Long-Distance Trips and Mode Choice in Illinois

Joshua Auld
Transportation Research and Analysis Computing Center
Argonne National Laboratory
9700 S. Cass Ave.
Argonne, IL 60439
Tel: (630) 252-5460
Email: jauld@anl.gov

Behzad Karimi
Postdoctoral Scholar
Center For Urban Transportation Research (CUTR)
University of South Florida
4202 E Fowler Avenue, CUT100
Tampa, FL 33620
Tel: (813) 974-5324
Email: behzad@cutr.usf.edu

Zahra Pourabdollahi
RS&H, Inc.
1715 N Westshore Blvd
Tampa, FL, 33607
Tel: (813) 996-0962
Email: zahra.pourabdollahi@rsandh.com

Abolfazl (Kouros) Mohammadian
Professor
Department of Civil and Materials Engineering
University of Illinois at Chicago
842 W. Taylor Street
Chicago, IL 60607-7023
Tel: (312) 996-9840
Fax: (312) 996-2426
Email: kouros@uic.edu

Kazuya Kawamura
Associate Professor
Department of Urban Planning & Policy
University of Illinois at Chicago
412 S. Peoria Street
Suite 215 Room 234, MC 348
Chicago, Illinois 60607
Tel: (312) 413-1269
Email: kazuya@uic.edu

Word count: 5508 + 8 × 250 = 7508
Paper submitted for presentation and publication in the 95th annual meeting of TRB in Washington D.C.
ABSTRACT

Estimating demand for high speed rail systems in the U.S. is increasingly important as more resources and attention are given to the development of such systems. However, estimating the potential ridership poses unique challenges. Obtaining behavioral data to accurately predict potential customers’ market responses to a new and unfamiliar product is of paramount importance. However, there is a limited amount of travel survey data regarding potential high speed rail usage. This data is typically the main source of information for estimating travel demand for long distance trips that extend beyond urban area boundaries or planning regions for metropolitan planning organizations. Therefore, in order to conduct a study of high speed rail demand requires both understanding both how individuals currently travel and how they would travel if the new system was available. The focus of this paper is on collecting such data for a hypothetical high-speed rail system in the state of Illinois. A new mixed-mode stated-choice intercept survey was designed and implemented as both a paper and pencil instrument (PAPI) for use when pre-trip intercept was possible, and as a combination screener and follow-up internet-based Computer Aided Self Interview (CASI) survey. The stated-choice survey was used to estimate hypothetical mode choice behavior. This was paired with a Computer-Aided Telephone Interview (CATI) personal travel survey, which is used to reveal respondents current long-distance travel characteristics.

Keywords: Long-distance travel, Mode Choice, Travel Survey, Stated Choice Survey
INTRODUCTION

Long-distance travel demand models require a large amount of data on the various processes involved in long-distance trip making in order to provide useful results. This includes data regarding traveling frequency, travel purpose, party composition, travel destinations, and mode choice processes. The data needs are similar to those found in standard urban travel demand models. However, the travel covers a much wider geographic range and occurs with much lower frequency. This makes many existing data sources unusable for estimating long-distance travel due to the low-frequency of observation (1). Some attempts have been made to collect long-distance travel information on a national level, such as the 1995 American Travel Survey (ATS) and its successors in the National Household Travel Surveys (NHTS) of 2001 and 2009. However, these data sources are either fairly outdated or suffer from small sample sizes for the area of interest in this as in the ATS survey and the NHTS survey, respectively. In addition, the existing long-distance surveys contain no information on individual trip-making behavior under the addition of a new alternative mode, such as High-Speed Rail (HSR) system. This paper documents several large scale data collection efforts undertaken in order to understand long-distance travel behavior in Illinois under the assumption of a new HSR system being implemented. These surveys include several mode-specific stated-preference intercept surveys and a general personal long-distance travel demand survey. Each survey would serve a specific purpose in the development of long distance travel models, with the intercept surveys providing mode choice behavior and the travel survey providing trip frequencies, destination distributions and estimates of existing mode choice distributions.

BACKGROUND

Systematic collection of information from a sample of entities is called survey (2). Surveys are conducted to obtain better understanding of entities’ attitudes, behavior or knowledge and build quantitative descriptor of the attitudes, behavior or knowledge of the larger population from which entities are selected (3). From a long time ago, survey has been a key tool used commonly in many branches of science such as economic and social sciences to depict a better picture of how things work. In other words, surveys are fundamental tools available to a modern information-based society (3). Every survey procedure is composed of four interrelated steps including defining the precise purpose of survey, designing the appropriate questionnaire, designing suitable data collection method, and selecting the sample of target respondent (2). The duality of data collected through survey is highly affected by the way that each of the above steps is carried out.

Some of the highlights about any of the four perspectives of survey are briefly described here based on the review of literature on the history of survey research (2, 3, and 4). Understanding social problems by gaining more information about social behaviors was the main reason for early surveys. Obtaining a systematic view of people’s opinion and measuring attitudes and opinion became surveys’ purpose in journalism and marketing research. The importance of questionnaire design and wording questions was affected by increased interest in measuring subjective information such as unobserved characteristics. Since full population surveys are impractical for large population, other sampling methods such as probability samples were developed and used as standard survey sample by which other samples are judged. The data collection step was the more organized step in the survey process.
The data collection method used in early surveys was face-to-face qualitative interviews. These interviews resulted in a set of verbal notes that needed to be summarized with statistics. Development of new systematic data collection methods that were less expensive and faster affected the popularity of surveys. Mailed paper questionnaires were first used in 1960 in a formal successful test of census procedure. It was proved that mail questionnaires are much less expensive than face-to-face interviews. Using telephone service as the data collection medium was noticed by researchers with its usage spread throughout the country. Telephone surveys are much faster than mail questionnaire surveys and still less expensive than interview surveys. However, the method has some shortages which are discussed in the following. By the 1990’s face-to-face interview surveys were almost abandoned in most of scientific research fields. The invention of computers made a breakthrough in efficiencies of surveys by reducing the large amount of human resources in every step of conducting a survey and speeding the survey process. Web-based surveys are the new generation of data collection methods that are growing very fast in most of scientific research areas and becoming an alternative to traditional mail methods. The two methods have been compared in several studies and the advantages and disadvantages of each are discussed thoroughly. Along with all the mentioned developments, the development of performance guidelines has also evolved using the lessons learned from the past experiences. However, as this short history shows, the most progress in defining the fundamentals of a good survey methodology go back to the first half of 20th century.

Every individual involved in transportation planning will encounter the need of data collection at some stage of his/her professional life. Collected data about travel demand, travel behavior and transportation systems performance through surveys are indispensable input to decision making process in transportation. Taylor et al. identifies the general purposes for conducting surveys and applications of collected data in transportation. These include Transportation policy formulation, Transportation planning, Forecasting and modeling, Project evaluation, Transport system design and operation and assessment of impacts on environment, traffic network, equity and distribution, and community.

One of the main applications of collected data through travel surveys is to develop travel demand models than can forecast and identify future transportation needs and problems. The state of the art of transportation modeling has been evolving fast and the state of practice travel demand models are going to be replaced with more sophisticated models. The emerging and cutting edge practical transportation models including simulation, activity-based and tour-based models have greater demands on the data collected. However, it should be noted that, all models, whether advanced or not need accurate and complete data free of any biases and missing segments. If available data sources are not sufficient for the model development, appropriate surveys must be designed to collect required data.

Richardson et al. discussed the steps in the process of conducting a transportation survey and propose a systematic framework for travel survey which is similar to the process presented by Groves et al. and reviewed some of the existing trade-offs and challenges in survey conduction. They also classified different data collection methods employed in travel surveys. These eight classes include documentary searches, observational surveys, household self-completion surveys, telephone surveys, intercept surveys, household personal interview surveys, group surveys and in-depth surveys. This list can be extended to include the newer methods of data collection such as web-based, GPS-based and smart phone-based travel surveys.
which are used and discussed in several studies (6, 11, 12, 13, and 14). The task of selecting the
appropriate data collection method is highly effective on the efficiency of the survey and quality
of the data gathered. The selected method is usually resulted from a trade-off between the
survey’s purposes and the available resources (10). Brög and Meybur (1983) explain the
appearance of travel diary survey and its introduction to the United States. They also discussed
the influence of data collection method on the overall quality of survey results in travel surveys
(15).

Similar to other research fields, face-to-face interviews with large sample sizes were the
dominant mode of data collection in transportation planning in early years (9). With the spread of
telephone service, this medium was used as a travel survey method in North America and face-
to-face interview method was almost abandoned due to very high costs and safety concerns for
interviewers. However, it is shown that considerable number of short trips is missed in travel
diaries completed through telephone surveys (16 and 17). The amount of omitted short trips in
telephone survey is reported up to 60% by Wolf et al. 2003. This amount of missing trips put a
big question on the accuracy of the data gathered through telephone surveys. Moreover, in
another study it was found that face-to-face interview surveys also suffers from a shortfall of
about 7-12% in reporting trips (18). Another issue that has to be considered in telephone survey
is the use of “DO Not Call” registry service and call screening devices to avoid or screen out
marketing calls. In addition, the increasing number of household that have no land lines, but only
mobile phones (9).

INTERCEPT SURVEYS

Travel demand models typically have as a central component a model to determine modal
splits for overall travel demand. Such models often take the form of statistically estimated
models, such as the multinomial logit model, which are estimated from data on observed mode
choice typically found in a personal travel survey. Data on observed trips are referred to as
revealed preference data, as the survey respondents preferences are revealed from the actual
choices they make in real-world situations. Revealed preference data are typically considered to
be the preferred data when estimating choice models. Unfortunately, such revealed preference
data is not obtainable in situations involving a hypothetical alternative, such as the high-speed
rail system in this study. In such cases, it is necessary to instead ask respondents what they
would do IF the hypothetical choice was actually available. In this case the data collected are
referred to as stated-preference observations. Stated-preference data are often considered less
reliable as their collection requires individuals to think about what they "would" do under the
hypothetical situation and these responses can differ significantly from what the respondent
would actually do. However, many of the shortcomings of stated-preference data collection can
be mitigated with proper design. Therefore, the team decided to develop a series of intercept
surveys in order to determine how travelers would behave with the addition of the new HSR
mode.

The team designed the stated-preference surveys to capture mode choice behavior with an
available HSR system. In order to provide data with as realistic responses as possible, the
surveys were administered as intercept survey. This means that the individuals make their stated
preference decisions based on the experiences of the trip that they are engaged in when
intercepted. This enhances the reliability of the stated-preference data discussed in the literature (19).

The ridership team designed a total of four separate intercept surveys for the study. A separate survey instrument and sampling plan was developed for each of the primary long distance travel modes. These include automobile, airplane, intercity bus and intercity rail. The four intercept surveys were implemented either as a face-to-face interview survey as in the case of the air and rail modes, or as internet-based surveys for the auto and bus modes. The motivating factor for using different survey administration modes was the amount of time during which respondents could be contacted prior to travel. For instance, at the airport individuals tend to arrive long enough before the scheduled departure time so that a team of surveyors could reasonably conduct the required intercept interview. The same situation held to an extent for the rail survey. However, for the bus survey the team observed during pilot testing that individuals tend to arrive at bus stops much closer to the departure time. In addition, the bus stops at which the team was able to survey generally did not have seating or lounge areas in which to conduct the interview, further impeding the administration of a face-to-face survey. The auto mode survey was even less amenable to the interview survey, as auto-drivers were intercepted at either highway rest areas, service stations or highway oases as discussed in the following section. For this reason an alternative administration mode was developed which involved prescreening the intercepted respondents for eligibility and then distributing a post-card with survey completion information for a web-administered intercept survey. The contents of the questionnaire do not significantly differ between the face-to-face interview and the web-based self interview so only the face-to-face interviews are discussed in the following sections.

3.1. Stated Preference Intercept Survey Design

Both the paper-and-pencil instrument (PAPI) and the internet based survey had four main sections:
1. General questions about the trip the respondent was taking that day,
2. A section asking the respondent to consider alternative modes of transportation that could be used to take the trip the respondent was making,
3. Choice exercises asking the respondent to state his/her preference from a variety of hypothetical mode, cost, and schedule scenarios, and
4. Demographic questions.

Since all interviews were done from hubs in either Chicago or Champaign-Urbana, the team tailored the questionnaires to the mode (air, train, bus) and city of origin, as the city of origin and destinations possible from those cities defined the response categories. Therefore, for each city, there are three questionnaire versions for the PAPI version of the survey. The internet version was tailored automatically based on respondents location. The remainder of the paper will focus solely on the internet-based survey design, as it followed the same design principles as the PAPI design, but allowed for a much higher degree of flexibility in attribute levels and more accurate estimates of such factors as access/egress travel time through the connection with the Google Maps API.

The survey began, after appropriate screening of respondents to ensure that a valid baseline trip was being intercepted, by querying the respondent about the details of the
intercepted trip. This would include such characteristics as the trip purpose, duration, origin, destination, travel mode, travel party, etc. Details of the basic trip questionnaire are shown in Figure 1. In the case of the internet intercept surveys, the respondents answer the questions after the completion of a trip, at their convenience, by entering a supplied code given by the survey workers at the point of initial contact.

**FIGURE 1.** Details of currently intercepted trip

After the details of the current trip were entered, a series of hypothetical questions regarding potential access/egress trips for various alternative modes were asked. In these questions, the person is asked to consider if they were making the current trip by all of the possible shared modes, i.e. air, bus, rail. The respondent indicates which access and egress stations in the origin/destination cities would have been selected, as well as what local mode would have been used to get to and from those stations. The responses for this section are then used to initialize the access / egress travel times for the respondents to consider in the stated response experiments through the use of the Google Maps API route finding routines based on the home location, selected station location and travel mode (auto or transit). The intention here is to provide more realistic conditions under which the individual would make the choices for long-distance mode. This process is demonstrated in Figure 2 for the online intercept survey. The routines had to be simplified for the PAPI version of the survey as the access/egress travel times cannot be calculated, but were rather estimated by the respondents for the selected station. This questionnaire shown in Figure 2 was repeated six times per respondent, once for each access and egress selection for each of the three shared modes.
FIGURE 2. Access / Egress Choices For Alternative Modes

The core of the state choice intercept survey consists of a set of experimental runs for the respondent to consider alternative travel modes for making the current trip under varying conditions. Each respondent sees a set of eight stated choice experiment runs with the attribute levels for the existing modes (auto, air, bus, rail) and the hypothetical mode (hsr) randomly varied according to a design matrix. The attributes are all based on the current observed trip. The choice experiments start by displaying to the respondent the mode characteristics screen (or card). This screen gives the respondent context for each mode and provides characteristics that are not randomly varied, in order to capture mode-specific effects such as comfort, convenience, etc. which are difficult to survey directly. For this reason, a picture of a typical example of each mode, as well as general operating characteristics like service frequency, pre-departure arrival time at station, and a description of the seating and entertainment options. The idea is to get the respondent to think realistically about each mode and factor the information provided into the later decisions.

In stated preference exercises, variables of interest for the potential choice options are randomly varied for each individual across multiple choice situations to determine individual sensitivity to the variables for use in a choice modeling context. In the case of the long-distance travel mode choice there were ten variables of interest i.e. the travel time and cost for each of the
five modes, each of which had three levels, i.e. low, medium and high. The attribute levels of
are based on the observed existing conditions for each mode based on publicly available travel
times and posted fares which are used to set the medium attribute level for each trip, and are
varied by +/- 33% to set the low and high levels. For the HSR mode the travel times were based
on a speed of 180 miles per hour with medium fares based on 75% of the average air fare. An
example of a stated-choice run is shown in Figure 3. The display repeats some of the key
information from the modal characteristics table, i.e. picture of mode, frequency and pre-
derparture arrival time requirement. Additionally, the display shows the varied factors for each
mode, i.e. travel time and cost, as well as a fixed access/egress time which was estimated earlier
from the access/egress questionnaires. Care is taken here to emphasize the travel time, access
egress time and waiting time separately, as well as to ensure that respondents are aware that fares
are per person.

<table>
<thead>
<tr>
<th>CHOICE SITUATION 1:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please select which mode you would choose to make a trip from Chicago to Champaign, similar to the trip you recently made, in the following scenario:</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Service Frequency:</td>
</tr>
<tr>
<td>Station arrival time:</td>
</tr>
<tr>
<td>Travel Time:</td>
</tr>
<tr>
<td>Total Cost for all trip legs:</td>
</tr>
</tbody>
</table>

CLICK HERE to view mode descriptions in a new window

FIGURE 3. Example of a mode choice scenario from the bus passenger online survey

Unfortunately due the large number of combinations of factors and levels, this means that
a full-factorial design which would enable the complete determination of all variable effects
would require respondents to evaluate $3^{10}$ separate choice situations, which is infeasible.
Therefore a fractional factorial design was selected for implementation. The design matrix for
the fractional factorial design is based on the Taguchi L-27 orthogonal design array (Taguchi et
al. 2005). The 27 experimental runs are organized into 9 blocks based on increasing quality of
the new high-speed rail mode, and respondents are randomly assigned to see a run from eight of
the nine blocks. The use of the orthogonal fractional factorial array allows for the efficient
estimation of the first order effects of the factors although many of the interaction effects cannot
be separately estimated.

3.2. Survey Administration

All intercept surveys took place in the Chicago and Urbana/Champaign regions. Due to
the resource constraints, it was not possible to deploy the survey team to St. Louis and
Indianapolis urban areas to conduct data collection. The team retained the University of Illinois
Survey Research Laboratory to carry out the intercept surveys at O’Hare, Midway, and
Champaign airports and Union Station (AMTRAK riders). The Survey Research Laboratory and
the ridership team jointly conducted intercept surveys at the Megabus stops at the Chicago Union
Station, Suburban Express stop at the Woodfield Mall in Schaumburg, and the IL Terminal in
Champaign. The Survey Research Laboratory and the ridership team also jointly conducted
automobile traveler intercept surveys at several rest areas and service stations along the I-55, I-
57 and I-80/294 corridors.

For air and train passengers, the survey team approached travelers waiting in AMTRAK
terminals and departure gates at O’Hare and Midway airports, and asked them to answer
questions regarding their backgrounds and choice of modes under hypothetical situations. We
offered five dollar gift card to travelers who participated in the survey. The survey team used
paper-and-pencil instrument (PAPI) that will be described in the next section.

For bus travelers, it was impossible to conduct surveys at the bus stops since they do not
congregate at stops during the winter time. Typically, bus travelers wait in the car or inside a
nearby establishment where they can see the bus stop until the bus arrives. Therefore, the team
distributed cards with a URL address that the respondents can access over the Internet to
complete a web-based survey. We offered ten dollar gift card to the bus passengers to encourage
participation. The content of the questions for the bus passengers are essentially the same as that
for the paper-and-pencil instrument.

Intercepting automobile travelers also proved to be difficult. In addition to the reluctance
of the driers to participate in a survey that may take up to 20 minutes, many of the drivers are not
eligible to participate in the survey even though drivers were eligible if they had made a trip in
the past six months to an eligible destination. The team distributed postcards to eligible drivers
screened by the interviewers. The instruction on the postcard promised drivers a $10 gift card for
completing a Web survey within the following five days. Since the team was not successful in
obtaining a permission to conduct a survey along Interstate 65 in Indiana, the team was limited to
data collection at rest stops on I-55 and I-57 in Illinois. Interviewers made a total of 29 trips to
rest stops. Chicago interviewers made 4 trips to the Prairie View rest stop on I-57 just south of
the Chicago area. Champaign-Urbana interviewers made 19 trips to the Funks Grove rest stop
south of Bloomington on I-55 and 6 trips to the Illini Prairie rest stop just south of Champaign-
Urbana on I-57.

To develop a schedule for the survey, the team created a comprehensive database of all
known routes between Chicago and Indianapolis, St. Louis, and Champaign-Urbana, and
between Champaign-Urbana and Chicago, St. Louis, and Indianapolis for air, bus, and rail travel.
Routes from Champaign-Urbana were somewhat limited on all three modes. Willard Airport
departures flew to Chicago only. There were no Amtrak departures from Champaign-Urbana to St. Louis or directly between Champaign-Urbana and Indianapolis. There were bus departures to both Chicago and Indianapolis, but not to St. Louis; all Champaign-Urbana to St. Louis routes were Greyhound routes.

Following are the criteria the team used to identify the flights and trains that should be targeted for a survey.

For Flights out of Willard Airport:
- Direct flight from Willard to O’Hare
- Connecting flight through Willard from anywhere else, terminating at O’Hare
- Flight originating in Willard and connecting through O’Hare, destined for another city

For Flights out of ORD or MDW:
- Direct flight from Chicago to St. Louis, Indianapolis, or Willard
- Connecting flight through Chicago terminating in St. Louis, Indianapolis, or Willard
- Flight originating in Chicago and connecting through St. Louis, Indianapolis, or Willard

For all train and bus, regardless of origin city, if the sampled route did not end in the destination city, the passenger was screened with the following question: “When the train/bus arrives in STL/IND/CMI/CHI today, are you getting off the train/bus or continuing on?” If he/she is continuing on to a city not served by the HSR, the respondent is ineligible to participate in the survey. The goal in sampling passengers was to attempt the approach of passengers to be as random as possible. In crowded waiting rooms, instructions to approach “every nth passenger” are not adequate, and the procedure needed to ensure that interviewers would not bias the sample by selecting potential respondents based on who looked available or easily approachable.

Overall, the team averaged eight completed interviews per flight at O’Hare and 8.6 completes per flight at Midway. In Champaign-Urbana, the average number of completes per sampled flight at Willard was 2.7. Intercepting Amtrak train passengers proved to be a more laborious task. The average number of passengers with whom the team was able to complete interviews per sampled route was 5.7. The train average is lower than the air average because intermediate stops are made along train routes. For example, on the January 23 7:21 p.m. train to STL, only 23 of 117 ticketed passengers had a final destination of STL, according to data provided by customer service. Thus, there was a lot of screening required to find the passengers eligible for the study. In Champaign-Urbana, the average number of completes per train route was 3.8. The lower averages in Champaign-Urbana for both train and air are due partly to lower passenger volume but also to shorter passenger wait times, especially at the airport. In Chicago, passengers are much more likely to arrive at the gate well in advance of their flight and will be at the gate. The travel experience at Willard is quite different because only one flight leaves the airport at a time.

INTERCEPT SURVEY RESULTS

The intercept surveys collected basic demographic data and detail regarding the trip the respondent was making when intercepted, and also his/her choice of mode among long-distance bus, rail (AMTRAK), air, automobile, and HSR under various hypothetical combinations of cost
and travel time. The latter type of questions is called Stated Preference responses and is often used to collect mode choice data for non-existent modes. The respondents were asked to consider the trip in its totality, including the cost and time associated with accessing each mode. shows the relevant statistics related to the intercept survey efforts. A total of 1,752 respondents across all modes successfully completed the survey yielding a total of 1629 valid questionnaires and 6,318 valid stated-preference responses. Comparisons show the sample matches well with census results in terms of key demographic variables. Table 1 presents the distribution of sample across various modes. The greatest number of respondents were air travelers, however many of these were connecting passengers, which is a somewhat different market, so only the 138 direct flight passengers are considered in the following tables. The response rates, where estimable, where acceptable, with between 23% to 29% of intercepted individuals completing the follow up internet surveys.

| Table 1  Long-Distance Traveler Intercept Survey Sample Sizes |
| Sample size | Bus | Rail | Air (Connecting) | Auto |
| 292 | 434 | 138 (401) | 502 |
| Response rate | 23% | – | – | 29% |

1. Using AAPOR response rate formula. Response rates for rail and air passenger could not be calculated

Table 2 presents the detailed descriptive statistics of the long-distance travel intercept surveys.
### TABLE 2 Descriptive Statistics of the Long-Distance Traveler Intercept Survey by Mode

<table>
<thead>
<tr>
<th>Trip Purpose</th>
<th>Air Passengers</th>
<th>Rail Passengers</th>
<th>Car Passengers</th>
<th>Bus Passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business trip</td>
<td>59.2%</td>
<td>16.0%</td>
<td>3.6%</td>
<td>5.6%</td>
</tr>
<tr>
<td>Personal emergency</td>
<td>1.7%</td>
<td>1.6%</td>
<td>0.3%</td>
<td>0.8%</td>
</tr>
<tr>
<td>School/college break</td>
<td>3.3%</td>
<td>5.3%</td>
<td>0.2%</td>
<td>28.6%</td>
</tr>
<tr>
<td>Vacation/recreation</td>
<td>15.8%</td>
<td>27.4%</td>
<td>6.9%</td>
<td>11.9%</td>
</tr>
<tr>
<td>Visit friends/family/relatives</td>
<td>14.5%</td>
<td>40.4%</td>
<td>85.8%</td>
<td>49.2%</td>
</tr>
<tr>
<td>Work commute</td>
<td>1.7%</td>
<td>2.3%</td>
<td>0.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Other</td>
<td>3.9%</td>
<td>7.0%</td>
<td>3.1%</td>
<td>4.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Party Size</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>78.4%</td>
<td>73.5%</td>
<td>19.7%</td>
<td>69.4%</td>
</tr>
<tr>
<td>2+</td>
<td>21.6%</td>
<td>26.5%</td>
<td>80.3%</td>
<td>30.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>43.5%</td>
<td>53.8%</td>
<td>44.9%</td>
<td>60.7%</td>
</tr>
<tr>
<td>Male</td>
<td>56.5%</td>
<td>46.2%</td>
<td>55.1%</td>
<td>39.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Household Size</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.7%</td>
<td>23.0%</td>
<td>9.7%</td>
<td>24.4%</td>
</tr>
<tr>
<td>2</td>
<td>33.6%</td>
<td>34.0%</td>
<td>35.9%</td>
<td>22.4%</td>
</tr>
<tr>
<td>3</td>
<td>15.9%</td>
<td>17.9%</td>
<td>23.0%</td>
<td>22.0%</td>
</tr>
<tr>
<td>4</td>
<td>21.7%</td>
<td>14.4%</td>
<td>20.2%</td>
<td>19.9%</td>
</tr>
<tr>
<td>5+</td>
<td>11.1%</td>
<td>10.8%</td>
<td>11.3%</td>
<td>11.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No of motor vehicles</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5.1%</td>
<td>16.1%</td>
<td>0.8%</td>
<td>21.1%</td>
</tr>
<tr>
<td>1</td>
<td>19.6%</td>
<td>30.0%</td>
<td>17.1%</td>
<td>26.8%</td>
</tr>
<tr>
<td>2</td>
<td>37.8%</td>
<td>31.5%</td>
<td>45.1%</td>
<td>24.4%</td>
</tr>
<tr>
<td>3</td>
<td>21.5%</td>
<td>14.9%</td>
<td>24.7%</td>
<td>17.1%</td>
</tr>
<tr>
<td>4+</td>
<td>16.0%</td>
<td>7.5%</td>
<td>12.3%</td>
<td>10.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>75.9%</td>
<td>59.6%</td>
<td>87.0%</td>
<td>71.9%</td>
</tr>
<tr>
<td>Black or African American</td>
<td>9.9%</td>
<td>25.2%</td>
<td>4.1%</td>
<td>5.0%</td>
</tr>
<tr>
<td>other</td>
<td>14.2%</td>
<td>15.1%</td>
<td>8.9%</td>
<td>23.1%</td>
</tr>
</tbody>
</table>

The distributions of the stated choice responses, by intercepted mode, are shown in Table 3. The intercepted modes are shown across the rows, with the distributions of the modes the respondents chose shown in the columns. The results show that there is generally a degree of endogeneity in the selected mode, with air passengers tending to stay on the air mode (or switch to HSR), Auto travelers staying with auto mode, etc, meaning that the collected sample data is a type of choice-based sample, which needs to be accounted for on model estimation. Overall, it
appears that many respondents state that they would switch to the HSR mode, although it is very important to note that the weighted distribution does not represent expected modal shares as the average across many stated-choice responses does not necessarily correspond to what would be observed in reality. In order to more realistically gauge potential mode shares a mode choice modeling process, which combines the stated preference responses with revealed mode choice data is necessary.

### TABLE 3. Stated Mode Choice Distributions by Surveyed Mode

<table>
<thead>
<tr>
<th>Mode</th>
<th>Stated Mode Choice with HSR Present</th>
<th>Sampled Distribution</th>
<th>Actual Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Auto</td>
<td>Air</td>
<td>HSR</td>
</tr>
<tr>
<td>Observed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air</td>
<td>2%</td>
<td>40%</td>
<td>57%</td>
</tr>
<tr>
<td>Bus</td>
<td>11%</td>
<td>3%</td>
<td>30%</td>
</tr>
<tr>
<td>Auto</td>
<td>39%</td>
<td>5%</td>
<td>37%</td>
</tr>
<tr>
<td>Rail</td>
<td>4%</td>
<td>3%</td>
<td>67%</td>
</tr>
<tr>
<td>Weighted Total</td>
<td>35%</td>
<td>5%</td>
<td>39%</td>
</tr>
</tbody>
</table>

**PERSONAL TRAVEL SURVEY**

The personal survey was conducted in order to capture general long-distance travel characteristics for individuals in the study area, primarily focusing on both annual travel characteristics and specific results for up to the last three trips conducted by the individual. The survey was conducted as a computer aided telephone interview (CATI). The questionnaire consists of several components:

- **Introduction**
  - Verify correct contact information
- **Household roster and demographic information**
  - Demographics (age race, employment status, marital status, income, etc.)
  - Demographics for other household members
  - Housing information
- **Trip screening questions**
  - Number of trips in the past 12 months
  - Trip count by quarter, mode, purpose, party size
  - Commuting trips
- **Trip detail for last trips**
  - Previous work-related and two previous non-work related, if any
    - Start date, duration
    - Origin, destination, station access
    - Travel Modes
    - Purpose
The questionnaire was pilot tested in late-2011 and reviewed by the SRL questionnaire committee. The personal travel survey was conducted primarily by the ridership team. From a commercial vendor, the team purchased a random sample of information on 45,000 individuals living in the study area. A team of trained callers, mainly consisting of UIC students, placed calls to the random numbers listed in the database.

5.1. Personal Travel Survey Results

The final data set contained responses from 1,217 individuals. The overall response rate was 2.7%. Since the database contained basic socioeconomic information for each individual that the team tried to contact, it was possible to assess the non-response bias and to determine appropriate weighting factors based on census distributions. The demographic distributions are shown in Table 4, compared to the Census 2010 American Community Survey results.

<table>
<thead>
<tr>
<th>Age</th>
<th>ACS</th>
<th>Survey</th>
<th>MSA</th>
<th>ACS</th>
<th>Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 40</td>
<td>56%</td>
<td>23%</td>
<td>Champaign</td>
<td>1%</td>
<td>2%</td>
</tr>
<tr>
<td>Over 40</td>
<td>44%</td>
<td>77%</td>
<td>Chicago</td>
<td>65%</td>
<td>68%</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>Decatur</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White tot</td>
<td>70%</td>
<td>62%</td>
<td>Indianapolis</td>
<td>13%</td>
<td>11%</td>
</tr>
<tr>
<td>Black tot</td>
<td>17%</td>
<td>15%</td>
<td>Springfield</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Other tot</td>
<td>13%</td>
<td>23%</td>
<td>St. Louis</td>
<td>19%</td>
<td>17%</td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$50k</td>
<td>38%</td>
<td>39%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$50-100k</td>
<td>33%</td>
<td>36%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;$100k</td>
<td>29%</td>
<td>26%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results from the above table show that the surveyed demographic distributions generally matched somewhat closely to the distributions found in the Census ACS for the area. The primary difference between the survey sample and the ACS was the over-representation of individuals over 40 who comprised 77% of the sample. The distributions for race, income level and geographic area matched very closely to targets. The survey sample was reweighted to match the Census target distributions through an iterative proportional fitting procedure which matched the age, race, income and region distributions in the ACS as well as the joint-distributions amongst these variables.

The key information obtained in the personal survey included estimates of yearly long-distance travel frequency which in average was 3.8 round-trips per person. In addition to the general demographic and travel information, the individuals provided data regarding 1136 business and non-business trips made in the year prior to the survey, in order to understand the location and mode choices relating to long-distance trips. The trips went to areas external to the modeled region 64% of the time and went to areas directly served by the proposed HSR 20% of the time. Figure 4 shows a map of the trip origin and destination points reported in the survey. The figure shows the trip origin points clustered around the sampled MSAs while the destinations are distributed throughout the region and to external areas as shown in the inset.
CONCLUSION

This paper presents the process of design, implementation, and the results of a long distance travel survey in Illinois. The aim of the study was to collect an accurate sample representing respondents’ choices and behavior while facing new transportation alternatives. In particular, the study attempted to collect behavioral data to support planning of a high speed rail system and to accurately estimate the potential ridership of rail system in the state of Illinois. The survey collected data on both how individuals currently travel and how they would travel if the new system was available. A new mixed-mode stated-choice intercept survey was designed and implemented as both a paper and pencil instrument (PAPI) for use when pre-trip intercept was possible, and as a combination screener and follow-up internet-based Computer Aided Self Interview (CASI) survey. The stated-choice survey was used to estimate hypothetical mode
choice behavior. This was paired with a Computer-Aided Telephone Interview (CATI) personal travel survey, which is used to reveal respondents current long-distance travel characteristics. A total of 1,752 respondents across all modes successfully completed the survey yielding a total of 1629 valid questionnaires and 6,318 valid stated-preference responses.

REFERENCES


