Internet Based Travel Surveys

By

Dr. Michael D. Anderson
Department of Civil and Environmental Engineering
University of Alabama in Huntsville
Huntsville, Alabama

Prepared by

UTCA

University Transportation Center for Alabama
The University of Alabama, The University of Alabama at Birmingham, and
The University of Alabama in Huntsville

UTCA Report Number 05302
September 2008

UTCA Theme: Management and Safety of Transportation Systems

Internet Based Travel Surveys

By

Dr. Michael D. Anderson
Department of Civil and Environmental Engineering
University of Alabama in Huntsville
Huntsville, Alabama

Prepared by

UTCA

University Transportation Center for Alabama
The University of Alabama, The University of Alabama at Birmingham, and
The University of Alabama in Huntsville

UTCA Report Number 05302
September 2008

Technical Report Documentation Page

1.Report No FHWA/CA/OR-	2.Government Accession No.	3.Recipient Catalog No.

4.Title and Subtitle Internet Based Travel Surveys	**5.Report Date** May 2008; published September 2008
7.Authors Michael Anderson	**8. Performing Organization Report No.** UTCA Report 05302
9.Performing Organization Name and Address Civil and Environmental Engineering Department University of Alabama in Huntsville Huntsville, AL 35899	**10.Work Unit No.** **11.Contract or Grant No.**
12.Sponsoring Agency Name and Address University Transportation Center for Alabama 207 Shelby Hall; P. O. Box 870205 Tuscaloosa, AL 35487-0205	**13.Type of Report and Period Covered** Final Report **14.Sponsoring Agency Code**
15.Supplementary Notes	

16.Abstract

.

The project tests the applicability of using Internet-based travel survey tools to collect information regarding transportation activities of individuals within the urban cities in Alabama. These travel data are vital as inputs to develop and improve travel modeling within the state. The travel information collected will be the basis for updating parameters in urban planning models that predict future roadway traffic volumes. Therefore, improving the accuracy of the urban planning models essentially improves the ability to forecast future traffic volumes and allows transportation professionals in Alabama to make better decisions regarding investment in new roadway facilities and infrastructure. This report documents the development of the Internet based survey and presents the results of a case study performed for the Calhoun County Metropolitan Planning Organization.

17.Key Words Travel Survey, Internet, Modeling Travel	18.Distribution Statement		
19.Security Class (of this report)	**20.Security Class (of this page)**	**21.No. of pages** 44	**22.Price**

Contents

List of Tables

List of Figures

Executive Summary

The purpose of this project was to determine if Internet travel surveys could effectively collect usable data for the development of trip generation models. In order to test this assumption, an Internet questionnaire was designed and made available to the public. The creation of the online survey guaranteed real time data collection and thus real time data analysis. A case study of the Internet travel survey was performed for the Calhoun County Metropolitan Planning Organization and a validation of the users and results was performed to determine if the survey respondents were similar to the entire study area population.

After conducting the case study, the factors that influenced the total daily trips made by the people were income, number of vehicles owned and the household size. Also, there was dependence between income and the number of vehicles owned, between the household size and the number of vehicles owned; and there is independence between the household size and income. Further, it can also be concluded that the Internet survey was effective in sampling for a particular category of households but it is not good for representing the entire population.

However, it is theorized that better results and a higher response rate could be obtained through a combination of surveys techniques. Essentially, the Internet-based survey could be used as a low cost means to collect data from a specific segment of the population, while targeted surveying techniques could be used to collect data from those who have been shown to not respond to this type of survey.

Section 1.0 Introduction

As the world entered the 21st century, society's necessity to use more accurate techniques to be able to predict future situations increased. In the transportation industry, forecasting and predicting are essential abilities directly affecting planning processes and projects. All transportation projects must be planned to satisfy both present needs and future demands. To accomplish this goal, current data related to the project's environment need to be collected and future data need to be forecasted. Hence, it is in the data collection aspect of a transportation project where travel surveys take center stage. More precisely, development of a traffic or transport model has relied on travel surveys since the mid 1940s as the method used to collect travel data, thus becoming a fundamental part of the transportation planning process (Alder and Rimmer, 2002). Also, as more research is directed toward the improvement of questionnaires, travel surveys become more direct, cost effective, dynamic and employed (Alder and Rimmer, 2002). However, the basic travel survey methodology has not experienced any drastic improvements and past techniques are still used.

The reliance on outdated methods to collect travel survey data limits their use and effectiveness. The incorporation of technology can drastically improve the process. Technological tools are numerous and could be used to enhance travel questionnaires. One major technological tool is the Internet, which has the ability to reach almost all people and if applied as a survey medium, could potentially improve survey management, convenience and capacity (Alder and Rimmer, 2002). Other technologies, such as geographic coding, could also be incorporated into the survey's design in order to make the questionnaire easier for respondents to answer and less complicated for interviewers to analyze. Consequently, with the intention of investigating the usefulness and promising application of these technologies, the University Transportation Center of Alabama (UTCA), the Alabama Department of Transportation (ALDOT) and the University of Alabama in Huntsville (UAH) have joined efforts to develop a project to test if Internet surveys can be used to collect travel data, ultimately upgrading the current trip generation models and friction factors used in the State. The household demographics and travel data needed for the trip generation models include the average number of trips per household per day. Also, data collected from the survey can be used for air quality monitoring, activity based modeling and traffic delay studies.

However, the scope of this report includes only the implementation of Internet features in a travel survey and the verification of the results with respect to data quality. Issues such as response rate, quality of responses and characteristics of respondents are mainly studied in order to investigate the efficacy of the Internet survey.

Report Objective

The purpose of this report is to determine if Internet travel surveys can effectively collect usable data for the development of trip generation models. In order to verify this assumption, an Internet questionnaire was designed and made available to the public. The creation of the online survey guarantees real time data collection and thus real time data analysis. With the aim of verifying the survey's effectiveness, the data collected was tested and organized. Hence, possible errors can be recognized and a proper assessment of the data's usefulness can be established.

Research Outline

For the appropriate development of this report the subsequent steps were closely studied in the research process.

Literature Review

A detailed literature review about various transportation concepts is needed to provide a general background for travel surveys. To have some familiarity with transportation issues, transport models and trip generation models are briefly introduced. Also, knowledge of existing travel surveys and the common methodology used are essential in the development of this report. Therefore, an evaluation of the existing trend regarding Internet applications is necessary to identify similarities and advantages of using the Internet in travel data collection.

Questionnaire and Web Page Building

With the purpose of reaching the report's objective, a real time collection of data is required. To perform this task, an on-line questionnaire was built. The construction of the survey involves issues such as design, format, organization, objectives, reduction of sources of error, and technical aspects. A more detailed explanation of the questionnaire building is included in Chapter Three. In the same way, in order to make the Internet questionnaire available to the public, a web page and a database need to be created. The special requirements and characteristics of the web page need to be carefully developed to present a pleasant format to the survey user. Details about the web page and the database are also found in chapter three and four.

Case Study and Analysis

A case study for the Calhoun County Metropolitan Planning Organization is presented with all pertinent information about the planning, pretests, advertising, observation and project supervising. After collecting a sample of household trip data, the inputs were examined with respect to their usefulness on developing a trip generation model. The validation process is a key step of the report since it will establish if the online travel survey was successful or if the data contain numerous errors and/or bias. In order to see how representative of the populations the

results are, analysis of the response distribution with respect to demographics, economics and spatial characteristics is performed. Using geographic information systems (GIS), maps, figures and tables of the results are developed and the data are studied to further detail.

Conclusions and Recommendations

The last part of the report deals with the analysis and critique of the results obtained from the previous sections. Hence, the online survey and the results are examined and discussed at a higher extent. Recommendations are also offered.

Document Organization

The report has been divided into six chapters. In the first chapter a general preamble to the report and its topic is given. Moreover, an introduction about the project is presented to give a general idea about the contents and scope of the report. The second chapter explains the theory behind transport and trip generation models as well as analysis methods. In addition, travel surveys are discussed and existing trends in Internet applications are analyzed. Chapter three explains the details and key points of the questionnaire, the web page and the database used for the survey. Chapter four is a User's Guide to the Internet survey. The fifth chapter deals with the case study and the analysis of its results. Hence, issues such as data retrieving, testing and validation are overviewed. Lastly, chapter five provides a review of the survey's performance. Chapter six presents conclusions and recommendations.

Section 2.0 Background

This literature review chapter for this report provides a brief description of the following topics:

- Transport Models
  - Trip generation models
- Travel Surveys (in general)
  - Statistical Obstacles
  - Advantages
  - History
- Current Methodology for Travel Surveys
- Methods for Travel Surveys
- Travel Surveys using Internet Applications
  - Modern and Future Directions

A better understanding of the present status and use of travel surveys can be achieved by analyzing these topics. Consequently, a more adequate study of the possible use, benefits, and drawbacks of Internet travel surveys can be addressed.

Transport Models

Transport issues have become a major issue for the majority of industrialized and developing countries. Increasing levels of traffic affect most major cities around the world. Thus, there is a need for traffic models that can predict real time conditions and produce possible solutions to congestion/delay problems. Ortuzar and Willumsen (1994) defines a model as a simplified representation of a part of the real world that concentrates on certain elements considered important for its analysis from a particular point of view. For transportation applications, models use mathematical approached to replicate the system of interest (Ortuzar and Willumsen, 1994). Also models and planning studies, in their majority, use data taken at a single point in time, which is known as a cross-sectional approach.

These theories lead to the basic four-stage transport model or urban transportation modeling system (UTMS) that has been used for several years and is the framework for all the transport models used. Meyer and Miller (2001) describe the used of the four-stage model by saying that it predicts the number of trips made within the zone. Also, it forecasts the time of day (peak hour, daily, etc.), the zonal origin-destination (OD) pair, the mode of travel used to make these trips, and the routes taken through the transportation network by these trips. Therefore, the final result of the model is a set of modal flows on links in a network that represents that represents the equilibrium processes that assigns the traffic demand into the networks within the transportation system (Ortuzar and Willumsen, 1994).

The model is a sequential decision process that starts by selecting the zones and networks. Next, present and future data are collected, calibrated and analyzed. Then, trip generation and distribution models are created to produce a trip matrix. Afterwards, the trips are divided according to their respective mode and then assigned to their corresponding network (Ortuzar and Willumsen, 1994).

Trip Generation Models

As stated above, the classic model is formed by a set of sub models, distribution, generation, modal split, and trip assignment. The scope of this report covers only the relationship between travel surveys and trip generation models. A trip generation model defines the total number of trips generated by households in a zone (Ortuzar and Willumsen, 1994). The trips could be home based (HB) when home is the origin or destination of the trip maker or non-home-based (NHB) when home is neither the origin nor destination of the trip maker. In the gathering of HB and NHB trips, travel surveys play a major role since they are the tools for collecting household activity data.

When a trip is the home end of a HB trip or the origin of a NHB trip, it is defined as a production trip. On the other hand, when a trip is the non-home end of a HB trip or the destination of a NHB trip, it is classified as an attraction trip (Meyer and Miller, 2001). Prediction of attraction trips requires data such as zonal employment levels and zonal floor space. On the contrary, predictors for production trips include household characteristics, income, car ownership, and distance to the central business district (CBD). Consequently, separate models are used to predict productions and attractions (Meyer and Miller, 2001). For this report the main focus is to ultimately collect data for a production model. Hence, household data are vital and a travel survey is the tool to collect these data.

Trips are also classified in the following ways: by trip purpose, by time of day, and by person type. Classifying trips by purpose and modeling each separately has proven to be more efficient (Ortuzar and Willumsen, 1994). Some categories commonly used to classify trips are trips to work, educational trips, shopping, recreational and social trips, and others. Due to their characteristics, the first two categories are said to be mandatory and the others optional (Ortuzar and Willumsen, 1994). When distinguishing trips by time of day, a differentiation is made for peak and off-peak period because the amount of trips usually varies with time of day. The last classification takes into account the socioeconomic condition of the person, such as income level, car ownership and household size. Once the trip classification is performed, trip generation is based on the relationships between trip ends and the activity characteristics of the land use attracting the trips (Meyer and Miller, 2001). Therefore, for various purposes, trip generation is expressed as the number of trips per x, where x represents the land use activity. For instance, for hospitals, the number of trips is taken per patient beds (Ortuzar and Willumsen, 1994). To obtain a complete list of vehicle trip rates refer to the Institute of Transportation Engineers' (ITE) Trip Generation Manual.

Two mathematical approached are used to analyze data and build up trip generation models. The first technique involves the use of a linear regression model and the second used the cross-

classification model. This last procedure, since the late 1960s, has quickly become the preferred method in the Federal Highway Administration. Linear regression models are relatively easy and inexpensive (Meyer and Miller, 2001). Regression models are mostly used for attraction trips and the basic idea behind the model is the formulation of trip generation as a linear function of one or more variables. On the other hand, for production trips the cross-classification method is most commonly utilized. The main idea behind cross-classification is the estimation of the number of trips as a function of specific household attributes (auto ownership, income, household size, etc.) (Ortuzar and Willumsen, 1994). The purpose of this grouping method is to establish a more homogenous data set and to calculate the average trip production rate from observed data specifically for each set. However, this requires a large amount of data, thus the number of households is a key factor in this procedure. As Meyer and Miller (2001) state, the category division of cross-classification analysis avoids the regression model's assumption of a linear, additive relationship between trip generation and its explanatory variables, as well as the pitfalls inherent in spatially aggregated models.

Travel surveys become important at this point of the trip generation process. The need for a larger quantity of data makes data collection techniques an essential part of the development of the production trip models. Consequently, one of the most important issues in the initial stage of the modeling process is the selection of the type of travel survey. A good selection could save great amounts of time and money. Using the Internet as tool to enhance travel survey capabilities could open new opportunities and advantages of trip generation modelers.

Travel Surveys

Travel Surveys constitute the main tool for traffic data collection. Using survey data, transportation models are planned and forecasted. Thereafter, the results obtained from these models are used in investment programs, urban and rural development, evaluation of existing facilities and many other applications. That is the reason why surveys have been effectively used for the past three decades by the transportation industry in the transportation planning process. The following subsections described some of the major disadvantages as well as advantages of travel surveys. Also, a short history and explanation of the different surveying methods is included.

Statistical Obstacles

Stopher and Lee-Gosselin (1997) define travel surveys as a fundamental part of the travel-behavior and travel-demand modeling process. Also, Stopher and Lee-Gosselin (1997) assure that in recent years travel surveys have achieved progress in fields such as administration of the survey, proper design of questionnaires, interviewer abilities, and results analysis. Nevertheless, matters such as the validation of the questionnaire, the choice of appropriate sampling procedures, and interpretation of the results, depend on the statistical approach taken to analyze and solve these issues. Consequently, a technically perfect survey may be useless if the statistical method used to analyze the data does not measure the accuracy of the results. The

statistical design and analysis of travel surveys have the following difficulties (U.S. Department of Transportation, 2000):

- The sampling plan becomes too complicated.
- The population totals and means often do not refer to a specific date but rather for a long period of time.
- The required parameters are mostly ratios, i.e. number of trips per person per day, both the numerator and denominator are unknown.
- There is a need for corrections between different mobility levels of responding and non-responding households.

Consequently, for the development of an effective travel survey, these issues should be taken into account and the statistical analysis procedure is a key factor for data validation. Therefore, the statistical method used for the survey's analysis is a difficult task requiring proper planning and examination.

Advantages

Travel surveys are widely used to gather current information related to household and individuals' characteristics, thus providing detailed data about demographics, socioeconomic, and trip-making properties. Similarly, Richardson et al. (2000) state that travel survey utilization allows researchers to increase the understanding of travel in relation to the choice, location, and scheduling of daily activities. Consequently, by implementing travel questionnaires, the methods and tools used to predict traffic conditions will be improved as well as the capability to establish trends in daily travel patterns. In addition, responses to current social and economical tendencies could be better addressed. Also, as new technologies and investments affect transportation systems and services, the use of current forecasted data will increase abilities to analyze changes in transportation supply, distribution and regulation. George Schoener from the United States Department of Transportation (U.S. Department of Transportation, 2000) recognizes an extensive list of trends in America's travel that have been identified by the use of existing data provided by travel surveys. Such data have beneficiated the transportation industry; the state DOT's municipal planning organization (MPO) and society in general. Some of the issues according to Schoener are as follows (U.S. Department of Transportation, 2000):

- Growth in average trip length especially for social and recreational purposes
- Mix of the household vehicle fleet and the fleet's aging, with some impacts on safety. Older drivers are driving older cars
- Impact of the baby boomer generation on travel
- Increases in women's travel. Between 1969 and 2001, the number of women of driving age grew by 47% while the number of women drivers grew by 113%
- Long-distance or "stretch" commuting. About 3.3 million Americans travel 50 miles or more one way to get to work. Of the 61.6 billion commutes to and from work each year, just less than one out of every 200 trips is a stretch commute.

History

Transportation predictions and models are usually based on household and trip data. However, there are also other types of survey data that could be used for modeling, such as land use, employment and demographic. For this report the data used for research and modeling analysis requires only households and trip data. According to the National Highway Institute (NHI), household travel surveys have been used in the United States since the 1940s. Before the implementation of travel surveys, the commonly used method for gathering data was to intercept the traveler in the course of the trip. The inefficiency of this method pushed the federal government to take the first step and formulate the Federal Highway Act of 1944 in which the federal funds were committed towards the creation of travel surveys. Afterwards, the Federal Highway Administration (FHWA) published several guidebooks to be used as references for conducting household travel surveys. Some of these guidebooks are the "Yellow Book" (1973) and the Travel Survey Manual (1996).

The period between the Federal Highway Act of 1944 and the mid-1970s was characterized by what is known as first generation travel surveys. These types of surveys are mostly in-home interviews that require large sample sizes (3,000 to 30,000) and retrospective data collection. A major problem with first generation surveys is the elevated cost. Stopher and Metcalf (1996) specify the cost for the last known home interview survey, conducted in Baltimore in 1976, to be $300 per household in 1976 dollars. After this period, travel surveys were improved and the need for large samples of data was reduced because household characteristics were divided into specific sets. Since then the home interview has been replaced with telephone and mail-back procedures. Also, travel dairies for each member of the household have been used to reduce dependence on trip recall (Stopher and Metcalf, 1996).

Nevertheless, according to Lysaker (1989) in order to design a travel survey, forecasters have to take into account not only the transportation related issues but also society's trends. The major tendencies that have been affecting travel surveying in the U.S. are the reduction of public co-operation, the need for better analytical techniques and most importantly the evolution of new technologies. This last trend directly influences the ability to improve survey efficiency by applying new tools and travel surveying such as computer assisted interviewing and geographic information system (GIS). Consequently, travel surveying is in search for revolutionary ideas and methods to mange developing complex trends.

Current Methodology for Travel Surveys

In a report prepared for the Transportation Research Board, Richardson et al. (2000), state that to collect household and daily trip information about a sample population, the current surveying techniques rely basically on the combination of two methods, telephone and mail-back. When two methods are combined, the type of survey is known as hybrid survey. In these types of surveys, households are randomly chosen, contacted by phone and mailed travel dairy to each trip-making member of the family. The use of telephones is widely accepted since at least 93% of the household nationwide have telephone service (Alder and Rimmer, 1999). Two systems

are used to contact people by phone. The first technique involves gathering samples of telephone numbers from available telephone numbers list (either published directories or lists from previous surveys). The second method, known as Random Digit Dialing (RDD) survey method, consists of collecting a sample of numbers randomly from the telephone number lists (U.S. Department of Transportation, 1996).

Afterwards, having selected a random one or two-day period, each person has to record his or her complete driving activities during that period. To retrieve the information, the interviewer either contacts the participant by telephone or the participant mails back the travel dairy. Once the data are collected, it is checked for missing or invalid entries as well as for obvious errors. Thus, if erroneous or missing inputs are identified, participants are requested to provide the necessary information to complete the survey. After the data have been reviewed and corrected, it is uploaded into computer files. Next, the data are tabulated, given geographic coordinates (geocoded) and then analyzed. An alternate procedure is to use a computer-assisted telephone interviewing system (CATI) that enables the interviewer to directly upload the information to the computer in order to be tabulated and analyzed. CATI surveys usually save time and that is why they have been used in 39% of the surveys conducted between 1989 and 1995 (Resource Systems Group, 2000). In the same way, if the CATI system is interactively linked to a geographic information system (GIS), the geocoding process is made instantaneously.

The last step is to organize and summarize the findings into reports and forms. The outputs include data dictionaries for all survey files produced, a complete description of all survey procedures, and full disclosure of survey response rates at each stage of the survey interviewing process (Richardson et al., 2000). Some of the drawbacks of the current methodology could arise if the CATI system cannot be used and the data needs to be entered after the survey team gets the responses by mail, also if the response rate for Telephone-Mail out-Mail back survey is very low (Richardson et al., 2000).

Other Methods for Travel Surveys

Depending on the type and circumstances of the travel survey required, there are several methods available to the interviewer. The more general methods are

- Personally administrated interview (face-to-face)
- Self-administrated surveys distributed by intercept methods
- Self-administrated surveys distributed to groups
- Telephone interviews
- Self-administrated mail surveys
- Combinations of methods (hybrid surveys)

These methods are not always reasonable choices for all purposes and survey conditions. For instance, it can be seen that in order to collect Household Travel Activity information the most used survey methods are, just as Richardson et al. (2000) have stated, telephone mail back

techniques. By using hybrid methods, the advantages of the selected first method are employed and a second method is utilized to overcome the drawback of the first chosen surveying method.

Personally administrated interviews are habitually used for household travel surveys in the past. But nowadays it is not feasible to conduct face-to-face interviews for household travel surveys in the United States. The reason for this is the requirement for highly trained personnel and large amounts of data entry and analysis. Thus, these surveys have become very expensive, labor intensive and time consuming. Therefore, despite advantages such as ease to collect information about the entire household and high response rates, these types of surveys are not currently used for collecting household travel activity.

Richardson et al. (2000) also introduced the idea that regardless of the type or method of survey chosen, emphasis should be increased in the concept of total design surveys. The following steps should be considered: preliminary planning; design and sample selection; development and pretest of the survey instruments; administration of the survey; coding, editing, validation, and processing of the data collected; the analysis of the weighting of survey data; presentation of survey results; and the documentation and archiving of the cleaned data. Hence, by conducting surveys under a total design structure, quality control issues are enhanced an accuracy in the results is increased.

Travel Surveys Using Internet Applications

Internet applications are currently used for several types of surveys such as health related studies, user satisfaction analysis, marketing and advertising, etc. The basic layout for the development of an Internet survey consists of the design of the web page and the survey questionnaire. The respondent's task is to complete and submit the questionnaire. A database is created to store the submitted data, which can be then analyzed. Richardson et al. (2000) state some of the advantages of using the Internet, such as, the ability to e-mail survey materials, increases use of multimedia to simplify complex questions and improved response time. Hence, some consequences are greater choice for respondents of when and where to be interviewed, use of interactive tools and graphics and use of multimedia methods that could simplify complicated survey questions and make it much easier and faster. Also, by improving the appearance and understandability of the survey, response rates could increase. Likewise, in the past years Internet usage has increased considerably and its influence on the American people is everyday bigger, thus the Internet could increase survey response rates (Alder and Rimmer, 2002).

However, Internet survey methods are still not widely used because the Internet users are not representative of the entire population; hence a wider distribution of household Internet use is needed.

According to Adler and Rimmer (2002) in their review of a Computer-Based Intelligent Travel Survey System, five sets of Internet Technologies can be used to support travel surveys. These technologies are

- Respondent-interactive geocoding – software and database that allow respondents the ability to interactively identify geographic locations of their trips
- Web-based survey templates – Software templates that support web use of a wide variety of travel surveys
- Online analytics – A set of server-based analytical tools to support simultaneous processing of survey responses, thus enabling sophisticated dynamic adaptation of interviews
- Authoring tools – a suite of tools developed to support efficient design and error-free implementation of web-based surveys
- Administration tools – database-driven approaches used to develop a full set of tools to support web-based and survey administration

NOTE: "As provided by the SBIR funding used to develop these technologies, the tools themselves and the details of their design are proprietary property of Resource Systems Group, Inc." (Alder and Rimmer, 2002).

In the same way, Resource Systems Group, Inc. is also running the project Surveycafe.com that host surveys for various purposes, including travel behavior surveys. Some of the cities that use Surveycafe.com for travel surveys are Austin, Texas and Chicago, Illinois. Other cities that have used Internet applications are summarized in Table 2-1.

Table 2-1. Surveys That Used Internet Applications, Adler and Rimmer (2002)

Place where survey was conducted	Type	Survey Instrument	Sample Size	Response rate
Las Cruses, NM	Household Travel diary survey	Internet	3000	30.6%
Central Florida	Origin/Destination Survey	Mail with Internet Options	100,960 Handed over, 66,189 mailed	12.7% (2.5% through Internet)
Bloomington, IN	Indiana University Travel Demand Survey	Emailed to selected students	5000	11.6%
Volusia County, FL	Household Travel Survey	Household travel diary web template	6313	25.0%
Chattanooga Urban Area	Household Travel Survey	Household travel diary web template	3700	27.4%

For the Las Cruces, New Mexico household travel diary survey, emailing was employed for the distribution of questionnaires. The survey also used mail back and incentive methods ($2 cash incentive) (Alder and Rimmer, 2002). Likewise, the distributions for the surveys in the Indiana (Travel Demand survey) and Central Florida (OD travel survey) Universities were sent by email and included cash rewards. Similarly for the Volusia County the incentives included free pen and raffle with 75 respondents winning annual passes to County Parks, while for Chattanooga there was $1 included with per-notice letter and $1 in each survey packet mailing. Also, the first 500 respondents received a free Ben & Jerry's Ice Cream coupon (Alder and Rimmer, 2002).

However, there are also difficulties for posting web surveys. These include errors in the source code for the questionnaire or issues dealing with the setup of the web page design. These obstacles have to be taken care and revised in order to provide the used with the appropriate web page layout that will enable him/her to accurately resolve the questionnaire and also facilitate

data retrieving. Nevertheless Internet travel surveys are becoming gradually more popular because of their low cost, simple usage, easy access, software availability faster response rates, faster intercommunication and interaction, high geographical technologies and administrative capabilities.

Modern and Future Directions

The top of the line technologies that have been recently developed and are currently used for travel surveys include Computer-Assisted Self Interviewing (CASI), Internet administration and interactive geocoding. Once again, Resource Systems Group is one of the pioneers putting into practice these revolutionary techniques. CASI methodologies include the wide use of graphics and interactive systems. These guarantee graphical user interface by means of photos and icons, which has proven to increase response rates and accuracy (Alder and Rimmer, 1999). Likewise, Internet administration deals with high levels of graphical inputs making it a great alternative to telephone surveys. The reasons for this are, besides graphical options, better response rates and cost effectiveness (Alder and Rimmer, 1999). Lastly, interactive geocoding has also several advantages such as improvements in the accuracy of location information, reduction in post processing and cost and increased geographic precision (Alder and Rimmer, 1999). Therefore, travel surveys employing these trends could effectively increase response rates as well as reduce administrative and operational time and cost.

Consequently, important benefits obtained specifically from CASI surveys are the abilities to use computer and Internet services for error checking, internal branching and, as stated above, use of pictures and graphics (Alder and Rimmer, 1999). The three types of error that can be checked using CASI surveys are logical, procedural and missing information error. Logical errors are those that are not probable or incoherent. Procedural errors are possible but not probable given the existing household info. Lastly, missing information errors are those where the respondent does not give the information (Alder and Rimmer, 1999). To facilitate the correct development of CASI travel survey software, Resource systems Group has identified the following items needed for a household-activity survey questionnaire and included them on their survey (Alder and Rimmer, 1999):

- Household structure
- Number of adults and children
- Number of licensed drivers
- Socio-economic status
- Trip information – for all trips made by all members of the household
- Trip start and end times
- Trip purposes
- Travel mode
- For auto – occupancy, operating and parking costs
- For transit – sub mode, access/egress modes, fare, breakdown of travel time (access, ride, egress), party size
- Location of trip origin/destination

The survey developed for this report included the above items, but disregards the transit specific values. Also, the survey will generate table summaries of the data inputs with the aim of alerting the respondent and making him/her double-check their answers. Other differences between the Resource Systems Group's survey and the proposed survey are that Resource Systems Group asks for the duration of the trips as well as parking info while our survey ignores test facts. There are also various format dissimilarities, but overall the main issues concerning household and trip data are found in both surveys.

Summary of Advantages and Disadvantages of Internet Surveys

The major advantages of an Internet survey include:

- Fast response and fast data retrieving
- Response time can be reduced and number of responses increases if the survey is posted in a high visited web page
- Very low operational and maintenance cost
- Interactive geocoding can help interviewers obtain geographical data very easily and rapidly
- Contacting time can also be reduced, due to fast email interaction. Hence, respondents can easily get feedback about the survey.
- To increase or decrease the sample size does not have an influence on the cost of the survey
- Graphics and multimedia tools can be used to make the survey more attractive and more user-friendly
- Internet surveys use special database and question logics that enable complex question skipping, randomization and other features not possible with paper questionnaires or most email surveys

Nevertheless, Internet surveys also have disadvantages, which consist mainly of the following:

- Internet use is not widespread enough to reach most of the population. Internet surveys may not reflect the entire population.
- Possible erroneous data, including incomplete inconsistent fields.
- Since everyone can access Internet, there is no control over who replies and individuals responding multiple times.

Consequently the applications of Internet features into travel surveys could have great impacts on the survey's capability of collecting graphical data, increase response rate, facilitate data analysis and reduce operations and maintenance cost. These characteristics could dramatically improve the survey's performance and guarantee successful results. However, the drawbacks of using the Internet need to be considered since they could also affect the survey's outcome.

3.0 System Design

The objective of this project is to determine the factors that influence the travel behavior in a particular location, including the number of trips made per person, the time at which they travel, the number of miles they travel, and other related information. This information is used to develop models for trip generation. Then the secondary objective of the project is to determine whether an Internet-based survey is suitable for collecting this type of data. The major cities of Alabama are selected for the survey. The mode of data collection used was the web questionnaire tool.

The website for the travel survey is http://travelsurvey.uah.edu. The web pages are a combination of HTML and ASP.NET formats. The website is hosted on the Microsoft Windows Server 2000. The initial intent was to configure a web server dedicated for the website; however, due to difficulties with web technology at the university, it was placed on the UAH main server.

The web page editor that has been used is the Microsoft Front Page 2000. The functionalities of the tool has made it easy to embed the desired java script code on the web pages, thus paving way for dynamic and static web pages on the web site. The website is divided into several functional modules which are as follows:

- Module 1: Information and City Selection- a static HTML page
- Module 2: Household Characteristics- HTML page with java script embedded
- Module 3: Trip Summary- ASP page with java script embedded

Module 1: Information and City Selection

The first web page of the Travel Survey web site has a detailed description of why this project has been developed. Also, there is a set of instructions provided to direct the user on how to complete questionnaire. At the bottom of the web page is a list of cities from which the user selects their location. The city selection is hyperlinked to connect to the appropriate next page.

Module 2: Household Characteristics

The Household Characteristics page is Part A of the questionnaire pages and contains household demographics, trips and vehicle questions. In the address section, the user is expected to enter street, city and zip code information. The only mandatory entry in this section is the zip code; the user must enter a 5-digit zip code without which one cannot navigate to the next web page. Next, the user is expected to select an appropriate income, which is presented in ranges using

radio buttons. Then, the age of the people living in the household are entered. Both the income and the age sections have been created according to survey standards. A user can add as many people in the household using the "add another family member feature". The final section would be choosing the type of car and the model year, here again the user can choose more than one vehicle by using the "add another car option". Java Script has been used to allow the user to add or remove vehicle and family member with functions such as Add Family Member ( ), Remove Family Member ( ), Add Vehicle ( ) and Remove Vehicle ( ) in the Charactersitcs.htm page. When all the data have been entered by the user, the submit button needs to be clicked by the user. The submit button checks if the zip code datum has been entered, if not it prompts the user to type in the zip code again. After submitting, two ASP.NET pages never seen by the user, process_savecontents.asp and process_saveaddresscontent.asp, are used to save the data to the database.

Module 3: Trip Summary

The third and the final module would be the trip summary page, in this page the user is provided with a map and on this map the user needs to select the starting and ending points for the first trip made by their vehicle by simply clicking on the map. The user needs to enter the time, the purpose and the number of people traveling in that car. If the user clicks the Submit button without choosing the start and the end locations, a message will prompt them to do a selection. In the case where the user feels the selection made is wrong, they can re-select by using the Clear button. On clicking the submit button the travel survey concludes by displaying a "Thank You" web page. The imagemap.asp page is the page that deals with this module. The details of the page such as the x and y coordinates of the map area are captured by the process.savemapcontents.asp page.

Database

The database that has been used in this project is Microsoft Access 2000, which is a part of the MS Office 2000 package. Since the project does not deal with a large number of data, MS Access has been preferred over SQL. Each city has its own database in the form of "cityname.mdb" all the databases have three tables:

- Userinfo
- Useripinfo
- Tripdetails_map

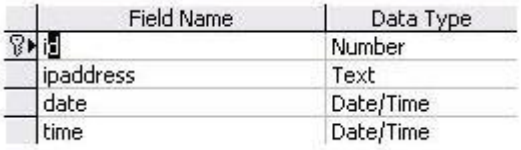
The userinfo table is the primary table of the three; it captures the details furnished by the user from the characteristics.htm page. The structure of the table is shown in Figure 3-1.

	Field Name	Data Type
	noofvehicles	Number
	street	Text
	apt	Text
	city	Text
	state	Text
	zip	Number
	gender1	Text
	gender2	Text
	age1	Text
	age2	Text
	race1	Text
	race2	Text
	gender3	Text
	gender4	Text
	age3	Text
	age4	Text
	race3	Text
	race4	Text
	gender5	Text
	gender6	Text
	age5	Text
	age6	Text
	race5	Text
	race6	Text
	income	Text
	type1	Text
	type2	Text
	make1	Text
	make2	Text
	year1	Text
	year2	Text
	nooftrips1	Text
	nooftrips2	Text
	type3	Text
	type4	Text
	make3	Text
	make4	Text
	year3	Text
	nooftrips3	Text
	nooftrips4	Text
	year4	Text
	type5	Text
	type6	Text
	make5	Text
	make6	Text
	year5	Text
	year6	Text
	nooftrips5	Text
	nooftrips6	Text
🔑	id	Number

Figure 3-1. Data definitions for userinfo table

The useripinfo table contains information about the computer used to complete the survey and the day and time the survey is completed. The ip address for the computer was collected to ensure the same computer was not being used to enter a large amount of data. In addition, the

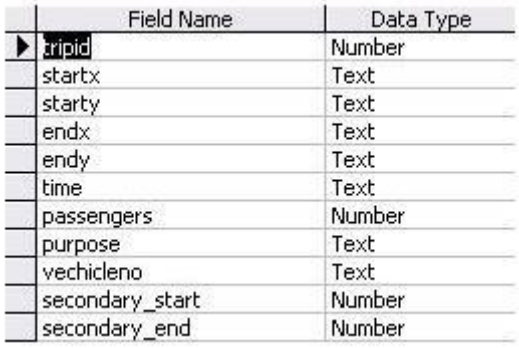
time for the entry was intended to ensure legitimacy in the data by removing suspicious entries entered at unexpected times. This information was collected by the system without knowledge of the user. The structure of the table is shown in Figure 3-2.



Field Name	Data Type
id	Number
ipaddress	Text
date	Date/Time
time	Date/Time

Figure 3-2. Data structure for useripinfo table

The tripdetails_map table collects the coordinates and first trip information from the map page. The structure of the table is shown in Figure 3-3.



Field Name	Data Type
tripid	Number
startx	Text
starty	Text
endx	Text
endy	Text
time	Text
passengers	Number
purpose	Text
vechicleno	Text
secondary_start	Number
secondary_end	Number

Figure 3-3. Data structure for tripdetails_map table

Section 4.0 User's Guide

This chapter provides a user's guide to the internet-based travel survey system. As mentioned previously, the first page displayed is the welcome/information page. This page provides the purpose for the survey and general instructions. (See Figure 4-1.) After the general instructions, the user selects the appropriate city to continue. (See Figure 4-2.)

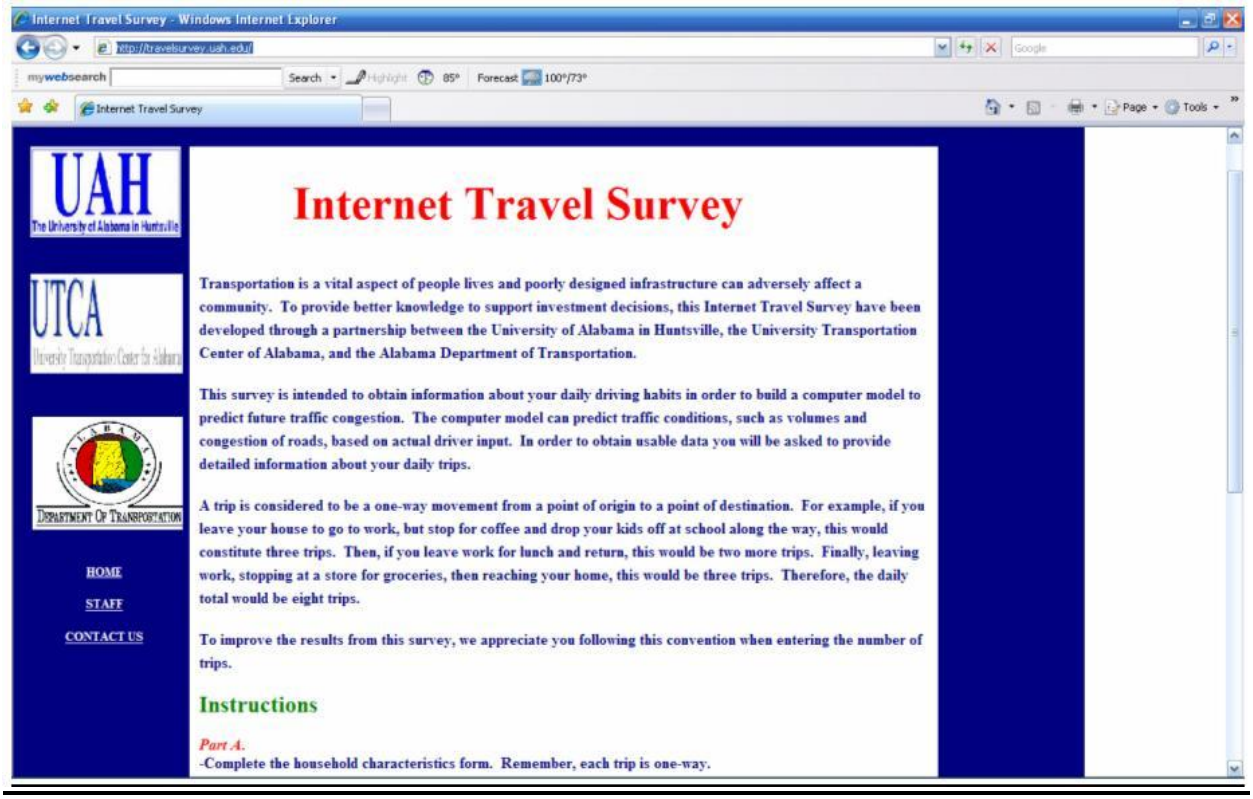


Figure 4-1. Welcome page

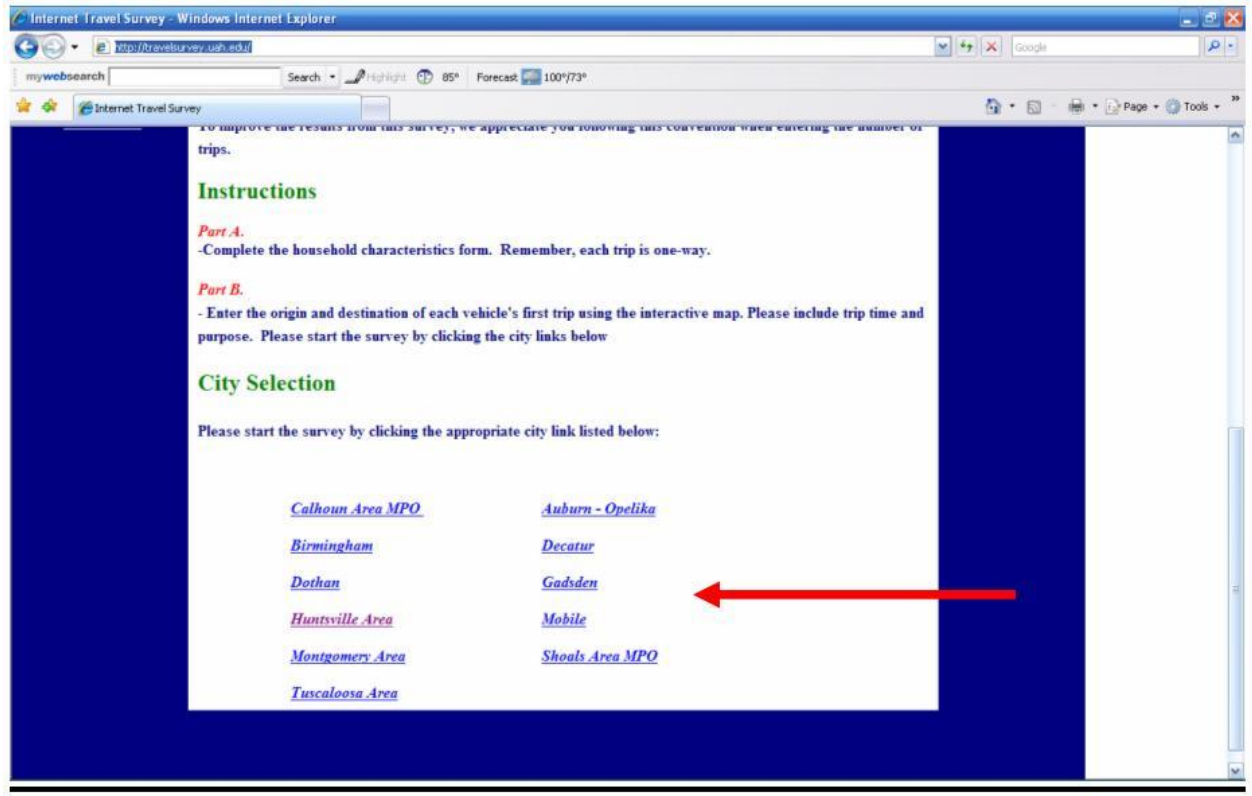


Figure 4-2. City selection

After selecting the appropriate city, the user must furnish their household details on the characteristics page. A zip code is required within the address as a means to ensure individuals are familiar with the area; all unrelated zip codes are removed from the databases. The user enters the appropriate income, identifies the people in the household and vehicles/trip information. (See Figures 4-3 and 4-4.) The user can add/remove family members and vehicles by clicking on page where indicated by arrows.

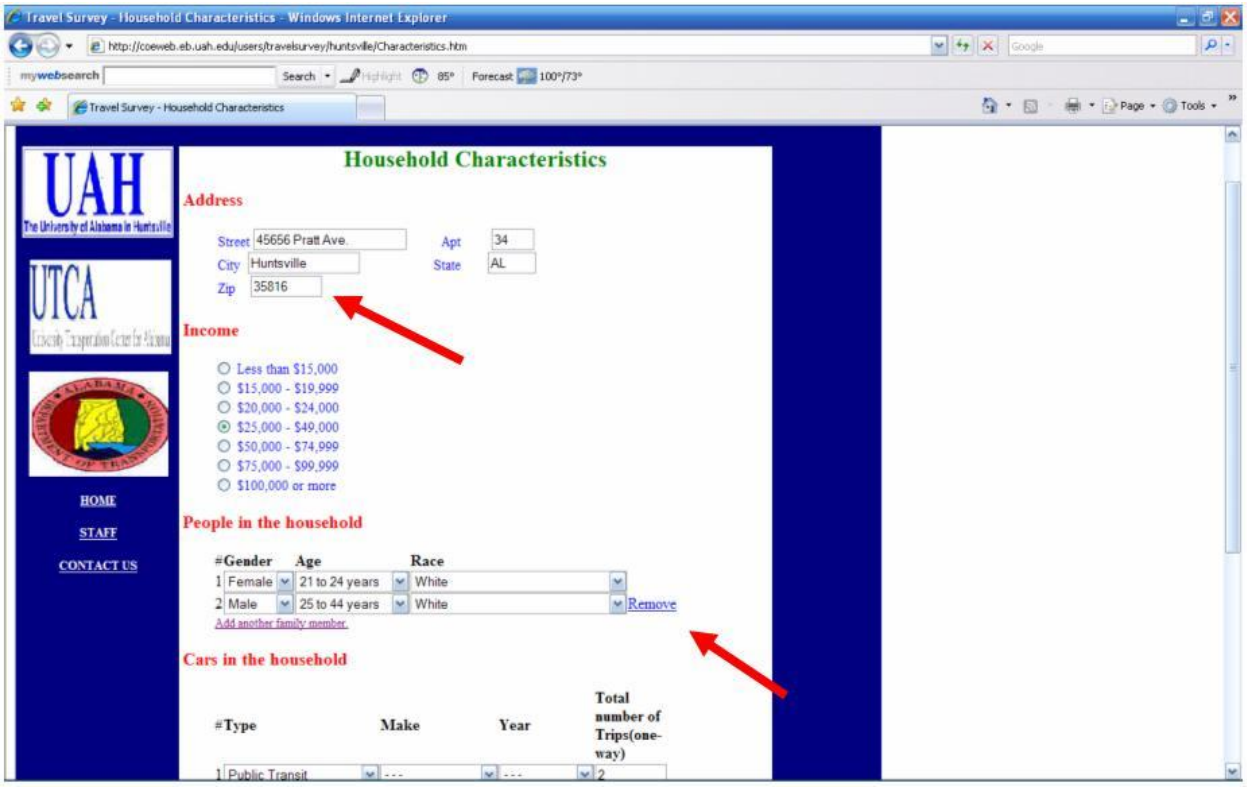


Figure 4-3. Household characteristics

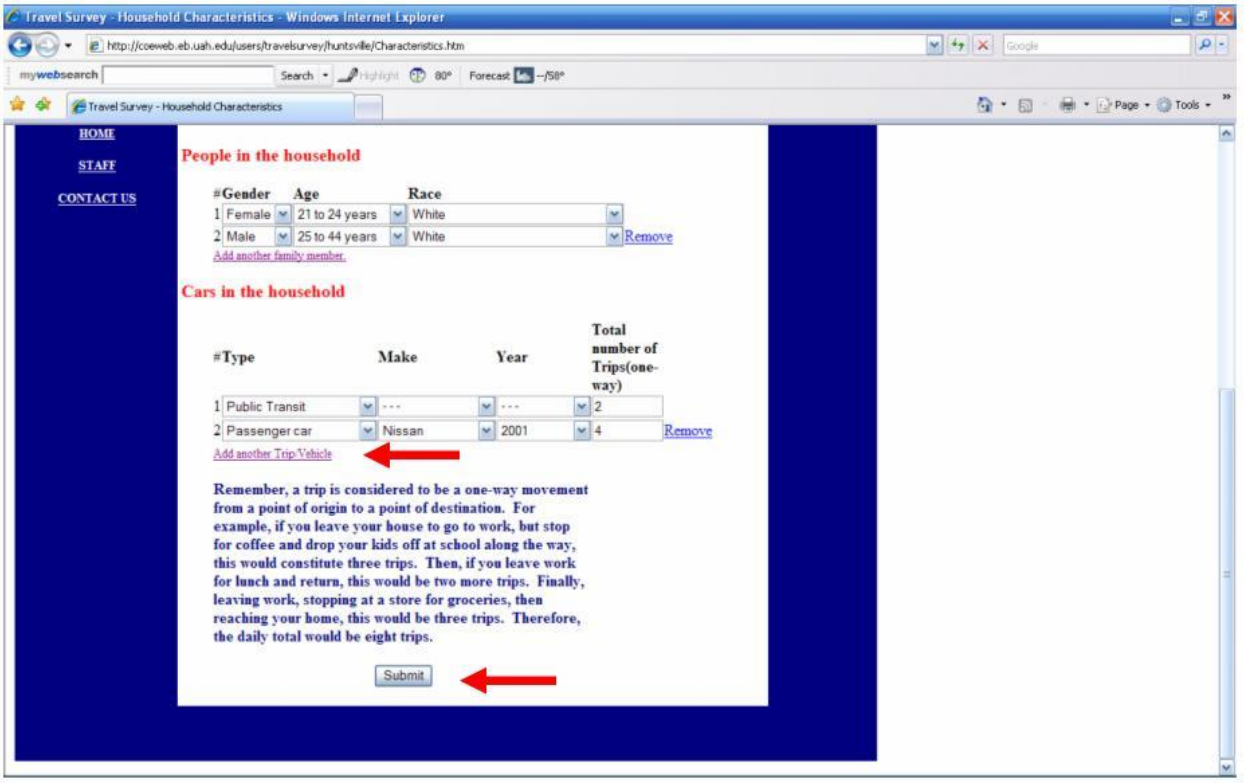


Figure 4-4. Vehicle characteristics

Once the user is finished entering the household and vehicle information, clicking the submit button will continue the survey. The next page displays a map on the city selected on the welcome page with instructions for the user to enter specific trip information. The user is expected to identify a starting and ending trip location on the map and enter travel time information, purpose of travel and the number of people traveling in the vehicle making this particular trip. (See Figure 4-5.) After clicking the submit button, the thank you page is shown, ending the survey. (See Figure 4-6.)

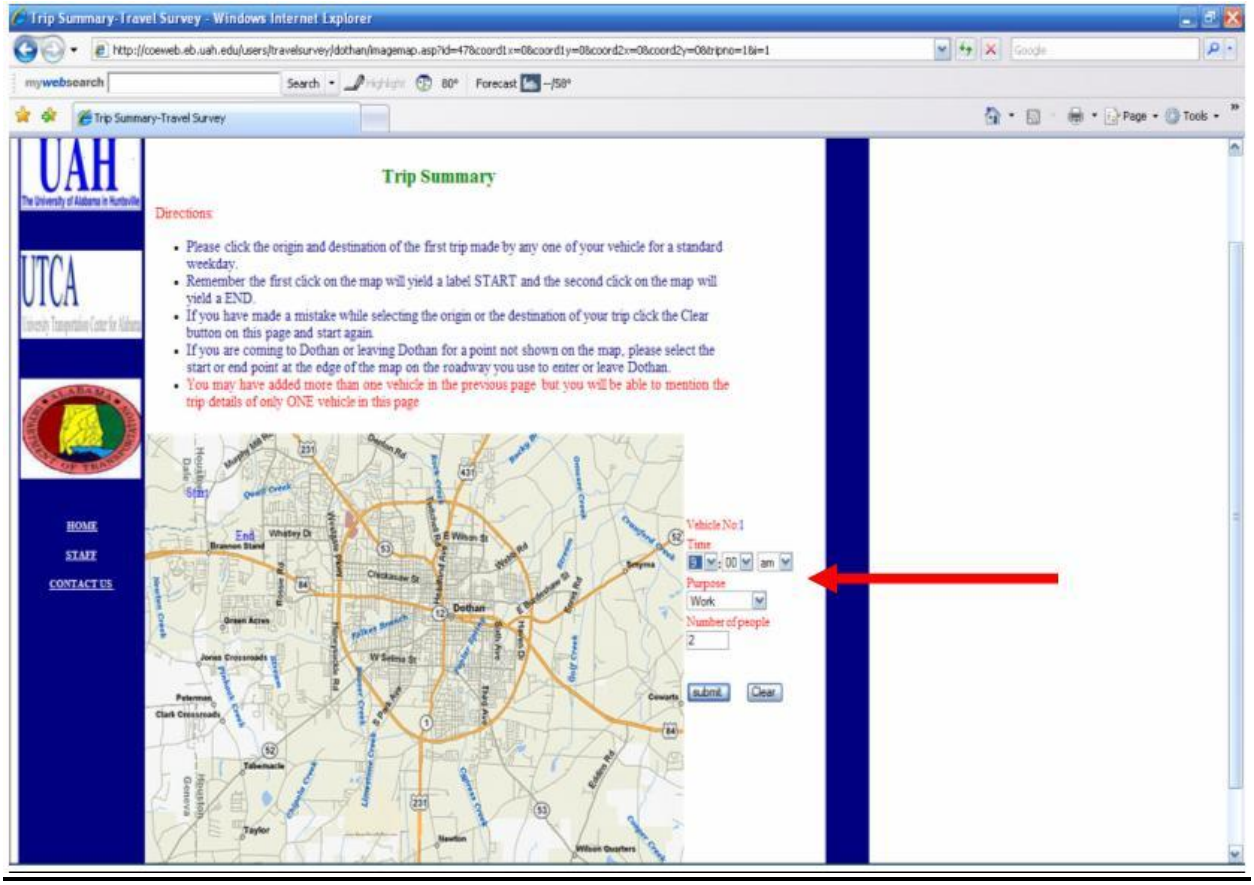


Figure 4-5. Trip characteristics

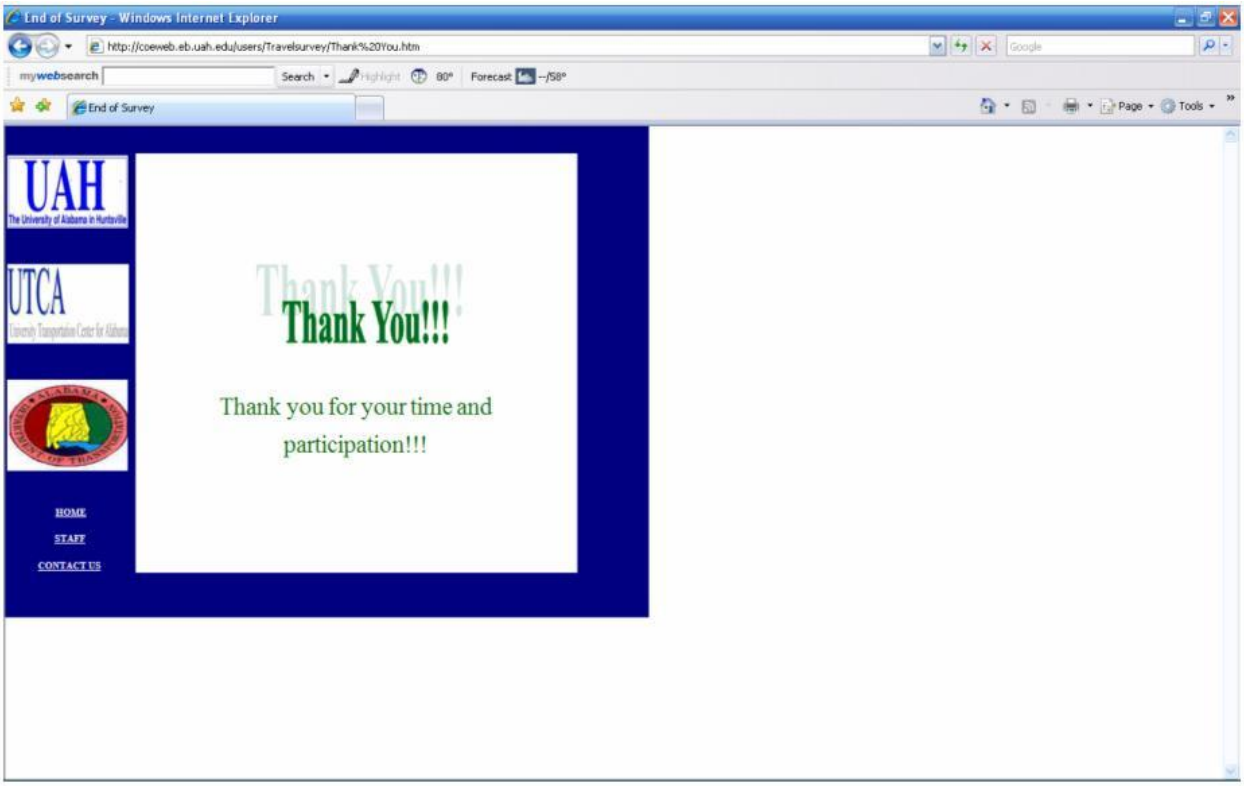


Figure 4-6. Final thank you page

Section 5.0 Case Study

The case study presented in this work is provided to examine the potential of the Internet to collect travel survey data for a small community in Alabama. Previous studies indicated that a drawback to Internet surveys is the respondents do not reflect the population, however, it is of interest to determine which populations respond to these types of surveys. The case study analysis presented is specific for the Calhoun County Area Metropolitan Planning Organization (MPO). Similar approaches can be followed any of the other cities included in the web pages. Calhoun County has a high level of automobile dependence and MPO budget limitations prohibit a traditional survey, therefore, it was hypothesized that an Internet-based approach can provide population specific data elements to enhance transportation modeling.

For Calhoun County, a total of 244 household responses were entered after eliminating the erroneous test entries using the IP address table. To determine the geographic scope of the survey, a zip code dot density plot was developed where each dot represents a respondent by the entered zip code, shown in Figure 5-1. From Figure 5-1, although not statistically validated, it was felt that the survey provided decent coverage geographically.

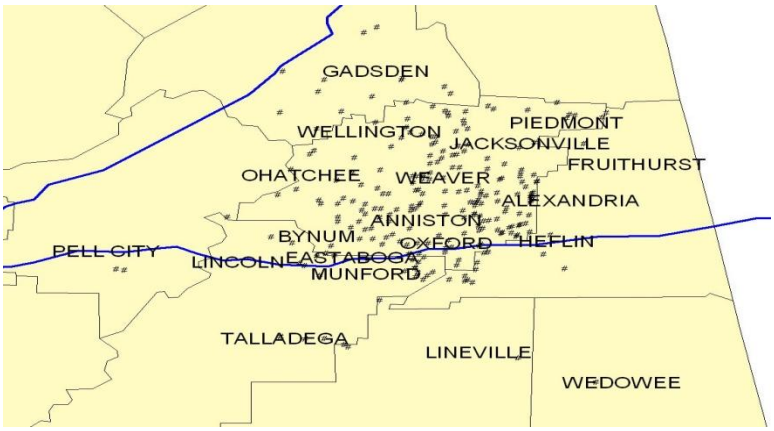
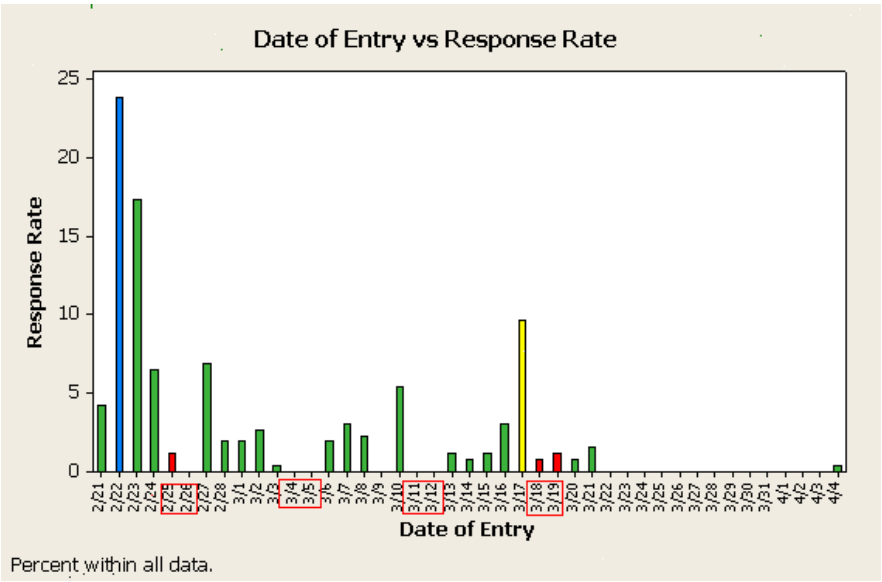


Figure 5-1. Location of survey respondents

The Internet survey was advertised to local businesses through a memorandum prepared by MPO staff and was announced in the local newspaper through a lengthy article. To check for the effectiveness of this advertising, a graph was developed showing the date and number of respondents per day. Figure 5-2 shows the number of responses per day and the timing of specific events related to the announcement and direction of people to the survey.



■ - The survey was announced by the Calhoun County MPO (local business).
■ - The survey was published in the newspaper article.
■ - Weekends

Figure 5-2. Day on which people responded to the survey

Figure 5-2 clearly indicates that the initial distribution by the MPO staff had the largest impact on responses. The newspaper article also generated a number of responses the day it was published. Conversely, weekends were not productive with respect to getting responses, which implied that a majority of the people responded from work. When the time-of-day in which the survey was completed was analyzed, it becomes obvious that most people responded to the survey from work. (See Figure 5-3.) This implies that a majority of people responding to the survey have computers and Internet facility at work. Therefore, the survey is limiting with respect to reaching people in lower income and service sector jobs.

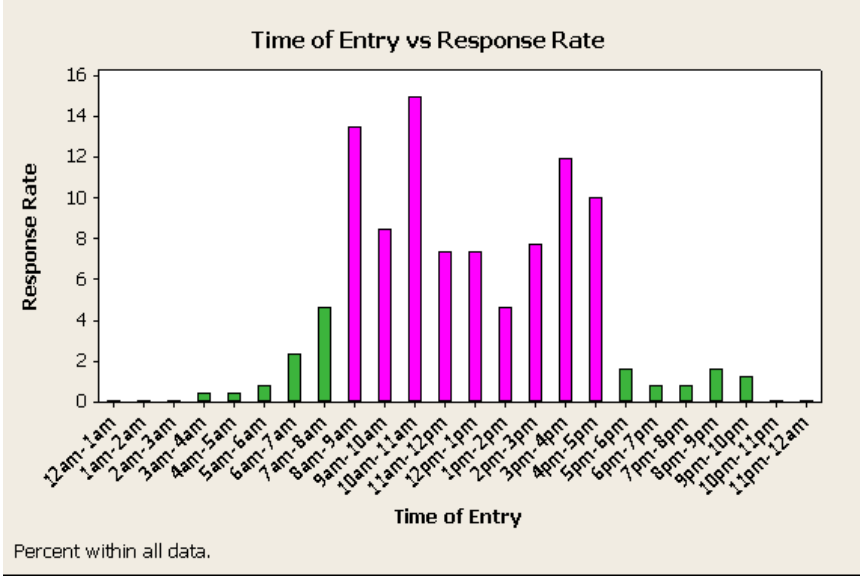
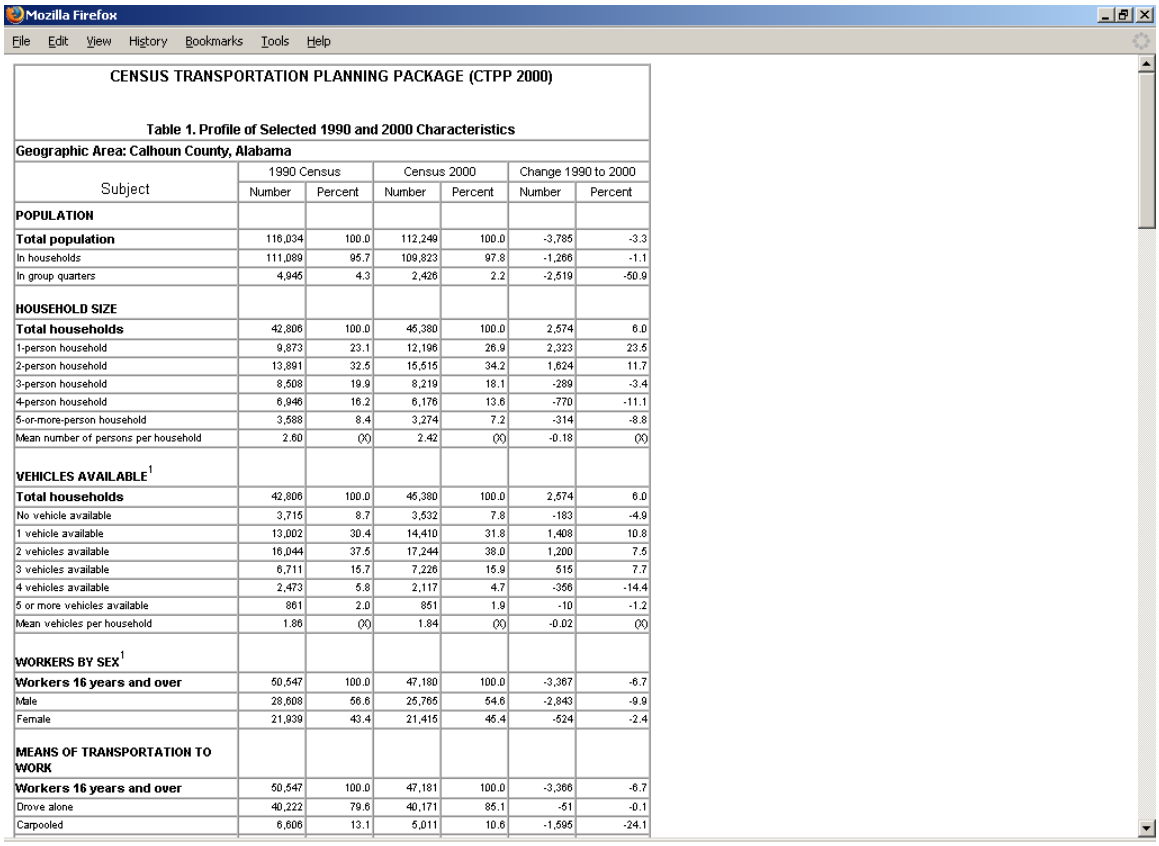


Figure 5-3. Time at which people responded to the survey

Demographic Comparison

The purpose of the demographic comparison was to establish if the Internet survey was able to reach a good cross section of the population. For this purpose, statistical tests such as the chi - square were used to test whether the population and sample have the same probability distributions. This was done to ensure that the sample gives true picture of the population. The census data used for the chi-square analysis was obtained from the Census Transportation Planning Package (CTTP), which is shown in Figure 5-4.


Mozilla Firefox

File Edit View History Bookmarks Tools Help

CENSUS TRANSPORTATION PLANNING PACKAGE (CTPP 2000)

Table 1. Profile of Selected 1990 and 2000 Characteristics

Geographic Area: Calhoun County, Alabama

Subject	1990 Census		Census 2000		Change 1990 to 2000	
	Number	Percent	Number	Percent	Number	Percent
POPULATION						
Total population	116,034	100.0	112,249	100.0	-3,785	-3.3
In households	111,089	95.7	109,823	97.8	-1,266	-1.1
In group quarters	4,945	4.3	2,426	2.2	-2,519	-50.9
HOUSEHOLD SIZE						
Total households	42,806	100.0	45,380	100.0	2,574	6.0
1-person household	9,873	23.1	12,196	26.9	2,323	23.5
2-person household	13,891	32.5	15,515	34.2	1,624	11.7
3-person household	8,508	19.9	8,219	18.1	-289	-3.4
4-person household	6,946	16.2	6,176	13.6	-770	-11.1
5-or-more-person household	3,588	8.4	3,274	7.2	-314	-8.8
Mean number of persons per household	2.60	(X)	2.42	(X)	-0.18	(X)
VEHICLES AVAILABLE[1]						
Total households	42,806	100.0	45,380	100.0	2,574	6.0
No vehicle available	3,715	8.7	3,532	7.8	-183	-4.9
1 vehicle available	13,002	30.4	14,410	31.8	1,408	10.8
2 vehicles available	16,044	37.5	17,244	38.0	1,200	7.5
3 vehicles available	6,711	15.7	7,226	15.9	515	7.7
4 vehicles available	2,473	5.8	2,117	4.7	-356	-14.4
5 or more vehicles available	861	2.0	851	1.9	-10	-1.2
Mean vehicles per household	1.86	(X)	1.84	(X)	-0.02	(X)
WORKERS BY SEX[1]						
Workers 16 years and over	50,547	100.0	47,180	100.0	-3,367	-6.7
Male	28,608	56.6	25,765	54.6	-2,843	-9.9
Female	21,939	43.4	21,415	45.4	-524	-2.4
MEANS OF TRANSPORTATION TO WORK						
Workers 16 years and over	50,547	100.0	47,181	100.0	-3,366	-6.7
Drove alone	40,222	79.6	40,171	85.1	-51	-0.1
Carpooled	6,606	13.1	5,011	10.6	-1,595	-24.1

Figure 5-4. Census data from the Census Transportation Planning Package

Number of Vehicles Owned

In order to test the distribution of the number of vehicles owned in the sample against the population, a chi–square test for testing probability distributions was performed. Initially, a frequency table showing the sample and population vehicle ownership was developed. (See Table 5-1.)

Table 5-1. Frequency Table of Number of Vehicles Owned

No. of vehicles	Population frequency	Sample frequency	Population percent	Sample percent
0	3532	0	7.8	0
1	14410	86	31.8	35.25
2	17244	100	38	40.98
3	7226	46	15.9	18.85
4	2117	10	4.7	4.1
5	851	2	1.9	0.82

However, some of the sample frequencies were less than five. This required combining to perform the test. The adjacent frequencies were pooled. (See Table 5-2.)

Table 5-2. Revised Frequency Table for Number of Vehicles Owned

Number of vehicles	Population frequency	Sample frequency	Population Percent	Sample Percent
1	14410	86	35.00	35.00
2	17244	100	41.00	41.00
3	7226	46	18.00	19.00
4 or more	2698	12	6.00	5.00
Total	41578	244	100.00	100.00

The chi-square test was performed using MINITAB Software. The p-value obtained from this test is greater than 0.05, which implies that H_0 cannot be rejected and the probability distribution for the number of vehicles owned in the sample is not significantly different from the probability distribution for the number of vehicles owned in the population. Therefore, the sample gives a sampling of the population. This can also be seen from the Figure 5-5 where the population distribution and sample distribution are very similar, with the exception of the Zero Vehicle Ownership category, which was not included in the sample.

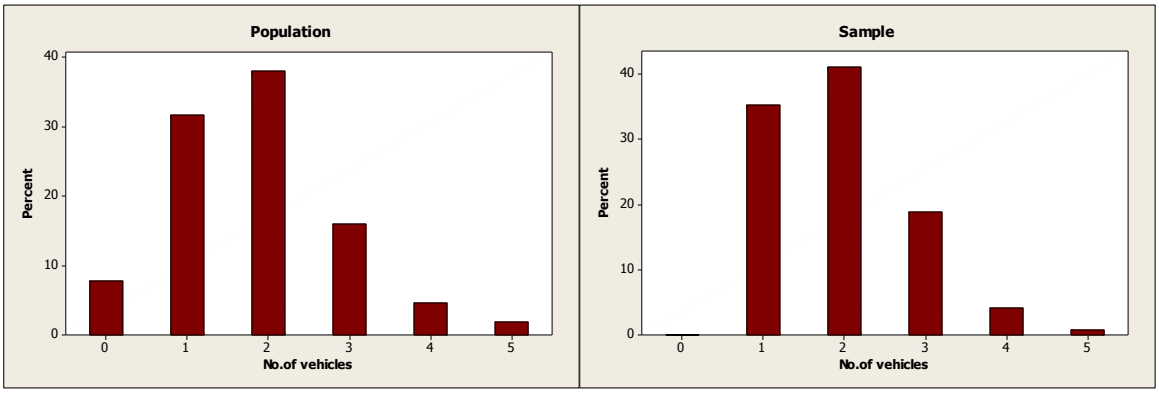


Figure 5-5. Distribution of number of vehicles owned, population, and sample

Household Size

In order to test the distribution of the household size in the sample against the population, a chi-square test for testing probability distributions is performed. Initially, a frequency table showing the sample and population household size was developed. (See Table 5-3.)

Table 5-3. Frequency Table for Household Size

Number of Persons	Population frequency	Sample frequency	Population percent	Sample percent
1	12196	96	26.9	39.6
2	15515	72	34.2	29.4
3	8219	42	18.1	17.1
4	6176	26	13.6	10.7
5	3274	8	7.2	3.2
Total	45380	244	100.0	100.0

All the sample frequencies are greater than five so the chi-square test can be applied directly. The p-value obtained from this test is less than 0.05, which tells that H_0 should be rejected and the probability distribution of the household size in the sample is not the same as the probability distribution of the household size in the population. This dissimilarity arose from the fact that one-person households were over sampled and five or more persons per household were under sampled. This can also be seen from the Figure 5-6.

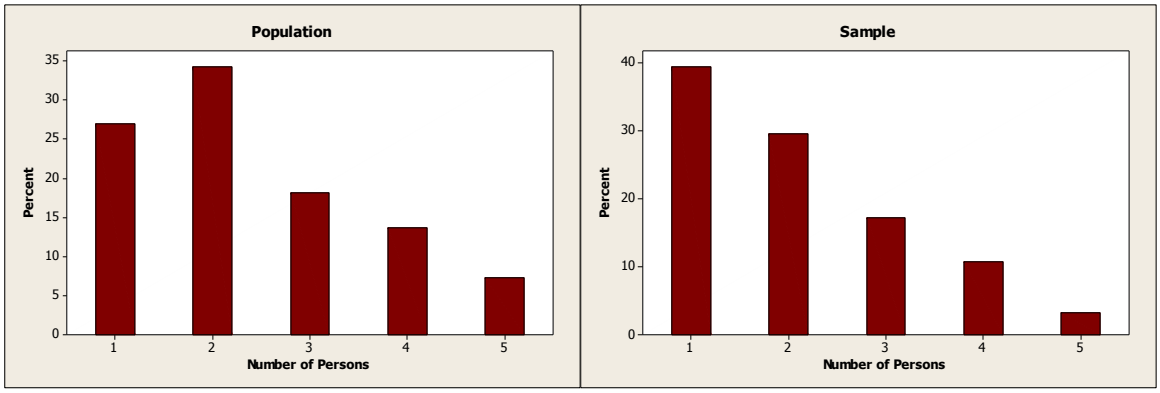


Figure 5-6. Distribution of household size, population, and sample

Secondarily, examining the dissimilar portions of the sample and population, which are single person and five or more person households, these households were excluded from the sample and again a chi-square test was performed to check the distributions. The p-value from this test was greater than 0.05 which tells that H_0 cannot be rejected and therefore, the probability distribution of the sample and population are the not significantly different for this portion of the sample.

Respondents Income

In order to test the distribution of income in the sample against the population, another chi-square test is performed. Initially, a frequency table showing the sample and population household income was developed. (See Table 5-4.)

Table 5-4. Frequency Table for Income

Income	Population frequency	Sample frequency	Population Percent	Sample Percent
Less than 15000	10773	8	23.7	3.3
15000 - 30000	10662	29	23.5	11.9
30000 - 50000	11016	80	24.3	32.8
50000 - 70000	7386	47	16.3	19.3
>70000	5543	80	12.2	32.8
Total	45380	244	100.0	100.0

The p-value obtained from this test is less than 0.05, which tells that H_0 can be rejected at 5% level of significance and it can be concluded that the probability distribution of the income in the sample is significantly different from that of the probability distribution of income in the population. This is because, the income categories of less than 15000 and 15000 - 30000 were under sampled and the income category of more than 70000 was over sampled. Therefore, the Internet survey could not capture correctly the distribution of the income in the population. This can be further supported by the histograms presented in the Figure 5-7.

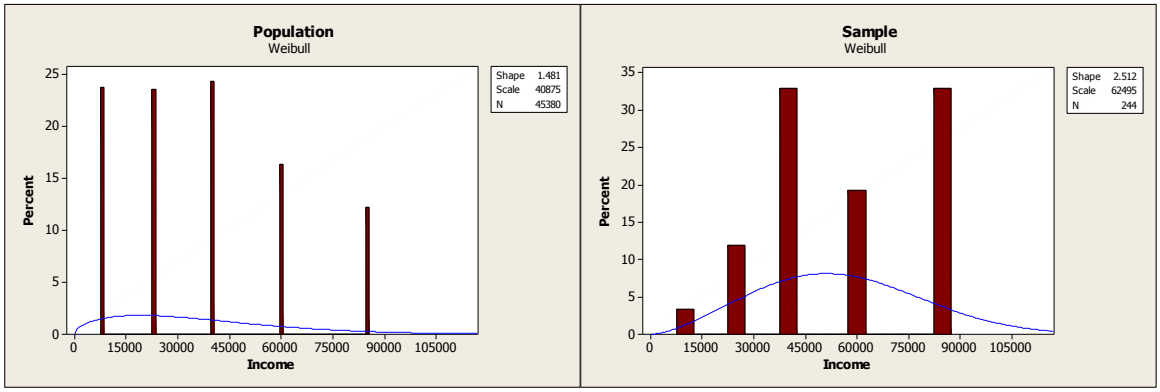


Figure 5-7. Distribution of income, population, and sample

Further, the areas where there are problems of under sampling and over sampling were recognized and excluded in the test and then the chi- square test is run again in order to evaluate whether the sample performs well within the specified classes. In this instance, the p-value for this test is greater than 0.05 we do not reject H_0 at 5% level of significance and conclude that the probability distribution of the income range 30000 – 70000, in the sample is the same as that of the probability distribution of the income range 30000 - 70000 in the population. Therefore, the

Internet sample has correctly captured the distribution of income in the population in these specified ranges.

Age of the Respondent

In order to test the distribution of 'age' in the sample against the population a chi-square test was performed. Initially, a frequency table showing the sample and population household income was developed. (See Table 5-5.)

Table 5-5. Frequency Table for Age

Age	Population Frequency	Sample Frequency	Population Percent	Sample Percent
0 - 5	6926	39	6.2	7.6
6-15	14879	50	13.3	9.7
16 - 25	16309	61	14.5	11.9
26 - 35	14432	81	12.8	16.1
36 - 45	16824	107	15	20.8
46 – 55	15970	100	14.2	19.4
56 - 65	11037	63	9.8	12.4
66 - 75	8949	9	7.9	1.7
>75	6923	2	6.3	0.4

The p-value obtained from this test was less than 0.05, which tells that H_0 can be rejected at 5% level of significance and it can be concluded that the probability distribution of the age of the respondents in the sample is significantly different from that of the probability distribution of age respondents in the population. This is further supported by the histograms presented in Figure 5-8.

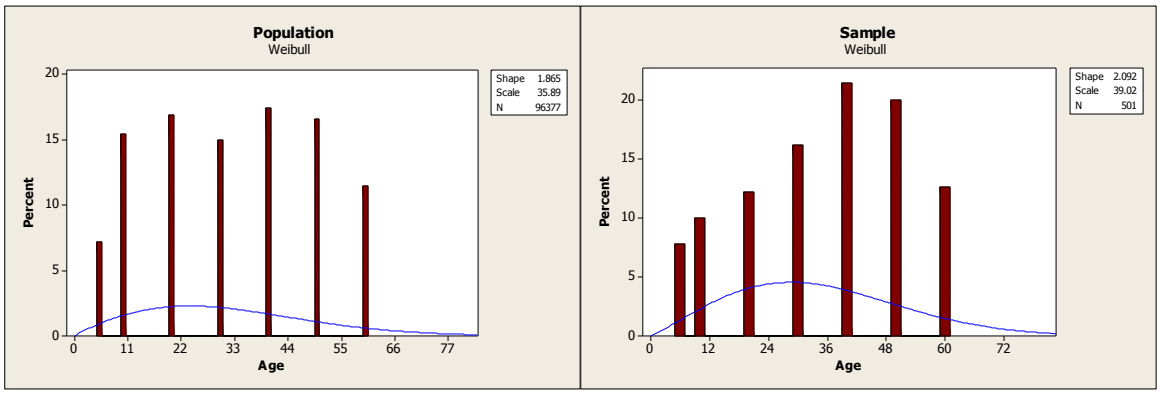


Figure 5-8. Distribution of age, population, and sample

Data Analysis

Data analysis was performed to examine how the trip rates collected varied when compared with different household characteristics. The first test of the data was to determine the factors that influence the number of trips made by the households. The variables that relate to the number of trips were recognized through common sense and based on past experience. These variables were income, number of vehicles owned and number of persons per household. Tables were generated to evaluate the average number of trips made by category. Initially, Table 5-6 compares income to trips. From Table 5-6, it is observed that the average number of trips made in one day is greater for the income category '70000 & above' and fluctuates around 6 for all other categories. This observation reveals that the income may not have a significant impact on the number of trips.

Table 5-6. Average Number of Trips by Income, Per Vehicle Owned

Income	Avg. # of trips per day	Variance	Std. Dev.	95% LCL	95% UCL	90% LCL	90% UCL
Less than 15000	5.8	13.9	3.7	4.2	7.3	4.4	7
15000 -30000	5.7	10.7	3.3	4.3	7	4.5	6.8
30000- 50000	5.6	37.3	6.1	3	8.1	3.4	7.7
50000-70000	5.9	15	3.9	4.3	7.6	4.6	7.3
70000 & above	7	27.4	5.2	4.8	9.2	5.2	8.8

Table 5-7 compares the number of people residing the household to the number of trips made per day. From Table 5-7, it can be observed that, as the household size increases the number of trips made also increases, which is not unexpected.

Table 5-7. Average Number of Trips by Persons Per Household (HH Size)

HH Size	Avg. # of trips per day	Variance	Std. Dev.	95% LCL	95% UCL	90% LCL	90% UCL
1	4.3	21.1	4.6	2.4	6.3	2.7	6
2	6.2	18.4	4.3	4.4	8	4.7	7.7
3	7	29.5	5.4	4.7	9.3	5.1	8.9
4	9.3	26.8	5.2	7.2	11.5	7.5	11.1
5	9.8	37.1	6.1	7.3	12.3	7.7	11.9
6	13.3	33.2	5.8	10.9	15.7	11.7	15.3

Table 5-8 compares the number of vehicles owned to the number of trips made per day. From Table 5-8, it is apparent that, as the number of vehicles increase, the average number of trips made also increases, which is also quite obvious.

Table 5-8. Average Number of Trips by Vehicles Owned

# of Vehicles	Avg. # of trips per Day	Variance	Std. Dev.	95% LCL	95% UCL	90% LCL	90% UCL
1	3.7	19	4.3	1.9	5.6	2.2	5.2
2	6.4	21.9	4.7	4.5	8.4	4.8	8.1
3	7.5	22.6	4.8	5.6	9.6	5.9	9.2
4	12.4	26	5.1	10.3	14.5	10.6	14.2
5	19.5	0.1	0.4	19.4	19.6	19.4	19.6

Putting this information into a format to be used for trip generation within Alabama required the development of a table showing the average number of trips expected for a household of a specific income and a known number of automobiles owned by the household. Table 5-9 provides a trip generation table developed from the results of the Internet sample that could be used in Alabama to support trip generation.

Table 5-9. Average Number of Trips by Income and Auto Ownership

	Number of Vehicles				
Income	**1**	**2**	**3**	**4**	**5**
Less than 15000	0	6.67	3	0	0
15000-30000	3.75	7.18	7.25	0	0
30000-50000	4.07	6.22	9	18	0
50000-70000	3.55	6.43	6.54	9.5	0
Above 70000	2.72	6.35	8.05	10.16	19.5

It deserves mentioning that there are many missing entries – or zero entries – in the table. These have been entered into Table 5-9 but indicate that there are no household who responded to the survey meeting these particular criteria. The lack of entries imply the specific type of household does not exist (income less than $15,000 and five vehicles) or more sampling would be needed to identify these households and collect their data as they are not likely candidates to response to an Internet-based survey.

Analyzing the time of day data for individual travel resulted in the graph, which is shown in Figure 5-9. This figure shows that the majority of the people start their first trips between 6.30 am and 8.00 am, which most likely represent work trips or travel to take children to school.

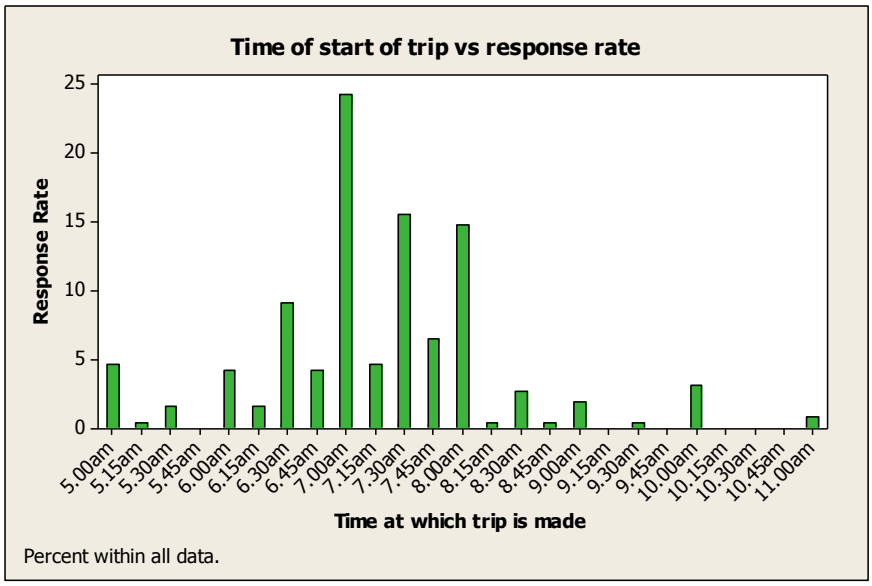


Figure 5-9. Time at which people started the first trip

From the pixel coordinates recorded through interactive geocoding, real world coordinates were developed and the distance from origin to destination was calculated using the standard distance equation. To further the analysis, only work trips were included. Figure 5-10 shows the miles to work distribution of the respondents. The distribution follows a Weibull pattern with respect to the number of miles traveled to work each day. The scale and shape parameters were calculated and using these, the mean and variance of the Weibull distribution were calculated. The data resulted in an average travel distance of 7.325 miles to work each day in Calhoun County.

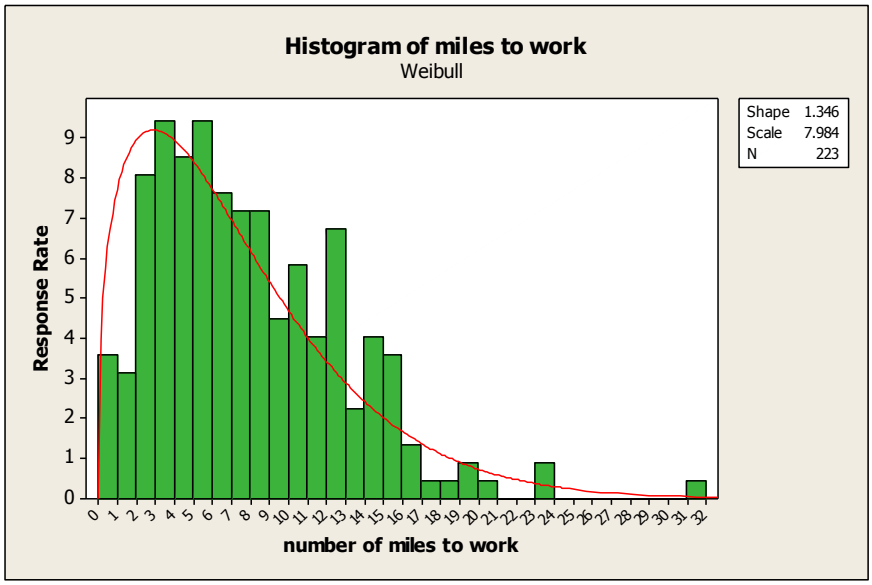


Figure 5-10. Distribution of number of miles to work

Further, the chi-square tests for independence were conducted on the factors that influence trips and to avoid any multi-co-linearity. The chi-square test for independence was performed to test the relationship between income and daily trips. The data are displayed in Table 5-10. The p-value obtained from this test is less than 0.05, which tells that H_0 can be rejected at 5% level of significance and it can be concluded that income and daily trips are not independent. Therefore, income has a significant effect on the total daily trips made.

Table 5-10. Contribution of Chi-Square to Income vs. Trips

Income	Trips 1 to 4	Trips 5 to 12	Trips 13 to 40
Less than 30000	0	0.189	0.594
30000-50000	2.563	5.013	0.002
50000-70000	0.432	1.58	0.361
Above 70000	1.167	0.925	1.077

The chi-square test for independency was performed to the relationship between the number of vehicles owned and number of trips made in a day. The data are displayed in Table 5-11. The p-value obtained from this test is less than 0.05, which tells that H_0 can be rejected at 5% level of significance and it can be concluded that the number of vehicles and the total trips are not independent. Therefore, the number of vehicles owned has a significant effect on the total trips made.

Table 5-11. Chi-Square Contributions to Number of Vehicles vs. Trips

Vehicles	Trips 1 to 4	Trips 5 to 12	Trips 13 to 40
1	8.541	7.843	5.972
2	0.018	0.035	0.482
3 or more	11.576	13.5	4.352

In order to test the dependence of the number of trips on the household size, a chi-square test for independency was performed. (See Table 5-12.) The p-value obtained from this test is less than 0.05, which signifies that H_0 can be rejected at 5% level of significance and it can be concluded that household size and the total trips are not independent. Therefore, the number of persons per household has a significant effect on the total trips made by the household.

Table 5-12. Chi-Square Contributions to Household Size vs. Daily Trips

Household Size	Trips 1 to 4	Trips 5 to 12	Trips 13 to 40
One person per household	4.441	4.432	2.575
Two person per household	0.064	0.214	0.037
Three or more persons per household	4.616	3.765	4.057

In order to test the interdependency between income and number of vehicles owned by household, a chi-square test for independency was performed. (See Table 5-13.) The p-value obtained from this test is less than 0.05, which signifies that H_0 can be rejected at 5% level of significance and it can be concluded that the income and number of vehicles are not independent. Therefore, there is a correlation between the household income and the number of vehicles owned by the household.

Table 5-13. Chi-Square Contributions for Number of Vehicles Owned vs. Income

Vehicles	Less than 30000	30000-50000	50000-70000	Above 70000
One vehicle	1.462	8.898	0.454	10.767
Two vehicles	0.13	0.797	0.06	1.752
Three or more vehicles	1.018	6.187	1.312	5.241

In order to test the interdependency between household size and number of vehicles owned, a chi-square test for independency was performed> (See Table 5-14.) The p-value obtained from this test is less than 0.05, which signifies that H_0 can be rejected at 5% level of significance and it can be concluded that the number of family members and the number of vehicles are not independent. Therefore, there is some correlation among the number of family members and the number of vehicles owned by the households. As the household size increases the number of vehicles owned increases as well.

Table 5-14. Chi-Square Contributions for Number of Vehicles Owned vs. Household Size

Vehicles	1 person household	2 person household	3 or more persons household
One vehicle	44.916	13.581	16.088
Two vehicles	13.861	16.248	0.081
Three or more vehicles	11.183	0.567	20.618

In order to test the interdependency between household size and income, a chi-square test for independence was performed (see Table 5-15). The p-value obtained from this test is greater than 0.05, which tells that H_0 cannot be rejected at 5% level of significance and it can be concluded that there is no significant dependence between number of family members and income.

Table 5-15. Chi-Square Contributions to Income vs. Household Size

Household Size	Less than 30000	30000-50000	50000-70000	Above 70000
one-person household	0.639	0.849	0.005	0.184
Two person household	0.055	0.551	0.593	1.388
Three person household	0.408	0.155	0.001	0.006
Four or more person household	0.094	0.002	0.917	0.822

In summary, the factors that influence the total daily trips made by the people within the case study locations were income, number of vehicles owned and the household size. Also, there is dependence between income and the number of vehicles owned, between the household size and the number of vehicles owned; and there is independence between the household size and income. Further, it can also be concluded that the Internet survey was effective in sampling for a particular category of households but it is not good for representing the entire population.

Section 6.0 Conclusions

Overall, the Internet-based survey achieved success in some aspects of travel data collection, but failed to provide an all-encompassing solution. From the case study, the geographic distribution of respondents was considered adequate. For the Calhoun County Area, there are significant responses from all locations within the county. Unfortunately, the respondents seem to come from mostly from small sized households, with relatively high incomes. This indicated that the respondents were mostly white-collar professionals who responded to the survey from a work computer. This had the adverse impact of leaving out a significant number of potential households in the Calhoun County Area.

However, it is theorized that better results and a higher response rate could be obtained through a combination of surveys techniques. Essentially, the Internet-based survey could be used as a low cost means to collect data from a specific segment of the population, while targeted surveying techniques could be used to collect data from those who have been shown to not respond to this type of survey.

In general, given the cost it took to develop, advertise, publicize and analyze, the Internet travel survey was successful. Thus, considering the total cost, having some minor lack of representation or inconsistencies in trip data could become insignificant and acceptable. Furthermore, the survey still available online and data are still being collected.

Section 7.0 References

Alder, T., and L. Rimmer (Resource Systems Group). *SBIR Final Project Report: Computer-Based Intelligent Travel Survey System: CASI/Internet Travel Diaries with Interactive Geocoding.* U.S. Department of Transportation, 1999.

Alder, T., and L. Rimmer (Resource Systems Group). *SBIR Phase II Final Report: Computer-Based Intelligent Travel Survey System.* U.S. Department of Transportation, 2002.

Lysaker, R. L. Data Collection Methods in the U.S. *Journal of the Market Research Society,* Vol. 31, N0. 4, 1989, pp. 477-488.

Meyer, M. D. and E. J. Miller. *Urban Transportation Planning*, 2nd ed. McGraw-Hill, New York, 2001.

Ortuzar, J.D. and L.G. Willumsen. *Modeling Transport,* 2nd ed. John Wiley & Sons, New York, 1994.

Richardson, A.J., M. Lee-Gosselin, and R. Griffiths. Travel Surveys. In *Transportation in the New Millennium*. Transportation Research Board of the National Academies, Washington, DC, 2000. Available online at http://onlinepubs.trb.org/onlinepubs/millennium/00135.pdf.

Stopher, P., and M. Lee-Gosselin, eds. *Understanding Travel Behavior in an Era of Change,* 1st ed. Pergamon Press, Oxford, UK, 1997.

Stopher, P., and H. Metcalf. Methods for Household Travel Surveys. In *NCHRP Synthesis of Highway Practices, No.236.* Transportation Research Board of the National Academies, Washington, DC, 1996.

U.S. Department of Transportation. *NIH Course No. 151034: Development and Implementation of Travel Surveys*, October 2000.