ABSTRACT

EOM, JIN KI. Incorporating Activity-Based Special Generator Data into A Conventional Planning Model. (Under the direction of Dr. John R. Stone)

Special generators need special attention in developing travel demand models since the standard trip generation and distribution model in the conventional four-step approach do not provide reliable estimates of their travel patterns. New modeling approaches such as activity-based and tour-based models, considering travel behavior of individual household or person, seem to be more appropriate for those special generators. However, only a few practical applications have been made since these approaches usually require a lot of data resources and computing time to solve their complicated model structure.

The primary objectives of this research are to improve the trip generation and trip distribution of special generators (e.g., university) by applying an activity-based approach, and to provide a transitional methodology for practically incorporating the activity-based data into a conventional planning model. The research developed a spatial and temporal activity-based model dealing with special generator data of North Carolina State University (NCSU). Also, the research tested the transferability of university student travel data by using statistical approaches that indicated that the university students' travel data can be transferred for the two cases considered. The NCSU activity-based model provided the estimates of trip generation at the disaggregated level of individual buildings by hours of the day - a disaggregation was not obtainable from a conventional planning model. The model estimates, student building presence and trip generation, compared well to field data from student registration records and student trips observed at sample buildings. The results revealed that the activity-based model well replicated both building presence and trip generation. In addition, the research compared the estimated trip generation of the activity-based model to that of a traditional planning model and discussed findings in terms of model accuracy, structure, data requirements, and capability of model application. The insights gained from this study will serve as the basis of an activity-based Triangle Regional model in North Carolina and elsewhere.

INCORPORATING ACTIVITY-BASED SPECIAL GENERATOR DATA INTO A CONVENTIONAL PLANNING MODEL

by

Jin Ki Eom

A dissertation submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

CIVIL ENGINEERING

Raleigh, NC

2007

APPROVED BY:

Dr. Sujit K. Ghosh	**Dr. Joseph E. Hummer**
Dr. Nagui M. Rouphail **(Co-chair of Advisory Committee)**	**Dr. John R. Stone** **(Co-chair of Advisory Committee)**

To my Father in Heaven

and

my family

Biography

Jin Ki Eom was born in Donghae, Republic of Korea on March 27, 1969. He entered the Department of Transportation Engineering at Hanyang University in 1988 and earned his B.S. in 1995. He earned M.C.P.(Master of city planning) in 1997 in graduate school of urban planning and environmental study at Seoul National University and earned another M.S. in 2003 at department of Civil Engineering, Pennsylvania State University. He joined the Civil Engineering, North Carolina State University for the pursuit of Ph.D. degree in 2004. He received his Doctor of Philosophy degree in Civil engineering in February 2007.

Acknowledgements

I would like to express my sincere thanks to my advisor, Dr. John R. Stone. This expression is not enough to him regarding his guidance, support, mentorship and encouragement. Without his help, this research would not have been possible. I would like to express thanks to my committee members: Dr. Rouphail, Dr. Joe Hummer, and Dr. Ghosh for providing their many productive comments on the way to improve the contents of this dissertation. I would also like to express gratitude to Leta F. Huntsinger for giving me her warm advice and support at ITRE. She always shares with me her modeling skills, experience and knowledge that was essential to this research.

I would like to express special thanks to ITE student chapter and Korean student folks in Civil Engineering at NCSU for their participation and assistance in the data collection. I would like to express gratitude to Slade McCalip and Sarah Smith for their help in data acquisition and editorial assistance. I would like to thank all my team members in TRM Service Bureau at ITRE, specially Mei Ingram, Bing Mei, and Nita Bhave. I would like to express special thanks to Dr. Kee-Yeon Hwang at Hong-Ik University in Korea for his mentorship and encouragement.

I would like to express my warmest gratitude and love to Pastor Byeong Jin Woo and my church family and prayer group members. Their prayers made me strong and positive when I was tired, depressed and about to lose my motivation. I would like to give my profound love and thanks to my mother and father in heaven, brother and sister who provide priceless support, their unconditional and endless love, patience and absolute faith in me. I would also like to thank to my parents-in-law and their families for their patience and their prayers for me and my family.

I would like to give my love and thanks to my wife Mijin, who was always with me, who was stronger than me when I needed her to lean on, who has been so patient, who has been so supportive and sacrificed and who loves me. I also would like to give my deepest love and thanks to my daughter, Soyeon, who is the source of happiness in my life.

Finally, I would like to express my sincere thankfulness to God who made it possible for me to finish this dissertation and to my father who is always looking upon me in heaven. I would like to thank to all those who I may have forgotten.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background

The key deficiency of the traditional four-step travel demand model is its lack of behavioral consideration at the individual level because the four-step model assumes trip-making behavior is uniform within a given traffic analysis zone (TAZ). Except for the aggregate modal choice step, these models have little or no behavioral basis, and they cannot be used to evaluate disaggregate effects of Transportation Demand Management (TDM) strategies and other programs, such as improvements in public transit [1]. Specifically, the four-step model does not fully capture the travel demand of special generators, such as universities. The trip generation component in a conventional travel demand model usually considers trip rates with socio-economic information such as household size, auto ownership, and household income. Thus, the traditional model structure does not accommodate university travel patterns based on the household model.

A university is a special community consisting of students, faculty, and staff. Yet, students living on campus are considered as single person households in current trip generation models even though such single person households on-campus are not likely to have the same travel patterns as single person households in residential areas. The usual travel patterns of conventional households do not exist in universities because the trip makers, which are faculty and students, have flexible schedules. Further, the conventional aggregated trip rates by TAZ are too coarse to replicate the actual number of trips and the time when the travel is made at the individual level. This causes a significant difference in university traffic on a subarea network. It is, therefore, important to give special attention to the characteristics of special generators like universities in travel demand models.

Over the past decade, many researchers have developed new disaggregated modeling approaches to overcome the deficiencies of traditional four-step models. The major concern is that travel demand models should consider the travel behavior of individual trip makers throughout the entire day. Although researchers believe that the new approaches would be better than the existing model, they usually require significant data resources and computing time to solve their complicated model structures, and only a few applications to community or regional models have been made [2, 3, 4].

1.2 Research Needs

Even as researchers develop improved travel demand models, most practitioners do not have significant theoretical knowledge or resources for the new data and model development.

Activity surveys for people and households are becoming more widely used, but metropolitan planning organizations (MPOs) are not ready to discard their current models and start over with activity-based models. This is because the resources for data collection, model development, and calibration are not available in many areas. Another reason is that not enough information or demonstrations have been given to practitioners to build their confidence in alternative approaches. Therefore, incremental approaches for transitioning from traditional models to behavior-based models are needed. An example of this incremental approach is the use of an activity-based sub-model for special generator data (such as a university) and the incorporation of this data into a regional transportation model while not changing the aggregate generation of other TAZs. This approach can potentially demonstrate better travel estimates overall in a regional model by providing better input data for special generator analysis. If successful such an approach will demonstrate that activity-based approach are practical while preserving current investments in traditional travel models.

1.3 Research Scope

While the structure of current activity-based models is theoretically logical, the methodologies are still far beyond what can be applied by practitioners due to the complexity of models, data availability and high computation time. In previous research tests were conducted to replicate observed activity choices of travel, but traffic has not been assigned to a transportation network - a critical concern for regional planning agencies. Thus, a practical bridge between activity-based theory and travel demand modeling practice is needed, and its application must be demonstrated to be cost-effective and accurate. In order for transportation practitioners to have confidence in the activity-based model and to use it to their regional models, further empirical research is required. Rossi (2004) [5] listed future directions of activity-based travel demand models to satisfy the necessity for practical use. Three main research areas relate to this study:

- Comparisons with traditional travel demand model
- Finer spatial and temporal resolutions
- Integration with land use model

This research will focus on comparisons with traditional travel demand models. By limiting the scope to a university special generator within a regional model, finer spatial and temporal resolution for a particular land use will be achieved. In order to do this, a spatial and temporal activity-based travel demand model will be developed for North Carolina State University (NCSU). The estimated activity-based spatial and temporal travel demand assigned to the university facility will be compared to observed student trips to ascertain whether this model works well in a university special generator. For the comparison between traditional and activity-based approach, first the estimated trips from both approach are compared to one another to examine which approach has a better accuracy and several issues such as data requirement, model validation and calibration, and the ability of model applications for policy analysis will be discussed. This approach will determine if the activity-based approach can practically replace a conventional travel demand model for special generator analysis.

1.4 Research Objectives

Three main objectives are identified for this research based on the hypothesis that incorporating activity information in a traditional four-step model will improve the accuracy of the model, especially for special generators. The proposed research objectives are:

- To provide a transitional methodology for practically incorporating activity-based data into a conventional planning model
- To improve the trip generation of a special generator (e.g., university) by applying an activity-based approach
- To compare the activity-based travel demand model against the traditional four-step model in terms of model structure, data requirement, and capability of model application.

1.5 Organization of this Thesis

Besides this introductory chapter describing the problem, scope and objectives of the research, the thesis has the following chapters. The second chapter is a review of transportation travel demand modeling as used in most planning agencies. The second chapter also reviews activity-based travel demand models. The third chapter focuses on the data that are used in this study such as the student trip diary and geographic information of the NCSU. The data are briefly described and the descriptive statistics are given. The fourth chapter describes the detailed activity-based model structure and data analysis required in this study. The application of particular phases of the model development is described by using a step-by-step approach. First an introduction to the components of activity-based model is provided. Then, classification of activity types and a methodology for developing an activity profile are presented. The fifth chapter discusses the activity-based trip generation results with respect to the observed trips on campus facilities sampled. The sixth chapter describes the results of model comparison between a conventional trip generation approach and an activity-based trip generation. Finally, the seventh chapter concludes the thesis with a summary and discussion of results and suggestions for future research that will improve the existing regional travel demand models by incorporating an activity-based model structure.

CHAPTER 2

LITERATURE REVIEW

This chapter illustrates the background and basic structure of the traditional four step travel demand model. The limitations of the traditional the four step travel demand model are identified and alternative travel demand modeling approaches such as activity-based travel demand models are introduced. Although most activity-based travel demand models still remain within the area of academic field, a few activity-based models practically applied to real context are introduced and reviewed with respect to suitability of incorporating activity-based special generator data into a conventional travel demand model.

2.1 Travel Demand Models

Since the 1950's the role of travel demand models has focused on long-range transportation planning associated with new infrastructure such as highway, rail, and transit. The travel demand models mostly provide the future travel demand and the affects on level of service of transportation infrastructure within a specific study context. More specifically, travel demand models answer whether the current transportation infrastructure will be sufficient if the travel demand is expected to increase in the future. Transportation planners want to know what would happen if transportation supply is not enough to accommodate increased travel demand. Since early model systems were expected to provide ”the big picture” of what would happen in the future, detailed accuracy of models was not a great concern.

2.2 Conventional Four-step Travel Demand Model

The conventional travel demand model was created in early 1950's. The model was called a Four-Step Model (Ortuzar and Willumsen 1994, 6). This model sequentially comprises four sub-steps: *trip generation, trip distribution, mode choice, and traffic assignment* as shown in Figure 2.1. The initial step, trip generation, requires land use and socioeconomic data such as population, household size, and household income, as well as information from a travel diary survey which describes trip making characteristics and patterns. After going through these four steps, the method assigns the predicted traffic volume on the modeled network. These sub-steps are briefly described in the following sections.

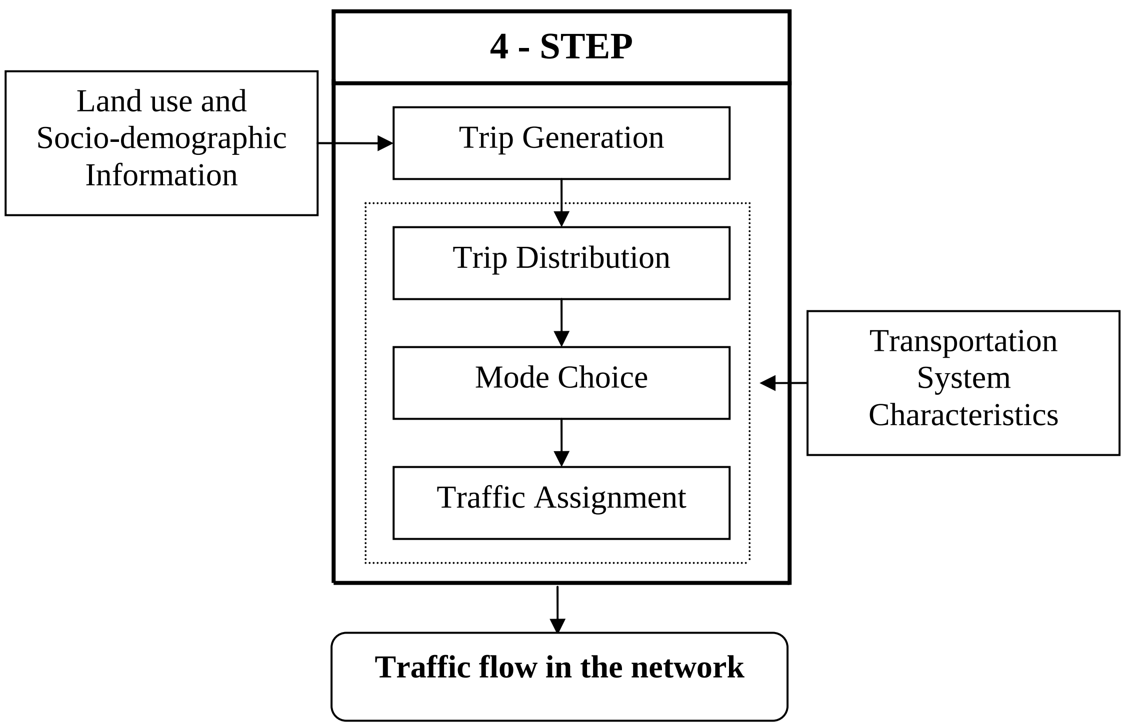


Figure 2.1 The Schematic Principle of Four-step Model

2.2.1 Trip Generation

Trip generation, the first step of the four-step planning process, estimates the total number of trips produced and attracted to each traffic analysis zone (TAZ). The basic idea is that the trips produced are estimated from the households and their socio-economic characteristics and the trips attracted are estimated from the employment classified by type. The trip generation model includes the application of both trip production and trip attraction models for each different trip purpose (home-based work, home-based other, non-home-based etc). The classification of the trip purposes varies among regional models and it depends on the modeling approach, level of aggregation and desired precision of the model. The variables usually used in the trip generation step are personal and household characteristics (age, household size, household income, car ownership), and land use characteristics (type of land uses, the number of employees, the value of the land and others). Two basic modeling approaches have been used in most applications of the trip generation step: category analysis used for trip production and regression models used for trip attraction.

Trip generation is an extremely important part of the model because any error generated in this module is carried on and multiplied through the whole system [7].

2.2.2 Trip Distribution

Trip distribution estimates the number of trips from each TAZ to each of the other TAZs while the trip generation only finds the number of trips that begin or end at a particular zone. Trip distribution links the trip ends to form an origin-destination pattern of trips. The trips are mostly estimated using a gravity model. The gravity model takes the trips produced at one zone and distributes them to other zones based on the size of the other zones (as measured by their trip attractions) and on the basis of the distance to other zones. A zone with a large number of trip attractions will receive a greater number of distributed trips than one with a small number of trip attractions and the farther away a zone is, the fewer trips it will attract. In the gravity model, the estimated number of trips between zones i and j, T_{ij}, is computed based on the following equation:

$$T_{ij} = P_i \left(\frac{A_j F_{ij} K_{ij}}{\sum_{j=1}^{zones} A_j F_{ij} K_{ij}} \right)$$

where,

T_{ij} : Number of trips from zone i to zone j
P_i : Number of trip productions in zone i
A_j : Number of trip attractions in zone j
F_{ij} : Friction factor relating the spatial separation between zone i and zone j
K_{ij} : Trip distribution adjustment factor for interchanges between zone i and zone j

The friction factors are inversely related to spatial separation of zones. The travel time between zones is usually used in the function of friction factors. If the travel time increases, the friction factor decreases. The output of this step is an Origin-Destination table (OD table) that shows the number of trips between zones in the study area.

2.2.3 Mode Choice

The role of mode choice in a model is to decide what modes of travel (car, public transportation. walking, or others) are used for each trip between the origins and destinations determined in the trip distribution step. Two types of discrete choice models are prevalent today: multinomial logit models (MNL) and nested logit models (NL). A multinomial logit model assumes equally competing alternatives while a nested logit model recognizes the potential for something other than equal competition among modes. Discrete choice models have been estimated either on aggregate (zone-level) data or disaggregate (household-level) data, and the most recent modeling efforts have focused on disaggregate nested logit models.

The variables required in a utility function typically include alternative-oriented variables such as transit travel time (waiting time, fare, in-vehicle and out-of-vehicle time), highway travel time, and auto costs, and they also include personalized traveler variables such as age, gender, income, etc. The model estimates the probabilities of transportation modes that a user might choose between an origin and destination.

2.2.4 Traffic Assignment

Traffic assignment is the last step in the traditional four-step travel demand model. This step assigns trips to the highway or transit network and the output becomes the basis for validating the models ability to replicate observed traffic volume in the base year as well as to evaluate transportation system improvements in the future years. The inputs for this step are the OD tables constructed by the previous step and appropriate modeled networks for the available modes.

The way trips are loaded onto the highway network is based on path-building algorithms. Two major path-building algorithms - all-or-nothing and stochastic assignment - have been used in traditional travel demand models. The all-or-nothing method assigns all of the trips to the minimum path which is based on link travel time or cost. This is the simplest method and is the foundation of several other methods such as capacity restraint and equilibrium assignment. Capacity restraint assignment considers the link capacity iteratively in which travel cost (time) increases if traffic on that link increases. The magnitude of a travel cost increases as assigned traffic approaches the capacity on the link. The increased travel cost (time) shifts traffic to other less congested links. The basic idea of capacity restraint is applied to equilibrium assignment. The equilibrium assignment process first involves the calculation of the shortest path from each origin to all destinations (usually the minimum time path is used). Trips for each O-D pair are then assigned to the links in the minimum path and the trips are added up for each link. The assigned trip volume is then compared to the capacity of the link to see if it is congested. If a link is congested, the speed on the link decreases which results in a longer travel time on that link. The shortest path may change due to the change of travel time. The whole process is iterated equilibrium when until no driver can improve travel time by changing paths. Trips on congested links will be shifted to uncongested links until this equilibrium condition occurs.

The stochastic algorithm estimates the probability that a trip will consider some other reasonable paths, and it loads proportions of the total trips to various paths based on the estimated probabilities. This method emphasizes the variability in drivers' perceptions of travel cost which consequently causes drivers' non-optimal route choice behaviors. The stochastic user equilibrium method deals with both capacity constraint and non-optimal route choice behaviors.

Among these assignment techniques, the proper method can be chosen based on the subject (person, freight, transit) for travel demand analysis and the level of detail in the modeled network. For instance, the analysis of transit and truck travel demand uses all-or-nothing assignment because of known routes, while personal travel demand typically uses equilibrium or stochastic assignment techniques since individual travelers may change their routes.

2.2.5. Limitation of UTPS

Traditional travel demand models take into account trips as a starting point, and model trips with respect to the systematic steps of trip generation, trip distribution, mode split, and roadway and transit assignment. These models are widely used in planning analysis, especially for the analysis of new infrastructure. However, traditional travel demand models have often been criticized for their limitations. The limitations of conventional travel demand models have been discussed by many authors and may be briefly summarized as follows [8, 9, 10]:

- *Traditional travel demand models ignore travel as a demand derived from activity participation decisions.*
- *They focus on individual trips, ignoring the spatial and temporal interrelationship between all trips and activities comprising individual activity patterns.*
- *They misrepresent overall behavior as an outcome of a true choice process, rather than as defined by a range of complex constraints that limit choice.*
- *Traditional models have inadequate specification of the interrelationships between travel and activity participation and scheduling, including activity linkages and interpersonal constraints.*
- *They misspecify individual choice sets, resulting from the inability to establish distinct choice alternatives available to the decision-maker in a constrained environment.*

- *Models are based strictly on the concept of utility maximization, and neglect substantial evidence relative to alternate decision strategies involving household dynamics, information levels, choice complexity, discontinuous specifications, and habit formation.*

2.3 Activity-based Approach

2.3.1 Overview

The basic idea of an activity-based model is that travel is a derived demand which is generated from people's activity participation. When people are engaged in different activities, the interrelationships among different activities with respect to temporal and spatial constraints are considered. The activity-based approach models the activities and the travel of people based on given characteristics of the households with respect to available time, transportation modes, transportation networks, and other constraints [11].

The roots of the innovative idea of an activity-based model go back to Hägerstrand (1970) with his time-space prism and to Chapin (1974) with his opinion that *"The observed patterns of trip making should be seen as a consequence of an individual's desire to participate in activities"* [12, 13]. Therefore, the analysis of travel behavior should be based on an understanding of the sequence of activities in which people engage during a day [14, 15].

Hägerstrand introduced three constraints associated with the activities such as "capability constraints", "coupling constraints", and "authority constraints". The capability constraints are biological constraints such as eating and sleeping. The coupling constraints reflect a certain activity for people to be at the same place at the same time. The authority constraints are various kinds of regulations such as opening hours of shops and working hours of employees. With the combination of these constraints over space and time, individuals can participate in activities within the prisms of time-space. Hägerstrand concluded that *"Each person lives in a time-space continuum and can only function in different points in time by experiencing the time and the cost of movement between the locations in a time-space prism"* [13].

An example of a simple activity-travel pattern is shown in Figure 2.2. The individual in this example went from home to work at about 8am. After 12pm he went to a bank and returned to his work. About 6pm this person left work, and on the way home stopped at a post office.

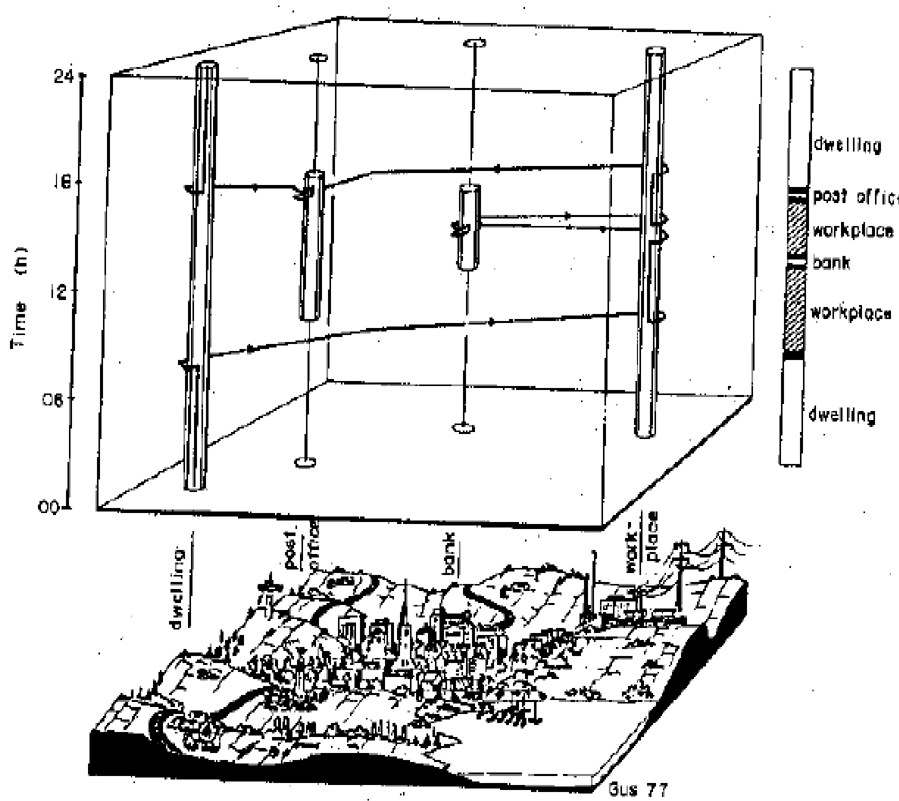


Figure 2.2 Example of a daily activity pattern in time and space (3D view)
Source: Papageorgiou (1991)

Understanding of the activity-travel patterns in this three-dimensional perspective is very important for all modeling efforts. The observed patterns of trip making should be seen as a consequence of an individual's desire to participate in activities. And the analysis of travel behavior should, therefore, be based on an understanding of the linked sequence of activities in which people engage during a day. This formulation was heavily influenced by the work of geographers such as Hagerstrand (1970) and Chapin (1974) and by work carried out on the use of time budgets during the 1960s (Szalai 1972).

Thus, within the activity approach, the basic focus is on the activity pattern, and the associated, contingent travel behavior to which it gives rise. The activity patterns of individuals are considered to be formed on the basis of their role within the household, as a result of the joint allocation and scheduling of tasks (e.g., working, shopping, child care) and resources (e.g., cars) and taking into account the short run time-space constraints of the current situation. Although time-space constraints have important effects in the short term, over the longer term the activity approach acknowledges that they are open to conscious change by the household. Activity patterns are also influenced by the activity space of the household which is formed through the personal experiences of its members and of their information gathering exercises.

A summary of the activity- based approach is as follows:

- Travel is derived from the demand for activity participation.
- Sequences or patterns of behavior, and not individual trips, are the relevant unit of analysis.
- Household and other social structures influence travel and activity behavior.

- Spatial, temporal, transportation, and interpersonal interdependencies constrain both activity-travel behaviors.
- Activity based approaches reflect the scheduling of activities in time and space.

2.3.2 Activity-based Model Applications

Many new frameworks for replacing conventional travel demand models have been proposed but only a few have been implemented and tested in real applications because the state-of-art methods remain firmly within the area of academic research [9,10]. Among recent operational activity-based models, four major classes are distinctive:

- Random-utility-based models [2, 16] (Portland model by Bowman and Ben-Akiva, 2000),
- Computational process approaches [19] (PCATS by Kitamura and Fujii, 1998, and AMOS by Pendyala et al., 1998),
- Rule-based approaches [5] (e.g. ALBATROSS by Arentze and Timmermans, 2000), and
- Activity presence-based approach [18, 23] (CentreSIM).

Algers et al. (2001) summarized well the recent practical activity-based models in their workshop report to figure out an appropriate activity-based modeling approach for supporting transportation planning in the Stockholm region [4]. The followings are brief descriptions that summarize the models introduced here based on the Algers' workshop report and other research related to activity-based models.

Portland Activity-based Model

Bowman and Ben-Akiva applied a daily activity schedule model in Portland, Oregon in 1998. It is closely related to traditional travel demand models [2], except that it is tour-based rather than trip-based. In the model system, a tour defined as the round-trips starting and ending at home is modeled with respect to the spatial and temporal connections between the activities. Each tour creates a daily activity pattern in which the individual's activity and travel demand are modeled. The tours are divided into primary tours, secondary tours, etc., for which the choices of destination, mode and time of day are modeled. The nested logit models with random utility maximization are used in model estimations [4]. The advantage of this model is that the modeling approach reveals the mechanism of how each activity is connected to the tour, and that it tries to implement all possible choices people make when they are traveling. Nested logit models are used to determine the probability of 114 choices making this model complicated in the practical application. Model calibration and validation are time-consuming, and fast, expensive computers are needed to run the model efficiently. Special generators were not considered in this model since an individual in a household is the basis of the travel demand model.

Prism-Constrained Activity-Travel Simulator (PCATS) Model

PCATS as implemented by Kitamura and Fujii in 1997 tries to represent Hägerstrand's concept of a time-space prism in a model system of activity engagement and travel simulation [19]. In the model system, activities are classified as fixed or flexible type. The fixed activity is defined as the activity is carried out at predetermined locations while the flexible can be occurred between the fixed activities. The time-space prisms are developed based on the speed of travel, locations and timing of the fixed activities in which the flexible activities are constrained. The individual is assumed to select those flexible activities and modes of travel that maximize the sum of the utilities of the activities and of the associated trips. The advantage of this model is that the models can be estimated from a travel diary by simply registering activities. However, the model system is composed of many sub-models that result in increased model complexity and computing time. PCATS does not include special generator modeling since it is focusing on how to replicate individual travel behavior.

AMOS (Activity-Mobility Simulator) Model

The AMOS model, developed by Kitamura and others, was designed to specifically examine travelers' responses to a set of travel demand management (TDM) strategies such as parking cost, improved bicycle and pedestrian facilities, and congestion pricing [20]. This prototype application was applied in the Washington, DC, area in the early 1990s. The data from an activity diary and stated preference survey of regular commuters was used to create a baseline travel pattern and generate a modified pattern as an adaptation to some change in the external conditions. Then, the dynamic micro-simulator replicated household responses to the proposed TDM measures. The simulator continued the search process until an acceptable pattern was found. The main advantage of this model is that the decision criterion is bounded rationality instead of using utility maximization. Thus, the decision-making process from individuals during their adapting periods to the TDM policy changes is investigated.

The critical component of this model is a neural network that predicts how a traveler's activity pattern will be changed by a particular policy. The dynamic micro-simulation and neural network makes the model more complex and needs excessive computing time. The model does not discuss travel behavior of the people in special land use categories. Hence, this is not helpful to develop a special generator activity-based model.

TRANSIMS

TRANSIMS is the activity-based model developed by the Los Alamos National Laboratory in 1995 [3]. The main strength of this model system is that TRANSIM has a microscopic simulation tools that are useful in the analysis of travel behavior affected by the changes in transportation policy or infrastructure [20]. The advanced micro-simulation tool combines activity-based trip generation and associated models for mode and route choice [21]. In the system, all populations within a study area can be synthesized and activity patterns are generated for the individuals in the population and these activities are later transformed into individual trip plans. These plans are inputs to the travel micro-simulation model. The simulations assign trips to the network based on different criteria like travel cost, travel time, departure time, congestion and safety. TRANSIMS has been tested in Dallas-Fort Worth, Texas, and Portland, Oregon. Costly surveys were needed to produce the library of activity schedules. Model development requires significant technical data, computer, and financial resources. TRANSIMS also does not consider the travel behavior affected by special land use.

ALBATROSS (A Learning Based Transportation Oriented Simulation System)

ALBATROSS is a prototype of a rule-based system for predicting travel demand. A team at Eindhoven University of Technology in the Netherlands developed it in 1998 [5]. Unlike models based on the utility-maximization principle, ALBATROSS schedules individuals' activities in a priority-based, rather than time-sequential, way. The purpose of this model is to develop a tool that can better assess the impact of changes from external factors such as shifting work hours or extending opening hours of shopping malls. In the model system, it is assumed that, with the given spatial and temporal constraints, individuals choose their daily activity patterns that satisfy their preferences. Thus, ALBATROSS can be used for policy analysis based on the behavioral changes of activity choice [4, 5]. The database in which activity-based approaches over recent years are accumulated is used within a system. However, this model also requires significant computing time and needs a large database from an expensive survey. For this reason, it is unknown whether the system actually has been applied to any other place outside the original study area of Rotterdam, Netherlands. The system is not concerned with the travel behavior of special generators.

CentreSIM (Centre County Simulation Model)

The CentreSIM, Center County Transportation Simulation model in Pennsylvania, is the first application of a spatial and temporal activity-based approach that uses activity participation within a traditional four-step planning framework [18]. The approach used in the CentreSIM is characterized as a disaggregated “presence-based” activity model. Unlike the conventional approach of activity-based modeling, CentreSIM assigns an individual to a spatial location and assigns a series of activities at the same time. The difference of changes in spatial distribution in consecutive hour is the travel demand. The main advantage of CentreSIM is to use activity data for a geography-based travel demand model and to provide the capabilities of more complex models without the need for significant investments in model development and data collection [18, 23]. Although the model structure of the CentreSIM seems to be much more practical for implementing activity data to real application, the elements in each step were not fully modeled. Further, an empirical case study is required. The data for the Centre County Study area, in particular the Penn State University special generator, is available.

Triangle Regional Model (TRM)

While the Triangle Regional Model is not an activity-based model, it is included here for comparison purposes. Furthermore, it is the type of traditional four-step model that may benefit from an activity-based model. The TRM is a state-of-the-practice regional travel demand model that is a tool for planning regional infrastructure improvements, transit services, and for evaluating air quality. [26]. The model component of the TRM is developed within TransCAD, which is a commercial transportation software that includes GIS tools for data processing and geographical resolutions of the travel demand model. This provides a good environment for accomplishing an empirical study by developing a university activity-based mode, incorporating activity-based university data into the TRM, and assessing the model from a view of regional perspectives. TransCAD and the TRM will be used by this research.

2.3.3 Summary

Table 2.1 shows the comparison of the activity-based models based on issues for practical application. Seven items such as model complexity, data requirement, computing time, etc, are evaluated for each activity-based model. All models require a large database for activity patterns for model development as well as long computing times caused by model complexity. Among the presented activity-based models, CentreSIM is found to be an appropriate model for practical application since the model is relatively easy to understand and does not require large data and computing time. Moreover, the model was originally developed for using activity data in a traditional travel demand framework, i.e., TransCAD**.** The activity data is available, and the data includes Penn State University which may transfer to the NC State University special generator.

Although the overall model system seems to be appropriate for this study in many respects, the model system of CentreSIM needs to be improved since the critical components of CentreSIM such as the synthesis of traveler groups, synthetic activity schedule, spatial activity capacity, and probability of trips were not entirely modeled. The CentreSIM model only used the data tables developed from the raw dataset. This is mainly because the main function of CentreSIM is to only provide future highway travel demand based on the assigned traffic volumes on the new highway system. Consequently, the model system does not have the capability of analyzing various transportation policies and TDM scenarios as most activity-based models do.

In this dissertation, the following model components are modeled and will enhance the potential ability of model applications to meet various demands in behavioral travel demand research.

- Synthetic activity schedule
- Destination choice model for developing spatial activity capacities
- Conditional probability of trips for estimating hourly travel demand

Table 2.1 Activity-based Travel Demand Model Application

Model	Model complexity	Available software	Data requirement	Available data	Computing time	Practical application (bridge)	Special generator analysis
ALBATROSS	High	Yes	High	No	High	No	No
AMOS	High	No	High	No	High	No	No
PCATS	High	No	Medium	No	High	No	No
Portland Model	Medium	No	Medium	No	High	Yes	No
TRANSIMS	High	Yes	High	No	High	No	No
CentreSIM	Low	No	Medium	Yes	Low	Yes	Yes
TRM (Traditional TDM)	Low	Yes	Medium	Yes	Low	Yes	Yes

CHAPTER 3

NORTH CAROLINA STATE UNIVERSITY ACTIVITY DATA

This chapter describes the activity survey data used for the analysis of university students' travel behavior and activity patterns. Two university student datasets, North Carolina State University (NCSU) and Pennsylvania State University (PSU), are compared to see whether they have similar travel patterns and to test data transferability. This chapter describes the method of data collection as well as the data representation that is essential for the comparison travel patterns between two school students. Also, the chapter provides a descriptive statistics analysis of the dataset.

3.1 Activity Travel Survey Data

North Carolina State University (NCSU) is located in Raleigh, North Carolina. The area of 2,110 acres consists of six regions: 1) Centennial Campus, 2) South Campus, 3) Central Campus, 4) North Campus, 5) the College of Veterinary Medicine, and 6) the Stadium, as shown in Figure 1. According to year 2004 statistics from the NCSU official website, the total population of students is 29,637 and the university has nearly 7,000 employees, including approximately 1,800 faculty members.

In the spring of 2000, a travel survey was conducted by the staff of the NCSU Department of Transportation [24]. The goal of survey was to improve the estimation of university-related travel in the Triangle Regional Model (TRM). In addition, the data set was expected to be used to help design, estimate, and calibrate a model for travel within and between the NC State campuses. Although the NCSU staff originally organized and performed the survey specifically for the purpose of estimating a campus travel demand model for the TRM, they expected it to be used in developing a travel demand model for NCSU campus. Moreover, they hoped that if the NCSU model is developed, it could be used in conjunction with the larger Triangle Regional Model to assess future conditions and to evaluate the impacts of various alternative transportation scenarios [24].

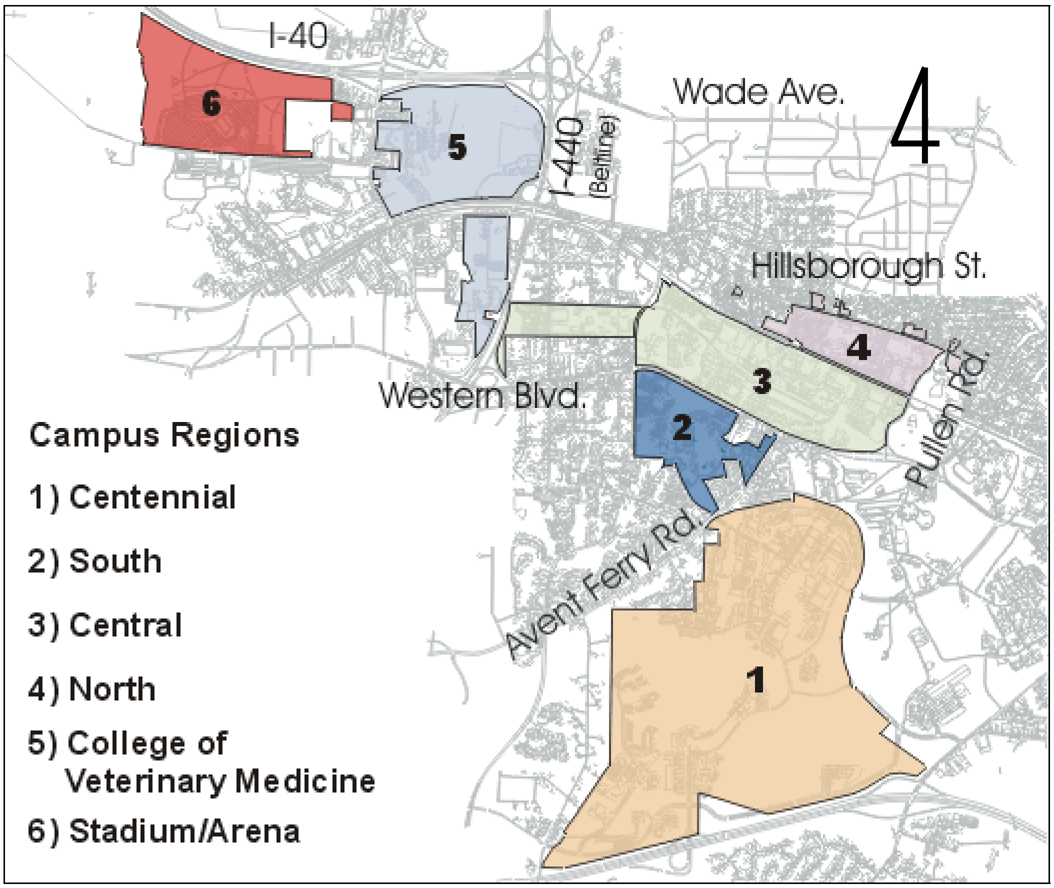


Figure 3.1 Map of North Carolina State University
(Source: NCSU, http://www.ncsu.edu/campus_map/)

NCSU DOT staff constructed a random sample of students representing the total NCSU student population whom they asked to complete a travel diary for a particular day. As shown in Table 3.1, 843 students returned survey questionnaires. The questionnaires collected data on student residential location (on-campus, off-campus), student characteristics (gender, age, and university status – freshman, sophomore, junior, senior, graduate student), location of daily activities, and travel mode to activities. Samples by student status were almost evenly distributed among the on-campus and off-campus students. Since freshmen, sophomores, and juniors mainly live on-campus, they returned more survey questionnaires than seniors and graduate students who usually live off-campus.

Table 3.1 Summary of Students Surveyed

Classification	Residence location		Total
	On-Campus (%)	Off-Campus (%)	
Freshman	159 (90.3%)	17(9.7%)	176 (21.0%)
Sophomore	130(79.8%)	33(20.2%)	163 (19.3%)
Junior	78(51.0%)	75(49.0%)	153 (18.1%)
Senior	46(30.1%)	107(69.9%)	153 (18.1%)
Graduate	16(8.1%)	182(91.9%)	198 (23.5%)
Total	429(50.9%)	414(49.1%)	843 (100.0%)

Table 3.1 describes the number of students surveyed, categorized by class year and their residence location. The post-graduate and special categories are classified into ‘Graduate’ category based on assumption that these travelers have similar trip-making characteristics as graduate students. Of all the class years, the freshman category has the largest representation in the survey data [24]. Somewhat more students who lived on-campus were surveyed (freshmen, sophomores and juniors) than students who lived off-campus (most seniors and graduate students). The responded total number of students living on- and off-campus is almost equal.

3.2 Travel Survey and Descriptive Statistics

The travel survey questionnaires are designed with four major categories (Appendix A):

- Student household survey,
- Personal survey,
- Place , and
- Trip survey.

The student household survey includes questionnaire about students' residential locations classified by on-and off-campus. If students live on-campus, they need to provide the residence hall name, otherwise (off-campus resident) address with zip code is provided. This information is critical for geocoding the address on the map. Geocoding is the process that relates a given address or building name to a on a map indicating individual residence longitude and latitude. The personal survey includes questions about the students' individual characteristic such as age, gender, student status, employer's address etc. The place and trip survey consists of questions to identify where trips originate and terminate and what student activities relate to individual trips. The activity types are classified by 11 main activities: 1) School/Class, 2) Study/Research, 3) Work/Volunteer, 4) Drop off/Pick up, 5) Eat meals, 6) Social/Recreational , 7) Shop, 8) Doctor/Other professional, 9) Family/Personal activities, 10) Sleep, and 11) Others. For each trip the questionnaires asked for the main and secondary activity related to the trip.

The descriptive summary statistics in Table 3.2 show that more male (58.4%) and undergraduate students (76.5%) were sampled compared to female and graduate students.

Table 3.2 Descriptive Statistics of Surveyed Student Characteristics

Variable	Description	Coding	Statistics
Age	Average age	Continuous	23.5
Gender*	If the person is male, then 0	0: Male	492 (58.4%)
	Otherwise, 1	1: Female	351 (41.6%)
Student status*	Educational status	0: Undergraduate	645 (76.5%)
		1: Graduate	198 (23.5%)
License*	Licensed to drive	0: yes	812 (96.3%)
		1: no	31 (3.7%)
Vehicle	Average number of vehicles available	Discrete	0.92 (per HH)
Employment*	Employment status	1: full-time	114 (13.5%)
		2: part-time	325 (38.6%)
		3: volunteer	9 (1.1%)
		4: unemployed	395 (46.9%)

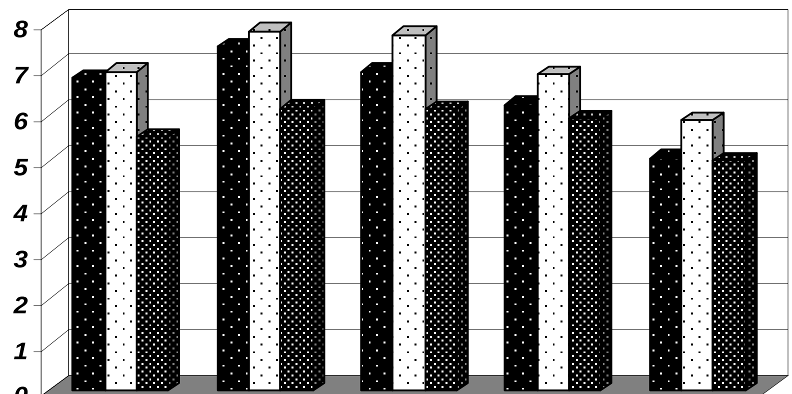
Note: The variables with an asterisk are categorical.
Statistics of categorical variables are denoted as frequency (percentage).

Of the 843 students, 812 students (96.3%) had a driver's license, and the average number of vehicles available was 0.92 (per household or residence). More than half of the students were employed: 114 (13.5%) students were full-time workers; 325 (38.6%) students were part-time workers; and other 9 (1.1%) students were volunteers.

3.3 Observed Travel Characteristics

3.3.1 Trip rates

Figure 3.2 shows the average number of trips by student group classified by residence status (on- and off-campus) and educational status (freshman, sophomore, junior, senior, and graduate). One of interesting findings is that students living on-campus make more trips than off-campus students. One possible reason is that on-campus students are more likely exposed to the opportunity of having diverse university activities than off-campus students.



	Freshman	Sophomore	Junior	Senior	Graduate
Combined	6.81	7.49	6.95	6.22	5.07
On-campus	6.94	7.82	7.74	6.89	5.88
Off-campus	5.53	6.15	6.12	5.93	5.00

Figure 3.2 Average Daily Trip Rates by Student Group

Table 3.3 shows whether the student groups have statistically different numbers of trips per day. The negative binomial model is used to analyze the discrete trip frequency for each individual in student groups. The dependent variable is the number of trips and independent variables are student groups which are coded with categorical format as shown in Table 3.3.

The reason to choose negative binomial model instead of Poisson model is that count data often show greater variability than the Poisson model allows. Poisson distributions have identical mean and variance. If the variances of data are much larger than the means, the greater variability is called over-dispersion which can be handled with the negative binomial model. Therefore, most researchers modeling count data have used the negative binomial model [27].

A negative binomial model can be written as

$$ln Y_i = X_i \beta + \varepsilon_i \tag{3.1}$$

Where X is the vector of covariates, β is the vector of coefficients, and ε is the error term. The negative binomial distribution has the probability mass function:

$$f(y;k,\mu)=\frac{\Gamma(y+k)}{\Gamma(k)\Gamma(y+1)}\left(\frac{k}{\mu+k}\right)^k\left(1-\frac{k}{\mu+k}\right)^y, \quad y=0,1,2;; \tag{3.2}$$

where, k and μ are parameters. This distribution has E$(Y) = \mu$, $var(Y) = \mu + \mu^2/k$ and k^{-1} is called a dispersion parameter. As k^{-1} closes to zero, then var(Y) closes to μ and the negative binomial distribution converges to the Poisson. Usually, k^{-1} is unknown so estimating it helps summarize the extent of overdispersion [27]. Overdispersion is not an issue in ordinary regression with normally distributed Y, since that distribution has a separate parameter to describe variability. In contrast to ordinary regression, binomial and Poisson distribution has the variance as a function of the mean.

Table 3.3 shows the results from modeling trips by classified student groups. The parameter estimated in each student group present a comparative value against one base (or reference) variable indicated by (+) in the group.

Table 3.3 Analysis of Student Group Effect on Number of Daily Trips

	Number of Trips in a Day		
Variable	Coefficient	χ^2	Pr> χ^2
Residence Status			
Off-campus	-0.0874	5.58	0.0183**
On-campus+	--	--	--
Gender			
Female	0.0323	0.91	0.3406
Male+	--	--	--
Driver's License			
No	-0.2014	5.01	0.0252**
Yes+	--	--	--
Student Status			
Graduate	-0.0973	3.29	0.0696*
Undergraduate+	--	--	--
Undergraduate Class			
Freshmen	0.0977	3.47	0.0626*
Sophomore	0.1208	5.01	0.0252**
Junior	0.0832	2.45	0.1173
Senior+	--	--	--
Employee Status			
Full-time	-0.2021	11.97	0.0005***
Part-time	0.0227	0.43	0.5143
Volunteer	0.0813	0.31	0.5772
Unemployed+	--	--	--
Model Goodness of Fit			
N	822		
Pearson- χ^2	805.98 (df: 811)		
Log-likelihood	4496.42		

Note: (+) indicates that this variable is a base variable in the group.
Significant: 0.001 `***', 0.05 `**', 0.1`*'.

For example, the coefficient of the off-campus variable is compared to the on-campus base variable which does not have an estimate. From the results in Table 3.3, the sign of the coefficient of off-campus is negative which means that off-campus students are less likely to make a trip compared to on-campus students. Also, the undergraduate students and students who have a driver’s license are more likely to travel than graduate students and students who do not have a driver’s license, which corresponds to trip rates in Figure 3.2. Unlike the variable such as residence status, driver’s license, and student status, gender is not statistically significant to students’ trip-making behavior. Among undergraduate students, sophomores are more likely to travel than seniors, and freshmen and juniors seem to make more trips, but freshmen and juniors are not statistically significantly different to seniors. The full-time working students are less likely to travel compared to part-time, volunteer, and unemployed students. For comparison to the regional average of total trip rates, the TRM household survey showed that the average daily person trip rate for household members was about 6 (trips/day) [26]. Thus, the average trip rates of NCSU students (5.0 to 7.8 trips/day, Figure 3.2) are higher than those of households in the TRM. The magnitudes of the trip rates and the number of students (28,000) create significant NCSU trips within the TRM model, similar to a small city.

3.3.2 Travel mode characteristics

The students’ mode choice characteristics are analyzed from tables of descriptive statistics of modes by trips (primary, secondary, and tertiary in Table 3.4) and residential locations (on-campus and off-campus students in Table 3.5).

Table 3.4 presents the overall percentage of travel classified by six modes: auto, city transit (taxi, bus), campus shuttle ("Wolfline"), walk, bicycle, and other. As expected, about 60% of campus trips are made by walking and 30% by automobile. The walk mode is mostly associated with on-campus trips from class to class, class to research lab, etc. In contrast, students living off-campus mainly use automobiles for their commutes from home to campus, whereas students living on-campus do not use automobiles as frequently as off-campus students. NCSU students used the campus shuttle for commuting from the main campus to a nearby research campus. Bicycles are not as frequently used as walking.

Table 3.4 Descriptive Statistics of Travel Characteristics

Variable	Description	Coding	Statistics
NTRIP	Average daily number of trips	Discrete	6.13/day
TREMETH (n=6021)*	Travel Method	1: Auto	1823 (30.28%)
		2: Auto-passenger	262 (4.35%)
		3: Taxi	2 (0.03%)
		4: Motorcycle	0 (0.0%)
		5: TTA	4 (0.07%)
		6: CAT	2 (0.03%)
		7: Wolfline**	261 (4.33%)
		8: Other bus/van	11 (0.18%)
		9: Walk	3597 (59.74%)
		10: Bicycle	48 (0.80%)
		11: Intercity mode	0 (0.0%)
		12: Other	1 (0.02%)
		13: Skateboard	10 (0.17%)
AUTOCC (n=2083)	Average number of person in vehicle	Discrete	1.40/veh
NTIME (n=3647)	Walk/Bike Time	Continuous	8.51 (minutes)

Note: '*' total observations valid for each variable, '**' NCSU Campus shuttle bus

Mode share may vary among universities due to topography and building locations. The NCSU main campus is flat and buildings are close together promoting walking. The research campus is hilly and separated from main campus by a ridge and major two major arterials – factors promoting campus shuttle use.

The auto occupancy was 1.4 (persons/vehicle), which is slightly higher than that the 1.1 to 1.4 reported from by the TRM regional household survey [26]. This suggests that on-campus students are more likely to carpool. The average walk/bike time was about nine minutes. It was seen that most students walk on campus because classes and facilities are relatively close and student cars on campus are restricted except during 5 pm to 8 am.

For further investigation, the travel modes by trip type and by residence location were categorized as shown in Table 3.5. The trips were segmented into three groups: primary, secondary, and tertiary trips. A primary trip is a commute trip from residence to a campus. For example, if an off-campus student makes a trip to campus using a car, then the primary mode is an automobile. If the student parks at a park-and-ride lot and rides a shuttle to campus, then the second trip uses a secondary mode transit. Then if the trip requires walking from a bus stop to the ultimate destination, walking is considered a tertiary mode.

Table 3.5 shows that the most frequent mode of primary trips for on-campus students is walking (79.9%), and the automobile is the most frequent mode for off-campus students (68.9%). Off-campus students use the campus shuttle as much as on-campus students. However, the overall ridership on the shuttle bus is only 3.5% for the primary trip to campus. Secondary trips are mainly associated with a transfer to another mode or a final trip to a destination. Therefore, walk and shuttle dominate secondary trips with more than 95% of the secondary trips. Other modes such as city transit, bicycle, and automobile do not seem to be highly used for secondary trips. Finally, walking is shown as the only mode for tertiary trips. Yet, walking does not seem to be frequently used as a tertiary trip with only 69 walking tertiary trips made by on- and off-campus students. Most students ended their trips at secondary trip, which means that students do not seem to frequently transfer from one mode to another.

Table 3.5 Travel Modes by Trips and Residence Location (percentage)

	Primary trip			Secondary trip			Tertiary trip		
Mode	On-campus	Off-campus	combined	On-campus	Off-campus	combined	On-campus	Off-campus	combined
Auto	491 (15.7)	1583 (68.9)	2074 (38.2)	4 (3.4)	7 (1.8)	11 (2.1)	--	--	--
Transit	3 (0.1)	12 (0.5)	15 (0.2)	--	4 (1.0)	4 (0.8)	--	--	--
Shuttle	100 (3.2)	92 (4.0)	192 (3.5)	6 (5.2)	63 (15.8)	69 (13.4)	--	--	--
Walk	2507 (79.9)	596 (25.9)	3103 (57.1)	105 (90.5)	319 (79.8)	424 (82.2)	25 (100.0)	44 (100.0)	69 (100.0)
Bicycle	25 (0.8)	16 (0.7)	41 (0.8)	1 (0.9)	6 (1.5)	7 (1.4)	--	--	--
Other	10 (0.3)	0 (0.0)	10 (0.2)	--	1 (0.3)	1 (0.2)	--	--	--
Total	3136	2299	5435	116	400	516	25	44	69

3.4 Activity Characteristics

3.4.1 Activity participation characteristics

Student activities were classified as follows: school-class, study-research, work-volunteer, drop off-pick up, meals, social-recreation, shop, doctor-other professional, family-personal, and sleep. The cleaned sample of 843 student surveys showed the total frequency of

4883 activities under taken by 698 students during the day. The frequency of activity types by trip purpose is presented in Table 3.6. As expected, the most frequent activity in daily student life is school-class. Of 4883 activities, 2222 (45.5%) corresponded to university work-related activities including all three purposes such as school-class, study-research, and work-volunteer. Besides work-related activities, meals and social-recreation had high frequencies.

Table 3.7 summarizes activities by average number per day, daily time allocations by students, and average travel time devoted to the activity. The average number for each activity is computed across all students in the sample. The average activity duration represents the mean duration across all activities of the particular activity. The average travel time for each activity type is defined as the average travel time to the location for the activity.

Table 3.6 Frequencies of Activities

Activity	Frequency	Percent (%)
School/Class	1160	23.76
Meals	847	17.35
Study/Research	798	16.34
Social/Recreational	562	11.51
Family/Personal activities	519	10.63
Sleep*	341	6.98
Work/Volunteer	264	5.41
Shop	219	4.48
Drop off/Pick up someone	152	3.11
Doctor/Other professional	21	0.43

Note: The total frequency is 4883 after cleaning no-responded answers.
*Sleep does not include night time sleep, rather mid-day sleep.

As expected, the average number of the "school/class" activity is the most frequent activity (1.61 per person per day) illustrating that students usually attend classes more than once per day. Also, students go to library or lab for studying or doing research at least once daily. The "meals" activity is the second most frequent. However the frequency is lower than expected. Another interesting finding is "sleep" activity. This sleep activity may be explained by students returning to their residences (mostly on-campus dormitory) and taking naps between classes.

Table 3.7 also summarizes the average activity duration. The travel time is the time difference between the departure time and arrival time. On the other hand, activity duration is the time difference between arrival time and departure time for a next activity. The following equations illustrate the definition of travel time and activity duration.

$$TT_j = time_a^j - time_d^{j-1} \tag{3.3}$$

$$AD_j = time_d^j - time_a^j \tag{3.4}$$

Where,

TT_j : travel time for activity j

$time_d^{j-1}$: departure time at previous activity j-1

$time_d^j$: departure time at activity j

$time_a^j$: arrival time at activity j

AD_j : duration of activity j

From the results, school, study, and work activities were allocated large proportions of daily time budgets. Note that the average duration of work-volunteer activity (about 4 hours per day) is about equal to the duration of study-research plus school-class activities. This is explained by some students (52%) working either full-time or part-time. Drop off/Pick up is the shortest whereas school related activities are the longest overall. Social-recreation also occupied a substantial portion of the time budgets (86.6 minutes per day). Shopping activity occurred ten times more frequently than the "Doctor-other professional activity", but Doctor-other professional activity took longer.

The overall average travel time for activities is 12 minutes. This indicates that the majority of trips, trip origins and trip destinations are within the university or immediate area. Activities such as class, study, meals, social-recreation, and shopping occurred mainly on or near campus and they have similar travel times of about 12 minutes. Also it was found that students spent more time traveling for Work-volunteer and Doctor-other professional activities suggesting that students mostly work off campus.

Table 3.7 Time Allocations for Daily Activities

Activity	Average number of activities[1)]	Average duration (min)[2)]	Average Travel Time (min)[3)]
School/Class	1.61	93.7	11.41
Study/Research	1.06	139.9	11.52
Work/Volunteer	0.40	238.5	15.16
Drop off/Pick up someone	0.19	14.9	12.03
Meals	1.10	50.5	11.48
Social/Recreational	0.73	86.6	11.84
Shop	0.27	21.1	10.80
Doctor/Other professional	0.03	51.1	20.48
Family/Personal activities	0.67	59.3	12.02
Sleep[4)]	0.68	186.8	12.01
Overall Average	6.76	93.6	12.00

Note: 1) Number of activities per person per day
2) Average duration per activity (Total activity time/total frequency of activity)
3) Average minutes spent on travel for each activity per day
4) Sleep does not include 'sleep' at night, rather mid-day naps.

3.4.2 Activity sequencing

The sequencing of activities is examined by creating a transition matrix of one activity to another. The transition matrix represents the likelihood that a subsequent activity of a certain type will occur given an activity of a current type [35]. In the matrix, all surveyed 5435 trips were allocated by their current and consequent activities, where activity is defined as the main activity associated with the trip as shown in Table 3.8. The rows in Table 3.8 indicate the current activity, and the columns represent the subsequent activity. The value in each cell indicates the percentage of occurrence of a subsequent activity after a current activity.

One of most interesting findings from the transition matrix is that the most frequent subsequent activity following a current activity type is meal activity. This activity appears five times as the most frequent activity in the columns of ten subsequent activity types As shown in the transition matrix, the current school/class activity followed by meal activity is 23% and work and social-recreation activities followed by meal activity are 24% and 21% respectfully. Also, the school/class and research activities are the most frequent subsequent activity following a meal activity. This reveals that students usually have meals between classes, research, or social-recreational activities. However, other personal, sleep and drop off/pick up activities are less likely to follow a meal activity. Except for the meal activity, academic related activities such as class, research, and work activities are more likely to follow academic activities. This makes sense because students often go to another class, library or research lab after class or research. Another interesting finding is that the most frequent activity among all possible combinations in the transition matrix is a school/class activity undertaken immediately after a sleep activity (52.6%). This could be explained by the fact that students wake up each day and take morning classes before eating breakfast or lunch. The frequency (4.4%) of a sleep activity occurring after a school/class activity suggests some students living on-campus take a nap between classes.

3.5 Summary

The major findings from the analysis of the NCSU student activity-travel behavior survey can be summarized as follows.

- Undergraduate students and on-campus residents are more likely to be involved in various activities than graduate students and off-campus residents. Consequently, the trip rates of on-campus students are higher than the off-campus students.
- Among undergraduate students, sophomores are more likely to make a trip. However, a statistical analysis indicates that there is no significant difference between their number of daily trips.
- Walk is a primary mode for on-campus residents while the automobile is the major mode for off-campus residents. Walking secondary trips are the most frequent for both on- and off-campus students. Auto use on campus is not as high as for a standard household, but auto occupancy is higher.
- Among all activities meal activity is the most frequent subsequent activity.
- The most frequent activity transition is school/class undertaken immediately after sleep.

Findings indicate that activity-travel behavior of university students is greater than standard households in terms of trip rates. As a result, activity-sequence is likely more diverse.

Table 3.8 Activity Transition Matrix (percentage)

Current Activity	Subsequent Activity: Doctor's appointment	Drop/ pick up	Meals	Other personal	School/ Class	Shop	Sleep	Social/ Recreation	Research	Work/ Volunteer
Doctor's appointment	0.0	4.5	**27.3**	9.1	13.6	9.1	9.1	13.6	9.1	4.5
Drop/pick up	0.6	6.1	15.3	11.0	10.4	8.6	11.0	**16.0**	14.1	6.7
Meals	0.5	3.3	2.3	10.9	22.2	3.3	8.6	14.3	**26.4**	8.3
Other personal	0.4	3.5	**20.8**	8.7	20.6	8.7	3.7	17.1	11.1	5.4
School/Class	0.2	2.2	**23.0**	12.2	22.5	2.2	4.4	9.9	19.8	3.5
Shop	0.4	6.1	14.0	**16.2**	6.6	9.6	12.2	11.4	16.2	7.4
Sleep	0.6	3.0	10.6	3.9	**52.6**	2.7	0.1	6.1	7.6	12.8
Social/Recreation	0.2	3.2	**21.3**	11.9	12.1	3.8	14.1	16.4	13.7	3.4
Research	0.5	1.5	22.8	8.6	**26.7**	4.0	8.6	9.2	14.6	3.2
Work/Volunteer	0.6	3.9	**24.3**	13.9	14.8	7.1	10.7	8.0	12.2	4.5

Although the majority of trips occurred on campus between campus buildings, their travel patterns may affect regional travel demand due to their flexible schedule and higher trip rates. This effect should be reflected in a regional travel demand model. It is not sufficient to classify university trips as home-based school trips or the aggregated trip rates of one-person households with low-income – typical practices.

CHAPTER 4

MODELING ACTIVITY–BASED TRIP GENERATION USING ACTIVITY PROFILE

This chapter introduces a spatial-temporal activity-based model for trip generation and trip distribution and provides practical methodologies in which the required dataset and models are discussed. A university special generator activity-based model is developed, and the transferability of university travel data between North Carolina State University (NCSU) and Pennsylvania State University (PSU) is investigated. The model output is presented.

4.1 Activity-based Model Components and Modeling Procedures

The activity-based approach in travel demand modeling is derived from the activities people do at certain times. The derived travel demand is the result of the linkage among activities that happened at spatially different locations at certain times of day. Thus, it is critical to model the interactions among activities and people (travelers) in the activity-based approach.

In order to model the interactions, the activity types and groups of people have to be clarified since not all activities can be included in the model due to limited data, time and budget. The activity types are grouped into five activities: home/maintenance, work/school, shopping, recreation, and other/services. This classification of activities varies depending on the purpose of model application and its level of aggregation. People as members of households are classified based on the level of income, household size, employment type, and car ownership (as classified in traditional travel demand models). By considering this socio-demographic information, people in a same group are assumed to be relatively homogenous in activity patterns and trip-making behaviors.

Another critical model component of activity-based approach is activity patterns or schedule which is defined as a sequence of activities or a schedule of activities in space and time. Activity patterns are illustrated by various elements that have been widely modeled in research to create synthetic activity schedules by traveler group. Following is the list of the elements forming an activity schedule:

- Activity frequency,
- Activity duration and time allocation,
- Spatial (space) allocation,
- Departure time decision, and
- Stop and trip chain characteristics.

The complexity of the activity-based approach depends on how detailed these elements are modeled and how disaggregated the model is. As mentioned before, this complexity has been a challenge for planning agencies to apply this approach in state-of-practice applications.

This study introduces the transitional activity-based approach which can be developed based on standard household travel surveys, and this approach can be incorporated into a conventional travel demand model within the trip generation and distribution steps.

Figure 4.1 depicts the summarized procedure of the activity-based NCSU travel demand model. Five main steps describe the activity-based model development:

1. Disaggregate the university population into several groups,
2. Synthesize activity schedules by each traveler group,
3. Develop spatial activity capacity,
4. Allocate the university population on-campus, and
5. Estimate travel demand.

First, the university travelers need to be aggregated. For example, they can be aggregated into several groups like students, faculty, and staff. An analysis of activity-travel pattern by group is required to confirm that the people in a same group have similar activity-travel patterns.

Next, the activity profiles representing travelers' activity sequencing and the probability of following one activity by another for twenty-four hours are created based on an activity diary.

Then, the spatial activity capacities of university buildings (destinations) are created by developing destination choice models to distribute travelers over a spatial activity dimension (e.g., study area). This represents the attractiveness of each facility by activity type in the university, which is also used in calculating relative attractiveness of activities.

Next, the spatial-temporal allocation of the university population determines the presence of travelers in each building, namely building presence, which represents the number of people present in each building in every hour segment.

Finally, modeling conditional probabilities of travel for a given time and given activity type for each traveler group estimates the number of trips (productions and attractions) for the people in each facility.

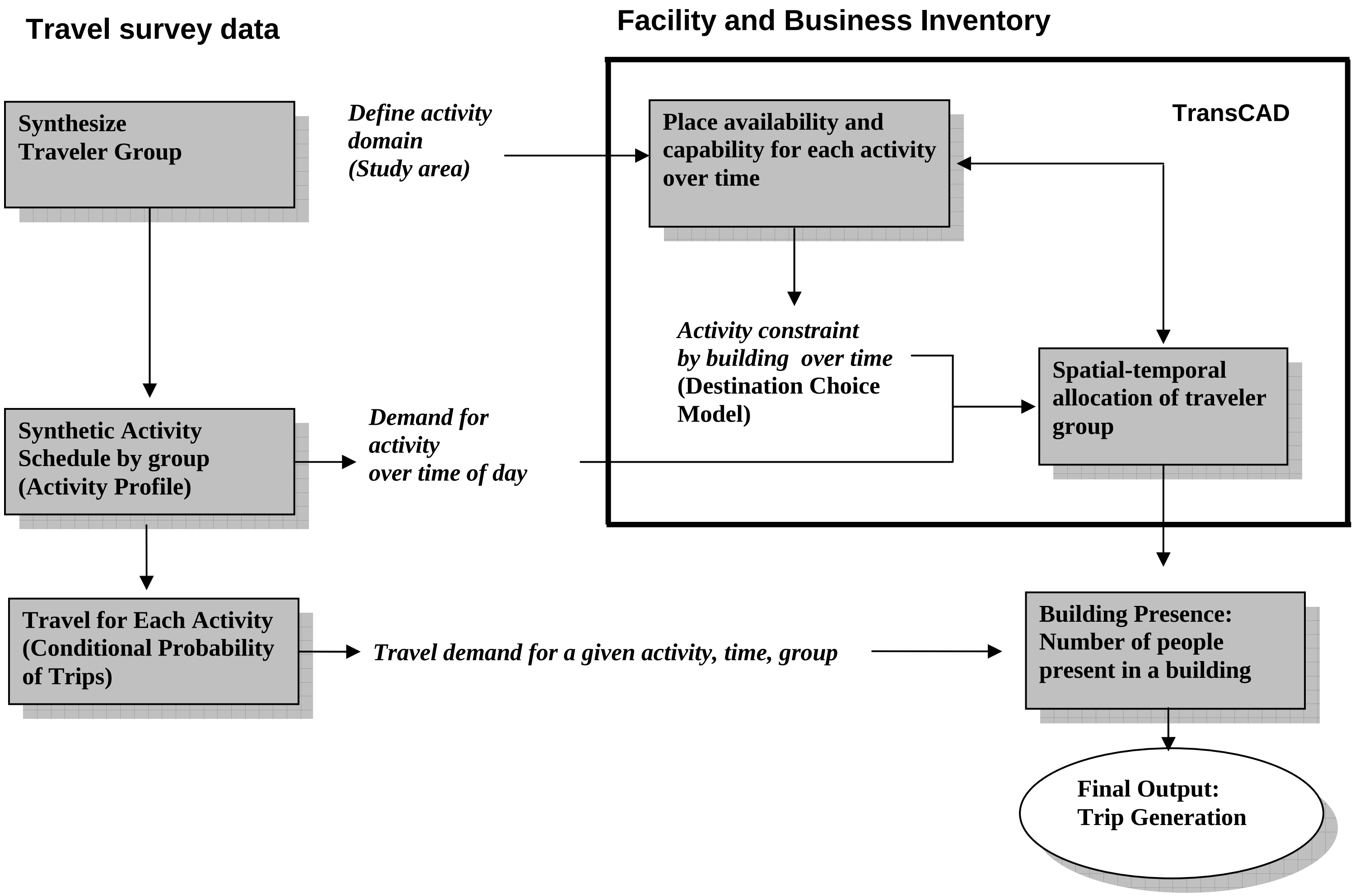


Figure 4.1 A Spatial and Temporal Activity-based Model for Trip Generation

4.2 University Traveler Groups

The trip generation step in a traditional four-step travel demand model usually uses trip rates classified by household size and auto ownership or income level such as one-person low income, two-person mid-income, etc. Unlike conventional travel demand models that model a community with households as units of analysis, clustering traveler groups in a university is much clearer because of the well-defined demographics of students, faculty, and staff. In order to classify university traveler groups, an a priori classification of persons is chosen based on the characteristics of population in the university. Then, the similarities in the activity patterns of each class are investigated to decide whether the groups are appropriately classified.

Table 3.3 in Chapter 3 is recalled here to provide a better idea of clustering university demographics.

Table 4.1 Analysis of Student Group Effect on Number of Trips in a Day

	Number of Trips in a Day		
Variable	Coefficient	χ^2	Pr> χ^2
Residence Status			
Off-campus	-0.0874	5.58	0.0183**
On-campus+	--	--	--
Driver's License			
No	-0.2014	5.01	0.0252**
Yes+	--	--	--
Student Status			
Graduate	-0.0973	3.29	0.0696*
Undergraduate+	--	--	--

Note: (+) indicates that this variable is a base variable in the group.
Significant: 0.001 `***', 0.05 `**', 0.1`*'.

Table 3.3 shows that the students' trip-making behavior is significantly different within the groups classified by residential status (on-and off-campus students), student status (undergraduate and graduate student), and driver's license. The residential status and student status are considered in synthesizing the traveler group for university students for model simplicity. The driver's license is not considered because the number of surveyed students with no driver's license is too small (31 students out of 843) as shown in Table 3.2.

Faculty and staff need to be considered in other traveler groups which are supposed to be different from students in terms of trip-making behavior. As mentioned in Chapter 3, the travel data for faculty and staff are not available from the NCSU survey while the data for both faculty and staff are available from the PSU activity survey. Therefore, the two traveler groups, faculty and staff, may be included in the activity-based model based on the assumption that faculty and staff at PSU have similar in trip-making behavior as those at NCSU.

The statistical test for similarity of activity pattern and trip-making behavior between two university students will be informative to ascertain the hypothesis that the university special generator data may be transferable. The six traveler groups are finally defined in this study as shown in Table 4.2.

Table 4.2 Traveler Group in University

Student Status	Residential Status	Traveler group
Undergraduate	On-campus	1
	Off-campus	2
Graduate	On-campus	3
	Off-campus	4
Faculty		5
Staff		6

4.3 Definition of Activity Type

The student activity survey conducted by the NCSU Department of Transportation classifies eleven main activities: 1) School/Class, 2) Study/Research, 3) Work/Volunteer, 4)

Drop off/Pick up, 5) Eat meals, 6) Social/Recreational , 7) Shop, 8) Doctor/Other professional, 9) Family/Personal activities, 10) Sleep, and 11) Other. In this study, these eleven activities are regrouped into five major activities to simplify model development regarding activity schedule and activity constraint.

The five main activity types are defined here as home/maintenance, work/school, shopping, recreation, and other/services. Table 4.3 shows these activity functional classes with different activities, which are commonly used in the activity-based approaches in travel demand modeling [10]. Thus, the eleven activities in the NCSU survey and the student responses can be reclassified as shown in Table 4.3. This classification could be changed if necessary depending on the scope of activity-based model development or with various model applications. For example, the activity 'volunteer work' is included in the category of 'Other/Service' in the original classification for a standard household activity-based model, but this study allocates that activity into the 'Work/School' category since 'volunteer work' mainly happens on campus.

A simple program of if-then-else statement was written to recode the five main activities in the NCSU data file with a number system of activity types: Home (1) denotes all in-home activities including maintenance; Work (2) corresponds to all work related or school related activities; Shopping (3) relates to all shopping related activities; Recreation (4) focuses on leisure activities and recreation; and Other (5) includes all activities not in the four other activities [31].

Table 4.3 Activity Functional Classes

Class	Activities	NCSU survey
(1) Home/Maintenance	Cleaning/maintenance at home Meal preparation Meals at home Shower/dress	10)* Sleep Other home activities
(2) Work/School	Work School (Student only) Volunteer work	1) School/class, 2) Study/research, 3) Work/volunteer
(3) Shopping	Department store Grocery/drug store Convenience store Video rental store Personal service (hair, nail, etc)	7) Shop
(4) Recreation	Fitness center Exercise Movies in theaters Sports Watching videos Walk dog Meal outside Relaxation/rest Hobbies at somewhere Bars Party Jogging, biking, etc	5) Eat meals (outside) 6) Social/Recreational
(5) Other/Service	Seeing a doctor Pick-up/drop-off kids Banking/ATM Shipping Errand Visiting Religious events	4) Drop off/pick up 8) Doctor/other professional 9) Family/Personal activities 11) Others

Note:'*' is the activity number classified in the original NCSU database.

4.4 Synthetic Activity Schedule for Each Traveler Group

In this study, the activity profile defined as the sequence of all activities and trips is obtained from observed daily activity participation of students in hourly intervals as shown in Figure 4.2. The activity profile shows how many people engage in each activity in each hourly time frame during a day.

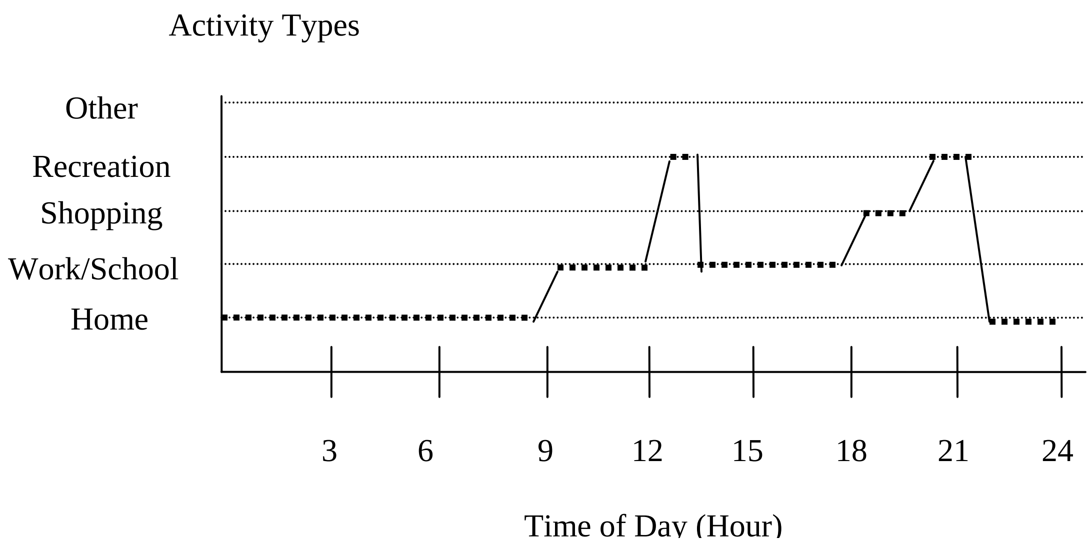


Figure 4.2 Sample Student Activity Profile

The advantage of the activity profile representation is that it facilitates understanding the daily activity sequence of each student group and helps embed the type of activities, their timing, and sequencing of activities into an activity-based model. Moreover, this representation does not require the estimation of many as parametric models as do econometric approaches [10, 26, 28]. The activity profile representation will provide activity-based travel demand models with less complexity in model development.

4.4.1 Development of activity profiles

The activities of NCSU students are stored in the database called 'Place' which includes sample number, place type, departure time, arrival time, type of major activity, etc. as shown in Figure 4.3. This activity information is joined with the 'Personal survey' database (Appendix A2) based on sample number (ID) to sort activities into the four student traveler groups previously defined.

	B	C	D	E	F	G	H	I	J	Y	Z	AA	AB
1	col2	col3	col4	col5	col6	col7	col8	col9	col10	col25	col26	col27	col28
2	SAMPNO	PLACENO	PLATYP	KINDPLA	PCODE	PNAME	ARRIVE	ARRAMPM	MAINACT	DEPART	PGEOMET	PGEOCIT	PGEOZIP
3	10009	1	1	HOME	0	E.S. King Village E -Onslow Hall			10	8:00 AM	E.S. King Village E -Onslow Hall	Raleigh	27607
4	10009	2	4	DAYCARE	9	CAMPUS CHILD CARE CENTER	8:10 AM	1	4	8:30 AM	900 TRAILWOOD DR	RALEIGH	27606
5	10009	3	1	HOME	0	E.S. King Village E -Onslow Hall	8:40 AM	1	2	9:30 AM	E.S. King Village E -Onslow Hall	Raleigh	27607
6	10009	4	2	WORK	9	UNIVERSITY	10:00 AM	1	3	12:10 PM	WILLIAMS HALL	RALEIGH	27607
7	10009	5	1	HOME	0	E.S. King Village E -Onslow Hall	12:30 PM	2	5	4:00 PM	E.S. King Village E -Onslow Hall	Raleigh	27607
8	10009	6	2	WORK	9	UNIVERSITY	4:20 PM	2	3	5:30 PM	WILLIAMS HALL	RALEIGH	27607
9	10009	7	3	NCSU Campus	10	CHILD CARE CENT	5:45 PM	2	4	6:00 PM	CHILD CARE CENT	RALEIGH	
10	10009	8	1	HOME	0	E.S. King Village E -Onslow Hall	6:50 PM	2	9		E.S. King Village E -Onslow Hall	Raleigh	27607
11	10011	1	1	HOME	0	E S King Village			10	8:40 AM	E S King Village	Raleigh	27607
12	10011	2	3	NCSU Campus	10	HARRELSON HALL	8:55 AM	1	3	9:15 AM	HARRELSON HALL	RALEIGH	27607
13	10011	3	3	NCSU Campus	10	DANIELS HALL	9:19 AM	1	4	9:22 AM	DANIELS HALL	RALEIGH	27607
14	10011	4	3	NCSU Campus	10	HARRELSON HALL	9:27 AM	1	2	9:50 AM	HARRELSON HALL	RALEIGH	27607
15	10011	5	2	WORK	9	university	9:53 AM	1	3	12:50 PM	2578 HILLSBOROUGH STREET	RALEIGH	27607
16	10011	6	3	NCSU Campus	10	D H HILL LIBRARY	12:55 PM	1	2	1:05 PM	D H HILL LIBRARY	RALEIGH	27607
17	10011	7	3	NCSU Campus	10	HARRELSON HALL	1:06 PM	2	3	1:30 PM	HARRELSON HALL	RALEIGH	27607
18	10011	8	2	WORK	9	university	1:35 PM	2	3	5:25 PM	2578 HILLSBROUGH STREET	RALEIGH	27607
19	10011	9	1	HOME	0	E S King Village	5:40 PM	2	10		E S King Village	Raleigh	27607
20	10013	1	1	HOME	0	E.S. King Village F -Northhampton Hall			10	9:00 AM	E.S. King Village F -Northhampton Hall	Raleigh	27607
21	10013	2	3	NCSU Campus	10	VENTURE BUIILDING	9:09 AM	1	1	10:30 AM	VENTURE BUIILDING	RALEIGH	
22	10013	3	1	HOME	0	E.S. King Village F -Northhampton Hall	10:40 AM	1	5	1:50 PM	E.S. King Village F -Northhampton Hall	Raleigh	27607
23	10013	4	3	NCSU Campus	10	DANIELS HALL	2:15 PM	2	1	3:20 PM	DANIELS HALL	RALEIGH	27607
24	10013	5	1	HOME	0	E.S. King Village F -Northhampton Hall	3:30 PM	2	2	7:00 PM	E.S. King Village F -Northhampton Hall	Raleigh	27607
25	10013	6	3	NCSU Campus	10	D H HILL LIBRARY	7:10 PM	2	2	11:00 PM	D H HILL LIBRARY	RALEIGH	27607
26	10013	7	1	HOME	0	E.S. King Village F -Northhampton Hall	11:10 PM	2	10		E.S. King Village F -Northhampton Hall	Raleigh	27607
27	10019	1	1	HOME	0	Becton Residence Hall			10	9:53 AM	Becton Residence Hall	Raleigh	27607
28	10019	2	3	NCSU Campus	10	DABNEY HALL	10:04 AM	1	1	10:45 AM	DABNEY HALL	RALEIGH	27607
29	10019	3	3	NCSU Campus	10	POLK HALL	10:50 AM	1	1	12:05 PM	POLK HALL	RALEIGH	27607
30	10019	4	3	NCSU Campus	10	TALLEY STUDENT CENTER	12:10 PM	2	5	12:24 PM	TALLEY STUDENT CENTER	RALEIGH	27607
31	10019	5	1	HOME	0	Becton Residence Hall	12:32 PM	2	2	6:20 PM	Becton Residence Hall	Raleigh	27607
32	10019	6	3	NCSU Campus	10	NCSU BOOKSTORES	6:32 PM	2	6	6:35 PM	NCSU BOOKSTORES	RALEIGH	27607
33	10019	7	3	NCSU Campus	10	FOUNTAIN DINING HALL	6:45 PM	2	5	7:10 PM	FOUNTAIN DINING HALL	RALEIGH	27607
34	10019	8	3	NCSU Campus	10	POLK HALL	7:26 PM	2	6	8:25 PM	POLK HALL	RALEIGH	27607
35	10019	9	1	HOME	0	Becton Residence Hall	8:32 PM	2	4	9:00 PM	Becton Residence Hall	Raleigh	27607
36	10019	10	4	FOOD - RESTAURA	7	ICE CREAM SHOP	9:02 PM	2	5	9:12 PM	2018 CAMERON STREET	RALEIGH	27605
37	10019	11	1	HOME	0	Becton Residence Hall	9:17 PM	2	10		Becton Residence Hall	Raleigh	27607
38	10034	1	1	HOME	0	Lee Residence Hall			10	8:55 AM	Lee Residence Hall	Raleigh	27607
39	10034	2	3	NCSU Campus	10	HARRELSON HALL	9:10 AM	1	1	10:00 AM	HARRELSON HALL	RALEIGH	27607
40	10034	3	3	NCSU Campus	10	CARMICHAEL GYM	10:08 AM	1	1	11:00 AM	CARMICHAEL GYM	RALEIGH	27607
41	10034	4	1	HOME	0	Lee Residence Hall	11:10 AM	1	9	12:30 PM	Lee Residence Hall	Raleigh	27607
42	10034	5	3	NCSU Campus	10	FOUNTAIN DINING HALL	12:35 PM	2	5	1:00 PM	FOUNTAIN DINING HALL	RALEIGH	27607
43	10034	6	1	HOME	0	Lee Residence Hall	1:05 PM	2	9	5:45 PM	Lee Residence Hall	Raleigh	27607
44	10034	7	3	NCSU Campus	10	FOUNTAIN DINING HALL	5:50 PM	2	5	6:30 PM	FOUNTAIN DINING HALL	RALEIGH	27607

Figure 4.3 Sample Database of 'Place survey'

In order to construct the activity profile by student groups in hourly time segments, the activity duration of each activity needs to be calculated with the associated trip. The activity duration is used to decide what activity in each hour is most dominant by each student. Then the chosen activity type creates a daily activity profile by aggregating all activities of students within a same group during twenty-four hours. The aggregation is required to estimate hourly travel demand. The algorithm for creating the activity profile follows.

Algorithm for Activity Profile

Step 1*: Develop a dataset including individual ID and sequence of activities by time, Then Calculate activity duration of each activity type*

$$AD_j = time_d^j - time_a^j$$

Where,

$time_d^j$: Departure time at activity j

$time_a^j$: Arrival time at activity j

AD_j : Duration of activity j

Step 2*: Identify the major activity in each hour frame based on activity duration of each activity (when the person participated in more than on activity, the major activity with the greatest duration is chosen. In case there are two activities with the same greatest duration, the one that is different from the previous hour is chosen)*

Activity	***Home***	***Work/School***	***Shopping***
Duration	***15 (min)***	***35(min)***	***10(min)***
Chosen		***X***	

60 minutes

***Step3**: Derive probabilities of students participating in particular activities (the activity profile) for 24-hour activity sequences*

Notation;

AN: Number of activities

ID(i): ID of person i

A(t): activity type at time t

h: the hour of activity observed

Pseudo-code (Visual Basic Application):

```
For i = 1: AN
     Initialization;
           ID(i) = 0, t=1

     Data read;
           ID, A, h
           ID(i) = ID,  A(t)=A;

           Do While (ID(i) = ID)
                       If (t<h)
                             Do While (t<=h)
                                   A(t) = A
                                   t = t+1
                             Loop
                       Else if (t=h)
                                   A(t) = A
                                   t = t+1

                       Else
                             Do Until (t=24)
                                   A(t) = A
                                   t = t+1
                             Loop
                       End If
                 Data read; ID, A, h
           Loop
End
```

This procedure results in short activities being overlooked, but the shorter time segments will increase model complexity. This deficiency was unavoidable. However, all short activities will be considered in the travel demand model in Chapter 5 to estimate an appropriate number of trips associated with the activities.

From the algorithm of the activity profile, the output is obtained as shown in Figure 4.4. The program records a major activity in every hour for a twenty four-hour time frame. The total number of records is 843, the total students who responded to the survey. The derived probabilities (proportions) of students participating in particular activities are calculated by the number of each activity in each hour divided by the total number of students. Therefore, the sum of the probabilities of the activities is equal to 1 at every hour.

ID	**1AM**	**2AM**	…………..	**10AM**	**11AM**	**12PM**	…………..	**11PM**	**12AM**
1	H	H	…………..	W	W	R	…………..	H	H
2	W	H	…………..	S	S	W	…………..	R	H
3	H	H	…………..	W	W	W	…………..	O	H
4	H	H	…………..	H	W	R	…………..	S	H
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
843	H	H	…………..	H	H	R	…………..	H	H

Activity	**1AM**	**2AM**		**10AM**	**11AM**	**12PM**		**11PM**	**12AM**
H	0.962	0.901	…………..	0.372	0.335	0.325	…………..	0.911	0.969
W/S	0.015	0.071	…………..	0.533	0.519	0.480	…………..	0.034	0.011
S	0.001	0.002	…………..	0.013	0.013	0.023	…………..	0.004	0.001
R	0.008	0.007	…………..	0.055	0.105	0.137	…………..	0.039	0.012
O	0.013	0.018	…………..	0.027	0.029	0.036	…………..	0.012	0.007
SUM	1.000	1.000	…………..	1.000	1.000	1.000	…………..	1.000	1.000

Figure 4.4 Sample Output of Activity Profile

4.4.2 Student's daily activity patterns

Figures 4.5, 4.6, and 4.7 help to illustrate daily activity-travel patterns of students and the sequence of activities (activity profile) by student groups (undergraduate and graduate, male and female, on-campus and off-campus students.)

In this study, for simplicity of analysis, each activity type is recorded every hour to create a daily activity profile. In the case of several activities occurring in an hour, one major activity is identified based on the activity duration (end time minus start time). In addition, all activities are aggregated to only five major activities: Home (all in-home activities); Work (class and research related activities); Shopping; Recreation (dining out and leisure activities); and Other activities.

As shown in Figures 4.4, 4.5, and 4.6 the daily activity profiles are developed and compared based on each pair of student groups: undergraduate/graduate, male/female, and on-campus/off-campus students, respectively. Comparisons show that daily activity patterns have an a.m. and a p.m. peak in work activity like that of typical urban employees. However, the peak hours appear later in the morning around 10 a.m. (instead of 8 a.m.) and earlier in the afternoon about 2 p.m. (instead of 5 p.m.) due to class schedules.

In Figure 4.5 a comparison of daily activity patterns between undergraduate and graduate students indicates that more undergraduate students are involved in Work in both morning and afternoon hours than graduate students. Graduate students spend more time for Work including Research in the afternoon than in the morning. Further, some graduate students undertake Work/research late at night causing fewer students to appear in the Work activity in the morning hours (e.g., graduate students undertake more home activity from 9 a.m. to noon). The Social-recreation activity is commonly observed at noon in both student groups since Meals is a major activity in the Recreation activity category. However, Recreation shows only one distinctive peak at noon and moderately decreases in the afternoon. In contrast to Recreation at noon, Recreation in the afternoon consists of other social-recreational activities such as going to a gym, meeting friends, having party according to comments in survey responses. A small number of students in both groups undertake Shopping and Other activities during a day although graduate students are more involved in these activities. This may be explained by graduate students undertaking these activities for their family members.

Figure 4.6 comparing daily activity patterns between male and female students shows that male students are more likely to engage in Recreation at noon, and they are slightly less involved in Shopping compared to female students.

Unlike other student groups, the daily activity profile between on-campus and off-campus students is distinctive. Figure 4.7 shows that students living on-campus are more likely to undertake Work and Recreation activities compared to students living off-campus. This corresponds to the result from a comparison between undergraduate and graduate students since most graduate students live off-campus. This also corresponds to the result that undergraduate students make more trips than graduate students (Figure 3.3).

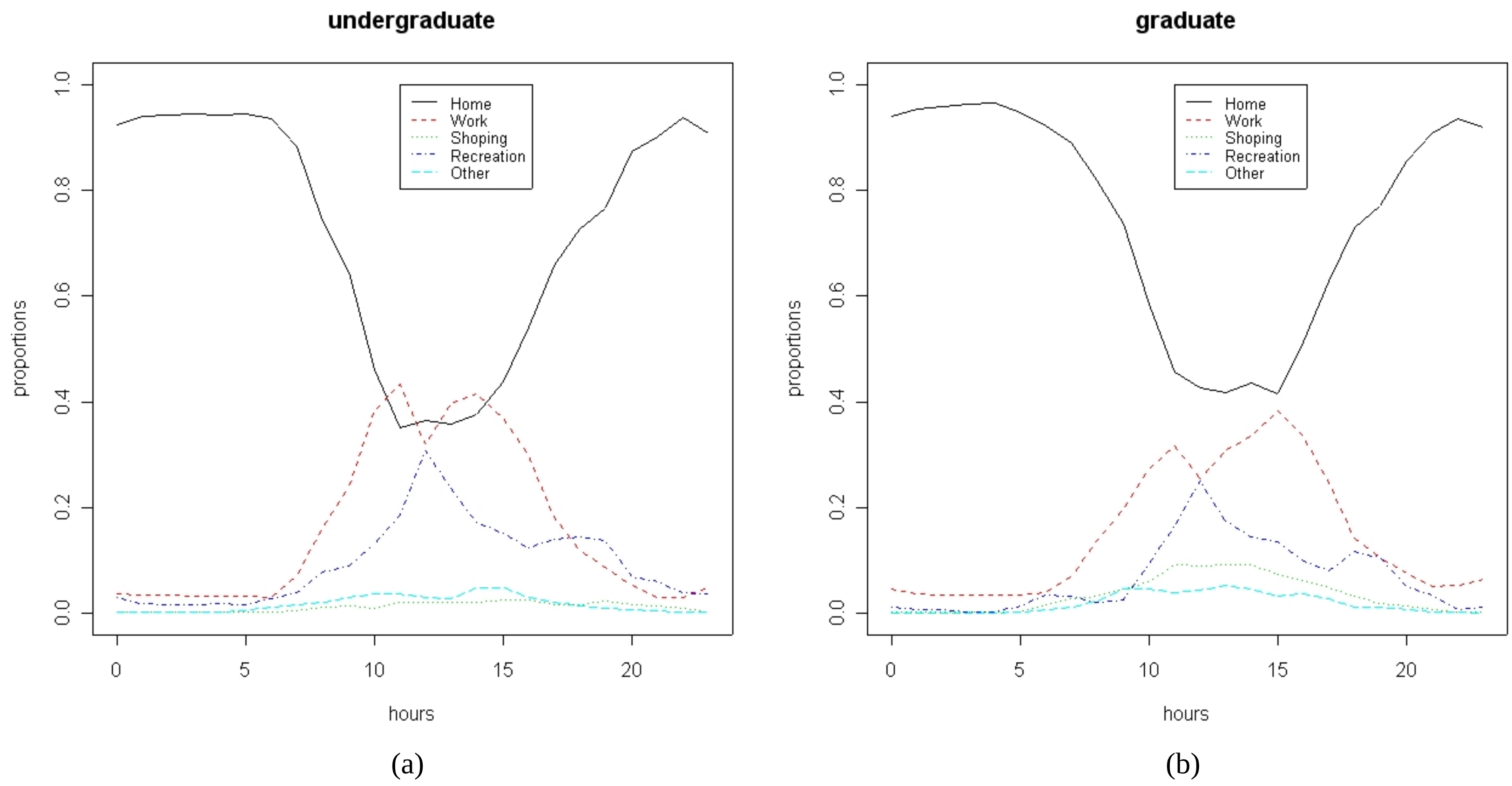
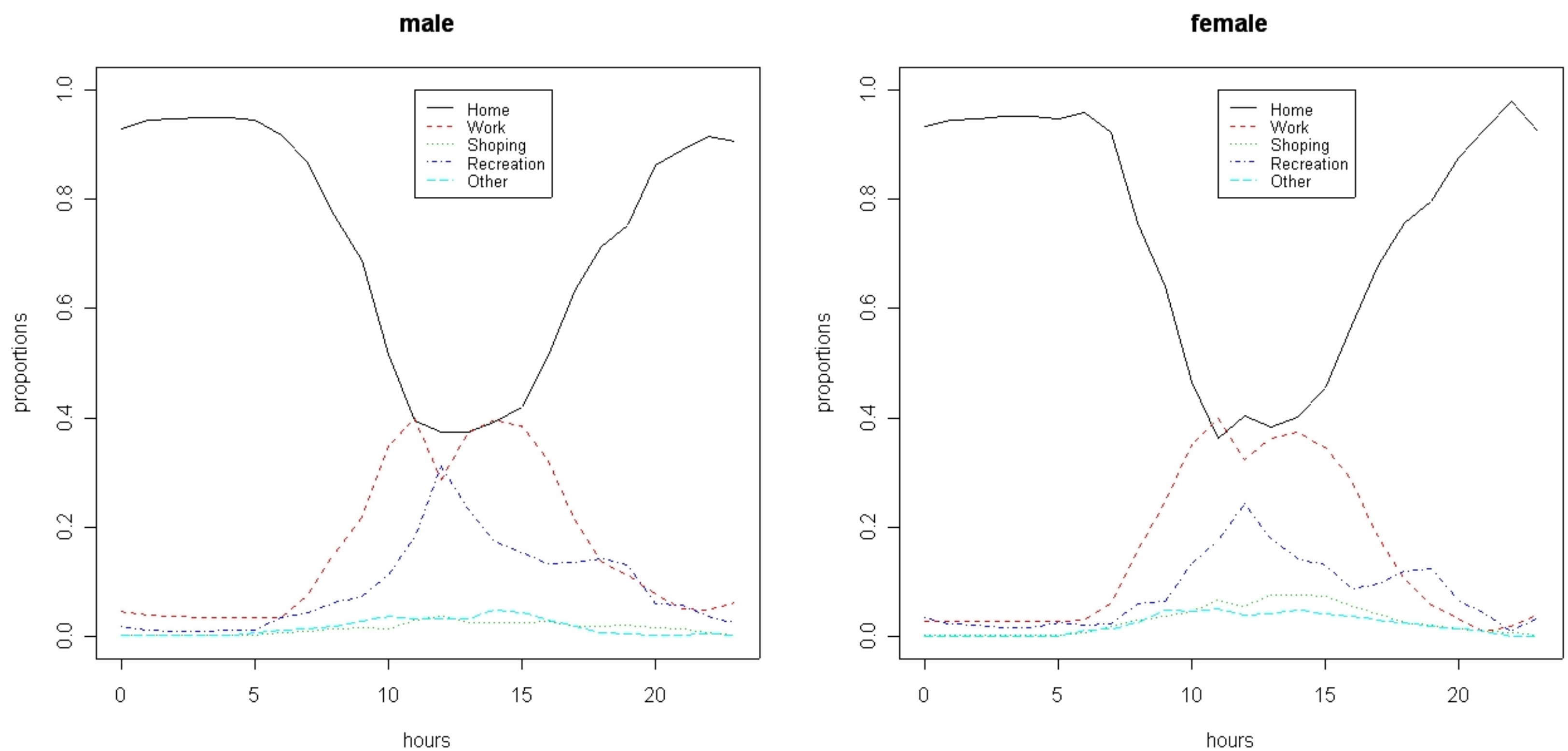
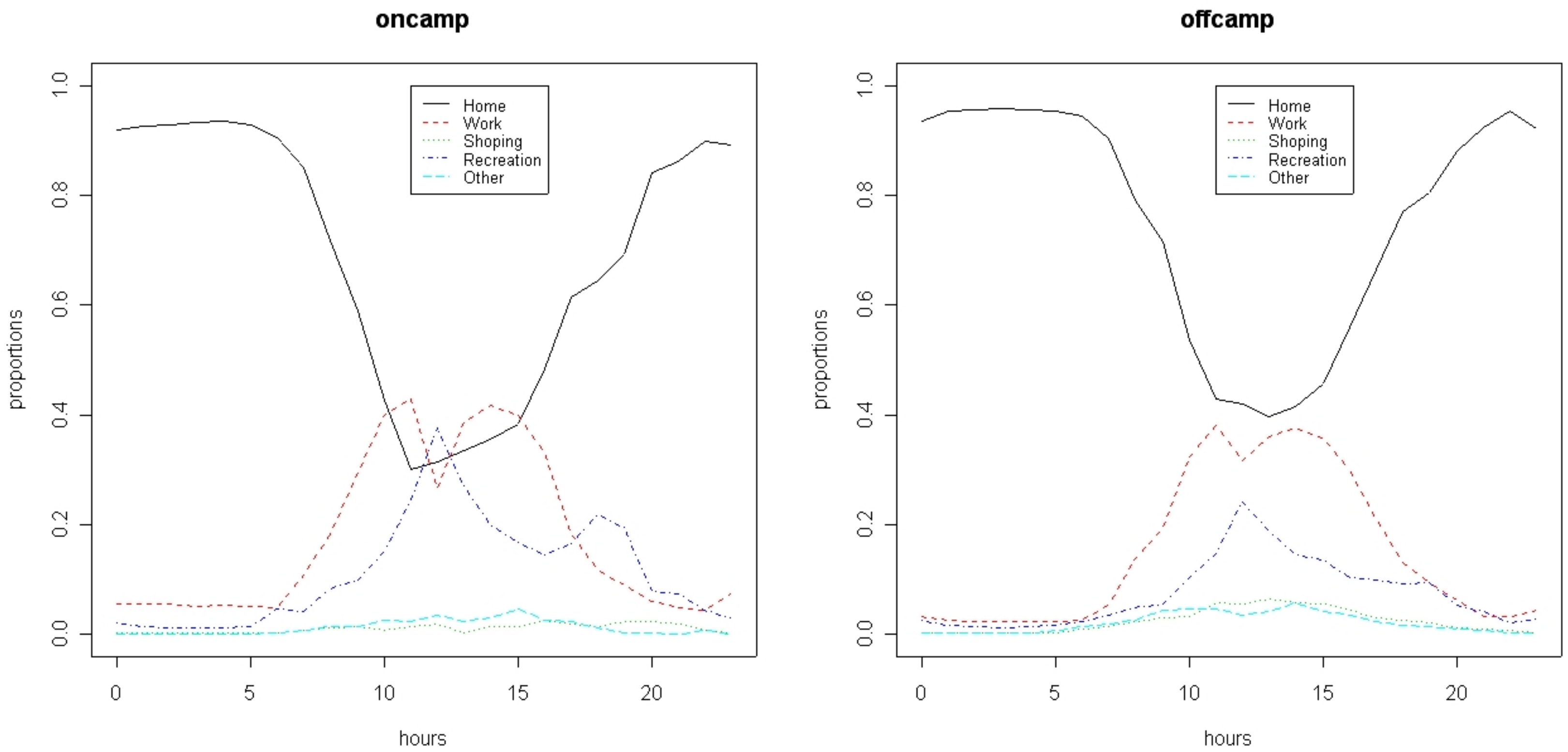


Figure 4.5 Activity Profile by Student Group: a) Undergraduate, b) Graduate



(a) (b)

Figure 4.6 Activity Profile by Student Group: a) Male, b) Female



(a) (b)

Figure 4.7 Activity Profile by Student Group: a) On-campus, b) Off-campus resident

4.4.3 Similarity test of activity profiles by students' group

In this section, the overall activity profile by student group is statistically examined to see whether students' activity profiles are similar to each other. From a visual comparison between the student groups shown in Figures 4.5 through 4.7, each pair of groups appears to be similar. However, the statistical analysis provides better information which could turn out to be useful in university travel demand modeling or as a sub-model within a regional travel demand model. In order to test the similarity of activity profiles between student groups, the functional representation of activity profile and test hypothesis is defined as follows.

$$\mathrm{H}_0 : \mathrm{P}_j^A(t) = \mathrm{P}_j^B(t),\ \ j = 1,\cdots,5\ \ and\ t = 1,2,3,\cdots,24$$

where,

$$\mathrm{P}_j^A(t) = proportion\ of\ \ group\ A\ people\ peforming\ \ j\text{ - }th\ activity\ at\ time\ t$$

$$N_j(t) = number\ \ of\ \ people\ \ \ in\ j\ activity\ \ at\ \ time\,(t)$$

$$N(t) = (N_1(t),\cdots,N_5(t)) \sim Multinomial(N_T(t),\mathrm{P}_1(t),\cdots,\mathrm{P}_5(t))$$

where,

$$N_T(t) = \sum_{j=1}^{5} N_j(t),$$

$$\hat{p}_j(t) = \frac{N_j(t)}{N_T(t)},\quad t = 1,2,\cdots,24$$

However, if this hypothesis is so strong that the test-statistic rejects it, then the alternative hypothesis which is not as strong is.

$$\mathrm{H}_0^* : \sum_{j=1}^{5} (\hat{p}_j^A(t) - \hat{p}_j^B(t))^2 = D(t),\ \ j = 1,\cdots,5$$

The first hypothesis (strong one), recursively accomplishes the test of analysis of variance (ANOVA) for each category (educational status, gender, and residence type). Table 4.4 shows the summary of the hypothesis test results between the students groups: undergraduate and graduate, male and female, on-and off-campus. The test results indicate that, overall, the student groups are not significantly different in terms of activity participation (proportion) and sequence of hourly activities. In other words, the daily trends of activity sequences between undergraduate and graduate student are not statistically different over twenty-four hours. However, the activity types are significantly different at all time frames irrespective of student groups. While grouping students by educational status may not be critical with respect to activity travel demand modeling, the activity type and its sequence are a critical component in university travel demand modeling. This trend is also observed in the comparison of other student groups.

Table 4.4 Summary of ANOVA for Activity Profile by Student Group

Source	DF	F-value	Prob(>F)
Educational Status (Undergraduate & Graduate)	1	0.31	0.5754
Gender	1	0.35	0.5555
Residence Type (On-campus & Off-campus)	1	0.76	0.3843
Hour	23	0.17	1.0000
Activity Type	4	6260	<0.0001***

Note: Significant: 0 `***'.

Table 4.4 also shows that there is no significant difference in activity profile between male and female students over twenty-four hours. The activity type, however, is found to be significant in all frames. Residential status does not seem to be a statistically significant factor with respect to the proportion of students involved in a certain activity at any time of day. In summary, students' hourly activity participation does not seem to be different although the proportions are slightly different among student groups. In contrast to student groups, hourly activity type is different which means that the activity type is highly correlated to time of day.

4.5 Transferability of University Student Trip Rates and Activity Data

The transferability of university student activity data is based on the data comparison between Pennsylvania State University (PSU) and North Carolina State University. The two university surveys had different scopes, purposes and questionnaires. However, this analysis will compare the two critical activity-based model components trip rates and activity profiles to test for transferability.

In order to do this, first the PSU student database of is briefly introduced, and students' trip rates between the two universities are compared for similarity. For the comparison of trip rates, two approaches are used; 1) negative binomial model approach, 2) test of spatial structural change approach. The first approach treats the trips with categorical explanatory variables while the second approach is a multiple regression model used in a Chow-test which provides statistics that show whether two databases observed in two different regions are similar or not. For the comparison of activity profiles, the summary of analysis of variance (ANOVA) is presented.

4.5.1 Pennsylvania State University data

The PSU student trip data used for this study was part of a database obtained from the CentreSIM survey, conducted 2002 [25]. This dataset is the product of a five-year initiative to improve modeling and simulation for transportation decisions in Centre County, Pennsylvania. CentreSIM is a regional simulation of Centre County, a region of approximately 136,000 persons that includes Pennsylvania State University with about 40,000 students and more than 12,000 faculty and staff. The characteristics of the Penn State population in terms of time and space choices need to be accounted for in the simulation of the county population. To accomplish this, a household survey was designed and conducted by PSU researchers. The research also included a complete inventory of businesses: address, business type classified by standard industry classification (SIC) code, and number of employees of each business.

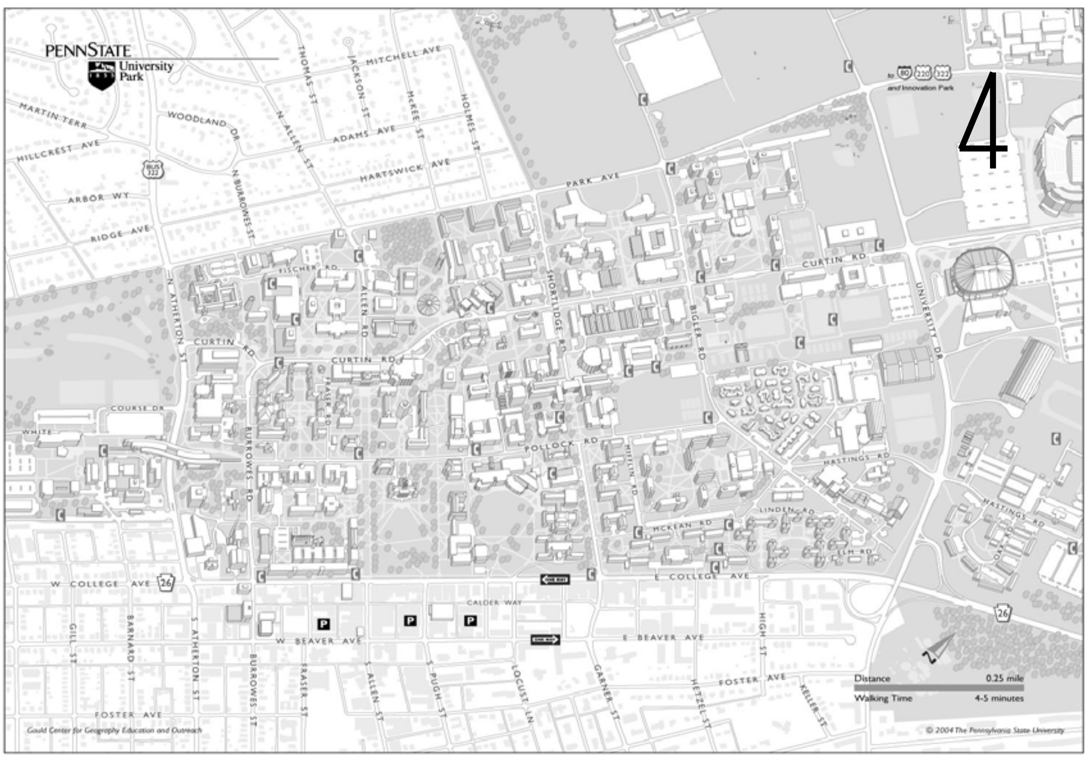


Figure 4.8 Map of Pennsylvania State University
(Source: www.psu.edu)

The survey was conducted by contacting households by telephone and by mail. Activity-travel diaries were mailed to each household as they returned completed household questionnaires. There are two major parts of the resulting data in the survey; a data table from the household survey questionnaire organized for each household member and a data table for each individual representing a two-day activity/travel diary organized activity by activity with associated trips reported. For the questionnaire each participating household provided voluntary information about household composition and house facilities available to the household members. Also, each household member provided personal information such as employment, driving ability, education, etc. The activity and travel data collected from each member in the household using the two-day complete record of the activities included the types of activities and the travel modes taken to them. The respondents were asked to record beginning and ending time of each activity, and reported starting and ending location of the trip. These questions captured a variety of contexts in the decision-making aspects of each household member in activity and travel scheduling.

Unlike the NCSU student activity survey, the PSU survey was not for students only. Thus, the PSU students were not clearly classified by students' characteristics such as student status (undergraduate and graduate), residence status (on-campus and off-campus). In order to compare the students' activity data between the two universities, the PSU student data were extracted from the original dataset based on the assumptions that the students whose main travel mode is 'Walk/Bike' are on-campus residents and others are off-campus students. Student status was based on the response for current educational level. The students who already had a bachelor degree or master degree were assumed graduate students and others undergraduate students. Although this assumption seems a little broad (post baccalaureate and adult continuing education students are classed as graduate students), it will identify the NCSU-like student groups required for the analysis.

Table 4.5 shows the summary of datasets between two universities. Students' daily trip rates for both universities are higher than normal households and the trip rates of PSU are slightly higher than those of NCSU.

Table 4.5 Summary of Student Dataset between Two Universities

Variable	Description	Coding	Statistics	
			NCSU	PSU
Trips	Number of trips	Continuous	6.35	6.72
Gender*	Gender of student	1: Male	492 (58.4%)	56(45.9%)
		0: Female	351 (41.6%)	66(54.1%)
Student status*	Educational status	1: Undergraduate	645 (76.5%)	80(65.6%)
		0: Graduate	198 (23.5%)	42(34.4%)
Residence*	Living location	1: On-campus	429 (50.9%)	55(45.1%)
		0: Off-campus	414 (49.1%)	67(54.9%)

Note: The variables with an asterisk are categorical.
Statistics of categorical variables are denoted as frequency (percentage).
Total PSU students responded are 121 for two days. The trips are coded with the average of two days.

The descriptive summary statistics of PSU show that more female (54.1%), undergraduate students (65.6%), and off-campus students (54.9%) responded compared to male, graduate students, and off-campus students. Although the sample size of PSU is relatively small compared to NCSU, the comparison will be important in order to determine transferability of student trip-making behavior and activities between the two universities.

4.5.2 Similarity test of trip rates between two universities

a. Negative binomial model approach

As noted before, the test of similarity of trip rates between the two universities is first undertaken by applying a negative binomial model. The PSU students' trip rates with respect to the students' characteristics (as considered in the analysis of NCSU students' trip-making behavior) show that overall trip-making behavior is similar to that of NCSU students.

Table 4.6 shows the results of the analysis of the negative binomial model applied to the PSU student groups: student status, gender, and residence status. The results indicate that the off-campus students are less likely to make a trip compared to on-campus students and undergraduate students are more likely to travel than graduate students. Also, female students are more likely to make trip compared to male students. It is very interesting to note that these results correspond to NCSU students' trip-making behavior (Table 3.3).

Table 4.6 PSU Student Group Effect on Daily Trip Rate

	Number of Trips in a Day		
Variable	Coefficient	χ^2	Pr> χ^2
Residence Status			
Off-campus	-0.0846	1.08	0.2985
On-campus+	--	--	--
Gender			
Female	0.1147	2.01	0.1564
Male+	--	--	--
Student Status			
Graduate	-0.0552	0.41	0.5235
Undergraduate+	--	--	--
Model Goodness of Fit			
N	121		
Pearson- χ^2	121.08 (df: 117)		
Log-likelihood	575.72		

Note: (+) indicates that this variable is a base variable in the group.

However, all student groups are not statistically significantly different from each other in trip-making behavior. In other words, all students have statistically similar trip-making behavior at PSU. This is somewhat different from the result for NCSU students (refer to Table 3.3) A possible reason could be that the format of the PSU questionnaire and database which did not clarify student status and residence status.

Table 4.7 shows the results of the negative binomial model applied to combined data including PSU and NCSU data together. From the results, it is found that the students' trip-making behavior is not significantly different between two universities. This illustrates that one university's student trip rates could be transferable to the other university if there are no travel data available.

Table 4.7 The Effect of Two Universities on Daily Trip Rate (Combined Data)

	Number of Trips in a Day		
Variable	Coefficient	χ^2	Pr> χ^2
<u>Residence Status</u>			
Off-campus	-0.1494	22.92	0.0001***
On-campus+		--	--
<u>Gender</u>			
Female	0.0361	1.33	0.2480
Male+	--	--	--
<u>Student Status</u>			
Graduate	-0.1325	16.35	0.0001***
Undergraduate+	--	--	--
<u>University</u>			
PSU	0.0814	3.16	**0.0756**
NCSU+	--	--	--
<u>Model Goodness of Fit</u>			
N	943		
Pearson- χ^2	937.76 (df: 938)		
Log-likelihood	5053.37		

Note: (+) indicates that this variable is a base variable in the group.
Significant: 0 `***'.

b. Regression model approach (Spatial structural stability test)

The Chow-test is a well-known approach in economics to determine the spatial structural change of model parameters between two models estimated for two regions separately located. Chow [34] provided a methodology for testing structural stability to determine whether the coefficients in a regression model are the same in separate sub-samples. The strength of the Chow test is that the spatial region difference can be assessed by means of testing the stability of regression coefficients. This study uses the Chow-test to assess the similarity of travel data between two universities.

Chow-test

The Chow-test is an application of the Wald-type *F*-test, and it requires the sum of squared errors from three regressions; one for each sample of spatial domain and one for pooled data. For example, the following two linear regression models are assumed to be different across two different datasets:

$$y_i = \alpha_0 + \alpha_1 X_{1i} + \cdots + \alpha_k X_{ki} + \varepsilon_i, \quad i = 1, 2, \ldots, n_1, \text{ for Region 1} \qquad (1)$$

$$y_j = \beta_0 + \beta_1 X_{1j} + \cdots + \beta_k X_{kj} + \varepsilon_j, \quad j = 1, 2, \ldots, n_2, \text{ for Region 2} \qquad (2)$$

where y_i, y_j are the observed dependent variables with n_1 and n_2 observations, $X_{ki,} X_{kj} \in \Re^p$, $k = 1,2,\ldots, p$, and $\alpha = (\alpha_0, \alpha_1, ..., \alpha_p)$, $\beta = (\beta_0, \beta_1, ..., \beta_p)$ are the corresponding parameter coefficients for each regression model. Note that it is important that the Chow-test requires ε_i and ε_j are normally distributed with mean 0 and variance σ^2, because Chow-test results are valid under the assumption of homoskedastic errors. If the disturbance variances are different, statistical inferences based on this F distribution will be misleading. The ordinary least squares (OLS) estimates of α and β from each unrestricted regression model can be calculated by $\alpha = (X_i^T X_i)^{-1} X_i^T y_i$ and $\beta = (X_j^T X_j)^{-1} X_j^T y_j$, which are unbiased estimators for α and β, respectively. The residuals ε_i and ε_j are also obtained from equation (1) and (2) which are known as the unrestricted models. If the observation from regions 1 and 2 are combined, the regression model becomes:

$$y_c = \gamma_0 + \gamma_1 X_{1c} + \cdots + \gamma_k X_{kc} + \varepsilon_c, \quad c = 1, 2, \ldots, n, \text{ for all regions} \tag{3}$$

where y_c represents the pooled observations, and n is total number of observations ($n_1 + n_2$). The residual ε_c from equation (3) is known as the restricted model because it restricts the parameters to equality across the two regions. The hypothesis to test is:

$$H_0 : \alpha = \beta, \tag{4}$$

where $\alpha = (\alpha_0, \ldots, \alpha_p)^T$ and $\beta = (\beta_0, \ldots, \beta_p)^T$. If the null hypothesis is true, then the restricted model will do just as well in explaining all observations as the unrestricted model. It is sensible to investigate whether one regression model applies to both regions and whether two models are needed. The test is a standard *F*-test that compares the residual sum of squares between the restricted and unrestricted models. The Chow-test compares the residual sum of squares of two linear regression models. The Chow-test statistic is:

$$F = \frac{\left[SSE_c - (SSE_1 + SSE_2)\right]/k}{(SSE_1 + SSE_2)/(n-2k)}, \tag{5}$$

where SSE_c is the sum of squared error of pooled samples for both regions, SSE_1 is the sum of squared error for region 1, SSE_2 is the sum of squared error for region 2, n is the total number of observations and k is the number of explanatory variables in the unrestricted model including the intercept. The test statistic is F distributed with k and n - 2k degrees of freedom. If the computed F-value exceeds the critical F, then the hypothesis that the two regressions are the same can be rejected. If the null hypothesis is true, then the Chow-test statistic may be small, and the null hypothesis should not be rejected as follows:

$F > F$ (k, n- 2k) Reject H_0

$F < F$ (k; n- 2k) Do not reject H_0

By using the Chow-test for parameter stability, the estimated structure of this “pooled" model can then be compared with the estimated structures of separate statistical models for each region. A significant difference indicates a structural stability in the relationship. The Chow-test can obviously be extended to more than two regions.

An alternative solution consists in using Bootstrap methods to yield approximations to the distribution of the Chow-test statistics considered. The Bootstrap, introduced by Efron [35], is a computer-intensive method for estimating the distribution of an estimator or test statistic by re-sampling the data at hand. In many instances the sampling distribution of a statistic may not be analytically available, while the Bootstrap, on the other hand, obtains the sampling distribution of the statistic via repeatedly re-sampling from the sample at hand. The Bootstrap method has been shown to be successful in many situations, which is accepted as an alternative to the asymptotic approach. Statistical inferences based on the Bootstrap applied to asymptotically pivotal statistics will generally be more accurate than inferences based on asymptotic theory [36, 37, 38]. By using the Bootstrap technique, it is possible to calculate either a critical value and a significance level, or P-value, associated with it.

Bootstrap Chow- test

The basic principle of the Bootstrap test is to draw a number of Bootstrap samples from the model under the null hypothesis and to calculate the Bootstrap test statistic F^*. The Bootstrap test statistic F^* can be calculated by repeating this step B times. As regards the number of Bootstrap samples used to estimate the Bootstrap critical value, a reasonable rule of thumb is that power loss will very rarely be a problem when B equals 999 [39].

For linear models of the form $Y = X\beta + \varepsilon$, Mammen [40] proved the asymptotic correctness of a Bootstrap test based on the residual Bootstrap. The Bootstrap samples are generated as $Y^* = X\hat{\beta}_0 + \varepsilon^*$, where $\hat{\beta}_0$ is the restricted least squares estimator under the null hypothesis and ε^* is a resample taken with replacement drawn from the empirical distribution function putting mass $1/n$ to the residuals $\hat{\varepsilon}_t, t = 1,...,n$. The residual Bootstrap method has been successfully used for structural stability testing [41, 42]. If the Bootstrap procedures are repeated, say B times, then a critical value may be obtained as the relevant percentage point, say, $F^*_{critical}$, from the empirical distribution of the Bootstrap test statistic and the null hypothesis is rejected if $F > F^*_{critical}$. Alternatively, a Bootstrap P-value of the test may be estimated as the fraction of times the value of the Bootstrap statistic exceeds F . For one-tailed test with a rejection region in the upper tail, the Bootstrap P-value is given by

$$P - value = \frac{1}{B}\sum_{i=1}^{B} I(F > F_i^*), \qquad (6)$$

where $I(F_i^* > F)$ is an indicated function that takes the value 1 if the argument is true or 0 otherwise.

Bootstrap Chow Test Algorithm

Step 1*: Estimate the model as though data were generated under the null hypothesis and obtain the sample values of the Chow statistic (F), $\hat{\beta}_0$, and the associated residuals* $\hat{\varepsilon}_t, t = 1,..., n$.

Step 2*: Generate B times the samples y^* as $Y_i^* = X\hat{\beta}_0 + \varepsilon_i^*, i = 1,..., B$, where ε_i^* is generated through simple random sampling with replacement from the empirical distribution putting mass $1/n$ and obtain the Bootstrap Chow-F statistic (F_i^*) using Bootstrap samples*

Step3*: Calculate the Bootstrap P-value as given by*

$$P - value = \frac{1}{B}\sum_{i=1}^{B} I(F > F_i^*), i = 1,2,..., B$$

***Step4**: The null hypothesis is rejected if the P-value is less than the significance level* α $(P - value < \alpha)$

The Bootstrap Chow-test algorithm is coded by using the statistical software package R that provides a convenient tool for matrix algebra and loop procedures [43].

Test results

Two regression models are constructed with the trip rate as the dependent variable and three independent variables that are dummy variables to represent qualitative variables such as gender, educational status, and residential area. Box-Cox transformation is applied to the trip rate to satisfy normality condition. Consider the following two unrestricted models for the two universities:

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \alpha_3 x_{3i} + \varepsilon_i, \quad i = 1, 2, \ldots, n_1, \text{ for PSU} \qquad (1)$$

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \beta_3 x_{3j} + \varepsilon_j, \quad j = 1, 2, \ldots, n_2, \text{ for NCSU} \qquad (2)$$

The hypothesis test shows whether the trip rate for NCSU is the same as for PSU, as shown in below.

$$H_0 : \alpha_0 = \beta_0,\ \alpha_1 = \beta_1,\ \alpha_2 = \beta_2, \text{ and } \alpha_3 = \beta_3$$

To investigate whether the trip rates differs between Model 1 (PSU) and Model 2 (NCSU), the Chow-test was employed. That is, Chow-test examined the equality of the pooled dataset of Model 1 and Model 2 regression coefficients to determine whether Model 1 and Model 2 have significantly different trip rates with respect to the independent variables.

As shown in Table 4.8, estimates are presented for two universities. The results of the OLS estimation are listed in the first column. The coefficients of the dummy variables are significant (P-value < 0.05) for the NCSU model, but the PSU model estimates are not significant. The estimates of the combined model are also significant.

Chow-test statistics of trip rates are computed between the two universities by running separate regressions, then pooling the data. In order to test whether there is a significant difference in the models for the two universities, the F statistic is calculated as follows.

$$F_{4,935} = \frac{(115.91 - 15.67 - 99.42)/4}{(15.67 + 99.42)/(943 - 8)} = 1.678 < 2.35$$

The critical value at the 0.05 level is about 2.35. Since the F statistic is smaller than the critical value, the null hypothesis of equality of parameter estimators cannot be rejected at the 5% level by the Chow-test. Therefore, the null hypothesis cannot be rejected such that the gender, on-off campus and under/grad values are statistically same for both universities. This is a test on the null hypothesis that the coefficients are the same in two universities, and the test is implemented for all coefficients jointly as well as for each coefficient separately.

The test for structural stability of regression coefficients of two universities does not reject the null hypothesis that coefficients are equal for the two universities, confirming that the trip-making pattern of two universities might be same. Moreover, the tests for coefficients of each university show that most explanatory variables do not exhibit significantly different effects. Coefficients for all three explanatory variables were found to be jointly significant at the 95% confidence level when applying the above test to the pooled data set.

Table 4.8 Parameter Estimates and Chow-test Statistics for Trip Rate Model

Model	DF	Parameter Estimates	Standard Error	T for H0: Parameter=0	Prob >|T|
Model: PSU					
Intercept	1	1.9166	0.0692	27.69	<.0001
Student Status	1	0.0383	0.0715	0.54	0.594*
Residential Status	1	0.0733	0.0683	1.07	0.285*
Gender	1	-0.0861	0.0675	-1.28	0.205*
Model: NCSU					
Intercept	1	1.8208	0.0268	68.01	<.0001
Student Status	1	0.0507	0.0264	1.92	0.0547*
Residential Status	1	0.0613	0.0262	2.34	0.0196
Gender	1	-0.0790	0.0258	-3.07	0.0022
Model: Combined					
Intercept	1	1.8372	0.0249	73.66	<.0001
Student Status	1	0.0482	0.0247	1.95	0.0516*
Residential Status	1	0.0694	0.0236	2.85	0.0045
Gender	1	-0.0883	0.0238	-3.71	0.0002

Summary of Chow-test and Bootstrap test

SSE_c (Comb)	115.91	Bootstrap Test	P-value = 0.32 (iteration: 999)
SSE_1(NCSU)	15.67		
SSE_2(PSU)	99.42		
df	4		
No. sample	943		
Chow-test statistic (OLS)	1.678		
F-value	F(4, 943, 0.05) =2.35		
Test: $H_0 : \alpha = \beta$	1.678<2.35 (Do not reject)		

Note: '*' is not significant at the .05 level
df = degree of freedom
Box-Cox transformation of 'Trip Rate' value is the power of 0.384.

Aside from the general Chow-test, the Bootstrap method also tests for structural stability between the two universities. Bootstrap generated 999 subsamples according to the algorithm previously given. The estimated structure of the restricted model can then, using the Bootstrap test for parameter stability, be compared with the estimated structures of separate statistical models for each unrestricted model (i.e., on separate regressions estimated for NCSU and PSU.) The empirical P-value of the test is obtained as the fraction of times the value of the Bootstrap statistics exceeds the values obtained from the original samples. It is important that Bootstrap samples, B, be sufficiently large because the power of a Bootstrap test depends on the number of bootstrap samples. The Bootstrap P-value with 999 iterations is obtained as 0.32 which cannot reject the null hypothesis. This means that the two universities are not statistically different from one another.

Consequently, both tests verify that there was a structural stability in the trip rate of travel demand model between NCSU and PSU and concludes that both university trip rate can be transferable from one another.

4.5.3 Similarity test of activity profiles between two universities

In section 4.4.3 the activity profiles developed for NCSU student groups are tested and the results indicated that the activity profiles among student groups are similar to one another. This section tests the similarity in activity profiles between the two universities by comparing the students' activity profiles from both universities and the NCSU students' profile against faculty and staff in PSU. This test will provide very useful information for developing an activity travel demand model for a university since the university travel demand model should consider the activities of all university populations. The activity profile for faculty and staff are developed from PSU data because NCSU survey did not collect activity data for this population. Hence, this would be an interesting comparison across the university with student data in NCSU and faculty and staff in PSU. The methods for similarity tests in activity profiles between students, faculty, and staff in both universities are the same as the one used in section 4.4.3.

Figure 4.9 shows students' activity profiles of both NCSU and PSU. At first glance, the overall patterns of activity profiles seem to be similar each other and the magnitude of each activity at each hour is slightly different. The proportions of ‘Work’ activities students do during a day in PSU are relatively higher than those in NCSU, while the proportions of ‘Home’ activity during a day in NCSU are slightly higher than those in PSU. This does not mean that NCSU students spend lesser time on ‘Work’ activity because any ‘Work’ related activities at home was classified as ‘Home’ activity. The other four activities seem to be similar and their magnitudes of difference in proportion are not significantly different.

Figure 4.10 shows the activity profiles of all university population consist of students, faculty, and staff. For cross comparison, the NCSU students’ activity profile is compared to PSU faculty and staff activity profiles. The overall activity patterns among traveler groups are comparable. However, it is of interest to note that faculty are more likely to be involved in ‘Work’ activity (above 70%) between 9am and 5pm compared to students (about 55%) and staff (about 60%). In contrast, a high increase of ‘Home’ activity for faculty is observed after 3pm while students’ ‘Home’ activity is gradually increased.

Figure 4.11 shows the difference of proportion of each activity in which people are involved during a day. The figure showing difference between students and faculty illustrates that students are more likely involved in various activities. Compared to students, faculty are highly involved in only two activities such as ‘Work’ activity during the working hours (9am to 5pm) and ‘Home’ activity after 5pm. Students are more likely to be engaged in all other activities specifically at evening hours (i.e. work, recreation, shopping). Similar activity patterns are also found in the difference between students and staff such that staff are more likely to be involved in ‘Work’ activity during a day and ‘Home’ activity in the evening. Faculty and staff have similar activity patterns but staff start ‘Work’ activities earlier than faculty do in the morning.

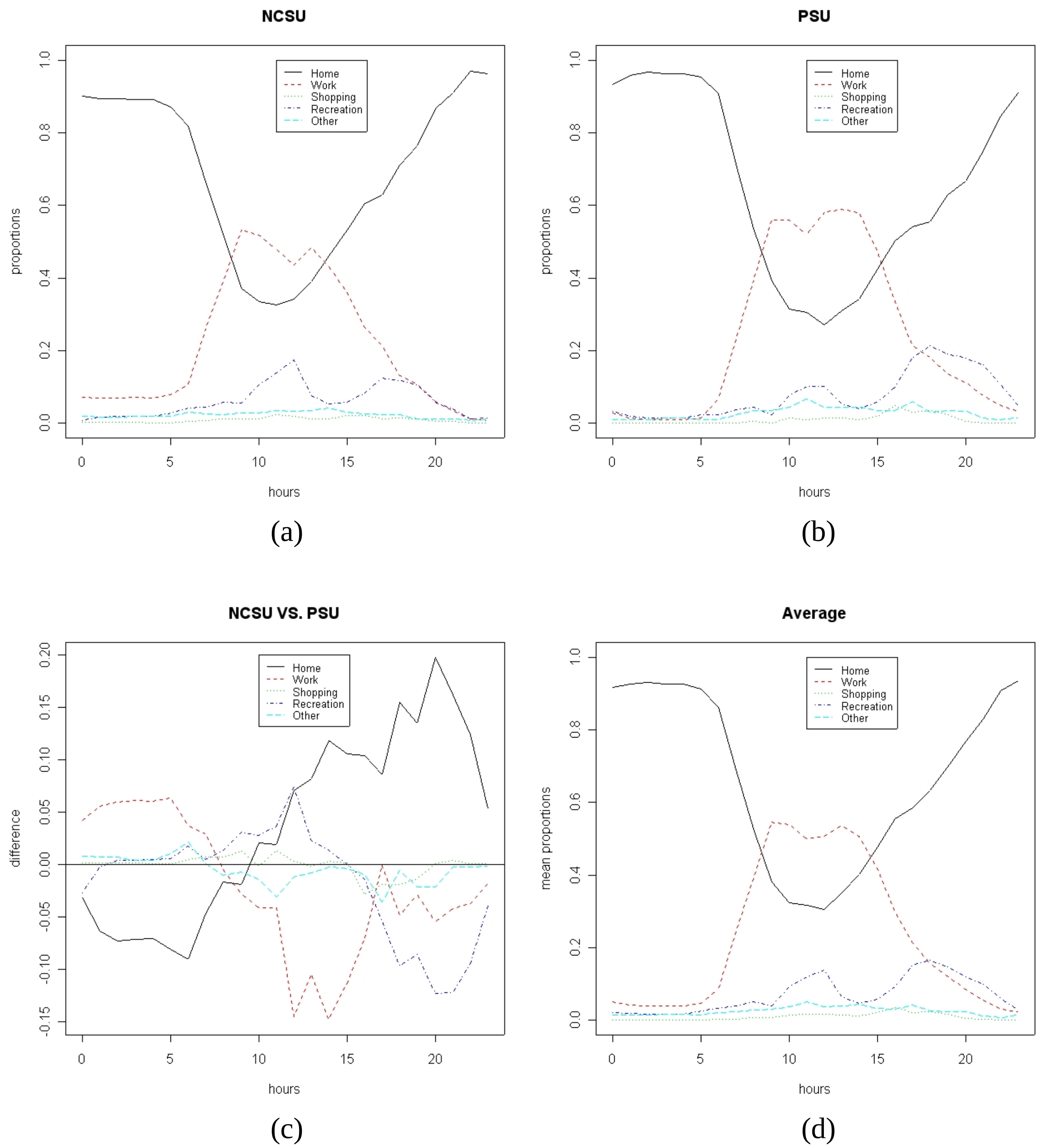


Figure 4.9 NCSU and PSU Students' Activity Profiles: a) NCSU, b) PSU, c) Difference between NCSU and PSU, d) Average of Two Universities

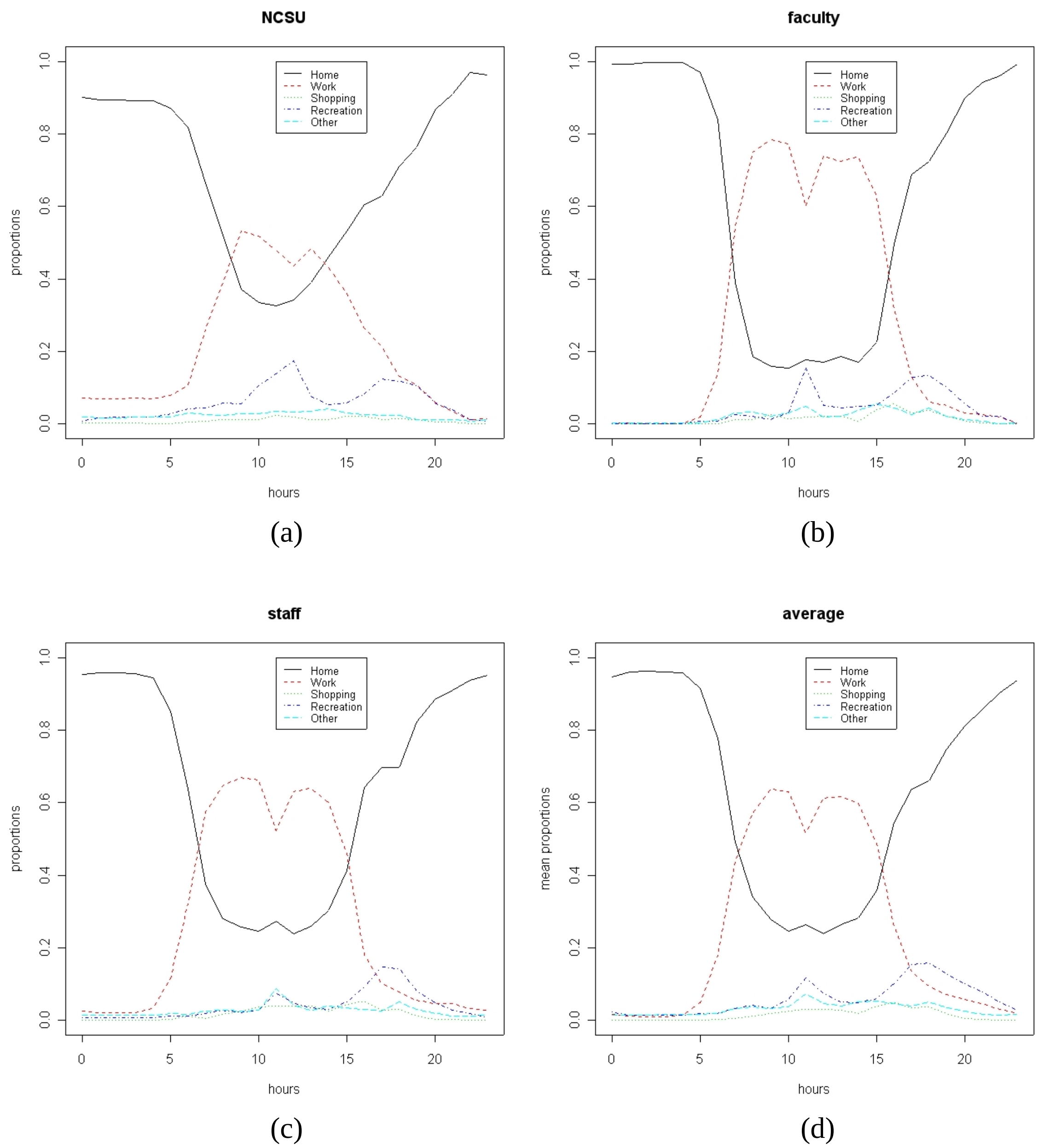


Figure 4.10 Activity Profiles of University Traveler Groups: a) NCSU Student, b) PSU Faculty, c) PSU Staff, d) Average of All Groups

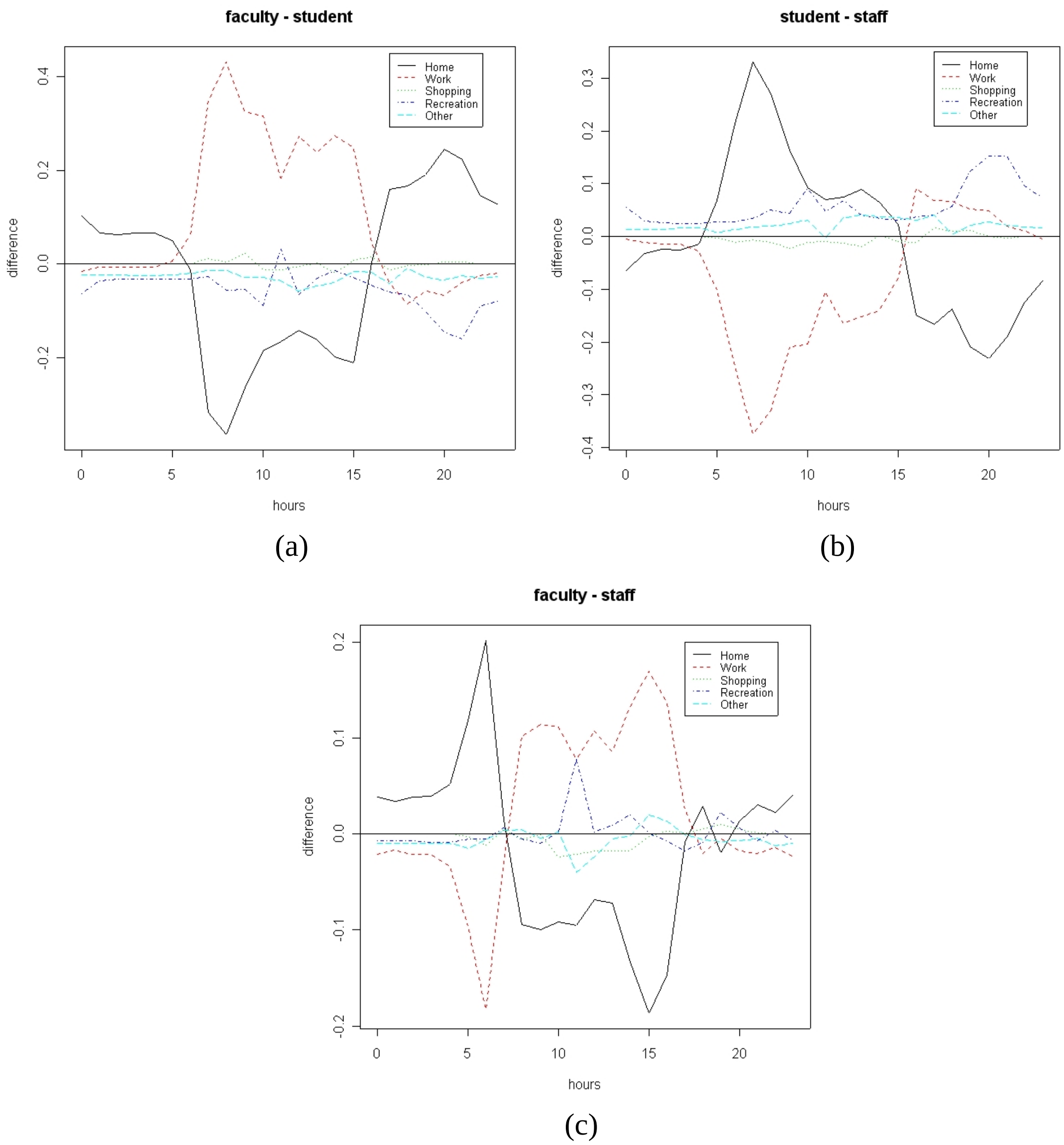


Figure 4.11 Differences of Activity Profiles between Traveler Groups: a) PSU Faculty vs. NCSU Student, b) NCSU Student vs. PSU Staff, c) PSU Faculty vs. PSU Staff

Faculty are more likely to be involved in ‘Work’ and ‘Recreation’ activity during a day but the difference of proportion in both activities is fairly small. The other activities are seen to be very small in the difference of proportion.

Table 4.9 summarizes the results of the statistical test of similarity in activity profiles between students at the two universities. In summary, the students’ activity profiles at PSU and NCSU are not significantly different. The students’ hourly activity participation does not seem to be different although the proportions are slightly different between activities among groups.

Table 4.9 Summary of ANOVA for Activity Profile by University Population

Source	DF	F-value	Prob(>F)
Group (NCSU, PSU)	1	0.01	0.9386
Hour	23	0.00	1.0000
Activity Type	4	142.88	<0.0001***

Note: Significant: 0 `***'.

4.5.4 Summary

With respect to the overall results of the statistical tests, it is of interest to note that the students’ activity data can be transferable from one university to another. The students’ daily trip rates do not seem to be different and this was verified by the statistical test of both negative binomial model and ordinary regression model approaches. As expected, student households are more likely to make a trip compared to regular households since students at universities are involved in a variety of activities on campus. It should come as no surprise that the results from the negative binomial model and ordinary regression model approach comparing two university students’ trip rates indicate exactly the same result. The statistics cannot reject the null hypothesis that the daily student trip-making behavior in both universities is not statistically different.

In the comparison of students' activity profiles between two universities, the overall activity pattern during a day looks similar and the test statistics verify that the activity profiles are not statistically different from one another. The only difference in the students' activity profile is 'Home' and 'Work' activity in which NCSU students are more likely to be involved in 'Home' activity and less likely to be engaged in 'Work' activity compared to PSU students. This does not mean that NCSU students spend lesser time on 'Work' activity because any 'Work' related activities at home were classified as 'Home' activity. The profile shows that the 'Home' activity may be opposite of 'Work' activity in terms of the activity participation. The statistical test also shows that the students between the two universities have similar activity patterns during the day although the proportion of each activity in every hour is different from one another.

From a statistical point of view, there is not a problem in transferring PSU activity data to NCSU. However, the proportion of activity participation is critical in activity travel demand model development and the proportion of each activity in every hour will decide how many people engaged in what activities. This turns out to be important later where the number of people calculated by using this proportion of activity participation is spatially allocated on the study area. Consequently, if activity survey data are available and the sample size is high enough, it would be better to develop the activity profile separately for each traveler group.

4.6 Modeling Activity Profiles

As shown in the ANOVA table, it appears that there is no significant group effect (i.e., students, faculty and staff behave almost the same). However there is a significant activity effect and such effects vary significantly with time. Therefore, the students, faculty, and staff' activity profile data are combined into one dataset since each group has only one observation (i.e., proportion of activity participation) for each activity at a certain time. This causes the model not to be appropriately estimated. With the combined dataset, the variables, activity type and time, are considered as explanatory variables in the model of an activity profile. Moreover, the model needs to take into account whether the interaction effect between activity type and time and the interaction variable, activity $\otimes$ time, need to be included in the model. Therefore, the candidate model structure is as follows:

$Y_{ijt} = \mu + \alpha_j + \beta_{j,t} + \varepsilon_{ij,t}$ i=1,2,3 represent group (faculty, students, and staff)

where,

Y_{ijt} : $log \dfrac{\hat{p}_{ijt}}{1-\hat{p}_{ijt}}$ ($\hat{p}_{ijt}$: proportion of group *i* performing activity *j* at time *t*)

μ : grand mean

α_j : activity type (j=1,..,5 represent activity)

$\beta_{j,t}$: interaction (activity $\otimes$ time, t=1,...,24 represent time)

One way to decide whether the interaction effect exists is visual inspection. If the effect of activity type is the same for all time periods, the lines in an interaction plot (i.e., activity profile graph) are parallel. However, the activity participations (proportions) vary according to time changes so that there is an interaction effect and the interaction term, activity $\otimes$ time, should be included in the model.

Table 4.10 and 4.11 show the summary results of estimated models. In comparison, the estimated model shown in Table 4.10 only includes the variable of activity type while activity type and an interaction variable (activity $\otimes$ time) are considered in the model shown in Table 4.11. Usually, the values of the interaction terms are difficult to interpret, but the important point is that they allow the y_{ijk} 's to change in a non-additive way. Both of models are seen to be statistically significant based on the model goodness-of-fit. The model with activity type and the interaction variable looks much better than the model only with the activity type variable such that the R-square values are 0.965 and 0.634, respectively. The model with the interaction variable seems to be fitted very well. Figure 4.12 shows the fitted activity profile for university travelers and the differences between observed and the estimated values. The differences between observed and estimated proportion of activity participation are very small so that the model is able to successfully replicate the actual activity profile.

In this study, the proposed model is targeted to develop a model for an aggregated activity profile that is much easier than the model for disaggregated components of activity schedule. There is no doubt that less complexity in model development using such a method would be better with respect to practical model application. Otherwise, many appropriate models have to be developed independently for every element that constructs the activity schedule. This, of course, highly increases the complexity of whole travel demand model structure and is impractical in practice. Consequently, modeling an aggregated activity profile will provide many benefits with respect to less complexity in the activity-based travel demand models for practical use.

Table 4.10 Activity Profile Model (Model I: With no interaction term)

Source	DF	Sum of square	Mean Square	F-value	Prob(>F)	
Model	5	34.421	6.8841	211.39	<0.001	***
Error	355	11.56	0.03256			
Uncorrected Total	360	45.982				

Signif. codes: `***' 0.001

Model goodness of fit

R-Square	Coeff Var	Root MSE	y Mean
0.6339	90.2351	0.180461	0.19999

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Intercept	1	14.39855644	14.39855644	442.13	<.0001
Activity	4	20.02242936	5.00560734	153.71	<.0001

Table 4.11 Activity Profile Model (Model II: with interaction term)

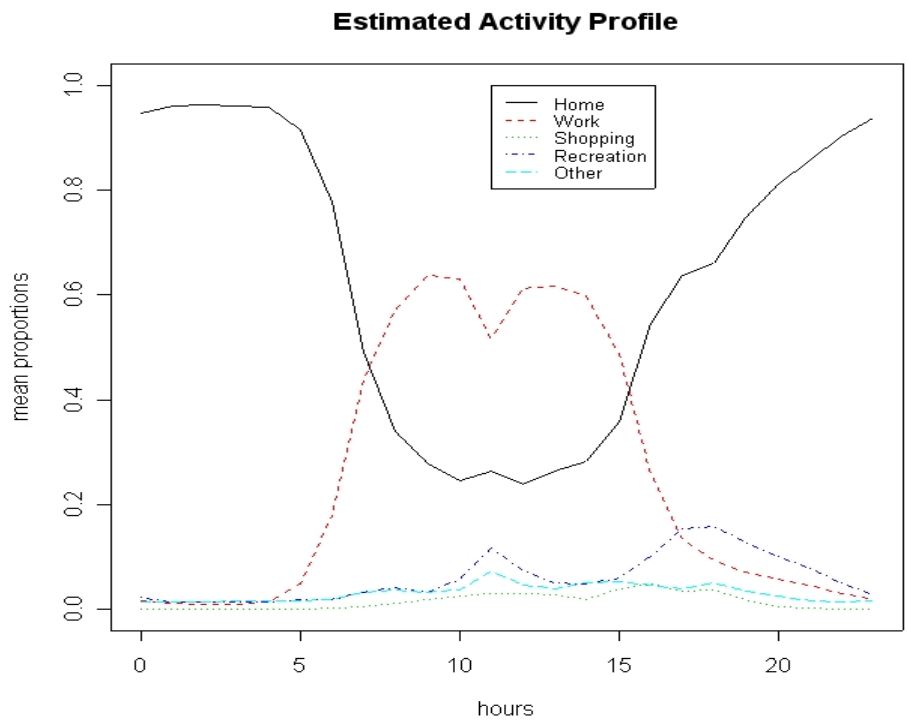
Source	DF	Sum of square	Mean Square	F-value	Prob(>F)	
Model	120	44.8862	0.3740	81.93	<0.001	***
Error	240	1.095	0.00456			
Uncorrected Total	360	45.982				

Signif. codes: `***' 0.001

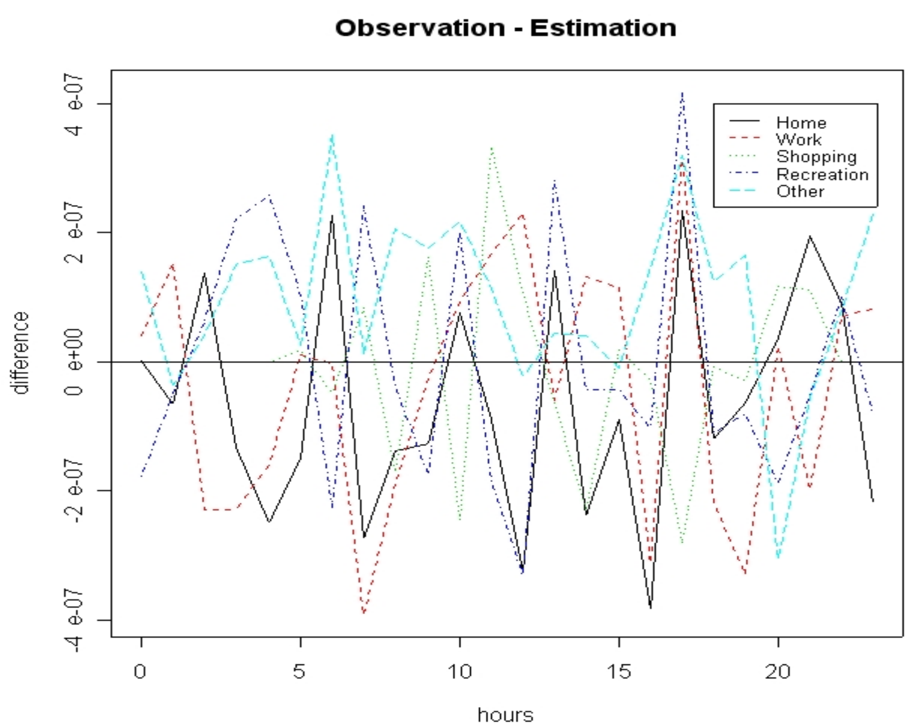
Model goodness of fit

R-Square	Coeff Var	Root MSE	y Mean
0.9653	33.7859	0.067568	0.19999

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Intercept	1	14.39855644	14.39855644	3153.77	<.0001
Activity	4	20.02242936	5.00560734	1096.40	<.0001
Activity*time	115	10.46529127	0.09100253	19.93	<.0001



(a)



(b)

Figure 4.12 Modeled Activity Profile: a) Fitted Activity Profile, b) Difference of Activity Profile between Observation and Estimation

CHAPTER 5

DEVELOPING SPATIAL ACTIVITY CONSTRAINTS USING A DESTINATION CHOICE MODEL

This chapter develops spatial activity constraints that are used in the estimation of the presence of university travelers on a campus. The spatial activity capacity is developed from a destination choice model based on a Multinomial Logit (MNL) model. The review of destination choice models applied in other studies is accomplished and the appropriate MNL model structure for NCSU data is introduced. Finally, the estimated probability of building choice by student is presented.

5.1 Activity Constraints (Capacity)

The “activity constraint” is another activity-based model component which takes into account the activity and trip with limited time and space. Each activity and trip independently exist since not all activities can be placed into a schedule at all times. As Hägerstrand (1970) introduced, three constraints, can be considered in the activity-based model:

- capability constraints
- coupling constraints
- authority constraints

The capability constraints are biological constraints such as eating and sleeping. The coupling constraints reflect a certain activity happening at the same place at the same time. The authority constraints are regulations such as opening hours of shops and working hours of employees. With the combination of these constraints over space and time, people can participate in activities. Modeling activity constraints is not easy since the required data are not normally available so an additional special survey may be required. This is another model component that could increase model complexity and difficulties in practical applications. As a result, no activity-based approach has been developed that would entirely replace the four-step travel demand model [8].

In this study, the university facility-based destination choice model developed from activity dairy survey data will take into account the activity constraints since the choice behavior for destination choice can reflect the combined constraints over space and time.

5.2 Activity Constraints on University Special Generator

In activity-based models, it is assumed that each person is engaged in one activity during any time period of a day. Subsequently, the model has to consider where they pursued what activities within a study area. The activity capacity is defined here as the value of attractiveness of each place (building, business or TAZ) where people can participate in a certain activity in a certain time period at a specific location. A place where a higher activity capacity exists can attract more people than the place with a lower capacity in the model.

The attractiveness of the spatial dimension where people pursued a certain activity can be obtained based on a complete inventory of facilities (buildings) and businesses containing the address, business type classified by standard industry code (SIC), and, if available, number of employees of each business or building.

The spatial dimension of a university student's activities for this study is limited to North Carolina State University. Therefore, the activity capacity will be developed only for the university campus. While private business information (i.e., business type and employees) is used in Triangle Regional Model, the number of employees in each university building is not available. Instead, the information such as building use, building square footage and number of seats (academic/teaching facility), if available, can be useful in developing the activity capacity in terms of attractiveness of buildings by activity types.

The university building information was obtained from the university facility inventory database at North Carolina State University. As shown in Figure 5.1, the geographic layer of university facilities on the main campus shows all facilities existing in 2001. Table 5.1 also shows the summary of NCSU facilities operating in 2001. The facility type is classified based on actual building use that is defined by the designated purpose of each building. All campus buildings are classified into ten types. As expected, the facility for research and teaching is the major use of buildings on campus. Of the 332 buildings, 93 buildings are used for research and 86 buildings are used for teaching. The housing facilities such as dormitories and graduate apartments represent 62 buildings. The highest square footage is seen at teaching facilities due to the number of classrooms in a building, while the total number of research buildings is higher than the teaching facilities.

In order to calculate activity capacity for a university facility, it is important to know what facility is associated with what activities and if a certain activity is possibly related to a certain facility, and how many of each activity can be generated. Sometimes, this would be simple, but in most cases, it is not easy. For example, the teaching and research facilities seem to be mostly associated with ‘Work’ activities rather than ‘Shopping’ and ‘Recreation’ activities.

Figure 5.1 Map of NCSU Main Campus Building Facility in 2001

Table 5.1 Summary of University Facility Type

Facility Type	Buildings		Area	
	Number	%	Square-foot	%
Academic/Teaching Facility	86	25.9	4,224,064	43.4
Administration/Office Facility	30	9.0	454,657	4.7
Athletic Facility	31	9.3	345,460	3.6
Housing Facility	62	18.7	2,394,163	24.6
Library	2	0.6	138,977	1.4
Maintenance/Warehouse	6	1.8	3,161	0.03
Parking Facility	3	0.9	507,720	5.2
Research Facility	93	28.0	1,233,581	12.7
Student Service Facility	8	2.4	407,780	4.2
Unclassified	11	3.3	14,039	0.1
Total	332	100.0	9,723,602	100.0

Note: Square-footage is totaled only for the buildings with data. 140 buildings with the type of Athletic, Housing, Parking, Research, and Academic have no square footage.

However, the university student center is classified as ‘Student Service Facility’ and is associated with various activities such as ‘Work’, ‘Shopping’, ‘Recreation’, and ‘Others’. In this case, it is hard to determine the proportion of each activity.

For consistency with the modeled activity profile, the activity types considered in activity capacity are defined as five types; Home, Work/School, Shopping, Recreation, and Other. However, the ‘Home’ activity is excluded in calculation of activity capacity since the university facility inventory provided the actual number of university beds in dormitories and graduate housing facilities. The actual number of beds will be applied in the ‘home’ activity capacity.

Figure 5.2 shows an algorithm for spatial activity capacity proposed in this study. In the CentreSIM study, the activity capacity was developed based on the business information and ITE Trip Generation Manual [44]. The local business information provides the business type and number of employees. Based on the facility and business type, the ITE Trip Generation Manual roughly provides the number of auto trips at a given unit of area. These auto trips were considered in developing activity capacities by activity type and that procedure was manually done for all business types. For example, a grocery store has ‘Work’ activity capacity as 1 and 20 for ‘Shopping’ activity as shown in Table 5.2. Therefore, if a grocery store has 5 employees, then the ‘Work’ activity capacity is 5 (i.e., 1 ‘Work’ capacity * 5 employees) and ‘Shopping’ activity capacity is 100 (i.e., 20 ‘Shopping’ capacity *5 employees). This means the total number activities accommodated with this business in an hour is 105.

A major disadvantage of this procedure is that it requires too much effort due to the labor required to develop activity capacity for all types of businesses although the activity capacity is disaggregated. Further, this assumes that the same business type will have the same activity capacity no matter the actual business. In reality, one store attracts more customers than the others. This is affected by many factors including ones not measured such as store decoration, workers’ courtesy, and neighborhood.

Table 5.2 Example of Activity Capacity by SIC Code

SIC Code	Home	Work/School	Shopping	Recreation	Other
5411 (Grocery stores)	0	1	20	0	0
5812 (Restaurants)	0	1	0	20	0
8748 (Business consulting)	0	1	0	0	0

Source: CentreSIM (2002)

Considering all these factors is not an easy task so that, in this study, the behavioral approach that considers actual travelers' choice is proposed. It is assumed that travelers make a destination choice by considering both known and unknown factors simultaneously and the choice probability represents the comparable attractiveness of one particular place from another. This can be represented by a destination choice model (DCM) based on travel survey data. The destination choice model will reflect the combined effect of attractiveness of each facility or TAZ by activity type. The estimated choice probability of each facility or TAZ by activity type will be used as the activity capacity.

One of the concerns in the practical application of a destination choice model is that destination choice models usually do not provide the choice probability for all facilities or TAZs within a study area since the number of TAZs or facilities are too many to estimate choice probability for them. Moreover, travel survey data do not include choice behavior data for entire TAZs. Hence, clustering facilities or TAZs into several groups by considering the similarity of socio-demographic information such as household size and business type is often required.

This study clusters all university buildings into a few facility types based on the data available for DCM model. Once the choice probabilities for the facility types are obtained, then a certain value presenting the physical characteristic of individual buildings like square-footage is considered to develop an individual building activity capacity.

Figure 5.2 shows the proposed algorithm for developing an activity capacity. First, all university buildings need to be classified into a few types based on building use. Then, student campus activities need to be identified for a university facility. The next step develops the required data for the destination choice model by activity type.

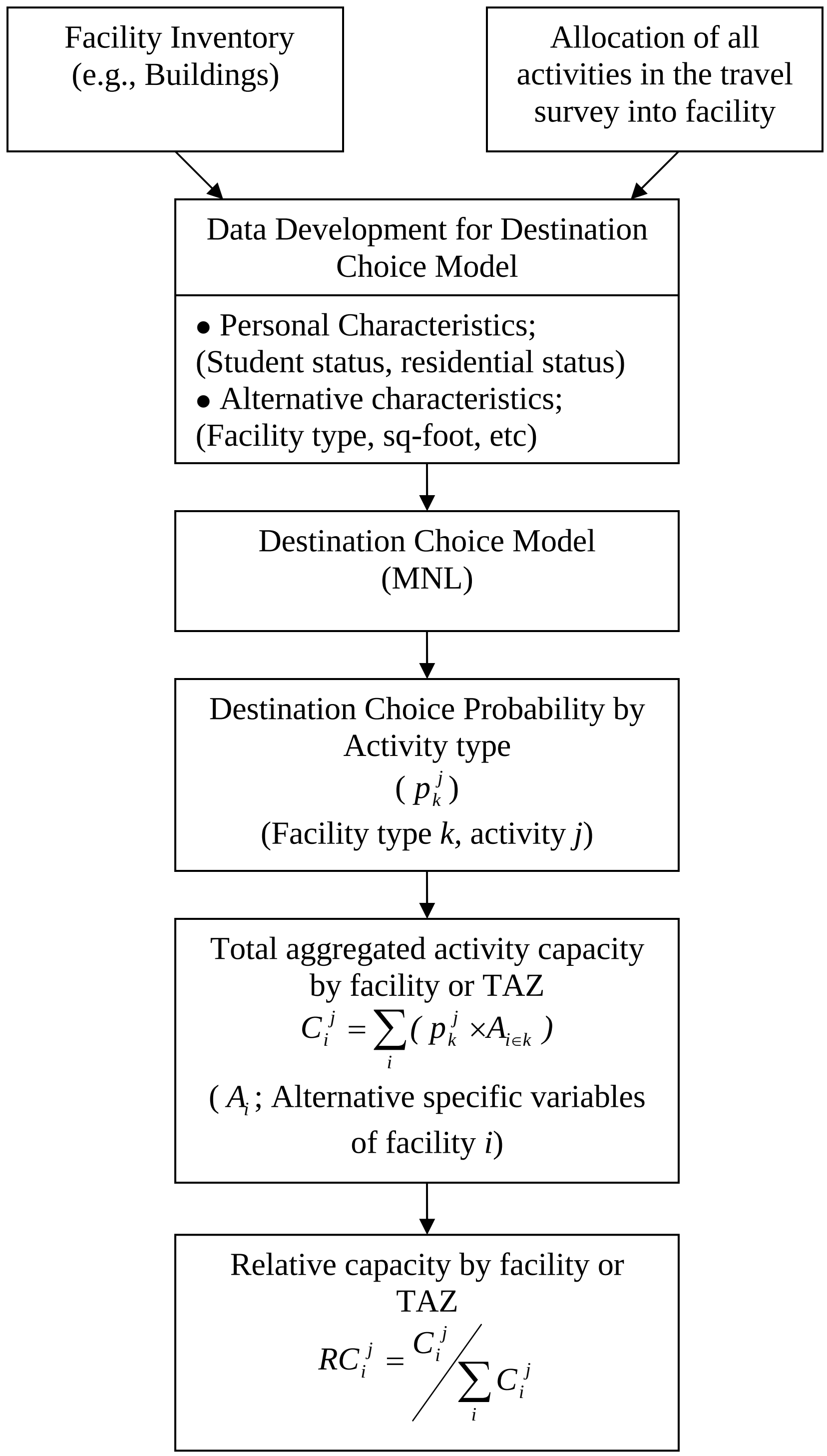


Figure 5.2 Activity Capacity Algorithm

The data structure varies depending on what destination choice model is used by considering data availability and the purpose of model development. This will be discussed in detail later this chapter. When the data are ready, the model runs and estimates choice probabilities by facility types by activities.

In order to develop a relative activity capacity, the facility specific data (i.e., square-footage) available at an individual building needs to be applied to a choice probability of that facility type and activity type. Then, the individual capacity value for each building will be totaled by activity type and the relative activity capacity will be obtained by calculating the ratio of individual building capacity divided by total capacity. This relative activity capacity represents the relative attractiveness of each building by activity within the study area. If activity travel survey data are available for all population groups and for model development, the relative activity capacity can be developed by traveler group.

5.3 Review of Destination Choice Models

5.3.1 Overview

In a conventional travel demand modeling approach, the gravity model has been most commonly used in the trip distribution step, but recent studies show that destination choice models are more frequently used when data are available. As described in Chapter 2, the gravity model uses one functional form with aggregate zonal impedances based on zone size and travel time, distance, and/or cost. This model does not take into account the travel behavior of travelers. Unlike the gravity model, destination choice models use disaggregate data including numerous variables inside the utility function. These models describe each traveler's choice of destination in probabilistic terms by simultaneously modeling effects of the destinations, origin-destination travel conditions, and personal characteristics [45].

Three main elements such as destination attractiveness, travel conditions between origin and destination pairs, and personal characteristics have been considered and modeled. The attractiveness of destination could be defined in terms of business category, total employment, and square footage by land use category. Travel conditions are measured by travel time or travel cost for travel modes. Personal characteristics are age, gender, household size, income, auto availability, and employment status.

Logit models are widely used in destination choice models. Although Probit models have the ability to capture correlations among alternatives, very few applications have been developed due to the high complexity of model estimation. Logit models are based on a probability distribution function that maximizes a series of random variables. Multinomial Logit (MNL), Nested Logit (NL), Cross-Nested Logit (CN), and General Extreme Value (GEV) assume that the error components of the utility function are identical and independently distributed (IID) with type I extreme-value distribution (Gumbel, 1958). Therefore, these models are called "Closed-form" logit models. Unlike closed-form logit model, "Open-form" logit models completely relax the IID error component assumption. The mixed Multinomial Logit (MMNL) model is a mixed model enhanced with mixing distributions [46].

Logit models have been successfully used in the mode choice step in conventional travel demand models with small number of alternatives, usually less than five travel modes. However, destination choice models have many alternatives for travel since the choices are made with respect to all TAZs in study area. Thus, destination choice models need to cluster the TAZs into several districts based on similarity of socioeconomic and demographic variables within a TAZ, or cluster households into several categories in most applications.

5.3.2 Discrete choice model

The model structure of destination choice models that is widely used is often called a "discrete choice model" that seeks to maximize the utilities of the destination choices. The decision makers can be any decision-making unit such as people, household, or firm. Within a discrete choice framework, the decision maker chooses only one alternative from the choice set which is mutually exclusive from one another. All available alternatives should be included and the number of alternatives needs to be finite [47].

The discrete choice models are usually derived under an assumption of utility maximization behavior by the decision maker. The utility function is the form of either alternative specific variables or alternative and personal specific variables together. Also, the error term is included for considering uncertainty that could be result from unobserved individual preferences, individual characteristics, alternative attributes, and measurement errors due to incomplete information.

The utility that a decision maker, i, faces a choice among D destination alternatives can be described as U_{id}, $d = 1, \cdots, D$. This utility is known to the decision maker who chooses the destination that provides the greatest utility. Therefore, the decision rule is mathematically shown as follows.

$$U_{id} \geq \max_{d=1,\ldots,D_i} U_{id} \qquad (5.1)$$

where,

U_{id} : Utility function of decision maker i and chosen destination d.
D_i : Number of destinations (TAZs) in decision maker i's available choice set.

The choice probability for the MNL is calculated precisely since it is closed-form as mentioned before. The choice probability of destination d being chosen by individual i is based on the utility of that destination verses the aggregate utility of all available destinations for the trip maker. The conditional choice probability of the MNL is below.

$$P_{id} = \frac{e^{U_{id}}}{\sum_{d \in D_i} e^{U_d}} \tag{5.2}$$

where,

P_{id} : Probability of trip maker i choosing destination d
U_{id} : Utility of destination d for individual i
D_i : Choice set of all available alternative destinations for trip maker i

The individual destination choices have been grouped by trip purpose and distribution models are developed for each trip purpose. Segmenting destination choices by trip purpose is an efficient way to develop trip distribution models and needs to be consistent with the trip purposes in trip generation [48]. However, it was recognized that, unlike mode choice models, establishing the dataset required in discrete choice model structure is not an easy task due to the number of alternatives (i.e., destinations). Consequently, no models have been developed for providing a choice probability for each of all destinations, especially for regions with more than a thousand TAZs.

5.4 Developing Destination Choice Model for Activity Capacity

5.4.1 Overview

The discrete choice model structure discussed herein is conceptually and theoretically superior to the gravity model in trip distribution in regional travel demand models. However, the discrete choice models that have been developed for destination choice models may be different from the model for estimating activity capacity on university buildings regarding several issues as follows.

- Students' destination choices on campus for all activity types are not actually dependent on their own choices, but depend on the given university environment. For example, the building choice for a 'Work/School' activity totally depends on class schedule.
- The choice probability may not change if the changes in personal and alternative characteristics such as student status, residence status, travel time, and distance are expected.
- The given specific alternative (building) information is not distinguishable among alternatives in terms of attractiveness of alternative.
- Some activities, such as 'Shopping' and 'Recreation', may have not enough data to estimate choice probability from discrete choice models because only a few buildings are related to these two activities.

For this reason, the purpose in developing a discrete choice model in this study is to estimate the attractiveness of each building by activity type rather than to analyze students' destination choice behavior corresponding to the changes of personal and alternative characteristics.

The discrete choice model selected here to analyze the destination (facility) choice of student is the multinomial logit (MNL) model. Choosing a destination of activity is a form of discrete choice. A random utility based MNL model for the destination choice of the students is specified. The utility function is defined as the linear form as follows.

$$U_{id} = \alpha_d + \beta_{dP} P_i + \beta_{dT} T_i + \beta_{dA} A_i + \varepsilon_{id} \qquad (5.3)$$

where

U_{id} : Utility of facility type d for student i,
α_d : Estimable alternative specific constants,
β_d : Estimable coefficients,
ε_{id} : Type I extreme value (Gumbel) distributed random error terms,
P_i, T_i, A_i : Vectors of personal, and trip information for student i, respectively.

If utility maximizing behavior is assumed, this utility leads to the MNL model:

$$P_{id} = \frac{e^{U_{id}}}{\sum_{d \in D_i} e^{U_d}} \qquad (5.4)$$

where,

P_{id} : Probability of student i choosing facility type d
U_{id} : Utility of facility type d for student i
D_i : Choice set of all available alternative facility type for student i

All variables are specific to the particular student. As mentioned before, this aspect is a limitation of the destination choice model, particularly as a predictive model of destination choice. However, the model will show how the different student specific variables affect destination choice. The unavailable alternative specific information is captured by the alternative specific constant.

5.4.2 MNL model estimation

In SAS, the choice set of facility type is a nominal response variable and the generalized logit model can fit this data. Two options are available to fit the model; one is ‘PROC CATMOD’ and the other is ‘PROC LOGISTIC’ which does analysis on nominal responses with ease [50].

In a generalized logit model, a particular destination of responses will be selected as the baseline reference that has the default utility of zero, because the model is based on differences in utility and because the variables are not alternative specific. The other utilities are interpreted as relative to the baseline reference. For example, if a facility ‘n’ is chosen as the baseline destination, then the probability of choice for an individual destination need to be estimated based on the following mathematical forms.

$$log\,it(D_1) = log(\frac{\pi_1}{\pi_n}),$$

$$log\,it(D_2) = log(\frac{\pi_2}{\pi_n})$$

Where,

π_1 = probability of choosing destination ‘1’
π_2 = probability of choosing destination ‘2’
$\vdots$
π_n = probability of choosing destination, ‘n’.

This model allows ‘n-1’ different sets of regression parameters, one for each logit.

In order to develop a dataset, first the students' personal information need to be connected to the database called 'Place' in which every activity is recorded with the place where activity happened. The place information in terms of university building, then, needs to be grouped into several facility type, based on the recorded frequency of activities. If there were not enough observations at a certain facility, the facility would not be included in destination choice set.

Table 5.3 shows the observed activity frequency by ten facility types. The recorded on-campus activities total 2293. The most visited facility is found to be academic buildings (66.4%) for 'Work/School' activity and very few observations are recorded for some facility types such as 'Bookstore' and 'Student Health Center'. Students are doing the 'Shopping' activity mostly at 'Bookstore' and 'Student Service' facilities although not so frequently. The student service facility is found to be most visited by students doing 'Recreation' activity and 'Academic/Teaching' facility is the second most visited facility.

Table 5.3 Observed Activity Frequency by Facility Type

Building Use	Frequency				Sum (%)
	Work/School	Shopping	Recreation	Other	
Athletic Facility	7	0	5	5	14(0.6)
Bookstore	0	4	2	1	7(0.3)
Academic/Teaching	1369	1	103	49	1522(66.4)
Student Health Center	3	0	0	4	7(0.3)
Housing Facility	21	3	42	20	86(3.8)
Library	110	0	5	7	122(5.3)
Office	9	0	25	3	37(1.6)
Research	15	0	0	1	16(0.7)
Student Service	13	9	433	6	461(20.1)
Other	6	1	12	2	21(0.9)
Total	1553 (67.7%)	18 (0.8%)	627 (27.3%)	95 (4.1%)	2293 (100%)

Note: 2293 activities are of total on-campus activities.

Since all activities do not occur at every facility, the destination choice model has to be developed by activity type in which different choice sets (facility) will be defined. For example, the ‘Recreation’ activity does not happen at ‘Research’ facility. Hence, this facility will not be included in the destination choice model for ‘Recreation’ activity.

In this study, eight facility types are defined as alternatives for both ‘Work/School’ and ‘Other’ activities and seven types (i.e., not including ‘Research’ facility) for ‘Recreation’ activity in the destination choice model. The alternatives are assumed to be the available destination options that a decision maker (student) is supposed to consider during the choice process. The choice set for both ‘Work/School’ and ‘Other’ activity consists of the eight types listed below:

Choice set:

1. Academic/Teaching Facility
2. Library
3. Student Service Facility
4. Housing Facility
5. Research Facility
6. Administration/Office Facility
7. Athletic Facility
8. Other (student health center, theater, unclassified)

Table 5.4 shows the explanatory variables that represent students’ personal information and their travel time for individual activities. Personal information includes age, gender, credits student registered, education status (undergraduate and graduate), student status (full-time and part-time), residential status (on and off-campus), license to drive, employment status (full-time, part-time, volunteer, and unemployed). These variables are tested and used in development of MNL destination choice model by activity type.

Table 5.4 Explanatory Variables Considered in the Analysis

Variable		Description	Coding
Personal Information	Age	Age in year	Continuous
	Credit	Credit registered	Continuous
	Gender	If the person is male, then 1 Otherwise, 0	1: Male 0: Female
	Estatus	Educational status	1: Under 0: Graduate
	Sstatus	Student status	1: Full-time 0: Part-time
	Residence	Residential status	1: On-campus 0: Off-campus
	License	Licensed to drive	1: Yes 0: No
	Empstatus	Employment status	Indicator ,1 if full-time worker Indicator ,1 if part-time worker Indicator ,1 if volunteer Indicator ,1 if unemployed
Travel	Ttime*	Travel time (minutes)	Continuous

5.4.3 MNL model results

Table 5.5 shows the summary of the model tested globally for the effect of each variable on the outcome variable, controlling for the other variables in the model. All the chi-square statistics are Wald statistics, not likelihood ratio statistics. Each chi-square is a test of the null hypothesis that the explanatory variable has no effect on the outcome variable.

In the destination choice modes for ‘Work/School’ and ‘Other’, there are seven degrees of freedom for each chi-square because each variable has seven coefficients. So the null hypothesis is that seven coefficients are zero. The log-likelihood ratio equals twice the positive difference between the log-likelihoods for the fitted model and the saturated model, and high p-values suggest a good fit. The p-value of 1.000 in all three models reassures that the models fitted well. The destination choice model for 'Shopping' activity was not estimated since the dataset is too small to estimate parameters. In order to get parameters, the choice set must include all datasets corresponding to every case that is a possible combination of data. For instance, if there are three alternatives with two explanatory variables such as gender and residential status (on-and off-campus), then the total number of possible combinations is twelve (i.e., 3*2*2). Hence, the dataset has to include all the possible combinations of data. Otherwise, the discrete choice model would not give any results of estimation. As shown in Table 5.5, the chi-square statistics of some variables were not estimated due to the lack of datasets that did not include all possible cases.

The SAS CATMOD procedure is used in the analysis. This analysis uses the largest value of dependent variables as a reference. Accordingly, the 'Other' facility is the reference variable to estimate the parameters for each facility type. Tables 5.6, 5.7, 5.8 show the estimation results for the MNL model of destination choice of students for 'Work/School', 'Recreation', and 'Other' activity respectively.

As shown in Table 5.6, a destination choice model for the 'Work/School' activity, the age in years shows that the older the student, the lower the propensity toward most of the facilities except for 'Library' and 'Student service' facility as compared with 'Other' facility. This would be a reasonable result, since with age a student is less likely to be involved in school activities. The older students are usually part-time and, as a result, they are registered for fewer credits. A related effect is seen for the 'Credit' variable, in which the more credits students are registered for, the lower the propensity toward 'Housing', 'Research', and 'Administration office'. On-campus residents have an increased propensity toward 'Housing' and ‘Athletic’ facilities. This is also a reasonable result, since the students living on-campus are more likely to choose a university housing facility and are more likely to use an athletic facility. It is also shown that, students are less likely to use library for longer travel times. However, travel time was not seen to be a critical effect on destination choice for other facilities. This is reasonable since the ‘Work/School’ activity usually happened at a classroom or research facility where students have to be there rather than choosing these facilities.

According to Table 5.7, three variables such as gender, residential status, and travel time are explainable by destination choice for the ‘Recreation’ activity. Interestingly, male students are more likely to choose ‘Library, ‘Student service’, ‘Housing’ and ‘Athletic’ facilities than female students as compared with ‘Other’ facilities. ‘Academic/teaching’ and ‘Administration office’ are not involved with ‘Recreation’ activity. This is reasonable since this facility does not facilitate any convenience for the students’ recreational activity. The on-campus residents are found to be less likely to choose library. However, the residential status is not quite as informative to explain the selection of facility for ‘Recreation’ activity. It is also seen that the longer the travel time, the lower the propensity to select ‘Student service’ and ‘Athletic’ facility.

Table 5.8 shows the destination choice model for ‘Other’ activity. Due to the lack of datasets to estimate parameters of all variables, only the age is selected as an explanatory variable. It is seen that the older the student, the lower the propensity toward ‘Academic/Teaching’, ‘Student service’, and ‘Housing’ facility as compared with ‘Other’ activity. This result is consistent with the result of ‘Work/School’ destination choice. The older students are less likely to be involved in ‘Other’ activities on campus.

Table 5.5 Summary of Maximum Likelihood Analysis of Variance

Variable	Activity							
	Work/School		Shopping		Recreation		Other	
	χ^2	**Pr>** χ^2	χ^2	**Pr>** χ^2	χ^2	**Pr>** χ^2	χ^2	**Pr>** χ^2
Age	21.31	0.0033**	--	--	--	--	12.74	0.0786
Credit	15.32	0.0321*	--	--	--	--	5.91	0.5499
Gender	5.38	0.6137	--	--	35.44	<0.0001**	--	--
Educational status	--	--	--	--	--	--	--	--
Student status	--	--	--	--	--	--	--	--
Residential status	18.77	0.0089**	--	--	59.87	<0.0001**	--	--
License	--	--	--	--	--	--	--	--
Employment status	--	--	--	--	--	--	--	--
Travel time	16.39	0.0218*	--	--	5.82	0.4432	5.43	0.6078
N	1,560		18		627		97	
DF	7		--		6		7	
Log-likelihood ratio	1431.10 (df:6e3)		--		140.26 (df:402)		265.99(df:581)	
p-value	1.0000		--		1.0000		1.0000	

Note: "—"indicates that the variable is not estimated. Significant: p< 0.05 `*', p<0.01 `**'.

Table 5.6 Estimation Results for MNL Model of Destination Choice of Students (‘Work/School’)

Variable	Destination Facility						
	Academic/ Teaching	Library	Student Service Facility	Housing Facility	Research Facility	Administration/ Office Facility	Athletic Facility
Intercept	8.8461**	5.3247**	--	7.8205*	8.3718*	7.9237*	8.0961*
Credit	--	--	--	-0.3409*	-0.2524	-0.2842*	--
Residence	--	--	--	2.8306*	--	--	1.8191
Age	-0.0753*	--	--	-0.1432	-0.1679*	-0.1342	-0.1972*
Travel Time	--	-0.0883*	--	--	--	--	--

Note: ‘Other’ building is defined as a reference variable. Coefficients that were not significant at the 90% level were restricted to zero and omitted from the table. Significant: p< 0.05 \`*', p<0.01 \`**'.

Table 5.7 Estimation Results for MNL Model of Destination Choice of Students ('Recreation')

Variable	Destination Facility					
	Academic/ Teaching	Library	Student Service Facility	Housing Facility	Administration/ Office Facility	Athletic Facility
Intercept	**2.9047****	--	**3.1179****	--	--	**2.5454****
Gender	--	2.8113*	**2.2159****	1.6872*	--	**2.4609****
Residence	--	-3.5128*	--	--	--	--
Travel time	--	--	-0.0657*	--	-0.0816	-0.0629

Note: 'Other' building is defined as a reference variable. 'Research' facility is omitted from destination facility due to no observations.
Coefficients that were not significant at the 90% level were restricted to zero and omitted from the table.
Significant: $p < 0.05$ \`*', $p < 0.01$ \`**'.

Table 5.8 Estimation Results for MNL Model of Destination Choice of Students (‘Other’)

Variable	Destination Facility						
	Academic/ Teaching	Library	Student Service Facility	Housing Facility	Research Facility	Administration/ Office Facility	Athletic Facility
Intercept	6.2477	--	21.7019*	**17.9458****	--	--	--
Age	-0.1351	--	-0.8008*	**-0.6019****	--	--	--

Note: ‘Other’ building is defined as a reference variable. Coefficients that were not significant at the 90% level were restricted to zero and omitted from the table. Significant: $p < 0.05$ `*', $p<0.01$ `**'.

Based on the results of destination choice models, the estimated probability of facility choice for each activity is reported as shown in Table 5.9. The probability of facility choice for 'Shopping' activity is not estimated from MNL model, the proportional values calculated from observed 'Shopping' activities replace the probability table. As seen in the table, the sum of the estimated probability of each facility by activity type is 1.0.

Table 5.9 Estimated Probability of Facility Choice by Activity Type

Facility Type	Activity			
	Work/School*	Shopping	Recreation*	Other*
Academic/Teaching Facility	0.8043	0.0000	0.2961	0.2904
Library	0.1259	0.0000	0.0404	0.0013
Student Service Facility	0.0071	0.8235	0.3614	0.0000
Housing Facility	0.0197	0.1765	0.0556	0.0000
Research Facility	0.0161	0.0000	0.0000	0.0320
Administration/Office Facility	0.0200	0.0000	0.0231	0.0016
Athletic Facility	0.0030	0.0000	0.2050	0.1599
Other	0.0039	0.0000	0.0184	0.5148
Total	1.0000	1.0000	1.0000	1.0000

Note: '*' illustrates that the probability obtained from MNL model

According to the estimated probability of facility choice, about 90% of students would expect to choose 'Academic/Teaching' facility and 'Library' for their 'Work/School' activity. Thirty-six percent of students would expect to choose 'Student Service' facility for their 'Recreation' activity and 20% of them may choose the 'Athletic' facility. With the 'Other' activity, the 50% of students would choose 'Other' facility and secondly, the 'Academic/Teaching' facility would be chosen by about 29% of students. As expected, this output of facility choice corresponds to the observed activity and students' facility choice.

5.4.4 Adjusting activity capacity

The estimated destination choice probability needs to be adjusted because the students' daily activities happen either inside or outside campus while the study area is limited to university campus. Table 5.10 shows the distribution of activity participation by place where activities were happened. About 88% of the 'Work/School' activity occurred on campus and only 13.6% of the 'Shopping' activity took place off campus. This makes sense because very few buildings on campus are associated with the 'Shopping' activity. Students make more 'Other' activities off campus (60.6%). This proportion will be applied to the estimated activity capacity to obtain actual activity capacity by building.

Table 5.10 Distribution of Student Activity Participation by Place

Place	Activity			
	Work/Class	Shopping	Recreation	Other
On-campus	87.7%	13.6%	68.3%	39.4%
Off-campus	12.3%	86.4%	31.7%	60.6%

5.5 Developing Relative Activity Capacity

The relative activity capacity of a building can be represented by the probability of destination choice of individual buildings. The MNL model developed herein estimated the probability of destination choice for the eight facility types. In order to develop the choice probability by individual building, it has to be disaggregated by individual building level to obtain relative capacity. The only available information that differentiates each building is a square-footage and this is used in developing a relative activity capacity.

The activity capacity of an individual building is a function of the probability of a facility chosen and the square-footage of a building included in that facility.

$$C_i^j = p_k^j \times (S_{i \in k})$$

where

C_i^j : Activity Capacity of building *i* for activity *j*

$S_{i \in k}$: 1,000 square-feet of building *i* included in facility type *k*

p_k^j : Choice probability for facility type *k* for activity *j*

The relative activity capacity of an individual building is then calculated from the ratio of individual building activity capacity to the total activity capacity as follows:

$$RC_i^j = C_i^j \Big/ \sum_i C_i^j$$

where

RC_i^j : Relative Activity Capacity of Building *i* in activity *j*

This relative activity capacity value represents that the higher relative activity capacity a building has, the more students that can be accommodated. The relative activity capacity is calculated from the campus building layer for each of 332 buildings and will be used in the estimation of the total number of students presented in a building for a certain activity. This will be discussed in detail in Chapter 6.

CHAPTER 6

ACTIVITY-BASED UNIVERSITY TRAVEL DEMAND ESTIMATION

This chapter estimates the spatial-temporal building presence and activity-based travel demand on NCSU campus. The “building presence” means the presence of students in the building or buildings of interest. The spatial-temporal activity-based model provides the building presence, the number of students accommodated in each building, and hourly travel demand in terms of trip productions and attractions. The estimated travel demand will be compared to the observed building presence and student trips counted at all entrances of sampled buildings for model validation. Finally, the estimated trips from the activity-based model are compared to the ones from a traditional four-step model to ascertain which model would be better based on trip generation.

6.1 Procedure of Activity-based Travel Demand Estimation

Figure 6.1 depicts the procedure for travel demand estimation in this Chapter. The travel demand estimation consists of two steps: 1) estimation of building presence, 2) estimation of trip generation (i.e., production and attraction). In order to estimate building presence, the number of people in each traveler group needs to be identified from the university demographic information and the number of people by group will be applied to the activity profile obtained in Chapter 4. The activity profiles multiplied by the number of persons in each traveler segment result in the number of persons engaged in each activity in each hour. Next, the activity capacity modeled in Chapter 5 will spatially distribute all travelers onto campus.

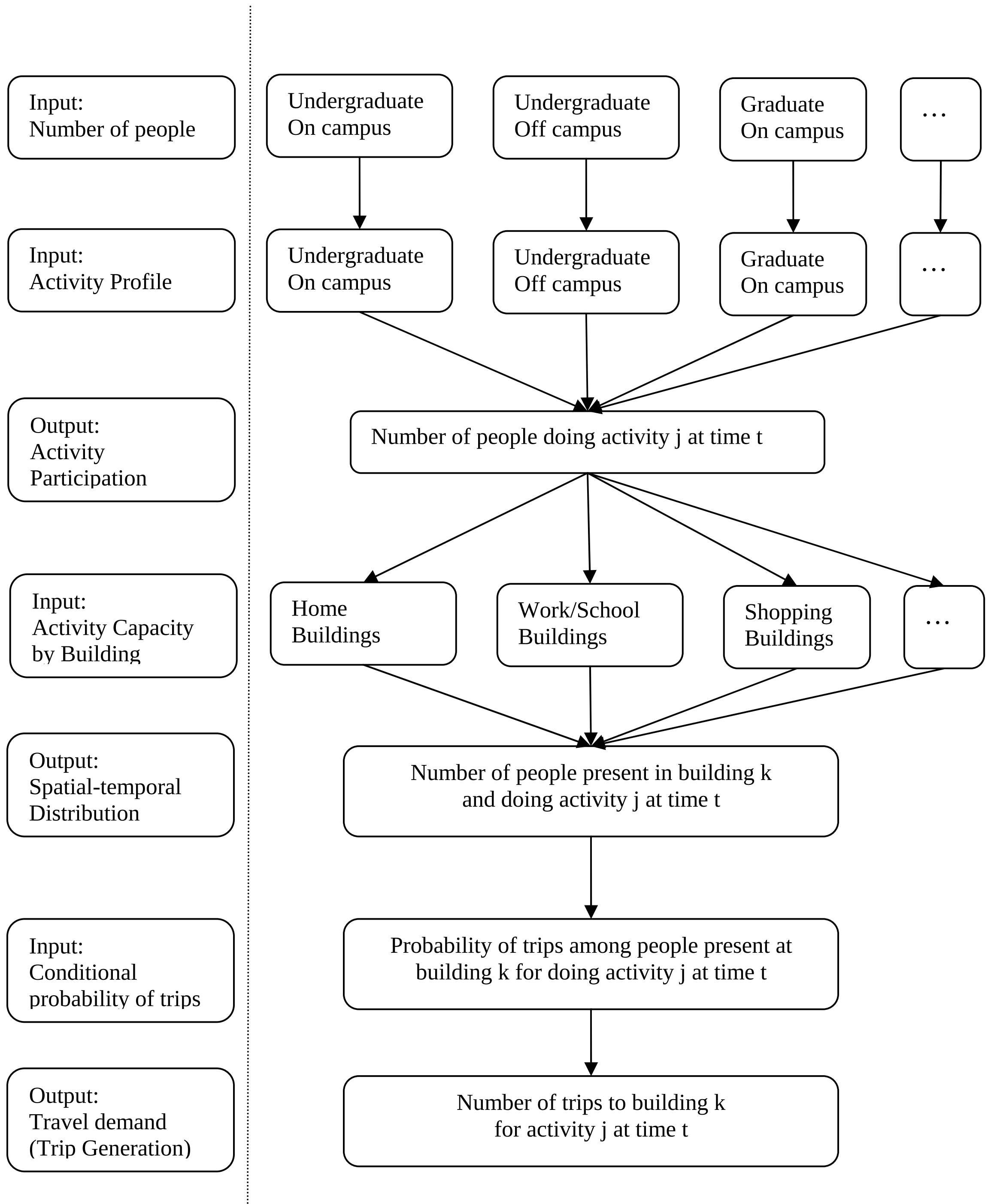


Figure 6.1 Procedure of Activity-based Travel Demand Estimation

The spatial distribution of the five activity types forms the basis of a 24-hour accounting system of the population called the building (or zone) presence model [17, 22]. These are the data that describe dynamic changes of activity demand for land over time and also the availability of land for activities. All the spatial data required for the model for activity demand and allocations are stored in the TransCAD. At each hourly period of allocation process, TransCAD retrieves the proper activity constraint and runs the calculation of hourly building presence.

Trip generation will then be estimated based on building presence data and the conditional probability of trips in which the departure time and current activity type are considered in the form of MNL model structure. The conditional probability of trips represents the probability of how many people make a trip among all who participated in a certain activity at a certain time. The conditional probability of trips at given time and given activity by traveler group is multiplied by the building presence. Then this results in trip productions and attractions.

6.2 Estimation and Validation of Student Presence in Buildings

6.2.1 Building presence estimation

Building presence is simply calculated from a function of the number of persons in each traveler group, their corresponding activity profiles, and individual building activity capacity. The hourly building presence is obtainable since the activity profile determines the number of activity participations in every hour during a day.

Equation 6.1 illustrates the function for building presence.

$$Z_{i,j,g}(t) = P_{jg}(t) \bullet N_g \bullet RC_{ij} \qquad (6.1)$$

Where:

$Z_{i,j,g}(t)$: Total number of persons in traveler group *g*, engaged in activity type *j* at time *t* presented at building *i*

$P_{jg}(t)$: Probability of participation in the given activity *j* at time *t*

N_g : Number of persons in traveler group *g*

RC_{ij} : Relative activity capacity of building *i* for activity *j*

The output of building presence is the number of people participating in each of the five activities in each building. Figure 6.2 through 6.5 show the example output of building presence created from TransCAD. TransCAD includes the estimated building presence within the database of a building layer and generates bar-graphs to show the number of people in each building on NCSU campus by activity and time segment.

For example, in the morning hours (i.e., 9 a.m. to 10 a.m.) students seem to be evenly distributed at ‘Housing’ and ‘Academic/Teaching’ buildings but the number of students at ‘Housing’ buildings decreases as the hour passes, and increases at ‘Academic/Teaching’ buildings for ‘Work/School’ activity. Conversely, the presence of students is high at ‘Housing’ buildings and low at ‘Academic/Teaching’ buildings after 3pm. Many students were presented at ‘Student Service’ and ‘Athletic’ buildings in the afternoon for recreational and other activities. This output makes sense since most class hours are before 4pm and students are doing recreational activities after then. Only a few buildings are related to ‘Shopping’ activity, which is reflected by the destination choice model capturing a small number of destinations available on campus for ‘Shopping’ activity from students’ activity diary. The building presence seems to successfully reflect the students’ daily spatial and temporal activity patterns on campus.

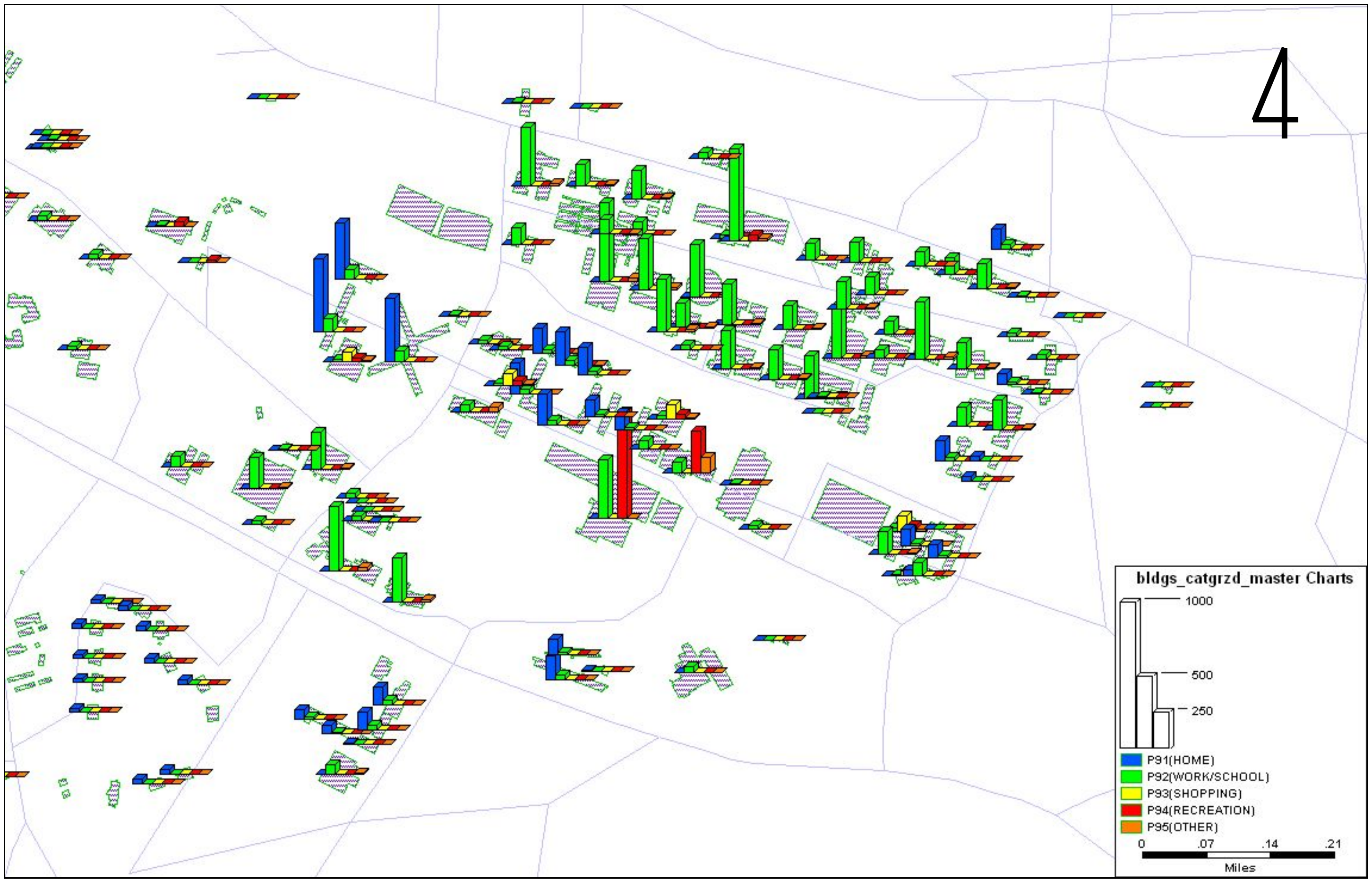


Figure 6.2 Number of Students Present in Each Building (Hour: 9 a.m. -10 a.m.)

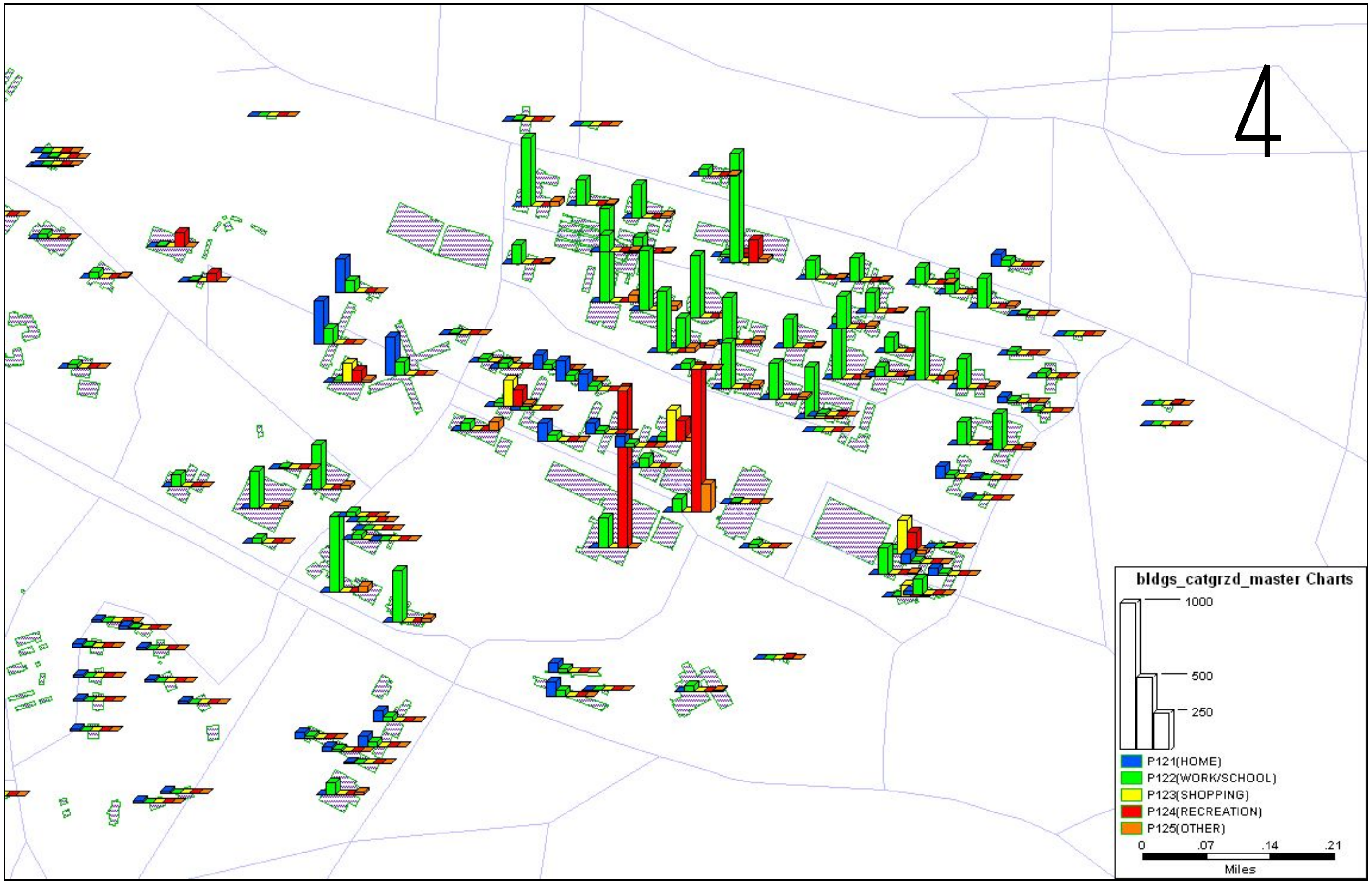


Figure 6.3 Number of Students Present in Each Building (Hour: 12 p.m.-1 p.m.)

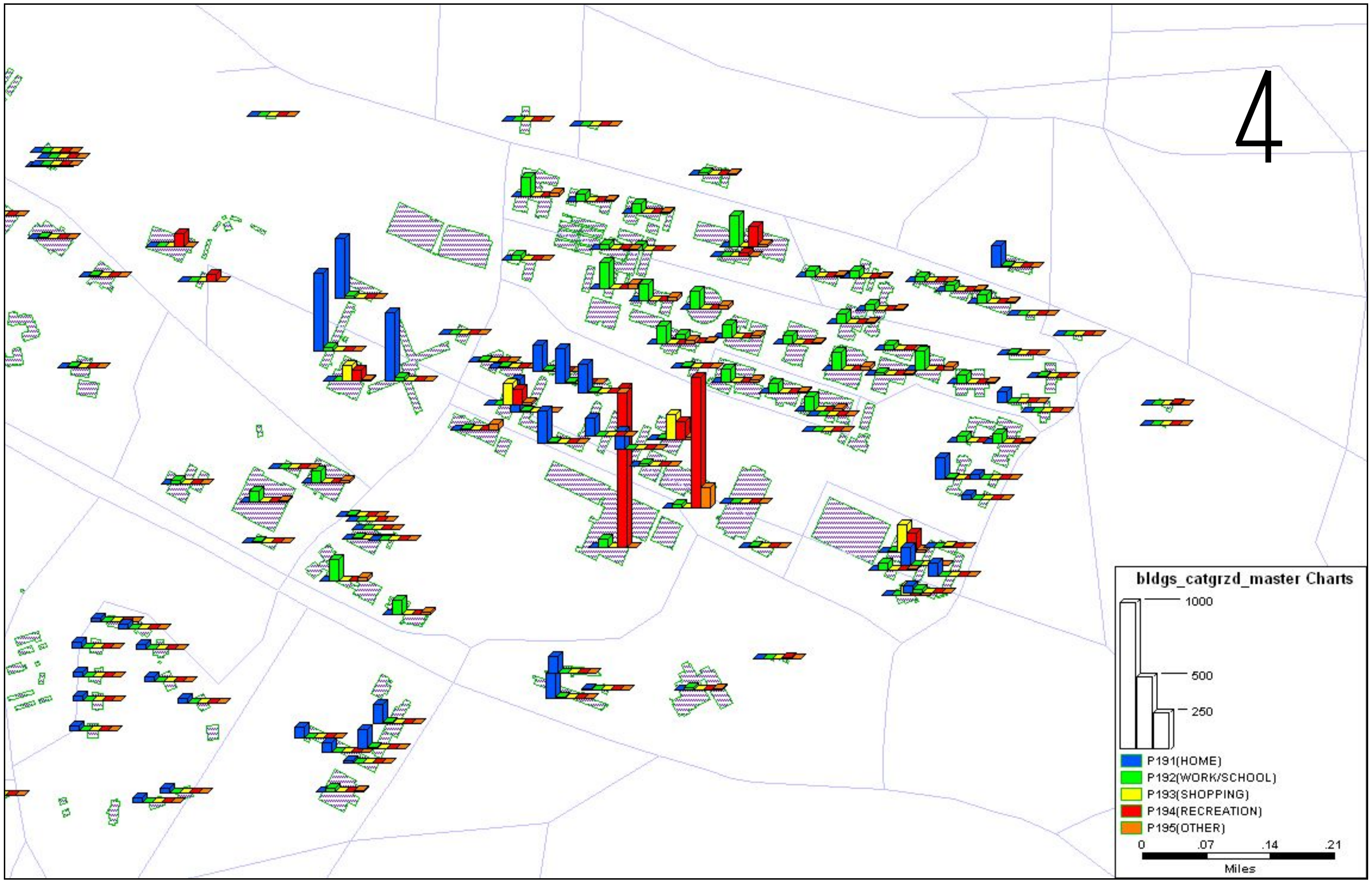


Figure 6.4 Number of Students Present in Each Building (Hour: 3 p.m.-4 p.m.)

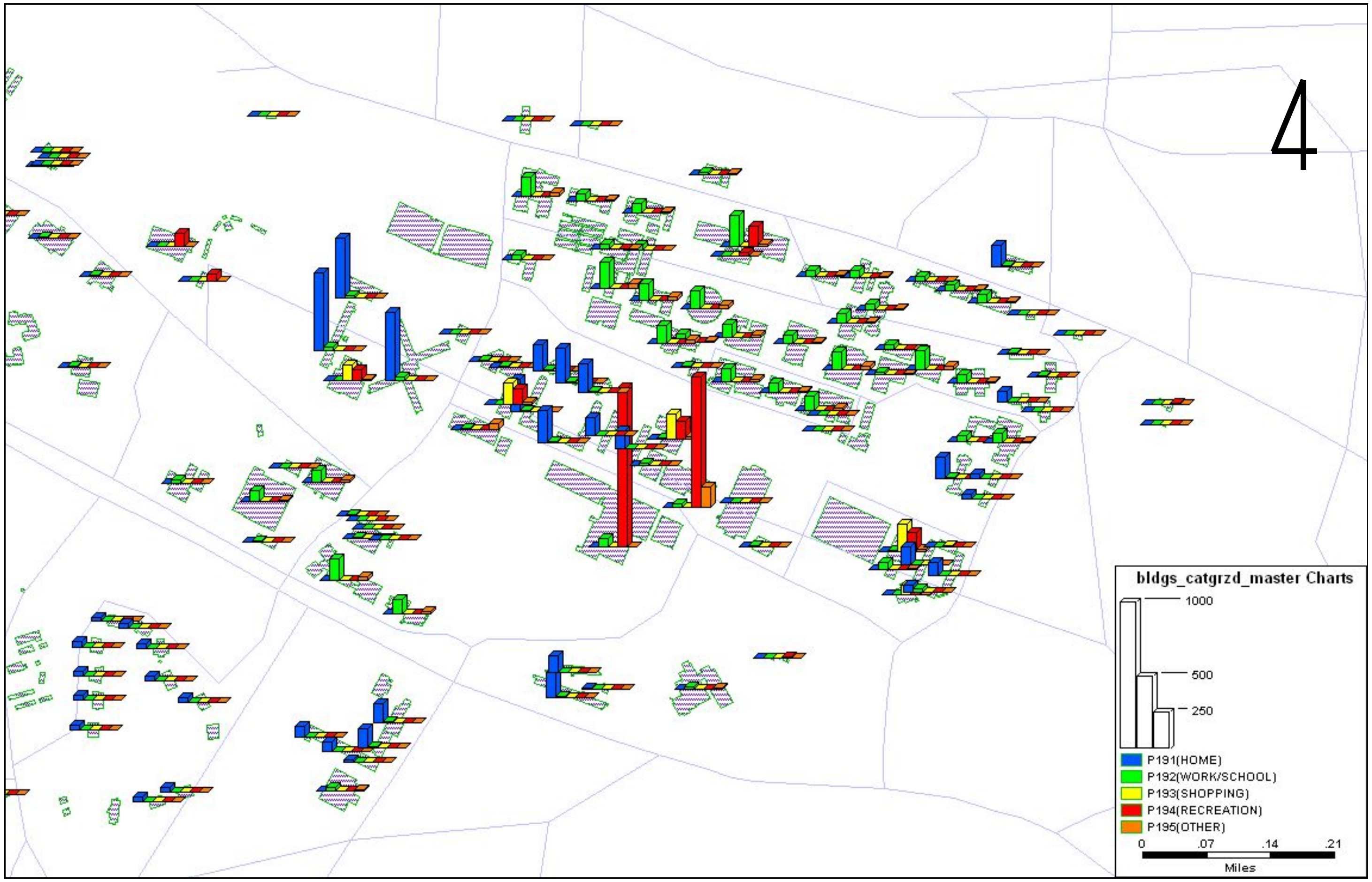


Figure 6.5 Number of Students Present in Each Building (Hour: 7 p.m.-8 p.m.)

6.2.2 Validation of student building presence

To validate the estimated student building presence, student class schedules on a Tuesday in the spring 2001 was obtained from NCSU student registration office and was used as the observed building presence data. Table 6.1 shows a list of the thirty selected NCSU buildings in which the modeled building presence compares to their observations.

Table 6.1 List of Buildings selected for Model Validation

ID	Building Name	Building Abbreviations
6	Jordan Hall	JOR
29	Biltmore Hall	BI
39	Schaub Food Science Building	SFS
41	D.S. Weaver Labs	DSW
47	Harrelson Hall	HA
49	Kilgore Hall	KI
50	Nelson Hall	N
51	Bostian Hall	BOS
52	Williams Hall	WMS
54	David Clark Laboratories	CL
66	Polk Hall	PK
67	Gardner Hall	GA
68	Broughton Hall	BR
70	Mann Hall	MN
71	Riddick Engineering Labs	RD
73	Daniels Hall	DAN
74	Withers Hall	WI
84	Kamphoefner Hall	KAM
85	Brooks Hall	BS
86	Leazar Hall	LEZ
99	Poe Hall	POE
189	Carmichael Gymnasium	CG
201	Textile Building	TEX
213	Winston Hall	WN
214	Caldwell Hall	CAL
215	Tompkins Hall	T
217	Reynolds Coliseum	COL
219	Price Music Center	PMC
222	Dabney Hall	DAB
223	Cox Hall	COX

The registration information includes: course code, number of students registered, class room and building, class time. A total of fifty-four buildings were recorded as classrooms in the spring 2001. As seen in Figure 6.6, thirty buildings are selected for model validation after comparing the registration information between 2001 and 2006. The buildings where the number of students is recorded in 2001 are highly different from the year 2006. During the 2001-2006 construction periods, a limited number of classrooms were available. The low number of students who registered at the buildings under construction would not appropriately reflect the student building presence regarding actual building capacity. The thirty buildings located on the main NCSU campus are seen as used to be in 2001.

To create the observed building presence data, the number of students registered was sorted by the building name and class time. The class time is separated again by hour and minute, a simple Visual Basic program was coded within Microsoft Excel to accumulate the number of students by buildings and hours. The way to accumulate students is that, for example, if 28 students registered for a class from 8:05 a.m. to 9:20 a.m., then the 28 students were counted into both 8 a.m. and 9 a.m. presence. All students registered in the same building are finally accumulated by time of day. It is assumed that the students registered are all attending the class.

Before comparing the estimated building presence against the observed presence, it is important to note that the building presence estimated from the model represents all students either in a class or not, while the observed presence only counts the students in a class.

As seen on Figure 6.6, all buildings are located on Main campus (North) of NCSU and they are the major buildings for ‘Academic/Teaching’ purpose.

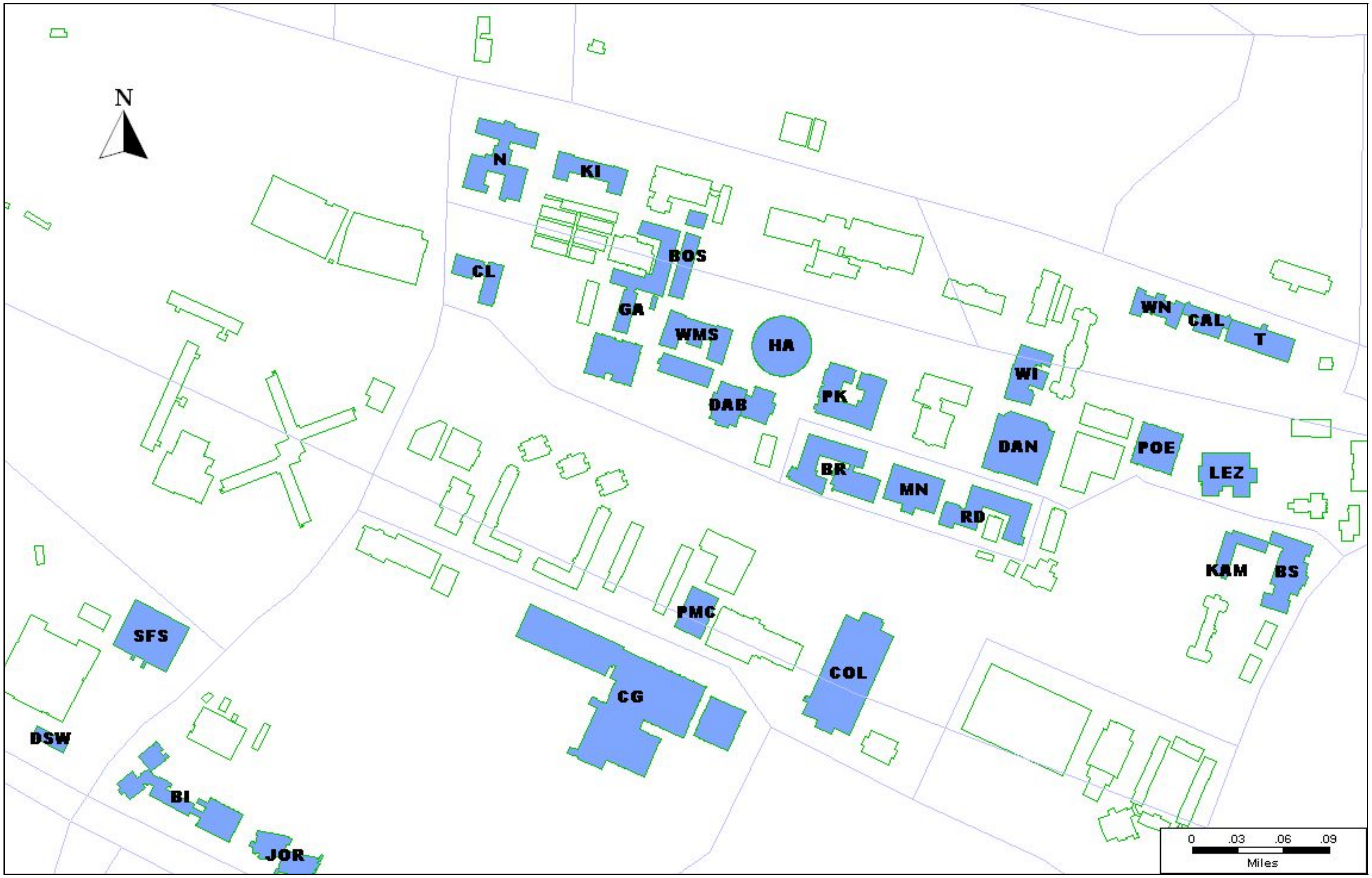


Figure 6.6 30 Buildings Selected for Building Presence Model Validation

Table 6.2 shows the direct comparison between the estimated building presence and the number of students registered in a class by buildings. The overall estimated presence of thirty buildings is seen to be higher than the observed as much as thirty five percent. This is mainly because the model considers not only the students in a classroom, but also the students doing research or other work related activities in the same building. The buildings classified for ‘Academic/Teaching’ consist of classroom, research lab, computer lab, etc. Hence, the students who are present and do not have a class would be doing work related activities at the building during the class hours. The model estimates all students supposed to be present in each building and the output consequently overestimates them. Thus, twenty five buildings out of thirty are overestimated with various ranges represented by the percent deviation that illustrates the variation of estimates regarding the observations.

Figure 6.7 visually shows the difference between estimated student presence in buildings and the number of students recorded. For example, Harrelson Hall has the highest student presence compared to any other building, which corresponds to the observed students registered. Reynolds Coliseum has the lowest student presence estimated and the second lowest from the observed records. As a result, the model seems to successfully replicate the spatial distribution of students in classroom buildings on campus.

The activity capacity in each building based on the destination choice model effectively works in the estimation of building presence. However, the building presence could be calibrated based on better observations, if data were available, for better representation of trip generation, since this is a critical component in estimation of trip generation like trip production and attraction.

Table 6.2 Comparison of Estimated Presence and Number of Students Registered

ID	Building Name	Estimates	Students registered	%Deviation
6	Jordan Hall	1570	1229	27.8%
29	Biltmore Hall	2840	1304	117.7%
39	Schaub Food Science Building	642	234	174.5%
41	D.S. Weaver Labs	1066	1011	5.4%
47	Harrelson Hall	19414	11438	69.7%
49	Kilgore Hall	1320	966	36.7%
50	Nelson Hall	7864	6338	24.1%
51	Bostian Hall	2821	3514	-19.7%
52	Williams Hall	3879	2162	79.4%
54	David Clark Laboratories	1313	553	137.2%
66	Polk Hall	2059	896	129.7%
67	Gardner Hall	2982	1190	150.5%
68	Broughton Hall	4134	2442	69.3%
70	Mann Hall	4217	2157	95.5%
71	Riddick Engineering Labs	5083	3266	55.6%
73	Daniels Hall	4267	4302	-0.8%
74	Withers Hall	3589	5997	-40.1%
84	Kamphoefner Hall	1599	334	378.3%
85	Brooks Hall	1333	814	63.8%
86	Leazar Hall	442	1261	-65.0%
99	Poe Hall	6023	5390	11.7%
189	Carmichael Gymnasium	4929	3782	30.3%
201	Textile Building	1968	1950	0.9%
213	Winston Hall	3991	4503	-11.4%
214	Caldwell Hall	2948	2092	40.9%
215	Tompkins Hall	4616	4236	9.0%
217	Reynolds Coliseum	44	273	-83.9%
219	Price Music Center	144	584	-75.4%
222	Dabney Hall	6439	2865	124.8%
223	Cox Hall	2808	1508	86.2%
Total		106342	78590	**35.3%**

Note: % Deviation is calculated as: $\%Deviation = \frac{(obs_i - est_i)}{obs_i} \times 100$.

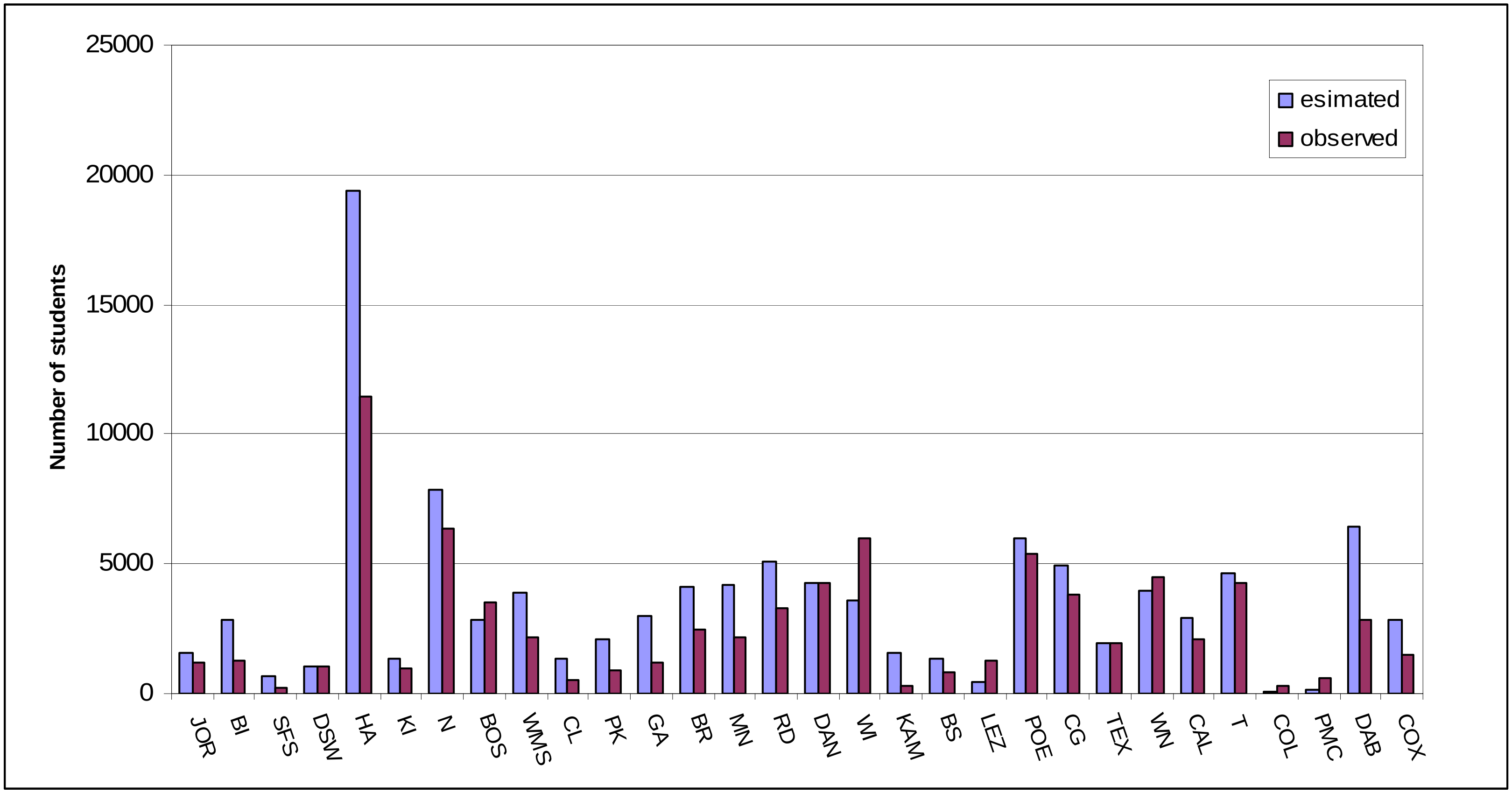


Figure 6.7 Comparison of The Estimated and Observed Student Presence in 30 Buildings During Class Hours (8a.m. – 8p.m.)

(Note: The estimation includes students in class and not in class. The observations count only students in class.)

This building presence estimated from the activity-based model developed herein will be used in the estimation of trip generation by developing the model like a conditional probability of trips. This will be discussed later in detail.

In Figure 6.8 a comparison between hourly modeled and the registered students in a class indicates that the model seems to successfully capture the temporal distribution of students during the class hours. However, a higher number of students were estimated compared to students registered in most hours. Again, this is mainly because the modeled hourly student presence also includes both students in a class and not in a class.

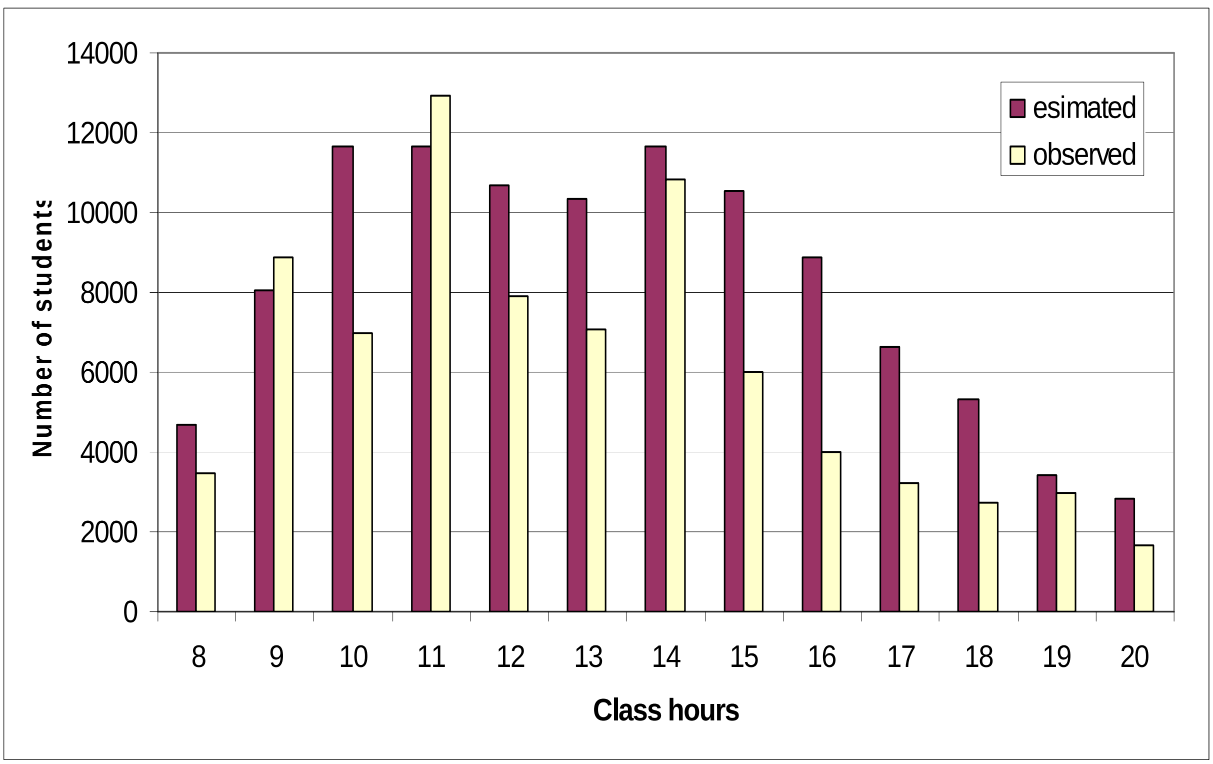


Figure 6.8 Comparison of the Hourly Modeled and Observed Student Presence

In order to see how well the modeled student presence in a building replicates the observations, R^2 is used in the analysis as a measure of effectiveness (MOE) to explain the overall accuracy of the modeled building presence. The equation for R^2 is as shown below:

$$R^2 = \left[\frac{\left(\sum_{I=1}^{n} (obs_i \times est_i) - \frac{(\sum_{I=1}^{n} obs_i \bullet \sum_{I=1}^{n} est_i)}{n} \right)}{\sqrt{\left((\sum_{I=1}^{n} est_i^2 - \frac{\sum_{I=1}^{n} (est_i)^2}{n}) \times (\sum_{I=1}^{n} obs_i^2 - \frac{\sum_{I=1}^{n} (obs_i)^2}{n}) \right)}} \right]^2 \quad (6.2)$$

where

obs_i : observed students in building i
est_i : estimated presence in building i
n : number of buildings in comparison

Using the equation, the R^2 is calculated as shown in Table 6.3. The overall modeled building presence explains approximately 81% of the variation based on thirty buildings during thirteen hour periods (from 8 a.m. to 8 p.m.), and the modeled hourly building presence explains about 67% of the variation. In both cases R^2 seems to be good because the models were not calibrated and the observed value includes only the students in a class.

Table 6.3 Calculated R^2 for 30 Classrooms and 13 class hours

	Number observations	**Modeled R^2**
Overall Classrooms	30	0.8132
Class hours	13	0.6649

6.3 Travel Demand Estimation

Travel demand, the amount of trips in the study area, is obtained from the function of the number of persons present in each building (building presence) multiplied by the conditional probability of trips that is estimated from the MNL model based on the student travel information.

6.3.1 MNL model for conditional probability of trips

The structure of multinomial logit model for the conditional probability of trips is defined in function 6.3:

$$P(Y = k \mid j, g, t) \qquad (6.3)$$

Where,

k : Travel activity (trips)
j : Activity type
g : Traveler group (student status and residential status)
t : Time frame (1-24hr)

This is simply defined as $\pi_k(x) = P(Y = k \mid x)$ at a fixed setting **x** as activity type (j), traveler group (g), and time (t) for explanatory variables, with $\sum_{k=1}^{5} \pi_k^t(x) = 1$. For observations at that setting, the counts (i.e., number of trips) are treated at the ***K*** categories of ***Y*** as multinomial with probabilities $\{\pi_1(x), \cdots, \pi_K(x)\}$. Logit models pair each response category with a baseline category, often the last one or the most common one. Equation 6.4 simultaneously describes the effects of **x** on these ***K-1*** logits.

$$log \frac{\pi_k(x)}{\pi_K(x)} = \alpha_k + \beta_k^{'} x, \quad k = 1, \cdots, K - 1 \qquad (6.4)$$

The effects vary according to the response paired with the baseline. These ***K-1*** equations determine parameters for logit with other pairs of response categories, since:

$$log\frac{\pi_a(x)}{\pi_b(x)} = log\frac{\pi_a(x)}{\pi_K(x)} - log\frac{\pi_b(x)}{\pi_K(x)} \tag{6.5}$$

Maximum-likelihood fitting of multinomial logit models maximizes the likelihood subject to $\{\pi_k(x)\}$ simultaneously satisfying the K-1 equations that specify the model.

In order to use MNL model in estimating the conditional probability of trips for given traveler group, activity, and time, the data need to be an aggregated form by traveler group and time of day so that the MNL model will fit the number of trips observed from the dataset. Table 6.4 shows the format of data constructed from the 'Place' database, which aggregates all trips corresponding to activity type and the hour when trips occurred.

Table 6.4 Observed Number of Trips Data for Fitting MNL Model

Hour	Student status	Residential status	Activity				
			Home	Work/School	Shopping	Recreation	Other
1	Under	On-campus	0	5	1	8	1
		Off-campus	0	2	0	2	0
:	Graduate	On-campus	0	1	0	0	0
		Off-campus	0	1	0	0	0
12	Under	On-campus	76	49	4	36	4
:		Off-campus	26	55	3	21	7
:	Graduate	On-campus	1	3	0	1	0
		Off-campus	11	18	6	10	3
24	Under	On-campus	2	4	2	6	1
		Off-campus	2	2	1	5	0
	Graduate	On-campus	0	2	1	0	0
		Off-campus	0	1	0	0	0

After cleaning the sample of activities, a total of 4883 activities are undertaken by 698 students during the day. Every trip recorded while students are doing a certain activity is reclassified by hour, student status, residential status, and activity type. The SAS LOGISTIC procedure was used in model estimation since this does analysis on nominal responses with relative ease.

6.3.2 Model output

Figure 6.9 shows the output of the overall goodness-of-fit of MNL model for the conditional probability of trips. The overall model goodness-of-fit is found to be good based on the statistics of 'Deviance' and 'Pearson' goodness-of-fit which illustrates that the statistics are statistically significant at 95% confidence level. Also, the global hypothesis that the parameters estimated are equal to zero (i.e., not meaningful) is rejected and the variables used in the analysis are found to be significant at 95% confidence level. The parameters estimated are shown in Appendix G.

Figures 6.10 through Figure 6.14 show the fitted number of conditional trips and observed trips for the activities of 'Home', 'Work/School', 'Shopping', 'Recreation', and 'Other'. Overall, the conditional trips for every activity are seen to be fitted well over the observed trips in most hours during a day in every student group. However, the fitted conditional trips for 'Shopping' activity do not seem to be good compared to those of other activities due to lack of observed trips. As expected, Shopping is not a major students' activity on-campus since a university usually does not have a large shopping facility on-campus other than a bookstore. So, the trips for shopping would be rare. Among student groups, the conditional trips for the group of graduate-on-campus with a few trips observed have low R^2 at all activity types.

```
                     Model Convergence Status

          Convergence criterion (GCONV=1E-8) satisfied.


         Deviance and Pearson Goodness-of-Fit Statistics

Criterion          Value       DF      Value/DF     Pr > ChiSq

Deviance        345.0761      252        1.3693         <.0001
Pearson         332.3785      252        1.3190         0.0005

                Number of unique profiles: 89


                      Model Fit Statistics

                                          Intercept
                            Intercept           and
            Criterion            Only     Covariates

            AIC             14228.356      13089.913
            SC              14254.596      13772.159
            -2 Log L        14220.356      12881.913


             Testing Global Null Hypothesis: BETA=0

   Test                 Chi-Square       DF     Pr > ChiSq

   Likelihood Ratio      1338.4434      100         <.0001
   Score                 1350.8347      100         <.0001
   Wald                  1101.9435      100         <.0001



                         The SAS System
                     The LOGISTIC Procedure

                   Type 3 Analysis of Effects

                                     Wald
         Effect        DF     Chi-Square    Pr > ChiSq

         HOUR          92       915.0066        <.0001
         STATUS         4        29.4364        <.0001
         RESIDENT       4       122.7708        <.0001
```

Figure 6.9 Model Goodness-Of-Fit Results of the Conditional Probability of Trips

The conditional trips to a 'Home' activity is seen to be fitted well over the observed travel in most hours during a day in every student group except the graduate-on-campus group as shown in Figure 6.10. The R^2 for undergraduate-on-campus and off-campus, graduate-on-campus and off-campus are 0.97, 0.96, 0.66, and 0.92, respectively. The conditional trips to a 'Home' activity generates two peaks around 10 a.m. and 6 .p.m. which can be explained by the trips from home to work/class in the morning and by the trips from home to other activities such as Shopping and Recreation.

Figure 6.11 shows the conditional trips to 'Work/School' activities. The R^2 for undergraduate-on-campus and off-campus, graduate-on-campus and off-campus are 0.98, 0.98, 0.63, and 0.86, respectively. These R^2 are similar to the R^2 values of the conditional trips to Home activities but the peaks are seen to be different. The highest peak is observed around noon and it is understood that many students make a trip for lunch from their work places or classes. Several peak hours are observed from the undergraduate students living either on-campus or off-campus between noon and 6 p.m. and the peaks get smaller as hours pass. This can be explained by undergraduate students having three or more classes in a day and making a trip from a classroom. Graduate students, on the other hand, have two travel peaks from their work places around at noon for lunch and 5 p.m. for returning home.

Figure 6.12 shows the conditional trips for the 'Shopping' activity. Due to the lack of trips observed at shopping places, most of the R^2 values are low, about 60%. The peak trips are observed in the afternoon between classes or after school and work.

The conditional trips for the 'Recreation' activity are shown in Figure 6.13. The R^2 values for undergraduate-on-campus and off-campus, graduate-on-campus and off-campus are 0.98, 0.91, 0.61, and 0.69, respectively.

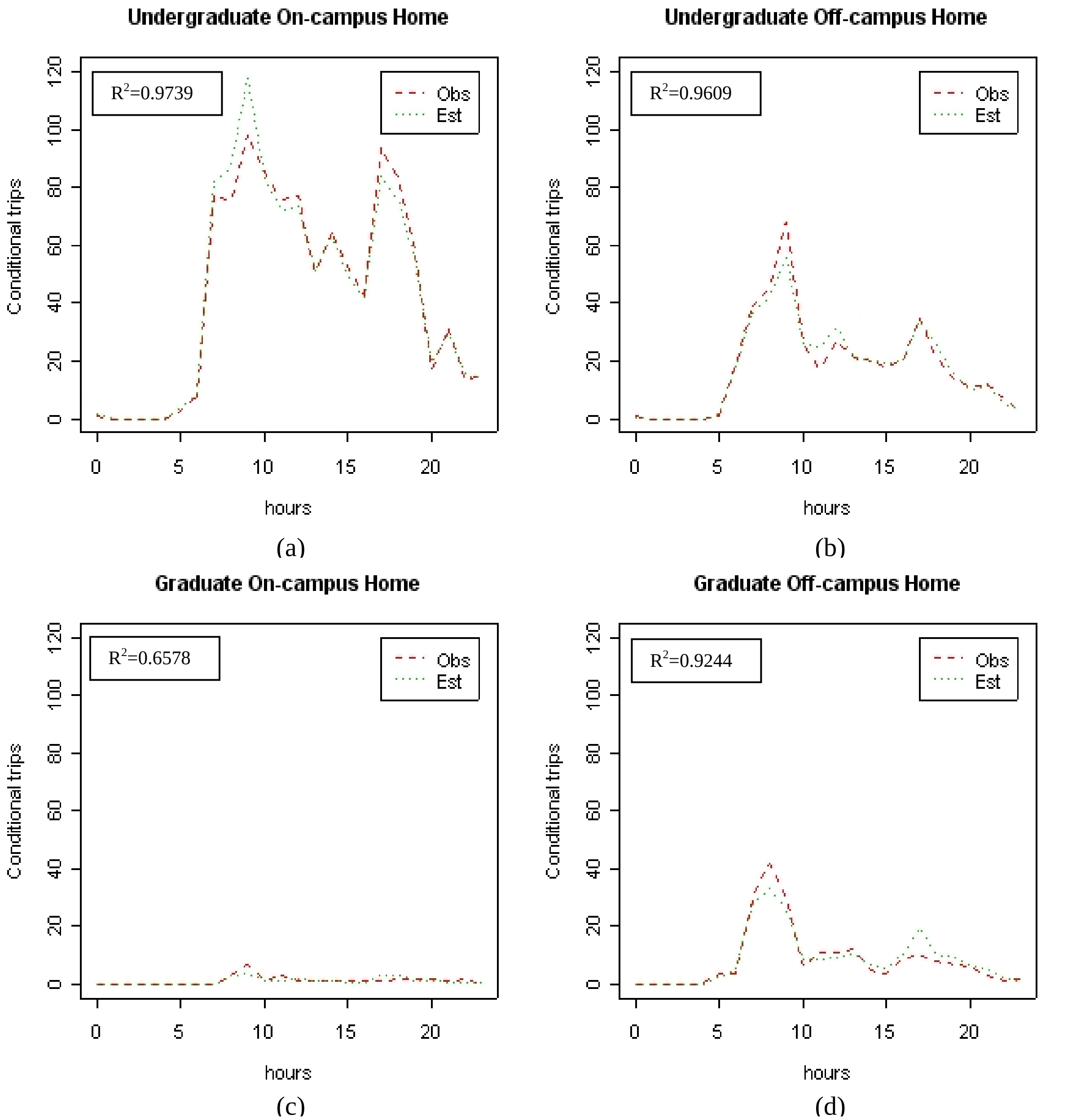


Figure 6.10 Observed and Estimated Number of Conditional Trips for Home Activity: a) Undergraduate and On-campus, b) Undergraduate and Off-campus, c) Graduate and On-campus, d) Graduate and Off-campus Students

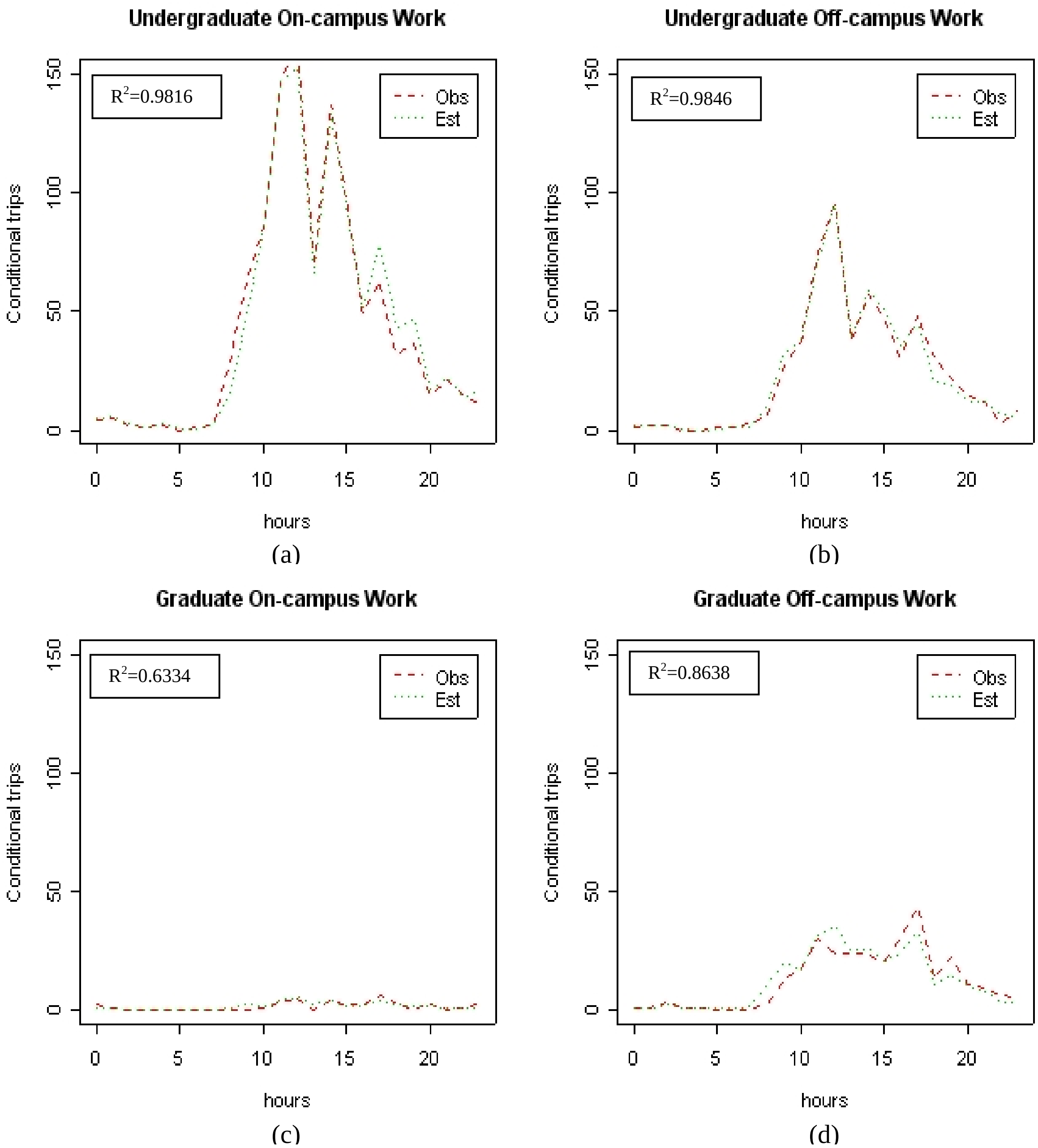
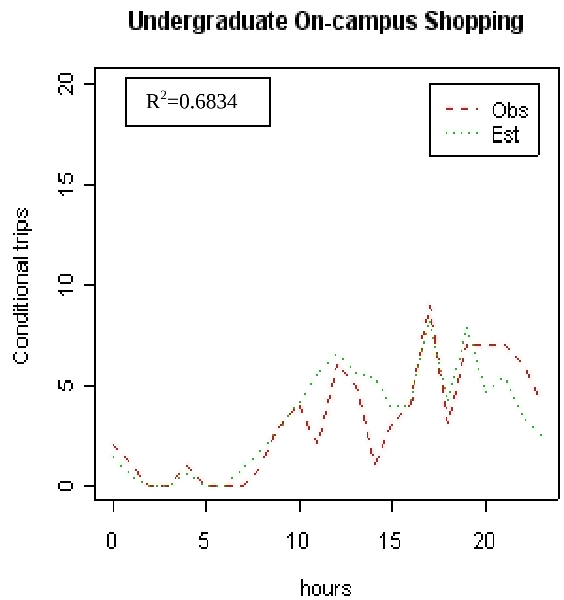
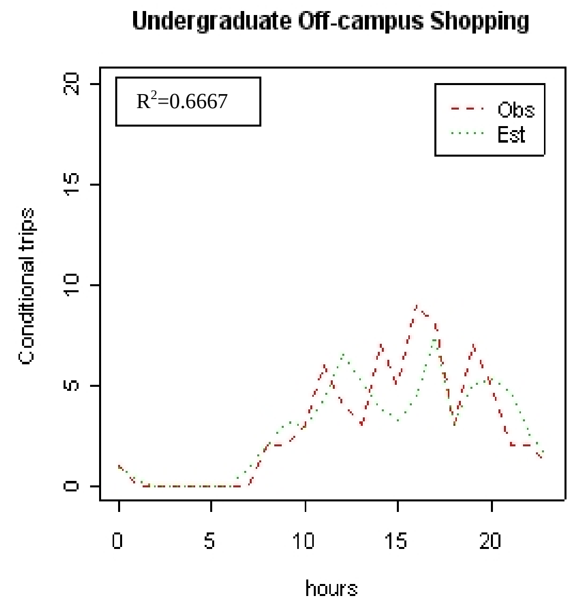


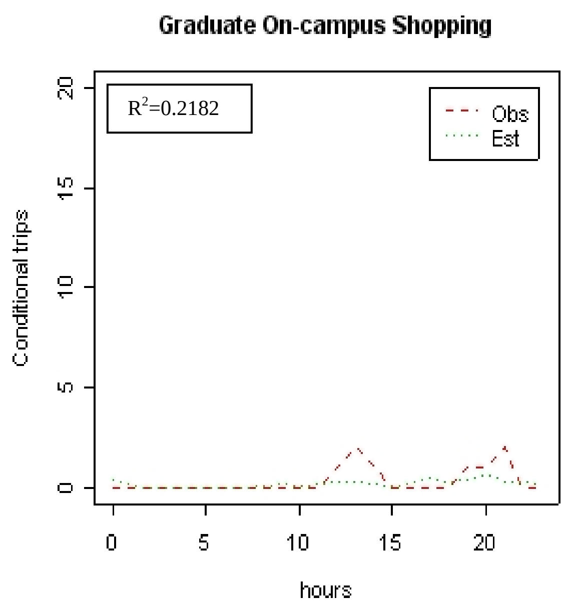
Figure 6.11 Observed and Estimated Number of Conditional Trips for Work/School Activity: a) Undergraduate and On-campus, b) Undergraduate and Off-campus, c) Graduate and On-campus, d) Graduate and Off-campus Students



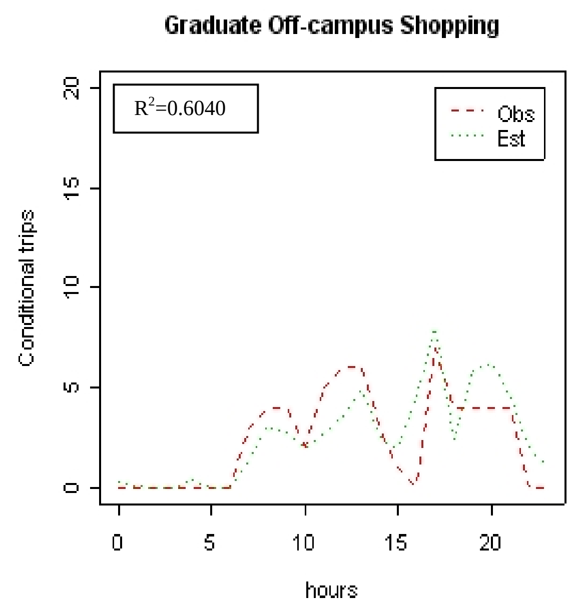
(a)



(b)



(c)



(d)

Figure 6.12 Observed and Estimated Number of Conditional Trips for Shopping Activity: a) Undergraduate and On-campus, b) Undergraduate and Off-campus, c) Graduate and On-campus, d) Graduate and Off-campus Students

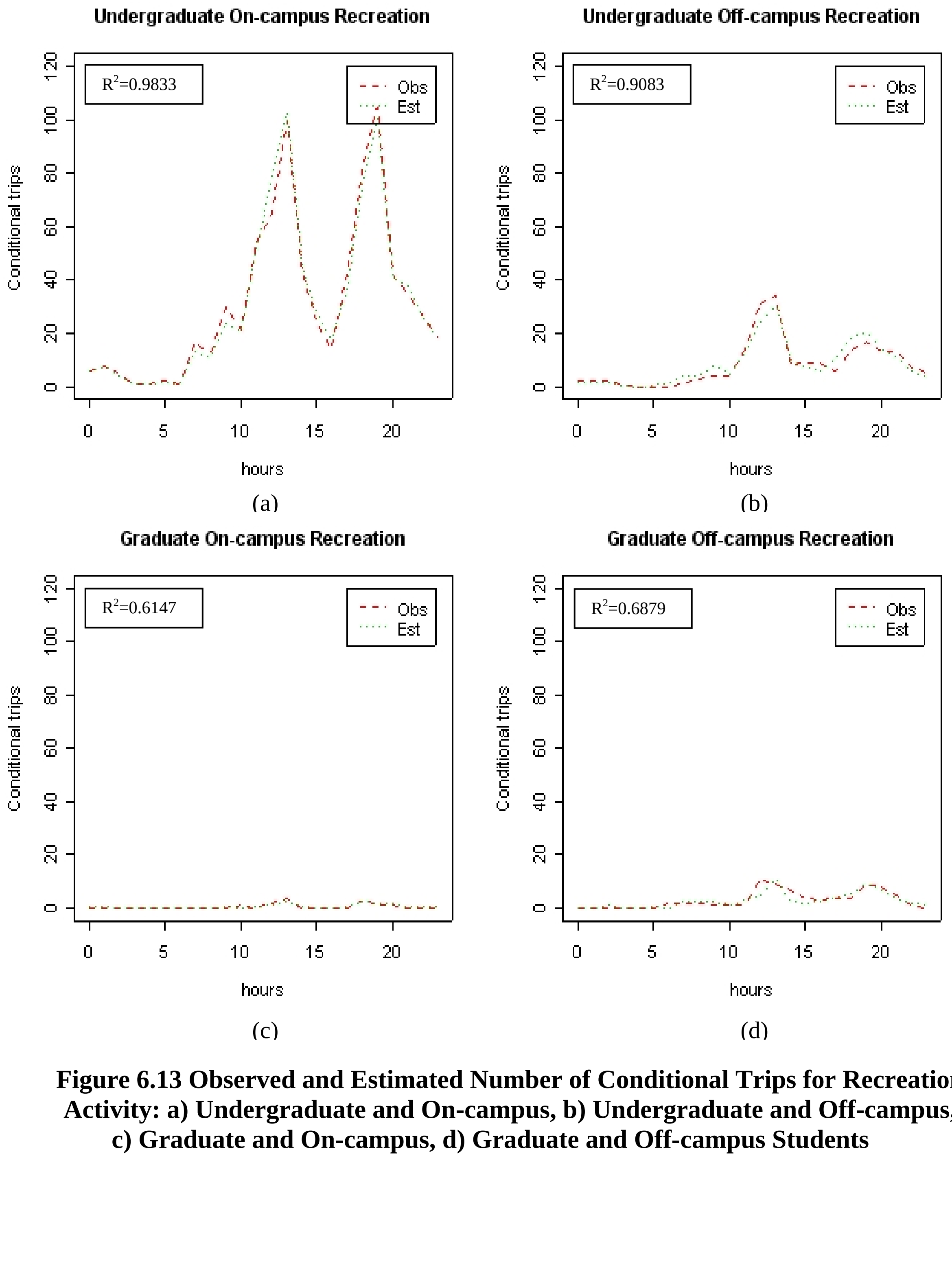


Figure 6.13 Observed and Estimated Number of Conditional Trips for Recreation Activity: a) Undergraduate and On-campus, b) Undergraduate and Off-campus, c) Graduate and On-campus, d) Graduate and Off-campus Students

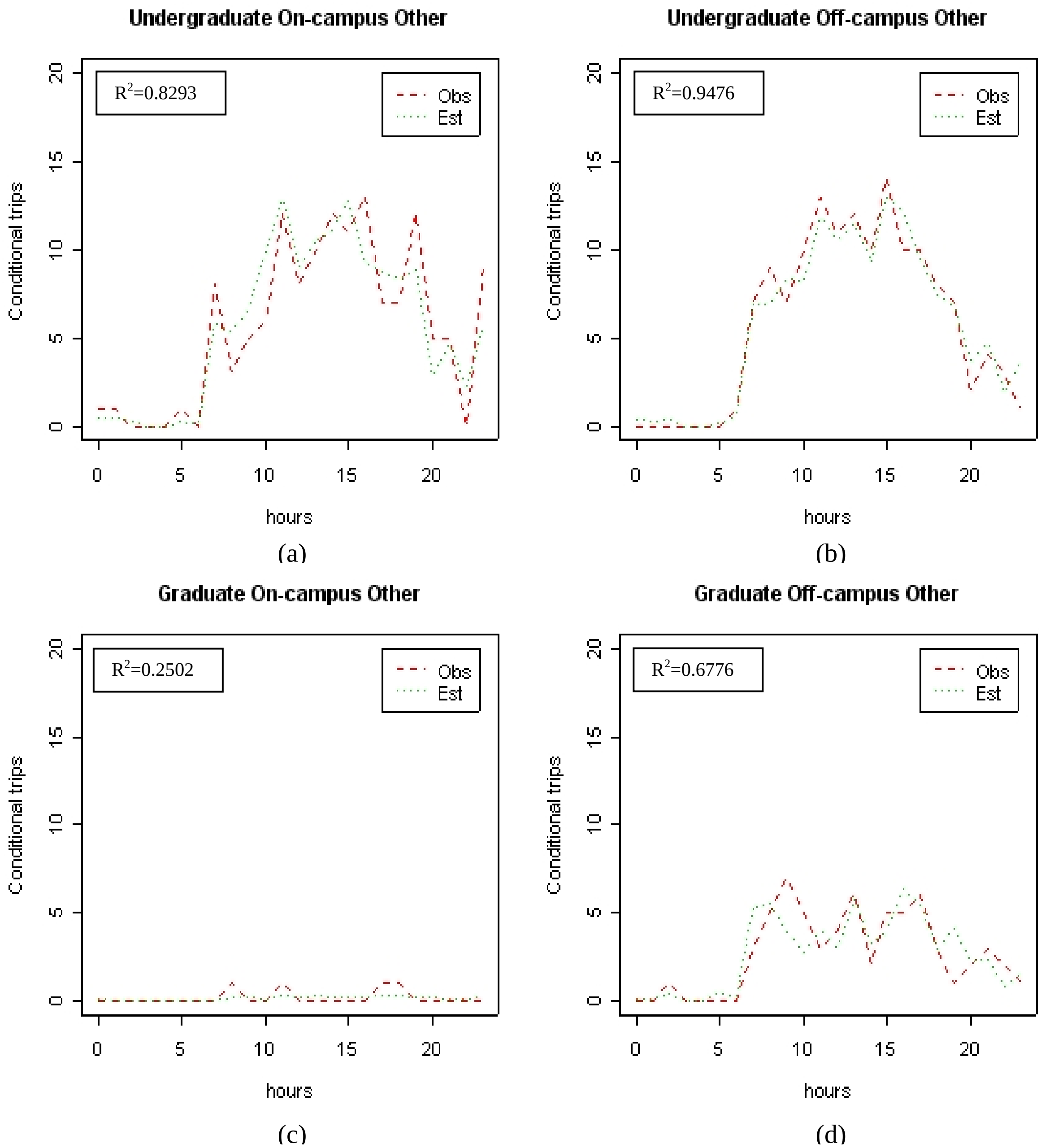


Figure 6.14 Observed and Estimated Number of Conditional Trips for Other Activity: a) Undergraduate and On-campus, b) Undergraduate and Off-campus, c) Graduate and On-campus, d) Graduate and Off-campus Students

The conditional trips for a ‘Recreation’ activity have two peak hours around noon and after 6 .p.m. because more students have ‘Recreation’ activities during lunch time and after school.

Figure 6.14 shows the conditional trips to ‘Other’ activity. The R^2 values for undergraduate-on-campus and off-campus, graduate-on-campus and off-campus are 0.83, 0.95, 0.25, and 0.68, respectively. Due to the lack of trips observed at ‘Other’ activities from graduate on-campus students, the R^2 is a relatively low as 25% as compared to other student groups. Unlike ‘Home’ and ‘Work/School’ activities, it is found that no specific peak hours are observed and travel for ‘Other’ activities happen all day long.

In summary, the MNL model fitted the observed conditional trips well but, for the better estimates, acquiring more observations is important.

6.3.3 Calculation of trip production and attraction

The final step for estimating trip generation is the calculation of trip production and attraction. The building presence and the conditional probability of trips estimated in the previous sections are used in a function of trip production and attraction as simple as shown in function 6.6 and 6.7. Trip production is calculated as all building presence (i) times the conditional probability of trips at a given time (t), given activity type (j), and given traveler group (g). The function of trip attraction is the same as the function of trip production except for the time (t-1). The assumption defined here is that some of the students in a current building at time (t) came from other buildings. So, the trip attraction at building (i) at time (t) is a matter of how to calculate the number of students actually moving from other buildings among existing students present at building (i). This is obtained by applying the conditional probability of trips at time (t-1) to the building presence at time (t) as shown in function 6.7:

$$P_{i,g}(t) = \left(\sum_{j} Z_{i,j,g}(t) \bullet P(trip \mid j, g, t) \right) \tag{6.6}$$

$$A_{i,g}(t) = \left(\sum_{j} Z_{i,j,g}(t) \bullet P(trip \mid j, g, t-1) \right) \tag{6.7}$$

Where:

$P_{i,g}(t)$: Trip productions of traveler group *g* at time *t* in building *i*

$A_{i,g}(t)$: Trip attractions of traveler group *g* at time *t* in building *i*

The estimated trip productions and attractions provide all student trips generated and attracted at every building on-campus by five activity types and for twenty four hours. The estimated trip production and attraction are compared to the observed student trips in sampled buildings for model validation in the next section.

6.4 Model Validation

To validate the estimated trip production and attraction at the building level, actual student trips were collected by counting students entering and exiting sampled buildings on the NCSU campus. The survey sheet for data collection is in Appendix G. Total thirty-two volunteers counted students entering and exiting buildings for one-hour periods between 9 a.m. and 6 p.m. either on Tuesday or Thursday. It is assumed that the student travel pattern on Tuesday is similar to that of Thursday because the class schedule is usually identical for those two days. The limitation in data collection is that during any data collection period an entire building needs to be covered (i.e. all entrances and exits). Every building has more than two entrances and exits. So, three or more volunteers were generally required to cover one building. Therefore, only fourteen buildings including twelve academic buildings, the library, and the Tally Student Center were sampled due to limited number of volunteers and time (Figure 6.15).

Table 6.5 presents the actual counts of students entering (in) and exiting (out) at fourteen buildings during one-hour periods and the estimated trip production (out) and attraction (in) for that time period. Overall, the modeled production and attraction seem to be fine with respect to two facts: (1) model estimates are not calibrated; (2) the student counts were collected in 2006 while the model output represents the year 2001. Although the year for data collection and for modeling base year are different, it is expected that there would be no large changes in students' travel patterns and the university facilities particularly the fourteen buildings between the two years. According to Table 6.5, the overall model production of fourteen buildings is slightly greater than the actual counts by 2.2% and the attraction is 13.4% greater than the counts. The total trips, sum of productions and attractions, are about eight percent greater than the total counts. The percent deviations at each building vary from – 48.2% to 267.0%, and are seen to be high at the buildings with low trips counted. Specifically, Broughton Hall, Mann Hall, and Patterson Hall have great variability in the percent deviation due to their low trips counted. This result suggests that the activity-based model developed in this study successfully captures the base year trips at both the aggregated (i.e., total trips) and the disaggregated level (i.e., individual building trips).

Figure 6.15 Sample Buildings and Entrances/Exits for Student Trip Counts

Table 6.5 Student Trip Comparisons between Observations and Estimates (Activity-based approach)

Building	Hour		Counts			Estimates			%Deviation		
	Start	End	In	Out	Total	In	Out	Total	In	Out	Total
Library	15:00	16:00	306	264	570	275	325	600	-10.1%	23.1%	5.3%
Talley Std Center	10:00	11:00	320	269	589	264	317	581	-17.5%	17.8%	-1.4%
Broughton Hall	15:00	16:00	51	52	103	120	142	262	136.0%	172.9%	154.7%
Winston Hall	13:00	14:00	179	156	336	130	118	248	-27.5%	-24.5%	-26.1%
Caldwell Hall	16:00	17:00	49	43	92	61	81	142	23.5%	88.1%	53.6%
Tomkins Hall	16:00	17:00	112	98	210	91	120	211	-18.8%	22.8%	0.6%
Mann Hall	13:00	14:00	31	84	115	113	109	222	267.0%	30.0%	93.7%
Patterson Hall	9:00	10:00	38	28	67	13	10	23	-66.2%	-64.6%	-65.5%
Poe Hall	11:00	12:00	320	331	652	224	240	464	-30.1%	-27.6%	-28.8%
Dabney	12:00	13:00	122	180	302	233	237	470	91.1%	31.3%	55.4%
Cox	12:00	13:00	60	89	149	103	107	210	71.3%	20.3%	40.8%
Polk Hall	11:00	12:00	93	83	176	81	85	166	-12.9%	2.4%	-5.7%
Daniels Hall	16:00	17:00	61	65	126	81	106	187	33.1%	62.5%	48.3%
Park shop	9:00	10:00	30	33	63	23	17	40	-22.6%	-48.2%	-36.0%
Total			1773	1776	3549	1812	2014	3826	**2.2%**	**13.4%**	**7.8%**

Note: % Deviation is calculated as: $\%Deviation = \frac{(obs_i - est_i)}{obs_i} \times 100$.

6.5 Comparisons of Activity-based and Traditional Approaches

6.5.1 Overview

As noted in the literature review, an interest of model comparison between conventional travel demand models and activity-based models increases to planning agencies in a decision of choosing an better approach. The research question herein is about whether the activity-based model is better than a traditional four-step travel demand model. Despite the theoretically superior basis of activity-based travel demand models, the lack of a detailed validation and assessment of this model system hinders their widespread adoption and application.

Many MPOs strongly ascertain that a detailed validation and assessment of activity-based models are necessary before they would consider implementing new approaches in their regions. MPOs may be motivated to adopt the new model if activity-based models are found to be better than traditional four-step travel demand models. However, there are no easy answers to the important question as: “what constitutes a “better” model?” This question has been raised and debated recently from MPOs. Pendyala and Bhat (2005) mentioned in their white paper that “if the definition of ‘better’ model is only related to higher standard of validation with the fewer adjustments to model components and parameters, then the activity-based model is likely to be better” [51]. However, there is no quantitative comparison between the trips estimated from activity-based approach and conventional approach especially for a university special generator.

This research addresses the questions of “better” model by limiting the scope to which model output is closer to actual counts.

6.5.2 Trip generation from a conventional model

In order to fairly compare the trip generation from a conventional approach to one from an activity-based approach, the conventional trip generation is developed by using the same student survey data and identical spatial resolution that are used in the activity-based approach (i.e., NCSU buildings). The activity travel survey is not a special survey anymore since most travel surveys include a questionnaire about daily activities trips. This does not require extra survey cost [1, 5, 10, 17]. The use of disaggregated spatial resolution (at the building level) in developing a conventional trip generation model enables comparison of its trip generation to the output of the activity-based approach.

The first step in a conventional approach for trip generation is determinating trip purposes to be used for the model development. From an analysis of the trips made by all NCSU students, the four trip purposes are defined as follows [52]:

- Home Based Class (HBC) – these are trips made from the student's home to class on campus including trips for other academic purposes, such as studying on campus.
- Home Based Other (HBO) – these include all other home-based trips, including shopping and recreation.
- Non-Home Based Class (NHBC) – these are trips associated with academic purposes, such as class-to-class or class-to-studying where neither end of the trip is at home.
- Non-Home Based Other (NHBO) – these include all other non-home based trips, including trips where one, but not both, ends is academic in nature, such as a trip from class to eat lunch.

All these trip purposes need to be recorded for every trip in the survey data. Then, the trip rates for trip production and attraction have to be estimated by trip purpose and facility type (building use). Since there was no statistical significance in the values obtained for any building type, student trip rates are developed based on four building types: 1) Administrative; 2) Other student space; 3) Academic; 4) University beds [52]. The trip rates developed for each category by a consultant are summarized in Table 6.6.

Table 6.6 Trip Production and Attraction Rates by Trip Purpose and Facility Type

Classification		HBC	HBO	NHBC	NHBO
Production	**Administrative**	--	--	6.32	19.17
	Other Student Facility	--	--	1.0	8.25
	Academic	--	--	1.84	6.51
	Beds	1.00	1.00	--	1.00
Attraction	**Administrative**	4.46	--	4.79	10.77
	Other Student Facility	1.00	8.52	1.00	5.60
	Academic	1.48	1.00	2.03	4.33
	Beds	--	2.88	--	1.00

Note: Trip rate represents the number of trips per day per 1,000 square feet of each facility.
Source: MAB (2004) [52].

As shown in Table 6.6, the highest student trip production rate is 19.17 (trips per day/1,000 square feet) of Administrative facility for NHBO, and the attraction rate at the Administrative facility is also the highest as 10.77. The NHBO trip production rate for Other Student Facility and Academic facility are also at high 8.25 and 6.51, respectively. The trip NHBC production rate for Administrative facility is 6.32, which suggests that students generate many trips from Administrative facilities to Academic facilities to attend class.

Conversely, the trip rate from one class to another is relatively low as 1.84 (trips per day/1,000 square feet). In trip attraction, Administrative facilities may attract four and one half times as many student HBC trips as Other Student Facility. The high HBO trips (8.52 trips per day /1000 ft^2) are seen to be attracted at Other Student Facility. The high HBO rate is almost twice as high as HBC trips attracted at Administrative facility. Only 1.48 trips per day /1000 ft^2 are attracted at an Academic facility. This suggests that the place where a majority of students go first from their home or dorm is not a classroom. They seem to do some other activities first rather than directly go to class when they arrive at campus. For instance, some students might do homework at library and then go to class, or do research and then go to class.

Using the trip rates developed, the student trip production and attraction can be calculated by trip purposes and buildings. Then, the hourly trip production and attraction are obtained by applying the hourly ratio of trip distribution of departure and arrival time during a day. In order to create the hourly ratio of trip distribution of departure and arrival time, all student trips are separately accumulated by departure and arrival time by hour and the hourly ratio is simply obtained by dividing hourly trips by total departure trips and arrival trips. Table 6.7 and Figure 6.16 show the distribution of daily trips by departure and arrival times.

It is interesting to note that the distribution of students' departure times is totally different from the one for typical households. The distribution of departure time shows that there are several peak hours during a day with the highest at 12 p.m. Unlike students' departure times, the most standard household travel survey shows two peak hours during a day - one in the morning (7 a.m. - 8 a.m.) and one in the afternoon (4 p.m. - 7 p.m.).

Table 6.7 Ratio of Departure and Arrival Time

Time interval	Departures	Arrivals	%Departure	%Arrival
1	21	24	0.39%	0.44%
2	17	15	0.32%	0.28%
3	5	6	0.09%	0.11%
4	6	9	0.11%	0.17%
5	16	10	0.30%	0.18%
6	45	26	0.83%	0.48%
7	205	152	3.80%	2.81%
8	252	250	4.67%	4.62%
9	374	391	6.93%	7.23%
10	325	317	6.02%	5.86%
11	481	**504**	8.91%	9.32%
12	**556**	399	10.30%	7.38%
13	415	455	7.69%	8.42%
14	415	430	7.69%	7.95%
15	334	278	6.19%	5.14%
16	267	312	4.95%	5.77%
17	401	378	7.43%	6.99%
18	332	354	6.15%	6.55%
19	356	360	6.60%	6.66%
20	176	195	3.26%	3.61%
21	171	184	3.17%	3.40%
22	98	112	1.82%	2.07%
23	87	90	1.61%	1.66%
24	41	155	0.76%	2.87%
total	**5396**	**5406**	**100.0%**	**100.0%**

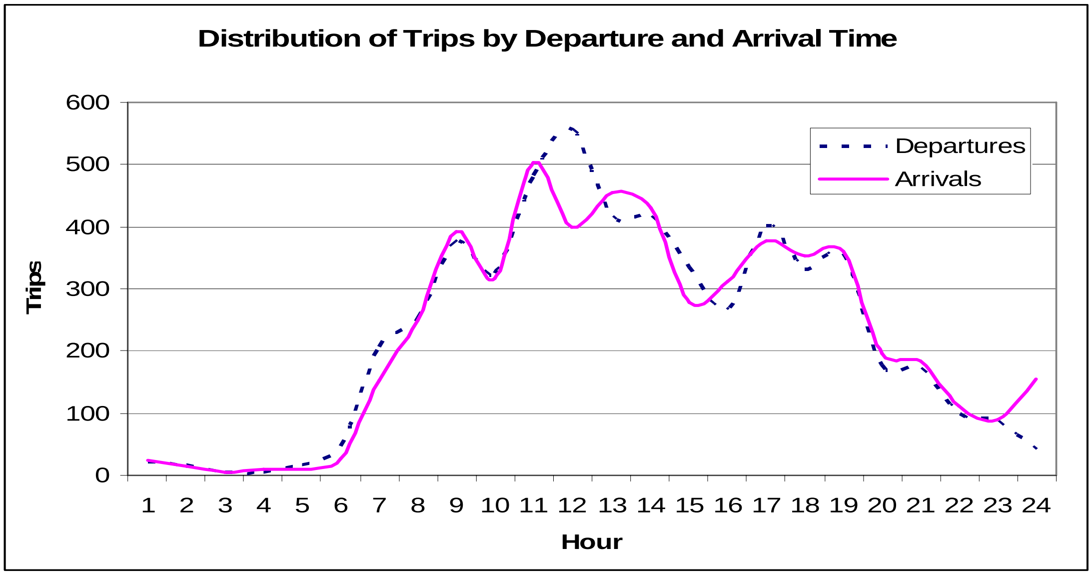


Figure 6.16 Distribution of Trips by Departure and Arrival Time

Home-Based-Work trips usually create those two peak hours in a standard household because HBW trips make up a majority of daily trips. Further, the morning peak is higher than the afternoon peak because the trips from work to home spread out over several hours in the afternoon. As a result, the students' departure time distribution clearly reveals that student trip making behavior needs to be considered separately from the standard household travel demand model. The distribution of arrival time is seen to follow the distribution of departure time, but the highest peak hour is 11 a.m.

Based on the trip rates and the ratio of distribution of departure and arrival time, the hourly trip production and attraction are estimated and summarized in Table 6.8. Overall, the modeled production and attraction are underestimated compared to observed counts. The total modeled production and attraction rates of fourteen buildings using the conventional trip rate approach are almost fifty percent below the observed counts. This clearly illustrates that the disaggregated activity-based approach is much better in terms of model accuracy.

The percent deviations at each building vary from – 86.6% to 202.8%, and are seen to be highest at the buildings with low trips observed. As a result, a conventional approach for a disaggregated travel demand model would not provide reliable trip generation. Although the trip rates are developed by trip purpose and facility type, the trip rates do not adequately replicate hourly student travel at the building level. This is because of two weaknesses in a conventional approach: firstly, the conventional approach uses trip rates to estimate trip generation based on surveyed trips associated with each facility. So, if the facility types are classified more in detail, the output would be better. However, this requires more survey data and time in producing trip rates. Secondly, the trip rate does not differentiate the attractiveness of one building from another.

Table 6.8 Students Trip Comparison between Observations and Estimates (Conventional Approach)

Building	Hour		Counts			Estimates			%Deviation		
	Start	End	In	Out	Total	In	Out	Total	In	Out	Total
Library	15:00	16:00	306	264	570	68	99	167	-77.7%	-62.6%	-70.7%
Talley Std Center	10:00	11:00	320	269	589	85	143	228	-73.6%	-46.7%	-61.3%
Broughton Hall	15:00	16:00	51	52	103	52	45	97	1.4%	-12.8%	-5.8%
Winston Hall	13:00	14:00	179	156	336	24	28	52	-86.6%	-82.1%	-84.5%
Caldwell Hall	16:00	17:00	49	43	92	18	22	40	-63.6%	-48.9%	-56.7%
Tomkins Hall	16:00	17:00	112	98	210	27	34	61	-75.9%	-65.2%	-70.9%
Mann Hall	13:00	14:00	31	84	115	50	58	108	62.4%	-30.8%	-5.8%
Patterson Hall	9:00	10:00	38	28	67	25	28	53	-35.0%	-0.8%	-20.5%
Poe Hall	11:00	12:00	320	331	652	111	123	234	-65.3%	-62.9%	-64.1%
Dabney	12:00	13:00	122	180	302	116	88	204	-4.9%	-51.2%	-32.5%
Cox	12:00	13:00	60	89	149	56	42	98	-6.9%	-52.8%	-34.3%
Polk Hall	11:00	12:00	93	83	176	77	85	162	-17.2%	2.4%	-8.0%
Daniels Hall	16:00	17:00	61	65	126	56	69	125	-8.0%	5.8%	-0.9%
Park shop	9:00	10:00	30	33	63	90	73	163	202.8%	122.6%	160.7%
Total			1773	1776	3549	855	937	1792	**-51.8%**	**-47.2%**	**-49.5%**

Note: % Deviation is calculated as: $\%Deviation = \frac{(obs_i - est_i)}{obs_i} \times 100$.

If two buildings fall into the same category of Academic facility, then the same trip rate is applied to estimate trips. The only different factor that affects trip generation is the square footage of a building. For example, the same trip rates and ratio of distribution of departure and arrival time for 4 p.m. are applied to Daniels and Tomkins Hall. Because Daniels Hall has a greater square footage than Tomkins Hall, the estimated trips of Daniels Hall are greater although the observed counts of Daniels Hall were lower.

6.5.3 Quantitative comparison of model accuracy between the two approaches

In the previous section, the trips estimated from a conventional approach were compared to actual counts at individual building level. This section addresses the quantitative comparison of model accuracy between two approaches based on two aggregated measures of effectiveness - R^2 and root mean square error (RMSE). Figure 6.17 also visually shows the trips estimated from both the activity-based and conventional approaches against the observed student counts by buildings on the main (north) NCSU campus. From the R^2 shown in Table 6.9, the overall modeled trips from the activity-based approach explains approximately seventy-nine percent of the variation for fourteen buildings for any one-hour period while the trips estimated from the conventional approach explains about forty-four percent of the variation.

Table 6.9 Model Validation using R^2 by Model Approach

Approach	Number observations	Modeled R^2
Activity-based Model	14	0.7915
Conventional Model	14	0.4443

Table 6.10 shows the root mean square error (RMSE) which is a measure of the relative error of the estimated trips as compared to the student counts. The %RMSE is commonly applied to volume (trips) groups rather than all volumes as a whole. As with the volume groupings, the %RMSE is usually lower for higher volume groups and higher for lower volume groups. Due to limited number of count data, the counts are regrouped into three categories: 1) low (less than 100 trips), 2) middle (greater than 100 and less than 200), and 3) high (over 200).

Table 6.10 Model Validation using %RMSE by Volume Group

Volume group	No. of counts	%RMSE	
		Activity-based Approach	Conventional Approach
<100	3	54.5	89.2
100 – 200	5	70.4	20.0
>200	6	24.7	70.1
Total	14	36.7	81.4

Note: The function of %RMSE is shown below:

$$\% RMSE = \frac{\sqrt{\frac{\sum_{I=1}^{n} (obs_i - est_i)^2}{n}}}{\frac{\sum_{I=1}^{n} (obs_i)}{n}}$$

Table 6.10 illustrates that the activity-based approach shows lower %RMSE at a higher volume group while the conventional approach shows much higher %RMSE at a higher volume group. The %RMSE from the conventional approach seems to be higher than the activity-based approach although the errors are low at the middle volume group (i.e., 100-200). According to two measures, it is concluded that the trips estimated from an activity-

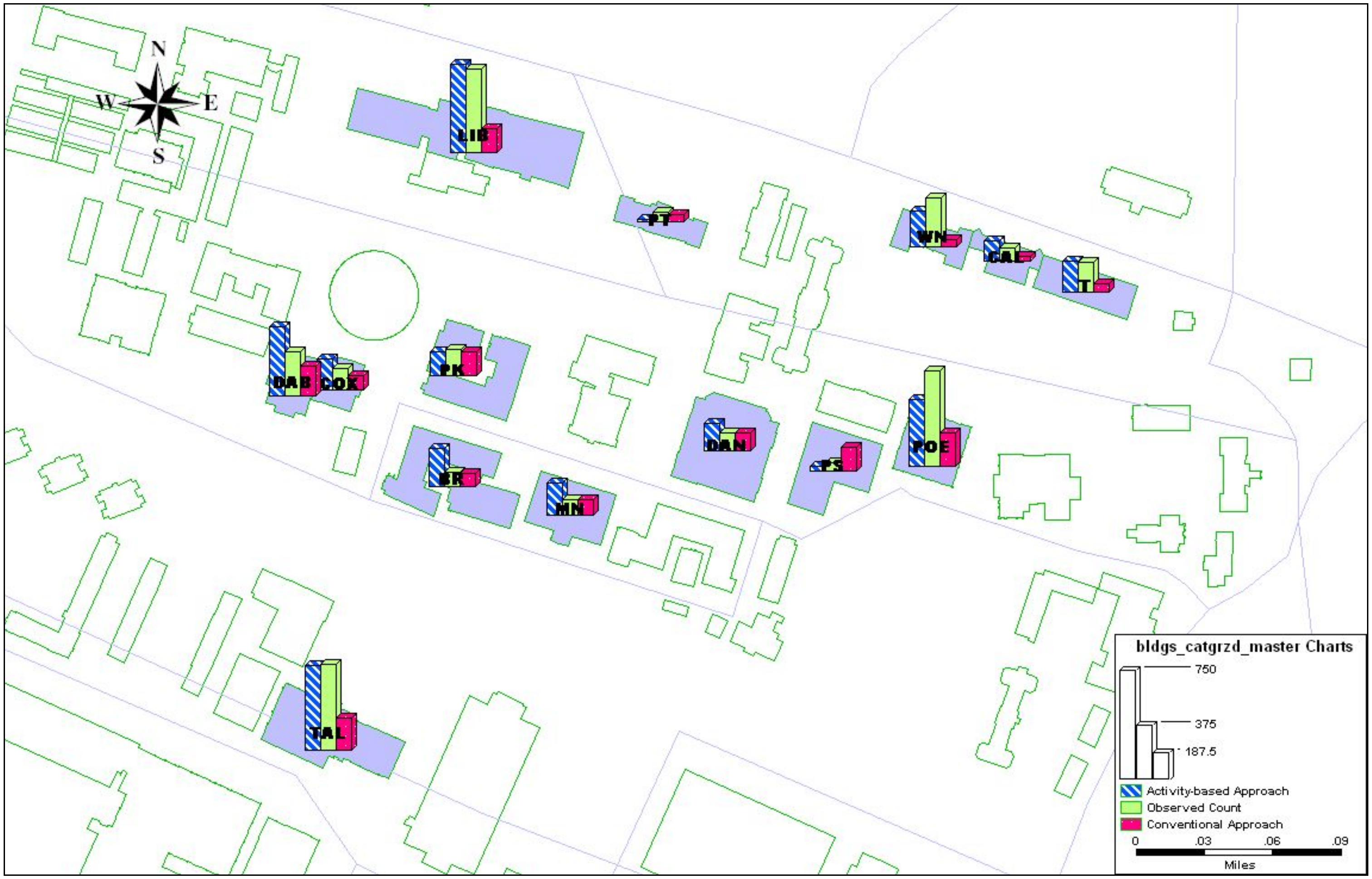


Figure 6.17 Comparison of Observations and Estimations for the Activity-based and Conventional Approaches

based approach show better accuracy as compared to the trips from a conventional approach even though the number of observed counts are not enough to say this with certainty.

6.5.4 Comparison of areas of model application

Aside from the quantitative analysis of trip generation in the previous section, it is also worth comparing the areas of model application between two approaches. One of the major motivations for developing an activity-based model was to meet the increased need for transportation policy analysis, which is addressed less well by the traditional travel demand models. Table 6.11 lists current research areas and the areas that a traditional approach can address. The traditional approach is mostly concerned with the impact of new infrastructure planned in the future and the analysis of emissions caused by traffic. However, the role of travel demand models has been shifted to meet the need for analyzing various transportation policies especially for transportation demand management (TDM).

Table 6.11 Analysis Topics

Current Analysis Topics	Traditional Analysis Topics
• Land development policies (smart growth, new urbanism) • Congestion pricing (HOV, HOT lanes) • Flexible working hours • Parking policies (pricing by hours) • Transit and walk access improvement • Ridesharing pricing and incentives • Telecommuting and related policies • Individualized marketing strategies	• Impact of new infrastructure (new major developments) • Demographic shifts (household size, employment) • Contribution to emission inventory

Source: Goulias (2004)

TDM policies such as congestion pricing, parking pricing, and telecommunication cause changes in trip-making behavior. Accordingly, the analysis of travel behavior needs an alternative model different from a conventional approach which is not able to capture the travel behavior regarding the TDM policies. The activity-based approach, therefore, becomes more important and its role is being expanded because of new study issues created as follows (Goulias, 2004):

- Homeland security
- Evacuation strategies
- Incorporating planning models into traffic operational studies
- Special event management

The one of the special features in the activity-based approach is a simulation of spatial-temporal presence of people which will be useful in planning homeland security and evacuation strategies since understanding the presence of people by time of day at specific locations is critical to develop strategies for evacuation plans. As discussed in the quantitative comparison between a conventional approach and activity-based approach, the more reliable travel demand estimated from an activity-based approach enhances the quality of analysis in the microscopic operational studies. Therefore, the effort to incorporate activity-based planning models into traffic operations studies is actively continuing and being using in real applications (e.g., TRANSIM). The activity-based approach is also expected to handle a short-term analysis of special events such as sports and religion events in the travel demand model.

6.5.5 Summary

According to the quantitative comparison of model accuracy between conventional and activity-based approaches, it is concluded that the conventional approach would not guarantee as much accuracy as the activity-based approach produces at a disaggregated level of model development. The developed activity-based model for the NCSU university special generator is seen to overcome the deficiencies of a conventional travel demand model for trip generation. Specifically: 1) the activity-based approach considers students' trip chains based on activity profiles, and 2) the building (destination) can be differentiated based on activity capacity estimated from destination choice models. Thus, the activity-based approach developed in this study provides more reliable estimates of trip generation although this approach requires extra effort that would not be necessary in a traditional approach where model accuracy may not be as critical.

Table 6.12 summarizes the model comparison between a conventional and activity-based approaches applied to a university special generator considering several perspectives such as data needs, model structure, modeling methods, model calibration and validation, etc.

The expected major benefits obtained from the activity-based approach over a conventional approach would be: 1) this approach can be applied to either TAZ or individual facility levels; 2) this approach does not require an extra data survey; 3) trip generation and distribution are considered simultaneously; 4) the model provides relatively high accuracy and the overall calibration process is superior to a conventional approach in terms of higher standard of validation with the fewer adjustments to model components and parameters. Further, the activity-based approach is able to meet the increased demand in the analysis of transportation policies which would not be afforded by a traditional approach.

Table 6.12 Summary of Comparisons between Conventional and Activity-based Approaches for the NCSU University Special Generator

Classification		Model		
		Conventional Approach		Activity-based Approach
Target Analysis Unit		TAZ	Individual facility	TAZ/ Individual facility
Data		-Socio-economic data -Travel survey	-Travel survey	-Socio-economic data -Travel survey
Modeling Structure		Aggregate/ Disaggregate	Disaggregate	Disaggregate
Model Unit		Person/ Household	Individual land use (facility type)	Person
Trip Generation Method		-Regression -Cross-classification (income, household size, etc)	-ITE Manual (No survey data) -Trip rates by land use or facility type (with data) -Departure & Arrival time distribution	- Daily activity pattern (Activity profile) - Spatial-temporal resolution (location & travel choice)
Trip Distribution Method		Gravity model (travel impedance)	Gravity model (travel impedance)	Choice-based (destination choice)
Consideration of Special Generator		-Trip Attraction Rates -Students considered in one-person household	Individual facility	Individual facility University population
Modeling Effort		Medium	Low (ITE Manual) Medium (Survey data)	High
Model Accuracy (not calibrated)		Low	Low	Medium
Expected Model Accuracy (calibrated)		Medium	Medium	High
Cost	Data Survey	Medium	Low (ITE Manual) Medium	Medium
	Model Calibration	High	Medium	Low
	Model Validation	Medium	Medium	Medium
TDM Policy Application		Low	Low	High

Note: 'Cost' represents time or budget required. The conventional approach with TAZ level is shown as a standard approach for comparison between activity-based and conventional approach with individual facility.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

7.1 Research Summary

Special generators are land uses that need particular attention in developing travel demand models because the travel patterns for special generators are different from standard land uses. The standard trip generation and distribution models in the conventional four-step approach are not expected to provide reliable estimates of special generator travel patterns. New modeling approaches such as activity-based and tour-based models, considering travel behavior of individual households or persons, seem to be more appropriate for special generators. However, only a few practical applications have been made since special generator approaches usually require significant data resources and computing time to solve their complicated model structure.

The primary objectives of this research are to improve the trip generation and trip distribution of a conventional approach by applying an activity-based approach to a university special generator, and to provide a transitional methodology for practically incorporating the activity-based data into a conventional planning model. To investigate the research objectives, this study developed several key components of the activity-based approach, which can be incorporated into a conventional approach with relative ease. These components are expected to enhance the potential ability of model applications to meet various needs in behavioral travel demand research. This research consisted of several tasks for developing a university activity-based travel demand model, and each are listed and briefly explained below.

- Literature review: The conventional four-step travel demand model and its deficiencies were reviewed. Then, the basic concept of activity-based models and the current practices of activity-based model were reviewed. Previous studies related to the practical application of activity-based models to special generators were examined. Then, a candidate model structure of an activity-based model was examined and discussed.

- Investigation of NCSU students' activity and travel patterns: The NCSU activity travel survey data were introduced and the students' activity and travel patterns were investigated. The data were used in the development of an activity-based model.

- Development of a university activity-based model: 1) A synthetic activity schedule was modeled by student groups based on the structure of an activity profile which represents hourly activity sequences and activity participation of students, 2) An on-campus destination choice model was developed by activity types for estimating spatial activity capacities, 3) Modeling a conditional probability of trips was accomplished for estimating hourly travel demand.

- Investigation of the transferability of university student activity data: NCSU student activity and travel data were statistically compared to PSU student activity data. The comparison of trip rates and activity profiles between the two universities were accomplished.

- Estimation of hourly presence of students in buildings (building presence) and travel demand: The spatial and temporal presence of students in buildings (building presence) was estimated based on the activity profile and activity capacity in each building on the NCSU campus. The travel demand, trip production and attraction, were estimated by applying the estimated conditional probability of trips to the building presence. The estimated trips were compared to actual student trips counted at entrances and exits at fourteen sample campus buildings.

- Comparisons of activity-based and traditional approaches: The quantitative comparison between the two approaches was accomplished to examine which approach is better in terms of model accuracy. In addition, the areas of current model application and the future need of activity-based models were discussed.

The activity-based approach used in this study successfully estimated trips at the level of university buildings and provided several findings that are expected to positively contribute to the application of activity-based approach to special generators, as explained in the next section.

7.2 Findings and Conclusions

The findings and conclusions are separately presented here following the procedure of this research as listed in the previous section.

Firstly, the major findings from the analysis of the NCSU student activity-travel behavior survey can be summarized as follows:

- Undergraduate students and on-campus residents are more likely to be involved in various activities than graduate students and off-campus residents. Consequently, the trip rates of on-campus students are higher than the off-campus students.

- NCSU student trip-making behavior is significantly different within the groups classified by residential status (on-and off-campus students), student status (undergraduate and graduate student), and driver's license.

- Walk is a primary mode for on-campus residents while the automobile is the major mode for off-campus residents. Walking secondary trips are the most frequent for both on- and off-campus students. Auto use on campus is not as high as for a standard household, but auto occupancy is higher.

- For all activities 'Meals' is the most frequent subsequent activity. The most frequent activity transition is School/Class undertaken immediately after Sleep.

Findings indicate that activity-travel behavior of university students is greater than for standard households in terms of trip rates. As a result, activity-sequence is likely more diverse. Although the majority of trips occurred on campus between campus buildings, their travel patterns may affect regional travel demand due to their flexible schedule and higher or lower trip rates during peak hour. This effect should be reflected in a regional travel demand model. It is not sufficient to classify university trips as home-based school trips or the aggregated trip rates of one-person households with low-income, which are typical practices to represent university special generators in regional models.

Secondly, the major findings from testing transferability of university activity-travel data from one university to another are summarized as follows.

- The students' daily trip rates between NCSU and PSU do not seem to be very different and this was verified by the statistical tests using the negative binomial model and ordinary regression model. As expected, the student households are more likely make a trip compared to regular households since both NCSU and PSU students are involved in a variety of campus activities.

- The activity profile consisting of hourly activity sequence and activity participation was not statistically different among student groups, and in the comparison of students' activity profiles between two universities, the overall activity pattern during a day looks similar, and statistical tests verify that the activity profiles are not statistically different from one another.

As a result, it is concluded that PSU and NCSU activity data can be transferred, and perhaps more generally to other large universities as well. However, the magnitude of the proportion of activity participation is critical in the estimation of how many people are engaged in what activities. Consequently, if activity survey data are available and the observations are plentiful, it would be better to develop the activity profile separately for each traveler group.

Thirdly, the major findings from the modeling the components of activity-based travel demand model are summarized as follows:

- Activity profiles were successfully modeled with the explanatory variables such as activity type, hour, and the interaction variable (activity $\otimes$ time).

- Students' daily activity profiles indicate that there are two peak hours for a 'Work' activity (including 'Class and Research') which appear later in the morning around 10 a.m. and early afternoon around 2 p.m. It is important to note that the peaks of university activities will be at off-peak times for the surrounding community.

- Daily activity participation is not significantly different between student groups when classified by gender, educational and residential status. Not being significantly different is advantageous because it simplifies activity-based travel demand modeling. Only activity types that are significant with respect to the proportion of students engaged in a certain activity at a certain time of day would have to be modeled. This means that time of day activity-travel behavior, as demonstrated in this paper, should be taken into account in activity-based travel demand modeling.

In this study, the proposed model is targeted for an aggregated activity profile that is much easier to model for the disaggregated components of an activity schedule. There is no doubt in that less complexity in model development would be better for practical model application.

- Destination choice models were successfully developed within the structure of the Multinomial Logit (MNL) approach based on discrete choices of university facilities. However, the MNL model for the 'Shopping' activity was not developed due to the lack of shopping opportunities and data associated with on-campus shopping trips. These models were the basis of developing relative activity capacity for each building on-campus. Based on this activity capacity, students were spatially distributed to the campus buildings with respect to their activities at every hour.

The purpose for developing discrete choice models in this study is to estimate the attractiveness of each building by activity type rather than to analyze students' destination choice behavior. The reason is that students' destination choices for campus buildings for all activity types were not dependent on their choices, but dependent on the university environment (i.e., class schedule, meeting or conference, etc).

Finally, the major findings from model validation of the activity-based travel demand model, and the results of quantitative comparisons between the trips estimated from an activity-based model and a conventional model are summarized below:

- The estimated number of students in each building (building presence) was directly compared to the actual students registered for classes in a building. The comparison illustrated that the overall estimated building presence replicated actual students relatively well even though the estimated presence included all students either in-class or not.

- The trip productions and attractions at each building were estimated based on the building presence and the MNL model for conditional probability of trips by activity types. Overall, the modeled trips were found to be slightly greater than the actual student trips entering and exiting at a building.

- The conventional approach did not provide reliable trips at a disaggregated level of model development for this university special generator.

As a result, developing an activity-based model for a university special generator seems to overcome the deficiencies of traditional travel demand model within the trip generation step. Even though the activity-based model development requires extra effort that would not be necessary in a traditional approach if model accuracy is not a critical issue in the analysis. However, if the model accuracy is a concern, the activity-based approach should be used for trip estimation. Further, the activity-based approach is able to meet the increased demand in the analysis of transportation policies which would not be afforded by the traditional travel demand models.

7.3 Recommendations for Future Research

This study provides valuable insights for activity-based university travel demand modeling. However, several issues that were not dealt with due to limited research scope and data should be addressed in future studies.

First, this study only considered student activity-travel behavior from one travel survey. This exploratory analysis should be extended to other university populations - faculty, staff, and visitors – and other universities.

Second, in order to apply unique student travel patterns in travel demand modeling, further comparisons are needed between university special generators and standard households.

Third, further study is needed on how to incorporate university travel patterns into a regional travel demand model as either conventional four-step or activity-based modeling. This study only addressed student trips on-campus and the estimated trips from the activity-based approach were found to provide better model accuracy compared to a conventional approach. The model evaluation should be carried out in a regional travel demand model either by incorporating the activity-based approach within the conventional four-step modeling or by developing an activity-based model for an entire region. It would be interesting to see how much the activity-based approach improves model accuracy in an aggregated level of model development, since the activity-based approach is also expected to provide better estimates of trips for a regional level,

Finally, further comparisons are needed between conventional and activity-based modeling to provide better insights on the activity-based approach to planning agencies and model practitioners. Activity-based models may improve forecasting future travel patterns when traffic conditions are changed from base year conditions. They can easily assess travel pattern shifts after the implementation of major changes in transportation services or polices, not just replicating base year patterns. Qualitative comparisons are also required based on an evaluation of activity models being able to respond to a range of scenarios and policies of interest.

REFERENCES

1. Activity-based modeling system for travel demand forecasting, Travel Model Improvement Program, U.S. Department of Transportation, 1995.

2. A system of activity-based models for Portland, Oregon, Travel Model Improvement Program, U.S. Department of Transportation, 1998.

3. Bush, B.W. (2000) "An algorithmic overview of TRANSIMS", http://transims.tsasa.lanl.gov.

4. Algers, S., Eliasson, J. and Mattsson, L., Activity-based model development to support transport planning in the Stockholm region, presented at the 5th Workshop of the TLE Network, Nynashamn, September 28-30, 2001.

5. Activity and Tour-based model seminar, Cambridge systematics, Inc, 2004.

6. Ortuzar and Willumsen (1994) *Modeling Transport 2nd edition*, Wiley, New York.

7. Travel demand model (1990), TMIP.

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

9. Timmermans, H.J.P., T.A. Arentze and C.H. Joh, Modeling learning and evolutionary adaptation processes in activity settings: Theory and numerical simulations, *Transportation Research Record,* 1718, 27-33. 2000

10. McNally, M. G. "The Activity-Based Approach." In Handbook of Transport Modelling. Edited by Hensher, D.A. and Button, pp. 113-128. 2000.

11. Ettema, D.F. and Timmermans, H.J.P. *Activity-Based Approaches to Travel Analysis,* Elsevier Science Ltd, 1997.

12. Kulkarni, A.A., and M.G. McNally, A Microsimulation of Daily Activity Patterns," Institute of Transportation Studies, University of California, Irvine,UCI-ITS-AS-WP-00-7, 2000.

13. Hägerstrand, T. What about people in regional science?, *Regional Science Association Papers*, 24:7-21, 1970.

14. Capin, F.S. (1974) *Human Activity Patterns in the City*, Wiley, New York.

15. Bowman, J.L. and Ben-Akiva, M.E. Activity based disaggregate travel demand model system with activity schedules, *Transportation Research Part A*, vol35, pp1-28, 2000.

16. Pendyala, R.M., Kitamura, R. and Reddy, D.V.G.P. Application of an activity-based travel demand model incorporating a rule-based algorithm, *Environment and Planning B*, vol25. pp753-772. 1998.

17. Ettema, D.F. and Timmermans, H.J.P. *Activity-Based Approaches to Travel Analysis,* Elsevier Science Ltd. 1997

18. Goulias K.G. (2003) *Tranportation Systems Planning*, CRC Press, Boca Raton.

19. Kitamura, R., Fujii, S. Yamamoto, T. and Kikuchi, A. "Application of PCATS/DEBNets to regional planning and policy analysis: Micro simulation studies for the cities of Osaka and Kyoto, Japan", Presented at the Transportation Research Board.2001.

20. Kitamura, R., Lula and Pas E. (1993) "AMOS: An activity-based, flexible and truly behavioral tool for evaluation of TDM measures", PTRC 1993.

21. Kosonen, I. and Ree, S. "The potential of microscopic simulation in traffic safety and conflict studies", Unpublished manuscript, Royal Institute of Technology, Stockholm, and Los Alamos National Laboratory.

22. Simon, P. and Nagel K. (1998) "Simplified cellular automaton model for city traffic", *Physical Review E*, 58:1286-1295.

23. Kuhnau, J.L., Activity-Based Travel Demand Modeling within the Urban Transportation Planning System, Master Thesis, The Pennsylvania State University, Department of Civil and Environmental Engineering, University Park, Pennsylvania, 2001.

24. MAB, The final report for analysis of NCSU travel demand survey, 2004.

25. South Central Centre County Transportation Study (SCCCTS), PENNDOT, 2003.

26. Triangle Regional Model, NC DOT, 1995.

27. Alan Agresti, *Categorical Data Analysis*, Wiley, New Jersey, 2002.

28. TransCAD User's Guide. Version 4.5. Caliper Corporation, Newton, MA. 2002.

29. Ming, S.L. and McNally, M.G. On the structure of weekly activity/travel patterns, *Transportation Research Part A*. vol37. pp823-839. 2003.

30. McNally, M. "The activity-based approach", in Hensher, D.A .and Button, K.J. *Handbook of Transport Modelling*, Elsevier Science Ltd, 53-69.2000.

31. Ondrej, P., A Microsimulation Model of Activity Patterns and within Household Interactions PhD Thesis, The Pennsylvania State University, Department of Civil and Environmental Engineering, University Park, Pennsylvania, 2004.

32. Data collection in the Portland, Oregon Metropolitan area case study, Travel Model Improvement Program, U.S. Department of Transportation, 1996.

33. Tomas F. Golob, A simultaneous model of household activity participation and trip chain generation, *Transportation Research Part B*. vol34. pp. 355-376. 2000.

34. Chow, G.C., Tests of equality between sets of coefficients in two linear regressions, *Econo-metrica*, 28, pp. 591-605. 1960.

35. B. Efron. and R.J. Tibshirani. *An Introduction to the Bootstrap*, Chapman and Hall, 1993.

36. Beran, R. *Prepivoting test statistics: a bootstrap view of asymptotic refinements,* J.Amer.Statist. Ass., 83, 1988, pp. 687-697.

37. Hall, P. *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York, 1992.

38. Davidson, R. and J.G. MacKinnon. The size distortion of bootstrap tests, Queen's university, I.E.R. Discussion Paper, *No. 936*,1998b.

39. Davidson, R. and J.G. MacKinnon. *Econometric Theory and Methods*, Oxford University Press, New York, 2003.

40. Mammen, E. Bootstrap and wild bootstrap for high dimensional linear models, In *Journal of Annals of Statistics*, Vol. 21, 1993, pp. 255-285.

41. Godfrey, L.G. and Orme, C. D. Controlling the significance levels of prediction error tests for linear regression models. In *Econometrics Journal*, Vol. 3, 2000, pp. 66-83.

42. Godfrey, L.G. and Orme, C. D. Using bootstrap methods to obtain nonnormality robust Chow prediction tests. In *Econometrics Letters*, Vol. 76, 2002, pp. 429-436.

43. http://www.r-project.org/

44. ITE Trip generation manual, 1997.

45. Zhao, F., L. F. Chow, and M.T. Li (2004) *Refinement of FSUTMS Trip Distribution Methodology,* Lehman Center for Transportation Research, FDOT.

46. Hess, S. and J. Polak (2005) *Mixed Logit Modeling of Airport Choice in Multi-Airport Regions*, Journal of Air Transport Management, Vol. 11, No. 2, pp.59-68.

47. Tran, K., (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, UK

48. Bhat, C. R. and H. Zhao (2003) *Transportation Control Effectiveness in Ozone Non-Attainment Areas: Final Report*, Research Report 1838-8, Center for Transportation Research, Austin, Texas.

49. McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," Frontiers in Econometrics, Zarembka ed., New York, Academic Press, pp. 105-142, 1974.

50. Allison, Paul David (1999) *Logistic regression using the SAS system : theory and application*, SAS Institute Inc, Cary

51. http://www.trb.org/Conferences/TDM/papers/BS3C%20-20Activity%20Model%20Validation_Pendyala%20and%20Bhat.pdf

52. MAB, NCSU travel demand model, 2004

APPENDICES

A. North Carolina State University Student Travel Survey

1. On-campus student

Draft NCSU Transportation and Parking Questionnaire
On Campus Student
February 28, 2000

Thank you for taking the time to fill out this survey. Only a sample of the NCSU community has received a copy, so your answers are very important. We will only use this information to improve the parking and travel options on and to campus. Refer to the campus map attached to the survey for questions as noted.

NCSU Origin/Destination Information

1. Where do you live?
   1) Dormitory ____________________

2. What building is your classroom "home"? (where you spend 51% or more of your time, refer to map)

   Building Name or Precinct (e.g., "Harrellson Hall" or North Precinct): ____________________

3. What days per week do you usually travel to this location (check all that apply)?

   Monday ☐ Tuesday ☐ Wednesday ☐ Thursday ☐ Friday ☐ Saturday ☐ Sunday ☐

4. What is your primary trip purpose to this location?
   1) Attend class/lab ☐
   2) Attend meeting ☐
   3) Other (specify: ____________) ☐

5. What time do you usually arrive and what time do you ultimately leave this location?

   1) Arrival Time ___ : ___ AM or PM 2) Depart Time ___ : ___ AM or PM

   5a. Do your arrival and departure times usually vary from day to day? Yes ☐ No ☐

   5b. If yes, describe how your times vary ____________________

Travel Information

6. How did you usually travel to your classes (check one)?
   1) Wolfline/Wolflink ☐
   2) Bicycle ☐
   3) Walk ☐
   4) Other (please specify: ____________) ☐

7. Would you prefer to travel via a different mode?
   1) Yes ☐
   2) No (skip to Question #8) ☐

   7a. If yes, what mode would you prefer?
   Preferred mode: ____________________

   7b. What is the primary reason that you did not travel this way (check one)?
   1) Parking not available ☐
   2) Parking too expensive ☐
   3) Transit not convenient ☐
   4) Transit too expensive ☐
   5) Transit unreliable ☐
   6) Too far to bike or walk ☐
   7) Biking/walking feels unsafe ☐
   8) Other (specify ____________) ☐

Page 1 of 3

8. What type of parking permit do you have (check one)?
   1) None ☐ 4) Permit D ☐ 7) Permit L ☐
   2) Permit DC ☐ 5) Permit E ☐
   3) Permit DD ☐ 6) Permit F ☐ 8) Other (________) ☐

9. If you drive to class, where did you park (check one)? Otherwise, skip to Q #10?
   1) Parked on street ☐ How many blocks away? ______
   2) Parking Deck ☐ Deck Name ______________
   3) Surface Parking Lot ☐ Lot Name or Location ________
   4) Parking meter on campus ☐ Location ________________
   5) Other Facility ☐ Specify ________________

   9a. How much does it cost you to park (fill in one response)?
   1) $ ______ per day; 3) $ ______ per semester; or
   2) $______ per month; 4) $______ per year

Travel During the Day

10. How many times per week do you travel between NCSU precincts (check one)?
    1) Never (skip to Question #12) ☐ 3) 4 – 9 trips per week ☐
    2) 1 - 3 trips per week ☐ 4) 10 or more trips per week ☐

11. How do you usually travel between precincts (check one)?
    (1) Walk ☐ (4) Bicycle ☐
    (2) Take Wolfline/Wolflink ☐ (5) Carpool/vanpool ☐
    (3) Drive ☐ (6) Other (______________) ☐

12. How many times per week do you travel **off** campus (check one)?
    1) Never (skip to Question #16) ☐ 3) 4 – 9 trips per week ☐
    2) 1 - 3 trips per week ☐ 4) 10 or more trips per week ☐

13. What is your trip purpose (check as many as apply)?
    1) Eating ☐ 4) Doctor's appointment ☐
    2) Shopping ☐ 5) Other (________________) ☐
    3) Visit friends ☐

14. How do you travel off-campus (check one)?
    1) Walk ☐ 4) Bicycle ☐
    2) Take Wolfline ☐ 5) Carpool/vanpool ☐
    3) Drive ☐ 6) Other (________________) ☐

15. How often do you ride the Wolfline/Wolflink/Werewolf bus routes?
    1) Every day ☐ 4) Only occasionally ☐
    2) 3-4 days per week ☐ 5) Never used ☐
    3) 1-2 days per week ☐

15. How well do the Wolfline routes and schedules meet your expectations (check one)?
    1) Much better ☐ 4) Worse than expected ☐
    2) Better than expected ☐ 5) Much worse ☐
    3) Just as expected ☐

Draft NCSU Transportation and Parking Questionnaire
On Campus Student
February 28, 2000

Organizational Information

16. Which category best represents your position at NCSU?
   1) Undergrad, full-time ☐
   2) Undergrad, part-time ☐
   3) Graduate/doctorial, full-time ☐
   4) Graduate/doctorial, part-time ☐
   5) Other (________________) ☐

17. Which Department do you work for? (check one):
   1) Agriculture & Life Sciences ☐
   2) Design ☐
   3) Education & Psychology ☐
   4) Engineering ☐
   5) Forest Resources ☐
   6) Graduate School ☐
   7) Humanities & Social Sciences ☐
   8) Management ☐
   9) Physical & Mathematical Sciences ☐
   9. Textiles ☐
   10. Veterinary Medicine ☐
   10) Other (____________________) ☐

18. Do you have any comments on transportation and parking conditions at NCSU?

____

____

If you would like to enter your name into a drawing for a **$50 gift certificate** from the NCSU bookstore, fill out your name below

Name: ______________________

Address: ______________________

Daytime Phone: ______________________

Thank you for completing this survey.
Your assistance is greatly appreciated.

2. Commuting student

Draft NCSU Transportation and Parking Questionnaire
Commuter Student
February 28, 2000

Thank you for taking the time to fill out this survey. Only a sample of the NCSU community has received a copy, so your answers are very important. We will only use this information to improve the parking and travel options on and to campus. Refer to the campus map attached to the survey for questions as noted.

NCSU Origin/Destination Information

1. Where do you live?
   1) Zip Code __ __ __ __ __

2. What building is your primary "home"? (where you spend 51% or more of your time, refer to map)

   Building Name or Precinct (e.g., "Harrellson Hall" or North Precinct): ____________________

3. What days per week do you usually travel to this location (check all that apply)?
   Monday ☐ Tuesday ☐ Wednesday ☐ Thursday ☐ Friday ☐ Saturday ☐ Sunday ☐

4. What is your primary trip purpose to this location?
   1) Attend class/lab ☐
   2) Attend meeting ☐
   3) Other (specify: ________________) ☐

5. What time do you usually arrive and what time do you ultimately leave this location?

   1) Arrival Time __ : __ AM or PM 2) Depart Time __ : __ AM or PM

   5a. Do your arrival and departure times usually vary from day to day? Yes ☐ No ☐

   5b. If yes, describe how your times vary ____________________________

Travel Information

6. How did you usually travel to NCSU (check one)?

   By Public Transportation
   1) CAT (please specify CAT lines: # ____________) ☐
   2) TTA (route __________) ☐
   3) TTA & CAT (routes __________) ☐
   4) Other (specify __________) ☐

   By Wolfline Shuttle
   1) From Blue Ridge Park & Ride ☐
   2) From Westgrove Park & Ride ☐
   3) From Varsity Park & Ride ☐
   4) From neighborhood stop ☐

   By Car, Van, or Motorcycle
   1) Car or van, drove alone ☐
   2) Carpool or vanpool, passenger ☐
   3) Carpool or vanpool, driver ☐
   4) Motorcycle ☐

   By Other Means of Travel
   1) Bicycle ☐
   2) Walk ☐
   3) Other (please specify: ________________) ☐

Draft NCSU Transportation and Parking Questionnaire
Commuting Student
February 28, 2000

7. Would you prefer to travel via a different mode?
   1) Yes ☐ 2) No (skip to Question #8) ☐

   7a. If yes, what mode would you prefer?
   Preferred mode: ______________________

   7b. What is the primary reason that you did not travel this way (check one)?
   1) Parking not available ☐ 6) Too far to bike or walk ☐
   2) Parking too expensive ☐ 7) Biking/walking feels unsafe ☐
   3) Transit not convenient ☐ 8) Carpool/vanpool options not available ☐
   4) Transit too expensive ☐ 9) Car in the shop ☐
   5) Transit unreliable ☐ 10) Other (specify ____________) ☐

8. What type of parking permit do you have (check one)?
   1) None ☐ 4) Permit D ☐ 7) Permit L ☐
   2) Permit DC ☐ 5) Permit E ☐
   3) Permit DD ☐ 6) Permit F ☐ 8) Other (________) ☐

9. If you drive to NCSU, where did you park (check one)? Otherwise, skip to Q #10?
   1) Parked on street ☐ How many blocks away? _____
   2) Parking Deck ☐ Deck Name ______________
   3) Surface Parking Lot ☐ Lot Name or Location ________
   4) Parking meter on campus ☐ Location ______________
   5) Other Facility ☐ Specify ______________

   9a. How much does it cost you to park (fill in one response)?
   1) $_______ per day; 3) $_______ per semester; or
   2) $_______ per month; 4) $_______ per year

Travel During the Day

10. How many times per week do you travel between NCSU precincts (check one)?
    1) Never (skip to Question #12) ☐ 3) 4 – 9 trips per week ☐
    2) 1 - 3 trips per week ☐ 4) 10 or more trips per week ☐

11. How do you usually travel between precincts (check one)?
    (1) Walk ☐ (4) Bicycle ☐
    (2) Take Wolfline/Wolflink ☐ (5) Carpool/vanpool ☐
    (3) Drive ☐ (6) Other (______________) ☐

12. How many times per week do you travel **off** campus (check one)?
    1) Never (skip to Question #16) ☐ 3) 4 – 9 trips per week ☐
    2) 1 - 3 trips per week ☐ 4) 10 or more trips per week ☐

13. What is your trip purpose (check as many as apply)?
    1) Eating ☐ 4) Doctor's appointment ☐
    2) Shopping ☐ 5) Other (______________) ☐
    3) Visit friends ☐

14. How do you travel off-campus (check one)?
    1) Walk ☐ 4) Bicycle ☐
    2) Take Wolfline ☐ 5) Carpool/vanpool ☐
    3) Drive ☐ 6) Other (______________) ☐

Draft NCSU Transportation and Parking Questionnaire
Commuting Student
February 28, 2000

15. How often do you ride the Wolfline/Wolflink/Werewolf bus routes?
   1) Every day ☐
   2) 3-4 days per week ☐
   3) 1-2 days per week ☐
   4) Only occasionally ☐
   5) Never used ☐

15. How well do the Wolfline routes and schedules meet your expectations (check one)?
   1) Much better ☐
   2) Better than expected ☐
   3) Just as expected ☐
   4) Worse than expected ☐
   5) Much worse ☐

Organizational Information

16. Which category best represents your position at NCSU?
   1) Undergrad, full-time ☐
   2) Undergrad, part-time ☐
   3) Graduate/doctorial, full-time ☐
   4) Graduate/doctorial, part-time ☐
   5) Other (________________) ☐

17. Which Department do you work for? (check one):
   1) Agriculture & Life Sciences ☐
   2) Design ☐
   3) Education & Psychology ☐
   4) Engineering ☐
   5) Forest Resources ☐
   6) Graduate School ☐
   7) Humanities & Social Sciences ☐
   8) Management ☐
   9) Physical & Mathematical Sciences ☐
   9. Textiles ☐
   10. Veterinary Medicine ☐
   10) Other (______________________) ☐

18. Do you have any comments on transportation and parking conditions at NCSU?

____

____

If you would like to enter your name into a drawing for a **$50 gift certificate** from the NCSU bookstore, fill out your name below

Name: ______________________________

Address: ______________________________

Daytime Phone: ______________________________

Thank you for completing this survey.
Your assistance is greatly appreciated.

B. SAS Data and Output of Negative Binomial Model

1. Example Data

The SAS System

Obs	id	Trips	residence	gender	undgrad	class	EMPSTAT	Licens
1	10009	7	1	0	0	5	1	1
2	10011	8	0	0	0	5	2	1
3	10013	6	1	0	0	5	4	1
4	10019	10	1	0	1	1	4	1
5	10034	9	1	0	1	1	2	1
6	10042	11	1	0	1	1	4	1
7	10076	6	1	0	1	1	4	1
8	10079	6	1	0	1	1	2	1
9	10086	2	1	0	1	1	4	1
10	10087	13	1	0	1	1	4	1

2. SAS Output

Model Information

Data Set	WORK.JINKI	
Distribution	Negative Binomial	
Link Function	Log	
Dependent Variable	Trips	Trips

Number of Observations Read	822
Number of Observations Used	822

Class Level Information

Class	Levels	Values
residence	2	0 1
gender	2	0 1
undgrad	2	0 1
class	5	1 2 3 4 5
EMPSTAT	4	1 2 3 4
Licens	2	0 1

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	811	844.5512	1.0414
Scaled Deviance	811	844.5512	1.0414
Pearson Chi-Square	811	805.2875	0.9930
Scaled Pearson X2	811	805.2875	0.9930
Log Likelihood		4499.0210	

The GENMOD Procedure

Algorithm converged.

Analysis Of Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept		1	1.8734	0.0528	1.7700	1.9769	1259.26	<.0001
residence	0	1	-0.0874	0.0370	-0.1599	-0.0148	5.57	0.0183
residence	1	0	0.0000	0.0000	0.0000	0.0000	.	.
gender	0	1	0.0323	0.0339	-0.0341	0.0988	0.91	0.3406
gender	1	0	0.0000	0.0000	0.0000	0.0000	.	.
undgrad	0	1	-0.0973	0.0536	-0.2024	0.0078	3.29	0.0696
undgrad	1	0	0.0000	0.0000	0.0000	0.0000	.	.
class	1	1	0.0977	0.0525	-0.0051	0.2006	3.47	0.0626
class	2	1	0.1208	0.0539	0.0151	0.2265	5.01	0.0252
class	3	1	0.0832	0.0531	-0.0209	0.1874	2.45	0.1173
class	4	0	0.0000	0.0000	0.0000	0.0000	.	.
class	5	0	0.0000	0.0000	0.0000	0.0000	.	.
EMPSTAT	1	1	-0.2021	0.0584	-0.3166	-0.0876	11.97	0.0005
EMPSTAT	2	1	0.0227	0.0348	-0.0455	0.0908	0.43	0.5143
EMPSTAT	3	1	0.0813	0.1459	-0.2046	0.3673	0.31	0.5772
EMPSTAT	4	0	0.0000	0.0000	0.0000	0.0000	.	.
Licens	0	1	-0.2014	0.0899	-0.3776	-0.0251	5.01	0.0252
Licens	1	0	0.0000	0.0000	0.0000	0.0000	.	.
Dispersion		1	0.0424	0.0100	0.0228	0.0621		

NOTE: The negative binomial dispersion parameter was estimated by maximum likelihood.

The GENMOD Procedure

LR Statistics For Type 1 Analysis

Source	2*Log Likelihood	DF	Chi-Square	Pr > ChiSq
Intercept	8916.7639			
residence	8942.6742	1	25.91	<.0001
gender	8942.8479	1	0.17	0.6768
undgrad	8971.8798	1	29.03	<.0001
class	8977.1486	3	5.27	0.1531
EMPSTAT	8992.8501	3	15.70	0.0013
Licens	8998.0420	1	5.19	0.0227

LR Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq
residence	1	5.54	0.0186
gender	1	0.91	0.3410
undgrad	0	0.00	.
class	3	5.57	0.1343
EMPSTAT	3	16.40	0.0009
Licens	1	5.19	0.0227

C. Statistical Test of Similarity of Activity Profile by Traveler Group

1. R-CODE

```
probplots <- function(){
faculty=read.table("faculty.txt",header=T)
faculty=faculty[,-1]
n=nrow(faculty)
m=ncol(faculty)
student=read.table("student.txt",header=T)
student=student[,-1]
worker=read.table("worker.txt",header=T)
worker=worker[,-1]

#par(mfrow=c(2,2),cex=0.65)
x11()
hours=c(0:23)
plot(hours,faculty[,1],type="l",ylim=c(0,1),ylab="proportions",col=1)
for(j in 2:m){
lines(hours, faculty[,j],lty=j, col=j) }
legend(list(x=11,y=1.0),legend=c("Home","Work","Shoping","Recreation","Other"),lty=1:5, col=1:5,cex=0.8)
title("faculty")

x11()
plot(hours,student[,1],type="l",ylim=c(0,1),ylab="proportions",col=1)
for(j in 2:m){
lines(hours, student[,j],lty=j, col=j) }
legend(list(x=11,y=1.0),legend=c("Home","Work","Shoping","Recreation","Other"),lty=1:5, col=1:5,cex=0.8)
title("student")

x11()
plot(hours,worker[,1],type="l",ylim=c(0,1),ylab="proportions",col=1)
for(j in 2:m){
lines(hours, worker[,j],lty=j, col=j) }
legend(list(x=11,y=1.0),legend=c("Home","Work","Shoping","Recreation","Other"),lty=1:5, col=1:5,cex=0.8)
title("worker")

average=matrix(0,n,m)
for(i in 1:n){
for(j in 1:m){
average[i,j]=(faculty[i,j]+student[i,j]+worker[i,j])/3
        }}

x11()
plot(hours,average[,1],type="l",ylim=c(0,1),ylab="mean proportions",col=1)
for(j in 2:m){
lines(hours, average[,j],lty=j, col=j) }
legend(list(x=11,y=1.0),legend=c("Home","Work","Shoping","Recreation","Other"),lty=1:5, col=1:5,cex=0.8)
title("average")

x11()
#par(ask=T)
#par(mfrow=c(2,2),cex=0.65)
diffFS=faculty-student
diffFW=faculty-worker
diffSW=student-worker

plot(hours,diffFS[,1],type="l",ylim=range(diffFS),ylab="difference",col=1)
for(j in 2:m){
lines(hours, diffFS[,j],lty=j, col=j) }
legend(list(x=17,y=0.45),legend=c("Home","Work","Shoping","Recreation","Other"),lty=1:5, col=1:5,cex=0.8)
abline(h=0)
```

```
title("faculty - student")

x11()
plot(hours,diffFW[,1],type="l",ylim=range(diffFW),ylab="difference",col=1)
for(j in 2:m){
lines(hours, diffFW[,j],lty=j, col=j) }
legend(list(x=17,y=0.2),legend=c("Home","Work","Shoping","Recreation","Other"),lty=1:5, col=1:5,cex=0.8)
abline(h=0)
title("faculty - worker")

x11()
plot(hours,diffSW[,1],type="l",ylim=range(diffSW),ylab="difference",col=1)
for(j in 2:m){
lines(hours, diffSW[,j],lty=j, col=j) }
abline(h=0)
legend(list(x=17,y=0.35),legend=c("Home","Work","Shoping","Recreation","Other"),lty=1:5, col=1:5,cex=0.8)
title("student - worker")

##Statistical analysis:
sink("D:\\Jinki files\\Other Files\\NCSU model\\Activity model\\profile comp\\analysis.txt")
for(i in 1:n){
cat(c("Time=",i-1),fill=T)
cat("------------------------------",fill=T)
proportion=c(as.numeric(faculty[i,]),as.numeric(student[i,]),as.numeric(worker[i,]))
group=as.factor(rep(1:3,each=m))
activity=as.factor(rep(1:m,3))
fit=lm(proportion ~ group+activity)
print(summary(fit))
print(anova(fit))
cat("------------------------------",fill=T)
        }
cat(c("Job finished on:",date()),fill=T)
sink()

return(cat(c("Job finished on:",date()),fill=T))
}
```

2. Example Output from R

```
Time= 0
------------------------------

Call:
lm(formula = proportion ~ group + activity)

Residuals:
      Min        1Q    Median        3Q       Max
-5.545e-02 -1.175e-02 -8.466e-17  8.743e-03  4.694e-02

Coefficients:
             Estimate Std. Error  t value Pr(>|t|)
(Intercept)  9.458e-01  2.195e-02    43.09 9.27e-11 ***
group2       4.000e-10  2.032e-02 1.97e-08        1
group3       2.000e-10  2.032e-02 9.84e-09        1
activity2   -9.295e-01  2.623e-02   -35.44 4.40e-10 ***
activity3   -9.458e-01  2.623e-02   -36.06 3.84e-10 ***
activity4   -9.224e-01  2.623e-02   -35.16 4.68e-10 ***
activity5   -9.313e-01  2.623e-02   -35.50 4.34e-10 ***
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 0.03213 on 8 degrees of freedom
Multiple R-Squared: 0.9961,  Adjusted R-squared: 0.9931
F-statistic: 336.9 on 6 and 8 DF,  p-value: 3.597e-09

Analysis of Variance Table

Response: proportion
          Df    Sum Sq   Mean Sq   F value    Pr(>F)
group      2 4.000e-19 2.000e-19 1.938e-16         1
activity   4   2.08677   0.52169    505.42 1.203e-09 ***
Residuals  8   0.00826   0.00103
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
```

D. Bootstrap Chow-test R-code and Data

1. R-code

```
library(boot)
library(car)
sample<- read.table("D:\\ bootdata.txt",header=T)
chow<- function(d,i){
        n <- nrow(d)
        rss<-lm(at^0.384~ug+off+sex,data=d) # at=y,
        y<-rss$fitted+rss$resid[i]
        set<-cbind(y,d[,2:5])
        set1<-set[set$pn==1,]
        set2<-set[set$pn==2,]
        brss1<-lm(y~ug+off+sex,data=set1)
        brss2<-lm(y~ug+off+sex,data=set2)
        brss<-lm(y~ug+off+sex,data=set)
        bssr1<-sum(brss1$resid^2)
        bssr2<-sum(brss2$resid^2)
        bssr<-sum(brss$resid^2)
        chow<-((bssr-bssr1-bssr2)/4)/((bssr1+bssr2)/(n-8))
        c(chow)
    }
    a1<-boot(sample, chow, R=999)
    pvalue<-sum(abs(a1$t-1) > abs(a1$t0-1))/(1+a1$R)  ## Bootstap P-value
    pvalue
    a1$t0
```

2. Example Data

at	ug	off	sex	pn
7.5	1	1	0	1
5	1	1	1	1
7.5	1	0	0	1
9	1	1	1	1
6.5	0	1	1	1
8.5	0	0	0	1
5.5	1	0	1	1
4	0	1	1	1
5	1	1	1	1
1	1	1	1	1
7	0	1	0	2
5	0	0	0	2
6	0	1	0	2
6	1	1	0	2
8	1	1	0	2
7	1	1	0	2

E. SAS Statistical Model and Output for Hourly Activity Profile

1. SAS Code and Data for Hourly Activity Profile

```
data faculty;
input time H W S R O ;
DATALINES;
1      0.992754     0.003623     0.000000     0.000000     0.003623
2      0.992754     0.003623     0.000000     0.000000     0.003623
3      0.996377     0.000000     0.000000     0.000000     0.003623
4      0.996377     0.000000     0.000000     0.000000     0.003623
5      0.996364     0.000000     0.000000     0.000000     0.003636
6      0.970480     0.018450     0.000000     0.007380     0.003690
7      0.839654     0.138077     0.000000     0.007464     0.011195
8      0.389706     0.544118     0.011029     0.025735     0.029412
9      0.185455     0.749091     0.010909     0.021818     0.032727
10     0.157509     0.783883     0.025641     0.010989     0.021978
11     0.153846     0.772894     0.014652     0.029304     0.029304
12     0.177122     0.601476     0.018450     0.154982     0.047970
13     0.169742     0.738007     0.022140     0.051661     0.018450
14     0.186813     0.725275     0.021978     0.043956     0.021978
15     0.170455     0.734848     0.007576     0.049242     0.037879
16     0.225092     0.627306     0.040590     0.051661     0.055351
17     0.496324     0.319853     0.055147     0.084559     0.044118
18     0.687500     0.128676     0.029412     0.128676     0.025735
19     0.724907     0.059480     0.037175     0.133829     0.044610
20     0.802920     0.051095     0.021898     0.102190     0.021898
21     0.897810     0.029197     0.007299     0.054745     0.010949
22     0.941818     0.025455     0.003636     0.021818     0.007273
23     0.960145     0.018116     0.000000     0.021739     0.000000
24     0.992754     0.003623     0.000000     0.000000     0.003623
;
run;
data faculty; set faculty; group=1;
*proc print data=faculty;run;
data student;
input time H W S R O ;
DATALINES;
1      0.890365     0.019934     0.000000     0.063123     0.026578
2      0.927393     0.009901     0.000000     0.036304     0.026403
3      0.933993     0.006601     0.000000     0.033003     0.026403
4      0.930693     0.006601     0.000000     0.033003     0.029703
5      0.930693     0.006601     0.000000     0.033003     0.029703
6      0.920792     0.013201     0.000000     0.039604     0.026403
7      0.852843     0.076923     0.000000     0.040134     0.030100
8      0.706667     0.196667     0.000000     0.053333     0.043333
9      0.550676     0.317568     0.006757     0.077703     0.047297
10     0.421769     0.459184     0.003401     0.064626     0.051020
11     0.338983     0.457627     0.027119     0.118644     0.057627
12     0.342282     0.419463     0.030201     0.124161     0.083893
13     0.312081     0.466443     0.026846     0.117450     0.077181
14     0.348123     0.488055     0.020478     0.075085     0.068259
15     0.369492     0.461017     0.027119     0.064407     0.077966
16     0.435811     0.378378     0.033784     0.081081     0.070946
17     0.494881     0.273038     0.040956     0.129693     0.061433
18     0.529010     0.170648     0.044369     0.187713     0.068259
```

```
19      0.559322     0.145763     0.040678     0.200000     0.054237
20      0.613559     0.108475     0.023729     0.203390     0.050847
21      0.653465     0.095710     0.003300     0.201320     0.046205
22      0.719472     0.066007     0.000000     0.181518     0.033003
23      0.813333     0.043333     0.000000     0.113333     0.030000
24      0.867110     0.023256     0.000000     0.079734     0.029900
;
run;

data student; set student; group=2;

data worker;
input time H W S R O ;
DATALINES;
1       0.954327     0.025240     0.000000     0.007212     0.013221
2       0.959085     0.020457     0.000000     0.007220     0.013237
3       0.957983     0.021609     0.000000     0.007203     0.013205
4       0.956783     0.021609     0.000000     0.008403     0.013205
5       0.944712     0.033654     0.000000     0.008413     0.013221
6       0.852619     0.114495     0.002436     0.012180     0.018270
7       0.637824     0.321384     0.011125     0.012361     0.017305
8       0.375309     0.572840     0.006173     0.019753     0.025926
9       0.279901     0.647349     0.017263     0.027127     0.028360
10      0.257426     0.669554     0.025990     0.021040     0.025990
11      0.245679     0.661728     0.038272     0.027160     0.027160
12      0.272503     0.524044     0.039457     0.076449     0.087546
13      0.237978     0.631319     0.039457     0.049322     0.041924
14      0.258663     0.639851     0.039604     0.034653     0.027228
15      0.303970     0.601737     0.024814     0.029777     0.039702
16      0.412346     0.458025     0.043210     0.050617     0.035802
17      0.643032     0.182152     0.052567     0.091687     0.030562
18      0.696341     0.102439     0.028049     0.146341     0.026829
19      0.696386     0.079518     0.031325     0.142169     0.050602
20      0.822464     0.055556     0.012077     0.079710     0.030193
21      0.884477     0.046931     0.002407     0.048135     0.018051
22      0.911164     0.045618     0.002401     0.028812     0.012005
23      0.937500     0.032452     0.000000     0.018029     0.012019
24      0.951923     0.027644     0.000000     0.007212     0.013221
;
run;

data worker; set worker; group=3;

data Eum;  set faculty student worker;
proc print; run;

proc univariate data=Eum;
var H W S R O;
histogram H W S R O;
run;

proc transpose data=eum out=eum1;
by  group time;
var H W S R O;
run;
proc print data=eum1;run;
data final(drop=_name_); set eum1; y=col1; subject=_name_; run;
```

```
proc print; run;


/* Activity Profile Modeling */

proc glm data=final;
class subject time;
*model y=  subject  time  subject*time/ intercept;
model y=  subject subject*time/intercept;
means subject subject*time;
lsmeans subject  subject*time;
run;
```

2. SAS Output

```
                              The GLM Procedure

Dependent Variable: y

                                         Sum of
      Source                   DF       Squares     Mean Square    F Value    Pr > F

      Model                   120    44.88627707     0.37405231      81.93    <.0001

      Error                   240     1.09572028     0.00456550

      Uncorrected Total       360    45.98199736


                      R-Square     Coeff Var      Root MSE        y Mean

                      0.965307      33.78594      0.067568      0.199990


      Source                   DF      Type I SS    Mean Square    F Value    Pr > F

      Intercept                 1    14.39855644    14.39855644    3153.77    <.0001
      subject                   4    20.02242936     5.00560734    1096.40    <.0001
      subject*time            115    10.46529127     0.09100253      19.93    <.0001


      Source                   DF    Type III SS    Mean Square    F Value    Pr > F

      Intercept                 1    14.39855644    14.39855644    3153.77    <.0001
      subject                   4    20.02242936     5.00560734    1096.40    <.0001
      subject*time            115    10.46529127     0.09100253      19.93    <.0001


                                                      Standard
          Parameter                 Estimate             Error    t Value    Pr > |t|

          Intercept             0.0181743333 B      0.03901069       0.47      0.6417
          activity      H       0.9190880000 B      0.05516944      16.66      <.0001
          activity      O       -.0025930000 B      0.05516944      -0.05      0.9626
```

activity	R	0.0108076667 B	0.05516944	0.20	0.8449
activity	S	-.0181743333 B	0.05516944	-0.33	0.7421
activity	W	0.0000000000 B	.	.	.
activity*time	H 1	0.0085530000 B	0.05516944	0.16	0.8769
activity*time	H 2	0.0224816667 B	0.05516944	0.41	0.6840
activity*time	H 3	0.0255220000 B	0.05516944	0.46	0.6441
activity*time	H 4	0.0240220000 B	0.05516944	0.44	0.6636
activity*time	H 5	0.0199940000 B	0.05516944	0.36	0.7174
activity*time	H 6	-.0226320000 B	0.05516944	-0.41	0.6820
activity*time	H 7	-.1604886667 B	0.05516944	-2.91	0.0040
activity*time	H 8	-.4467016667 B	0.05516944	-8.10	<.0001
activity*time	H 9	-.5985850000 B	0.05516944	-10.85	<.0001
activity*time	H 10	-.6583610000 B	0.05516944	-11.93	<.0001
activity*time	H 11	-.6910930000 B	0.05516944	-12.53	<.0001
activity*time	H 12	-.6732933333 B	0.05516944	-12.20	<.0001
activity*time	H 13	-.6973286667 B	0.05516944	-12.64	<.0001
activity*time	H 14	-.6727293333 B	0.05516944	-12.19	<.0001
activity*time	H 15	-.6559566667 B	0.05516944	-11.89	<.0001
activity*time	H 16	-.5795126667 B	0.05516944	-10.50	<.0001
activity*time	H 17	-.3925166667 B	0.05516944	-7.11	<.0001
activity*time	H 18	-.2996453333 B	0.05516944	-5.43	<.0001
activity*time	H 19	-.2770573333 B	0.05516944	-5.02	<.0001
activity*time	H 20	-.1909480000 B	0.05516944	-3.46	0.0006
activity*time	H 21	-.1253450000 B	0.05516944	-2.27	0.0240
activity*time	H 22	-.0797776667 B	0.05516944	-1.45	0.1495
activity*time	H 23	-.0336030000 B	0.05516944	-0.61	0.5430
activity*time	H 24	0.0000000000 B	.	.	.
activity*time	O 1	-.0011073333 B	0.05516944	-0.02	0.9840
activity*time	O 2	-.0011603333 B	0.05516944	-0.02	0.9832
activity*time	O 3	-.0011710000 B	0.05516944	-0.02	0.9831
activity*time	O 4	-.0000710000 B	0.05516944	-0.00	0.9990
activity*time	O 5	-.0000613333 B	0.05516944	-0.00	0.9991
activity*time	O 6	0.0005396667 B	0.05516944	0.01	0.9922
activity*time	O 7	0.0039520000 B	0.05516944	0.07	0.9430
activity*time	O 8	0.0173090000 B	0.05516944	0.31	0.7540
activity*time	O 9	0.0205466667 B	0.05516944	0.37	0.7099
activity*time	O 10	0.0174146667 B	0.05516944	0.32	0.7525
activity*time	O 11	0.0224490000 B	0.05516944	0.41	0.6844
activity*time	O 12	0.0575550000 B	0.05516944	1.04	0.2979
activity*time	O 13	0.0302703333 B	0.05516944	0.55	0.5837
activity*time	O 14	0.0235736667 B	0.05516944	0.43	0.6695
activity*time	O 15	0.0362676667 B	0.05516944	0.66	0.5116
activity*time	O 16	0.0384516667 B	0.05516944	0.70	0.4865
activity*time	O 17	0.0297896667 B	0.05516944	0.54	0.5897
activity*time	O 18	0.0246930000 B	0.05516944	0.45	0.6549
activity*time	O 19	0.0342350000 B	0.05516944	0.62	0.5355
activity*time	O 20	0.0187313333 B	0.05516944	0.34	0.7345
activity*time	O 21	0.0094870000 B	0.05516944	0.17	0.8636
activity*time	O 22	0.0018456667 B	0.05516944	0.03	0.9733
activity*time	O 23	-.0015750000 B	0.05516944	-0.03	0.9772
activity*time	O 24	0.0000000000 B	.	.	.
activity*time	R 1	-.0055370000 B	0.05516944	-0.10	0.9201
activity*time	R 2	-.0144740000 B	0.05516944	-0.26	0.7933
activity*time	R 3	-.0155800000 B	0.05516944	-0.28	0.7779
activity*time	R 4	-.0151800000 B	0.05516944	-0.28	0.7834
activity*time	R 5	-.0151766667 B	0.05516944	-0.28	0.7835
activity*time	R 6	-.0092606667 B	0.05516944	-0.17	0.8668

activity*time R 7	-.0089956667	B	0.05516944	-0.16	0.8706
activity*time R 8	0.0039583333	B	0.05516944	0.07	0.9429
activity*time R 9	0.0132340000	B	0.05516944	0.24	0.8106
activity*time R 10	0.0032363333	B	0.05516944	0.06	0.9533
activity*time R 11	0.0293873333	B	0.05516944	0.53	0.5948
activity*time R 12	0.0895486667	B	0.05516944	1.62	0.1059
activity*time R 13	0.0438290000	B	0.05516944	0.79	0.4277
activity*time R 14	0.0222493333	B	0.05516944	0.40	0.6871
activity*time R 15	0.0188266667	B	0.05516944	0.34	0.7332
activity*time R 16	0.0321376667	B	0.05516944	0.58	0.5608
activity*time R 17	0.0729976667	B	0.05516944	1.32	0.1870
activity*time R 18	0.1252613333	B	0.05516944	2.27	0.0241
activity*time R 19	0.1296840000	B	0.05516944	2.35	0.0196
activity*time R 20	0.0994480000	B	0.05516944	1.80	0.0727
activity*time R 21	0.0724180000	B	0.05516944	1.31	0.1906
activity*time R 22	0.0484006667	B	0.05516944	0.88	0.3812
activity*time R 23	0.0220516667	B	0.05516944	0.40	0.6897
activity*time R 24	0.0000000000	B	.	.	.
activity*time S 1	0.0000000000	B	0.05516944	0.00	1.0000
activity*time S 2	0.0000000000	B	0.05516944	0.00	1.0000
activity*time S 3	0.0000000000	B	0.05516944	0.00	1.0000
activity*time S 4	0.0000000000	B	0.05516944	0.00	1.0000
activity*time S 5	0.0000000000	B	0.05516944	0.00	1.0000
activity*time S 6	0.0008120000	B	0.05516944	0.01	0.9883
activity*time S 7	0.0037083333	B	0.05516944	0.07	0.9465
activity*time S 8	0.0057340000	B	0.05516944	0.10	0.9173
activity*time S 9	0.0116430000	B	0.05516944	0.21	0.8330
activity*time S 10	0.0183440000	B	0.05516944	0.33	0.7398
activity*time S 11	0.0266810000	B	0.05516944	0.48	0.6291
activity*time S 12	0.0293693333	B	0.05516944	0.53	0.5950
activity*time S 13	0.0294810000	B	0.05516944	0.53	0.5936
activity*time S 14	0.0273533333	B	0.05516944	0.50	0.6205
activity*time S 15	0.0198363333	B	0.05516944	0.36	0.7195
activity*time S 16	0.0391946667	B	0.05516944	0.71	0.4781
activity*time S 17	0.0495566667	B	0.05516944	0.90	0.3699
activity*time S 18	0.0339433333	B	0.05516944	0.62	0.5390
activity*time S 19	0.0363926667	B	0.05516944	0.66	0.5101
activity*time S 20	0.0192346667	B	0.05516944	0.35	0.7277
activity*time S 21	0.0043353333	B	0.05516944	0.08	0.9374
activity*time S 22	0.0020123333	B	0.05516944	0.04	0.9709
activity*time S 23	0.0000000000	B	0.05516944	0.00	1.0000
activity*time S 24	0.0000000000	B	.	.	.
activity*time W 1	-.0019086667	B	0.05516944	-0.03	0.9724
activity*time W 2	-.0068473333	B	0.05516944	-0.12	0.9013
activity*time W 3	-.0087710000	B	0.05516944	-0.16	0.8738
activity*time W 4	-.0087710000	B	0.05516944	-0.16	0.8738
activity*time W 5	-.0047560000	B	0.05516944	-0.09	0.9314
activity*time W 6	0.0305410000	B	0.05516944	0.55	0.5804
activity*time W 7	0.1606203333	B	0.05516944	2.91	0.0039
activity*time W 8	0.4197006667	B	0.05516944	7.61	<.0001
activity*time W 9	0.5531616667	B	0.05516944	10.03	<.0001
activity*time W 10	0.6193660000	B	0.05516944	11.23	<.0001
activity*time W 11	0.6125753333	B	0.05516944	11.10	<.0001
activity*time W 12	0.4968200000	B	0.05516944	9.01	<.0001
activity*time W 13	0.5937486667	B	0.05516944	10.76	<.0001
activity*time W 14	0.5995526667	B	0.05516944	10.87	<.0001
activity*time W 15	0.5810263333	B	0.05516944	10.53	<.0001

activity*time W 16	0.4697286667 B	0.05516944	8.51	<.0001
activity*time W 17	0.2401733333 B	0.05516944	4.35	<.0001
activity*time W 18	0.1157466667 B	0.05516944	2.10	0.0369
activity*time W 19	0.0767460000 B	0.05516944	1.39	0.1655
activity*time W 20	0.0535343333 B	0.05516944	0.97	0.3328
activity*time W 21	0.0391050000 B	0.05516944	0.71	0.4791
activity*time W 22	0.0275190000 B	0.05516944	0.50	0.6184
activity*time W 23	0.0131260000 B	0.05516944	0.24	0.8121
activity*time W 24	0.0000000000 B	.	.	.

F. SAS Output of Multinomial Logit Model for Destination Choice

WORK

The CATMOD Procedure
Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	7	43.50	<.0001
CREDIT	7	15.39	0.0313
RESIDENCE	7	18.55	0.0097
ttime	7	17.00	0.0174
age	7	22.95	0.0017
Likelihood Ratio	6E3	1280.13	1.0000

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	8.8461	2.5647	11.90	0.0006
	2	5.3247	2.6274	4.11	0.0427
	3	1.4022	3.1275	0.20	0.6539
	4	7.8205	3.6102	4.69	0.0303
	5	8.3718	3.3324	6.31	0.0120
	6	7.9237	3.3647	5.55	0.0185
	7	8.0961	3.3021	6.01	0.0142
CREDIT	1	-0.1285	0.1169	1.21	0.2717
	2	-0.1036	0.1207	0.74	0.3906
	3	-0.0296	0.1413	0.04	0.8339
	4	-0.3409	0.1445	5.57	0.0183
	5	-0.2524	0.1377	3.36	0.0667
	6	-0.2842	0.1439	3.90	0.0483
	7	-0.1392	0.1294	1.16	0.2821
RESIDENCE	1	0.9239	0.9399	0.97	0.3256
	2	1.0376	0.9646	1.16	0.2821
	3	1.1246	1.0903	1.06	0.3023
	4	2.8306	1.2354	5.25	0.0220
	5	-0.1425	1.1299	0.02	0.8996
	6	-1.3407	1.4660	0.84	0.3604
	7	1.8191	1.0235	3.16	0.0755
ttime	1	-0.0272	0.0327	0.70	0.4043
	2	-0.0883	0.0365	5.86	0.0155
	3	-0.00183	0.0398	0.00	0.9634
	4	-0.0303	0.0456	0.44	0.5060
	5	-0.00918	0.0410	0.05	0.8230
	6	-0.0241	0.0448	0.29	0.5906
	7	-0.0524	0.0393	1.78	0.1824
age	1	-0.0753	0.0362	4.34	0.0373
	2	-0.0218	0.0367	0.35	0.5514
	3	-0.0168	0.0469	0.13	0.7210
	4	-0.1432	0.0870	2.71	0.0997
	5	-0.1679	0.0764	4.83	0.0279
	6	-0.1342	0.0713	3.54	0.0598
	7	-0.1972	0.0841	5.50	0.0190

RECREATION

The CATMOD Procedure

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	6	24.00	0.0005
GENDER	6	35.44	<.0001
RESIDENCE	6	59.87	<.0001
ttime	6	5.82	0.4432
Likelihood Ratio	402	275.16	1.0000

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	2.9047	1.1176	6.76	0.0093
	2	1.0856	1.6104	0.45	0.5002
	3	3.1179	1.0820	8.30	0.0040
	4	1.2381	1.2085	1.05	0.3056
	5	0.4010	1.4960	0.07	0.7886
	6	2.5454	1.1177	5.19	0.0228
GENDER	1	0.8061	0.8521	0.89	0.3441
	2	2.8113	1.3865	4.11	0.0426
	3	2.2159	0.7941	7.79	0.0053
	4	1.6872	0.8463	3.97	0.0462
	5	0.9255	0.8920	1.08	0.2995
	6	2.4609	0.8391	8.60	0.0034
RESIDENCE	1	-1.4331	1.1057	1.68	0.1949
	2	-3.5128	1.5498	5.14	0.0234
	3	0.3339	1.0781	0.10	0.7568
	4	0.1482	1.1838	0.02	0.9003
	5	1.0255	1.4699	0.49	0.4854
	6	-1.5885	1.0956	2.10	0.1471
ttime	1	-0.0587	0.0379	2.39	0.1217
	2	-0.1445	0.1070	1.82	0.1770
	3	-0.0657	0.0297	4.88	0.0272
	4	-0.0617	0.0379	2.65	0.1035
	5	-0.0816	0.0473	2.98	0.0843
	6	-0.0629	0.0361	3.04	0.0814

OTHER

The CATMOD Procedure

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	7	12.83	0.0763
CREDIT	7	5.91	0.5499
ttime	7	5.43	0.6078
age	7	12.74	0.0786
Likelihood Ratio	581	265.99	1.0000

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	6.2477	3.7739	2.74	0.0978
	2	3.4735	5.8567	0.35	0.5531
	3	21.7019	9.3704	5.36	0.0206
	4	17.9458	6.2713	8.19	0.0042
	5	-9.7276	9.6589	1.01	0.3139
	6	8.2188	8.4522	0.95	0.3309
	7	7.2216	4.6559	2.41	0.1209
CREDIT	1	-0.1091	0.1481	0.54	0.4612
	2	0.0644	0.2179	0.09	0.7677
	3	-0.3115	0.2468	1.59	0.2068
	4	-0.2891	0.1842	2.46	0.1165
	5	0.1643	0.3489	0.22	0.6377
	6	-0.1931	0.2672	0.52	0.4699
	7	-0.2057	0.1748	1.38	0.2393
ttime	1	0.0498	0.0780	0.41	0.5234
	2	-0.0310	0.1061	0.09	0.7705
	3	-0.0222	0.0984	0.05	0.8216
	4	0.0323	0.0813	0.16	0.6910
	5	0.1951	0.1372	2.02	0.1551
	6	-0.0284	0.1295	0.05	0.8264
	7	-0.0242	0.0945	0.07	0.7974
age	1	-0.1351	0.0765	3.12	0.0775
	2	-0.1840	0.1737	1.12	0.2893
	3	-0.8008	0.3652	4.81	0.0284
	4	-0.6019	0.2128	8.00	0.0047
	5	0.0960	0.1451	0.44	0.5082
	6	-0.2587	0.2582	1.00	0.3165
	7	-0.1486	0.1016	2.14	0.1437

G. SAS Output of Multinomial Logit Model for Conditional Probability of Trips

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	345.0761	252	1.3693	<.0001
Pearson	332.3785	252	1.3190	0.0005

Number of unique profiles: 89

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	14228.356	13089.913
SC	14254.596	13772.159
-2 Log L	14220.356	12881.913

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1338.4434	100	<.0001
Score	1350.8347	100	<.0001
Wald	1101.9435	100	<.0001

The LOGISTIC Procedure

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
HOUR	92	915.0066	<.0001
STATUS	4	29.4364	<.0001
RESIDENT	4	122.7708	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		ACTIVITY	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		H	1	-1.6048	0.7944	4.0808	0.0434
Intercept		O	1	-1.6481	1.0666	2.3878	0.1223
Intercept		R	1	-0.4470	0.5116	0.7633	0.3823
Intercept		S	1	-0.7563	0.6912	1.1972	0.2739
HOUR	1	H	1	-11.3931	197.0	0.0033	0.9539
HOUR	1	O	1	-0.1105	1.4990	0.0054	0.9412
HOUR	1	R	1	0.0734	0.6869	0.0114	0.9149
HOUR	1	S	1	-1.1979	1.2568	0.9085	0.3405
HOUR	2	H	1	-11.3171	233.9	0.0023	0.9614
HOUR	2	O	1	-0.00807	1.5097	0.0000	0.9957
HOUR	2	R	1	0.2666	0.7484	0.1269	0.7217
HOUR	2	S	1	-13.9625	583.5	0.0006	0.9809
HOUR	3	H	1	-11.5061	478.7	0.0006	0.9808
HOUR	3	O	1	-12.5679	1002.6	0.0002	0.9900
HOUR	3	R	1	0.2477	1.1508	0.0463	0.8296

HOUR	3	S	1	-14.1201	1188.1	0.0001	0.9905
HOUR	4	H	1	-13.0323	766.0	0.0003	0.9864
HOUR	4	O	1	-13.8486	1759.2	0.0001	0.9937
HOUR	4	R	1	-1.1810	1.2691	0.8659	0.3521
HOUR	4	S	1	-0.1118	1.3540	0.0068	0.9342
HOUR	5	H	1	3.6356	1.3265	7.5114	0.0061
HOUR	5	O	1	1.9769	1.7713	1.2456	0.2644
HOUR	5	R	1	0.9861	1.3350	0.5456	0.4601
HOUR	5	S	1	-12.1436	609.2	0.0004	0.9841
HOUR	6	H	1	4.3158	1.0778	16.0348	<.0001
HOUR	6	O	1	1.1089	1.6231	0.4667	0.4945
HOUR	6	R	1	0.7617	1.0464	0.5299	0.4667
HOUR	6	S	1	-12.0870	428.0	0.0008	0.9775
HOUR	7	H	1	4.8669	0.9134	28.3929	<.0001
HOUR	7	O	1	3.2443	1.1779	7.5870	0.0059
HOUR	7	R	1	1.5662	0.7120	4.8394	0.0278
HOUR	7	S	1	0.2754	0.9996	0.0759	0.7829
HOUR	8	H	1	3.0019	0.8127	13.6440	0.0002
HOUR	8	O	1	1.2407	1.1021	1.2674	0.2603
HOUR	8	R	1	-0.5698	0.5852	0.9481	0.3302
HOUR	8	S	1	-0.8969	0.7974	1.2651	0.2607
HOUR	9	H	1	2.1413	0.8010	7.1465	0.0075
HOUR	9	O	1	0.2808	1.0925	0.0660	0.7972
HOUR	9	R	1	-0.9497	0.5424	3.0659	0.0799
HOUR	9	S	1	-1.5948	0.7651	4.3449	0.0371
HOUR	10	H	1	1.2322	0.8014	2.3638	0.1242
HOUR	10	O	1	0.1219	1.0888	0.0125	0.9108
HOUR	10	R	1	-1.6338	0.5476	8.9024	0.0028
HOUR	10	S	1	-1.8122	0.7627	5.6451	0.0175
HOUR	11	H	1	0.5467	0.8000	0.4670	0.4944
HOUR	11	O	1	-0.1531	1.0813	0.0201	0.8874
HOUR	11	R	1	-1.2668	0.5234	5.8568	0.0155
HOUR	11	S	1	-2.0475	0.7378	7.7014	0.0055
HOUR	12	H	1	0.5177	0.7993	0.4195	0.5172
HOUR	12	O	1	-0.5413	1.0853	0.2487	0.6180
HOUR	12	R	1	-0.9052	0.5179	3.0545	0.0805
HOUR	12	S	1	-1.9240	0.7254	7.0357	0.0080
HOUR	13	H	1	1.0001	0.8040	1.5473	0.2135
HOUR	13	O	1	0.4370	1.0837	0.1626	0.6868
HOUR	13	R	1	0.2097	0.5199	0.1628	0.6866
HOUR	13	S	1	-1.2544	0.7310	2.9448	0.0862
HOUR	14	H	1	0.5107	0.8014	0.4061	0.5240
HOUR	14	O	1	-0.1953	1.0849	0.0324	0.8571
HOUR	14	R	1	-1.3215	0.5266	6.2973	0.0121
HOUR	14	S	1	-1.9795	0.7425	7.1086	0.0077
HOUR	15	H	1	0.6229	0.8037	0.6007	0.4383
HOUR	15	O	1	0.2786	1.0818	0.0663	0.7967
HOUR	15	R	1	-1.4409	0.5367	7.2086	0.0073
HOUR	15	S	1	-2.0073	0.7622	6.9361	0.0084
HOUR	16	H	1	1.0458	0.8060	1.6834	0.1945
HOUR	16	O	1	0.5768	1.0845	0.2829	0.5948
HOUR	16	R	1	-1.3118	0.5510	5.6687	0.0173
HOUR	16	S	1	-1.3299	0.7422	3.2109	0.0731
HOUR	17	H	1	1.3257	0.8002	2.7449	0.0976
HOUR	17	O	1	0.0915	1.0859	0.0071	0.9328
HOUR	17	R	1	-1.0024	0.5304	3.5727	0.0587
HOUR	17	S	1	-1.0548	0.7157	2.1719	0.1405
HOUR	18	H	1	1.8210	0.8053	5.1135	0.0237
HOUR	18	O	1	0.6221	1.0938	0.3235	0.5695
HOUR	18	R	1	0.3134	0.5280	0.3522	0.5529
HOUR	18	S	1	-1.1327	0.7594	2.2249	0.1358
HOUR	19	H	1	1.4124	0.8071	3.0629	0.0801
HOUR	19	O	1	0.6208	1.0923	0.3230	0.5698
HOUR	19	R	1	0.5336	0.5250	1.0332	0.3094
HOUR	19	S	1	-0.5718	0.7271	0.6184	0.4317

HOUR	20	H	1	1.3587	0.8243	2.7170	0.0993
HOUR	20	O	1	0.4558	1.1257	0.1639	0.6856
HOUR	20	R	1	0.6229	0.5449	1.3069	0.2530
HOUR	20	S	1	-0.0796	0.7407	0.0116	0.9144
HOUR	21	H	1	1.5569	0.8198	3.6065	0.0576
HOUR	21	O	1	0.7449	1.1129	0.4481	0.5032
HOUR	21	R	1	0.3275	0.5469	0.3586	0.5493
HOUR	21	S	1	-0.1633	0.7450	0.0480	0.8265
HOUR	22	H	1	1.3784	0.8440	2.6673	0.1024
HOUR	22	O	1	0.4561	1.1719	0.1515	0.6971
HOUR	22	R	1	0.4028	0.5728	0.4945	0.4819
HOUR	22	S	1	-0.1914	0.7946	0.0580	0.8097
HOUR	23	H	1	1.0399	0.8466	1.5087	0.2193
HOUR	23	O	1	1.1618	1.1223	1.0717	0.3006
HOUR	23	R	1	-0.1596	0.5808	0.0755	0.7835
HOUR	23	S	1	-0.7331	0.8376	0.7660	0.3815
STATUS	grad	H	1	-0.2718	0.1107	6.0235	0.0141
STATUS	grad	O	1	-0.2616	0.1626	2.5888	0.1076
STATUS	grad	R	1	-0.5774	0.1435	16.1974	<.0001
STATUS	grad	S	1	0.3928	0.1825	4.6340	0.0313
RESIDENT	on	H	1	0.3492	0.0805	18.8421	<.0001
RESIDENT	on	O	1	-0.6292	0.1296	23.5657	<.0001
RESIDENT	on	R	1	0.6773	0.0949	50.9063	<.0001
RESIDENT	on	S	1	-0.4615	0.1670	7.6400	0.0057

Odds Ratio Estimates

Effect	ACTIVITY	Point Estimate	95% Wald Confidence Limits	
HOUR 1 vs 0	H	<0.001	<0.001	>999.999
HOUR 1 vs 0	O	0.895	0.047	16.901
HOUR 1 vs 0	R	1.076	0.280	4.136
HOUR 1 vs 0	S	0.302	0.026	3.544
HOUR 2 vs 0	H	<0.001	<0.001	>999.999
HOUR 2 vs 0	O	0.992	0.051	19.122
HOUR 2 vs 0	R	1.306	0.301	5.660
HOUR 2 vs 0	S	<0.001	<0.001	>999.999
HOUR 3 vs 0	H	<0.001	<0.001	>999.999
HOUR 3 vs 0	O	<0.001	<0.001	>999.999
HOUR 3 vs 0	R	1.281	0.134	12.223
HOUR 3 vs 0	S	<0.001	<0.001	>999.999
HOUR 4 vs 0	H	<0.001	<0.001	>999.999
HOUR 4 vs 0	O	<0.001	<0.001	>999.999
HOUR 4 vs 0	R	0.307	0.026	3.693
HOUR 4 vs 0	S	0.894	0.063	12.707
HOUR 5 vs 0	H	37.925	2.817	510.587
HOUR 5 vs 0	O	7.220	0.224	232.428
HOUR 5 vs 0	R	2.681	0.196	36.698
HOUR 5 vs 0	S	<0.001	<0.001	>999.999
HOUR 6 vs 0	H	74.872	9.056	619.049
HOUR 6 vs 0	O	3.031	0.126	72.971
HOUR 6 vs 0	R	2.142	0.275	16.655
HOUR 6 vs 0	S	<0.001	<0.001	>999.999
HOUR 7 vs 0	H	129.923	21.688	778.318
HOUR 7 vs 0	O	25.644	2.549	257.976
HOUR 7 vs 0	R	4.789	1.186	19.331
HOUR 7 vs 0	S	1.317	0.186	9.344
HOUR 8 vs 0	H	20.123	4.092	98.958
HOUR 8 vs 0	O	3.458	0.399	29.984
HOUR 8 vs 0	R	0.566	0.180	1.781
HOUR 8 vs 0	S	0.408	0.085	1.946
HOUR 9 vs 0	H	8.511	1.771	40.905
HOUR 9 vs 0	O	1.324	0.156	11.269
HOUR 9 vs 0	R	0.387	0.134	1.120
HOUR 9 vs 0	S	0.203	0.045	0.909

HOUR	10 vs 0	H	3.429	0.713	16.493
HOUR	10 vs 0	O	1.130	0.134	9.545
HOUR	10 vs 0	R	0.195	0.067	0.571
HOUR	10 vs 0	S	0.163	0.037	0.728
HOUR	11 vs 0	H	1.728	0.360	8.287
HOUR	11 vs 0	O	0.858	0.103	7.143
HOUR	11 vs 0	R	0.282	0.101	0.786
HOUR	11 vs 0	S	0.129	0.030	0.548
HOUR	12 vs 0	H	1.678	0.350	8.039
HOUR	12 vs 0	O	0.582	0.069	4.884
HOUR	12 vs 0	R	0.404	0.147	1.116
HOUR	12 vs 0	S	0.146	0.035	0.605
HOUR	13 vs 0	H	2.718	0.562	13.141
HOUR	13 vs 0	O	1.548	0.185	12.948
HOUR	13 vs 0	R	1.233	0.445	3.417
HOUR	13 vs 0	S	0.285	0.068	1.195
HOUR	14 vs 0	H	1.666	0.346	8.015
HOUR	14 vs 0	O	0.823	0.098	6.896
HOUR	14 vs 0	R	0.267	0.095	0.749
HOUR	14 vs 0	S	0.138	0.032	0.592
HOUR	15 vs 0	H	1.864	0.386	9.008
HOUR	15 vs 0	O	1.321	0.159	11.011
HOUR	15 vs 0	R	0.237	0.083	0.678
HOUR	15 vs 0	S	0.134	0.030	0.598
HOUR	16 vs 0	H	2.846	0.586	13.812
HOUR	16 vs 0	O	1.780	0.213	14.915
HOUR	16 vs 0	R	0.269	0.091	0.793
HOUR	16 vs 0	S	0.264	0.062	1.133
HOUR	17 vs 0	H	3.765	0.785	18.066
HOUR	17 vs 0	O	1.096	0.130	9.205
HOUR	17 vs 0	R	0.367	0.130	1.038
HOUR	17 vs 0	S	0.348	0.086	1.416
HOUR	18 vs 0	H	6.178	1.275	29.944
HOUR	18 vs 0	O	1.863	0.218	15.895
HOUR	18 vs 0	R	1.368	0.486	3.851
HOUR	18 vs 0	S	0.322	0.073	1.427
HOUR	19 vs 0	H	4.106	0.844	19.970
HOUR	19 vs 0	O	1.860	0.219	15.825
HOUR	19 vs 0	R	1.705	0.609	4.771
HOUR	19 vs 0	S	0.565	0.136	2.347
HOUR	20 vs 0	H	3.891	0.773	19.574
HOUR	20 vs 0	O	1.577	0.174	14.327
HOUR	20 vs 0	R	1.864	0.641	5.424
HOUR	20 vs 0	S	0.923	0.216	3.944
HOUR	21 vs 0	H	4.744	0.951	23.660
HOUR	21 vs 0	O	2.106	0.238	18.654
HOUR	21 vs 0	R	1.388	0.475	4.053
HOUR	21 vs 0	S	0.849	0.197	3.658
HOUR	22 vs 0	H	3.969	0.759	20.751
HOUR	22 vs 0	O	1.578	0.159	15.689
HOUR	22 vs 0	R	1.496	0.487	4.598
HOUR	22 vs 0	S	0.826	0.174	3.920
HOUR	23 vs 0	H	2.829	0.538	14.869
HOUR	23 vs 0	O	3.196	0.354	28.831
HOUR	23 vs 0	R	0.852	0.273	2.661
HOUR	23 vs 0	S	0.480	0.093	2.481
STATUS	grad vs Under	H	0.762	0.613	0.947
STATUS	grad vs Under	O	0.770	0.560	1.059
STATUS	grad vs Under	R	0.561	0.424	0.744
STATUS	grad vs Under	S	1.481	1.036	2.118
RESIDENT	on vs off	H	1.418	1.211	1.660
RESIDENT	on vs off	O	0.533	0.413	0.687
RESIDENT	on vs off	R	1.969	1.634	2.371
RESIDENT	on vs off	S	0.630	0.454	0.874

H. Pedestrian Data Collection Work Sheet

DATA COLLECTION WORK SHEET

DATE: ____________________ Name : ____________________

Building Name: ____________________

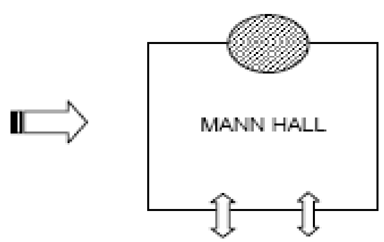
Time Period: ____________________

Sketch Survey Location

Counts

Hour	minutes	# IN	# OUT
	15		
	30		
	45		
	60		

Example sketch



* : Entrance (door) * : surveyed