Making the housing-transportation connection: University of Montana student residential patterns and commuting practices

John H. Harland
The University of Montana

Follow this and additional works at: https://scholarworks.umt.edu/etd
Let us know how access to this document benefits you.

Recommended Citation
https://scholarworks.umt.edu/etd/9214

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
Permission is granted by the author to reproduce this material in its entirety, provided that this material is used for scholarly purposes and is properly cited in published works and reports.

**Please check "Yes" or "No" and provide signature**

Yes, I grant permission  

No, I do not grant permission

Author's Signature: 

Date: 5/18/2001

Any copying for commercial purposes or financial gain may be undertaken only with the author's explicit consent.
MAKING THE HOUSING-TRANSPORTATION CONNECTION

UNIVERSITY OF MONTANA STUDENT RESIDENTIAL PATTERNS AND COMMUTING PRACTICES

by

John H. Harland, III

B.A. The University of Montana, USA, 1999

presented in partial fulfillment of the requirements for the degree of

Master of Arts

The Department of Geography

The University of Montana

May 2001

Approved by:

Chairperson

Dean, Graduate School

Date
College students are rarely the focus of transportation and housing research. The purpose of this thesis is to study the relationship between University of Montana student residential patterns and commuting practices in a small American city - Missoula Montana. This relationship is addressed by investigating four questions: a) Where do students live? b) Why do students live where they do? c) How important are commuting considerations in student residential decisions? d) How do residential patterns affect commuting practices? GIS was used to locate over four hundred off-campus respondents to a survey on housing and transportation, showing where students live. Linear regression was used to test how residence price is predicted by residence size and distance to campus. Analysis of variance was used to test whether some household types prefer proximity to campus, paying a premium per square foot, while other household types prefer to pay less per square foot sacrificing accessibility. The linear regression and analysis of variance procedures examine why students live where they do, and the importance of accessibility in student residential decisions. Logistic regression was used to assess the impacts of residential location on commuting mode. The logistic regression procedure investigated how residential patterns affect commuting practices.

The analysis found that residence price varies with residence size, increasing about twenty-six cents per square foot. Distance is not a significant predictor of residence price, probably due to non-random distribution of housing quality and subtle effects of distance overall. When analyzed by demographic group distance to campus is more important, with family household types preferring more residence for the dollar over accessibility and non-family household types preferring accessibility but paying a higher residence price per square foot. Thus, students live where they do primarily due to the size of the residence, but accessibility is important for certain demographic groups. Residential location affects commuting practices. Residential location was the most important variable for predicting alternative versus car commuting modes. In summary, the relationship between residential patterns and commuting practices is shown to be very strong.
Dedicated to Elysia
ACKNOWLEDGMENTS

There is one person, in particular, I would like to thank for the production of this document, Dr. Christiane von Reichert. Her dedication to education and research in general, and her dedication to my education and research in particular, have left me awestruck on several occasions. Above all else, she has shown me that fortitude and dedication will lead to personal satisfaction, and this is more than enough. I hope to work hard and profit from this lesson for years to come. I would also like to thank my other committee members, Dr. Evan Denney and Dr. Jeff Bookwalter.

There are a couple organizations I would also like to thank. I would like to thank the Missoula City-County Office of Planning and Grants and the Associated Students of the University of Montana Office of Transportation for their generous contributions, without which this research would not have been possible. I would also like to thank AZ TECH graphic design for donating time and energy to prepare the survey advertising.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................ iv  
LIST OF ILLUSTRATIONS ............................................................................................... vii  
LIST OF TABLES ................................................................................................................ viii  
1. INTRODUCTION ............................................................................................................ viii  
2. CONCEPTUAL BACKGROUND AND HYPOTHESES .............................................. 4  
   Conventional Conceptualizations .............................................................................. 4  
   Conceptualizing the Housing-Transportation Connection .......................................... 8  
   Hypotheses to be Tested ............................................................................................. 10  
   Applying the Alonso Model to Student Residential Patterns .................................. 10  
   Space versus Accessibility for Different Household Types ....................................... 13  
   Residential Location and Commuting Practices ....................................................... 16  
   Confounding Variables .............................................................................................. 18  
   Hypotheses Summarized ........................................................................................... 21  
3. METHODOLOGY ......................................................................................................... 22  
   Data Collection ........................................................................................................ 22  
   Sample Selection ..................................................................................................... 23  
   Survey Design .......................................................................................................... 28  
   Data Preparation ...................................................................................................... 31  
   Data Analysis ........................................................................................................... 33  
   Data Description ...................................................................................................... 33  
   Linear Regression .................................................................................................... 34  

v
LIST OF ILLUSTRATIONS

Figure 3.1 - Pared-down Registrar's Database Distance Distribution ..................... 27
Figure 3.2 - Survey Respondent Distance Distribution ........................................... 27
Figure 3.3 - Histogram of Residuals for Linear Regression ................................. 36
Figure 3.4 - Percent Car Distribution ................................................................. 45
Figure 3.5 - Linearity of Distance Plot ............................................................... 47
Figure 4.1 - Ownership Status ........................................................................... 51
Figure 4.2 - Demographic Groups ................................................................. 52
Figure 4.3 - Ownership Status by Demographic Group ...................................... 53
Figure 4.4 - Percent Distribution of Housing Types .............................................. 55
Figure 4.5 - Spatial Distribution of Housing Types .............................................. 55
Figure 4.6 - Commuting Practices by Mode ....................................................... 58
Figure 4.7 - Percent Distribution of Job Location ............................................... 61
Map 4.1 - Ownership Status ............................................................................ 66
Map 4.2 - Rent plus Utilities Map ................................................................. 67
Map 4.3 - Respondent's Residences Size Map ................................................... 68
Figure 4.8 - Basic Distance Model Probability Plot ............................................. 81
Map 4.4 - Respondent's Mode of Transportation Map ....................................... 83
Figure 4.9 - Distance and Parking Permit Probability Plot .................................. 88
Figure 4.10 - Presence of Children in the Household Probability Plot .................. 89
Figure 4.11 - No Bike Ownership Probability Plot ............................................ 89
Figure 4.12 - Distance Model versus Final Model Probability Plots ..................... 91
LIST OF TABLES

Table 4.1 - General Descriptive Statistics .............................................................. 50
Table 4.2 - Linear Regression Models for the Alonso Theoretical Model .......... 64
Table 4.3 - Residence Price Evaluations ............................................................... 70
Table 4.4 Demographic Group Analysis of Variance ............................................. 73
Table 4.5 - Bonferroni Post-Hoc Mean Difference Tests .................................... 75
Table 4.6 - Bonferroni Post-Hoc Mean Difference Tests
  (Live with Children Only Omitted) ................................................................. 76
Table 4.7 - Basic Distance Logistic Regression Model ........................................ 80
Table 4.8 - Calculated Probabilities for Commuting by Car at Selected Distances ... 82
Table 4.9 - Logistic Regression Model All Variables with a Possible Effect ........ 85
Table 4.10 - Final Model Development ............................................................... 87
Table 4.11 - Final Logistic Model ...................................................................... 90
CHAPTER 1

INTRODUCTION

Billions of dollars each year are spent on housing choice, employment siting, and transportation. There is enormous economic incentive to effectively manage the allocation of housing, location of employment opportunities, and transportation resources. A 1992 University of Montana study found that, “students are spending 13.6 percent more for their residences than they should. Students tend to search for housing for several weeks before finding suitable housing and are too likely to rent a unit because of availability rather than its appropriateness for them. Thus many students are living in units that are too large or too small, too hard to heat and obsolete to all but the poorest Missoulians (McQuiston 1992, 2).” Meanwhile, the Associated Students at the University of Montana have created an Office of Transportation whose main purpose is to combat the mounting traffic problems and congestion around campus.

University students comprised 7 percent of the United States population over the age of fifteen in 1998 (U.S. Bureau of the Census 1998). University students are unique in the study of housing choice, job location, and transportation practices. Schoolwork
and classes are a form of 'employment' for university students. Thus, for the
preponderance of this research school location will be treated as the primary 'job
location', and additional employment location will refer to the location of any job a
student maintains in addition to school. Another unique feature is that school location is
fixed to a much higher degree than 'regular job location'. Unless a student lives in a
major metropolitan area with several universities, school location is fixed during
enrollment. University students are also unique in that the job of school is a source of
expenditure, not a source of income. This presents different economic circumstances
than presumed by traditional research on housing choice. University students are also
unique in that school location is congruent with housing location for dorm residents.
Aside from the armed services, few individuals in this country now reside on the same
grounds as their 'job'.

Most research has rarely considered the university student in housing and
transportation research. One exception is a study of University of Montana students
(McQuiston 1992), which sought mainly to describe the particularities of student housing
within the Missoula area.

In addition to describing student housing choice, this thesis explains student
residential patterns and commuting practices. The purpose of this thesis is to understand
the housing-transportation relationship by identifying and analyzing critical variables that
shape University of Montana student housing and transportation.

A housing and commuting practices survey was designed specifically for the
research in this thesis, with over 400 off-campus respondents describing student
residential choices and commuting practices. Hypotheses will be developed from several sources including Alonso's *Location and Land Use* (1964), a 1992 study of student residential choices (McQuiston 1992), and contemporary transportation literature. Hypothesis development will construct theoretical expectations for the relationship between residential decisions and proximity to campus, and for the effects residential decisions have on commuting practices. Hypothesis testing will allow these relationships to be verified, explaining connections between student residential patterns and commuting practices.

Administrators at The University of Montana and City of Missoula officials confront a variety of challenges that stem from student residential decisions and commuting practices. There are number of groups and agencies in Missoula that would like to see changes in where students live and how they get around. In short, student residential decisions and congestion around campus are often treated as separate issues. The problem addressed by this research is to provide basic data and explanations for why students live where they do and why some students choose one commuting mode over another. By investigating the connections between student residential patterns and commuting practices, discussions of student housing and transportation policies may begin with a realistic view of how these two issues relate.
CHAPTER 2

CONCEPTUAL BACKGROUND AND HYPOTHESES

Many research approaches have conceptualized the relationships between housing choices and transportation practices. These approaches inform the research question: is there a connection between student residential patterns and commuting practices, and if so, how should this connection be understood? Three specific hypotheses have been developed and tested to address this question.

Conventional Conceptualizations

Alonso (1964), in *Location and Land Use*, examined urban residential patterns using the assumptions that accessibility to the city center was desirable and the supply of land was limited. This conceptualization of the relationship between employment location and residential location was an extension of the 'featureless plain' first applied to agriculture in the classical economic geography of von Thuened (Thuened 1842, 7), which described how transportation costs predict 'land rent'. Hence, the value of land increased as proximity to the city center increased. In turn, the increased value of land near the city center resulted in a greater intensity of land use. Density gradients generally follow a
downward-sloping convex curve from city centers (Muth 1969). The ideal city, from the perspective of transportation in the Alonso model, was a compact city. High intensity land uses result in a steep density gradient from the employment center outward, and both the countryside and the city center were accessible (Alonso 1964).

In 1987, Simpson introduced an alternative model based upon the converse of the assumptions made by Alonso. Simpson maintained that residential location was fixed and employment location was variable so that, "employment searches are spatially systematic; that is, workers prefer jobs closer to their place of residence, other factors such as skill or occupation being equal,"(Simpson 1987, 119-128). Still, the Simpson job search model does not change implications for commuting practices since a compact city still allows for greater employment opportunities within a small geographic area, and thus an employment change would not increase commute distance significantly.

Even though Alonso maintained employment location was fixed and residential location was variable and Simpson maintained the converse, both approaches recognize the important relationship between residential location and employment location. The relationship between residential location and employment location is the foundation for the jobs-housing balance debate. This debate has been prompted by the relocation of employment centers to suburban areas. Jobs-housing balance refers to the spatial relationship between the number of jobs and housing units within a given geographic area (Peng 1997, 1216). Where there are roughly equivalent distributions of housing and employment, places will have lower levels of traffic and housing markets will better suit the needs of all local employees. Jobs-housing balance has the potential to, "produce
well-defined commuter sheds where local neighborhood traffic is segregated from regional through traffic.”(Cervero 1989, 136-150)

The jobs-housing debate has sparked conflict associated with conventional approaches, but for the most part this conflict has focused on large metropolitan areas. For example, Hamilton rightly pointed out that the Alonso model predicted densities, but did not necessarily predict the associated rent gradient (Hamilton 1982). Using a standard monocentric optimization model, Hamilton also showed that, in fourteen major US cities, the Alonso model underestimates actual commuting by at least five times. Hamilton showed that random employment and residential patterns have a stronger predictive power than the monocentric city model. The problems identified by Hamilton tend to be the result of the decentralization of employment centers. These problems do not present a problem for this research, because the study area has not experienced widespread employment center and residential decentralization.

Conventional approaches have been refined as well, with many having predictive power under an array of real world conditions. Zhang-Ren Peng has applied, “geographical information system techniques in a piece wise, nonlinear model- spline functions- to analyze empirically the relationship between the jobs-housing ratio and urban commuting patterns,” (Peng 1997, 1215). Hai Yang and Qiang Meng have developed, “a mixed, combined, and stochastic user-equilibrium model for urban location and travel choices with variable origin and destination costs,”(Yang and Meng 1998, 575). These are examples of sophisticated models addressing the relationship between residential patterns and commuting practices, but they are aimed at metropolitan areas.
The relationships between residential patterns and commuting practices are characterized by two conceptual assumptions in the research literature. First, the geographies of residences and centers of employment will affect commuting practices. Second, economic considerations are crucial when analyzing the relationship between residential patterns, employment centers, and commuting practices. These two assumptions are central conceptual elements for this thesis.

However, two presumptions rarely mentioned in the literature exist. This first is that motorized transportation is the preeminent mode of commuting so non-motorized modes may be excluded. This presumption may be valid for large metropolitan areas where employment center and residential patterns have been decentralized creating imbalances between housing and employment. This assumption will not be employed because Missoula is not a large metropolis, and fits the monocentric city model. Moreover, students at the University of Montana use a variety of commuting modes. Each day hundreds of students walk to school. On a nice day, bike racks are crammed to capacity. The Associated Students at the University of Montana have created an office whose primary goal is to increase alternative transportation. All these considerations point to the importance of studying non-car transportation modes.

The second presumption in conventional conceptualizations is that economic considerations are paramount and household demographics are subordinate. Many University of Montana student households do not fit traditional notions of college student household demographics. These diverse household types result in important differences
in residential decision making (McQuiston 1992). This thesis will maintain that both economic considerations and household demographics are important.

There are two primary ways in which the connection between residential patterns and commuting practices could be analyzed. The first is to assume that transportation decisions are paramount and that all residential decisions are made in terms of the impacts they will have on commuting practices. None of the research, presented in the first section, approaches the problem in this way. The alternative approach expects residential decisions to be made first, and transportation to be one of many factors important in the residential choice process. This is the approach taken by the preponderance of research on this topic. This approach means residential patterns should be understood first, and commuting practices should be treated as a resultant of residential decisions.

**Conceptualizing the Housing-Transportation Connection**

The goal of this study is to identify if there is a connection between student residential patterns and commuting practices, and if so how this connection can be understood. Specifically: a) Where do students live? b) Why do students live where they do? c) How important are commuting considerations for residential decisions? d) What are the impacts of residential decisions on commuting practices?

A few definitions are helpful. Residential patterns refer to the spatial distribution of students. Residential decisions and residential choices are simply the individual student
choices that produce residential patterns. Commuting practices are defined as how students get to campus.

Some research has sought to quantify the importance of commuting in residential decisions. This has been accomplished by looking at communities with different jobs-housing ratios, and then building a model designed to assess whether the jobs-housing ratio was important in choosing an area in which to live. Jonathan Levine found that people of low income exhibited preferences for places where the jobs-housing balance was high (Levine 1998, 147). Students generally have lower incomes so students may be expected to be sensitive to transportation factors when making residential decisions. Conceptualizing the influence of accessibility on residential decisions is necessary, but not sufficient. It only explains why students live where they live, and how important transportation is for residential decision making. Assessing the effects of residential decisions on commuting practices is also necessary.

Connecting the impacts of residential patterns on commuting practices can be done in a number of ways, but two ways tend to dominate the literature. One is to look at transportation in terms of vehicle miles traveled. This is the approach used by Zhong-Ren Peng in his article, “The Jobs-Housing Balance and Urban Commuting,” (Peng 1997). However, this approach is best suited to large U. S. metropolitan areas where the vast majority of commuting is done by car.

Another approach is to analyze modal split, which is percent of travel by mode. Travel mode simply refers to the means of travel, such as: car, bus, walk, etc. This approach is better suited to places that resemble the compact monocentric city, because
the overall proximity of goods and services is higher, making transportation modes such as biking and walking easier. An effective example of modal split analysis, for a compact city, is given in a study of Oslo by Peter Naess and Synnove Lyssand Sandberg (Naess and Sandberg 1996). Thus modal split will be the measure of commuting practices employed in this thesis, because Missoula in general, and the University in particular, are well characterized by the compact central city.

In summary, investigation of the research question should begin with student residential patterns. This should be followed by an assessment of the impacts of residential patterns on commuting practices. The first step in the investigation is to describe where students live. Once the location of student residences has been described, hypotheses should attempt to answer three basic questions: a) Why do students live where they do? b) How important is commuting to campus for explaining where students live and why they live there? c) What are the impacts of residential patterns on commuting practices?

**Hypotheses to be Tested**

**Applying the Alonso Model to Student Residential Decisions**

The first hypothesis to be tested is essentially a reformulation of the Alonso theoretical model. The original Alonso model assumes that residential decisions are made on a featureless plain with a city in the center of the plain. It assumes that all land is of equal innate utility. Based on these assumptions, the cost per unit of land area can be calculated using transportation costs to the city center. Accessible land commands a
higher price due to lower transportation costs, while more distant parcels are less valuable due to lower accessibility and higher transportation costs.

Alonso rightly points out that residential decisions are not a matter of minimizing costs nor a matter of minimizing transportation costs. If this were the case then people would agree to live in as small a space as possible, maximally compacting a city. Alonso writes, “The minimum aggregate costs of friction are thus not the low point of a continuous function, but a unique value, once the ‘rules of the game’ are given (Alonso 1964, 103).”

Application of the Alonso model requires two steps. First, the ‘rules of the game’ refer to consumer behavior according with the expectations of the model. As distance increases or residence size decreases, ceteris paribus, residence price will diminish. For the purposes of this thesis, can student residential decisions be explained by distance to campus and residence size as the sole predictors? The second step is contingent upon the first. If this distance and residence size are good predictors of residence price, then how does residence price change with distance and residence size to maximize student satisfaction. Some modifications of the measurements used to develop the Alonso model are necessary.

This thesis is concerned with total residence price. Total residence price is appropriate, because this thesis is seeking to explain student residential patterns. It is important to be able to distinguish between the effects of a four-bedroom house and a one-bedroom apartment. If the price is strictly per square foot, no such distinction will persist. Alonso defines total price as, “the price of land times the quantity of land in a site
represents the payment for use of the site (Alonso 1964, 16).” This is the definition that will be employed for residence price.

Also, the Alonso model operates in terms of land. This thesis will use residence size instead of land for this thesis because usable data for residence size was easier to acquire. The Alonso model uses the land per occupant as the unit of measurement. This thesis will use the individual residence size as the unit of measurement. Practically speaking, this difference is fairly non-restrictive because price is still predicted by location and size. Also, these units of measure can be expected to generally vary together. Big houses are usually located on big lots, dividing the lot size or the residence size by a single household will result in a large number. Apartment buildings are rarely located on forty acres, dividing a typical apartment lot by a high number of households, such as 20, will result in a small number. Most all twenty-unit apartment complexes in Missoula, for instance, have small residences.

It is now possible to formulate the hypothesis concisely. Assuming residential quality is constant, **student household residential price can be predicted by distance to the university in miles and residence size in square feet.**

It should also be clear how testing this hypothesis will promote the discussion of the research question. It has been argued that examination of the research question should begin by understanding residential patterns. The Alonso model will be used to explain why students live where they do by testing how distance to campus and residence size effect residence price. Testing this hypothesis will also address the importance of commuting costs for explaining where students live and why they live there. The more
important distance is in predicting residence price the more important transportation costs are in describing student residential decisions.

Space versus Accessibility for Different Household Types

Unlike the hypothesis developed from the Alonso model, the hypothesis related to demographics will not be developed strictly through the use of a long-standing theory. This hypothesis will be developed from one part theory and one part empirical observation. The theoretical basis stems from a consequence of the Alonso model. The empirical basis is the product of a 1992 study on University of Montana student housing (McQuiston 1992).

If the general form of the Alonso model is correct, then one consequence of the model is that residential choice can be conceptualized as a trade-off between residential space and accessibility to the University, when controlling for residence price. Reformulated, this means that proximity to the university should vary with the ratio of residence price to residence size. The residence price to residence size ratio is a measure of the dollars per square foot of housing and will be abbreviated as the RP/RS ratio. In terms of student choices, the trade-off between space and accessibility presents an interesting situation for student household demographics. If demographic groups make different choices for residential value (RP/RS ratio in dollars per square foot) versus accessibility, these choices ought to have simultaneous impacts on spatial distribution. Demographic groups that choose to live close to the university ought to pay more per
square foot, and demographic groups that choose to live further from campus ought to enjoy a lower ratio of residence price to residence size.

There is reason to believe that different types of student households have different preferences for space versus accessibility. For example, it is reasonable to expect that a student who has a husband and three children require a certain size residence. A student in this situation might set a square footage necessary to house her family, and then set a limit on what she is willing to spend on housing. Once these two decisions are made, the trade-off between space and accessibility will translate into a predetermined range of locations.

Conversely, a single sophomore who has decided to move out of the dorms may place considerable value on remaining close to campus. This student does not need to provide space for dependent children, thus the amount of acceptable residential space may dependent on how close the residence is to the university. In this scenario, the solution may be to live alone in a small unit that is expensive per square foot but close to school, or the student may choose to live with several roommates spreading the high cost per square foot and only sacrificing modest accessibility.

McQuiston (1992) found that different household demographic groups exhibited very different preferences for residential space. McQuiston separated the student population into six household demographic groups: Dorm Residents, Live with Parents, Live Alone, Live with Adults Only, Live with Adults and Children, and Live with Children Only (McQuiston 1992, 6). Two of these demographic groups exhibited preferences for larger residences. The demographic group Live with Parents and the
demographic group Live with Adults and Children generally exhibited preferences for larger residences, but did not appear to pay a great deal more for them. In the McQuiston study, no further investigation into this anomaly was undertaken.

The hypothesis being proposed is precisely an investigation into this anomaly. Applying the Alonso model to the findings by McQuiston, it should only be possible that some student demographic groups pay less for a comparable size residence if the residence is in a less accessible location. **The hypothesis is that if a demographic group exhibits a preference for a low residence price to residence size (RP/RS) ratio this preference will be accompanied with low accessibility to the university, and if a demographic group exhibits a preference for high accessibility to the university this preference will be accompanied with a high RP/RS ratio.** This hypothesis is similar to the concentric rings of von Thuened, which hypothesized where certain types of crops would be grown based on land rent calculations (Thuened 1842).

The last matter of concern regarding the demographic hypothesis is how it relates to the research question. First, this hypothesis can extend the claims made by McQuiston that student household demographic groups make very different residential decisions. Second, testing this hypothesis can identify which types of students make certain types of residential decisions. This should make it possible to explain why different types of students live where they do. Specifically, it will be possible to understand why one type of student may to choose a low residence price per square foot over accessibility to the university, and why another student may choose accessibility even in the face of a higher residence price per square foot.
Residential Location and Commuting Practices

Making the connection between student residential patterns and commuting practices requires an assessment of the impacts of residential decisions on commuting practices. The initial form of the hypothesis in this subsection requires minimal development. The hypothesis is that student residential patterns can be expected to have systematic impacts on commuting practices.

The dependent variable, commuting practices, should be one that is a reasonable way to present commuting practices. Commuting practices will be operationalized as modal split (percent biking, percent walking, percent driving, etc.). There are two basic ways in which modal split can be used to represent commuting practices, by trip or by proportion of distance traveled. Naess and Sandberg provide a good example of modal split analysis utilizing the proportion of distance traveled by mode (Naess and Sandberg 1996). There are a couple of advantages to the proportion of distance approach. The miles distance travelled by mode approach makes it very easy to accommodate commuters who use several different modes, and it is easy to study modal split from the perspective of energy consumption across modes.

There is also a significant disadvantage to using the proportion of distance traveled by mode. The disadvantage is that distance becomes an outcome variable. If distance between 'work' and residence is expected to be important in explaining modal split, then distance should not be both an explanatory and an outcome variable. In the research for this thesis, it is expected that distance to the university will be the key explanatory
variable in describing the impacts of residential choices on commuting practices. For the purposes of modal split hypothesis testing, commuting practices will refer to trips by mode, not distance by mode.

The reason distance is expected to affect commuting mode is simplistic. Essentially, people are only willing to walk so far, and people are only willing to bike so far. For example, walking a half mile to school might be an enjoyable way for many students to begin the day. Biking two miles on a nice day might be a nice way for many students to begin the day. It is unlikely that very many students would think that biking or walking seven miles would be an enjoyable way to start the day. However, few students would consider driving seven miles to school a huge inconvenience. Hartshorne employs this reasoning when explaining how urban structures have changed as a result of changing transportation technology (Hartshorn 1992). Muller also employs this reasoning when he writes, “Before 1850 the American city was a highly compact settlement in which the dominant means of getting around was on foot, requiring people and activities to tightly agglomerate in close proximity to one another. This usually meant less than a 30-minute from the center of town to any given urban point (Muller 1986, 9).”

This modest example describes the reasoning behind distance to the university being expected to influence modal split. Modal split will be measured in terms of number of trips by mode. The mode is a dichotomous choice between alternative (bike, walk, bus) and car commuting. Using these two measures, the hypothesis can be stated succinctly. **Distance is expected to have significant effects on modal split.** In other words, whether students drive depends on distance from their residence to campus.
Confounding Variables

Several external variables have a reasonable chance to influence the logic applied in hypothesis development, such as additional employment location, commitment to school, bad weather, and separate residence price markets.

The first external variable with the potential to affect this research is additional employment location. For some students, proximity to work may be valued more than proximity to school. This presents a problem for the logic applied in the Alonso theoretical model, which is based on the city center as the single destination node. Understanding the control that must be applied begins by assessing the extent to which the external variable represents a problem. Additional employment location is essentially a location problem so it is necessary to analyze it as such. Many students work in downtown Missoula and on campus. For these students, valuing proximity to work over proximity to school is a moot point, because they are in the same area. The problem created by additional employment location is limited to students who have jobs that are some distance from campus, for instance at least a mile from campus (past downtown). The control variable for additional employment location will be categorical variable, identifying students with additional employment located more than a mile from campus.

The second external variable with the potential to affect the logic of this study is ‘commitment to school’. This variable also represents a problem with the assumption that proximity to school is important for students making residential decisions. If school is not a very important factor in a student’s daily activities then there is little incentive to
locate close to school. The problem with this external variable is that ‘commitment to school’ is not an easy thing to measure. For this reason, a proxy for commitment to school is important. The logical proxy is semester credit hours. If a student is taking more than fifteen credits, then it seems reasonable to proposition that they are very committed to school. Conversely, if a student is taking three credits, then it seems reasonable that the student may not be especially committed to school. Semester credits are not an ideal measure of ‘commitment to school’, but they are an acceptable proxy.

The third external variable that can present a problem is weather, by affecting commuting mode. Riding a bike on a nice day can be very enjoyable, riding a bike in freezing rain can be a harrowing experience. Weather’s potential to alter behavior can be subjective. For example, some students at the University of Montana spend their free time climbing some of the world’s tallest mountains. These students may be less likely to change their behavior because of weather than a student who spends their free time at the mall. An objective measurement of weather effects on commuting mode was not attempted. Achieving an objective measurement of weather’s capacity to alter commuting behavior would have increased the complexity of the survey considerably. As the complexity of a survey increases response rates have a tendency to diminish. Rather, students were given the opportunity to express how weather affects their commuting practices. This is an example of an area where further research could be conducted, and an objective connection between weather conditions and commuting mode could be established.
The last confounding factor is separate markets which can cause serious problems for the Alonso model. Price models require consumer decisions be made in the same market. Most all students are making decisions within the Missoula real estate market. However, real estate markets are more complex than simple geographic area. The main market distinctions important for this research are the rental residence price market, the homeowner residence price market, and the Live with Parents/Relatives anti-market.

The rental residence price market is the most prominent and dynamic market. Rental decisions are usually not as long term as home buying decisions, and can be expected to respond more quickly to changes in supply and demand. In this sense, it is expected that the Alonso model will work best when the rental market is the only market included in the model.

The temporal nature of the home owner residence price market presents some interesting considerations. Some students who just recently purchased a home can be expected to have residence prices that are influenced by many of the same factors as the rental residence price market. However, some students purchased their homes years ago. Many of these homes have doubled in value. Thus, the quality, size and location factors alone would predict a residence price twice that reported in the survey.

Students who live with their parents/relatives also introduce some intriguing considerations. The survey required that students put both the total household residence price and the portion of the residence price that they are responsible to pay. Thus, the household total should be expected to follow the same market principles as the home owner residence price market. However, this has little to do with explaining why
students choose to Live with Parents/Relatives. The location and size of the home are determined by the parents/relatives. These students can pay little or nothing for the size and location of a home chosen by someone else, or they can make their own residence choices subject to the rental residence price market. Thus, these students can be characterized as making an anti-rental-market decision.

**Hypotheses Summarized**

The research question for this thesis is whether there is a connection between student residential patterns and commuting practices, and how this connection should be understood. Residential patterns will be studied first, including study of the effects of accessibility to the university on residential decisions. Two hypotheses will be tested to explain residential patterns. The first hypothesis is that residence price can be expected to be predicted by residence size and distance from the university. The second hypothesis is that if a demographic group exhibits a preference for a low ratio of residence price to residence size this preference will be accompanied with low accessibility to the university, and vice versa. These hypotheses together will investigate two questions. Why do students live where they do, and how important is accessibility to campus in the residential decision? Once residential patterns have been examined, commuting practices will be studied by testing the hypothesis that residential location can be expected to significantly influence student commuting modes.
CHAPTER 3

METHODOLOGY

This thesis examines University of Montana student residential patterns and commuting practices. A survey was used to capture variables important for describing student residential patterns and commuting practices. The methodology employed falls under the quantitative social science paradigm, and accords with generally accepted quantitative social science research practices. The methodology is presented in three sections. The first section focuses on data collection such as sample selection and survey design. The second section focuses on data preparation addressing issues related to data input and geocoding. The last section will cover the analytical procedures for establishing the connection between student residential patterns and commuting practices.

Data Collection

The first problem associated with any quantitative social science research is procuring a representative sample. This is especially challenging given the various economic, temporal, and human capital restrictions that inevitably accompany research projects. This is one area in which this thesis diverges slightly from typical research. Often, it is the case that a desired sample size will be determined by a mathematical
formula designed to generate the appropriate minimum desired sample size. Sample size was determined in a somewhat different fashion for this thesis. The desired sample size was pursued as a maximization of the available economic and human capital resources. There is no a priori scientific argument against this method for determining the sample size, so long as the resulting sample exceeds the minimum requirements presented by a traditional mathematical approach.

Results of this research are based on over four hundred responses, given that a sample size of that size is sufficiently large, the sample selection and survey design represent the primary sources of error in the data collection process. It is important to note that sample selection and survey design are both affected by the decision to geocode the survey responses. This thesis examines residential patterns and commuting practices. Therefore, a precise measurement of the location of survey respondents is desirable. Geocoding is a GIS process that creates a geographic index in a database, allowing the information in the database to be linked to a digital map. The geocoding process makes it possible to locate survey responses precisely.

Sample Selection

In fall 2000, over 12,000 students attended the University of Montana, and this study is based on a sample of 1,243 students drawn from that population. The first step in developing the sample was to acquire a list of student names and addresses from the registrar's database. The sample could not be randomly drawn from the entire registrar's
database because of University of Montana confidentiality policy, and the requirement that each address be the correct geocodeable address.

University confidentiality policy allows students to declare their student information strictly confidential. Thus, the registrar was unable to provide these students’ information. The total number of student records received from the registrar was 11,797. Given, that total student enrollment was about 12,000 fall semester 2000 it is very unlikely that these two hundred students could result in a biased sample.

Students are encouraged to provide the registrar with two addresses: a permanent address (often the family home), and a current address in the Missoula area. Students who do not make the appropriate address changes during registration were eliminated from the registrar’s database, because a current geocodeable respondent address is critical. Thus, it was necessary to eliminate addresses that clearly did not represent a current address, such as an address in Glasgow. This process resulted in a few subjective decisions for data set elimination. Any student taking six credits or less, with an address located within an hour and a half of Missoula, and on the west side of the continental divide was kept in the data set, otherwise they were eliminated. Furthermore, if a student was taking more than six credits, then they were eliminated if they did not have an address that was within an hour from Missoula. Overall, decisions to eliminate student records needed to be reasonable, but an attempt was made to error toward leaving a student in the data set if there was a reasonable chance that the address listed did in fact represent a current address.
The third reason for elimination occurred because many students have post office mail boxes. If the address listed is not a street address, then it is not possible to geocode that student record. As a result, all post office mail boxes in the Missoula area were eliminated. Post office boxes in outlying areas were retained because location precision is not as critical beyond the Missoula urban area.

When the registrar’s database had been pared down using the methods above, the end result was a data set of 9,480 students. Therefore, a substantial number of students, about 2,500, were eliminated from the research project at the outset. The fact that some students were omitted from the analysis is not really the important question. The important question is whether the methods and reasons for eliminating students denote a serious threat to representative sampling, inducing bias. An answer to this important question is somewhat complex, thus an essay on this topic is available in appendix a. The essential conclusion of this essay is that the methods employed to pare the registrar’s database could provide a source of sample bias, but it is difficult to deduce what this bias might be.

Once the registrar’s database had been pared down, a standard random sampling process was undertaken. 1250 students were randomly selected. Seven students could not be geocoded effectively. Thus, the final sample consisted of 973 off-campus students and 270 dorm students. The dorm students were not sent surveys because the vast majority of the information in the survey can be obtained from campus resources, such as Residence Life. For the most part, dorm students were not included in the analysis. However, when dorm students were included a random sample of 111 dorm students
from the 270 students in the original sample was extracted based on the net survey response rate of 42 percent. This essentially makes dorm ‘responses’ proportionate to the rest of the survey responses.

For geographic research, it is important to get a good geographical distribution, and this is one of the most challenging issues related to a survey such as this. One could argue that if the data collection methods result in a geographically representative data set, then this is a good indicator of representativeness across the board.

In order to assess the geographic representativeness of the responses, the pared-down registrar’s database was geocoded. This allowed the mean distance differences to be compared between the responses and the original registrar’s database from which the sample for surveying was drawn. The differences, in mean distances to campus, for the pared-down registrar’s database and the responses were tested for significance with independent samples t-test, given in appendix b. For the most part, these t-tests show that the responses are geographically representative. Figures 3.1 and 3.2 provide some insight.
Figure 3.1 - Pared-down Registrar’s Database Distance Distribution

Pared-down Registrar’s Distance Distribution

<table>
<thead>
<tr>
<th>Distance Range</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mile</td>
<td>5000</td>
</tr>
<tr>
<td>1-2 Miles</td>
<td>4000</td>
</tr>
<tr>
<td>2-3 Miles</td>
<td>3000</td>
</tr>
<tr>
<td>3-4 Miles</td>
<td>2000</td>
</tr>
<tr>
<td>4-5 Miles</td>
<td>1000</td>
</tr>
<tr>
<td>5+ Miles</td>
<td></td>
</tr>
</tbody>
</table>

Number of Students with Missoula area address

Figure 3.2 - Survey Respondent Distance Distribution

Survey Respondent Distance Distribution

<table>
<thead>
<tr>
<th>Distance Range</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mile</td>
<td>300</td>
</tr>
<tr>
<td>1-2 miles</td>
<td>250</td>
</tr>
<tr>
<td>2-3 Miles</td>
<td>200</td>
</tr>
<tr>
<td>3-4 Miles</td>
<td>150</td>
</tr>
<tr>
<td>4-5 Miles</td>
<td>100</td>
</tr>
<tr>
<td>5+ Miles</td>
<td></td>
</tr>
</tbody>
</table>

Survey Respondents (Dorm Students Included)
The patterns depicted in figures 3.1 and 3.2 are almost identical. This lends credence to the geographic representativeness of the sample. Only for the five-mile-plus distance category do the proportions differ, and this difference may actually show truer geographic distribution. For example, when paring the registrar’s database, an address in Stevensville was not eliminated, but that address in the registrar’s database might be the student’s parent’s address. It would appear, from the last column in figures 3.1 and 3.2, that the survey design may have reduced errors of this type in the survey responses. In summary, it should be clear that great care was taken to achieve as representative a data set as possible, and it would appear this care has been rewarded.

Survey Design

This research utilized a questionnaire, displayed in appendix c, and will henceforth be referred to simply as ‘the survey’. The survey needed to acquire information in three areas, household demographics, household size, and commuting practices. The survey design had to meet information needs, while still generating a decent response rate. After several survey design revisions, a pilot survey was conducted. The pilot survey allowed the final survey to be fine tuned, reducing error.

Demographic information was collected by replicating the survey design of the McQuiston study (1992). The other information that was particularly difficult to acquire was the household square footage and the commuting practices. Each of these presented unique problems and will be addressed individually. The other questions that appear on the survey in appendix c are yes/no type questions designed to deal with potentially
confounding variables, and design of these questions do not present methodological issues worth discussion.

Accurate information on residence size is critically important for this research, and the biggest problem to overcome in survey design was getting reasonable data for residence size. The basic problem is that most students only have a rough idea how big their residence is. As a result, just asking for residence size would likely result in many aberrant answers. The solution utilized was a three-part question in which the respondent answers only the part of the question that applies to their residence. The three-part question was divided so that the first part pertained to one bedroom residences, the second part to two and three bedroom residences, and the third part applied to residences with more than three bedrooms. Respondents then circled the range of square footage that applied to their residence, and estimated the storage space at their residence. This was the other key factor in procuring the residence size.

This approach took advantage of statistics for grouped data. Usually grouped data is used when a data source has been aggregated and the original micro data is no longer available. The standard approaches when working with grouped data is to apply the midpoint of each group to each observation. This approach is generally accepted, because the law of averages dictates that as N grows the true mean of the sample will approach the midpoint of the group. In this case, the respondents grouped themselves. Some students placed their residence in the wrong group, but it is unlikely that this would have occurred systematically, and therefore the same logic may be applied.
While the approach to acquire residence size may be inventive, the reported data should be treated as estimates. Therefore, some caution must be exerted in over interpreting the raw numbers as being absolute. Achieving highly accurate absolute numbers would have required considerable effort for the survey respondent. This could have jeopardized survey response rates, because fewer people tend to respond as the effort to complete the survey increases. However, this survey was developed through the use of a pilot survey in which the respondents were required to physically measure their residences after taking the pilot survey. When the estimated size was compared to the measured size, the mean difference was within one hundred square feet, and statistically insignificant. This suggests that the size estimates are fairly reliable. Hence, a respondent who reported living in a larger residence should in fact live in a larger residence, on average.

The second obstacle in survey design was dealing with commuting practices, which can be notoriously difficult to measure in a short survey. Each mode of transportation requires a repetition of each question, and it was expected that most students utilize one or two modes exclusively. Weather can be expected to affect commuting practices, for instance. Other problems can occur due to underestimates of normal trip patterns. Resolving all these issues result in an excessive number of questions for the typical respondent. The solution employed was to use a table. Respondents placed the appropriate number of commutes in the appropriate cells and left the rest blank.

In general, the survey design was effective. The information required for the study was acquired. Most of the surveys appeared to be filled out correctly. Most all of the
surveys were completed with few missing values. Even some of the surveys that were ‘red flagged’, were done so due to extenuating circumstances. For example, some foreign students couldn’t even guess how many square feet their residence was, because they are used to the metric system.

**Data Preparation**

Through the course of this thesis, only a couple data preparation issues are worth mentioning. The two principal issues relate to the survey data input and how the distances from campus were calculated.

Data input from the surveys followed a strict adherence to the information provided on the survey. The only significant deviation from this approach was residence size for students in the University Villages, because a more accurate measure was available. The University Villages document residential space. Thus, the actual square footage was input from the documentation instead of the survey estimate.

Strict adherence to the surveys can increase data preparation work in the analysis stage. For example, many students do not alter transportation habits in bad weather, and failed to fill out the bad weather columns. Thus, the data had to be prepared prior to analysis by filling in the missing values with values from the ‘usually’ column of trips.

As the discussion in the conceptual background showed, residential location is treated as an important variable for understanding the connection between student residential patterns and commuting practices. ‘Residential location’ can be understood as the position of a student’s residence in relation to the university. Thus, distance to the
university will be used as the ‘residential location’ measurement. Generating a distance from campus for each respondent was the most important data preparation challenge. This was accomplished using ArcView GIS software and base maps created by Missoula County. The base map is a parcel map, shown in appendix f. Thus, the geocodes assigned to each student are simply the geocode that identifies the parcel polygon on the base map where the student residence is located.

Mapping the student’s residences was a matter of joining the geocoded student information database to the geocodes in the base map. A new shape file with the polygons that had a student joined to them was then created. The X-TOOLS extension was then used to convert these polygons to points at the centroids of the polygons. This process yields the maps throughout this thesis.

Locating students on the map does not assign a distance from campus that can be used for statistical analysis. Assigning distances required a somewhat more complex procedure. First, the Spatial Analyst extension in Arc View was used to create contours at fifty foot intervals from the center of the University. Second, a spatial join was performed using the geoprocessing wizard. Once this join is completed, a new file can be created and each student residence point includes a field with the closest contour value. This process yields a distance to the center of campus which is precise to plus or minus fifty feet, well within the tolerances necessary for the statistical analysis.

Every student in the analysis had then been assigned the correct distance with one exception. The base map covers most of the greater Missoula area, but some students live beyond the boundaries of the base map. For these students, a more crude distance
measure was used. The distance calculation for these students was assigned based on the mileage to the town in which they live, precise to about two miles. However, this does not present a real problem for the analysis, because the phenomena utilizing the distance calculation are not affected much once the distance is greater than about eight miles from campus, and the base map covered almost any student within eight miles of campus.

Care was taken in data collection to acquire a reliable data set. Judgement was exercised when making data preparation decisions, protecting the integrity of the data set. Most importantly, the data set derived from the survey includes demographic, residence size, residence price, residence location, and commuting practices data. This is the requisite data to make the connection between student residential patterns and commuting practices.

**Data Analysis**

Four analytical procedures will be discussed in this section, corresponding to the analytical procedures employed in the analysis chapter. For each procedure, a connection should be established between the hypotheses and the conceptual and technical appropriateness of the analytical procedures. Second, it is important that the analysis chapter advance adeptly so any unique methodological issues will be covered in this section.

**Data Description**

The general hypothesis is that there is a connection between student residential patterns and commuting practices and that this connection should be placed in the context
of the broader socio-economic framework. Data description summarizes the raw numbers and patterns forming the foundation upon which the connection between student residential patterns and commuting practices can be made. Data description also conveys cursory knowledge about the data set making the inferential procedures more transparent.

The data description consists of tables, charts and simple thematic maps. Data description techniques are a combination of convention and style. When different styles and conventions are employed, the data description may take on a somewhat different character. For this reason conclusions drawn from the data description will be subordinate to more powerful inferential procedures.

There is one other procedure that will be included in the data description section that is not technically a descriptive procedure. This procedure is the paired t-test. Paired t-tests are inferential statistics, but they will not be used as a hypothesis testing procedure. Instead, they will be used to compare commuting practices in good and bad weather as well as between what was reported for usual behavior versus what was reported for yesterday's behavior. In the case of the commuting practices data, the data collection was more precise than is required for the analysis. The additional precision was included to try and improve overall accuracy. The paired t-tests analyze the data at a higher level of precision, assessing the accuracy of the data to be aggregated.

Linear Regression

The hypothesis being tested with linear regression is a reformulation of the theoretical model developed by Alonso (1964). Assuming residential quality is constant,
student household residential price can be predicted by distance to the university in miles and residence size in square feet. The corresponding null hypothesis is that, if housing quality is held constant, then distance to the university and residence size are not effective as predictors of residence price.

The purpose of linear regression is to quantify the relationship between one or more continuous independent variables and a continuous dependent variable. In order to reject the null hypothesis, it must be established that two continuous dependent variables, distance to the university and residence size, exert a significant effect on a continuous dependent variable, residential price. Linear regression is the correct analytic procedure capable of detecting a relationship to reject the null hypothesis. The equation for the regression is as follows:

\[ y = b_0 + b_1 x_1 + b_2 x_2 \]

where \( b_0, b_1 \) and \( b_2 \) are the regression coefficients.

Linear regression is appropriate if the observed versus expected residuals are normally distributed. The histogram in figure 3.3 shows that the distributional requirements of linear regression are met.
The conceptual background designated separate markets as a significant problem for any price model. Thus, the linear regression model will be run in aggregate form, and then it will be disaggregated into renters only, owners only and live at home. The expectation being that the model is likely to work better for the renters only, because the rental market is subject to fewer confounding factors. If the model does not work at all for owners and students that live with parents/relatives, then these models will not be
Another factor that can confound the model is variability in housing quality. No attempt was made to measure housing quality because perceptions of housing quality can vary widely. The research was performed by survey, and perceptions of housing quality are such that an useful objective measure of housing quality would have been very difficult to acquire. However, the housing quality problem is less of a concern than it first appears. The model assumes that housing quality is constant and that residence price can be determined by distance to the university and size of the residence. Housing quality is the “odd man out”, whatever isn’t explained by distance to the university and residence size ought to occur because the constant housing quality assumption is violated, and bias is introduced. But bias is not a huge problem in this instance, because the bias comes from only one source. Thus, if neither distance nor residence size were significant in the model then housing quality would have to be the other important variable in explaining residence price.

Another factor that can confound the regression analysis is the presence of dorm students. Linear regression is the only analytic procedure in this thesis that will explicitly include dorm students in some parts of the analysis. The reason that they will be included is to assess the degree to which dorm student patterns accord with the rest of the residence price market. Identifying variation from normal market forces is useful in assessing how university policy can affect residential patterns. Some caution should be exerted when interpreting the analysis that includes dorm students, because dorm students
form a homogeneous group compared to off-campus students. A homogeneous group with a significant number of cases can have adverse impacts when a line is fit in linear regression.

The linear regression analytic procedure includes a series of models in which the general form of the model is unchanged, but alternative measures and subsets are interchanged. A separate model will be utilized for each of two measures of residence price: Total household rent, and rent plus total household utilities. Also, a separate model will be included for each of the two measures of residence size. Residence Size refers to the living space estimation. Total Size refers to the living space estimation plus the storage space estimation. The reason for these separate measures is for consistency. For example, two otherwise identical residences could be such that one residence has a garage and utilities are included in the rent, while the other unit has no storage space and utilities are not included. By using these different measurements, it may be possible to identify the most effective measurements of these variables. The following equations illustrate the four specific regression equations (the constant term omitted).

\[
\begin{align*}
\text{Total Household Rent} &= b(\text{Distance}) \text{ and } b(\text{Residence Size}) \\
\text{Total Household Rent} + \text{Total Household Utilities} &= b(\text{Distance}) \text{ and } b(\text{Residence Size}) \\
\text{Total Household Rent} &= b(\text{Distance}) \text{ and } b(\text{Total Size}) \\
\text{Total Household Rent} + \text{Total Household Utilities} &= b(\text{Distance}) \text{ and } b(\text{Total Size})
\end{align*}
\]

Finally, restrictions on distance to the university will be placed on some models. The intervals employed are: all distances, residences within ten miles, and residences
within three miles. All distances shows the complete data set. Ten miles represents the greater Missoula area. Three miles represents Missoula proper. The purpose of using these intervals is that the market forces exerted by proximity to the university may become more prevalent as proximity to the university increases.

In summary, linear regression will be used to accomplish one primary goal. The goal is to identify how well distance to the university and residence size predict residence price. Linear regression is the appropriate analytical technique for this type of analysis both conceptually and technically. In addition, by using different variable restrictions and measures it will be possible to assess which measurements are the most meaningful for residence price and residence size.

Analysis of Variance

The third analytical procedure employed in this thesis is analysis of variance (ANOVA). ANOVA and Bonferroni post-hoc tests will be used to test the hypothesis that for demographic groups exhibiting preferences for a low ratio of residence price to residence size (RP/RS) this preference will be accompanied with low accessibility to the university. And conversely, demographic groups exhibiting preferences for high accessibility to the university this preference will be accompanied with a high RP/RS ratio. The methodological discussion of these analytical procedures will proceed by looking at conceptual matters first and technical matters second.

This hypothesis has two components, and testing can be partitioned into two null hypotheses. The first null hypothesis is that household demographic groups will not
exhibit statistically significant preferences for accessibility and the RP/RS ratio. The second null hypothesis is that there is no direct relationship between demographic group’s preferences for accessibility to the university and demographic group’s preferences for the RP/RS ratio.

For each of these components, three measures will be employed. The measure, for accessibility to the university, is distance to the university. Two measures of the RP/RS ratio will be employed. The first measure will be the ‘personal ratio’. This ratio is equal to the respondent’s share of the residence price divided by the square footage per resident. The other measure of the residence size to residence price ratio is equal to the household residence price divided by household square footage.

ANOVA is the technique that will be used to test the first null hypothesis. ANOVA is used to examine whether two or more means are significantly different. Five demographic groups were outlined in the previous chapter, Live Alone, Live with Adults Only, Live with Adults and Children, Live with Children Only, and Live with Parents/Relatives. For each of these demographic groups, means can be computed for distance, ‘personal’ RP/RS ratio, and ‘household’ RP/RS ratio. ANOVA is a technique that tests the differences between means, “by analyzing the variance, that is, by partitioning the total variance into the component that is due to true random error and the components that are due to differences between means ” (Statsoft 1999). Thus, ANOVA is a method that provides the statistical power to discern whether apparent mean differences between demographic groups are “real” or just a matter of sampling error. If
the means computed for each demographic group are significantly different, then the first null hypothesis can be rejected.

Assuming that significant differences between demographic groups are found, the second component of the hypothesis is that there will be a relationship between demographic preferences for distance and the residence price to residence size ratio. The first step in testing this component of the hypothesis is the identification of variables that contribute to the differences between demographic groups.

Bonferroni post-hoc pairwise comparisons will be used to identify significant differences. The Bonferroni tests are conservative pairwise tests (Statsoft 1999). Meaning the test rarely indicates significant results that are the result of random error. The essential element of the test is that, unlike a t-test, the Bonferroni test accounts for the number of tests being performed (Green, Salkind, and Akey 1997, 460).

Once the significantly different groups are identified, rejection of the second null hypothesis can proceed by looking at whether there is a relationship between distance and the RP/RS ratio. Do demographic groups that exhibit preferences for a low RP/RS ratio also demonstrate significantly higher distances to campus, and do demographic groups that exhibit preferences for low distances to campus also demonstrate significantly higher RP/RS ratios. The Bonferroni post hoc statistics will also be used to compare the significant mean differences to analyze these relationships. The second null hypothesis can be rejected if there is a trade-off pattern between the RP/RS ratio and distance to campus.
There are a couple important technical matters that must be addressed for testing the demographic hypothesis. The first is a distance restriction. A distance restriction of ten miles will be employed for this analysis. Comparisons of distance are a crucial part of hypothesis testing, but if no distance restriction is employed then a few students in a particular demographic group living far from campus can skew the distance mean severely. A severely skewed mean can adversely affect an interpretation of the relationships being tested.

The other important technical matters are the measures of residence price and residence size. The linear regression analysis will utilize two separate measures of each of these phenomena. However, it is expected that the linear regression procedure will identify the best of the two measures. Assuming this is the case, the analysis of variance will be run with the best measures for residence price and residence size.

Logistic Regression

The fourth analytical procedure employed in this thesis is logistic regression. Logistic regression will be used to test the hypothesis that a student’s residential choices can be expected to have systematic impacts on his/her commuting practices. The methodological discussion for the logistic regression will proceed by looking at conceptual elements and then looking at technical elements.

The conceptual elements for the logistic regression model are best evaluated in the context of the hypothesis being tested. The hypothesis being tested is that student residential patterns will affect commuting modes. Consequently, the null hypothesis is
that variables important for describing student residential patterns, in particular distance, will not affect student commuting modes.

The conceptual background demonstrated modal split as an appropriate measure of commuting modes. This thesis will employ a mutually exclusive binary categorization of modal split into alternative modes and car modes. For the purposes of the logistic regression, modal split will be defined in terms of a variable coded to be zero or one. A respondent is coded as a zero if biking, walking, or bussing account for more than 50 percent of trips to campus in a typical week, and coded as a one for car commuters. Rejection of the null hypothesis requires at least one variable that is important for describing residential patterns to also exhibit systematic impacts on commuting modes. Distance to the university is expected to be the variable that exhibits both characteristics. The discussion of the previous two paragraphs provides the conceptual foundation for the basic logistic regression model. The specific form of the basic logistic regression model is as follows:

\[ y = \frac{e^{B_0 + B_1 x}}{1 + e^{B_0 + B_1 x}} \]

Where \( y \) is the expected probability that a respondent uses a car as the primary commuting mode,

\( B_0 \) is the model constant, \( B_1 \) is the maximum likelihood coefficient for distance in the model and \( x \) is distance.
The conceptual framework for the basic distance model is complete with two exceptions. Bus commuters and car poolers represent unique problems for the logistic regression model. Car poolers were coded as car commuters, even though car pooling is somewhat different than driving alone. Bus commuters needed to have the distance variable recalculated based on proximity to the bus route. A brief essay in appendix d deals with conceptual problems presented by car poolers and the 'bus distance correction' in detail. However, neither of these issues present serious problems for the analysis because these two groups together comprise less than ten percent of all respondents.

The last conceptual element important for the logistic regression is assessing the importance of distance relative to other variables. Through research development process, other variables were identified that could be important for describing modal split. A final model will investigate variables besides distance, making it possible to assess the relative importance of distance. From a hypothesis testing standpoint, the basic distance model is the critical model and the final model will be developed for comparative purposes. Therefore, extensive coverage of methodological issues related to the final model would be superfluous. Methodology for the final model can be coterminous with the final model development, and the remainder of this methodology chapter will concentrate on the basic distance logistic regression model.

Two main technical issues arise in determining the appropriateness of logistic regression for testing the hypothesis that distance can be expected to have systematic impacts on modal split. First, logistic regression is similar to linear regression except that the outcome variable is dichotomous instead of continuous, and the covariates can be
either continuous or categorical. Second, linearity in the logit is assumed in the model, and must be tested.

It has been shown that modal split will be defined by a coded variable, with alternative transportation (biking, walking, bussing) coded as a zero and car commuting coded as a one. Thus, a continuous variable, percent of trips per week by car, was coded into a dichotomous variable for logistic regression. A graphical display of the data structure will make the reasoning for 'data downgrading' from interval-ratio data to binary data clear.

Figure 3.4 - Percent Car Distribution

The scatter plot in figure 3.4 shows why it is reasonable to code the variable as categorical, and why linear regression would not be effective. The data points are clustered at 0 percent and 100 percent. Fitting a linear regression to a data set with such a non-normal distribution will result in a poor linear regression model. It is clear that it is
appropriate to code the transportation variable. Technically speaking, the data structure
for the basic model is ideally suited to logistic regression. The outcome variable is coded
dichotomous. The independent variable, distance, is continuous.

When building a logistic regression model it is assumed that any continuous
independent variable is linear in the logit. Some preliminary discussion regarding the
logit transformation will be instructive.

\[ y = \frac{e^{B_0 + B_1x}}{1 + e^{B_0 + B_1x}} \]

Where \( y \) is the expected probability that a respondent uses a car as the primary commuting mode,
\( B_0 \) is the model constant, \( B_1 \) is the maximum likelihood coefficient for
distance in the model and \( x \) is distance.
The logit transformation is as follows:
\[ g(x) = \ln\left(e^{B_0 + B_1x}\right) \]
\[ g(x) = B_0 + B_1x \]

The logit transformation makes it possible to treat the dichotomous dependent
variable as if it were continuous in the regression. This is evidenced by the last equation
above. Linear regression should be used for variables that are expected to have a linear
relationship. The same reasoning applies to logistic regression, but in the case of logistic
regression the variable must be linear in the logit.

The test for linearity of the distance variable in the logit was carried out in the
manner prescribed in *Applied Logistic Regression*.

The difference between the logits for two different groups is equal to the value of an
estimated coefficient for one of the design variables obtained by fitting a logistic
regression model that treats the grouped variables as categorical. Thus, we treat the
grouped independent variable as if it were categorical with the lowest group serving as the referent group. Following the fit of the model we plot the estimated coefficients versus the midpoints of the groups . . . (Hosmer and Lemeshow 1989, 90).

This procedure yields the following graph in figure 3.5:

**Figure 3.5 - Linearity of Distance Plot**

The above plot is very linear even when no distance restriction is made, which would limit the skewness of the data to the right. If a distance limitation were applied, the plot would become very linear. Thus, the assumption of distance being linear in the logit is satisfied.

In summary, logistic regression will be used to analyze whether residential patterns can be used to predict commuting practices. It is expected that one variable in particular, residential location measured as distance, will be important for describing commuting modes. Logistic regression will be used to determine if distance has systematic effects on modal split.
CHAPTER 4

RESULTS OF ANALYSIS

This thesis analyzes the connection between student residential patterns and commuting practices. The analysis of this connection is presented in four parts, consistent with the analytical procedures presented in the methodology. Data description will consist of charts, tables, and paired t-tests, giving a general sense of the data set. The linear regression section will test whether residence size and distance to campus are effective predictors of residence price. The analysis of variance section compares demographic groups' preferences for accessibility versus preferences for residence size for the dollar. The linear regression and ANOVA examine two questions: a) Why do students live where they do? b) How important is accessibility to campus in student residential decisions. The logistic regression section will quantify the impacts of residential location on commuting practices. This chapter will feature the most meaningful results, and additional results will be referenced in the appendices. Finally, the results in this chapter will apply a conservative approach in analyzing statistical differences, requiring an alpha level of .01 for most tests.

This is a geographical thesis and the use of maps is important for most any geographic work. GIS has been an integral part of this research, and without GIS this
research would have been difficult, if not impossible to complete. The particular analytic elements employed in this thesis will not rely heavily on map interpretation, but maps will be used to enhance the analytic presentation.

Also, before proceeding to the analysis, a general statement regarding dorm students is necessary. The linear regression section is the only section that will include any dorm students, and the presence of dorm students will be explicitly indicated in this analysis.

**Data Description**

The purpose of the data description section is to provide an overview of the data set. About four hundred responses to the survey are the primary source of data for this thesis. In order to show the connection between student residential patterns and commuting practices it is first necessary to show the basic characteristics of student residential patterns and commuting practices. The data set includes important variables such as household demographics, residence sizes, commuting practices, etc., as shown in table 4.1. These descriptive statistics also present a preview of the data patterns that will be used to establish connections with inferential statistics and through hypothesis testing. This preview will make it easier to understand the meaning of the inferential statistics, especially for those who have little experience with some of the procedures being employed.
Table 4.1 - General Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of Residence</td>
<td>1,212 square feet</td>
<td>974 square feet</td>
</tr>
<tr>
<td>Size of Residence and with Storage Space Included</td>
<td>1,468 square feet</td>
<td>1,193 square feet</td>
</tr>
<tr>
<td>Respondent’s Share of the Rent</td>
<td>$271.66</td>
<td>$124.43</td>
</tr>
<tr>
<td>Rent of Respondent’s Household</td>
<td>$608.44</td>
<td>$309.10</td>
</tr>
<tr>
<td>Rent of Respondent’s Household with Utilities</td>
<td>$692.85</td>
<td>$342.49</td>
</tr>
<tr>
<td>Distance to Campus</td>
<td>3.9 miles</td>
<td>8.7 miles</td>
</tr>
<tr>
<td>Distance to Campus for Respondent (10mi. limit)</td>
<td>1.9 miles</td>
<td>1.6 miles</td>
</tr>
<tr>
<td>Age of Respondent</td>
<td>26 years</td>
<td>8.94 years</td>
</tr>
<tr>
<td>Number of People per Respondent Household</td>
<td>2.93</td>
<td>3.08</td>
</tr>
<tr>
<td>Respondents that are College of Tech. Students</td>
<td>4.4 %</td>
<td>NA</td>
</tr>
<tr>
<td>Percent of Students with Parking Decal</td>
<td>33.5 %</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 4.1 shows a few variables of interest. First, it is interesting to see that the size variables have means and standard deviations that are reasonable, and no standard deviations that exceed the means. Few students live alone evidenced by the mean household size close to three. Also, the skewness of the distance variable when no distance restrictions are included is evident. A few students who commute great distances to campus skew the data. Distance restrictions will be applied and identified as required, but a categorical distance restriction will not be applied because some parts of the analysis would be compromised.
Figure 4.1 shows the percent distributions for off-campus ownership status. There is a fifteen percent home ownership rate among off-campus students. In 1992, 23 percent of off-campus respondents were homeowners (McQuiston 1992, 15). This indicates a precipitous drop of eight percent in just eight years. While this drop is not statistically significant at the standards set by this thesis, it does approach significance. Either a real drop in student home ownership rates has occurred, or one or both of the studies were skewed with respect to ownership. Even though this research is incapable of drawing definitive conclusions, it would appear that further study could be fruitful. The research could assess any latent student demand for ‘homes for sale’.

Figure 4.1- Ownership Status
Figure 4.2 shows household demographic percent distributions. Over 15 percent of respondents have dependent children in their household, probably single parents and families. Also, it is evident that relatively few respondents that live off-campus live alone, when compared to those that have at least one roommate. The majority of students share housing with other adults. The spatial distribution of household demographics is show in the Respondents by Demographic Group map in appendix f.

**Figure 4.2 - Demographic Groups**
Figure 4.3 shows how percent distribution of ownership status and household demographic groups relate. Live with Adults Only and Live with Adults and Children are the two demographic groups that are responsible for most all of the home ownership. And Live with Adults and Children, presumably families, is the only demographic group in which there are more homeowners than renters.

**Figure 4.3 - Ownership Status by Demographic Group**

![Ownership Status by Demographic Group](image)

Figure 4.4 shows percent distribution of housing types. Mid-sized housing types dominate University of Montana off-campus student housing. Two to three bedroom residences account for 53.3 percent of the housing for off-campus UM students. Another prevalent housing type is a four-plus-bedroom detached house. Thus, students at the
University of Montana appear to exhibit an overall preference for larger housing, perhaps atypical of university students in other areas. Figure 4.4 shows that most off-campus students share housing, which accords with a mean of 2.93 people per household. A map of housing types is displayed in appendix f.

**Figure 4.4 - Percent Distribution of Housing Types**
The Respondent's Housing Type map tends to show a ringlike pattern. The first ring extends from the university about a mile and quarter. A variety housing of types seem to occupy the first ring, including a number of single bedroom units. The next ring seems to be predominated by mid-sized apartments and houses, ranging from about a mile and quarter to about two and three quarter miles. The last ring is dominated by houses with many of them falling in the four-bedroom-plus category.

Figure 4.5 displays the spatial distribution of housing types. Showing there is some credence to the ringlike characterization of the housing type map. Two and three bedroom apartments dominate the middle distance interval. Detached houses dominate the housing that is more than 2.63 miles from campus. Housing types closer than 1.23 miles are varied, but there are a number of single bedroom apartments.

Figure 4.5 - Spatial Distribution of Housing Types
The next area of data description is commuting practices. Commuting practices are the one part of the data description that will include more than basic descriptive statistics. Some paired t-tests were employed to ascertain three things. First, commuting practice changes from good weather to bad weather was assessed. Second, how the respondents 'usual commuting practices’ compared to the computing practices that were reported for 'yesterday’ was assessed. Third, how respondent commuting practices change by day of the week was tested. All the paired t-tests for this part of the analysis are displayed in appendix e.

The survey data showed that commuting practices do change from bad weather to good weather for two commute modes. Respondents reported a reduction in the number of bicycle trips during bad weather for each day of the week. Respondents reported an increase in the number of bus trips in bad weather for each day of the week. These results have two implications. It shows that commuting practices do change with the weather, and that the changes are logical and reasonable, reinforcing confidence in the data.

Respondents were asked for 'usual commutes’ and 'commutes yesterday’, allowing a comparison of estimations of usual commuting practices with recollections from the previous day. These paired t-tests did not show significant differences between the commutes that were recalled from the day before and estimations of usual commuting practices. This lends support to the argument that the 'usual commute’ estimates are reliable and can be used with confidence.

Finally, when comparing between days of the week, not in bad weather, the results were not significant for the most part. A few significant comparisons were found, but,
there does not seem to be a discernable pattern to them. Thus, variability by day of the week may in fact be significant on certain days for certain transportation modes, but it is unlikely that this type of survey research is precise enough to come to any definitive conclusions on this matter.

Summarizing these paired t-tests, they support the notion that the commuting data collection was prosperous. The t-tests support the proposition that an aggregated weekly total of ‘usual commuting practices’ is a reasonable representation of the data. The tests only present two caveats to this proposition, busing and biking in bad weather. There are no theoretical reasons why all the commuting analysis could not proceed by doing good and bad weather analysis separately. However, this would double the amount of commuting analysis which would present logistical considerations beyond the scope of this thesis. However, this is an example of subsequent research that is available in this data set. Keeping weather issues in mind, all commuting analysis will henceforth be based on aggregated weekly totals of usual commuting practices reported for each day.

Figure 4.6 depicts commuting practices. Figure 4.6 is presented with a ten-mile distance restriction, preventing the miles traveled per week from being exaggerated by a few students that commute a very long way to campus.
Figure 4.6 offers the first look at the commuting practice data. The large percentage of weekly commutes not by car is interesting. Biking, walking, and bussing makes up almost half of all commutes. An institutional factor that is important for increasing alternative transportation is the scarcity of parking spaces on campus. The University of Montana has fewer parking spaces than there are people who wish to park on campus, especially at peak times. The University of Montana registers about 30 percent more...
vehicles than there are parking spaces. This situation creates a significant disincentive to drive to campus. Figure 4.6 also shows that biking, walking, and bussing account for a greater percentage of total commutes than commuting miles traveled, meaning more frequent commutes of a shorter length, which fits with expectations.

The last component of the data description section will cover some ancillary variables and data collected in the survey. These are important for the discussion, because through the research process they were identified as potential confounders. For the most part, they did not exhibit systematic influences on any of the analysis, but dealing with them now will allow the remainder of the analysis to proceed efficaciously.

One variable that was tracked through the research process was College of Technology students. College of technology students commute to a different campus. This difference had to be accounted throughout the research. Only seventeen College of Technology students responded, so this does not represent a large enough group to warrant a separate analysis, but surprisingly the residential patterns and commuting practices did not change much when they were included or excluded. For this reason, it was not necessary to exclude College of Technology students from the analysis. The map in appendix f, entitled Respondent's Campus Designation, depicts the location of respondents that attend the College of Technology.

There are two interesting things that should be noted about the College of Technology respondents. First, they are distributed similar to main campus respondents with the exception that none were located close the main campus, because respondents who do not attend on the main campus have no incentive to locate close to it. The second
interesting feature is that the College of Technology does not have excess demand for parking spaces, and parking at the College of Technology costs a nominal amount when compared to the main campus. Thus, there is not a strong disincentive to drive. When these two factors combine, the result is that the respondents from the College of Technology drive just like their main campus neighbors.

The next variable with the potential to confound the analysis was credit hours. The essential logic is that the importance of school in a student’s daily activities could affect residential and transportation preferences. Credits are depicted in the map entitled *Credits Respondent was Registered for at the Time of the Survey* which is displayed in appendix f. In this map, the red points represent respondents with credit loads that are at least one standard deviation below the mean, and black points represent respondents with credit loads that are at least one standard deviation above the mean. This data was not collected in the survey, but was obtained directly from the registrar’s database. This map is a good summary of the findings for this variable in the research. Basically, the map does not present any discernable pattern. Indeed, the number of credit hours for which a student was registered did not seem to affect any of the analytic procedures. The logic for controlling for this variable is rational, but this is an instance when empiricism trumps deduction.

The last topic for data description is the potential for additional employment to act as a confounding factor. Employment, in addition to school, could be important when explaining where students live and how they get to school. Maps in appendix f depict
some geographical attributes of additional employment. Figure 4.7 depicts employment location.

**Figure 4.7 - Percent Distribution of Job Location**

74.5 percent of respondents reported additional employment, but just having additional employment did not appear to be important at all. None of the statistics in this thesis found additional employment location to be an important explanatory variable either, but job location and residential location appear associated.

Figure 4.7 provides insight into how additional employment location relates to the location of the university. The map depicting employment location is given in appendix f. It shows residences, for respondents with a job within a mile of campus, clustered around the university. Among students that have additional employment, students with all jobs located within a mile of campus live significantly closer to campus (.001). Even
though none of the analysis found additional employment location to be a strong predictor variable, job location and residential location seem to associate. Thus, it seems reasonable that radical changes in student employment patterns could affect residential patterns and commuting practices.

**Linear Regression**

The hypothesis tested in this section is a derivative of the Alonso theoretical model. Linear regression will be used to test the hypothesis that if residential quality is held constant, then residence price can be predicted by location and residence size. The resulting null hypothesis is that distance to the university and residence size cannot be expected to be significant predictors of residence price, assuming housing quality is held constant or is randomly distributed.

This null hypothesis can be logically reformulated into two null hypotheses, making it easier to analyze. The first null hypothesis is that distance to the university is not a significant predictor of residence price, housing quality held constant and controlling for residence size. The second null hypothesis is that residence size is not a significant predictor of residence price, housing quality held constant and controlling for distance to the university. Dividing the null hypothesis into its two logical components makes it possible discuss the contribution of residence size and distance to the university separately even when they will be tested simultaneously in a multiple regression model.
The general form of the equation for the regressions is:

\[
\text{Residence Price} = B_0 + B_1x_1 + B_2x_2
\]

Where \(x_1\) is distance to the university in miles, \(x_2\) is size of the respondent's household residence in square feet, and \(B_1\) and \(B_2\) are the parameter estimates for the standard least squares regression.

Table 4.2 shows the parameters for this linear regression model and adjusted R squares. Two measures each are given for household residence size and household residence price. One measure for household residence size includes storage space and the other does not. One measure for residence price includes utilities and the other does not. Finally, the model is reported using three distance intervals, all distances, ten miles (greater Missoula urban area), and three miles (Missoula proper).

As noted in the methods chapter, the linear regression model was run with several different measures and subsets. The linear regression model is the only analysis in this thesis that includes dorm students. The dorm students are included to compare housing prices off-campus with housing prices on campus. The linear regression was also run for separate markets. The separate markets are rental residence price, homeowner residence price, and Live with Parents/Relatives residence price. The model does not work well for the separate homeowner residence price market and the Live with Parents/Relatives anti-market, so these results are not shown. Instead, the renters only residence price market is the only disaggregated market shown.
Table 4.2 - Linear Regression Models for the Alonso Theoretical Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Independent Variable</th>
<th>All Distances</th>
<th>Ten Mile Distance Restriction</th>
<th>Three Mile Distance Restriction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total Household Rent</td>
<td>Household + Utilities</td>
<td>Rent</td>
</tr>
<tr>
<td>Dorm Included</td>
<td>Size</td>
<td>0.157</td>
<td>0.197</td>
<td>0.188</td>
</tr>
<tr>
<td>N= 430-481</td>
<td>Total Size</td>
<td>0.131</td>
<td>0.165</td>
<td>0.155</td>
</tr>
<tr>
<td></td>
<td>R square</td>
<td>0.282</td>
<td>0.281</td>
<td>0.357</td>
</tr>
<tr>
<td>Dorm Included</td>
<td>Size</td>
<td>0.242</td>
<td>0.285</td>
<td>0.260</td>
</tr>
<tr>
<td>N= 317-268</td>
<td>Total Size</td>
<td>0.199</td>
<td>0.236</td>
<td>0.215</td>
</tr>
<tr>
<td></td>
<td>R square</td>
<td>0.493</td>
<td>0.491</td>
<td>0.548</td>
</tr>
<tr>
<td>No Dorm</td>
<td>Size</td>
<td>0.176</td>
<td>0.205</td>
<td>0.197</td>
</tr>
<tr>
<td>N= 345-404</td>
<td>Total Size</td>
<td>0.147</td>
<td>0.173</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>R square</td>
<td>0.292</td>
<td>0.291</td>
<td>0.329</td>
</tr>
<tr>
<td>No Dorm</td>
<td>Size</td>
<td>0.282</td>
<td>0.313</td>
<td>0.285</td>
</tr>
<tr>
<td>N= 281-236</td>
<td>Total Size</td>
<td>0.233</td>
<td>0.260</td>
<td>0.236</td>
</tr>
<tr>
<td></td>
<td>R square</td>
<td>0.548</td>
<td>0.542</td>
<td>0.556</td>
</tr>
</tbody>
</table>

Bold indicates coefficient significance meeting or exceeding alpha = .01. R squared values are adjusted.
A few general characterizations of the models in table 4.2 can be made. It is evident that the expectation that the model would perform better when the rental market is the only market in the model is upheld. Thus, it would seem that the residence price rental market is not subject to confounding factors such as those introduced when a student lives in a home that was purchased some years ago.

Looking only at the models that include renters only and no dorm students, the $R^2$ hovers around .55 regardless of the distance restrictions. Overall, this means that the size of a residence explains about 55 percent of the cost of a rental anywhere in the greater Missoula area. The effect of distance restrictions is not as large as one might expect in this type of model. If one had to choose a single distance restriction, it appears that the ten-mile restriction is most meaningful. The model performance between the ten mile and the three mile distance restriction is negligible, so the ten mile is preferable because it encompasses a larger area. Also, the distance parameters seem to be more stable across all models, for the ten-mile restriction.

The model seems to work best when residence price is measured as rent plus utilities, and the independent residence size variable in the model includes both living space and storage space. This suggests that rent plus utilities is a better measurement of residence price than rent alone. This also suggests that a measurement of residence size is most effective when both living space and storage space are included.

The next three pages display the spatial patterns of the variables included in the linear regression model.
Map 4.1 - Respondent’s Ownership Status

Legend

Ownership Status
- Live with Parents/Relatives
- Own
- Rent

Main Roads
Major Rivers
Streams

Mileage Contours
\( \wedge \) 1 mile
\( \wedge \wedge \) 3 miles

Jay Harland
March 12, 2001
Base Maps Courtesy
Missoula County
Map 4.3 - Respondent's Residence Size
(Classified by Natural Breaks)

Legend

Living Space plus Storage Space
- 115 - 400 Square Feet
- 400 - 900 Square Feet
- 900 - 2000 Square Feet
- 2000 - 5850 Square Feet

Main Roads
Major Rivers
Streams

Mileage Contours
\n\n\n1 mile
3 miles

Jay Harland
March 0, 2001
Base Maps Courtesy
Missouri County
The Respondent’s Residential Ownership map shows very few student homeowners within a mile of campus. Many homes sell for close to $200,000 in the university area, and the cost of housing may be prohibitively high. The Respondent’s Total Household Rent plus Utilities Map shows a distribution with two locations of particular interest. Overall, the residence prices seem distributed fairly randomly with the exception of the South Hills and the Rattlesnake Valley where residence prices are high. There are some lower priced clusters downtown and at the University Villages. Correspondingly, the residences in the Rattlesnake and South Hills are large, as shown in the Respondent’s Residence Size map. As the discussion returns to the linear regression models, these maps will provide insight into some of ‘the real world patterns behind the numbers’.

With the presentation of the model results and maps, it is possible to deal with the null hypotheses directly. With an $R^2$ between .285 and .564 the model is working. The model works modestly overall, and is decent if the rental market is the only market included. The two null hypotheses presented in the beginning of this section will be analyzed in turn.

Residence size is significant in all the models, and on this basis alone the null hypothesis, that residence size is not a significant predictor of residence price, can be rejected. However, the significance of residence size is not especially interesting. Everyone would expect a 3,000 square foot home to rent for more than an 800 square foot home, especially when controlling for location. The models provide a parameter estimate, which is more interesting. Out of all the models displayed, the most meaningful model is probably the model with the ten-mile distance restriction, renters only with no
dorm students, using measures of rent plus utilities and total size. The parameter estimate in this model for residence size is 0.262. The constant for this model is 387. This model yields the following equation for the student rental residence price market in the greater Missoula area.

\[
\text{Residence Price} = -7.968 \times \text{(Distance in Miles)} + 0.262 \times \text{(Residence Size in Square Feet)} + 387
\]

When this equation is evaluated for several size residences and distances, it yields the results in table 4.3.

<table>
<thead>
<tr>
<th>Miles</th>
<th>200 sq. ft.</th>
<th>500 sq. ft.</th>
<th>850 sq. ft.</th>
<th>1100 sq. ft.</th>
<th>2700 sq. ft.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mile</td>
<td>$431.43</td>
<td>$510.03</td>
<td>$601.73</td>
<td>$667.23</td>
<td>$981.63</td>
</tr>
<tr>
<td>3 Miles</td>
<td>$415.50</td>
<td>$494.10</td>
<td>$585.80</td>
<td>$651.30</td>
<td>$965.70</td>
</tr>
<tr>
<td>5 Miles</td>
<td>$399.56</td>
<td>$478.16</td>
<td>$569.86</td>
<td>$635.36</td>
<td>$949.76</td>
</tr>
</tbody>
</table>

The estimates in table 4.3 are reasonable. This model seems to establish an appropriate relationship between residence size and residence price.

What is interesting is that distance is not statistically significant for predicting residence price. Distance is only significant when dorm students are included in the model, and the influence of dorm students can be attributable to two things. First, dorm students are a homogeneous group, and this homogeneity may boost the regression in and of itself. The alternative explanation is that dorm students are paying a premium to live on campus, and left to market forces alone many dorm students that are required to live
on-campus would opt to live off-campus. The correct explanation is probably a combination of these two.

With respect to null hypothesis rejection, the model does not show a strong enough relationship between residence price and distance to campus, for off-campus students. Thus, the null hypothesis that distance will not be a significant predictor of residence size cannot be rejected. This finding is a bit surprising, but this finding does not mean that distance is not important in determining residence price. Rather it means, that the Alonso theoretical model, as it was tested in this thesis, failed to find a statistically significant relationship.

There are at least two possible explanations for this failure. The first, is that the assumption that housing quality is constant or random across the study area has been violated. There are some neighborhoods in Missoula that are generally considered to have better overall quality of housing. If this is the case, then the effects of distance on residential price may be suppressed by bias as the result of variability in housing stock across neighborhoods. This is an example of an area where this research could be extended. Further research may be able to assess the potential bias introduced by variability in housing quality.

The other explanation for the failure of distance to significantly predict residence price is that the effect of distance is subtle. Excluding the models that incorporate dorm students, the lowest parameter for distance is -26.900. Which means that residence price should decrease $26.90 per household for every mile from campus. Supposing a student lives four miles from campus, this translates into a savings of $107.60. The average rent
plus utilities per student household is \$692.85. Thus, at a distance of four miles, the \$107.60 only represents about 15 percent of the total average housing cost. This argument was generated using the lowest parameter estimate for distance, so the total contribution of distance is probably even smaller.

The correct explanation for failure to reject the second null hypothesis is probably a combination of the above two explanations. Housing quality spatial variability probably creates a bias and the effects of distance on residence price are subtle. When these two factors are combined, the model was incapable of establishing a significant relationship between residence price and distance to the university.

**Analysis of Variance**

Analysis of variance and Bonferroni post-hoc statistics are used to test the hypothesis that for demographic groups exhibiting preferences for a low residence price to residence size (RP/RS) ratio this preference will be accompanied with low accessibility to the university, and for demographic groups exhibiting preferences for high accessibility to the university this preference will be accompanied with a high RP/RS ratio.

This hypothesis will be segmented into two different null hypotheses for testing. The first null hypothesis is that household demographic groups will not exhibit significantly different preferences for accessibility to the university and the RP/RS ratio. The second null hypothesis is that there is no relationship between demographic preferences for accessibility to the university and demographic preferences for the RP/RS ratio.
Analysis of variance will be used to test the first null hypothesis. Table 4.4 shows the results of the analysis of variance. Additionally, a ten-mile distance restriction will be employed to reduce aberrant results in the distance comparisons due to a few students that live a far from campus. Also, the linear regression section identified rent plus utilities as the best measure of residence price, and living space plus storage space was identified as the best measure of residence size. In accordance with these findings, this section will use these measures exclusively.

Table 4.4 - Demographic Group Analysis of Variance
(10 mile distance restriction)

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Demographic Group</th>
<th>N</th>
<th>Mean</th>
<th>s</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>Alone</td>
<td>63</td>
<td>1.461</td>
<td>1.112</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adults Only</td>
<td>237</td>
<td>1.680</td>
<td>1.307</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Children Only</td>
<td>16</td>
<td>2.106</td>
<td>1.868</td>
<td>19.772</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Live with Adults and Children</td>
<td>32</td>
<td>2.841</td>
<td>1.664</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Parents</td>
<td>29</td>
<td>3.870</td>
<td>2.374</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>377</td>
<td>1.928</td>
<td>1.582</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residence Price</td>
<td>Alone</td>
<td>61</td>
<td>1087</td>
<td>0.743</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adults Only</td>
<td>205</td>
<td>0.768</td>
<td>0.538</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Children Only</td>
<td>12</td>
<td>0.540</td>
<td>0.299</td>
<td>9.456</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Live with Adults and Children</td>
<td>29</td>
<td>0.518</td>
<td>0.262</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Parents</td>
<td>15</td>
<td>0.325</td>
<td>0.220</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>320</td>
<td>0.775</td>
<td>0.576</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Residence Price to Residence Size Ratio</td>
<td>Alone</td>
<td>59</td>
<td>1.040</td>
<td>0.762</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adults Only</td>
<td>204</td>
<td>0.737</td>
<td>0.584</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Children Only</td>
<td>13</td>
<td>0.754</td>
<td>0.490</td>
<td>4.590</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Live with Adults and Children</td>
<td>29</td>
<td>0.798</td>
<td>0.422</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Parents</td>
<td>17</td>
<td>0.318</td>
<td>1.098</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>322</td>
<td>0.776</td>
<td>0.656</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The analysis of variance shows significant differences between demographic groups for all three variables. Based on the significance levels alone, the first null hypothesis can be rejected. These demographic groups exhibit very different preferences for accessibility and the RP/RS ratio.

Distance appears to be the variable in which the demographic groups are most different, with an F statistic of 19.772. This is interesting because distance is an important variable for describing the demographics of students' residential decisions, but this is one variable that McQuiston failed to measure (McQuiston 1992).

The personal RP/RS ratio is the variable in which the least differences between demographics are exhibited. It is obvious from the means that the significant variability for this variable comes from only two demographic groups, Live Alone and Live with Parents/Relatives. This disparity would be even greater except for a single Live with Parents/Relatives respondent who rents a room, probably at a relative's house. The calculation for the personal RP/RS ratio results in this one respondent paying a lot per square foot. Besides this anomaly, the rest of the respondents that Live with Parents/Relatives pay very little for their residences, whereas respondents that Live Alone pay the most per square foot on average.

The personal RP/RS ratio variable will not be used for hypothesis testing, but it is shown for completeness. The household RP/RS ratio is the appropriate measure because it is not confounded by dependent children. Dependent children present a problem because they are a factor in the personal residence price to residence size ratio calculation, even though they do not usually contribute to rent or mortgage payments.
Interestingly, the variables do exhibit the hypothesized relationship. In table 4.4 the demographic groups are arranged in ascending order by mean distance, and the mean household RP/RS ratios decrease sequentially. Table 4.5 shows the significance of these mean differences.

### Table 4.5 - Bonferroni Post-Hoc Mean Difference Tests

<table>
<thead>
<tr>
<th>Distance</th>
<th>Household Residence Price to Household Residence Size Ratio (dollars/sqft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live Alone v Adults Only</td>
<td>-.2186</td>
</tr>
<tr>
<td>Live Alone v Live with Children Only</td>
<td>-.6453</td>
</tr>
<tr>
<td>Live Alone v Live with Adults and Children</td>
<td>-1.3799*</td>
</tr>
<tr>
<td>Live Alone v Live with Parents/Relatives</td>
<td>-2.4094*</td>
</tr>
<tr>
<td>Live with Adults Only v Live with Children Only</td>
<td>-.4267</td>
</tr>
<tr>
<td>Live with Adults Only v Live with Adults and Children</td>
<td>-1.1613*</td>
</tr>
<tr>
<td>Live with Adults Only v Live with Parents/Relatives</td>
<td>-2.1908*</td>
</tr>
<tr>
<td>Live with Children Only v Live with Adults and Children</td>
<td>-.7346</td>
</tr>
<tr>
<td>Live with Children Only v Live with Parents/Relatives</td>
<td>-1.7641*</td>
</tr>
<tr>
<td>Live with Adults and Children v Live with Parents/Rel.</td>
<td>-1.0295</td>
</tr>
</tbody>
</table>

Bold Indicates Significant Mean Difference at alpha = .05
* Indicates Significant Mean Difference at alpha = .001

The second null hypothesis is that there is no relationship between demographic preferences for accessibility to the university and demographic preferences for the RP/RS ratio. The demographic group Live with Children Only will be omitted from the
remainder of the analysis in this section, because the number of observations is small.

There are less than twenty observations for all variables, and one cannot have much confidence in results based on fewer than twenty observations. For ease of interpretation, the remaining four demographic groups are shown again in table 4.6.

Table 4.6 - Bonferroni Post-Hoc Mean Difference Tests
10 mile distance restriction, and Live with Children Only Omitted

<table>
<thead>
<tr>
<th>Household Residence Price to Household Residence Size Ratio (dollars/sqft)</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live Alone v Adults Only</td>
<td>-.2186</td>
</tr>
<tr>
<td>Live Alone v Live with Adults and Children</td>
<td>-1.3799*</td>
</tr>
<tr>
<td>Live Alone v Live with Parents/Relatives</td>
<td>-2.4094*</td>
</tr>
<tr>
<td>Live with Adults Only v Live with Adults and Children</td>
<td>-1.1613*</td>
</tr>
<tr>
<td>Live with Adults Only v Live with Parents/Relatives</td>
<td>-2.1908*</td>
</tr>
<tr>
<td>Live with Adults and Children v Live with Parents/Rel.</td>
<td>-1.0295</td>
</tr>
</tbody>
</table>

Bold Indicates Significant Mean Difference at alpha = .05
* Indicates Significant Mean Difference at alpha = .001

The hypothesized relationship becomes more apparent in table 4.6. In three comparisons, significant differences in distance are accompanied by significant differences in the RP/RS ratio, and signs for the mean differences are opposite as expected. The demographic group Live Alone exhibits preferences for greater accessibility when compared to the Live with Adults and Children demographic group and the Live with Parents/Relatives group, while these groups exhibit a preference for a lower RP/RS ratio. The demographic group Live with Adults Only also exhibits
preferences for accessibility when compared to the demographic group Live with Parents/Relatives, preferring a lower RP/RS ratio.

There is a significant difference between Live with Adults Only and Live Alone, for the RP/RS ratio. This difference is not accompanied by a significant difference in distance. This difference is surprising because these demographic groups are similar in that they are non-family households. Respondents who Live Alone pay significantly more per square foot. Respondents who Live with Adults Only often take on roommates to lower costs per square foot. Thus, by taking on five roommates to live in a large residence it may be possible to get somewhat more space for less money without sacrificing much accessibility. Non-family households exhibit similar preferences for access, but respondents who Live Alone are willing to pay more per square foot for proximity to campus. In this case, the lack of a significant finding for distance does not support null hypothesis rejection.

According to table 4.6, respondents who Live with Adults Only do not pay significantly more per square foot than respondents who Live with Adults and Children, even though they live significantly closer to campus. The reason the RP/RS ratio is insignificant is that the sub-sample only has 29 observations for the Live with Adults and Children household demographic. N is only 29 because the distance restriction eliminates 11 usable members of this group. In this instance, the preference of the Live with Adults and Children demographic group for a low RP/RS ratio means that many of them locate outside the Missoula urban area. It is the position of this research that the hypothesized relationship is strong between these two groups.
A distance restriction had to be placed on the bulk of the analysis in this section to avoid emphasizing aberrant distance results. However, when looking specifically at these two groups relaxing the distance restriction is not a problem. For the purposes of comparing these two groups only the entire analysis of variance and Bonferroni post-hoc statistics, including all demographic groups and variables, were rerun and are given in appendix g. When the distance restriction is removed, N is sufficiently large for Live with Adults and Children. The mean distance differences remain highly significant with respondents that Live with Adults Only living significantly closer to the university than respondents that Live with Adults and Children. When the distance restriction is removed, the RP/RS ratio mean difference drops slightly and N increases to 40. Once the distance restriction is removed, respondents that Live with Adults and Children pay a significant (.020) amount less per square foot than respondents that Live with Adults Only.

Based on the results and discussion above, there is a trade-off between proximity to campus and residence size for the dollar. Respondents who Live Alone and respondents who Live with Adults Only enjoy greater accessibility to the university, but pay more per square foot. The parents/relatives of respondents who provide then housing and respondents who Live with Adults and Children pay less per square foot, but are located further from the university. On this basis, the second null hypothesis, that no such trade-off would exist, can be rejected. In explaining why students live where they do, it is evident that ‘family’ household types (Live with Adults and Children and Live with Parents/Relatives) prefer a low RP/RS ratio over high accessibility. Conversely, ‘non-
family' household types (Live with Adults Only and Live Alone) prefer accessibility even in the face of a high RP/RS ratio.

There is an important explanatory component to understanding different preferences of demographic groups. Ownership rates vary considerably across demographic groups. The demographic groups that exhibited preferences for low residence price to residence size ratios have higher ownership rates. Respondents that Live with Parents/Relatives do not own their residences, but their parents or relatives often do, and these students usually benefit by paying little or nothing to live in a large house. Over half of respondents that Live with Adults and Children own their residence. Based on the results in this section, one can expect to get more residence for the dollar if the residence is not close to the university. For students in family households, it may be in their best economic interest to purchase more residence for the dollar, at the cost of accessibility.

**Logistic Regression**

Logistic regression will be used to analyze student commuting practices in terms of a binary modal split between car commuting and alternative commuting. The hypothesis tested with logistic regression is that distance can be expected to have systematic impacts on modal split. The corresponding null hypothesis is that distance will not exhibit systematic impacts on commuting mode. The null hypothesis will be tested with the basic distance model. The dependent variable in the basic distance model is a dichotomous variable in which each respondent is coded as a zero if biking, walking, or bussing account for more than 50 percent of trips to campus in a typical week, and coded
as a one otherwise. The independent variable in the basic distance model is distance, which includes the correction for bus commuters. When the hypothesis testing is complete, logistic regression will also be used to develop a more complex multivariate model. The purpose of this model is to compare the relative importance of distance when other variables that affect modal split are included. These variables were identified throughout the research as variables that could be important in describing student commuting practices.

A technical explanation of the reported logistic regression statistics is given in appendix h. Also, a probability plot is displayed, to make model interpretation somewhat easier. Table 4.7 depicts the results of the basic distance logistic regression model.

**Table 4.7 - Basic Distance Logistic Regression Model**

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Wald</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>1.738</td>
<td>77.719</td>
<td>.000</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.432</td>
<td>72.432</td>
<td>.000</td>
</tr>
</tbody>
</table>

Hosmer-Lemeshow Goodness of Fit Statistic = .900, $\hat{B}_0$ Log Likelihood = 547.099, -2 Log Likelihood = 346.738, Model Chi Square = 200.361

The model fit is quite good. A perfect fitting model would result in a Hosmer-Lemeshow goodness-of-fit statistic of 1.00. At a goodness of fit statistic of .900, the model fits very well. The parameter has the expected sign, as distance increases the probability that a student will use a car as the primary mode of commuting increases. The Wald statistic follows a normal distribution so any value above three is statistically
significant. A Wald statistic of 77.719 is very high, indicating a very significant relationship.

The probability plot for the basic distance model is given in figure 4.8.

Figure 4.8 - Basic Distance Model Probability Plot

This plot is a scatter plot of the expected probability that a respondent would commute by car based on the distance variable. It may be instructive to explain the mathematics of this plot in order to make model interpretations transparent. For this, it is necessary to return to the formula for the logistic regression.

\[
y = \frac{e^{b_0 + b_1 x}}{1 + e^{b_0 + b_1 x}}
\]

Where \( y \) is the probability of commuting by car.

Using the Parameters from table 4.7

\[
y = \frac{e^{-2.432 + 1.738(x)}}{1 + e^{-2.432 + 1.738(x)}}
\]
Substituting distances for x in the above equation yields the values in table 4.8.

<table>
<thead>
<tr>
<th>Distance</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>.5 miles</td>
<td>17%</td>
</tr>
<tr>
<td>1.0 mile</td>
<td>33%</td>
</tr>
<tr>
<td>1.5 miles</td>
<td>54%</td>
</tr>
<tr>
<td>2.0 miles</td>
<td>74%</td>
</tr>
<tr>
<td>2.5 miles</td>
<td>87%</td>
</tr>
<tr>
<td>3.0 miles</td>
<td>94%</td>
</tr>
</tbody>
</table>

It should now be clear that the formula for the logistic regression is an approximation of the probability line for a well-fit model, as the calculated percentages in table 4.8 match the curve displayed in figure 4.8. Table 4.8 and figure 4.8 shows how quickly the probabilities change for commuting by car as opposed to commuting by biking, walking, and bussing. The probability that a respondent uses a car as the primary mode of commuting is just 17 percent for respondents that live a half mile from the university. This probability jumps to 54 percent with a modest one mile increase in distance.

This section has shown that distance to the university (also, distance to bus routes) is a significant predictor of modal split. Moreover, this section has shown precisely the rates at which distance effects modal split. The null hypothesis is that distance will not exhibit systematic impacts on commuting mode, so the null hypothesis can be unconditionally rejected.

The basic distance model not only illustrates the strong relationship between residential patterns and commuting practices, it also illustrates some important aspects of commuting behavior. This illustration is enhanced by the map on the following page, *Resident's Primary Mode of Transportation.*
Map 4.4 - Respondent’s Primary Mode of Transportation
(Based on a Weekly Summary)

Legend

- **CAR**
- **CAR POOL**
- **BIKE**
- **WALK**
- **BUS**

**Main Roads**

**Major Rivers**

**Streams**

Mileage Contours

- \(\approx\) 1 mile
- \(\approx\) 3 miles

Jay Harland
March 20, 2001
Base Maps Courtesy
Missouri County
Examining the map, the basic distance model can be tied to physical space in Missoula. A low percentage of students drive within a mile of campus. The space between the mileage contours corresponds to the section of the regression that is changing the fastest. As distance increases, this map shows the rapid decrease in biking and walking as distance ranges from one to three miles. Beyond the three-mile contour, the regression predicts that most all students will drive, and this is exactly what is represented on the map.

It is also important to keep in mind that this model depicts ‘usual’ commuting practices. The data description section showed that respondents tend to bike less when the weather is bad. The logistic regression model shows that at a distance of two miles there is an eighty percent probability that the respondent will get to campus by car. Two miles takes about ten to fifteen minutes to bike. Thus, students who live within easy biking distance drive, even in nice weather. This is shown by the proliferation of red triangles, representing car commuters, on the map beginning at about 1.5 miles.

Finally, the map and the model show that proximity to the university is important. As distance increases, the probability of commuting by car increases rapidly. Walking and biking are the dominant forms of alternative commuting, accounting for 88 percent of the alternative commuters in the model (coded as zero). For 88 percent of respondents using alternative transportation, proximity to the university is a significant predictor of commuting practices.

Having established the significance of distance in predicting modal split, it is important to assess the relative importance of distance when compared to other variables.
that effect commuting practices. This can be accomplished by building a model that
includes variables in addition to distance that are significant in predicting modal split.
The first step in this process is to fit a model to all the variables that could conceivably
affect modal split.

Some of these variables were discussed in the confounding section of the conceptual
background, such as semester credit hours. Credit hours are a proxy for ‘commitment to
school’, signifying the importance of school in a student’s daily activities. Additional
employment and additional employment location could constrain commuting options for
students. Other variables such as owning a car or not owning a bike could constrain
commuting options. The fitted model depicted in table 4.9 shows these variables.

**Table 4.9 - Logistic Regression Model**

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Wald</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>1.599</td>
<td>47.732</td>
<td>.000</td>
</tr>
<tr>
<td>Credits</td>
<td>-0.026</td>
<td>0.350</td>
<td>.554</td>
</tr>
<tr>
<td>Age</td>
<td>-0.010</td>
<td>0.153</td>
<td>.696</td>
</tr>
<tr>
<td>Children in Household</td>
<td>1.438</td>
<td>6.037</td>
<td>.014</td>
</tr>
<tr>
<td>Employment</td>
<td>0.073</td>
<td>0.411</td>
<td>.859</td>
</tr>
<tr>
<td>Additional Employment within a Mile</td>
<td>0.276</td>
<td>0.527</td>
<td>.468</td>
</tr>
<tr>
<td>Car Ownership</td>
<td>0.001</td>
<td>-.0548</td>
<td>.978</td>
</tr>
<tr>
<td>No Bike Ownership</td>
<td>1.047</td>
<td>6.364</td>
<td>.012</td>
</tr>
<tr>
<td>Park Permit</td>
<td>2.942</td>
<td>40.840</td>
<td>.000</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.848</td>
<td>1.052</td>
<td>.007</td>
</tr>
</tbody>
</table>

Hosmer-Lemeshow Goodness of Fit - .770, $B_0$ Log Likelihood = 547.099
$-2 \text{ Log Likelihood} = 262.283$, Model Chi Square = 284.816
Table 4.9 shows a model with at least three variables important for describing commuting practices, besides distance. Ownership of a parking permit is highly significant. Having dependent children in the household and not owning a bike are both moderately significant. Each of these variables will be used to develop a final model. Several things will need to be addressed before the final model can be completed. First, the justification and rationale for each variable must be developed. Second, each variable must be checked for interaction factors. Third, a probability plot for each variable with distance will be presented showing how model predictions deviate, only.

The significance of the parking permit variable in the model is unsurprising. Parking permits are over one hundred dollars, and the decision to purchase a permit can be conceptualized as a conscious decision to drive. Once this decision has been made, it is unlikely that other factors such as residential location will affect modal split. However, respondents that own a parking permit generally live farther from campus. The spatial patterns of parking permits are displayed in the map entitled Parking Permit Holders versus No Parking Permit in appendix f.

The presence of dependent children in the household was also found to be significant in the model. This finding supports the popular notion that child rearing results in increased vehicle trips. The justification for including this variable in the final model is that the presence of dependent children in the household can provide increased pressure to drive to campus, all other variables equal. Other household demographics may predict modal split, but because residential location differs significantly by demographic group most of the prediction is due to an interaction with distance. The
presence of dependent children is the only household demographic factor that results in modal split differences that are in addition to the relationship between household demographics and distance.

The last variable that will be included in the final model is bike ownership. 72 percent of respondents own a bike. Thus, owning a bike was coded as a one and not owning a bike was coded as a zero because it is the interesting trait. The reasoning behind including no bike ownership in the model is because it is difficult for a respondent to bike to campus if the respondent does not own a bike.

Table 4.10 shows a series of models. Each variable was fitted in a model as a lone variable, as a bivariate with distance, and with an interaction term with distance.

Table 4.10 - Final Model Development

<table>
<thead>
<tr>
<th>Model</th>
<th>Constant</th>
<th>Children Present in Household</th>
<th>No Bike Ownership * Distance</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.754</td>
<td>-1.718</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-2.616</td>
<td>1.496</td>
<td></td>
<td>1.770</td>
</tr>
<tr>
<td>3</td>
<td>-2.691</td>
<td>.174</td>
<td>-1.860</td>
<td>3.555</td>
</tr>
<tr>
<td>4</td>
<td>-.146</td>
<td></td>
<td>1.650</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-1.381</td>
<td>-1.270</td>
<td></td>
<td>1.715</td>
</tr>
<tr>
<td>6</td>
<td>-1.673</td>
<td>-1.673</td>
<td>-311</td>
<td>1.971</td>
</tr>
<tr>
<td>7</td>
<td>-.533</td>
<td></td>
<td></td>
<td>3.289</td>
</tr>
<tr>
<td>8</td>
<td>-2.962</td>
<td></td>
<td>2.960</td>
<td>1.660</td>
</tr>
<tr>
<td>9</td>
<td>.473</td>
<td></td>
<td>-3.533</td>
<td>1.243</td>
</tr>
</tbody>
</table>

Bold indicates variable significance at alpha = .01
There is no apparent interaction between distance and any of the other variables to be included in the final model. Thus, no interaction terms will be included in the final model. The only technical matter of importance is that the Hosmer-Lemeshow goodness-of-fit statistic plummets whenever the parking permit variable is introduced in the model. It is not entirely clear why this occurs, but because the variable is being added to an excellent fitting model it does not present a problem.

Because there is no interaction to be concerned with, the focus will be strictly on the bivariate models depicted in table 4.10. Focusing on these models makes it possible to isolate the relationship between distance and each of the other variables to be included in the final model. The parameters in table 4.10 exhibit the expected signs. Owning a parking permit increases the expected probability that a respondent will drive as does the presence of a dependent child in the household. Not owning a bike decreases the probability that a respondent will bike, increasing the probability that the respondent will drive. Figures 4.9- 4.11 are the probability plots for the three bivariate models

**Figure 4.9 - Distance and Parking Permit Probability Plot**
Figure 4.10 - Presence of Children in the Household Probability Plot

Figure 4.11 - No Bike Ownership Probability Plot
The above scatter plots show how these variables work in a bivariate model with distance. Each plot represents a shift factor in the data. Essentially, these variables represent a shift factor when added to the basic distance model. The shift factors operate according to expectations. Not owning a parking permit, not having a dependent child in the household, and not owning a bike all push the main curve down and to the right.

With the relationships between the individual variables and distance established, it is possible to introduce the final model. Table 4.11 shows the results of the fitted model and figure 4.12 shows probability plots for both the basic distance model and the final model.

**Table 4.11 - Final Model**

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Wald</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>1.649</td>
<td>55.149</td>
<td>.000</td>
</tr>
<tr>
<td>Children in Household</td>
<td>1.431</td>
<td>7.073</td>
<td>.008</td>
</tr>
<tr>
<td>No Bike Ownership</td>
<td>1.048</td>
<td>6.698</td>
<td>.010</td>
</tr>
<tr>
<td>Park Permit</td>
<td>2.894</td>
<td>42.874</td>
<td>.000</td>
</tr>
<tr>
<td>Constant</td>
<td>-3.277</td>
<td>80.702</td>
<td>.000</td>
</tr>
</tbody>
</table>

Hosmer-Lemeshow Goodness-of-Fit Statistic .344, $B_0$ Log Likelihood = 547.099, -2 Log Likelihood = 263.618, Model Chi Square = 283.481
The purpose of building the final model is to compare the relative effect of distance once the other important variables for predicting modal split are included. Distance is still the dominant variable. Ownership of a parking permit is significant, but still not as significant as distance.

The results of this section show two basic things. First, the most significant variable for predicting modal split is distance. Second, variables other than distance can be important for describing modal split. These variables seem to act as a shift factor for distance in the model, altering but not supplanting the power of distance. For 95 percent of the survey respondents, proximity to the university is equivalent to the distance variable. Modal split between car commuting and alternative commuting is a good
variable for describing commuting practices. Therefore, residential location, relative to the university, is the single most important predictor of commuting practices.

**Summary of Analysis**

The analysis in this chapter has shown three general features. First, the variable important for explaining respondent residential price is predominantly the residence size, and quality of the residence is probably quite important as well. The subtle effect of proximity to the university on residence price is likely drowned out by variability in the quality of Missoula’s housing stock when applying the Alonso model. Second, the effect of proximity to the university is not as subtle when residential patterns are analyzed by demographic group. Family household types sacrifice accessibility to the university in exchange for more residence for the dollar, while non-family household types pay a premium per square foot to live close to the university. Third, residential location is the most significant variable for predicting whether a respondent uses a car as the primary means of commuting under ‘usual’ commuting conditions.

A discussion of the research question will begin the conclusion chapter. Followed by a look at some general policy implications from these research findings. The last part of the conclusion chapter will offer a strategy for improving housing and commuting opportunities for students at the University of Montana.
CHAPTER 5

CONCLUSION

Discussion of the Research Question

The research question is whether there is a connection between student residential patterns and commuting practices, and if so, how should this connection be understood. The analysis chapter presented three connections between student residential patterns and commuting practices.

The conceptual background introduced the Alonso model to explain the economics of student residential choices. Accessibility to campus and residence size were used to test the strength of the Alonso theoretical model. Residence price for rentals was found to increase about twenty-six cents for every additional square foot of residence size. Accessibility was expected to measure the consideration given to commuting costs in explaining student’s residential decisions. In this restrictive model, accessibility to the university was not found to be an important variable.

An alternative explanation for the insignificance of distance is that variability in housing quality is not constant or random across the study area and that the effects of
commuting costs may be too small to measure effectively. Assuming the alternative explanation is the correct one, housing quality and housing size appear more important than commuting costs in determining a student's residential decisions.

The next hypothesis suggested that demographic groups exhibit different preferences for residential value (RP/RS Ratio was the measure) and different preferences for accessibility to the University of Montana. Analysis of survey results showed in particular that demographic preferences for a low RP/RS ratio accompanied low accessibility to the university, and vice versa. Thus, distance to the university appears to be more important in describing student residential patterns when analyzed in terms of demographic groups. Thus, the connection between student residential patterns and commuting practices implies different things for different types of student households. Non-family household types' residential decisions are sensitive to distance factors, whereas family household types' residential decisions are sensitive to residential value.

A central part of this thesis was that residential decisions would impact commuting practices. Modal split was the measure used to assess this impact, and distance was used as the measure for residential patterns. The analysis in the previous chapter showed a very strong relationship between residential patterns and commuting practices. Students who choose to live far from campus generally drive a car, and students who choose a residence close to campus generally bike or walk.

The above discussion and data analysis make it possible to assess the validity of the research question. The connection between student residential patterns and commuting practices is very strong. Residential location affects commuting mode. Accessibility is
also important for student residential decisions, mainly for respondents that Live Alone and Live with Adults Only. These two demographic groups accounted for 76.2 percent of respondents, spending the most per square foot, and living about a mile and a half to campus on average.

**General Implications**

What parts of this research can be generalized and offer insight into the connection between residential patterns and commuting practices beyond the study area?

First, commuting modes, other than the car, are very sensitive to distance factors less than two miles. Whether it is walking to the bus or walking to campus, as distance increases commuting mode changes rapidly. The strength of this finding is somewhat dependent on a disincentive to drive. In the case of the University of Montana, there are fewer parking spaces than there are students. This situation decreases the convenience of commuting by car, making other forms of transportation more desirable.

For planners and policy makers this means commuting considerations are important in residential decisions, but there are many other important factors. Thus, residential decisions would not appear to be nearly as sensitive to commuting considerations as commuting practices are to residential decisions. Two questions illustrate the dilemma. How do you alleviate traffic and congestion when residential patterns result in people living in places that a further than most people are willing to bike or walk? Furthermore, how can residential patterns be changed to mitigate congestion, when transportation is only one important factor among many for residential decisions?
This thesis has shown the connection between student residential patterns and commuting practices is different for different household types. Even when studying a narrow segment of society, the college student, there is enough diversity to lead to important differences in the connection between residential patterns and commuting. This means monolithic programs to improve housing and commuting are likely to overlook important household differences.

**Implications for Student Housing and Commuting in Missoula**

The research in this thesis has some important implications for University of Montana student housing, and commuting. This research shows solutions addressing student housing and transportation will likely impact land use and traffic around the university, if enrollment continues to increase. Increasing student densities will impact the university area with a higher intensity of land use. However, failing to increase densities will almost certainly result in serious vehicular traffic increases. This is especially likely if additional parking is added at the University of Montana. Based on this research, these are the realities of University of Montana student housing and commuting.

This does not mean that this research does not provide insight into some promising areas for the University of Montana. For example, this research suggests that the presence of a child in the household can affect both residential decisions and commuting practices resulting in more vehicle miles traveled. It would appear that a program, such
as a program to coordinate transportation to day care services and after school activities, could reduce parental dependence on the automobile.

Ownership of a parking permit is an important predictor of commuting practices. The purchase of a parking permit is a conscious decision to drive. There are many students who live close to direct bus routes who purchase parking permits. There are many students who live within easy biking distance and still purchase a parking permit. In terms of changing student behavior, these students present a considerable challenge. These students are committed to using a car, in spite of the limited parking on campus.

Also, it is evident that the bicycle is the most important alternative commuting mode for reducing vehicle miles traveled. However, this commuting mode has been shown to be adversely affected by bad weather. This means that there is an opportunity to try and increase the range of conditions in which a student is willing to ride a bike. The ASUM Office of Transportation is looking at a bike winterizing rally, which could include a subsidy allowing students to get studded bike tires and fenders at little or no cost. This would make it safer to bike on icy winter streets. Also, bike ownership is a slight indicator for commuting mode. This means that increasing bicycle availability alone could reduce vehicle miles traveled. This finding is upheld by the success of the ASUM Cruiser Co-Op bike check-out program.

A Strategy for Better Housing and Commuting Options

The University of Montana and the City of Missoula must work together to improve student housing and transportation in the future. The findings in this thesis show
that the connections between student residential patterns and commuting practices are extensive in scale, geography, and demography.

Independent decision making by these two entities can result in serious problems. In the early nineties, University enrollment outstripped the supply of housing in Missoula resulting in students protesting by camping in the oval. Currently, the City of Missoula is experiencing an affordable housing crisis. If enrollment swells, without a plan to provide housing for the increasing number of students, this affordable housing crisis is likely to deepen. Meanwhile, if the only housing units available to these new students are located more than 1.5 miles from campus there is going to be an enormous increase in vehicular traffic, pressuring already strained transportation infrastructures.

These examples show the vested interest that the City of Missoula and the University of Montana have to work together. However, as countless newspaper articles on planning attest, vested interest alone is rarely incentive enough to develop a successful plan. This thesis suggests housing near campus is one of the best strategies to reduce vehicle miles traveled in Missoula. However, there is no collaborative plan between the City of Missoula and the University of Montana to locate students near campus. There are at least two important obstacles to a successful collaborative plan to improve options for student housing, consequently improving student transportation.

The first obstacle is financial. The City of Missoula and the University of Montana have immediate financial reasons not work together to develop a plan to locate students near campus. Given current budget constraints, implementation of a plan that included housing development could not be fully funded by the University of Montana. If any
student housing development were to be successful, it would probably not be able to be developed for less than it could be rented. In order to meet these conditions, the City of Missoula would almost certainly have to subsidize the development. Thus, a successful plan would probably be a source of expenditure for the City of Missoula. The current budget constraints also present a problem for the University of Montana. Funding for the University of Montana is largely dependent on enrollment. The greater the enrollment the larger the budget. It is unlikely that the City of Missoula would agree to help fund student housing to reduce congestion and improve affordable housing, only to have the University of Montana immediately increase enrollment beyond capacity of the new housing. If the University of Montana limited enrollment, this would have serious impacts on education funding.

The second obstacle is that the main goals of the City of Missoula and the University of Montana are not coincidental. The City of Missoula's main goal is to provide basic local government functions for its citizenry. Students only represent about 10 percent of the Missoula citizenry. Thus, housing and transportation solutions for students must be balanced against the housing and transportation needs of other citizens. Alternatively, the main goal of the University of Montana is to provide a quality education for students in general, and Montanans in particular. Thus, investments in student housing and transportation are only necessary to the extent that they are adequate for a quality education. Therefore, student housing and commuting may not be a top priority for the University of Montana or the City of Missoula.
The obstacles in the previous two paragraphs are extensive. However, the citizens of Missoula and the students at the University of Montana can have considerable influence over these two entities. On a personal note, I have volunteered considerable time in the last year to improve student transportation, sitting on both the Associated Students at the University of Montana Office of Transportation Board and University of Montana Transportation Task Force. I can attest that any chance to develop a collaborative plan for student housing and transportation, must begin with a unified front between Missoula citizens and the student body. Before the City of Missoula and the University of Montana can be expected to work together, it is necessary for Grizzlies and Missoulians to lock arms and demand a realistic plan for student housing and transportation.
Appendix A - Representative Sampling Essay

There are two possible ways to answer the question as to how the elimination of students from the registrar’s database could have affected the sample. The first would be a rational approach. From a logical standpoint, there are two types of sampling error that could result from the methods employed to get an accurate address for all respondents. One could argue that students who were omitted from the analysis, because they lacked a reasonable address in the Missoula area, could represent a bias toward students that are not as well rooted in Missoula. Or put another way, the sample might be over selecting for juniors, seniors, grad students, and students who were residents of the Missoula area prior to enrollment. These students have addresses in Missoula that are stable and thus worth transmitting to the registrar. There is some reason to think that this effect is not likely to be substantial. In all likelihood, freshmen that live in the dorms probably comprise the largest component of this omitted group. While dorm students are included in a few select components of the research, they are not the primary focus of the research.
Thus, the main effect of omission is probably located in a group of students that are not especially important for the analysis. With respect to inaccurate addresses one could also argue the reverse, that students were kept in the data set with an address in Stevensville that is a parental address when in fact they live in Missoula. This could present a serious problem in the analysis because the student response would be associated with the Missoula address, but the analysis would have that response located in Stevensville. Some things in the survey were designed to combat this problem, and hopefully very few students could have slipped through causing this type of problem.

The other significant source of omission from the analysis stemmed from Missoula P.O. boxes. This problem appears to present fewer problems from a rationalistic standpoint. While a number of students were eliminated for this reason, it is not clear what systematic characteristics a student with a P.O. box would possess that a student with a standard address would not possess.

In summary, the methods employed to eliminate students with logistical barriers to geocoding could have produced some sources of sampling bias from a rationalist perspective. However, there does not appear to be any clear line of reasoning that would support a particular bias. Moreover, the only bias that would seem to have the power to result in serious errors would be if responses were tied to the wrong address. However, this source of error was recognized prior to surveying. Thus, attempts were made to address this potential source of error throughout the data collection process.

Alternatively, an empirical assessment could have been performed to determine the degree to which these potential sources of error could have resulted from sample selection
bias. This approach was not undertaken in this research for two reasons. Practicality was the primary reason that an empirical approach was not undertaken. This study was time sensitive. According to the registrar, Phil Bain, the registrar's database is most accurate in early October. Thus, the database was received in early October, and in order to complete data collection by the end of the semester the first mailing had to be sent by the end of October. A two-week window was available to select the sample, geocode the sample, and execute the mailing. This is a relatively small time frame, and as a result any attempt to empirically assess the effects of eliminating the aforementioned students from the analysis would have resulted in an introduction of other sources of error such as not being able to complete the study in a single semester. The second reason why no empirical assessment was made stems for the rational analysis above. There is little reason to think that the process used to eliminate students from the study would have contained a systematic bias strong enough to result in a data set that was grossly unrepresentative.
### Appendix B - Paired T-Tests for Geographics of Respondents

Comparison of Means

Means given in miles.

<table>
<thead>
<tr>
<th></th>
<th>Pared-down Registrar’s</th>
<th>Sample</th>
<th>Survey Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Data- All Distances</td>
<td>4.519</td>
<td>4.271</td>
<td>.536</td>
</tr>
<tr>
<td>All Data- All Distances</td>
<td>4.271</td>
<td>3.091</td>
<td>.043</td>
</tr>
<tr>
<td>All Data- All Distances</td>
<td>4.519</td>
<td>3.091</td>
<td>.035</td>
</tr>
<tr>
<td>All Data- Ten Miles Sample to Student Body</td>
<td>1.493</td>
<td>1.483</td>
<td>.846</td>
</tr>
<tr>
<td>All Data- Ten Miles Response to Student Body</td>
<td>1.483</td>
<td>1.502</td>
<td>.798</td>
</tr>
<tr>
<td>All Data- Ten Miles Sample to Response</td>
<td>1.493</td>
<td>1.502</td>
<td>.913</td>
</tr>
<tr>
<td>All Data- Three Miles Sample to Student Body</td>
<td>1.000</td>
<td>1.004</td>
<td>.902</td>
</tr>
<tr>
<td>All Data- Three Miles Response to Student Body</td>
<td>1.004</td>
<td>1.017</td>
<td>.759</td>
</tr>
<tr>
<td>All Data- Three Miles Response to Sample</td>
<td>1.000</td>
<td>1.020</td>
<td>.732</td>
</tr>
<tr>
<td>No Dorm-All Distances Sample to Student Body</td>
<td>5.763</td>
<td>5.596</td>
<td>.746</td>
</tr>
<tr>
<td>No Dorm- All Dist. Response to Student Body</td>
<td>5.596</td>
<td>3.923</td>
<td>.024</td>
</tr>
<tr>
<td>No Dorm- All Distances Response to Sample</td>
<td>5.763</td>
<td>3.923</td>
<td>.031</td>
</tr>
<tr>
<td>No Dorm- Ten Miles Sample to Student Body</td>
<td>1.932</td>
<td>1.976</td>
<td>.456</td>
</tr>
<tr>
<td>No Dorm- Ten Miles Response to Student Body</td>
<td>1.976</td>
<td>1.928</td>
<td>.582</td>
</tr>
<tr>
<td>No Dorm - Ten Miles Sample to Response</td>
<td>1.932</td>
<td>1.928</td>
<td>.967</td>
</tr>
</tbody>
</table>
Appendix C - Survey Questionnaire
The 2000 University of Montana Housing and Transportation Survey

1. How many people live at your residence, including you?

2. What are their ages? (list yours first)

3. How many of them are related to you?

4. Do you rent or own your residence?
   □ Rent □ Own □ Live with parents/relatives

5. About what is your rent or mortgage per month (without utilities)? Be sure to include the total regardless of whether you are responsible for any of it (government subsidy, live with parents, etc.).

   $ __________ Your
   $ __________ Total

   □ I do not have to pay

   For question 7 answer either section a, or section b, or section c. Only one of these sections will apply to your residence and the others will not.

6. About what are the utilities per month (without phone)?

   $ __________ Your
   $ __________ Total

7(a). Do you live in...
   □ A 1 bedroom apartment □ A UM Dorm
   □ A 1 bedroom detached house □ A rented room
   □ Other 1 bedroom, _________

   How big is your residence in square feet (circle one)?

   Less
   220 to 450
   450 to 600
   More
   □ Less
   □ 220 to 450
   □ 450 to 600
   □ More

7(b). Do you live in...
   □ 2-3 bedroom apartment □ 2-3 bedroom duplex
   □ 2-3 bedroom triplex □ Single wide mobile
   □ 2-3 bedroom detached house □ Other 2-3 _________

   How big is your residence in square feet (circle one)?

   Less
   220 to 450
   450 to 600
   More
   □ Less
   □ 220 to 450
   □ 450 to 600
   □ More

7(c). Do you live in...
   □ 4+ bedroom apartment □ 4+ bedroom duplex
   □ 4+ bedroom triplex □ Fraternity/Sorority
   □ 4+ bedroom detached house □ Other 4+ _________

   How big is your residence in square feet (circle one)?

   Less
   450 to 600
   600 to 800
   More
   □ Less
   □ 450 to 600
   □ 600 to 800
   □ More

   More
   □ Less
   □ 20 to 80
   □ 80 to 200
   □ More

   □ Less
   □ 20 to 80
   □ 80 to 200
   □ More

   □ Less
   □ 300 to 500
   □ 500 to 700
   □ More

   □ Less
   □ 300 to 500
   □ 500 to 700
   □ More

8. Do you have a job(s) in addition to school?
   □ Yes □ No

9. If you have a job(s) in addition to school, is it...
   (Check both if both apply.)
   □ On campus or within a mile of campus
   □ Greater than a mile from campus

10. Do you have a car in Missoula?
    □ Yes □ No

11. Do you have a bicycle in Missoula?
    □ Yes □ No

12. How many cars are at your household?

13. Do you have a UM parking permit?
    □ Yes □ No

14. Was this survey forwarded to you from a previous address?
    □ Yes, current address: _______________________
    □ No
If you live in a dorm on campus, then you are not required to answer questions 15-20. For questions 15-20, think as detailed as possible about how you typically commute to campus. Please write the correct number in each applicable box and leave the rest blank. Please use fractions to represent trips utilizing more than one mode (the park and ride at Dornblaser, is an example of a commute utilizing more than one mode). The purpose of this survey is to measure the relationship between housing choice and commuting practices. Thus, a commute occurs any time you travel from home and end up at campus. If you drove from your house to work and then drove to campus for class, then this would count as one commute. Once you are on campus, if you biked from campus to work and biked back to campus (but did not bike home for lunch), then this would not count as a commute. Also, remember to fill in only the boxes that apply to the commuting methods you utilize.

15. Circle the day of the week that you are completing this survey.

<table>
<thead>
<tr>
<th></th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usually</td>
<td>Usually</td>
<td>Usually</td>
<td>Usually</td>
<td>Usually</td>
<td>Usually</td>
<td>Usually</td>
<td>Usually</td>
</tr>
</tbody>
</table>

16. Number of commutes to campus in your car

17. Number of commutes to campus as a passenger in someone else's car

18. Number of commutes to campus on a bike

19. Number of commutes to campus by walking

20. Number of commutes to campus by bus
Appendix D - Essay on Methodology for Car-Poolers and Bus Users

Respondents whose primary mode of commuting was as a car passenger represent about 5 percent of the respondents in the analysis so this does not represent a serious problem for the logistic model. The problem is only conceptual in a very general sense. It simply depends on how ‘alternative transportation’ is defined. In terms of overall transportation and land use planning this does not seem to be a major issue. If 5 percent of students ride with a roommate instead of driving themselves, this is not going to have huge implications for infrastructures, etc. In a sense, car pooling is somewhere between alternative transportation and driving alone. This is an example of an area where the research in this thesis could be extended. One could research the variables that predict why only 5 percent of the car commuting students are car passengers.

The second issue of particular interest for the logistic regression is bus commuters. Respondents whose primary mode of commuting is by bus represent about 5 percent of the total responses. These students represent a little wrinkle for the model and the hypothesis testing procedure. For transit users, proximity to the university is not nearly as important as proximity to the bus. One would not expect distance to the university to be a good predictor of transit use for this reason. There are two ways to deal with this situation. The first is to leave these students out of the model. While there is nothing
wrong with this approach, it does diminish the predictive power of the model. The other way to deal with this issue is to include these students in the model, but to correct the distance variable so that it reflects proximity to a logical bus route. This strategy keeps the bus users in model, strengthening the overall predictive power of the model. This was the implemented strategy.

This strategy appears appropriate from a transportation perspective, but the implications for hypothesis testing must be addressed. Discussing the rejection of the null hypothesis, it has been proposed that if distance to the university is an effective predictor for categorizing a student as an alternative commuter then the null hypothesis may be rejected. However, this logic has now been compromised somewhat. The actual model will be testing whether distance to the university for 95 percent of the respondents and distance to the bus for 5 percent of the respondents is an effective predictor for categorizing a student as an alternative commuter. This does present a potential source of error with respect to null hypothesis rejection, but this error should not be over stated. If 95 percent of the model is based on distance to campus and the model works well, then the 5 percent bus students should not overwhelm a decision to reject the null hypothesis.

The other problem is that the ‘bus distance correction’ is an ad hoc prescription. There are hundreds of students that live close to a direct bus route that do not take the bus, but the distances were only corrected for students that use the bus. This is a slight contradiction when viewed from the perspective of residential patterns affecting commuting practices. There is no real way to avoid this contradiction in the model. Again, this issue is not damaging enough to warrant serious concern. It is better to put
the 'bus distance correction' in perspective in terms of the model. Using the 'bus distance correction' in the model is going to emphasize how proximity to a bus route affects commuting practices. Some residences are close to bus routes, others are not. Residential location still affects commuting practices, but the influence is somewhat less rigid. When viewed from this perspective, the 'bus distance correction' is less of a problem than it first appears.
### Appendix E - Paired T-Tests Comparing Commuting Trips

#### Comparison of Usual to Bad Weather Commuting Practices

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean Diff.</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday by Car</td>
<td>-0.0226</td>
<td>0.3264</td>
<td>-1.382</td>
<td>397</td>
<td>0.168</td>
</tr>
<tr>
<td>Tuesday by Car</td>
<td>-0.0126</td>
<td>0.3071</td>
<td>-0.816</td>
<td>397</td>
<td>0.415</td>
</tr>
<tr>
<td>Wednesday by Car</td>
<td>-0.0264</td>
<td>0.3133</td>
<td>-1.680</td>
<td>397</td>
<td>0.094</td>
</tr>
<tr>
<td>Thursday by Car</td>
<td>-0.0196</td>
<td>0.2847</td>
<td>-1.373</td>
<td>397</td>
<td>0.170</td>
</tr>
<tr>
<td>Friday by Car</td>
<td>-0.0295</td>
<td>0.3056</td>
<td>-1.733</td>
<td>321</td>
<td>0.084</td>
</tr>
<tr>
<td>Monday by Car Pool</td>
<td>-0.0151</td>
<td>0.1580</td>
<td>-1.904</td>
<td>397</td>
<td>0.058</td>
</tr>
<tr>
<td>Tuesday by Car Pool</td>
<td>-0.0126</td>
<td>0.1660</td>
<td>-1.510</td>
<td>397</td>
<td>0.132</td>
</tr>
<tr>
<td>Wednesday by Car Pool</td>
<td>-0.0075</td>
<td>0.1663</td>
<td>-0.904</td>
<td>397</td>
<td>0.366</td>
</tr>
<tr>
<td>Thursday by Car Pool</td>
<td>-0.0075</td>
<td>0.1504</td>
<td>-1.000</td>
<td>397</td>
<td>0.318</td>
</tr>
<tr>
<td>Friday by Car Pool</td>
<td>-0.0025</td>
<td>0.1328</td>
<td>-0.378</td>
<td>397</td>
<td>0.706</td>
</tr>
<tr>
<td>Monday by Bike</td>
<td>0.1476</td>
<td>0.4248</td>
<td>6.932</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Tuesday by Bike</td>
<td>0.1187</td>
<td>0.3934</td>
<td>6.020</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Wednesday by Bike</td>
<td>0.1476</td>
<td>0.4248</td>
<td>6.932</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Thursday by Bike</td>
<td>0.1246</td>
<td>0.3983</td>
<td>6.239</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Friday by Bike</td>
<td>0.1175</td>
<td>0.3783</td>
<td>6.194</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Monday by Walking</td>
<td>-0.0075</td>
<td>0.3871</td>
<td>-0.389</td>
<td>397</td>
<td>0.698</td>
</tr>
<tr>
<td>Tuesday by Walking</td>
<td>-0.0025</td>
<td>0.3566</td>
<td>-0.141</td>
<td>397</td>
<td>0.888</td>
</tr>
<tr>
<td>Wednesday by Walking</td>
<td>-0.0151</td>
<td>0.3633</td>
<td>-0.828</td>
<td>397</td>
<td>0.408</td>
</tr>
<tr>
<td>Thursday by Walking</td>
<td>-0.0121</td>
<td>0.3451</td>
<td>-0.697</td>
<td>397</td>
<td>0.486</td>
</tr>
<tr>
<td>Friday by Walking</td>
<td>0.0025</td>
<td>0.3806</td>
<td>0.132</td>
<td>397</td>
<td>0.895</td>
</tr>
<tr>
<td>Monday by Bus</td>
<td>-0.0371</td>
<td>0.2115</td>
<td>-3.495</td>
<td>397</td>
<td>0.001</td>
</tr>
<tr>
<td>Tuesday by Bus</td>
<td>-0.0396</td>
<td>0.2337</td>
<td>-3.378</td>
<td>397</td>
<td>0.001</td>
</tr>
<tr>
<td>Wednesday by Bus</td>
<td>-0.0408</td>
<td>0.2211</td>
<td>-3.685</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Thursday by Bus</td>
<td>-0.0383</td>
<td>0.2157</td>
<td>-3.543</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Friday by Bus</td>
<td>-0.0408</td>
<td>0.2211</td>
<td>-3.685</td>
<td>397</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### Number of Usual Trips by Day of the Week

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean Yest</th>
<th>Mean Usual</th>
<th>Mean Diff.</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yesterday Car - Monday Car</td>
<td>0.6267</td>
<td>0.581</td>
<td>0.0451</td>
<td>0.4422</td>
<td>1.225</td>
<td>143</td>
<td>0.223</td>
</tr>
<tr>
<td>Yesterday Car Pool - Monday</td>
<td>0.1169</td>
<td>0.097</td>
<td>0.0195</td>
<td>0.3519</td>
<td>0.687</td>
<td>153</td>
<td>0.493</td>
</tr>
<tr>
<td>Yesterday Bike - Monday Bike</td>
<td>0.3759</td>
<td>0.403</td>
<td>-0.0276</td>
<td>0.3716</td>
<td>-0.894</td>
<td>144</td>
<td>0.373</td>
</tr>
<tr>
<td>Yesterday Walk - Monday Walk</td>
<td>0.2119</td>
<td>0.238</td>
<td>-0.0265</td>
<td>0.3820</td>
<td>-0.852</td>
<td>150</td>
<td>0.396</td>
</tr>
<tr>
<td>Yesterday Bus - Monday Bus</td>
<td>0.0597</td>
<td>0.079</td>
<td>-0.0194</td>
<td>0.1791</td>
<td>-1.345</td>
<td>154</td>
<td>0.181</td>
</tr>
<tr>
<td>Yesterday Car - Tuesday Car</td>
<td>0.7727</td>
<td>0.663</td>
<td>0.1091</td>
<td>0.3146</td>
<td>2.571</td>
<td>54</td>
<td>0.013</td>
</tr>
<tr>
<td>Yesterday Pool - Tuesday Pool</td>
<td>0.1034</td>
<td>0.068</td>
<td>0.0345</td>
<td>0.3226</td>
<td>0.814</td>
<td>57</td>
<td>0.419</td>
</tr>
<tr>
<td>Yesterday Bike - Tuesday Bike</td>
<td>0.2679</td>
<td>0.348</td>
<td>-0.0804</td>
<td>0.3787</td>
<td>-1.588</td>
<td>55</td>
<td>0.118</td>
</tr>
<tr>
<td>Yesterday Walk - Tuesday Walk</td>
<td>0.3103</td>
<td>0.293</td>
<td>0.0172</td>
<td>0.2288</td>
<td>0.574</td>
<td>57</td>
<td>0.568</td>
</tr>
<tr>
<td>Yesterday Bus - Tuesday Bus</td>
<td>0.0759</td>
<td>0.060</td>
<td>0.0172</td>
<td>0.1313</td>
<td>1.000</td>
<td>57</td>
<td>0.322</td>
</tr>
<tr>
<td>Yesterday Car - Wednesday Car</td>
<td>0.5625</td>
<td>0.625</td>
<td>-0.0625</td>
<td>0.3536</td>
<td>-1.000</td>
<td>31</td>
<td>0.325</td>
</tr>
<tr>
<td>Yesterday Pool - Wednesday Pool</td>
<td>0.0810</td>
<td>0.067</td>
<td>0.0135</td>
<td>0.2496</td>
<td>0.329</td>
<td>36</td>
<td>0.744</td>
</tr>
<tr>
<td>Yesterday Bike - Wednesday Bike</td>
<td>0.4324</td>
<td>0.493</td>
<td>-0.0608</td>
<td>0.3302</td>
<td>-1.120</td>
<td>36</td>
<td>0.270</td>
</tr>
<tr>
<td>Yesterday Walk - Wednesday Walk</td>
<td>0.1892</td>
<td>0.189</td>
<td>0.0000</td>
<td>0.7454</td>
<td>0.000</td>
<td>36</td>
<td>1.000</td>
</tr>
<tr>
<td>Yesterday Bus - Wednesday Bus</td>
<td>0.0810</td>
<td>0.121</td>
<td>-0.0405</td>
<td>0.1817</td>
<td>-1.357</td>
<td>36</td>
<td>0.183</td>
</tr>
<tr>
<td>Yesterday Car - Thursday Car</td>
<td>0.6250</td>
<td>0.625</td>
<td>0.0000</td>
<td>0.3651</td>
<td>0.000</td>
<td>15</td>
<td>1.000</td>
</tr>
<tr>
<td>Yesterday Pool - Thursday Pool</td>
<td>0.0556</td>
<td>0.000</td>
<td>0.0556</td>
<td>0.2357</td>
<td>1.000</td>
<td>17</td>
<td>0.331</td>
</tr>
<tr>
<td>Yesterday Bike - Thursday Bike</td>
<td>0.0588</td>
<td>0.294</td>
<td>-0.2353</td>
<td>0.4372</td>
<td>-2.219</td>
<td>16</td>
<td>0.041</td>
</tr>
<tr>
<td>Yesterday Walk - Thursday Walk</td>
<td>0.2778</td>
<td>0.222</td>
<td>0.0556</td>
<td>0.2357</td>
<td>1.000</td>
<td>17</td>
<td>0.331</td>
</tr>
<tr>
<td>Yesterday Car - Friday Car</td>
<td>0.2000</td>
<td>0.300</td>
<td>-0.1000</td>
<td>0.5676</td>
<td>-0.557</td>
<td>9</td>
<td>0.591</td>
</tr>
<tr>
<td>Yesterday Pool - Friday Pool</td>
<td>0.2000</td>
<td>0.100</td>
<td>0.1000</td>
<td>0.3162</td>
<td>1.000</td>
<td>9</td>
<td>0.343</td>
</tr>
<tr>
<td>Yesterday Bike - Friday Bike</td>
<td>0.1000</td>
<td>0.200</td>
<td>0.1000</td>
<td>0.3162</td>
<td>-1.000</td>
<td>9</td>
<td>0.343</td>
</tr>
</tbody>
</table>

### Number of Usual Car Trips by Day of the Week

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean Diff.</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday - Tuesday</td>
<td>0.0300</td>
<td>0.5513</td>
<td>1.087</td>
<td>397</td>
<td>0.278</td>
</tr>
<tr>
<td>Monday - Wednesday</td>
<td>0.0251</td>
<td>0.3558</td>
<td>1.409</td>
<td>397</td>
<td>0.160</td>
</tr>
<tr>
<td>Monday - Thursday</td>
<td>0.0576</td>
<td>0.5803</td>
<td>1.987</td>
<td>397</td>
<td>0.048</td>
</tr>
<tr>
<td>Monday - Friday</td>
<td>0.0754</td>
<td>0.4496</td>
<td>3.345</td>
<td>397</td>
<td>0.001</td>
</tr>
<tr>
<td>Tuesday - Wednesday</td>
<td>-0.0049</td>
<td>0.5509</td>
<td>-0.177</td>
<td>397</td>
<td>0.859</td>
</tr>
<tr>
<td>Tuesday - Thursday</td>
<td>0.0278</td>
<td>0.3122</td>
<td>1.774</td>
<td>397</td>
<td>0.077</td>
</tr>
<tr>
<td>Tuesday - Friday</td>
<td>0.0454</td>
<td>0.5351</td>
<td>1.691</td>
<td>397</td>
<td>0.092</td>
</tr>
<tr>
<td>Wednesday - Thursday</td>
<td>0.0327</td>
<td>0.5876</td>
<td>1.109</td>
<td>397</td>
<td>0.268</td>
</tr>
<tr>
<td>Wednesday - Friday</td>
<td>0.0503</td>
<td>0.3789</td>
<td>2.646</td>
<td>397</td>
<td>0.008</td>
</tr>
<tr>
<td>Thursday - Friday</td>
<td>0.0176</td>
<td>0.5709</td>
<td>0.615</td>
<td>397</td>
<td>0.539</td>
</tr>
</tbody>
</table>
### Number of Usual Car Pool Trips by Day of the Week

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean Diff</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday - Tuesday</td>
<td>0.0000</td>
<td>0.2559</td>
<td>0.000</td>
<td>397</td>
<td>1.000</td>
</tr>
<tr>
<td>Monday - Wednesday</td>
<td>-0.0025</td>
<td>0.1664</td>
<td>-0.301</td>
<td>397</td>
<td>0.763</td>
</tr>
<tr>
<td>Monday - Thursday</td>
<td>0.0000</td>
<td>0.2459</td>
<td>0.000</td>
<td>397</td>
<td>1.000</td>
</tr>
<tr>
<td>Monday - Friday</td>
<td>0.0101</td>
<td>0.1584</td>
<td>1.266</td>
<td>397</td>
<td>0.206</td>
</tr>
<tr>
<td>Tuesday - Wednesday</td>
<td>-0.0025</td>
<td>0.2703</td>
<td>-0.185</td>
<td>397</td>
<td>0.853</td>
</tr>
<tr>
<td>Tuesday - Thursday</td>
<td>0.0000</td>
<td>0.1420</td>
<td>0.000</td>
<td>397</td>
<td>1.000</td>
</tr>
<tr>
<td>Tuesday - Friday</td>
<td>0.0101</td>
<td>0.2352</td>
<td>0.853</td>
<td>397</td>
<td>0.394</td>
</tr>
<tr>
<td>Wednesday - Thursday</td>
<td>0.0025</td>
<td>0.2509</td>
<td>0.200</td>
<td>397</td>
<td>0.842</td>
</tr>
<tr>
<td>Wednesday - Friday</td>
<td>0.0126</td>
<td>0.1322</td>
<td>1.896</td>
<td>397</td>
<td>0.059</td>
</tr>
<tr>
<td>Thursday - Friday</td>
<td>0.0101</td>
<td>0.2352</td>
<td>0.853</td>
<td>397</td>
<td>0.394</td>
</tr>
</tbody>
</table>

### Number of Usual Bike Trips by Day of the Week

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean Diff</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday - Tuesday</td>
<td>0.0678</td>
<td>0.4098</td>
<td>3.303</td>
<td>397</td>
<td>0.001</td>
</tr>
<tr>
<td>Monday - Wednesday</td>
<td>0.0025</td>
<td>0.2188</td>
<td>0.229</td>
<td>397</td>
<td>0.819</td>
</tr>
<tr>
<td>Monday - Thursday</td>
<td>0.0519</td>
<td>0.3916</td>
<td>2.646</td>
<td>397</td>
<td>0.008</td>
</tr>
<tr>
<td>Monday - Friday</td>
<td>0.0704</td>
<td>0.3689</td>
<td>3.804</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Tuesday - Wednesday</td>
<td>-0.0653</td>
<td>0.4312</td>
<td>-3.023</td>
<td>397</td>
<td>0.003</td>
</tr>
<tr>
<td>Tuesday - Thursday</td>
<td>-0.0159</td>
<td>0.2354</td>
<td>-1.348</td>
<td>397</td>
<td>0.179</td>
</tr>
<tr>
<td>Tuesday - Friday</td>
<td>0.0025</td>
<td>0.4214</td>
<td>0.119</td>
<td>397</td>
<td>0.905</td>
</tr>
<tr>
<td>Wednesday - Thursday</td>
<td>0.0494</td>
<td>0.4198</td>
<td>2.348</td>
<td>397</td>
<td>0.019</td>
</tr>
<tr>
<td>Wednesday - Friday</td>
<td>0.0678</td>
<td>0.3141</td>
<td>4.309</td>
<td>397</td>
<td>0.000</td>
</tr>
<tr>
<td>Thursday - Friday</td>
<td>0.0184</td>
<td>0.4133</td>
<td>0.889</td>
<td>397</td>
<td>0.375</td>
</tr>
</tbody>
</table>

### Number of Usual Walking Trips by Day of the Week

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean Diff</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday - Tuesday</td>
<td>-0.0126</td>
<td>0.2581</td>
<td>-0.971</td>
<td>397</td>
<td>0.332</td>
</tr>
<tr>
<td>Monday - Wednesday</td>
<td>0.0025</td>
<td>0.1664</td>
<td>0.301</td>
<td>397</td>
<td>0.763</td>
</tr>
<tr>
<td>Monday - Thursday</td>
<td>0.0013</td>
<td>0.2714</td>
<td>0.092</td>
<td>397</td>
<td>0.926</td>
</tr>
<tr>
<td>Monday - Friday</td>
<td>0.0000</td>
<td>0.1878</td>
<td>0.000</td>
<td>397</td>
<td>1.000</td>
</tr>
<tr>
<td>Tuesday - Wednesday</td>
<td>0.0151</td>
<td>0.2211</td>
<td>1.360</td>
<td>397</td>
<td>0.175</td>
</tr>
<tr>
<td>Tuesday - Thursday</td>
<td>0.0138</td>
<td>0.2080</td>
<td>1.326</td>
<td>397</td>
<td>0.186</td>
</tr>
<tr>
<td>Tuesday - Friday</td>
<td>0.0126</td>
<td>0.2377</td>
<td>1.054</td>
<td>397</td>
<td>0.292</td>
</tr>
<tr>
<td>Wednesday - Thursday</td>
<td>-0.0013</td>
<td>0.2367</td>
<td>-0.106</td>
<td>397</td>
<td>0.916</td>
</tr>
<tr>
<td>Wednesday - Friday</td>
<td>-0.0025</td>
<td>0.1809</td>
<td>-0.277</td>
<td>397</td>
<td>0.782</td>
</tr>
<tr>
<td>Thursday - Friday</td>
<td>-0.0013</td>
<td>0.2620</td>
<td>-0.096</td>
<td>397</td>
<td>0.924</td>
</tr>
</tbody>
</table>

### Number of Usual Bus Trips by Day of the Week

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean Diff</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday - Tuesday</td>
<td>0.0038</td>
<td>0.2023</td>
<td>0.372</td>
<td>397</td>
<td>0.710</td>
</tr>
<tr>
<td>Monday - Wednesday</td>
<td>-0.0013</td>
<td>0.0561</td>
<td>-0.447</td>
<td>397</td>
<td>0.655</td>
</tr>
<tr>
<td>Monday - Thursday</td>
<td>0.0025</td>
<td>0.2069</td>
<td>0.242</td>
<td>397</td>
<td>0.809</td>
</tr>
<tr>
<td>Monday - Friday</td>
<td>0.0138</td>
<td>0.1601</td>
<td>1.722</td>
<td>397</td>
<td>0.086</td>
</tr>
<tr>
<td>Tuesday - Wednesday</td>
<td>-0.0050</td>
<td>0.1943</td>
<td>-0.516</td>
<td>397</td>
<td>0.606</td>
</tr>
<tr>
<td>Tuesday - Thursday</td>
<td>-0.0013</td>
<td>0.1150</td>
<td>-0.218</td>
<td>397</td>
<td>0.828</td>
</tr>
<tr>
<td>Tuesday - Friday</td>
<td>0.0101</td>
<td>0.1807</td>
<td>1.110</td>
<td>397</td>
<td>0.268</td>
</tr>
<tr>
<td>Wednesday - Thursday</td>
<td>0.0038</td>
<td>0.1991</td>
<td>0.378</td>
<td>397</td>
<td>0.706</td>
</tr>
<tr>
<td>Wednesday - Friday</td>
<td>0.0151</td>
<td>0.1732</td>
<td>1.736</td>
<td>397</td>
<td>0.083</td>
</tr>
<tr>
<td>Thursday - Friday</td>
<td>0.0113</td>
<td>0.1858</td>
<td>1.214</td>
<td>397</td>
<td>0.225</td>
</tr>
</tbody>
</table>
Appendix F - Additional Maps
Base Map Used for Geocoding

Mileage Contours
1 mile
3 miles
10 miles

Border for Maps Displayed in Text

Base Map for Geocoding

Jay Harland
May 1, 2001
Base Map
Courtesy of Missoula County
Pared-down Registrar’s Database Map

Legend

Student Body Distribution
- 1 - 2
- 3 - 5
- 6 - 11
- 12 - 63

Main Roads

Major Rivers

Streams

Mileage Contours
- 1 mile
- 3 miles

Jay Harland
March 12, 2001
Base Maps Courtesy
Missoula County
Respondent’s Campus Designation
(as designated in the registrar’s database)
Credits Respondent was Registered for at the Time of the Survey

Legend

Fall 2000 Semester Enrollment
- 1 - 8 credits
- 9 - 16 credits
- 17 - 21 credits

Main Roads
Major Rivers
Streams

Mileage Contours
- 1 mile
- 3 miles

Jay Harland
March 12, 2001
Base Maps Courtesy
Missoula County
Parking Permit Holders versus No Parking Permit

Legend

Parking Permits
- No Permit
- Permit Holders

Main Roads

Major Rivers

Streams

Mileage Contours
- 1 mile
- 3 miles

Jay Hartland
March 22, 2001
Base Maps Courtesy Missoula County
Respondent's that Reported having a Job within a Mile of Campus

Legend

Job Location
- No Job or Job not within a Mile
- Job < Mile from Campus
- Both

Main Roads

Major Rivers

Streams

Mileage Contours

\(\n\) 1 mile
\(\n\) 3 miles

Jay Harland
March 12, 2001
Base Maps Courtesy Missoula County
Respondents with a Job in Addition to School

Legend

Respondents with Jobs
- No Job
- Job

Main Roads

Major Rivers

Streams

Mileage Contours

1 mile

3 miles

Jay Harland
March 12, 2001
Base Maps Courtesy
Missoula County
## Appendix G - ANOVA with no Distance Restriction

### Demographic Group Analysis of Variance

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Demographic Group</th>
<th>N</th>
<th>Mean</th>
<th>s</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alone</td>
<td>65</td>
<td>2.078</td>
<td>3.832</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adults Only</td>
<td>244</td>
<td>2.310</td>
<td>4.428</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Children Only</td>
<td>17</td>
<td>4.041</td>
<td>8.180</td>
<td>20.618</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Live with Adults and Children</td>
<td>45</td>
<td>13.651</td>
<td>19.959</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Parents</td>
<td>34</td>
<td>6.090</td>
<td>6.228</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>405</td>
<td>3.923</td>
<td>8.740</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>Alone</td>
<td>61</td>
<td>$1.080</td>
<td>$0.731</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adults Only</td>
<td>211</td>
<td>$0.775</td>
<td>$0.556</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Children Only</td>
<td>13</td>
<td>$0.559</td>
<td>$0.295</td>
<td>11.333</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Live with Adults and Children</td>
<td>40</td>
<td>$0.482</td>
<td>$0.238</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Parents</td>
<td>17</td>
<td>$0.296</td>
<td>$0.222</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>342</td>
<td>$0.763</td>
<td>$0.579</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Residence Price to Household Residence Size Ratio</td>
<td>Alone</td>
<td>61</td>
<td>$1.033</td>
<td>$0.751</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adults Only</td>
<td>210</td>
<td>$0.768</td>
<td>$0.789</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Children Only</td>
<td>14</td>
<td>$0.798</td>
<td>$0.499</td>
<td>4.338</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>Live with Adults and Children</td>
<td>41</td>
<td>$0.725</td>
<td>$0.400</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Live with Parents</td>
<td>21</td>
<td>$0.258</td>
<td>$0.990</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>347</td>
<td>$0.780</td>
<td>$0.766</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Bonferroni Post-Hoc Mean Difference Tests

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Distance</th>
<th>Household Residence Price to Household Residence Size Ratio (dollars/sqft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live Alone v Adults Only</td>
<td>-0.2326</td>
<td>$0.305</td>
</tr>
<tr>
<td>Live Alone v Live with Children Only</td>
<td>-1.9636</td>
<td>$0.521</td>
</tr>
<tr>
<td>Live Alone v Live with Adults and Children</td>
<td>-11.5735*</td>
<td>$0.598*</td>
</tr>
<tr>
<td>Live Alone v Live with Parents/Relatives</td>
<td>-4.0124</td>
<td>$0.784*</td>
</tr>
<tr>
<td>Live with Adults Only v Live with Children Only</td>
<td>-1.7310</td>
<td>$0.217</td>
</tr>
<tr>
<td>Live with Adults Only v Live with Adults and Children</td>
<td>-11.3409*</td>
<td>$0.294</td>
</tr>
<tr>
<td>Live with Adults Only v Live with Parents/Relatives</td>
<td>-3.7798</td>
<td>$0.479</td>
</tr>
<tr>
<td>Live with Children Only v Live with Adults and Children</td>
<td>-9.6099*</td>
<td>$0.077</td>
</tr>
<tr>
<td>Live with Children Only v Live with Parents/Relatives</td>
<td>-2.0488</td>
<td>$0.263</td>
</tr>
<tr>
<td>Live with Adults and Children v Live with Parents/Rel.</td>
<td>7.5611*</td>
<td>$0.186</td>
</tr>
</tbody>
</table>

*Bold Indicates Significant Mean Difference at alpha = .05*

*Indicates Significant Mean Difference at alpha = .001*
Appendix H - Essay on Logistic Regression Statistics

The Hosmer - Lemeshow statistic is a model fit statistic. Model fit involves two things. The calculation of the summary measures comparing the distance between the observed and the expected outcome, and a thorough examination of the individual components of these measures. The test breaks the expected model outputs and the observed values into deciles. The statistic is then calculated using a Pearson chi-square statistic for the two by ten contingency table. The statistic has been shown to follow a chi-square distribution with eight degrees of freedom. Thus for the Hosmer-Lemeshow test, the p-value for a perfectly fit model is 1.00 and 0.00 for a model which does not fit at all (Hosmer and Lemeshow 1989).

The Wald statistic is calculated by dividing the parameter estimate by the standard error of the parameter estimate. Using an alpha equal .05 significance level, a Wald of 3.0 is generally considered to be significant (Hosmer and Lemeshow 1989).

The logistic regression operates on the assumption that by estimating the parameters for the model using a likelihood function the maximum values for that function will correctly estimate the parameters. Mathematically, it is easier to use the log of the likelihood function. Three log likelihood statistics are given for the two main
models. The $B_0$ Log Likelihood is the valuation of the maximum likelihood function for the most likely constant. The -2 Log Likelihood statistic is just -2 times the log likelihood calculated for the maximum log likelihood function using the estimated parameters for the constant and the independent variable(s). For the purposes of model comprehension, it is sufficient to understand that if the -2 Log Likelihood decreases the model improves (Hosmer and Lemeshow 1989). Finally the model chi square is the difference between the -2 Log Likelihood and the $B_0$ Log Likelihood.

The statistic $B_0$ is the estimated parameter for the model. In linear regression, interpretation of the model parameter is quite straightforward. The parameter is simply the slope of the regression line. The larger parameter the more powerful the variable across all models. This logic does not hold for logistic regression, because the scale of the variable in the logit can affect the size of the parameter. In logistic regression, interpretation of the parameter is more complex. The first matter of concern is the nature of the independent variable. The independent variable for the basic model is continuous. Continuous variables can present practical problems for model interpretation, mainly because a meaningful change in the independent variable must be defined. Distance is linear in the logit, but the meaning of a one mile distance change in the heart of Missoula and a one mile change past Reserve Street is difficult to formulate. For this reason, interpretation of the model coefficients will focus strictly on the direction of the coefficient.
The next statistic is the G2 statistic. The G2 statistic is not shown in the tables, but is necessary to explain the significance statistic and is given by the following equation:

\[ -2 \ln \left( \frac{\text{likelihood without the variable}}{\text{likelihood with the variable}} \right) \]

The last statistic given in the logistic regression section is the significance of the variable in the model. This significance is the calculated p-value for the G2 statistic following a chi-square distribution with n-1 degrees of freedom. Thus, the significance test for logistic regression is similar in concept to the significance reported in linear regression, it is an assessment of the degree to which that variable contributes to the model (Hosmer and Lemeshow 1989).


