Modeling. Pergamon, Amsterdam, The Netherlands.

Gliebe, J.P. and F.S. Koppelman (2001). A Model Of Joint Activity Participation, forthcoming, *Transportation Research Record*.

Golledge, R.G., M.–P. Kwan and T. Gärling (1994). Computational Process Modeling of Household Travel Decisions Using a Geographical Information System, *Papers in Regional Science*, vol. 73, pp. 99–117.

Golob, T.F. and M.G. McNally (1997). A Model of Activity Participation and Travel Interactions Between Household Heads, *Transportation Research*, vol. 31B(3), pp. 177–194.

Golob, T.F. (1999). A Simultaneous Model of Household Activity Participation and Trip Chain Generation, *Transportation Research* , vol. 34B(5), pp. 355–376.

Goulias, K.G. and T. Kim (2001). A Multilevel Analysis of Activity and Travel–Patterns Accounting for Person–Specific and Household–Specific Observed and Unobserved Effects Simultaneously, forthcoming, *Transportation Research Record*.

Kitamura, R. and M. Kermanshah (1984). Sequential Model of Interdependent Activity and Destination Choices, *Transportation Research Record*, (987), pp. 81–89.

Kockelman, K.M. (1999). Application of a Utility–Theory–Consistent System–of–Demand Equations to Household Activity Travel Choice, *Transportation Research Record*, (1676), pp. 20–29.

Manser, M. and M. Brown (1980). Marriage and Household Decision–Making: A Bargaining Analysis, *International Economic Review*, vol. 21(1), pp. 31–44.

McElroy and Horney (1981). Nash Bargained Household Decisions, *International Economic Review*, vol. 22(2), pp. 333–349.

Miller, E.J. (2001a). Propositions. Department of Civil Engineering, University of Toronto, Toronto, Ontario, Working paper.

Miller, E.J. (2001b). Generating and Scheduling Episodes. Department of Civil Engineering, University of Toronto, Toronto, Ontario, Working paper.

Miller, E.J. (2001c). Episode Location Choice. Department of Civil Engineering, University of Toronto, Toronto, Ontario, Working paper.

Miller, E.J. (2001d). Episode Mode Choice. Department of Civil Engineering, University of Toronto, Toronto, Ontario, Working paper.

Miller, E.J. and W.R. Rhamey (1987). Modeling Household Mode Choice, *Fourth World Conference on Transport Research*, pp. 1551–1569.

Miller, E.J. and M. Roorda (2002). Estimating $CO_2$ Emissions and Other Transportation Impacts of Alternative Urban Forms: Summary of Base Network Model Test Results. Joint Program in Transportation, University of Toronto, Toronto, Ontario.

Montemarquette, C. and L. Monty (1987). An Empirical Model of a Household's Choice of Activities, *Journal of Applied Econometrics*, vol. 2(2), pp. 145–158.

Nash, J.F. (1950). The Bargaining Problem, *Econometrica*, vol. 18(1), pp. 155–162.

Ott, N. (1992). *Intrafamily Bargaining and Household Decisions*. Springer–Verlag, Berlin.

Recker, W.W., M.G. McNally and G.S. Root (1986a). A Model of Complex Travel Behavior: Part I: Theoretical Development, *Transportation Research*, vol. 20A(4), pp. 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*, vol. 20A(4), pp. 319–330.

Recker, W.W. (1995). The Household Activity Travel Pattern Problem: General Formulation and Solution, *Transportation Research*, vol. 29B(1), pp. 61–77.

Roorda, M., E.J. Miller and J. Guan (2002). A Microsimulation Model of 24–Hour Household Activity Schedules – Methodology. Joint Program in Transportation, University of Toronto, Toronto, Ontario.

Scott, D.M. (2001). A New Approach To Modeling Household Activity–Episode Generation That Captures Interactions Between Household Heads, forthcoming, *Transportation Research Record*.

Simma, A. and K.W. Axhausen (2001). Within–Household Allocation of Travel: The Case of Upper Austria, forthcoming, *Transportation Research Record*.

Stopher, P.R. and M. Lee–Gosselin, Eds. (1997). Understanding Travel Behaviour in an Era of Change. Pergamon, Oxford.

Townsend, T.A. (1987). The Effects of Household Characteristics on the Multi–Day Time Allocation and Travel/Activity Patterns of Households and Their Members, Ph.D. dissertation, Northwestern University, Evanston, Illinois.

Wen, C. and F.S. Koppelman (1999). An Integrated System of Stop Generation and Tour Formation for the Analysis of Activity and Travel Patterns, *Transportation Research Record*, (1676), pp. 136–144.

Wen, C. and F.S. Koppelman (2000). A Conceptual and Methodological Framework for the Generation of Activity–Travel Patterns, *Transportation*, vol. 27, pp. 5–23.

# APPENDIX A

## GLOSSARY OF TERMS

Following are definitions of selected terms used within the text. Their usage may or may not be original or consistent with that of other activity–based literature and are defined in terms of how they are utilised within this model.

| | |
|---|---|
| Activity | An action with specific purpose. Examples of activity types include work, shopping or entertainment. The level of definition of purpose, however, is arbitrary and may be at any level of precision warranted. |
| Activity Episode | A unit of activity, having specific duration, start time and location attributes. |
| Agenda | A list of activity episodes for a given project. A project's agenda may span any amount of time, even beyond that of the analysis period. Activity episodes may overlap between household members *within* agendas as well as for a single person *between* agendas. |
| Decision–Maker | A unit called upon to resolve a conflict. The unit may be comprised of a single household member, several household members or the household as an entity itself. |
| Duration | The length of time required to execute an activity or travel episode. |
| Event | A collection of one or more activity episodes of the same activity type that are logically tied together to represent common periods of time (e.g. the 'workday' often consists of several work episodes). More than one daily activity event may occur. |
| Frequency | The number of daily occurrences of an activity (i.e. the number of activity episodes). |
| Generation | The process of creating potential activity episodes for possible insertion into a schedule. |
| Home–Based Chain | A trip chain that begins upon leaving home and ends upon returning, consisting of all activity and travel episodes in between. |
| Joint Activity Episode | An activity episode involving two or more members of a household. Utility is produced through joint engagement in the activity as opposed to merely being at the same location. |

| | |
|---|---|
| Project | A set of related activity episodes, tied together by a common objective. Ultimately, these evolve from fundamental human needs such as sustenance, health, shelter, *et cetera*. Projects may involve one or more household members as well as differing activity types. |
| Schedule | The final sequence of feasible, non–overlapping activity and travel episodes generated for a person. The sum of all episode durations equals that of the analysis time period. |
| Scheduling | The process of resolving conflicts between generated activity episodes to produce a schedule. This may involve modifying or rejecting some generated activity episodes. |
| Start Time | The point in time that defines the beginning of an activity or travel episode. Trip start time plus travel episode duration plus activity episode duration equals the activity end time, or the start time of the next scheduled travel episode. |
| Travel Episode | A unit of travel from an origin to a destination (i.e. a trip), having start time and duration attributes. Duration is synonymous with travel time, which is conditional upon the travel mode. Frequency is determined by the scheduled number of activity episodes. |

# APPENDIX B

## DATA PREPARATION & RECORD DELETION

Following is a record of the process and assumptions made in extracting activity–based variables from the trip–based 1996 TTS database and of the criteria used to choose records for deletion.

### Original variables from TTS 'Trip Attributes' database:[1]

1. hhld_num
2. pers_num
3. trip_num
4. start_time
5. mode_prime
6. purp_orig
7. pd_orig
8. gta96_orig
9. purp_dest
10. pd_dest
11. gta96_dest
12. trip_km

### Variables created from original variables and other sources:

1. total_dur
   - the total duration of the trip and the activity, in minutes
   - equal to the start time of the subsequent trip minus the start time of the current trip
   - for the last trip of the day, the start time of the subsequent trip was assumed to be the end of the day

2. total_dur5
   - total_dur rounded to the nearest 5 minutes

3. trip_start5
   - the start time of the trip rounded to the nearest 5 minutes, with 4 a.m. as time zero

4. tmode_accs
   - the access mode if the trip was made by public transit, GO rail, or jointly between GO rail and public transit (i.e. mode_prime equals B, G, or J)

5. trip_dur
   - the estimated travel time (or trip duration) from origin (gta96_orig) to destination (gta96_dest)
   - see below for further detail

[1] For descriptions of TTS variables, see DMG (1996).

6. trip_dur5
   - trip_dur rounded to the nearest 5 minutes

7. eps_start5
   - the start time of the activity episode rounded to the nearest 5 minutes (by default), with 4 a.m. as time zero
   - equal to trip_start5 plus trip_dur5

8. eps_dur5
   - the activity episode duration rounded to the nearest 5 minutes (by default)
   - equal to total_dur5 minus trip_dur5

9. jt_pers
   - pers_num of the other person (within the household) who participated in the trip and activity jointly
   - determined by matching start_time, purp_dest, mode_prime, and gta96_dest of each trip record with those of other household members

10. jt_trip
    - trip_num of the other person who participated in the trip and activity jointly

11. ch_num
    - the number of the chain to which a particular trip belongs to, by person
    - a trip chain begins when the person leaves home and ends when the person returns home
    - chain numbering did not begin until the first trip from home

12. ch_tr_num
    - the number of the trip within its chain
    - starts at one for each new chain of a person

13. ch_ty
    - the type of chain
    - defined as a home–based chain (H), a work–based sub–chain (W) or a school–based sub–chain (S)

14. subch_num
    - the number of the sub–chain to which a particular trip belongs to, by person
    - a trip sub–chain begins when the person leaves work or school and ends when the person returns to the same, within the same chain

15. subch_tr_num
    - the number of the trip within its sub–chain
    - starts at one for each new sub–chain of a person

## Estimation of travel time (trip_dur):

Travel times had previously been estimated for:

1. Automobile (i.e. mode_prime equal to D, M, P, S, or T) for seven time periods spanning a 24-hour period,

2. Public transit (mode_prime equal to B) for same zone access and for automobile access (i.e. Kiss N' Ride), and

3. GO rail (mode_prime equal to G or J) for automobile access and for walk/transit access.

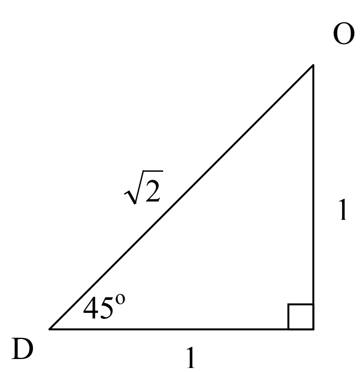
These times were produced from the EMME/2 route assignment procedure used in GTAModel, a trip–based travel demand model for the GTA. They are estimates from origin zone (gta96_orig) centroid to destination zone (gta96_dest) centroid. Issues and solutions related to travel time estimation include:

1. There were no travel time estimations for intra–zonal (i.e. within zone) trips. For these, a travel time was estimated by dividing the trip distance (trip_km) by an assumed average speed. The assumed speeds were 25 km/h for automobile and 15 km/h for transit. Intra–zonal GO rail trips were assumed not to occur (see below).

2. There were no travel time estimations for the remaining modes (walk, bicycle, and other). These were calculated as above, using assumed speeds of 4 km/h for the walk mode and 15 km/h for the bicycle and other modes.

3. The trip distance variable (trip_km) used in the above two calculations is a straight line distance. To avoid underestimating these travel times, this distance is multiplied by 1.4 to compensate. This assumes a Manhattan grid network layout, as explained below.

4. Where trip distance was needed for the calculation, some had a value of zero. As these values had been rounded and not truncated, an average value of 0.25 kilometres was used where required.




$$\sqrt{2}x = 1 + 1$$
$$x = \sqrt{2}$$
$$x \cong 1.4$$

5. Estimated GO rail travel times were poor and were usually were in excess of reasonable times. As such, these times were capped at 90 minutes.

6. Some estimated GO rail travel times were provided for one direction only, given that they were from a morning peak–period model that provided times based on the schedule for that period. As a result, many travel times for trips occurring out of this time period were missing (usually evening commute trips out of the city core). Thus, the morning period travel time was substituted.

7. For transit trips, the origin and destination zones were not always the same as the access and egress zones, resulting in some missing transit travel times. In these cases, travel times were searched for in the following order:
   - access zone to destination zone,
   - origin zone to egress zone, and
   - access zone to egress zone.

## Issues related to Joint Activity and Trip Chain Identification:

As the joint activity and trip chain fields were fairly simple approaches to quantifying possibly complex phenomenon, there exist some faults in these fields of the data set. The problem with both lies in overlapping information. It is impossible to represent joint activities involving more than two people or overlapping sub–chains in the manner explained above. As such, due to the algorithms used to identify these phenomenon, those occurring first in the data set are overwritten by those occurring later. For instance, for a three person joint activity, the first and second persons joint activity variables will both refer to the third person, and the third person's variable will refer to the second. Similarly, the variables of the second overlapping sub–chain will overwrite the first, possibly resulting in the complete erasure of the first if it is contained completely within the second.[2]

## Criteria for choosing records for deletion:

Once all activity episode durations were calculated, the data set had to be cleaned of faulty records for which no travel time had been found or, if one had been found, had resulted in a negative activity episode duration. As the model is a household level one, a faulty record meant that all records associated with it had to be deleted, as one could not arbitrarily remove people from a household and use it to capture household interdependencies. Thus, all households with records with the following characteristics were deleted:

1. An origin zone or destination zone equal to 4000. This code referred to trips that either left or entered the study area, and for which we had no information to calculate travel times.

[2] Refer to households 100211 and 101160 (person 1) of the final data set for examples of these joint activity and trip chain identification errors, respectively.

2. Public transit or GO trips that were reported as occurring outside of these service areas (i.e. origin zone or destination zone greater than 2, 670).

3. Intra–zonal GO rail trips.

4. A missing mode code (i.e. ‘9’).

5. Missing public transit and GO rail travel times.

6. The calculated total duration (total_dur) was null. This meant that two trips had been reported or incorrectly recorded as starting at the exact same time.

7. The calculated and rounded total duration (total_dur5) was null for activities (i.e. purp_dest) other than facilitate passenger or daycare. This meant that the reported activity had lasted either 1 or 2 minutes, a duration only reasonable for a facilitate passenger trip and activity (in this case, the trip and the activity may be the same entity). These records were hard–coded as having a trip duration and an activity episode duration of 5 minutes (the total duration was left as null).

8. Calculated activity durations (act_dur5) less than –15 and total durations (total_dur5) greater than 30. This retained trips that had small total durations and for which a slightly overestimated travel time would result in a negative activity episode duration. This also retained approximately half of the records with negative activity durations. These values were hard–coded positive according to the following table (as determined by the most likely values found in the rest of the data set):

| Total Duration | Trip Duration | Activity Episode Duration |
|---|---|---|
| 5 | 0 | 5 |
| 10 | 5 | 5 |
| 15 | 5 | 10 |
| 20 | 5 | 15 |
| 25 | 5 | 20 |
| 30 | 5 | 25 |

After completion of the data preparation and deletion of unwanted records, the data set consists of 434, 583 records, representing 148, 299 persons or 73, 034 households.

# APPENDIX C

## CODE USED FOR DATA PREPARATION

Following is the code written to extract activity–based information from the TTS data set. The code is written in the C language and assumes specific input file format and content. It is intended mainly as a guide to the rules and algorithms applied to the data.

### Program 1: Creation of Trip Chain Variables

```
#include "stdio.h"
#include "string.h"
#include "stdlib.h"

void main(void)
{
	FILE *f1, *f2, *f3;
	int hhld_num,pers_num,trip_num,start_time,pd_orig,gta96_orig,pd_dest,gta96_dest,trip_km;
	char mode_prime,purp_orig,purp_dest;
	int ch_num,ch_tr_num,subch_num,subch_num_tr,pers_chk=0,hhld_chk=0,x=0,y=1,i=0,w1z=0,
		w2z=0,w3z=0,w1c=0,w2c=0,w3c=0,s1z=0,s2z=0,s3z=0,s1c=0,s2c=0,s3c=0;
	long int n=0;
	char ch_ty,t;
	char (*str)[60] = new char[600000][60];

	if ((f1 = fopen("trip96.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}

	if ((f2 = fopen("middle trip chains.txt","w+")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}

	while (fscanf(f1, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d\n", hhld_num, &pers_num,
		&trip_num,&start_time,&mode_prime,&purp_orig,&pd_orig,&gta96_orig,&purp_dest,&pd_
		dest,&gta96_dest,&trip_km) != EOF) {
		n++;
		if (hhld_num != hhld_chk) ch_num=0, ch_tr_num=0, subch_num=0, subch_num_tr=0,
			w1z=0, w2z=0, w3z=0, w1c=0, w2c=0, w3c=0, s1z=0, s2z=0, s3z=0, s1c=0,
			s2c=0, s3c=0;

		else if (pers_num != pers_chk) ch_num=0, ch_tr_num=0, subch_num=0,
			subch_num_tr=0, w1z=0, w2z=0, w3z=0, w1c=0, w2c=0, w3c=0, s1z=0, s2z=0,
			s3z=0, s1c=0, s2c=0, s3c=0;

		hhld_chk = hhld_num;
		pers_chk = pers_num;

		w1c++; w2c++; w3c++; s1c++;   s2c++; s3c++;

		if (purp_orig == 'H') {
```

```
            ch_num++;
            ch_tr_num = 1;
            w1z = 0;w2z = 0;w3z = 0;s1z = 0;s2z = 0;s3z = 0;w1c = 0;w2c = 0;w3c = 0;s1c =
                        0;s2c = 0;s3c = 0;
}

else if (purp_orig != 'H' && ch_num) ch_tr_num++;

if (purp_orig == 'W') {
            if (!w1z) {
                        w1z = gta96_orig;
                        w1c = 1;
            }
            else if (!w2z && gta96_orig != w1z) {
                        w2z = gta96_orig;
                        w2c = 1;
            }
            else if (!w3z && (gta96_orig != w1z || gta96_orig != w2z)) {
                        w3z = gta96_orig;
                        w3c = 1;
            }
}

else if (purp_orig == 'S') {
            if (!s1z) {
                        s1z = gta96_orig;
                        s1c = 1;
            }
            else if (!s2z) {
                        s2z = gta96_orig;
                        s2c = 1;
            }
            else if (!s3z) {
                        s3z = gta96_orig;
                        s3c = 1;
            }
}

if (purp_dest == 'R' && (w1z == gta96_dest || w2z == gta96_dest || w3z == gta96_dest)
            &&(!(gta96_orig == gta96_dest && purp_orig == 'W'))) {
            subch_num++;
            ch_ty = 'W';
            if (w1z == gta96_dest) {
                        subch_num_tr = w1c;
                        w1c = 0;
            }
            else if (w2z == gta96_dest) {
                        subch_num_tr = w2c;
                        w2c = 0;
            }
            else if (w3z == gta96_dest) {
                        subch_num_tr = w3c;
                        w3c = 0;
            }

            fprintf(f2, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%c,%d,%d\n",
```

```
				hhld_num,pers_num,trip_num,start_time,mode_prime,purp_orig,pd_orig,
				gta96_orig,purp_dest,pd_dest,gta96_dest,trip_km,ch_num,ch_tr_num,ch
				_ty,subch_num,subch_num_tr);
		}

		else if (purp_dest == 'C' && (s1z == gta96_dest || s2z == gta96_dest || s3z ==
			gta96_dest) && (!(gta96_orig == gta96_dest && purp_orig == 'S'))) {
			subch_num++;
			ch_ty = 'S';
			if (s1z == gta96_dest) {
				subch_num_tr = s1c;
				s1c = 0;
			}
			else if (s2z == gta96_dest) {
				subch_num_tr = s2c;
				s2c = 0;
			}
			else if (s3z == gta96_dest) {
				subch_num_tr = s3c;
				s3c = 0;
			}

			fprintf(f2, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%c,%d,%d\n",
				hhld_num,pers_num,trip_num,start_time,mode_prime,purp_orig,pd_orig,
				gta96_orig,purp_dest,pd_dest,gta96_dest,trip_km,ch_num,ch_tr_num,ch
				_ty,subch_num,subch_num_tr);
		}

		else fprintf(f2, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,H,0,0\n", hhld_num,
			pers_num, trip_num, start_time, mode_prime, purp_orig, pd_orig, gta96_orig,
			purp_dest,pd_dest,gta96_dest,trip_km,ch_num,ch_tr_num);
	}

	fclose(f1);
	rewind(f2);

	for (i=0; i<n; i++) {
		fgets(str[i], 60, f2);
	}
	rewind(f2);

	for (i=n; i>0; i--) {
		fputs(str[i-1], f2);
	}
	fclose(f2);

	if ((f2 = fopen("middle trip chains.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}

	if ((f3 = fopen("trip chains.txt","w+")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}
```

```
	for (i=0; i<n; i++) {

		fscanf(f2, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%c,%d,%d\n",
			&hhld_num,&pers_num,&trip_num,&start_time,&mode_prime,&purp_orig,&pd_or
			ig,&gta96_orig,&purp_dest,&pd_dest,&gta96_dest,&trip_km,&ch_num,&ch_tr_nu
			m,&ch_ty,&subch_num,&subch_num_tr);

		if (y > 1) {
			y--;
			ch_ty = t;
			subch_num = x;
			subch_num_tr = y;
		}

		if (subch_num_tr && y == 1) {
			t = ch_ty;
			x = subch_num;
			y = subch_num_tr;
		}

		if (!subch_num_tr && y == 1) {
			subch_num = 0;
		}

		fprintf(f3, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%c,%d,%d\n",
			hhld_num,pers_num,trip_num,start_time,mode_prime,purp_orig,pd_orig,gta96_o
			rig,purp_dest,pd_dest,gta96_dest,trip_km,ch_num,ch_tr_num,ch_ty,subch_num,
			subch_num_tr);
	}

	fclose(f2);
	rewind(f3);

	for (i=0; i<n; i++) {
		fgets(str[i], 60, f3);
	}

	rewind(f3);
	fprintf(f3, "hhld_num,pers_num,trip_num,start_time,mode_prime,purp_orig,pd_orig,gta96_orig,
		purp_dest,pd_dest,gta96_dest,trip_km,ch_num,ch_tr_num,ch_ty,subch_num,subch_tr_n
		um\n");

	for (i=n; i>0; i--) {
		fputs(str[i-1], f3);
	}

	fclose(f3);

}
```

## Program 2: Creation of Joint Activity Variables

```
#include "stdio.h"
#include "string.h"
#include "stdlib.h"

void main(void)
{
	FILE *f1, *f2, *f3;
	int hhld_num, pers_num, trip_num, start_time, pd_orig, gta96_orig, pd_dest, gta96_dest, trip_km,
		hhld_chk, pers, p1, t1, p2,t2,i,j,max_trip;
	char mode_prime,purp_orig,purp_dest,tmode_accs;
	int total_dur,total_dur5,trip_start5,trip_index,trip_dur,trip_dur5,act_start5,act_index,act_dur5;
	int hhld_num2, pers_num2, trip_num2, start_time2, pd_orig2, gta96_orig2, pd_dest2,
		gta96_dest2, trip_km2;
	char mode_prime2,purp_orig2,purp_dest2,tmode_accs2;
	int total_dur2, total_dur52, trip_start52, trip_index2, trip_dur2, trip_dur52, act_start52, act_index2,
		act_dur52;
	int (*hhld_array)[9][99][10] = new int[2][9][99][10];
	char str[300];

	if((f1 = fopen("Activity Durations Final.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}

	if((f2 = fopen("Activity Durations Final2.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}

	if((f3 = fopen("Joint and Durations.txt","w")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}

	fscanf(f1, "%s\n", str);
	fscanf(f2, "%s\n", str);

	fscanf(f1, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%d,%d,%c,%d,%d,%d,%d,
		%d\n",   &hhld_num, &pers_num, &trip_num, &start_time, &mode_prime, &purp_orig,
		&pd_orig, &gta96_orig, &purp_dest, &pd_dest, &gta96_dest, &trip_km, &total_dur,
		&total_dur5, &trip_start5, &trip_index, &tmode_accs, &trip_dur, &trip_dur5, &act_start5,
		&act_index, &act_dur5);

	hhld_array[0][pers_num][trip_num][0] = mode_prime;
	hhld_array[0][pers_num][trip_num][1] = trip_start5;
	hhld_array[0][pers_num][trip_num][2] = purp_dest;
	hhld_array[0][pers_num][trip_num][3] = gta96_dest;
	hhld_array[0][pers_num][trip_num][4] = act_start5;
	hhld_array[0][pers_num][trip_num][5] = purp_orig;
	hhld_array[0][pers_num][trip_num][6] = 0;
	hhld_array[0][pers_num][trip_num][7] = 0;
	hhld_array[0][pers_num][trip_num][8] = 0;
	hhld_array[0][pers_num][trip_num][9] = 0;
```

```
fprintf(f3, "hhld_num, pers_num, trip_num, start_time, mode_prime, purp_orig, pd_orig,
gta96_orig, purp_dest, pd_dest, gta96_dest, trip_km, total_dur, total_dur5, trip_start5, trip_index,
tmode_accs, trip_dur,trip_dur5,act_start5,act_index,act_dur5,jt_pers,jt_trip\n");

do {

        hhld_chk = hhld_num;
        max_trip = 0;
        i=0;

        while(hhld_num == hhld_chk && !feof(f1)) {

                pers = pers_num;
                i++;

                if(trip_num > max_trip) max_trip = trip_num;

                fscanf(f1, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%d,%d,%c,
                        %d,%d,%d,%d,%d\n", &hhld_num, &pers_num, &trip_num, &start_time,
                        &mode_prime, &purp_orig, &pd_orig, &gta96_orig,&purp_dest, pd_dest,
                        &gta96_dest, &trip_km, &total_dur, &total_dur5, &trip_start5, trip_index,
                        &tmode_accs, &trip_dur, &trip_dur5, &act_start5, &act_index,
                        &act_dur5);

                if(mode_prime == 'P') mode_prime = 'D';

                if(hhld_num == hhld_chk) {
                        hhld_array[0][pers_num][trip_num][0] = mode_prime;
                        hhld_array[0][pers_num][trip_num][1] = trip_start5;
                        hhld_array[0][pers_num][trip_num][2] = purp_dest;
                        hhld_array[0][pers_num][trip_num][3] = gta96_dest;
                        hhld_array[0][pers_num][trip_num][4] = act_start5;
                        hhld_array[0][pers_num][trip_num][5] = purp_orig;
                        hhld_array[0][pers_num][trip_num][6] = 0;
                        hhld_array[0][pers_num][trip_num][7] = 0;
                        hhld_array[0][pers_num][trip_num][8] = 0;
                        hhld_array[0][pers_num][trip_num][9] = 0;
                }

                else {
                        hhld_array[1][pers_num][trip_num][0] = mode_prime;
                        hhld_array[1][pers_num][trip_num][1] = trip_start5;
                        hhld_array[1][pers_num][trip_num][2] = purp_dest;
                        hhld_array[1][pers_num][trip_num][3] = gta96_dest;
                        hhld_array[1][pers_num][trip_num][4] = act_start5;
                        hhld_array[1][pers_num][trip_num][5] = purp_orig;
                        hhld_array[1][pers_num][trip_num][6] = 0;
                        hhld_array[1][pers_num][trip_num][7] = 0;
                        hhld_array[1][pers_num][trip_num][8] = 0;
                        hhld_array[1][pers_num][trip_num][9] = 0;
                }
        }

        if(pers > 1) {

                for(p1=1; p1<pers; p1++) {
```

```
                for(t1=1; t1<max_trip+1; t1++) {

                    for(p2=p1+1; p2<pers+1; p2++) {

                        for(t2=1; t2<max_trip+1; t2++) {

                            if(hhld_array[0][p1][t1][0] ==
                                hhld_array[0][p2][t2][0] &&
                                hhld_array[0][p1][t1][1] ==
                                hhld_array[0][p2][t2][1] &&
                                hhld_array[0][p1][t1][2] ==
                                hhld_array[0][p2][t2][2] &&
                                hhld_array[0][p1][t1][3] ==
                                hhld_array[0][p2][t2][3] &&
                                (hhld_array[0][p1][t1][2] == 'O' ||
                                hhld_array[0][p1][t1][2] == 'M')) {

                                hhld_array[0][p1][t1][6] = p2;
                                hhld_array[0][p1][t1][7] = t2;
                                hhld_array[0][p2][t2][6] = p1;
                                hhld_array[0][p2][t2][7] = t1;
                            }
                        }
                    }
                }
            }
        }

        for(j=0; j<i; j++) {

            fscanf(f2, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%d,%d,%c,%d,
                %d,%d,%d,%d\n",&hhld_num2, &pers_num2, &trip_num2, &start_time2,
                &mode_prime2, &purp_orig2, &pd_orig2, &gta96_orig2, &purp_dest2,
                &pd_dest2, &gta96_dest2, &trip_km2, &total_dur2, &total_dur52,
                &trip_start52, &trip_index2, &tmode_accs2, &trip_dur2, &trip_dur52,
                &act_start52, &act_index2, &act_dur52);

            if(pers > 1) fprintf(f3, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,
                %d,%d,%c,%d,%d,%d,%d,%d,%d,%d\n", hhld_num2,pers_num2,
                trip_num2,start_time2,mode_prime2,purp_orig2,pd_orig2,gta96_orig2,pu
                rp_dest2,pd_dest2,gta96_dest2,trip_km2,total_dur2,total_dur52,trip_start
                52,trip_index2,tmode_accs2,trip_dur2,trip_dur52,act_start52,act_index2,
                act_dur52,hhld_array[0][pers_num2][trip_num2][6],hhld_array[0][pers_nu
                m2][trip_num2][7]);

            else fprintf(f3, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%d,%d,
                %c,%d,%d,%d,%d,%d,%d,%d\n", hhld_num2, pers_num2, trip_num2,
                start_time2, mode_prime2, purp_orig2, pd_orig2, gta96_orig2,
                purp_dest2, pd_dest2, gta96_dest2, trip_km2, total_dur2, total_dur52,
                trip_start52, trip_index2, tmode_accs2, trip_dur2, trip_dur52, act_start52,
                act_index2,act_dur52,0,0);
        }

        for(p1=1; p1<pers+1; p1++) {
```

```
                    for(t1=1; t1<max_trip+1; t1++) {

                              hhld_array[0][p1][t1][0] = 0;
                              hhld_array[0][p1][t1][1] = 0;
                              hhld_array[0][p1][t1][2] = 0;
                              hhld_array[0][p1][t1][3] = 0;
                              hhld_array[0][p1][t1][4] = 0;
                              hhld_array[0][p1][t1][5] = 0;
                              hhld_array[0][p1][t1][6] = 0;
                              hhld_array[0][p1][t1][7] = 0;
                              hhld_array[0][p1][t1][8] = 0;
                              hhld_array[0][p1][t1][9] = 0;
                    }
          }

          hhld_array[0][pers_num][trip_num][0] = hhld_array[1][pers_num][trip_num][0];
          hhld_array[0][pers_num][trip_num][1] = hhld_array[1][pers_num][trip_num][1];
          hhld_array[0][pers_num][trip_num][2] = hhld_array[1][pers_num][trip_num][2];
          hhld_array[0][pers_num][trip_num][3] = hhld_array[1][pers_num][trip_num][3];
          hhld_array[0][pers_num][trip_num][4] = hhld_array[1][pers_num][trip_num][4];
          hhld_array[0][pers_num][trip_num][5] = hhld_array[1][pers_num][trip_num][5];
          hhld_array[0][pers_num][trip_num][6] = hhld_array[1][pers_num][trip_num][6];
          hhld_array[0][pers_num][trip_num][7] = hhld_array[1][pers_num][trip_num][7];
          hhld_array[0][pers_num][trip_num][8] = hhld_array[1][pers_num][trip_num][8];
          hhld_array[0][pers_num][trip_num][9] = hhld_array[1][pers_num][trip_num][9];

     } while(!feof(f1));

     fscanf(f2, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%d,%d,%c,%d,%d,%d,%d,
          %d\n", &hhld_num2, &pers_num2, &trip_num2, &start_time2, &mode_prime2,
          &purp_orig2, &pd_orig2, &gta96_orig2, &purp_dest2, &pd_dest2, &gta96_dest2,
          &trip_km2, &total_dur2, &total_dur52, &trip_start52, &trip_index2, &tmode_accs2,
          &trip_dur2, &trip_dur52, &act_start52, &act_index2, &act_dur52);

     if(pers > 1) fprintf(f3, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%d,%d,%c,%d,%d,
          %d,%d,%d,%d,%d\n", hhld_num2,pers_num2,   trip_num2, start_time2, mode_prime2,
          purp_orig2, pd_orig2, gta96_orig2, purp_dest2, pd_dest2, gta96_dest2, trip_km2,
          total_dur2, total_dur52, trip_start52, trip_index2, tmode_accs2, trip_dur2, trip_dur52,
          act_start52, act_index2, act_dur52, hhld_array[0][pers_num2][trip_num2][6],
          hhld_array[0][pers_num2][trip_num2][7]);

     else fprintf(f3, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%d,%d,%c,%d,%d,%d,
          %d,%d,%d,%d\n", hhld_num2, pers_num2, trip_num2, start_time2, mode_prime2,
          purp_orig2, pd_orig2, gta96_orig2, purp_dest2, pd_dest2, gta96_dest2, trip_km2,
          total_dur2, total_dur52, trip_start52, trip_index2, tmode_accs2, trip_dur2, trip_dur52,
          act_start52, act_index2, act_dur52,0,0);

     fclose(f1);
     fclose(f2);
}
```

## Program 3: Creation of Activity Episode Variables

```
#include "stdio.h"
#include "string.h"
#include "stdlib.h"

void main(void)
{
	FILE *f1, *f2, *f3, *f4, *f5, *f6, *f7;
	int hhld_num1, pers_num1, trip_num1, start_time1, pd_orig1, gta96_orig1, pd_dest1,
		gta96_dest1, trip_km1;
	char mode_prime1,purp_orig1,purp_dest1;
	int hhld_num2, pers_num2, trip_num2, start_time2, pd_orig2, gta96_orig2, pd_dest2,
		gta96_dest2, trip_km2;
	char mode_prime2,purp_orig2,purp_dest2;
	int trans_orig, trans_dest, tst1_int, tst2_int, tst1_min, tst2_min, total_dur, trip_start5, trip_index,
		act_index,b_tt, ab_tt, twgo_tt, ago_tt, act_start5, act_dur5, car_orig, car_dest, car_tt1,
		car_tt2, car_tt3, car_tt4, car_tt5, car_tt6, car_tt7, orig, dest, trip_dur, total_dur5, trip_dur5,
		go_time, go_orig, go_dest;
	int (*trans_array)[4] = new int[188791][4];
	int (*car_array)[7] = new int[188791][7];
	int (*index_array)[4000] = new int[4000][4000];
	int (*trans2_array)[4000] = new int[4000][4000];
	int thhld_num, tpers_num, ttrip_num, tpd_accs, tgta96_accs, tpd_egrs, tgta96_egrs;
	char troute_1[4], tmode_accs, tmode_egrs;
	int i, index, car_index, trans_index, master_index;
	int trans2_orig, trans2_dest;
	float trans2_tt;

	if((f1 = fopen("trip96.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}
	if((f2 = fopen("auto_time_index.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}
	if((f3 = fopen("transit_time_index.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}
	if((f4 = fopen("activity_duration.txt","w")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}
	if((f5 = fopen("tran96.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}
	if((f6 = fopen("od_pairs_index.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
	}
	if((f7 = fopen("transit_time.txt","r")) == NULL) {
		printf("Cannot open file.\n");
		exit(1);
```

```
}

while(fscanf(f2, "%d,%d,%d,%d,%d,%d,%d,%d,%d,%d\n", &car_index, &car_orig, &car_dest,
        &car_tt1, &car_tt2, &car_tt3, &car_tt4, &car_tt5, &car_tt6, &car_tt7) != EOF) {

        car_array[car_index-1][0] = car_tt1;
        car_array[car_index-1][1] = car_tt2;
        car_array[car_index-1][2] = car_tt3;
        car_array[car_index-1][3] = car_tt4;
        car_array[car_index-1][4] = car_tt5;
        car_array[car_index-1][5] = car_tt6;
        car_array[car_index-1][6] = car_tt7;
}

while(fscanf(f3, "%d,%d,%d,%d,%d,%d,%d\n", &trans_index, &trans_orig, &trans_dest, &b_tt,
        &ab_tt, &twgo_tt, &ago_tt) != EOF) {

        trans_array[trans_index-1][0] = b_tt;
        trans_array[trans_index-1][1] = ab_tt;
        trans_array[trans_index-1][2] = twgo_tt;
        trans_array[trans_index-1][3] = ago_tt;
}

while(fscanf(f6, "%d,%d,%d\n", &master_index, &orig, &dest) != EOF) index_array[orig-1][dest-1]
        = master_index;

while(fscanf(f7, "%d %d %f\n", &trans2_orig, &trans2_dest, &trans2_tt) != EOF)
        trans2_array[trans2_orig-1][trans2_dest-1] = trans2_tt + 0.5;

fscanf(f1, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d\n", &hhld_num1, &pers_num1,
        &trip_num1, &start_time1,&mode_prime1, &purp_orig1, &pd_orig1, &gta96_orig1,
        &purp_dest1, &pd_dest1,&gta96_dest1,&trip_km1);

fprintf(f4, "hhld_num, pers_num, trip_num, start_time, mode_prime, purp_orig, pd_orig,
        gta96_orig, purp_dest, pd_dest, gta96_dest, trip_km, total_dur, total_dur5, trip_start5,
        trip_index, tmode_accs, trip_dur, trip_dur5, act_start5, act_index, act_dur5\n");

for(i=0; i<500313; i++) {

        index = index_array[gta96_orig1-1][gta96_dest1-1];

        if(hhld_num1 != hhld_num2 && pers_num1 != pers_num2) go_time = 999;

        fscanf(f1, "%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d\n", &hhld_num2, &pers_num2,
                &trip_num2, &start_time2,&mode_prime2, &purp_orig2, &pd_orig2,
                &gta96_orig2, &purp_dest2, &pd_dest2, &gta96_dest2, &trip_km2);

        tst1_int = (start_time1/100);
        tst1_min = (tst1_int*60) + (start_time1 - (tst1_int*100));
        tst2_int = (start_time2/100);
        tst2_min = (tst2_int*60) + (start_time2 - (tst2_int*100));

        if(feof(f1)) total_dur = 1680 - tst1_min;
        else if(hhld_num1 == hhld_num2 && pers_num1 == pers_num2) total_dur = tst2_min –
                tst1_min;
        else total_dur = 1680 - tst1_min;
```

```
total_dur5 = ((int)((float)total_dur+2.5)/5)*5;

trip_start5 = ((int)(((float)tst1_min - 240)+2.5)/5)*5;

trip_index = ((trip_start5/5)+1);

if(mode_prime1 == 'B' || mode_prime1 == 'G' || mode_prime1 == 'J') {
        fscanf(f5, "%d,%d,%d,%c,%d,%d,%c,%d,%d,%s\n", &thhld_num, &tpers_num,
        &ttrip_num, &tmode_accs, &tpd_accs,&tgta96_accs, &tmode_egrs, &tpd_egrs,
        &tgta96_egrs, &troute_1);
}
else tmode_accs = 'X';

if(mode_prime1 == 'B') {
        if(gta96_orig1 == gta96_dest1 && trip_km1 > 0) trip_dur =
                (((float)trip_km1/15)*60*1.4)+0.5;
        else if(gta96_orig1 == gta96_dest1 && trip_km1 == 0) trip_dur =
                ((0.25/15)*60*1.4)+0.5;
        else trip_dur = trans_array[index-1][0];
        if(trip_dur == 999 && (tmode_accs == 'D' || tmode_accs == 'M' || tmode_accs ==
                'P' || tmode_accs == 'T'))
                trip_dur = trans_array[index-1][1];
        if(trip_dur == 999) {
                trip_dur = trans2_array[tgta96_accs-1][gta96_dest1-1];
                if(trip_dur > 1000 || trip_dur < -1000) trip_dur = trans2_array[gta96_orig1-
                        1][tgta96_egrs-1];
                if(trip_dur > 1000 || trip_dur < -1000) trip_dur =
                        trans2_array[tgta96_accs-1][tgta96_egrs-1];
        }
        if(trip_dur > 1000 || trip_dur < -1000) trip_dur = 999;
}

else if(mode_prime1 == 'G') {
        if(tmode_accs == 'C' || tmode_accs == 'O' || tmode_accs == 'W' || tmode_accs ==
                '9') {
                trip_dur = trans_array[index-1][2];
                if(trip_dur > 90 && trip_dur != 999) trip_dur = 90;
        }
        else if (tmode_accs == 'D' || tmode_accs == 'M' || tmode_accs == 'P' ||
                tmode_accs == 'T') {
                trip_dur = trans_array[index-1][3];
                if(trip_dur > 90 && trip_dur != 999) trip_dur = 90;
        }
        if(trip_dur != 999) {
                go_time = trip_dur;
                go_orig = gta96_orig1;
                go_dest = gta96_dest1;
        }
        else if(trip_dur == 999 && go_orig == gta96_dest1 && go_dest == gta96_orig1)
                trip_dur = go_time;
}

else if(mode_prime1 == 'J') {
        trip_dur = trans_array[index-1][2];
        if(trip_dur > 90 && trip_dur != 999) trip_dur = 90;
```

```
		if(trip_dur != 999) {
			go_time = trip_dur;
			go_orig = gta96_orig1;
			go_dest = gta96_dest1;
		}
		else if(trip_dur == 999 && go_orig == gta96_dest1 && go_dest == gta96_orig1)
			trip_dur = go_time;
	}

	else if(mode_prime1 == 'C' || mode_prime1 == 'O') {
		if(trip_km1 == 0) trip_dur = ((0.25/15)*60*1.4)+0.5;
		else trip_dur = (((float)trip_km1/15)*60*1.4)+0.5;
	}

	else if(mode_prime1 == 'W') {
		if(trip_km1 == 0) trip_dur = ((0.25/4)*60*1.4)+0.5;
		else trip_dur = (((float)trip_km1/4)*60*1.4)+0.5;
	}

	else if(mode_prime1 == '9') trip_dur = 999;

	else if(mode_prime1 == 'D' || mode_prime1 == 'P' || mode_prime1 == 'T' || mode_prime1
		== 'S' || mode_prime1 == 'M') {

		if(gta96_orig1 == gta96_dest1 && trip_km1 > 0) trip_dur =
			(((float)trip_km1/25)*60*1.4)+0.5;
		else if(gta96_orig1 == gta96_dest1 && trip_km1 == 0) trip_dur =
			((0.25/15)*60*1.4)+0.5;

		else if(start_time1 >= 400 && start_time1 < 600) trip_dur = car_array[index-1][0];
		else if(start_time1 >= 600 && start_time1 < 700) trip_dur = car_array[index-1][1];
		else if(start_time1 >= 700 && start_time1 < 900) trip_dur = car_array[index-1][2];
		else if(start_time1 >= 900 && start_time1 < 1500) trip_dur = car_array[index-
			1][3];
		else if(start_time1 >= 1500 && start_time1 < 1800) trip_dur = car_array[index-
			1][4];
		else if(start_time1 >= 1800 && start_time1 < 2000) trip_dur = car_array[index-
			1][5];
		else if(start_time1 >= 2000 && start_time1 < 2400) trip_dur = car_array[index-
			1][6];
		else if(start_time1 >= 2400 && start_time1 < 2800) trip_dur = car_array[index-
			1][0];

		if(trip_dur == 999 && gta96_orig1 >= 3000 && gta96_dest1 >= 3000) {
			if(trip_km1 > 0) trip_dur = (((float)trip_km1/25)*60*1.4)+0.5;
			else trip_dur = ((0.25/15)*60*1.4)+0.5;
		}
	}

	if(trip_dur == 999) trip_dur5 = 999;
	else trip_dur5 = ((int)((float)trip_dur+2.5)/5)*5;

	if(trip_dur5 == 999) act_start5 = trip_start5;
	else act_start5 = trip_start5 + trip_dur5;

	act_index = ((act_start5/5)+1);
```

```
			if(trip_dur == 999) act_dur5 = total_dur5;
			else act_dur5 = total_dur5 - trip_dur5;

		fprintf(f4,"%d,%d,%d,%d,%c,%c,%d,%d,%c,%d,%d,%d,%d,%d,%d,%d,%c,%d,%d,%d,%d,%d\n",
			hhld_num1,pers_num1,trip_num1,start_time1,mode_prime1,purp_orig1,pd_orig1,gta96_
			orig1,purp_dest1,pd_dest1,gta96_dest1,trip_km1,total_dur,total_dur5,trip_start5,trip_inde
			x,tmode_accs,trip_dur,trip_dur5,act_start5,act_index,act_dur5);

			hhld_num1 = hhld_num2;
			pers_num1 = pers_num2;
			trip_num1 = trip_num2;
			start_time1 = start_time2;
			mode_prime1 = mode_prime2;
			purp_orig1 = purp_orig2;
			pd_orig1 = pd_orig2;
			gta96_orig1 = gta96_orig2;
			purp_dest1 = purp_dest2;
			pd_dest1 = pd_dest2;
			gta96_dest1 = gta96_dest2;
			trip_km1 = trip_km2;

		}
		fclose(f1);
		fclose(f2);
		fclose(f3);
		fclose(f4);
		fclose(f5);
		fclose(f6);
		fclose(f7);
}
```

## Program 4: Creation of Auto Availability Variables

```
#include "stdio.h"
#include "string.h"
#include "stdlib.h"

void main(void)
{
		FILE *f1, *f2;

		int hhld_num, trip_start5, eps_start5, pers_num, trip_num, ch_num, ch_tr_num, max_ch_tr_num,
			total_dur5, trip_dur5, jt_pers, jt_trip, n_vehicle, hhld_chk=0, n_veh,time;
		char str[1000],pers_array[9],mode_prime,purp_orig,purp_dest,driver_lic,auto_avail;

		if((f1 = fopen("Auto Availability.txt","r")) == NULL) {
			printf("Cannot open file 1.\n");
			exit(1);
		}

		if((f2 = fopen("Auto Availability Final.txt","w")) == NULL) {
			printf("Cannot open file 2.\n");
			exit(1);
```

```
        }

        fscanf(f1, "%s\n", str);

        fprintf(f2, "hhld_num, trip_start5, mode_prime, eps_start5, pers_num, trip_num, ch_num,
                ch_tr_num, MaxOfch_tr_num, purp_orig, purp_dest, total_dur5, trip_dur5, jt_pers, jt_trip,
                driver_lic, n_vehicle, auto_avail, n_veh_at_home\n");

        while(fscanf(f1, "%d,%d,%c,%d,%d,%d,%d,%d,%d,%c,%c,%d,%d,%d,%d,%c,%d\n",
                &hhld_num, &trip_start5, &mode_prime, &eps_start5, &pers_num, &trip_num, &ch_num,
                &ch_tr_num, &max_ch_tr_num, &purp_orig, &purp_dest, &total_dur5, &trip_dur5,
                &jt_pers, &jt_trip, &driver_lic, &n_vehicle) != EOF) {

                if(hhld_num != hhld_chk) n_veh = n_vehicle, time = 0;

                hhld_chk = hhld_num;

                if(ch_num == 0) auto_avail = 'X';

                if(ch_tr_num == 1) {

                        if(mode_prime != 'D' && trip_start5 < time) n_veh--;

                        if(n_veh > 0 && driver_lic == 'Y') auto_avail = 'Y';
                        else auto_avail = 'N';

                        pers_array[pers_num] = auto_avail;

                        if(mode_prime != 'D' && trip_start5 < time) n_veh++;
                }
// Does not consider auto access to transit
                else if(ch_tr_num > 1) auto_avail = pers_array[pers_num];

                if(ch_tr_num == 1 && mode_prime == 'D') n_veh--;

                if(ch_tr_num == max_ch_tr_num && mode_prime == 'D' && purp_dest == 'H') {
                        n_veh++;
                        time = eps_start5;
                }

                if (n_veh < 0 || n_veh > n_vehicle) auto_avail = 'X';

                fprintf(f2, "%d,%d,%c,%d,%d,%d,%d,%d,%d,%c,%c,%d,%d,%d,%d,%c,%d,%c,%d\n",
                        hhld_num,trip_start5,mode_prime,eps_start5,pers_num,trip_num,ch_num,ch_tr_
                        num,max_ch_tr_num,purp_orig,purp_dest,total_dur5,trip_dur5,jt_pers,
                        jt_trip,driver_lic,n_vehicle,auto_avail,n_veh);

        }
}
```

# APPENDIX D

## DURATION, START TIME & FREQUENCY T–TEST RESULTS

Following are the mean, standard deviation, number of observations and result of the statistical t–tests of each selected explanatory variable as compared to the overall value for the duration, start time and frequency of each activity type. Note that, here, frequency is calculated assuming participation in at least one episode.

### Work Activity Duration

| Variable | Overall | General Offi | Manufacturii | Professiona | Retail Sales | Full Time W | Part Time W | Work at Hon |
|---|---|---|---|---|---|---|---|---|
| Mean | 511.5 | 494.0 | 541.8 | 514.0 | 490.0 | 531.3 | 395.7 | 371.0 |
| S.D. | 146.8 | 120.2 | 140.8 | 145.0 | 161.3 | 130.8 | 164.5 | 230.2 |
| N | 84314 | 11702 | 19344 | 34657 | 18059 | 72594 | 9769 | 1648 |
| T-stat | 0.00 | 12.33 | -26.06 | -2.69 | 17.57 | -28.04 | 72.83 | 37.96 |

| Variable | 15 to 19 | 20 to 24 | 25 to 29 | 30 to 34 | 35 to 39 | 40 to 44 | 45 to 49 | 50 to 54 |
|---|---|---|---|---|---|---|---|---|
| Mean | 360.7 | 495.9 | 525.6 | 520.9 | 515.8 | 514.1 | 515.2 | 514.8 |
| S.D. | 168.4 | 155.7 | 137.2 | 140.8 | 144.1 | 146.1 | 144.2 | 142.5 |
| N | 1422 | 6043 | 11105 | 13996 | 13511 | 11912 | 10613 | 7554 |
| T-stat | 38.32 | 7.94 | -9.58 | -7.06 | -3.18 | -1.84 | -2.46 | -1.87 |

| Variable | 55 to 59 | 60 to 64 | 65+ | Male | Female | PD 1 | PD 2,3,4,6 | PD 5,7,8,9,1 |
|---|---|---|---|---|---|---|---|---|
| Mean | 509.0 | 498.5 | 445.0 | 533.1 | 484.3 | 509.4 | 501.0 | 520.8 |
| S.D. | 143.4 | 155.9 | 173.1 | 151.1 | 136.4 | 138.3 | 157.3 | 133.0 |
| N | 4532 | 2302 | 1307 | 46972 | 37339 | 13825 | 8596 | 10919 |
| T-stat | 1.11 | 4.17 | 16.21 | -25.31 | 30.44 | 1.58 | 6.25 | -6.29 |

| Variable | PD 11,12,13 | PD 17 to 12 | With Childre | No Children | Full Time St | Part Time Student |
|---|---|---|---|---|---|---|
| Mean | 511.3 | 515.0 | 511.3 | 511.6 | 321.8 | 497.9 |
| S.D. | 140.4 | 147.2 | 147.5 | 146.5 | 144.5 | 142.6 |
| N | 9750 | 38618 | 24936 | 59378 | 1749 | 3612 |
| T-stat | 0.14 | -3.92 | 0.20 | -0.11 | 53.52 | 5.46 |

### School Activity Duration

| Variable | Overall | Full Time S | Part Time S | 11 to 15 | 16 to 18 | 19 to 25 | 26 to 30 | 31+ |
|---|---|---|---|---|---|---|---|---|
| Mean | 393.6 | 404.7 | 258.9 | 415.5 | 404.5 | 377.1 | 353.5 | 290.1 |
| S.D. | 115.4 | 104.5 | 141.5 | 70.6 | 88.3 | 158.4 | 174.0 | 158.8 |
| N | 29633 | 27584 | 1810 | 13088 | 7260 | 5996 | 1159 | 2130 |
| T-stat | 0.00 | -11.97 | 47.54 | -20.15 | -7.52 | 9.44 | 11.34 | 38.85 |

| Variable | Full Time V | Part Time V | Work at Home |
|---|---|---|---|
| Mean | 213.9 | 370.4 | 315.3 |
| S.D. | 124.8 | 137.2 | 156.6 |
| N | 754 | 4515 | 87 |
| T-stat | 42.14 | 12.25 | 6.32 |

## Market Episode Duration

| Variable | Overall | Driver's Lic | No Driver's | Full Time W | Part Time W | Not Employ | Work at Ho | 1 Person |
|---|---|---|---|---|---|---|---|---|
| Mean | 79.2 | 75.2 | 102.3 | 65.0 | 77.4 | 89.2 | 72.5 | 81.6 |
| S.D. | 77.7 | 75.0 | 87.8 | 65.6 | 69.8 | 84.7 | 73.4 | 80.6 |
| N | 35679 | 30384 | 5294 | 12293 | 3402 | 18865 | 1099 | 4659 |
| T-stat | 0.00 | 6.74 | -19.82 | 18.17 | 1.30 | -13.80 | 2.81 | -1.98 |

| Variable | 2 Persons | 3 Persons | 4 Persons | 5+ Persons | 11 to 19 | 20 to 29 | 30 to 39 | 40 to 49 |
|---|---|---|---|---|---|---|---|---|
| Mean | 81.3 | 78.9 | 74.8 | 78.0 | 89.4 | 85.2 | 70.7 | 69.0 |
| S.D. | 77.8 | 77.3 | 77.0 | 75.4 | 82.7 | 84.0 | 71.4 | 68.7 |
| N | 12871 | 6597 | 7024 | 4528 | 1351 | 3596 | 7950 | 6803 |
| T-stat | -2.62 | 0.29 | 4.34 | 0.96 | -4.72 | -4.37 | 8.99 | 10.16 |

| Variable | 50 to 59 | 60+ |
|---|---|---|
| Mean | 78.3 | 88.7 |
| S.D. | 73.6 | 84.2 |
| N | 4688 | 11221 |
| T-stat | 0.72 | -11.01 |

## Other Episode Duration

| Variable | Overall | Full Time W | Part Time W | Not Employe | Work at Hor | Driver's Lice | No Driver's I | With Childre |
|---|---|---|---|---|---|---|---|---|
| Mean | 134.4 | 125.1 | 131.0 | 143.6 | 125.6 | 130.4 | 154.7 | 111.8 |
| S.D. | 140.2 | 131.7 | 137.3 | 146.6 | 146.0 | 138.0 | 149.4 | 129.2 |
| N | 62396 | 24395 | 6537 | 29283 | 2143 | 52148 | 10246 | 15204 |
| T-stat | 0.00 | 8.96 | 1.89 | -9.12 | 2.87 | 4.83 | -13.45 | 18.15 |

| Variable | No Children | 11 to 19 | 20 to 29 | 30 to 39 | 40 to 49 | 50 to 59 | 60+ |
|---|---|---|---|---|---|---|---|
| Mean | 141.7 | 147.3 | 156.1 | 120.0 | 119.6 | 134.6 | 140.2 |
| S.D. | 142.9 | 140.8 | 149.3 | 132.3 | 134.2 | 144.1 | 141.0 |
| N | 47192 | 5832 | 9678 | 14006 | 11182 | 7453 | 14245 |
| T-stat | -8.47 | -6.72 | -14.05 | 11.15 | 10.35 | -0.07 | -4.41 |

## Work Activity Start Time

| Variable | Overall | General Offi | Manufacturi | Professiona | Retail Sales | Full Time W | Part Time W | Work at Hon |
|---|---|---|---|---|---|---|---|---|
| Mean | 315.8 | 306.3 | 298.5 | 301.8 | 364.8 | 296.6 | 445.0 | 452.5 |
| S.D. | 186.2 | 152.5 | 224.8 | 149.2 | 212.7 | 170.2 | 238.2 | 215.0 |
| N | 84314 | 11702 | 19344 | 34657 | 18059 | 72594 | 9769 | 1648 |
| T-stat | 0.00 | 5.31 | 11.18 | 12.44 | -31.26 | 21.18 | -62.86 | -29.42 |

| Variable | 15 to 19 | 20 to 24 | 25 to 29 | 30 to 34 | 35 to 39 | 40 to 44 | 45 to 49 | 50 to 54 |
|---|---|---|---|---|---|---|---|---|
| Mean | 604.8 | 383.6 | 321.4 | 310.8 | 305.2 | 302.0 | 298.5 | 292.8 |
| S.D. | 242.9 | 233.3 | 180.2 | 180.8 | 172.9 | 170.8 | 173.7 | 168.4 |
| N | 1422 | 6043 | 11105 | 13996 | 13511 | 11912 | 10613 | 7554 |
| T-stat | -57.68 | -26.84 | -2.99 | 2.99 | 6.25 | 7.67 | 9.10 | 10.37 |

| Variable | 55 to 59 | 60 to 64 | 65+ | Male | Female | With Childre | No Children | Full Time St |
|---|---|---|---|---|---|---|---|---|
| Mean | 292.4 | 294.4 | 319.9 | 308.4 | 325.2 | 314.5 | 316.4 | 665.8 |
| S.D. | 175.0 | 177.5 | 156.9 | 195.6 | 173.3 | 188.7 | 185.2 | 207.2 |
| N | 4532 | 2302 | 1307 | 46972 | 37339 | 24936 | 59378 | 1749 |
| T-stat | 8.29 | 5.46 | -0.79 | 6.85 | -8.29 | 0.96 | -0.55 | -77.60 |

| Variable | Part Time Student |
|---|---|
| Mean | 331.4 |
| S.D. | 187.7 |
| N | 3612 |
| T-stat | -4.92 |

## School Activity Start Time

| Variable | Overall | Full Time S | Part Time S | 11 to 15 | 16 to 18 | 19 to 25 | 26 to 30 | 31+ |
|---|---|---|---|---|---|---|---|---|
| Mean | 309.6 | 291.1 | 563.2 | 268.5 | 276.0 | 363.8 | 425.7 | 461.4 |
| S.D. | 124.9 | 82.2 | 260.6 | 37.6 | 60.2 | 148.5 | 215.2 | 236.5 |
| N | 29633 | 27584 | 1810 | 13088 | 7260 | 5996 | 1159 | 2130 |
| T-stat | 0.00 | 20.84 | -76.77 | 36.95 | 22.31 | -29.62 | -29.94 | -50.00 |

| Variable | Full Time W | Part Time W | Work at Home |
|---|---|---|---|
| Mean | 687.4 | 335.3 | 419.8 |
| S.D. | 249.5 | 144.1 | 229.9 |
| N | 754 | 4515 | 87 |
| T-stat | -79.13 | -12.59 | -8.18 |

## Market Episode Start Time

| Variable | Overall | Driver's Lic | No Driver's | Full Time W | Part Time W | Not Employ | Work at Ho | 1 Person |
|---|---|---|---|---|---|---|---|---|
| Mean | 632.61 | 633.13 | 629.57 | 743.63 | 640.92 | 559.03 | 629.55 | 608.91 |
| S.D. | 212.03 | 213.29 | 204.64 | 203.97 | 208.25 | 184.40 | 211.03 | 199.09 |
| N | 35679 | 30384 | 5294 | 12293 | 3402 | 18865 | 1099 | 4659 |
| T-stat | 0.00 | -0.31 | 0.98 | -50.55 | -2.19 | 40.28 | 0.47 | 7.23 |

| Variable | 2 Persons | 3 Persons | 4 Persons | 5+ Persons | 11 to 19 | 20 to 29 | 30 to 39 | 40 to 49 |
|---|---|---|---|---|---|---|---|---|
| Mean | 594.06 | 655.08 | 674.63 | 668.63 | 798.02 | 716.62 | 682.94 | 673.80 |
| S.D. | 199.36 | 214.66 | 220.08 | 220.63 | 161.49 | 208.24 | 219.99 | 216.64 |
| N | 12871 | 6597 | 7024 | 4528 | 1351 | 3596 | 7950 | 6803 |
| T-stat | 17.96 | -7.89 | -15.09 | -10.72 | -28.36 | -22.68 | -19.01 | -14.63 |

| Variable | 50 to 59 | 60+ |
|---|---|---|
| Mean | 630.84 | 525.34 |
| S.D. | 206.80 | 158.48 |
| N | 4688 | 11221 |
| T-stat | 0.54 | 49.42 |

## Other Episode Start Time

| Variable | Overall | Full Time V | Part Time \ | Not Employ | Work at Hc | Driver's Lic | No Driver's | With Childr |
|---|---|---|---|---|---|---|---|---|
| Mean | 689.92 | 770.53 | 718.82 | 617.97 | 669.53 | 688.98 | 694.67 | 692.50 |
| S.D. | 249.95 | 238.42 | 255.48 | 235.95 | 252.48 | 252.55 | 236.22 | 245.53 |
| N | 62396 | 24395 | 6537 | 29283 | 2143 | 52148 | 10246 | 15204 |
| T-stat | 0.00 | -43.26 | -8.87 | 41.36 | 3.71 | 0.63 | -1.80 | -1.15 |

| Variable | No Childrei | 11 to 19 | 20 to 29 | 30 to 39 | 40 to 49 | 50 to 59 | 60+ |
|---|---|---|---|---|---|---|---|
| Mean | 689.09 | 807.07 | 779.31 | 714.63 | 700.70 | 654.78 | 566.85 |
| S.D. | 251.35 | 202.46 | 253.80 | 247.53 | 246.71 | 245.45 | 214.58 |
| N | 47192 | 5832 | 9678 | 14006 | 11182 | 7453 | 14245 |
| T-stat | 0.54 | -34.74 | -32.67 | -10.59 | -4.21 | 11.49 | 54.37 |

## Work Episode Frequency

| Variable | Overall | General Of | Manufactui | Profession | Retail Sale | Full Time V | Part Time \ | Work at Hc |
|---|---|---|---|---|---|---|---|---|
| Mean | 1.15 | 1.08 | 1.11 | 1.17 | 1.17 | 1.14 | 1.10 | 1.56 |
| S.D. | 0.59 | 0.38 | 0.55 | 0.62 | 0.68 | 0.58 | 0.47 | 1.15 |
| N | 84314 | 11702 | 19344 | 34657 | 18059 | 72594 | 9769 | 1648 |
| T-stat | 0.00 | 11.24 | 7.91 | -7.74 | -4.94 | 1.04 | 7.59 | -27.54 |

| Variable | 15 to 19 | 20 to 24 | 25 to 29 | 30 to 34 | 35 to 39 | 40 to 44 | 45 to 49 | 50 to 54 |
|---|---|---|---|---|---|---|---|---|
| Mean | 1.03 | 1.07 | 1.11 | 1.14 | 1.16 | 1.16 | 1.18 | 1.16 |
| S.D. | 0.20 | 0.36 | 0.51 | 0.57 | 0.63 | 0.63 | 0.67 | 0.59 |
| N | 1422 | 6043 | 11105 | 13996 | 13511 | 11912 | 10613 | 7554 |
| T-stat | 7.67 | 9.85 | 6.02 | 1.29 | -2.49 | -2.97 | -5.40 | -1.60 |

| Variable | 55 to 59 | 60 to 64 | 65+ | Male | Female |
|---|---|---|---|---|---|
| Mean | 1.17 | 1.18 | 1.22 | 1.18 | 1.11 |
| S.D. | 0.63 | 0.66 | 0.71 | 0.67 | 0.48 |
| N | 4532 | 2302 | 1307 | 46972 | 37339 |
| T-stat | -2.47 | -3.00 | -4.73 | -8.43 | 10.88 |

## School Episode Frequency

| Variable | Overall | Full Time S | Part Time S | 11 to 15 | 16 to 18 | 19 to 25 | 26 to 30 | 31+ |
|---|---|---|---|---|---|---|---|---|
| Mean | 1.03 | 1.03 | 1.03 | 1.02 | 1.03 | 1.03 | 1.04 | 1.05 |
| S.D. | 0.18 | 0.18 | 0.18 | 0.16 | 0.19 | 0.19 | 0.21 | 0.25 |
| N | 29633 | 27584 | 1810 | 13088 | 7260 | 5996 | 1159 | 2130 |
| T-stat | 0.00 | 0.56 | 0.25 | 3.85 | -1.87 | -0.55 | -2.22 | -4.10 |

## Market Episode Frequency

| Variable | Overall | Driver's Lic | No Driver's | Full Time V | Part Time \ | Not Employ | Work at Hc | 11 to 19 |
|---|---|---|---|---|---|---|---|---|
| Mean | 1.25 | 1.27 | 1.18 | 1.22 | 1.27 | 1.27 | 1.34 | 1.15 |
| S.D. | 0.61 | 0.62 | 0.50 | 0.58 | 0.65 | 0.61 | 0.70 | 0.46 |
| N | 28480 | 23979 | 4500 | 10104 | 2687 | 14853 | 820 | 1179 |
| T-stat | 0.00 | -2.66 | 8.03 | 5.21 | -1.08 | -2.82 | -4.05 | 5.98 |

| Variable | 20 to 29 | 30 to 39 | 40 to 49 | 50 to 59 | 60+ |
|---|---|---|---|---|---|
| Mean | 1.19 | 1.24 | 1.26 | 1.28 | 1.28 |
| S.D. | 0.50 | 0.60 | 0.64 | 0.65 | 0.61 |
| N | 3072 | 6411 | 5411 | 3670 | 8737 |
| T-stat | 5.24 | 1.52 | -0.49 | -2.29 | -4.24 |

## Other Episode Frequency

| Variable | Overall | Full Time V | Part Time \ | Not Employ | Work at Hc | Driver's Lic | No Driver's | With Childr |
|---|---|---|---|---|---|---|---|---|
| Mean | 1.38 | 1.33 | 1.43 | 1.40 | 1.57 | 1.41 | 1.24 | 1.40 |
| S.D. | 0.85 | 0.74 | 0.88 | 0.92 | 1.00 | 0.89 | 0.60 | 0.85 |
| N | 45299 | 18345 | 4586 | 20975 | 1367 | 37048 | 8250 | 10883 |
| T-stat | 0.00 | 6.65 | -3.63 | -2.56 | -8.11 | -4.95 | 13.85 | -2.16 |

| Variable | No Childre | 11 to 19 | 20 to 29 | 30 to 39 | 40 to 49 | 50 to 59 | 60+ |
|---|---|---|---|---|---|---|---|
| Mean | 1.37 | 1.27 | 1.36 | 1.41 | 1.39 | 1.38 | 1.40 |
| S.D. | 0.85 | 0.67 | 0.79 | 0.85 | 0.84 | 0.84 | 0.96 |
| N | 34416 | 4597 | 7108 | 9960 | 8021 | 5412 | 10201 |
| T-stat | 1.02 | 8.41 | 1.48 | -3.06 | -1.62 | 0.02 | -1.99 |

# APPENDIX E

## DEFINITION OF PROBABILITY DISTRIBUTION CATEGORIES

Following are definitions of the distribution categories used for activity episode generation, including the distribution identification number, whether it is individual or joint, the activity type, explanatory variables, and the number of observations used in the distribution.[1] Explanatory variables are defined in DMG (1996), except as noted.

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Age | Occ | Emp Stat | | Duration | Start Time | Frequency |
| 0 | Ind | Primary W | 11 to 18 | G | F | | 9 | 9 | 9 |
| 1 | Ind | Primary W | 11 to 18 | G | P | | 76 | 76 | 76 |
| 2 | Ind | Primary W | 11 to 18 | M | F | | 47 | 47 | 47 |
| 3 | Ind | Primary W | 11 to 18 | M | P | | 84 | 84 | 84 |
| 4 | Ind | Primary W | 11 to 18 | P | F | | 12 | 12 | 12 |
| 5 | Ind | Primary W | 11 to 18 | P | P | | 26 | 26 | 26 |
| 6 | Ind | Primary W | 11 to 18 | S | F | | 44 | 44 | 44 |
| 7 | Ind | Primary W | 11 to 18 | S | P | | 536 | 536 | 536 |
| 8 | Ind | Primary W | 19 to 25 | G | F | | 1093 | 1093 | 1093 |
| 9 | Ind | Primary W | 19 to 25 | G | P | | 362 | 362 | 362 |
| 10 | Ind | Primary W | 19 to 25 | M | F | | 1553 | 1553 | 1553 |
| 11 | Ind | Primary W | 19 to 25 | M | P | | 307 | 307 | 307 |
| 12 | Ind | Primary W | 19 to 25 | P | F | | 2082 | 2082 | 2082 |
| 13 | Ind | Primary W | 19 to 25 | P | P | | 332 | 332 | 332 |
| 14 | Ind | Primary W | 19 to 25 | S | F | | 1815 | 1815 | 1815 |
| 15 | Ind | Primary W | 19 to 25 | S | P | | 1297 | 1297 | 1297 |
| 16 | Ind | Primary W | 26 to 64 | G | F | | 8710 | 8710 | 8710 |
| 17 | Ind | Primary W | 26 to 64 | G | P | | 1218 | 1218 | 1218 |
| 18 | Ind | Primary W | 26 to 64 | M | F | | 15723 | 15723 | 15723 |
| 19 | Ind | Primary W | 26 to 64 | M | P | | 958 | 958 | 958 |
| 20 | Ind | Primary W | 26 to 64 | P | F | | 29004 | 29004 | 29004 |
| 21 | Ind | Primary W | 26 to 64 | P | P | | 1920 | 1920 | 1920 |
| 22 | Ind | Primary W | 26 to 64 | S | F | | 11477 | 11477 | 11477 |
| 23 | Ind | Primary W | 26 to 64 | S | P | | 2243 | 2243 | 2243 |
| 24 | Ind | Primary W | 65 to 98 | G | F | | 77 | 77 | 77 |
| 25 | Ind | Primary W | 65 to 98 | G | P | | 43 | 43 | 43 |
| 26 | Ind | Primary W | 65 to 98 | M | F | | 116 | 116 | 116 |
| 27 | Ind | Primary W | 65 to 98 | M | P | | 54 | 54 | 54 |
| 28 | Ind | Primary W | 65 to 98 | P | F | | 289 | 289 | 289 |
| 29 | Ind | Primary W | 65 to 98 | P | P | | 120 | 120 | 120 |

[1] Note that some distributions are composed of few or no observations. This is acceptable as this is also an indication of the likelihood of such a classification existing in the population and being used to generate an activity episode. Also note that, by definition, the duration and start time observation totals are only required to equal the observed frequency totals for event distributions. Totals for duration and start time are based on the number of episodes/events, whereas frequency involves the total number of observed persons.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 30 | Ind | Primary W | 65 to 98 | S | F | | 114 | 114 | 114 |
| 31 | Ind | Primary W | 65 to 98 | S | P | | 110 | 110 | 110 |
| | | | | | | | 81851 | 81851 | 81851 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Occ | Emp Stat | | | Duration | Start Time | Frequency |
| 32 | Ind | Secondary W | G | F | | | 74 | 74 | 74 |
| 33 | Ind | Secondary W | G | P | | | 26 | 26 | 26 |
| 34 | Ind | Secondary W | M | F | | | 173 | 173 | 173 |
| 35 | Ind | Secondary W | M | P | | | 35 | 35 | 35 |
| 36 | Ind | Secondary W | P | F | | | 524 | 524 | 524 |
| 37 | Ind | Secondary W | P | P | | | 55 | 55 | 55 |
| 38 | Ind | Secondary W | S | F | | | 261 | 261 | 261 |
| 39 | Ind | Secondary W | S | P | | | 85 | 85 | 85 |
| | | | | | | | 1233 | 1233 | 1233 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Age | Occ | Emp Stat | | Duration | Start Time | Frequency |
| 40 | Ind | Work-Bus | 11 to 18 | G | F | | 0 | 0 | 0 |
| 41 | Ind | Work-Bus | 11 to 18 | G | P | | 6 | 6 | 6 |
| 42 | Ind | Work-Bus | 11 to 18 | M | F | | 9 | 9 | 5 |
| 43 | Ind | Work-Bus | 11 to 18 | M | P | | 8 | 8 | 8 |
| 44 | Ind | Work-Bus | 11 to 18 | P | F | | 1 | 1 | 1 |
| 45 | Ind | Work-Bus | 11 to 18 | P | P | | 2 | 2 | 2 |
| 46 | Ind | Work-Bus | 11 to 18 | S | F | | 2 | 2 | 2 |
| 47 | Ind | Work-Bus | 11 to 18 | S | P | | 17 | 17 | 16 |
| 48 | Ind | Work-Bus | 19 to 25 | G | F | | 59 | 59 | 46 |
| 49 | Ind | Work-Bus | 19 to 25 | G | P | | 23 | 23 | 22 |
| 50 | Ind | Work-Bus | 19 to 25 | M | F | | 186 | 186 | 147 |
| 51 | Ind | Work-Bus | 19 to 25 | M | P | | 24 | 24 | 23 |
| 52 | Ind | Work-Bus | 19 to 25 | P | F | | 174 | 174 | 134 |
| 53 | Ind | Work-Bus | 19 to 25 | P | P | | 50 | 50 | 44 |
| 54 | Ind | Work-Bus | 19 to 25 | S | F | | 143 | 143 | 100 |
| 55 | Ind | Work-Bus | 19 to 25 | S | P | | 89 | 89 | 71 |
| 56 | Ind | Work-Bus | 26 to 64 | G | F | | 438 | 438 | 324 |
| 57 | Ind | Work-Bus | 26 to 64 | G | P | | 106 | 106 | 87 |
| 58 | Ind | Work-Bus | 26 to 64 | M | F | | 2411 | 2411 | 1726 |
| 59 | Ind | Work-Bus | 26 to 64 | M | P | | 220 | 220 | 165 |
| 60 | Ind | Work-Bus | 26 to 64 | P | F | | 4143 | 4143 | 2915 |
| 61 | Ind | Work-Bus | 26 to 64 | P | P | | 414 | 414 | 323 |
| 62 | Ind | Work-Bus | 26 to 64 | S | F | | 2135 | 2135 | 1302 |
| 63 | Ind | Work-Bus | 26 to 64 | S | P | | 329 | 329 | 241 |
| 64 | Ind | Work-Bus | 65 to 98 | G | F | | 3 | 3 | 3 |
| 65 | Ind | Work-Bus | 65 to 98 | G | P | | 1 | 1 | 1 |
| 66 | Ind | Work-Bus | 65 to 98 | M | F | | 22 | 22 | 18 |
| 67 | Ind | Work-Bus | 65 to 98 | M | P | | 19 | 19 | 13 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 68 | Ind | Work-Bus | 65 to 98 | P | F | | 58 | 58 | 39 |
| 69 | Ind | Work-Bus | 65 to 98 | P | P | | 34 | 34 | 30 |
| 70 | Ind | Work-Bus | 65 to 98 | S | F | | 19 | 19 | 13 |
| 71 | Ind | Work-Bus | 65 to 98 | S | P | | 28 | 28 | 21 |
| | | | | | | | 11173 | 11173 | 7848 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Age | Occ | Emp Stat | | Duration | Start Time | Frequency |
| 72 | Ind | WAH-Bus | 19 to 25 | G | H, J | | 3 | 3 | 3 |
| 73 | Ind | WAH-Bus | 19 to 25 | M | H, J | | 22 | 22 | 21 |
| 74 | Ind | WAH-Bus | 19 to 25 | P | H, J | | 35 | 35 | 22 |
| 75 | Ind | WAH-Bus | 19 to 25 | S | H, J | | 20 | 20 | 14 |
| 76 | Ind | WAH-Bus | 26 to 64 | G | H, J | | 72 | 72 | 52 |
| 77 | Ind | WAH-Bus | 26 to 64 | M | H, J | | 614 | 614 | 434 |
| 78 | Ind | WAH-Bus | 26 to 64 | P | H, J | | 1083 | 1083 | 700 |
| 79 | Ind | WAH-Bus | 26 to 64 | S | H, J | | 579 | 579 | 329 |
| 80 | Ind | WAH-Bus | 65 to 98 | G | H, J | | 2 | 2 | 1 |
| 81 | Ind | WAH-Bus | 65 to 98 | M | H, J | | 13 | 13 | 8 |
| 82 | Ind | WAH-Bus | 65 to 98 | P | H, J | | 76 | 76 | 37 |
| 83 | Ind | WAH-Bus | 65 to 98 | S | H, J | | 31 | 31 | 13 |
| | | | | | | | 2550 | 2550 | 1634 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Age | Stu Stat | | | Duration | Start Time | Frequency |
| 84 | Ind | School | 11 to 15 | S | | | 13069 | 13069 | 13069 |
| 85 | Ind | School | 11 to 15 | P | | | 16 | 16 | 16 |
| 86 | Ind | School | 16 to 18 | S | | | 7184 | 7184 | 7184 |
| 87 | Ind | School | 16 to 18 | P | | | 74 | 74 | 74 |
| 88 | Ind | School | 19 to 25 | S | | | 5471 | 5471 | 5471 |
| 89 | Ind | School | 19 to 25 | P | | | 511 | 511 | 511 |
| 90 | Ind | School | 26 to 30 | S | | | 817 | 817 | 817 |
| 91 | Ind | School | 26 to 30 | P | | | 316 | 316 | 316 |
| 92 | Ind | School | 31 to 98 | S | | | 1018 | 1018 | 1018 |
| 93 | Ind | School | 31 to 98 | P | | | 889 | 889 | 889 |
| | | | | | | | 29365 | 29365 | 29365 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Occ | Emp Stat | | | Duration | Start Time | Frequency |
| 94 | Ind | Return Home | G | F | | | 130 | 130 | 125 |
| 95 | Ind | Return Home | G | P | | | 30 | 30 | 27 |
| 96 | Ind | Return Home | M | F | | | 141 | 141 | 128 |
| 97 | Ind | Return Home | M | P | | | 60 | 60 | 45 |
| 98 | Ind | Return Home | P | F | | | 481 | 481 | 454 |
| 99 | Ind | Return Home | P | P | | | 45 | 45 | 37 |
| 100 | Ind | Return Home | S | F | | | 198 | 198 | 182 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 101 | Ind | Return Home | S | P | | | 74 | 74 | 64 |
| | | | | | | | 1159 | 1159 | 1062 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Age | Licence | Emp Stat | | Duration | Start Time | Frequency |
| 102 | Ind | Other | 11 to 25 | Y | F | | 1799 | 1799 | 1373 |
| 103 | Ind | Other | 11 to 25 | Y | P | | 1577 | 1577 | 1164 |
| 104 | Ind | Other | 11 to 25 | Y | O, H, J | | 1854 | 1854 | 1397 |
| 105 | Ind | Other | 11 to 25 | N | F | | 136 | 136 | 111 |
| 106 | Ind | Other | 11 to 25 | N | P | | 379 | 379 | 304 |
| 107 | Ind | Other | 11 to 25 | N | O, H, J | | 2384 | 2384 | 2012 |
| 108 | Ind | Other | 26 to 49 | Y | F | | 12884 | 12884 | 9837 |
| 109 | Ind | Other | 26 to 49 | Y | P | | 2175 | 2175 | 1478 |
| 110 | Ind | Other | 26 to 49 | Y | O, H, J | | 5854 | 5854 | 3873 |
| 111 | Ind | Other | 26 to 49 | N | F | | 505 | 505 | 407 |
| 112 | Ind | Other | 26 to 49 | N | P | | 227 | 227 | 172 |
| 113 | Ind | Other | 26 to 49 | N | O, H, J | | 807 | 807 | 629 |
| 114 | Ind | Other | 50 to 64 | Y | F | | 2648 | 2648 | 2057 |
| 115 | Ind | Other | 50 to 64 | Y | P | | 616 | 616 | 427 |
| 116 | Ind | Other | 50 to 64 | Y | O, H, J | | 3881 | 3881 | 2633 |
| 117 | Ind | Other | 50 to 64 | N | F | | 95 | 95 | 82 |
| 118 | Ind | Other | 50 to 64 | N | P | | 54 | 54 | 44 |
| 119 | Ind | Other | 50 to 64 | N | O, H, J | | 499 | 499 | 420 |
| 120 | Ind | Other | 65 to 98 | Y | F | | 127 | 127 | 94 |
| 121 | Ind | Other | 65 to 98 | Y | P | | 170 | 170 | 119 |
| 122 | Ind | Other | 65 to 98 | Y | O, H, J | | 4811 | 4811 | 3431 |
| 123 | Ind | Other | 65 to 98 | N | F | | 4 | 4 | 2 |
| 124 | Ind | Other | 65 to 98 | N | P | | 10 | 10 | 7 |
| 125 | Ind | Other | 65 to 98 | N | O, H, J | | 1221 | 1221 | 1054 |
| | | | | | | | 44717 | 44717 | 33127 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables[2] | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Children | # Adults | Auto Acc | | Duration | Start Time | Frequency |
| 126 | Jnt | Other | Y | 1 | 0 | | 14 | 14 | 13 |
| 127 | Jnt | Other | Y | 1 | < | Impossible | 0 | 0 | 0 |
| 128 | Jnt | Other | Y | 1 | >= | | 86 | 86 | 74 |
| 129 | Jnt | Other | Y | 2 | 0 | | 50 | 50 | 42 |
| 130 | Jnt | Other | Y | 2 | < | | 520 | 520 | 399 |
| 131 | Jnt | Other | Y | 2 | >= | | 1342 | 1342 | 1078 |
| 132 | Jnt | Other | Y | 3+ | 0 | | 19 | 19 | 18 |
| 133 | Jnt | Other | Y | 3+ | < | | 543 | 543 | 441 |

[2] Here, the 'Children' variable indicates the presence of persons 15 years of age and under, the 'Number of Adults' variable indicates the number of persons aged 16 and above and the 'Auto Accessibility' variable indicates zero household automobiles, fewer automobiles than drivers (i.e. number of licenced persons in the household) and automobiles greater than or equal to drivers.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 134 | Jnt | Other | Y | 3+ | >= | | 415 | 415 | 337 |
| 135 | Jnt | Other | N | 2 | 0 | | 184 | 184 | 167 |
| 136 | Jnt | Other | N | 2 | < | | 1351 | 1351 | 999 |
| 137 | Jnt | Other | N | 2 | >= | | 2073 | 2073 | 1665 |
| 138 | Jnt | Other | N | 3+ | 0 | | 46 | 46 | 43 |
| 139 | Jnt | Other | N | 3+ | < | | 892 | 892 | 725 |
| 140 | Jnt | Other | N | 3+ | >= | | 698 | 698 | 563 |
| | | | | | | | 8233 | 8233 | 6564 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Age | Emp Stat | Gender | | Duration | Start Time | Frequency |
| 141 | Ind | Market | 11 to 25 | F | M | | 226 | 226 | 195 |
| 142 | Ind | Market | 11 to 25 | F | F | | 289 | 289 | 256 |
| 143 | Ind | Market | 11 to 25 | P | M | | 145 | 145 | 128 |
| 144 | Ind | Market | 11 to 25 | P | F | | 294 | 294 | 266 |
| 145 | Ind | Market | 11 to 25 | O, H, J | M | | 319 | 319 | 290 |
| 146 | Ind | Market | 11 to 25 | O, H, J | F | | 477 | 477 | 430 |
| 147 | Ind | Market | 26 to 49 | F | M | | 3027 | 3027 | 2591 |
| 148 | Ind | Market | 26 to 49 | F | F | | 3351 | 3351 | 2789 |
| 149 | Ind | Market | 26 to 49 | P | M | | 230 | 230 | 192 |
| 150 | Ind | Market | 26 to 49 | P | F | | 1352 | 1352 | 1067 |
| 151 | Ind | Market | 26 to 49 | O, H, J | M | | 731 | 731 | 615 |
| 152 | Ind | Market | 26 to 49 | O, H, J | F | | 3181 | 3181 | 2522 |
| 153 | Ind | Market | 50 to 64 | F | M | | 709 | 709 | 597 |
| 154 | Ind | Market | 50 to 64 | F | F | | 808 | 808 | 647 |
| 155 | Ind | Market | 50 to 64 | P | M | | 97 | 97 | 74 |
| 156 | Ind | Market | 50 to 64 | P | F | | 331 | 331 | 259 |
| 157 | Ind | Market | 50 to 64 | O, H, J | M | | 935 | 935 | 728 |
| 158 | Ind | Market | 50 to 64 | O, H, J | F | | 1780 | 1780 | 1396 |
| 159 | Ind | Market | 65 to 98 | F | M | | 45 | 45 | 36 |
| 160 | Ind | Market | 65 to 98 | F | F | | 32 | 32 | 24 |
| 161 | Ind | Market | 65 to 98 | P | M | | 39 | 39 | 37 |
| 162 | Ind | Market | 65 to 98 | P | F | | 53 | 53 | 47 |
| 163 | Ind | Market | 65 to 98 | O, H, J | M | | 1716 | 1716 | 1383 |
| 164 | Ind | Market | 65 to 98 | O, H, J | F | | 2447 | 2447 | 1972 |
| | | | | | | | 22614 | 22614 | 18541 |

| ID | Ind/Jnt | Activity Type | Explanatory Variables | | | Comments | Number of Observations | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Children | # Adults | Auto Acc | | Duration | Start Time | Frequency |
| 165 | Jnt | Market | Y | 1 | 0 | | 12 | 12 | 8 |
| 166 | Jnt | Market | Y | 1 | < | Impossible | 0 | 0 | 0 |
| 167 | Jnt | Market | Y | 1 | >= | | 42 | 42 | 34 |
| 168 | Jnt | Market | Y | 2 | 0 | | 35 | 35 | 32 |
| 169 | Jnt | Market | Y | 2 | < | | 379 | 379 | 312 |
| 170 | Jnt | Market | Y | 2 | >= | | 853 | 853 | 696 |
| 171 | Jnt | Market | Y | 3+ | 0 | | 14 | 14 | 13 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 172 | Jnt | Market | Y | 3+ | < | | 341 | 341 | 289 |
| 173 | Jnt | Market | Y | 3+ | >= | | 261 | 261 | 219 |
| 174 | Jnt | Market | N | 2 | 0 | | 139 | 139 | 117 |
| 175 | Jnt | Market | N | 2 | < | | 1195 | 1195 | 935 |
| 176 | Jnt | Market | N | 2 | >= | | 1785 | 1785 | 1391 |
| 177 | Jnt | Market | N | 3+ | 0 | | 27 | 27 | 27 |
| 178 | Jnt | Market | N | 3+ | < | | 641 | 641 | 531 |
| 179 | Jnt | Market | N | 3+ | >= | | 523 | 523 | 415 |
| | | | | | | | 6247 | 6247 | 5019 |

# APPENDIX F

## THEMATIC MAPS OF ACTIVITY LEVELS WITHIN THE GTA

Following are thematic maps showing the results of activity level assignment for the *work*, *market* and *other* location choice models based on the stated criteria used for activity level definition.

## Figure F–1 *Work* Activity Level by GTA Zone

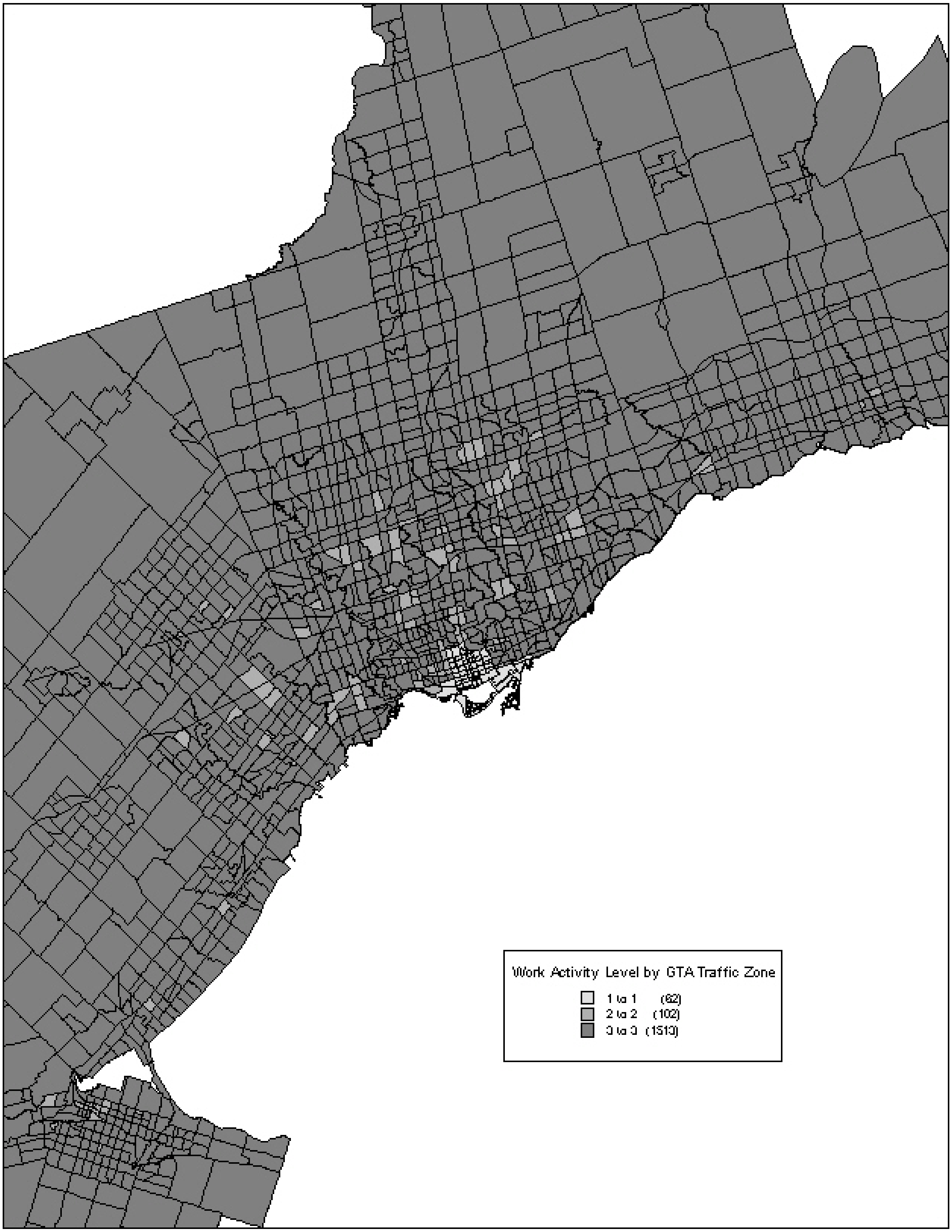
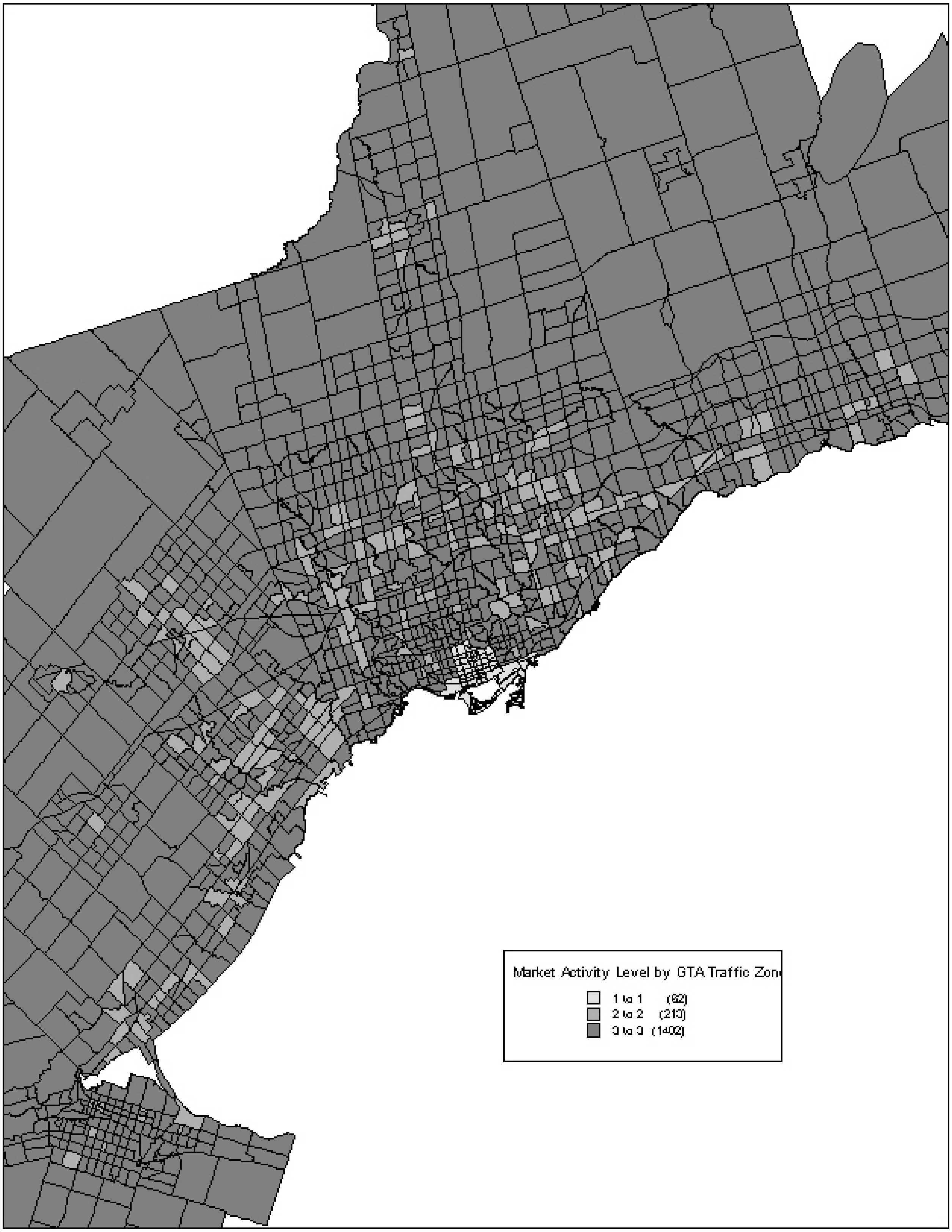
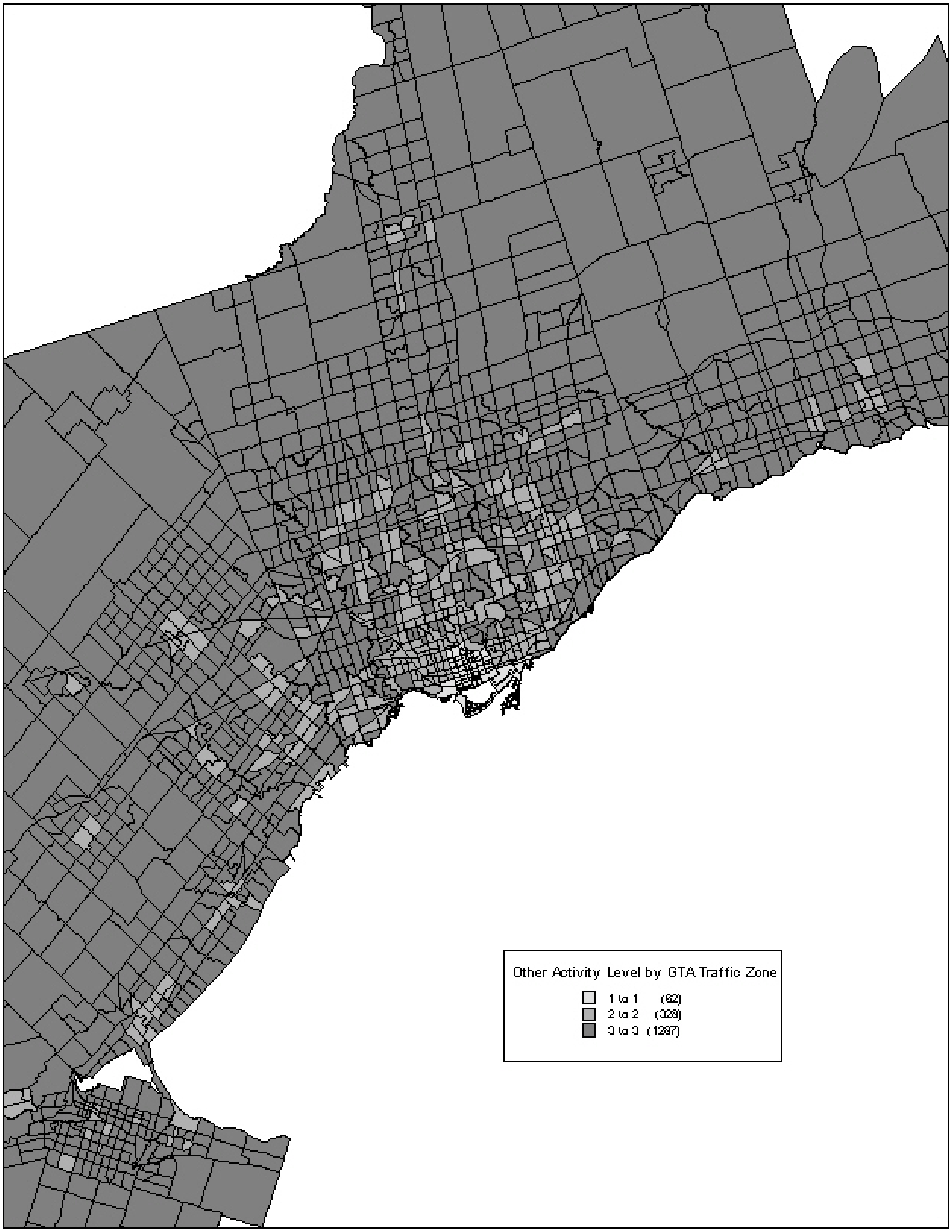


**Figure F–2 *Market* Activity Level by GTA Zone**



## Figure F–3 *Other* Activity Level by GTA Zone



# APPENDIX G

## LOCATION CHOICE MODEL PARAMETER ESTIMATION

Following is a definition of the explanatory variables used in the location choice models as well as output from the parameter estimation software and the maximum distance allowable constraint for each model.

### Explanatory Variable Definitions

| | |
|---|---|
| dist | Distance from reference zone to destination zone (straight line kilometres) |
| empT | Total employment (1000's) |
| empG | Total employment, general office/clerical (1000's) |
| empP | Total employment, professional/management/technical (1000's) |
| empM | Total employment, manufacturing/construction/trades (1000's) |
| empR | Total employment, retail sales and service (1000's) |
| pop | Total population (1000's) |
| sh–*x* | Short distance dummy variable, = 1 if the distance lies between *x* and *x*+1; = 0 otherwise |
| variable*y* | Indicates association of an explanatory variable with an activity level, where *y* is the activity level ($1 \leq y \leq 3$) |

### Work–Business (Usual Place of Work), Occupation Code G

```
********************************************************************************
*                                                                              *
* RUN NUMBER wbus5g     ENTROPY.F VERSION NUMBER 1.00                          *
*                                                                              *
* 96TTS Work-Business (Usual Place of Work) Location Choice - Gen: Test 5      *
*                                                                              *
********************************************************************************


 ******     MODEL RUN PARAMETERS     ******

 Max. No. of Iterations:                         20
 Max. % Error for Convergence:               0.0010

Observed O-D trip file= data/work_bus_usual_choices_gen.txt

     CONVERGENCE AFTER  3 ITERATIONS.


 ****** PARAMETER ESTIMATION RESULTS ******

 Parameter Parameter Parameter     TSTAT  Observed      Predicted    (P-O)/Obs
 Number    Name      Value                Total         Total            (%)

     1     dist      -0.111100E+00 -17.82  0.114451E+03  0.114448E+03   -0.003
     2     empP       0.362647E+00   4.03  0.253862E+01  0.253858E+01   -0.001
     3     empG       0.332294E+00   4.19 -0.996314E+01 -0.996321E+01    0.001
     4     sh-0       0.381626E-01   0.17  0.519060E+00  0.519028E+00   -0.006
     5     sh-1       0.583421E+00   3.55  0.120257E+01  0.120248E+01   -0.007
     6     sh-2       0.271133E+00   1.54  0.866330E+00  0.866866E+00    0.062
```

```
 ****** GOODNESS-OF-FIT STATISTICS ******

 Total observations =        523
 Total trips (1000's) =    11.207
 Log-likelihood, B=0 =   -0.832070E+02
 Log-likelihood, conv. = -0.593519E+02
 Rho-square =             0.2867
 R-square =               0.0384
```

Maximum distance: 50 kilometres

## Work–Business (Usual Place of Work), Occupation Code P

```
********************************************************************************
*                                                                              *
* RUN NUMBER wbus1p     ENTROPY.F VERSION NUMBER 1.00                          *
*                                                                              *
* 96TTS Work-Business (Usual Place of Work) Location Choice - Prof: Test 1     *
*                                                                              *
********************************************************************************


 ******     MODEL RUN PARAMETERS     ******

 Max. No. of Iterations:                           20
 Max. % Error for Convergence:                 0.0010

Observed O-D trip file= data/work_bus_usual_choices_prof.txt

     CONVERGENCE AFTER  3 ITERATIONS.


 ****** PARAMETER ESTIMATION RESULTS ******

 Parameter Parameter Parameter     TSTAT  Observed      Predicted    (P-O)/Obs
 Number    Name      Value               Total         Total           (%)

     1     dist      -0.922957E-01 -48.27  0.101636E+04  0.101636E+04    0.000
     2     empP       0.482183E+00  14.80  0.173948E+02  0.173951E+02    0.002
     3     empG       0.245875E+00   8.85 -0.765417E+02 -0.765410E+02   -0.001
     4     sh-0      -0.828902E-01  -0.86  0.264347E+01  0.264379E+01    0.012
     5     sh-1       0.330525E+00   4.82  0.568726E+01  0.568723E+01   -0.001


 ****** GOODNESS-OF-FIT STATISTICS ******

 Total observations =       3766
 Total trips (1000's) =    82.099
 Log-likelihood, B=0 =   -0.609564E+03
 Log-likelihood, conv. = -0.473336E+03
 Rho-square =             0.2235
 R-square =               0.0586
```

Maximum distance: 50 kilometres

## Work–Business (Usual Place of Work), Occupation Code M

```
********************************************************************************
*                                                                              *
* RUN NUMBER wbus1m     ENTROPY.F VERSION NUMBER 1.00                          *
*                                                                              *
* 96TTS Work-Business (Usual Place of Work) Location Choice - Man: Test 1      *
*                                                                              *
********************************************************************************


 ******     MODEL RUN PARAMETERS     ******

 Max. No. of Iterations:                         20
 Max. % Error for Convergence:               0.0010

Observed O-D trip file= data/work_bus_usual_choices_man.txt

     CONVERGENCE AFTER  3 ITERATIONS.


 ****** PARAMETER ESTIMATION RESULTS ******

 Parameter Parameter Parameter     TSTAT  Observed      Predicted    (P-O)/Obs
 Number    Name      Value                Total         Total           (%)

     1     dist      -0.837838E-01 -30.61  0.466864E+03  0.466850E+03   -0.003
     2     empM       0.544748E+00  28.93 -0.276586E+02 -0.276584E+02   -0.001
     3     sh-0       0.338273E+00   1.91  0.717340E+00  0.717287E+00   -0.007
     4     sh-1       0.486238E+00   3.85  0.153168E+01  0.153320E+01    0.099
     5     sh-2       0.405243E+00   3.45  0.177541E+01  0.177566E+01    0.014
     6     sh-3       0.207568E+00   1.67  0.151511E+01  0.151500E+01   -0.007


 ****** GOODNESS-OF-FIT STATISTICS ******

 Total observations =       1470
 Total trips (1000's) =    32.680
 Log-likelihood, B=0 =   -0.242641E+03
 Log-likelihood, conv. = -0.195202E+03
 Rho-square =             0.1955
 R-square =               0.0335
```

Maximum distance: 60 kilometres

## Work–Business (Usual Place of Work), Occupation Code R

```
********************************************************************************
*                                                                              *
* RUN NUMBER wbus1r     ENTROPY.F VERSION NUMBER 1.00                          *
*                                                                              *
* 96TTS Work-Business (Usual Place of Work) Location Choice - Ret: Test 1      *
*                                                                              *
********************************************************************************


 ******     MODEL RUN PARAMETERS     ******

 Max. No. of Iterations:                         20
 Max. % Error for Convergence:               0.0010

Observed O-D trip file= data/work_bus_usual_choices_ret.txt
```

```
    CONVERGENCE AFTER  2 ITERATIONS.


 ****** PARAMETER ESTIMATION RESULTS ******

 Parameter Parameter Parameter     TSTAT  Observed      Predicted    (P-O)/Obs
 Number    Name      Value                Total         Total           (%)

     1     dist      -0.837206E-01 -32.69  0.548286E+03  0.548314E+03    0.005
     2     empR       0.716674E+00  34.73 -0.204670E+02 -0.204674E+02    0.002
     3     sh-0       0.131481E+00   0.84  0.923000E+00  0.922985E+00   -0.002
     4     sh-1       0.498474E+00   4.80  0.236363E+01  0.236360E+01   -0.001
     5     sh-2       0.326770E+00   3.17  0.229486E+01  0.229485E+01   -0.001


 ****** GOODNESS-OF-FIT STATISTICS ******

 Total observations =       1869
 Total trips (1000's) =    41.035
 Log-likelihood, B=0 =   -0.304675E+03
 Log-likelihood, conv. = -0.242193E+03
 Rho-square =              0.2051
 R-square =                0.0393
```

Maximum distance: 55 kilometres

## Work–Business (No Usual Place of Work)

```
********************************************************************************
*                                                                              *
* RUN NUMBER wbusn4     ENTROPY.F VERSION NUMBER 1.00                          *
*                                                                              *
* 96TTS Work-Business (No Usual Place of Work) Location Choice: Test 4         *
*                                                                              *
********************************************************************************


 ******     MODEL RUN PARAMETERS     ******

 Max. No. of Iterations:                          20
 Max. % Error for Convergence:                0.0010

Observed O-D trip file= data/work_bus_no_usual_choices.txt

    CONVERGENCE AFTER  4 ITERATIONS.


 ****** PARAMETER ESTIMATION RESULTS ******

 Parameter Parameter Parameter     TSTAT  Observed      Predicted    (P-O)/Obs
 Number    Name      Value                Total         Total           (%)

     1     dist      -0.808636E-01 -41.58  0.854784E+03  0.854780E+03    0.000
     2     empT       0.692335E+00  43.79  0.529464E+02  0.529463E+02    0.000
     3     sh-0       0.115990E+01  10.75  0.204700E+01  0.204756E+01    0.027
     4     sh-1       0.530249E+00   5.57  0.263989E+01  0.263985E+01   -0.001
     5     sh-2       0.378443E+00   4.40  0.326829E+01  0.326825E+01   -0.001


 ****** GOODNESS-OF-FIT STATISTICS ******
```

```
 Total observations =       2736
 Total trips (1000's) =    60.113
 Log-likelihood, B=0 =   -0.446328E+03
 Log-likelihood, conv. = -0.356296E+03
 Rho-square =             0.2017
 R-square =               0.0387
```

Maximum distance: 60 kilometres

## Work–at–Home Business

```
********************************************************************************
*                                                                              *
* RUN NUMBER wahb4      ENTROPY.F VERSION NUMBER 1.00                          *
*                                                                              *
* 96TTS WAH-Business Location Choice: Test 4                                   *
*                                                                              *
********************************************************************************


 ******     MODEL RUN PARAMETERS     ******

 Max. No. of Iterations:                          20
 Max. % Error for Convergence:                0.0010

Observed O-D trip file= DATA/wah_bus_choices.txt

     CONVERGENCE AFTER  6 ITERATIONS.


 ****** PARAMETER ESTIMATION RESULTS ******

 Parameter Parameter Parameter     TSTAT  Observed      Predicted    (P-O)/Obs
 Number    Name      Value                Total         Total           (%)

     1     dist      -0.847437E-01 -38.32  0.666297E+03  0.666297E+03    0.000
     2     empT       0.769965E+00  43.12  0.500513E+02  0.500512E+02    0.000
     3     sh-0       0.162646E+01  15.99  0.243090E+01  0.243091E+01    0.000
     4     sh-1       0.112602E+01  13.44  0.373911E+01  0.373910E+01    0.000


 ****** GOODNESS-OF-FIT STATISTICS ******

 Total observations =       2249
 Total trips (1000's) =    49.630
 Log-likelihood, B=0 =   -0.368493E+03
 Log-likelihood, conv. = -0.287909E+03
 Rho-square =             0.2187
 R-square =               0.0364
```

Maximum distance: 60 kilometres

## Other

```
******************************************************************************
*                                                                            *
* RUN NUMBER other4     ENTROPY.F VERSION NUMBER 1.00                        *
*                                                                            *
* 96TTS Other Purpose Location Choice: Test 4                                *
*                                                                            *
******************************************************************************


 ******     MODEL RUN PARAMETERS     ******

 Max. No. of Iterations:                          20
 Max. % Error for Convergence:                0.0010

Observed O-D trip file= data/other_choices.txt

     CONVERGENCE AFTER  5 ITERATIONS.


 ****** PARAMETER ESTIMATION RESULTS ******

 Parameter Parameter Parameter     TSTAT  Observed      Predicted    (P-O)/Obs
 Number    Name      Value                Total         Total           (%)

     1     dist      -0.223809E+00-265.11  0.826030E+04  0.826341E+04    0.038
     2     empT       0.534774E+00 140.67  0.659067E+03  0.659031E+03   -0.006
     3     pop        0.913737E-01  52.26  0.667984E+03  0.668068E+03    0.013
     4     sh         0.763785E+00  48.29  0.107352E+03  0.107448E+03    0.090


 ****** GOODNESS-OF-FIT STATISTICS ******

 Total observations =      35170
 Total trips (1000's) =  1231.148
 Log-likelihood, B=0 =   -0.914098E+04
 Log-likelihood, conv. = -0.619983E+04
 Rho-square =             0.3218
 R-square =               0.3814
```

Maximum distance: 35 kilometres

## Market

```
******************************************************************************
*                                                                            *
* RUN NUMBER shop7      ENTROPY.F VERSION NUMBER 1.00                        *
*                                                                            *
* 96TTS Shopping Location Choice: Test 3, SHORT & DMAX                       *
*                                                                            *
******************************************************************************


 ******     MODEL RUN PARAMETERS     ******

 Max. No. of Iterations:                          20
 Max. % Error for Convergence:                0.0010

Observed O-D trip file= data/market_choices.txt

     CONVERGENCE AFTER  4 ITERATIONS.
```

****** PARAMETER ESTIMATION RESULTS ******

| Parameter Number | Parameter Name | Parameter Value | TSTAT | Observed Total | Predicted Total | (P-O)/Obs (%) |
|---|---|---|---|---|---|---|
| 1 | dist1 | -0.215486E+00 | -39.10 | 0.247350E+03 | 0.247347E+03 | -0.001 |
| 2 | empR1 | 0.111729E+01 | 38.42 | 0.255646E+02 | 0.255639E+02 | -0.003 |
| 3 | sh1-0 | 0.152106E+01 | 12.79 | 0.272130E+01 | 0.272364E+01 | 0.086 |
| 4 | sh1-1 | 0.103685E+01 | 10.78 | 0.476410E+01 | 0.476345E+01 | -0.014 |
| 5 | sh1-2 | 0.469059E+00 | 5.00 | 0.363426E+01 | 0.363375E+01 | -0.014 |
| 6 | dist2 | -0.271879E+00 | -115.82 | 0.189766E+04 | 0.189764E+04 | -0.001 |
| 7 | const2 | 0.177907E+01 | 26.44 | 0.392564E+03 | 0.392560E+03 | -0.001 |
| 8 | empR2 | 0.108578E+01 | 105.79 | 0.331610E+02 | 0.331608E+02 | -0.001 |
| 9 | sh2-0 | 0.141252E+01 | 44.50 | 0.425184E+02 | 0.425180E+02 | -0.001 |
| 10 | sh2-1 | 0.103826E+01 | 41.30 | 0.778518E+02 | 0.778509E+02 | -0.001 |
| 11 | sh2-2 | 0.679137E+00 | 28.10 | 0.683854E+02 | 0.683846E+02 | -0.001 |
| 12 | dist3 | -0.290998E+00 | -107.45 | 0.123349E+04 | 0.123345E+04 | -0.003 |
| 13 | const3 | 0.619005E+00 | 9.20 | 0.288534E+03 | 0.288536E+03 | 0.001 |
| 14 | empR3 | 0.665791E+00 | 78.76 | -0.317128E+03 | -0.317132E+03 | 0.001 |
| 15 | sh3-0 | 0.159716E+01 | 56.63 | 0.551817E+02 | 0.551887E+02 | 0.013 |
| 16 | sh3-1 | 0.113225E+01 | 45.41 | 0.708839E+02 | 0.708829E+02 | -0.002 |

****** GOODNESS-OF-FIT STATISTICS ******

Total observations = 14855
Total trips (1000's) = 717.687
Log-likelihood, B=0 = -0.532866E+04
Log-likelihood, conv. = -0.281576E+04
Rho-square = 0.4716
R-square = 0.4966

Maximum distance: 50, 40 and 40 kilometres for activity centres 1, 2 and 3, respectively

# APPENDIX H

## MODE CHOICE MODEL PARAMETER ESTIMATION

Following are definitions of the explanatory variables used in the mode choice models as well as output from the parameter estimation software for each model. Relevant non–drive modes for each level are listed in the output.

### Explanatory Variable Definitions[1]

| | |
|---|---|
| adist | chain distance (sum of straight line trip distances for the chain in kilometres) |
| adjzon | = 1 if origin and destination zones are adjacent; = 0 otherwise |
| age<16 | = 1 if person's age is less than 16 years; = 0 otherwise |
| age>15 | = 1 if person's age is greater than 15 years; = 0 otherwise |
| age>25 | = 1 if person's age is greater than 25 years; = 0 otherwise |
| atime | auto travel time, trip and chain level (minutes) |
| bictt | bicycle travel time, calculated assuming 15 km/h speed and straight line distance correction factor of 1.4 (minutes) |
| hz>tor | = 1 if home zone is outside of the City of Toronto; = 0 otherwise |
| indtrp | number of independent trips |
| intraz | = 1 if the trip origin zone is the same as the trip destination zone; = 0 otherwise |
| km<5 | = 1 if straight–line trip distance is less than or equal to 5 kilometres; = 0 otherwise |
| lsmndr | logsum term as calculated for the non–drive chain |
| manu | = 1 if person's occupation is manufacturing; = 0 otherwise |
| ndrive | number of licensed drivers in the person's household |
| njoint | number of joint trips in the chain in which the trip occurs |
| ntrips | number of trips in the chain in which the trip occurs |
| ntrp2+ | = 1 if the number of trips in the chain is greater than 2, = 0 otherwise |
| nveh2 | = 1 if the number of household vehicles is 2 or greater; = 0 otherwise |
| pkcost | daily parking cost of destination zone (1996$) |
| pkct/4 | = pkcost divided by 4 |
| prof | = 1 if person's occupation is professional; = 0 otherwise |
| taxcst | taxi fare, calculated assuming $2.5 + $1/km with straight line distance correction factor of 1.4 (1996$) |
| tfare | transit fare (1996$) |
| tivtt2 | sum of transit in–vehicle travel time and access/egress walk time (minutes) |
| trip2+ | =1 if the number of trips in the chain is greater than 2; = 0 otherwise |
| twait | transit wait and transfer time (minutes) |
| tsteve | = 1 if the trip start time occurs between 1900 and 559; = 0 otherwise |
| walktt | walk travel time, calculated assuming 4 km/h speed and straight line distance correction factor of 1.4 (minutes) |
| wsub | = 2 if both trip origin and trip destination are within 1 kilometre of a subway station; = 1 if either trip origin or destination are within 1 kilometre of a subway station; = 0 otherwise |

[1] Note that level of service explanatory variables used in the non–drive trip models are trip–level and those used in the drive/non–drive chain models are chain–level (e.g. travel time is that for the entire chain).

## Level 4 Non–Drive Trips

```
****************************************************************************
*                                                                          *
*                    LOGIT MODEL ESTIMATION VERSION 5.1                    *
*                                                                          *
*            DATA SET: School 18- Trips                                    *
*            MODEL:    17                                                  *
*                                                                          *
*                                                                          *
*    PAR:     c-apas c-tran c-walk c-bike cost   ivtt   ntrips ndrive      *
*                                                                          *
*    a-pass   1      0      0      0      pkct/4 atime  ntrips ndrive      *
*    transit  0      1      0      0      tfare  tivtt2 ntrips 0           *
*    walk     0      0      1      0      0      walktt 0      0           *
*    bicycle  0      0      0      1      0      bictt  0      0           *
*    sch-bus  0      0      0      0      0      0      0      0           *
*                                                                          *
*    PAR:     hz>tor age<16 intraz age>15 tadjzn tintra twait  wsub        *
*                                                                          *
*    a-pass   0      age<16 0      0      0      0      0      0           *
*    transit  0      0      0      age>15 adjzon intraz twait  wsub        *
*    walk     0      0      intraz 0      0      0      0      0           *
*    bicycle  0      0      intraz 0      0      0      0      0           *
*    sch-bus  hz>tor age<16 0      0      0      0      0      0           *
*                                                                          *
*    PAR:     sadjzn sintra km<5   njoint tsteve                           *
*                                                                          *
*    a-pass   0      0      0      njoint 0                                *
*    transit  0      0      0      0      tsteve                           *
*    walk     0      0      km<5   0      0                                *
*    bicycle  0      0      km<5   0      0                                *
*    sch-bus  adjzon intraz 0      0      0                                *
*                                                                          *
*                                                                          *
****************************************************************************


****************************************************************************
*                                                                          *
*                          FINAL PARAMETER VALUES                          *
*                                                                          *
*                NUMBER        NAME          VALUE         T-STAT          *
*                                                                          *
*                   1          c-apas    -.11666E+01    -12.4600           *
*                   2          c-tran    0.13832E+01     13.9088           *
*                   3          c-walk    0.34953E+01     13.6475           *
*                   4          c-bike    -.86369E+00     -3.4586           *
*                   5          cost      -.15647E-01     -1.1236           *
*                   6          ivtt      -.47666E-01    -58.8657           *
*                   7          ntrips    0.95473E+00     25.0791           *
*                   8          ndrive    0.44455E+00     31.4569           *
*                   9          hz>tor    0.15248E+01     36.9979           *
*                  10          age<16    0.54909E+00     17.4802           *
*                  11          intraz    0.22020E+00      5.5663           *
*                  12          age>15    0.44558E+00     13.1290           *
*                  13          tadjzn    -.17724E+01    -26.4671           *
*                  14          tintra    -.29374E+01    -38.8734           *
*                  15          twait     -.40470E-01    -15.2422           *
*                  16          wsub      0.44813E+00     17.5072           *
*                  17          sadjzn    -.12071E+00     -1.7324           *
```

```
*               18          sintra    -.40572E+00     -7.2923            *
*               19          km<5      0.80892E+00      3.2521            *
*               20          njoint    0.13134E+01     14.6503            *
*               21          tsteve    -.57433E+00     -7.7643            *
*                                                                        *
**************************************************************************
```

```
**************************************************************************
*                                                                        *
*                         GOODNESS-OF-FIT STATISTICS                     *
*                                                                        *
*                    No. of weighted observations= 38949                 *
*                    No. of cases=                149404                 *
*                    No. of parameters estimated=     21                 *
*                    Degrees of freedom=          149383                 *
*                                                                        *
*                    Log likelihood at B=0,=    -61259.3                 *
*                    Log likelihood at conv.=   -40519.3                 *
*                    Log likelihood ratio=       41479.9                 *
*                    RHO-square=                  0.3386                 *
*                    Adjusted RHO-square=         0.3385                 *
*                    Horowitz' RHO-square=        0.3384                 *
*                                                                        *
*                    Expected percent right=        44.9                 *
*                                                                        *
**************************************************************************
```

## Level 4 Drive/Non–Drive Chains

```
**************************************************************************
*                                                                        *
*                     LOGIT MODEL ESTIMATION VERSION 5.1                 *
*                                                                        *
*             DATA SET: School 18- Chains                                *
*             MODEL:    32                                               *
*                                                                        *
*                                                                        *
*    PAR:     logsum aivtt  nveh2+ hz>tor dist   ntrp2+                  *
*                                                                        *
*    drive    0      atime  nveh2  hz>tor adist  ntrp2+                  *
*    nondrive lsmndr 0      0      0      0      0                       *
*                                                                        *
*                                                                        *
**************************************************************************
```

```
**************************************************************************
*                                                                        *
*                           FINAL PARAMETER VALUES                       *
*                                                                        *
*                 NUMBER        NAME         VALUE         T-STAT        *
*                                                                        *
*                    1          logsum    0.28978E+00     16.9585        *
*                    2          aivtt     -.36875E-01     -3.4383        *
*                    3          nveh2+    0.10532E+01      8.1013        *
*                    4          hz>tor    -.21387E-01     -0.2021        *
*                    5          dist      0.34896E-01      2.6638        *
*                    6          ntrp2+    0.29539E+01     14.7096        *
*                                                                        *
**************************************************************************
```

```
****************************************************************************
*                                                                          *
*                        GOODNESS-OF-FIT STATISTICS                        *
*                                                                          *
*                   No. of weighted observations=  2697                    *
*                   No. of cases=                  2697                    *
*                   No. of parameters estimated=      6                    *
*                   Degrees of freedom=            2691                    *
*                                                                          *
*                   Log likelihood at B=0,=     -1869.7                    *
*                   Log likelihood at conv.=    -1325.1                    *
*                   Log likelihood ratio=        1089.2                    *
*                   RHO-square=                  0.2913                    *
*                   Adjusted RHO-square=         0.2896                    *
*                   Horowitz' RHO-square=        0.2897                    *
*                                                                          *
*                   Expected percent right=        68.1                    *
*                                                                          *
****************************************************************************
```

## Level 5 Non–Drive Trips

```
****************************************************************************
*                                                                          *
*                    LOGIT MODEL ESTIMATION VERSION 5.1                    *
*                                                                          *
*            DATA SET: Non-Work/School Trips                               *
*            MODEL:    17                                                  *
*                                                                          *
*                                                                          *
*    PAR:     c-apas c-tran c-walk c-bike ivtt   ndrive wsub   tadjzn      *
*                                                                          *
*    a-pass   1      0      0      0      atime  ndrive 0      0           *
*    transit  0      1      0      0      tivtt2 0      wsub   adjzon      *
*    walk     0      0      1      0      walktt 0      0      0           *
*    bicycle  0      0      0      1      bictt  0      0      0           *
*    taxi     0      0      0      0      atime  0      0      0           *
*                                                                          *
*    PAR:     tintra twait  hz>tor cost   njoint km<5   tsteve             *
*                                                                          *
*    a-pass   0      0      hz>tor pkct/4 njoint 0      0                  *
*    transit  intraz twait  0      tfare  0      0      tsteve             *
*    walk     0      0      0      0      0      km<5   0                  *
*    bicycle  0      0      0      0      0      km<5   0                  *
*    taxi     0      0      0      taxcst 0      0      0                  *
*                                                                          *
*                                                                          *
****************************************************************************
```

```
****************************************************************************
*                                                                          *
*                          FINAL PARAMETER VALUES                          *
*                                                                          *
*                NUMBER        NAME         VALUE         T-STAT           *
*                                                                          *
*                   1          c-apas    0.60881E+00     10.7525           *
*                   2          c-tran    0.26730E+01     48.8673           *
*                   3          c-walk    -.80757E+00     -5.7497           *
*                   4          c-bike    -.16444E+01    -13.2016           *
*                   5          ivtt      -.20399E-01    -20.5237           *
```

```
*                6          ndrive    0.50999E+00     39.8011             *
*                7          wsub      0.36970E+00     18.8808             *
*                8          tadjzn    -.14607E+01    -20.5506             *
*                9          tintra    -.22116E+01    -26.7926             *
*               10          twait     -.54598E-01    -15.8099             *
*               11          hz>tor    0.81851E+00     26.2060             *
*               12          cost      -.97944E-01    -16.7750             *
*               13          njoint    0.15684E+01     67.2696             *
*               14          km<5      0.12203E+01     10.5269             *
*               15          tsteve    -.10468E+01    -35.3966             *
*                                                                         *
***************************************************************************


***************************************************************************
*                                                                         *
*                         GOODNESS-OF-FIT STATISTICS                      *
*                                                                         *
*                    No. of weighted observations= 48548                  *
*                    No. of cases=                175391                  *
*                    No. of parameters estimated=     15                  *
*                    Degrees of freedom=          175376                  *
*                                                                         *
*                    Log likelihood at B=0,=    -73940.5                  *
*                    Log likelihood at conv.=   -29281.6                  *
*                    Log likelihood ratio=       89317.7                  *
*                    RHO-square=                  0.6040                  *
*                    Adjusted RHO-square=         0.6039                  *
*                    Horowitz' RHO-square=        0.6039                  *
*                                                                         *
*                    Expected percent right=        68.9                  *
*                                                                         *
***************************************************************************
```

## Level 5 Drive/Non−Drive Chains

```
***************************************************************************
*                                                                         *
*                     LOGIT MODEL ESTIMATION VERSION 5.1                  *
*                                                                         *
*             DATA SET: Non-Work/School Chains                            *
*             MODEL:    24                                                *
*                                                                         *
*                                                                         *
*    PAR:     d-drv  logsum aivtt  nveh2+ hz>tor ntrp2+  pkcost age>25    *
*                                                                         *
*    drive    1      0      atime  nveh2  hz>tor ntrp2+  pkcost age>25    *
*    nondrive 0      lsmndr 0      0      0      0       0      0         *
*                                                                         *
*    PAR:     prof   manu                                                 *
*                                                                         *
*    drive    prof   manu                                                 *
*    nondrive 0      0                                                    *
*                                                                         *
*                                                                         *
***************************************************************************


***************************************************************************
*                                                                         *
*                           FINAL PARAMETER VALUES                        *
*                                                                         *
```

```
*              NUMBER        NAME         VALUE         T-STAT             *
*                                                                          *
*                1          d-drv     0.23118E+01     45.9233              *
*                2          logsum    0.30565E+00     55.4230              *
*                3          aivtt     -.92683E-02    -12.4661              *
*                4          nveh2+    0.42874E+00     17.4151              *
*                5          hz>tor    0.49820E+00     19.3244              *
*                6          ntrp2+     0.21907E+01    41.6773              *
*                7          pkcost    -.11112E+00    -26.7614              *
*                8          age>25    0.77038E+00     21.4454              *
*                9          prof      0.24352E+00      8.0938              *
*               10          manu      0.56777E+00     12.9204              *
*                                                                          *
****************************************************************************
```

```
****************************************************************************
*                                                                          *
*                        GOODNESS-OF-FIT STATISTICS                        *
*                                                                          *
*                    No. of weighted observations= 57807                   *
*                    No. of cases=                 57807                   *
*                    No. of parameters estimated=     10                   *
*                    Degrees of freedom=           57797                   *
*                                                                          *
*                    Log likelihood at B=0,=    -40069.2                   *
*                    Log likelihood at conv.=   -23438.7                   *
*                    Log likelihood ratio=       33261.0                   *
*                    RHO-square=                  0.4150                   *
*                    Adjusted RHO-square=         0.4149                   *
*                    Horowitz' RHO-square=        0.4149                   *
*                                                                          *
*                    Expected percent right=        74.9                   *
*                                                                          *
****************************************************************************
```