

2012 ON-BOARD CAMPUS BUS SURVEY

FINAL REPORT

March 21, 2012



This report was produced by staff at the University of Wisconsin – Madison, Department of Transportation Services.



Photo 1: Riders board campus bus route 80. *Credit: University Communications, UW-Madison*

Staff at the following organizations provided support through informal interviews:

- *Transportation Services*
- *The Center for Urban Transportation Research (CUTR) - University of South Florida*
- *The National Center for Freight & Infrastructure Research & Education (CFIRE) - University of Wisconsin – Madison*
- *The City of Madison and Madison Metro*
- *Cambridge Systematics*

TABLE OF CONTENTS

I. INTRODUCTION	1
A. UNIVERSITY OF WISCONSIN - MADISON.....	1
B. TRANSPORTATION SERVICES.....	1
C. CAMPUS TRANSIT SYSTEM	1
D. SURVEY PURPOSE AND SCOPE.....	2
E. SURVEY OBJECTIVES	2
II. SURVEY METHODOLOGY.....	3
A. SURVEY DESIGN.....	3
B. SAMPLING PLAN	4
C. SURVEY TARGETS.....	9
D. PRE-TEST	9
E. FIELD IMPLEMENTATION.....	9
F. DATA ENTRY AND CLEANING.....	10
G. DATA WEIGHTING	11
III. RESULTS.....	12
A. RESULTS BY QUESTION.....	13
B. ADDITIONAL SUMMARY RESULTS.....	18
C. CROSS TABBED RESULTS	20
IV. CONCLUSION.....	22
A. SURVEY OBJECTIVES AND ACCOMPLISHMENTS	23
B. SURVEY LIMITATIONS.....	23
C. POTENTIAL SOURCES OF SAMPLE BIAS.....	24
D. FUTURE DATA AND ANALYSIS NEEDS	25
REFERENCES	i
APPENDIX A: On-Board Survey Instrument	ii
APPENDIX B: Survey Sample Size and Statistical Confidence	iii
APPENDIX C: Survey Crew Proposed Schedule.....	iv
APPENDIX D: Data Cleaning Process	v
APPENDIX E: Potential Sources of Sample Bias	viii
APPENDIX F: Selected 2008 Madison Metro on-board survey results.....	x
APPENDIX G: Madison Metro 2008 On-Board Survey Methodology	xi
APPENDIX H: Campus Bus Route Maps (Routes 81, 82, and 84 not included).....	xiii

TABLE OF TABLES

Table 1: Calculation of Estimated Sample Sizes for Individual Strata	8
Table 2: Response Rates by Survey Question.....	12
Table 3: Results by Occupation.....	13
Table 4: Results by Residence.....	14
Table 5: Results by Student Status.....	15
Table 6: Results by Academic Credit.....	16
Table 7: Survey Responses by Time	17
Table 8: Survey Responses by Route	18
Table 9: February 2011 Campus Bus Rides by Route	18
Table 10: Survey Responses by Day.....	19
Table 11: Results by Route/Occupation	20
Table 12: Results by Day/Occupation.....	21
Table 13: Results by Day/Residence	21
Table 14: Results by Occupation/Time of Day.....	21
Table 15: Comparison of Preliminary and Final Estimated Sample Sizes and Confidence Intervals for Individual Strata	iii
Table 16: Survey Crew Proposed Schedule.....	iv
Table 17: Error types, examples and diagnostic steps.....	vi
Table 18: Survey errors and data cleaning	vii
Table 19: Survey Results by Occupation.....	x
Table 20: Summary of Madison Metro 2008 On-Board Survey Schedule of UW Routes	xii

TABLE OF FIGURES

Figure 1: Results by Occupation.....	13
Figure 2: Results by Residence.....	14
Figure 3: Results by Student Status	15
Figure 4: Results by Academic Credit.....	16
Figure 5: Survey Responses by Time.....	17
Figure 6: Comparison of February 2011 Rides and Survey Responses by Route	18
Figure 7: Survey Responses by Day.....	19
Figure 8: Results by Route/Occupation	20
Figure 9: On-Board Survey Instrument.....	ii

EXECUTIVE SUMMARY

This is the final report of the 2012 on-board campus bus survey administered by Transportation Services (TS) at the University of Wisconsin – Madison (UW). This report includes:

- the purpose of the survey,
- a review of survey methodology,
- presentation of selected findings,
- and general conclusions.

The survey was administered on campus bus routes during four days between February 8th and February 14th, 2012. The purpose of the survey was to collect basic ridership data that was representative of all campus routes during the standard service schedule (when classes are in session). Five campus bus routes provided approximately 2.5 million rides during 2011. Service is funded by the Associated Students of Madison (ASM), Transportation Services, and the Division of University Housing.

Results are reported in a cleaned, non-weighted format to maintain consistency with the reporting of results by the 2008 Madison Metro survey. This, coupled with similarities in methodology, allows for comparison between surveys of certain results of interest, including ridership by occupation.

2,775 completed surveys were received from 8,371 total boarding passengers yielding a response rate of 33%. Regarding question one, 81.6% of question respondents indicated they were a UW student, 14.3% of respondents identified themselves as UW faculty or staff, and 4.1% of respondents identified themselves as a UW Hospital employee, UW campus visitor, or other. Four (4) surveys had multiple responses. Nine (9) surveys did not answer this question.

22.1% of question two (2) respondents indicated they lived in Eagle Heights/University Houses. 23% of respondents said they lived in University Residence Halls and 55% indicated they lived elsewhere. 74.2% of student respondents indicated their status as undergraduate while 21.8% said they were graduate students. 4% of respondents said they were special, professional, guest or other student and two (2) surveys had multiple responses while ten (10) surveys did not answer question three.

81.8% of student respondents indicated they were taking 12 or more credits while 18.2% of respondents said they were taking 11 credits or less. Two (2) surveys had multiple responses for question four and twenty (20) surveys did not answer question four.

55.5% of surveys were collected on route 80 while 26.8% of surveys were collected on route 85. 8.8% were collected on route 82, 6.6% on route 81 and 2.4% on route 84. 34.4% of surveys were collected on Wednesday, February 8th, 26.3% on Friday, February 10th, 14.3% were collected on Saturday, February 11th, and 25% on Tuesday, February 14th. Further data and analysis can be found in the Results section.

A post-survey analysis found no significant sources of sample bias and while general conclusions and study limitations with future data needs are offered, recommendations and policy conclusions are not discussed in this report.

I. INTRODUCTION

This section provides background information on the University of Wisconsin – Madison, Transportation Services, Madison Metro, and the UW campus transit system. Also discussed are the survey purpose, scope, and objectives.

A. UNIVERSITY OF WISCONSIN - MADISON

In achievement and prestige, the University of Wisconsin–Madison has long been recognized as one of America's great universities. A public, land-grant institution, UW–Madison offers a complete spectrum of liberal arts studies, professional programs and student activities. Spanning 935 acres along the southern shore of Lake Mendota, the campus is located in Madison, Wisconsin. The University has a student enrollment of 42,099, faculty and staff numbering 18,524 and nearly 375,000 living alumni. More information about the University can be found here: www.wisc.edu. The 2005 campus master plan can be accessed at the following website: www.uc.wisc.edu/masterplan/. A technical Transportation Element can be found here: <https://fpm-www3.fpm.wisc.edu/campusplanning/LinkClick.aspx?fileticket=YqDyYgzeZQg%3d&tabid=66&mid=461>.

B. TRANSPORTATION SERVICES

The mission of Transportation Services is to “provide innovative transportation solutions that serve and support the University of Wisconsin-Madison.” Our objective is to “sustain and strengthen our cutting edge transportation programs” through our strategic priorities, which include amplifying customer service, accelerating multimodal transportation initiatives, and advancing technology to stay on the cutting edge. As an auxiliary enterprise, UW Transportation Services does not receive any state funding. TS funding sources include parking revenues, student fees as designated by ASM, program revenue and fees, citation revenue and other income totaling \$17.2 million in FY 2010-11. Expenditures include capital and interest expense, salaries and fringe benefits, supplies and services, campus bus service, and the UW employee bus pass program.

Charged with overseeing transportation and parking on the University campus, TS administers numerous programs that strive to meet the needs of campus visitors, employees, and students. TS manages about 13,000 parking stalls and much of the associated infrastructure. These include annual base lot, reserved, metered, visitor and service stalls. TS also manages commuter solutions programs such as carpooling, walking, biking and an employee bus pass program. Infrastructure such as bike parking is also managed by TS.

C. CAMPUS TRANSIT SYSTEM

The main component of the campus transit system is the campus bus program. The campus bus program includes five (5) fixed routes operating on standard and recess schedules. The standard schedule typically operates when University classes are in session and recess service operates when classes are not in session. The campus bus system provides approximately 2.5 million rides per year in a fixed route, unlimited access system.

More route information, including maps, can be found by scrolling to the bottom of this webpage: <http://www.cityofmadison.com/metro/schedules/schedules.cfm>. An interactive map of the University of Wisconsin-Madison campus can be found here: <http://map.wisc.edu/>. Campus bus routes can be displayed by clicking the “show me” tab. Appendix H also contains maps of routes 80 and 85.

Campus bus service is provided unlimited to all riders. No fare is charged upon boarding and no proof of University association is required, such as the showing of a student ID. UW campus bus service is funded through a memorandum of understanding agreement between multiple stakeholders including Transportation Services, Associated Students of Madison, and the Division of University Housing. Service is provided by Madison Metro (more information <http://www.cityofmadison.com/metro/>).

Madison Metro offers door-to-door paratransit service for registered users within the Metro service area, including the UW campus. TS and ASM fund paratransit service on a per-ride basis for those who are eligible. More information about Metro’s paratransit service can be found here: <http://www.cityofmadison.com/metro/paratransit/>. UW – Madison does not directly offer a similar service to students, faculty, staff or visitors. Buses operating on campus bus routes are wheelchair accessible.

D. SURVEY PURPOSE AND SCOPE

The purpose of the 2012 on-board campus bus survey was to better understand transit trip characteristics of campus bus customers during the standard service schedule (during the academic semester). The scope of the survey was narrow in that no satisfaction, socio-economic, or origin-destination data was gathered.

Characteristics pertinent to understanding transit trip characteristics on the campus bus system included UW affiliation (i.e. student, staff, visitor), residence (i.e. residence hall, off-campus), UW student status (i.e. undergraduate, professional, etc.), and current UW student credit hours. Route, day, and time data were also collected to ensure a representative sample and improve reporting of results.

E. SURVEY OBJECTIVES

1. Study population includes all campus bus trips.
2. Survey all campus bus routes (routes 80, 81, 82, 84, & 85).
3. Achieve an approximate 30% response rate.
4. Achieve statistically significant, representative samples stratified by routes & days.

II. SURVEY METHODOLOGY

This section introduces the survey methodology and includes discussions of survey design, sampling plan, field implementation, and data entry. Some caveats and limitations are presented here, but most study limitations are discussed in the Conclusions section.

Transportation Services considered best practices among public transit surveys, conducted literature reviews, and consulted transportation and survey professionals to ensure that proper survey methodology was observed and results from the survey were reliable. Goals of survey methodology and planning were to:

- minimize sampling and nonresponse error,
- establish an appropriate sample size, and
- minimize sample and question bias.

It is important to note that while general guidelines for conducting on-board transit surveys exist, a variety of factors influence the methodology and implementation of individual projects. These factors include survey goals, budgetary constraints, transit system characteristics, ridership loads, scheduling, and even weather. Thus, it is difficult to compare methodology, implementation or results between on-board transit surveys.

A. SURVEY DESIGN

Because the average trip time on a campus bus route is relatively short, it was necessary to create a survey that would take minimal time to complete and be easy to understand. Also, trip origin-destination, trip purpose, rider socio-economic data, and satisfaction questions were outside the scope of this survey and therefore not included.

Survey design requires careful consideration of wording, meaning and accessibility to minimize confusion and promote correct responses. Careful consideration was given to survey wording including the avoidance of ambiguous phrases, uncommon acronyms, or verbose questions. The survey was printed on white, single-sided, half-sheet, card-stock with 12-point text in Ariel Rounded MT Bold font and dark blue lettering. The heavier card stock allowed for easier survey completion on a moving bus and the text size and font allowed for easier reading. Surveys were printed on scannable forms to ensure quick and reliable data entry. See Appendix A for the on-board survey instrument.

The on-board campus bus survey included five (5) questions aimed at collecting data related to transit trip characteristics. A sixth question at the beginning of the survey asked respondents to indicate if they had taken the survey on a previous trip and, if they had taken the survey, instructed them to continue to complete another survey. Answers to questions 1, 2, 3 & 5 were nominal while question 4 was numerical. Answers to all questions were exhaustive and mutually exclusive except for questions 1 and 3 which asked for University affiliation¹ and student status respectively. Respondents could possess more than one campus affiliation but were asked to choose the best answer.

¹ Campus affiliation may also be referred to as campus/University occupation in this report.

B. SAMPLING PLAN

The sampling plan was developed to meet the following conditions:

- All campus bus routes will be sampled (routes 80, 81, 82, 84, & 85).
- All routes will be sampled during standard service schedule (when classes are in session).
- Each route will be sampled during a typical weekday and, if applicable, weekend.
- A minimum target sample will be identified for each route and day.
- The statistical confidence for the entire campus bus system will be at least 95% (+/-5%).
- The statistical confidence for each route (excluding 84) should be at least 90% (+/-5%).
- The statistical confidence for each day (i.e. Wed & Fri) should be at least 90% (+/-5%).
- Statistical confidence of other strata of significance will be determined and reported.

Study Population and Sample Selection

The study population is defined as all campus bus trips except those trips taken by young children. Defining the study population as transit trips rather than transit riders is an important distinction. Indeed, the Transportation Research Board (TRB) reports that,

“[f]ocusing on trips is most appropriate when the information will be used to profile characteristics of trips such as O [origin] & D [destination] patterns and trip purposes. The objective is to obtain a completed survey for each customer trip in the sample frame. Thus, a rider who is encountered by survey workers twice is asked to complete two surveys” (TRB 2005 pp. 13-14).

Focusing on transit riders is more appropriate for surveys that ask demographic, satisfaction or attitudinal questions, in which case each rider is surveyed only once (TRB 2005).

Simple random sampling, where each trip has an equal chance of being surveyed, is sometimes used in on-board transit surveys. However, a simple random sample is not the ideal method for this on-board transit survey. If trips were randomly sampled, the survey may omit entire routes, days or time periods of interest. Because statements need to be made about routes, days or time periods, each segment must be represented. This results in single sampled bus trips being clustered into groups by route, day, and time-of-day. This process is called stratification. Stratification aims to maintain representation of key subgroups in the broader population (USDOT 1996, TRB 2005). Stratified random sampling and the consideration of statistical significance and error in the stratified sample allows for inferential statements to be made regarding a particular stratum of the population (TRB 2005, CUTR 2002).

Based on the sampling plan conditions and campus bus system characteristics strata can be identified and sample frames for each stratum can be determined. For example, the sample frame of one stratum is all trips on a Tuesday on route 80. Eighteen (18) sample frames were identified and described for seventeen (17) strata. The eighteenth sample frame is all routes and all days equaling all survey responses. Each of these strata can be sampled and reported on independently. The sampling plan for each stratum was influenced by the sample size required of each stratum. Determination of the sample size of each stratum is discussed in the sample size section below.

Thus, a stratified sampling plan consisting of a two-phase approach was chosen. The first phase included determining a selection of transit trips from all campus bus route trips taken and, second, sampling passengers of the chosen transit trips. During selection of the transit trip sample, consideration was given to:

- budgetary and scheduling constraints,
- service schedules (recess vs. standard),
- historical trip counts,
- day-of-the-week,
- time-of-day, and
- peak ridership times.

Budgetary and scheduling constraints did not allow sampling of all transit trips or sampling during both recess and standard service hours. Historical trip counts suggested that ridership is highest and most consistent in February during the spring semester. Because weather events have an impact on travel mode selection, a range of dates spanning at least three days was selected to ensure possible impacts of weather events were mitigated as much as possible. Consideration of class schedules, peak vs. off-peak ridership, and day-of-the-week resulted in the decision to sample on a Wednesday, Friday, Saturday and Tuesday. Tuesdays and Thursdays (as well as Mondays, Wednesdays and Fridays) exhibit similar ridership characteristics due to class scheduling. Weekends exhibit lower ridership. Off-peak and weekend routes are typically surveyed less frequently during on-board transit surveys (FTA 2005). Therefore, the first phase included certain transit trips on all routes during a Wednesday, Friday, Saturday and Tuesday.

The second phase of sampling included determining a selection of passenger trips on chosen transit trips. The sampling plan called for 100% sampling of passenger trips on sampled transit trips. This meant each boarding passenger was offered a survey. Because passenger trips were being sampled, a single passenger was allowed to take the survey multiple times so long as it was on a different transit trip. Multiple responses by passengers were recorded on subsequent surveys.

Sample size

To determine how many samples were needed from each stratum to maintain an acceptable level of statistical confidence a sample size was estimated. Sample size is a function of error rate (or confidence interval) and confidence level. A 95% or 90% confidence level and an error of +/- 5-10% are commonly accepted among transit surveys (CUTR 2002, TRB 2005, USDOT 1996). For populations that are very large or unknown, sample sizes can be calculated using the following equation²:

$$SS = [Z^2 * p * (1-p)] / CI^2$$

Where,

SS = Sample Size

Z = Confidence level (i.e. 95% = 1.96)

p = Proportion of sample elements having particular attribute (50%)³

CI = Confidence interval (i.e. 0.05 for 5%)

However, for studies where populations are known or when it can be reasonably expected that 5% or more of the population will be sampled, then a finite population correction can be applied after the initial sample size calculation to obtain a better estimate. In the case of this study, the population could be reliably estimated and it was assumed that 5% or more of the population would be sampled from many of the stratum.⁴ Thus the following finite population correction was applied,

$$NEW\ SS = SS / [1 + (SS - 1) / N]$$

Where,

NEW SS = New Sample Size

SS = "Old" Sample Size

N = Population

The finite population correction better estimates sample sizes for routes, days and time of day when ridership is low. Without finite population correction a stratum with a relatively small population would have the same sample size requirements as a stratum with a very large population (USDOT 1996). For example, at a 95% confidence level and +/- 5% error the sample requirements for the route 80 on a weekday would be 384 from a population of over 12,000. This would not be a problem. However, without the finite population correction the same sample size would be required of the route 84 which only averages 130 trips per day. At an expected response rate of 30% a total of 39 surveys would be expected to be returned each day on route 84. Achieving a sample of 384 on the route 84 would then take 10 days of sampling. Multiplying this across all strata of routes, days, and times this amount of sampling becomes

² This equation assumes that the population is diverse and sampling errors and biases are minimized.

³ The proportion of sample elements having a particular attribute was kept at 50% (0.50) to ensure the estimated sample size calculated was large enough to maintain representation.

⁴ February 2011 ridership data from Madison Metro was used to estimate ridership. Average ridership for each stratum was calculated by dividing ridership by operational days of that stratum.

impossible. Therefore, when the population is known a sample size correction can be applied while maintaining high levels of statistical confidence.

Table 1 shows the estimated sample size with and without finite population correction. The confidence level and error of each stratum is assigned based on the sampling plan conditions identified at the beginning of section B.

Response Rate

A response rate of between 20-40% is typical of on-board transit surveys (CUTR 2002) but ranges have been reported between 13-90% (TRB 2005). A response rate is often measured as the ratio of returned surveys to distributed surveys. In this case, the response rate is measured as the ratio of returned surveys to boarding passengers. This is because surveyors were instructed to ask each boarding passenger to take a survey and return it when finished. If a passenger refused to take the survey, this was considered a non-response and a paper survey was not wasted or discarded uncompleted on the bus. This method has been practiced among other on-board transit surveys (FTA 2002, TRB 2005).

Sampling Bias

Sample bias refers to some members of the population being less likely to participate in the survey than other members. Three common forms of sample bias exist among transit surveys including non-coverage bias, self-selection bias, and non-response bias (CUTR 2002). These types of bias are best controlled by planning and implementing an appropriate sampling plan that includes surveying all routes across all times over many days (CUTR 2002, USDOT 1996). Further, all boarding passengers were asked to complete a survey and surveyors were instructed to assist any persons who requested help completing a survey. However, at least a small amount of bias exists in nearly all surveys. Potential sources of sample bias in this survey are discussed further in the conclusions section. Further, bias can be partially controlled for by weighting responses when reporting results (TRB 2005). Data weighting is discussed further in section G.

Table 1: Calculation of Estimated Sample Sizes for Individual Strata

Strata	Operational Days	Ridership*	Average Feb. Ridership (N)[†]	Confidence Level	Error Rate	Est. SS	Est. SS w/ Finite Population Correction
All	27	340,198	12,600	95%	+/- 5%	384	373
Route 80	27	250,475	9,277	90%	+/- 5%	271	263
Route 81	27	15,327	568	90%	+/- 5%	271	183
Route 82	27	27,291	1,011	90%	+/- 5%	271	214
Route 84	19	2,468	130	90%	+/- 5%	271	88
Route 85	19	44,637	2,349	90%	+/- 5%	271	243
All Routes Tue	4	69,411	17,353	95%	+/- 5%	384	376
All Routes Wed/Fri	8	108,845	13,606	95%	+/- 5%	384	374
All Routes Weekday	19	309,007	16,264	95%	+/- 5%	384	375
All Routes Weekend	8	31,191	3,899	95%	+/- 5%	384	350
80 Weekday	19	234,015	12,317	90%	+/- 5%	271	265
80 Weekend	8	16,460	2,058	90%	+/- 5%	271	239
81 Weekday	19	12,106	637	90%	+/- 5%	271	190
81 Weekend	8	3,221	403	90%	+/- 5%	271	162
82 Weekday	19	15,781	831	90%	+/- 5%	271	204
82 Weekend	8	11,510	1,439	90%	+/- 5%	271	228
84 Weekday	19	2,468	130	90%	+/- 5%	271	88
85 Weekday	19	44,637	2,349	90%	+/- 5%	271	243

* February 2011 ride data from Madison Metro was used to determine ridership by strata.

[†] Average ridership for each stratum was determined by dividing ridership by operational days of that stratum.

C. SURVEY TARGETS

The survey targets were trips of the campus bus system. As noted earlier in the sampling plan section, an attempt was made to ask all boarding passengers on a sampled transit trip to complete a survey, even if they had completed a survey on a previous trip. However, sampling all riders during class change times was difficult due to the crowding on the bus. It is important to note that the goal of the survey was to identify characteristics about the transit trip rather than the rider. Although riders were asked to answer questions about themselves, what was really being surveyed was the transit trip.

D. PRE-TEST

A pre-test was not conducted outside of consulting with Transportation Services staff, other related University staff, Madison Metro staff, and administrators and consultants of the 2008 Madison Metro on-board survey. Pre-tests can help identify implementation problems, survey design flaws or other important issues to address before implementation. However, careful planning, review and research can mitigate many of these concerns and a pre-test was determined not to be necessary.

E. FIELD IMPLEMENTATION

The on-board survey was advertised by TS through the internet, a press release, and posters at campus bus shelters. Madison Metro also sent rider alerts and communicated to drivers that a survey would take place. No survey identification materials were placed on the bus itself.

Four (4) of the five (5) campus bus routes serve stops adjacent to the Memorial Union. A TS office near this location was utilized as the administration and survey storage location. Transportation Services full-time, limited term, and student staff were enlisted to conduct the survey. Surveyors were offered cards and lanyards identifying them as survey crew members. Survey crew members were asked to arrive at this location approximately 30 minutes before the start of their shift to meet their crew partner, pick up supplies, and find the appropriate bus stop. Crew members boarded the appropriate route with the survey materials, which included:

- sequentially numbered survey batches,
- clipboards,
- golf pencils,
- survey return boxes,
- clicker counters, and
- pencil sharpener and erasers.

Once boarded, the crew members found an appropriate location to store survey materials and began surveying boarding passengers. One crew member typically would locate themselves near the front of the bus and approach passengers as they boarded. Boarding passengers were asked to complete the survey and if they agreed were given a survey and pencil and instructed to return the survey to a crew member or place in the return box. No mail-in return option was provided. The other member would sit towards the middle of the bus, count boarding passengers, collect completed surveys and pencils, and distribute additional surveys if

necessary. This process continued for the duration of the shift. Shifts ranged from 2-5 hours in duration. Some shifts only required staffing of one person due to low ridership numbers.

When shifts ended, surveyors were instructed to exit the bus and return surveys and materials to the administration location. Surveyors were also instructed to complete a surveyors trip log that recorded trip information, including the start and end times, start and end survey numbers, the total boarding passengers and additional notes. After the survey was completed, crew members were asked to complete questions pertaining to survey bias and sampling.

Completed and blank surveys were separated and kept organized by day, route and shift. A preliminary hand-count of completed surveys was conducted by shift. This data was used to compare with the scanned data to identify any potential inconsistencies. Surveys were then inspected and delivered for scanning.

F. DATA ENTRY AND CLEANING

Data entry was accomplished by scanning survey forms. Survey questions were answered by filling in a bubble with a #2 pencil next to the appropriate answer. Completed surveys were sorted by day, route and shift. Surveys were inspected for physical damage, stray pencil marks, or incompletely filled bubbles that would have prevented proper scanning. Surveys were prepared for scanning by recoding physically damaged surveys, filling in bubbles completely, and/or erasing stray pencil marks. A preliminary count of completed surveys was also conducted during this time. No responses were changed during this step.

Completed, prepared surveys were delivered to and scanned by UW - Madison Testing and Evaluation Services. The scanned data results were returned in .csv format. Raw data files in .csv and .xlsx format were saved. The raw, original data was preserved by creating a separate, "cleaned" data file. A total of 2,775 surveys were scanned.

Surveys with robust design and data collection can still contain data irregularities and errors. Some of these errors can be the result of imperfect survey design and/or implementation. Other errors can be the result of respondent fatigue, scanning errors, respondent misrepresentation, or other causes. These errors required review and identification, and potential correction through the process of data cleaning. Data cleaning is the process by which these secondary types of errors are identified and eliminated or minimized, resulting in a "cleaner" and more accurate data set. Data cleaning is based on a technical understanding of the survey and an expected range of accepted or possible values (Kavanaugh (no date), Van den Broeck et. al. 2005). However, it is important to note that data cleaning is not a remedy for an improperly designed or conducted survey. If conducted purposefully and transparently, data cleaning can allay concerns of data manipulation (USDOT 1996, Van den Broeck et. al. 2005).

The purpose of data cleaning is to correct and/or mitigate responses that are either; a) impossible, b) suspect or unlikely, c) erroneous inliers, or d) missing. Cleaning is conducted with the goal of ensuring and, indeed, enhancing data integrity (USDOT 1996, Van den Broeck et. al. 2005). Once errors are defined and identified, corrective action is taken and the data is cleaned. A preliminary review of the data suggested that data cleaning would be appropriate. A five step data-cleaning process was employed and included:

1. Preserve raw, original data by saving a new copy of the cleaned data.
2. Define problematic responses.
3. Define mitigation actions.
4. Identify and flag defined errors.
5. Clean errors.

See Appendix D for a table of errors, decision rules and error count information.

G. DATA WEIGHTING

Data weighting is the process of reassigning weight to segmented data in order to more equally represent strata within the total. This process helps correct the potential over- or under-representation of data (CUTR 2006). Weighting is based off the actual responses and the total riders by route, day or other segment. Data was not weighted in this report to maintain consistency with the 2008 Madison Metro bus survey and to diminish the potential for misinterpretation of survey results.

III. RESULTS

Results presented include results by question, basic summary data and cross-tabbed results of particular interest. The data are cleaned and non-weighted. Each table includes a brief summary of results and, if necessary, further explanation of the survey question or other special considerations. Strata sample sizes, confidence levels and error rates can be found in Appendix B.

Total responses: **2,775**
Total boarding passengers: **8,371**
Response rate: **33%**

Table 2: Response Rates by Survey Question

Question	Responses	Response Rate
Completed Previously (YES)	189	6.8%
Q1	2,762	99.5%
Q2	2,758	99.4%
Q3	2,318	83.5%
Q4	2,245	80.9%
Q5	2,708	97.6%

*Table does not include responses with multiple entries.
N = 2,775

Table 2 shows response rates by survey question. 6.8% of respondents indicated they had completed a survey on a previous transit trip. Response rates were high (>97%) for questions that all respondents were asked to complete (Q1, Q2, & Q5). Q3 and Q4 began “If you are a UW student...” and were intended to be answered only by UW students, thus slightly lower response rate is reported.

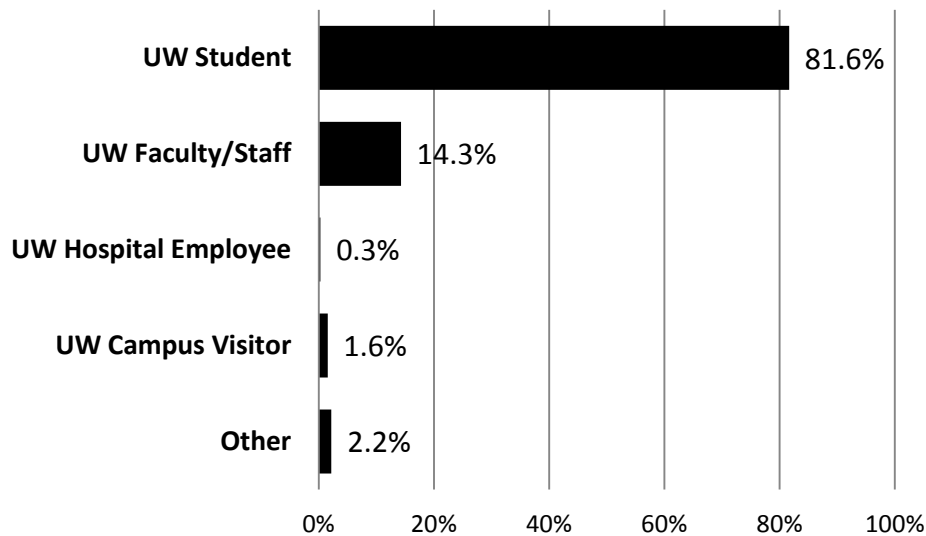
A. RESULTS BY QUESTION

Q1: Please check which best describes you:

Table 3: Results by Occupation

Occupation	Frequency	Percent
UW Student	2,255	81.6%
UW Faculty/Staff	395	14.3%
UW Hospital Employee	8	0.3%
UW Campus Visitor	43	1.6%
Other	61	2.2%
	2,762	100.0%

Figure 1: Results by Occupation



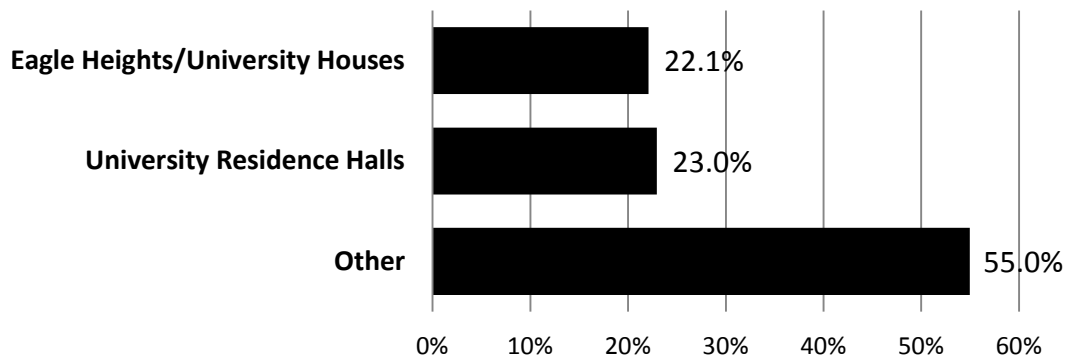
81.6% of question one (1) respondents identified themselves as UW students. 14.3% of respondents identified themselves as UW faculty or staff. 4.1% of respondents identified themselves as a UW Hospital employee, UW campus visitor, or other. Four (4) surveys had multiple responses. Nine (9) surveys did not answer this question. Therefore, 2,762 of 2,775 responses are included in Table 3 and Figure 1.

Q2: I currently live in:

Table 4: Results by Residence

Residence	Frequency	Percent
Eagle Heights/University Houses	609	22.1%
University Residence Halls	633	23.0%
Other	1,516	55.0%
	2,758	100.0%

Figure 2: Results by Residence



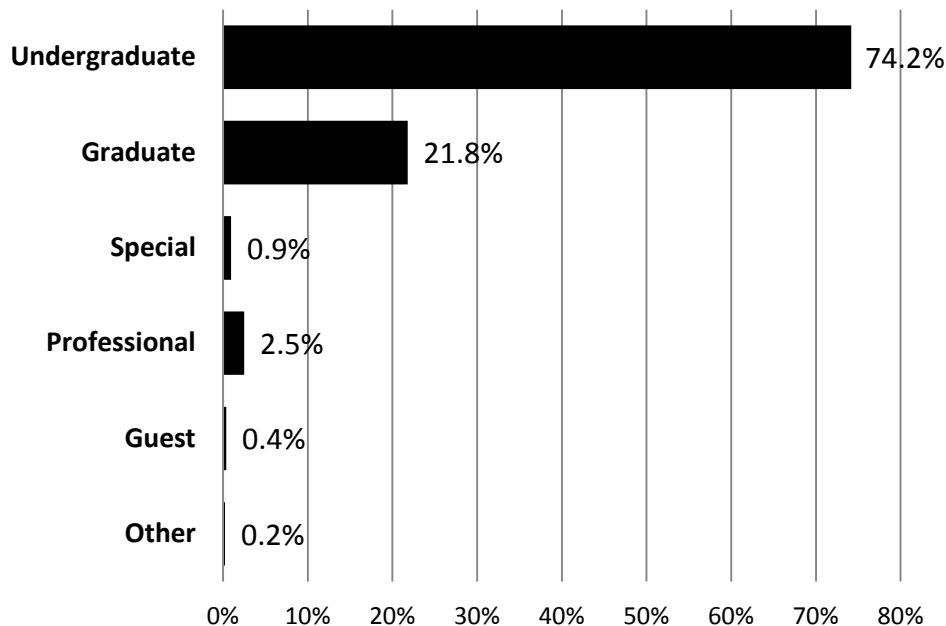
22.1% of question two (2) respondents indicated they lived in Eagle Heights/University Houses. Eagle Heights and University Houses refer to a specific community that is geographically separated from most other University Housing, such as the Lakeshore or Southeast residence halls. Additionally, Eagle Heights and University Houses cater specifically to graduate students, many of who may have families. 23% of respondents said they lived in University Residence Halls and 55% indicated they lived elsewhere. One (1) survey had multiple responses and sixteen (16) surveys did not answer this question. Therefore, 2,758 of 2,775 responses are included in and Figure 1.

Q3: If you are a UW student, the status that best describes you is:

Table 5: Results by Student Status

Student Status	Frequency	Percent
Undergraduate	1,664	74.2%
Graduate	489	21.8%
Special	21	0.9%
Professional	56	2.5%
Guest	8	0.4%
Other	5	0.2%
	2,243	100.0%

Figure 3: Results by Student Status



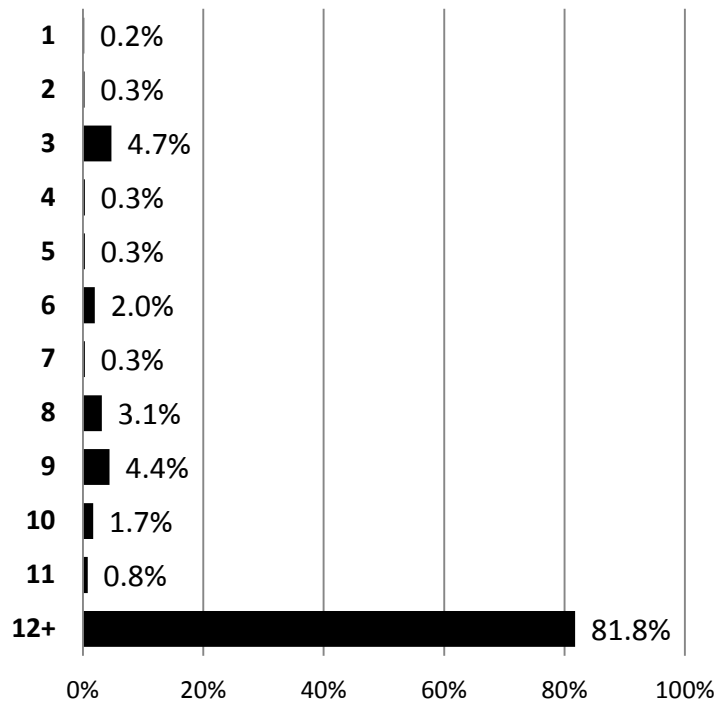
Only responses indicating they were a student in question one and answered question three are included in this table (2,243 of 2,775). 74.2% of respondents indicated they were undergraduate students while 21.8% said they were graduate students. 4% of respondents said they were either special, professional, guest or other students. Two (2) surveys had multiple responses for question three and ten (10) surveys answered question one as students but did not answer question three.

Q4: If you are a UW student, how many credits are you currently registered for?

Table 6: Results by Academic Credit

Credits	Frequency	Percent
1	5	0.2%
2	6	0.3%
3	106	4.7%
4	7	0.3%
5	7	0.3%
6	44	2.0%
7	7	0.3%
8	70	3.1%
9	99	4.4%
10	38	1.7%
11	18	0.8%
12+	1,826	81.8%
	2,233	100.0%

Figure 4: Results by Academic Credit



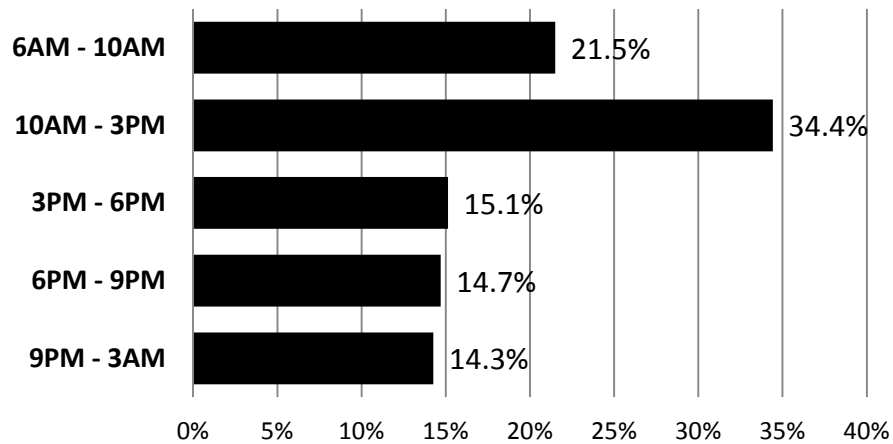
Only responses indicating they were a student in question one and answered question four are included in this table (2,233 of 2,775). 81.8% of respondents indicated they were taking 12 or more credits. 18.2% of respondents said they were taking 11 credits or less. Two (2) surveys had multiple responses for question four and twenty (20) surveys answered question one as students but did not answer question four.

Q5: What time is it?

Table 7: Survey Responses by Time

Time	Frequency	Percent
6AM - 10AM	582	21.5%
10AM - 3PM	932	34.4%
3PM - 6PM	410	15.1%
6PM - 9PM	398	14.7%
9PM - 3AM	386	14.3%
	2,708	100.0%

Figure 5: Survey Responses by Time



17 responses had multiple answers and 50 respondents did not answer the question. 2,708 of 2,775 responses were included in Table 7 & Figure 5.

B. ADDITIONAL SUMMARY RESULTS

Table 8: Survey Responses by Route

Route	Frequency	Percent
80	1,539	55.5%
81	182	6.6%
82	243	8.8%
84	66	2.4%
85	745	26.8%
	2,775	100.0%

Table 9: February 2011 Campus Bus Rides by Route

Route	Frequency	Percent
80	250,475	73.6%
81	15,327	4.5%
82	27,291	8.0%
84	2,468	0.7%
85	44,637	13.1%
	340,198	100.0%

Source: Madison Metro ridership data

Figure 6: Comparison of February 2011 Rides and Survey Responses by Route

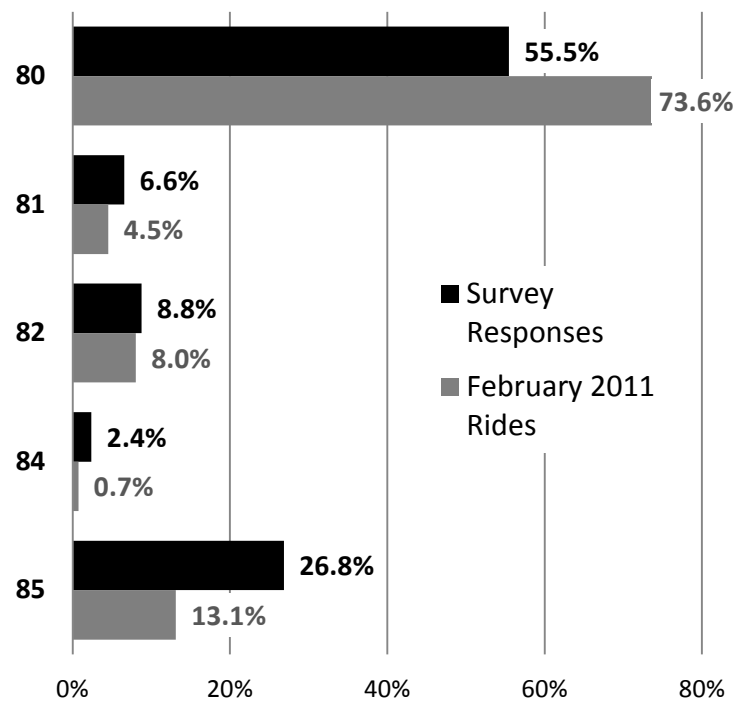


Table 8 shows surveys collected by route. Table 9 shows February 2011 campus bus ridership by route and Figure 6 shows a comparison of the two data sets. 55.5% of surveys were collected on route 80 while 26.8% of surveys were collected on route 85. 8.8% were collected on route 82, 6.6% on route 81 and 2.4% on route 84. The number of surveys collected by route is a result of many factors, including ridership, survey scheduling, and response rates.

Table 10: Survey Responses by Day

Day	Frequency	Percent
Wednesday (2/8)	955	34.4%
Friday (2/10)	729	26.3%
Saturday (2/11)	396	14.3%
Tuesday (2/14)	695	25.0%
	2,775	100.0%

Figure 7: Survey Responses by Day

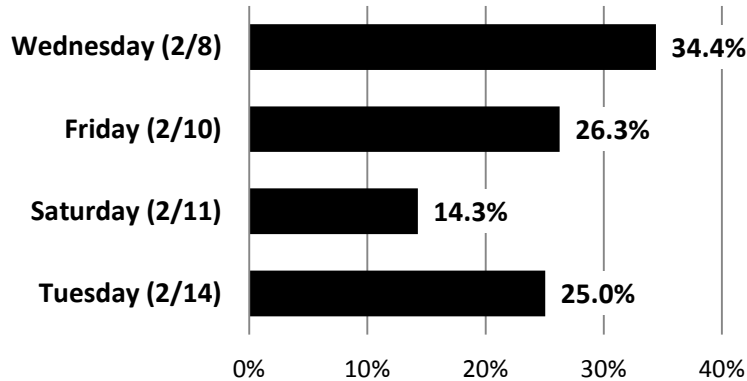


Table 10 shows surveys collected by day and associated percentages. The number of collected surveys is a result of ridership, survey scheduling and response rates.

C. CROSS TABBED RESULTS

Table 11: Results by Route/Occupation

	80	81	82	84	85	<i>total</i>
UW Student	1,207	171	228	33	616	2,255
UW Faculty/Staff	242	4	3	32	114	395
UW Hospital Employee	7	1	-	-	-	8
UW Campus Visitor	30	4	5	1	3	43
Other	40	2	7	-	12	61
	1,526	182	243	66	745	2,762

Figure 8: Results by Route/Occupation

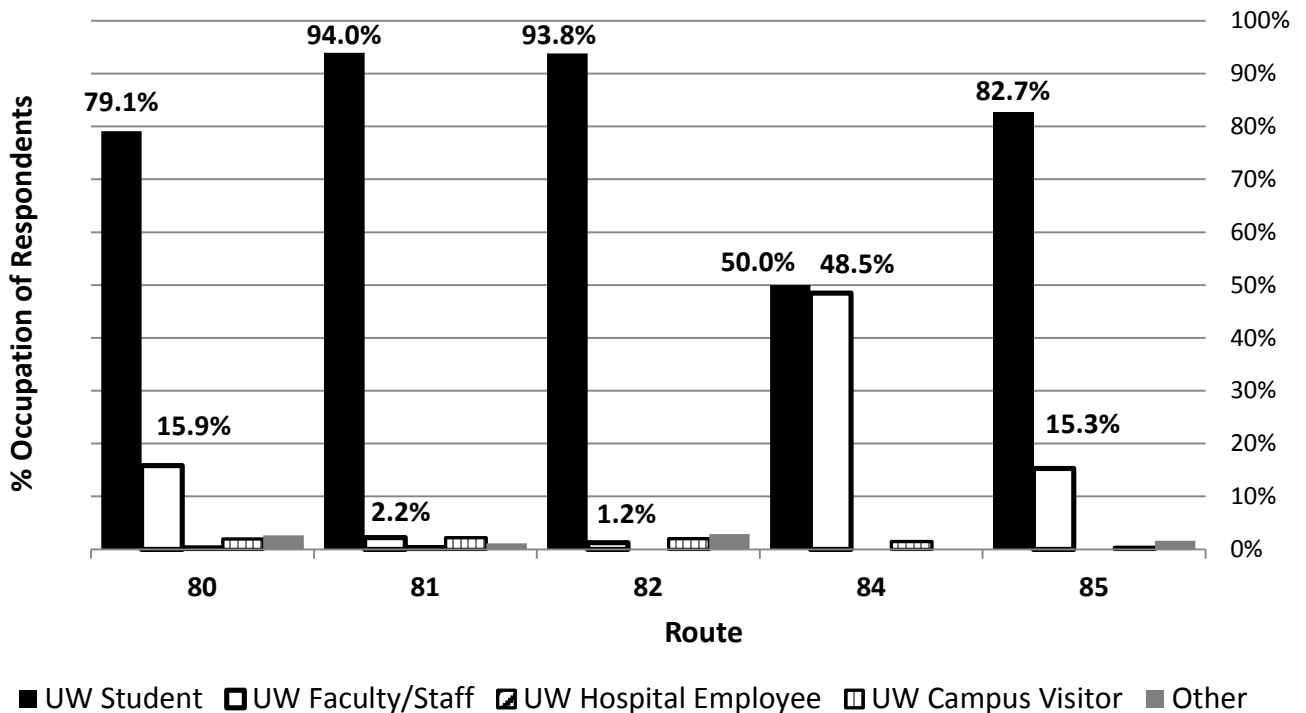


Table 11 shows results by route and occupation. The majority of survey respondents on all routes indicated they were UW students. Expectedly, a particularly high percentage (>93%) of respondents were UW students on routes 81 and 82, which are late night service routes. Also notable, the percentage of respondents on route 84 (Eagle Heights Express) was approximately 50% UW students and 50% UW faculty/staff.

Table 12: Results by Day/Occupation

	WED 2/8		FRI 2/10		SAT 2/11		TUE 2/14		<i>total</i>
		%		%		%		%	
UW Student	781	82.1%	570	78.4%	307	77.5%	597	86.8%	2,255
UW Faculty/Staff	140	14.7%	131	18.0%	50	12.6%	74	10.8%	395
UW Hospital Employee	3	0.3%	-	0.0%	3	0.8%	2	0.3%	8
UW Campus Visitor	8	0.8%	13	1.8%	16	0.8%	6	0.9%	43
Other	19	2.0%	13	1.8%	20	5.1%	9	1.3%	61
	951	100.0%	727	100.0%	396	96.7%	688	100.0%	2,762

Table 13: Results by Day/Residence

	WED 2/8		FRI 2/10		SAT 2/11		TUE 2/14		<i>total</i>
		%		%		%		%	
Eagle Heights/University Houses	215	22.6%	170	23.5%	137	34.6%	87	12.7%	609
University Residence Halls	191	20.0%	177	24.4%	142	35.9%	123	18.0%	633
Other	547	57.4%	377	52.1%	117	29.5%	475	69.3%	1,516
	953	100.0%	724	100.0%	396	100.0%	685	100.0%	2,758

Table 14: Results by Occupation/Time of Day

	6AM - 10AM		10AM - 3PM		3PM - 6PM		6PM - 9PM		9PM - 3AM		<i>total</i>
		%		%		%		%		%	
UW Student	419	72.0%	783	84.2%	335	81.9%	333	83.7%	347	89.9%	2,217
UW Faculty/Staff	143	24.6%	115	12.4%	57	13.9%	42	10.6%	21	5.4%	378
UW Hospital Employee	5	0.9%	-	0.0%	-	0.0%	1	0.3%	2	0.5%	8
UW Campus Visitor	5	0.9%	8	0.9%	10	2.4%	12	3.0%	7	1.8%	42
Other	10	1.7%	24	2.6%	7	1.7%	10	2.5%	9	2.3%	60
	582	100.0%	930	100.0%	409	100.0%	398	100.0%	386	100.0%	2,705

IV. CONCLUSION

This section presents a summary of the on-board campus survey results, a discussion of survey limitations, reports on potential sources of sample bias, and covers further data needs. This conclusion does not offer analysis or recommendations regarding campus bus service.

Results were reported in a cleaned, non-weighted format to maintain consistency with the reporting of results by the 2008 Madison Metro survey. This, coupled with similarities in methodology, allows for comparison between surveys of certain results of interest, including ridership by occupation.

A response rate of 33% yielded 2,775 completed surveys. 81.6% of respondents who answered question one (1) identified themselves as students while 14.3% were UW faculty or staff, and 4.1% said they were UW Hospital employees, visitors or other. This is consistent with the 2008 on-board survey conducted by Madison Metro that also found 81% of trips were consumed by students.

22.1% of question two (2) respondents indicated they lived in Eagle Heights/University Houses. 23% of respondents said they lived in University Residence Halls and 55% indicated they lived elsewhere. 74.2% of student respondents indicated their status as undergraduate while 21.8% said they were graduate students. 4% of respondents said they were special, professional, guest or other student and two (2) surveys had multiple responses while ten (10) surveys did not answer question three.

81.8% of student respondents indicated they were taking 12 or more credits while 18.2% of respondents said they were taking 11 credits or less. Two (2) surveys had multiple responses for question four and twenty (20) surveys did not answer question four.

55.5% of surveys were collected on route 80 while 26.8% of surveys were collected on route 85. 8.8% were collected on route 82, 6.6% on route 81 and 2.4% on route 84. 34.4% of surveys were collected on Wednesday, February 8th, 26.3% on Friday, February 10th, 14.3% were collected on Saturday, February 11th, and 25% on Tuesday, February 14th. Further data and analysis can be found in the Results section.

A. SURVEY OBJECTIVES AND ACCOMPLISHMENTS

The following discusses survey accomplishments in relation to the objectives identified in the Introduction.

1. Sample 100% of campus bus trips on selected bus routes.
Operationally, this goal means sampling 100% of boarding passengers on routes that were being surveyed. While we thought this may be possible at the beginning of survey planning, we also understood that the campus busses become crowded and 100% sampling may not be possible. Ultimately, survey crew members reported that they were not able to survey all boarding passengers, particularly during class change times. While an attempt was made to survey all boarding passengers we do not believe this adversely effected results. However, this may have contributed small amounts of sample bias in results due to the inherent characteristics of riders during class change times and is discussed in more detail below.
2. Survey all campus bus routes (routes 80, 81, 82, 84, & 85).
All campus bus routes were surveyed.
3. Achieve an approximate 30% response rate.
A response rate of 33% was achieved.
4. Achieve statistically significant, representative samples stratified by routes & days.
Statistically significant samples were obtained for routes 80, 81, 82, and 85. A less statistically confident sample was obtained from route 84 because of very low ridership limiting survey responses. Statistically significant samples were obtained for the campus transit system as a whole and for each day that was sampled. Appendix B contains a table with actual sample sizes, confidence levels and error rates.

B. SURVEY LIMITATIONS

This section discusses limitations of the 2012 on-board campus bus survey. The first limitation is that the survey is not necessarily representative of trip characteristics during academic breaks (or recess service). For example, ridership by route or day may vary in the month of July when classes are not in session and busses run less frequently.

However, this is not a significant barrier to applying survey results to the larger campus bus rider population. This is for two reasons. First, the results that were obtained during the survey are of high statistical confidence which would presumably compensate for not surveying during recess service. Second, although approximately one third of days during the year are in recess service it is inaccurate to assume that the survey does not comprehensively reflect campus bus trips. Most notably, this is because nearly 9 out of 10 trips taken on campus bus routes are taken during standard service hours. While surveying during recess service would fill small data gaps, considering time and effort commitments, there is no robust argument to survey during these times. There is strong evidence that applying the survey results to the broader campus bus rider population is appropriate.

The second limitation of the survey concerns low statistical confidence during specific strata, such as off-peak times or low-ridership routes. For example, because the route 84 has

relatively low average daily ridership it is difficult to acquire enough surveys to obtain high statistical confidence. There is little to remedy this situation besides accepting lower levels of statistical confidence.

C. POTENTIAL SOURCES OF SAMPLE BIAS

This section discusses sample bias and potential sources in the on-board survey. While it is assumed that small amounts of bias exist in almost all surveys, without controlling for large amounts of bias survey results can be useless or misleading. To control for bias a sampling plan was produced that identified and defined potential sources of bias before surveying. The survey then incorporated bias mitigation into implementation. Potential sources of bias can then be reviewed at the conclusion of the survey. It was determined that the sampling plan and implementation for this on-board survey was sufficient to control for major sources of bias.

Sample bias refers to some members of the population being less likely to participate in the survey than other members. This produces bias in the final results if one group, route, or time is over- or under-represented. As introduced in the methods section, three common forms of sample bias exist among on-board transit surveys including non-coverage bias, self-selection bias, and non-response bias (CUTR 2002). A more detailed discussion of the types of sample bias can be found in Appendix E.

Although no major sources of bias were discovered, surveyors were asked questions relating to sample bias after the conclusion of the survey. The purpose of the questions was to help assess potential sources of bias not identified in the sampling plan. Responses from surveyors combined with knowledge from previous on-board transit surveys and supervisor observations informed the discussion of sample bias.

Small amounts of sample bias were identified in two places. First, sampling all boarding passengers during class change times was difficult because of crowded busses and 100% sampling was not accomplished. This led to a segment of the population not being surveyed as accurately and may have contributed to sample bias. This segment of the population can be defined as “peak”, “overload” or “class-change” riders. While not all trips taken during these times were taken due to class changes, it can be conservatively estimated that at least 50% of trips had an origin or destination of a UW class. Those trips would have necessarily been taken by UW student riders and thus because they were not sampled may have resulted in UW students being under-represented in the final results. This form of sample bias is classified as non-coverage bias. However, there is no way to determine the amount of UW student under-representation and it is therefore not reported.

The second form of bias worth noting concerns the self-selection bias of distracted riders. Riders more likely to be wearing headphones, using a mobile phone, or talking with colleagues or friends had an additional barrier preventing them from participating in the survey. It is very difficult to measure or estimate the characteristics of distracted riders and thus this effect on results cannot be reported.

D. FUTURE DATA AND ANALYSIS NEEDS

No major data needs can be identified besides surveying longer during off-peak ridership times to increase confidence levels.

Further analysis needs of the current data could include a discussion of data that is weighted and any other cross-tabbing or visualization of results that would help better understand campus bus ridership.

REFERENCES

- Baltes, M. 2003. *Conducting a Successful On-Board Survey of Public Transit Customers*. Presented at the Proceedings of the 2003 Mid-Continent Transportation Research Symposium.
- Center for Urban Transportation Research (CUTR). 2002. *Customer Surveying For Public Transit: A Design Manual for On-Board Surveys*. Available to the public through the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, VA 22161, <http://www.ntis.gov>, and through the NCTR web site at <http://www.nctr.usf.edu>.
- Federal Transit Administration. (FTA). 2002-2004. Transit Performance Monitoring System (TPMS) Results, Summary Report Phase I, II, & III.
- Kavanaugh, A., Evered, S., & Krizan, M. (no date). *Data cleaning primer with worked examples*. Prepared for Blue Cross Blue Shield Minnesota. Center for Prevention.
- City of Madison. 2000. Madison Metro 2000 On-Board Passenger Survey
- City of Madison. 2008. Madison Metro 2008 On-Board Passenger Survey
- Transportation Research Board (TRB). 2005. TCRP Synthesis 63. *On-Board and Intercept Transit Survey Techniques: A Synthesis of Transit Practice*. Transit Cooperative Research Program.
- United States Department of Transportation (USDOT). 1996. *Travel Survey Manual*. Travel Model Improvement Program. TRB continuously updated document available at <http://www.travelsurveymanual.org/>
- Van den Broeck, J., Argeseanu Cunningham, S., Eeckels, R. & Herbst, K. (2005). *Data cleaning: Detecting, diagnosing, and editing data abnormalities*. PLoS Med 2(10): e267.

APPENDIX A: On-Board Survey Instrument

Figure 9: On-Board Survey Instrument

2012 CAMPUS BUS PASSENGER SURVEY

Fill bubble if you already filled out a questionnaire on another trip and continue to complete this form.

Fill bubbles completely with #2 pencil.

1. Please check which best describes you (fill one):

UW Student UW Faculty/Staff UW Hospital Employee UW Campus Visitor Other

2. I currently live in (fill one):

Eagle Heights/University Houses University Residence Halls Other

3. If you are a UW student, the status that best describes you is (fill one):

Undergraduate Graduate Special Professional Guest Other

4. If you are a UW student, how many credits are you currently registered for (fill one)?

1 2 3 4 5 6 7 8 9 10 11 12+

5. What time is it (fill one)?

6AM - 10AM 10AM - 3PM 3PM - 6PM 6PM - 9PM 9PM - 3AM

Please return the completed form to a survey crew member or drop-off box.

ACCU-SCAN™ 01005-0112 (Yellow/Red) APPERSON PRINT RESOURCES 04125-1

APPENDIX B: Survey Sample Size and Statistical Confidence

Table 15: Comparison of Preliminary and Final Estimated Sample Sizes and Confidence Intervals for Individual Strata

Strata	Preliminary C.L.	Preliminary Error Rate	Est. SS	Est. SS w/ Finite Population Correction	Actual SS	Actual C.L.	Actual Error Rate
All	95%	+/- 5%	384	373	2,775	99%	+/- 2.2%
Route 80	90%	+/- 5%	271	263	1,539	99%	+/- 3.0%
Route 81	90%	+/- 5%	271	183	182	90%	+/- 5.0%
Route 82	90%	+/- 5%	271	214	243	90%	+/- 4.6%
Route 84	90%	+/- 5%	271	88	66	90%	+/- 7.1%
Route 85	90%	+/- 5%	271	243	745	99%	+/- 3.9%
All Routes Tue	95%	+/- 5%	384	376	695	99%	+/- 4.8%
All Routes Wed/Fri	95%	+/- 5%	384	374	1,684	99%	+/- 2.9%
All Routes Weekday	95%	+/- 5%	384	375	2,379	99%	+/- 2.4%
All Routes Weekend	95%	+/- 5%	384	350	396	95%	+/- 4.7%
80 Weekday	90%	+/- 5%	271	265	1,205	99%	+/- 3.5%
80 Weekend	90%	+/- 5%	271	239	334	95%	+/- 4.9%
81 Weekday	90%	+/- 5%	271	190	160	90%	+/- 5.6%
81 Weekend	90%	+/- 5%	271	162	22	90%	+/- 17.1%
82 Weekday	90%	+/- 5%	271	204	203	90%	+/- 5.0%
82 Weekend	90%	+/- 5%	271	228	40	90%	+/- 12.9%
84 Weekday	90%	+/- 5%	271	88	66	90%	+/- 7.1%
85 Weekday	90%	+/- 5%	271	243	745	99%	+/- 4.6%

APPENDIX C: Survey Crew Proposed Schedule

Table 16: Survey Crew Proposed Schedule

ROUTE	ARRIVE	START SHIFT	END SHIFT	# HRS	Wednesday, 8th	Friday, 10th	Saturday, 11th	Tuesday, 14th	Subtotal Shifts	Total Hrs
80	7:30	8:00	12:00	4	█			█	1	4
80	6:00	6:30	10:30	4			█		3	12
80	10:00	10:30	14:00	3.5			█		3	10.5
80	11:30	12:00	16:30	4.5	█			█	1	4.5
80	13:30	14:00	16:30	2.5			█		3	7.5
80	16:00	16:30	21:30	5					4	20
80	21:00	21:30	1:00	3.5		█			2	7
80	21:00	21:30	2:30	5	█			█	2	10
85	6:30	7:00	10:30	3.5	█		█	█	3	10.5
85	10:00	10:30	14:00	3.5	█		█	█	3	10.5
85	13:30	14:00	17:30	3.5	█		█	█	3	10.5
81	18:30	19:00	23:00	4					4	16
81	22:30	23:00	1:30	2.5		█			2	5
81	22:30	23:00	3:00	4	█			█	2	8
82	18:30	19:00	23:00	4	█		█	█	4	16
82	22:30	23:00	1:30	2.5	█	█		█	2	5
82	22:30	23:00	3:00	4	█	█	█	█	2	8
84	16:15	16:40	18:40	2			█		3	6
<i>Hr totals</i>					44	48.5	34.5	44	47	171

APPENDIX D: Data Cleaning Process

Completed surveys were sorted by day, route and shift. Surveys were inspected for physical damage, stray pencil marks, or incompletely filled bubbles that would have prevented proper scanning. Surveys were prepared for scanning by recoding physically damaged surveys, filling in bubbles completely, and/or erasing stray pencil marks. A preliminary count of completed surveys was also conducted during this time. No responses were changed during this step.

Completed, prepared surveys were scanned by UW Madison Testing and Evaluation Services. The scanned data results were returned in .csv format on February 17, 2012. Raw data files in .csv and .xlsx format were saved. The raw, original data was preserved by creating a separate, “cleaned” data file. A total of 2,775 surveys were scanned.

Surveys with robust design and data collection can still contain data irregularities and errors. Some of these errors can be the result of imperfect survey design and/or implementation. Other errors can be the result of respondent fatigue, scanning errors, respondent misrepresentation, or other causes. Data cleaning is the process by which these secondary types of errors are identified and eliminated or minimized, resulting in a “cleaner” and more accurate data set. Data cleaning is based on a technical understanding of the survey and an expected range of accepted or possible values. However, it is important to note that data cleaning is not a remedy for an improperly designed or conducted survey. If conducted purposefully and transparently, data cleaning can allay concerns of data manipulation.

The purpose of data cleaning is to correct and/or mitigate responses that are either; a) impossible, b) suspect or unlikely, c) erroneous inliers, or d) missing. Cleaning is conducted with the goal of ensuring and, indeed, enhancing data integrity. Once errors are defined and identified, corrective action is taken and the data is cleaned. A five step data-cleaning process was employed and is described below:

- 1. Preserve raw, original data by saving a new copy of the cleaned data.**

Saved here: P:\Trans\TDM\Programs\Transit\campus
bus\PublicInputProcess\On_Board_Bus_Survey\Data\RAW_Data(DO_NOT_EDIT)

- 2. Define problematic responses**

The purpose of this step is to define erroneous data. Four types of errors can occur including responses that are a) impossible, b) suspect or unlikely, c) erroneous inliers, or d) missing. Table 1 shows error types, examples, and the diagnostic steps taken to identify the errors.

Table 17: Error types, examples and diagnostic steps

Error Type	Erroneous Example	Diagnostic Step
A) Impossible	A respondent indicated that they were on the 85 bus route between 9PM and 3AM. The bus does not run during this time.	Examine entire dataset, errors are obvious
B) Suspect or unlikely	Respondent identifies themselves as a UW Hospital Employee living in the University Residence Halls.	Examine entire dataset
C) Erroneous inliers	Respondent indicates they are a UW student but erroneously checks they live in Eagle Heights/University Houses when they live in University Residence Halls.	Example samples of dataset, errors are difficult/impossible to identify
D) Missing	Respondent failed to complete a required answer.	Examine entire dataset

These errors were identified by sorting the spreadsheet and cross tabbing responses. Once problematic responses were identified decision rules were created.

3. Define mitigation actions

Once problematic observations were identified, a decision needed to be made regarding corrective actions. Three actions are possible, including 1) leaving unchanged, 2) deleting, and 3) correcting.

4. Identify and flag defined errors

The spreadsheet of responses was sorted and cross tabbed to identify errors. Erroneous response examples, error type, occurrence and decision rules for each error are shown in Table 18 below. The corrective action was dependent on the type of error encountered.

Table 18: Survey errors and data cleaning

Error	Error description	Error Type*	Error Occurrence	Decision Rule
E1	No responses, survey not completed	D	0	Delete
E2	'Completed previously' bubble filled, no other bubbles filled	D	8	Unchanged. Response only included in multiple ridership analysis
E3	Q1 multiple answer	B	4	Unchanged. Excluded Q1 from related analysis
E4	Q2 multiple answer	B	1	Unchanged. Excluded Q2 from related analysis
E5	Q3 multiple answer	B	2	Unchanged. Excluded Q3 from related analysis
E6	Q4 multiple answer	B	2	Unchanged. Excluded Q4 from related analysis
E7	Q5 multiple answer	B	17	Unchanged. Excluded Q5 from related analysis
E8	Answered student credits (Q4) but not UW Student (Q1)	B	37	Corrected. IF Q3 = 1, 2, 3, 4, or 5 & Q4 = any, THEN change Q1 to 1 (28 count). IF Q3 = 6 & Q4 = any, THEN no change (6 count). IF Q3 = any & Q4 = blank, THEN no change (3 count).
E9	Answered student status (Q3) but not UW Student (Q1)	B	99	Corrected. IF Q3 = 1, 2, 3, 4, or 5 & Q4 = any, THEN change Q1 to 1 (28 count). IF Q3 = 6 & Q4 = any, THEN no change (6 count). IF Q3 = any & Q4 = blank, THEN no change (3 count).
E10	Missing Q1	D	9	Unchanged. Q1 not included in analysis
E11	Missing Q2	D	16	Unchanged. Q2 not included in analysis
E12	Missing Q5	D	31	Unchanged. Q5 not included in analysis
E13	Answered impossible time for route 81	A	5	Corrected. Time (Q5) was removed from response
E14	Answered impossible time for route 82	A	4	Corrected. Time (Q5) was removed from answer
E15	Answered impossible time for route 84	A	8	Corrected. Time (Q5) was removed from answer
E16	Answered impossible time for route 85	A	2	Corrected. Time (Q5) was removed from answer
E17	Undetermined erroneous inliers	C	?	Unchanged, difficult to identify.

*See Table 18 for error type and description.

5. Errors were cleaned.

Based on decision rules in Table 19, the flagged dataset was cleaned. A final, cleaned copy of the dataset was saved and used for all analysis.

APPENDIX E: Potential Sources of Sample Bias

This section discusses sample bias and potential sources in the on-board survey. While it is assumed that small amounts of bias exist in almost all surveys, without controlling for large amounts of bias survey results can be useless or misleading. To control for bias a sampling plan was produced that identified and defined potential sources of bias before surveying. The survey then incorporated bias mitigation into implementation. Potential sources of bias can then be reviewed at the conclusion of the survey. It was determined that the sampling plan and implementation for this on-board survey was sufficient to control for major sources of bias.

Sample bias refers to some members of the population being less likely to participate in the survey than other members. This produces bias in the final results if one group, route, or time is over- or under-represented. As introduced in the methods section, three common forms of sample bias exist among on-board transit surveys including non-coverage bias, self-selection bias, and non-response bias (CUTR 2002).

Non-coverage bias occurs when results from a census of the population would differ significantly from results of the sample and is difficult to determine because a census of all campus bus trips was not done. However, to control for non-coverage bias all boarding passengers were asked to complete a survey and surveyors were instructed to assist any persons who requested help completing a survey. All routes were surveyed on at least one trip all the hours they operated during at least one weekday and weekend to ensure that no segment of the population was systematically excluded. Self-selection bias occurs when the survey is voluntary in nature and some segment of the population systematically chooses not to take the survey. This is difficult to control for and, indeed, surveyors indicated this bias may be present. This was controlled for by asking all boarding passengers to take the survey and removing barriers that may make certain segments of the population less likely to take the survey. Survey characteristics that may be perceived as barriers include lengthy forms, improperly worded questions, or the exclusion of certain possible answers. Non-response bias occurs when it is known who or what segment of the population does or does not respond to the survey and is the most difficult to control for (CUTR 2002). An example of non-response bias would be if it was known that women were more likely to respond to the survey than men. It is not known that any particular segment of campus bus trips systematically did or did not respond to the survey.

Although no major sources of bias were discovered, surveyors were asked questions relating to sample bias after the conclusion of the survey. The purpose of the questions was to help assess potential sources of bias not identified in the sampling plan. Questions were worded as objectively as possible. Eight (8) responses from surveyors were received. Summarized responses to each question are presented below.

Q1. Did you notice different groups of people were more or less likely to complete a survey?

Four surveyors noted that certain groups of people were less likely to respond to a survey during their shifts. These groups or characteristics included passengers that appeared to be

students and passengers that were distracted either by wearing headphones, texting or otherwise using a mobile phone.

Q2. Did the time of day seem to affect response rates?

Five surveyors responded that the time of day did not affect response rates while three surveyors indicated that the time of day did affect response rates. These surveyors noted that the time of day when classes changed resulting in crowded buses reduced survey response rates.

Q3. Did it seem that riders wearing headphones were more or less likely to respond to the survey?

Five surveyors indicated that passengers wearing headphones were less likely to respond a survey while three surveyors indicated that it did not matter.

Q4. Did a particular part or geography of the route seem to affect response rates?

Four surveyors noted that geography did play a role in response rates and indicated that busy stops, including those near the intersection of Park St. and Johnson St. and Park St. and University Ave. were particularly lower response rates.

Q5. Did you notice other patterns or characteristics during the survey that seemed to affect response rates?

Four surveyors indicated other patterns existed including that some respondents who appeared to be non-native English speakers were less likely to take the survey, that peer pressure played a role (if the first boarding passenger took a survey all passengers took a survey and vice-versa), and standing passengers were less likely to take a survey than those sitting.

Q6. Please describe actions you took if or when the bus became crowded.

Six surveyors responded that they stopped surveying because it was either unsafe or not possible to reach everyone. During this time a number of respondents were able to keep counting boarding passengers while others were not able to. This may have reduced the response rate slightly because those boarding passengers who were counted were not offered a survey. Two other surveyors reported they did not have crowding issues because they were on late night routes.

Q7. Please describe any other difficulty you had in distributing surveys or eliciting responses.

Seven surveyors indicated they had no other specific problems. One surveyor reported that boarding passengers who said they had already completed a survey were less likely to complete another survey, even after being told they could complete another.

APPENDIX F: Selected 2008 Madison Metro on-board survey results

Table 19: Survey Results by Occupation.

Occupation	Frequency	Percent
Student	3,705	81%
Faculty/Staff	626	14%
Hospital Employee	48	1%
Campus Visitor	41	1%
Other	105	2%
No Answer	47	1%
TOTAL	4,572	100%

Source: Madison Metro 2008 on-board survey results. Adapted from Table B8: Occupation. This table included surveys with “no answer” while the 2012 Campus Bus Survey did not include non-responses to the associated question of occupation.

APPENDIX G: Madison Metro 2008 On-Board Survey Methodology

Information collected from the Madison Metro 2008 On-Board Survey Report, Jason Bittner (CFIRE Deputy Director), David Trowbridge (Planner, City of Madison), Tim Sobota (Madison Metro), Anurag Komanduri (Cambridge Systematics), and Transportation Services staff.

Survey Design

A short-form survey was developed for UW Campus routes because trip times tend to be shorter. See Figure 2.2 in the Madison Metro Survey Report for a copy of the survey instrument. The survey asked for the time, origin type and location, destination type and location, routes taken, trip frequency, occupation, gender, race, past use of Madison Metro, and ratings for service, cleanliness, etc. The survey was one page.

Scheduling

Routes 80, 81, 82, and 85 were surveyed between April 15, 2008 and April 21, 2008. Service for route 84 began on August 25th, 2008 which was after the survey was completed and was not surveyed. Surveying was scheduled to capture all service hours but not all days. Weekend routes were not surveyed. A total of approximately 120 hours were spent surveying. General schedule design and scheduling was performed by Madison Metro, the City of Madison, and Cambridge Systematics. Budget considerations also impacted survey schedule design. See table below for a summary of the 2008 campus bus survey schedule.

Survey Logistics

At least two crew members were on the bus to survey. Crew members approached bus passengers after they had entered the bus and had sat or found a place to stand. This was not systematic among routes, times or crew members. Generally, one crew member was at the front of the bus and one at the back, randomly selecting passengers. When the bus was crowded, all passengers could not be approached to take a survey. This was an accepted condition for the 2008 Madison Metro survey. Surveys were returned through the mail, at drop off locations other than the bus, or most often, handed to the crew members.

Response Rate

The 2008 Madison Metro on-board survey did not report a response rate but discussions with survey administrators revealed an approximate response rate of 30% which is consistent with other on-board bus surveys. Survey crew workers counted the number of passengers that boarded at each stop while they were surveying. The final report used the average daily ridership and completed surveys to determine response rate for other Metro routes. While refusals to take the survey were recorded, survey administrators acknowledge that this number could not be used to determine a response rate. This is because many bus passengers were not asked to take the survey due to bus overcrowding or other reasons.

Table 20: Summary of Madison Metro 2008 On-Board Survey Schedule of UW Routes

ROUTE	DAY	START DATE	START TIME	END DATE	END TIME	TOTAL HRS
80	Tuesday	4/15/2008	6:15AM	4/15/2008	9:36PM	15.344
80	Tuesday	4/15/2008	6:14PM	4/16/2008	2:31AM	8.283
80	Wednesday	4/16/2008	6:27AM	4/16/2008	6:19PM	11.867
80	Thursday	4/17/2008	7:36AM	4/17/2008	5:16PM	9.667
80	Friday	4/18/2008	6:12AM	4/18/2008	8:52PM	14.667
80	Monday	4/21/2008	6:14PM	4/22/2008	1:35AM	7.35
81	Wednesday	4/16/2008	6:36PM	4/17/2008	1:51AM	7.25
81	Thursday	4/17/2008	6:37PM	4/18/2008	3:18AM	8.683
82	Wednesday	4/16/2008	6:19PM	4/17/2008	2:06AM	7.783
82	Thursday	4/17/2008	6:19PM	4/17/2008	2:06AM	7.783
85	Thursday	4/17/2008	7:06AM	4/17/2008	6:20PM	11.233
85	Tuesday	4/22/2008	7:15AM	4/22/2008	5:55PM	10.667
						120.577

*Adapted from data supplied by Jason Bittner, Deputy Director, National Center for Freight and Infrastructure Research and Education

Data Entry

Surveys had sequential numbering to facilitate data entry. Data was entered manually into software supplied by prime contractor Cambridge Systematics.

APPENDIX H: Campus Bus Route Maps (Routes 81, 82, and 84 not included)

