Southeast Michigan Council of Governments

Regional On-Board Transit Survey

Final Report

July 2012

NUSTATS RESEARCH SOLUTIONS

206 Wild Basin Rd., Suite A-300

Austin, Texas 78746

Contact: Fred G'sell, Project Manager

(512) 306-9065

www.nustats.com

Table of Contents

List of Tables and Figures

Executive Summary

Background

In the fall of 2010 and spring of 2011, the Southeast Michigan Council of Governments (SEMCOG), with consultant support from NuStats and CDM Smith, conducted a regional on-board origin-destination (O/D) survey of all fixed-route transit systems in Southeast Michigan. The transit systems surveyed were those operated by the Detroit Department of Transportation (DDOT), Suburban Mobility Authority for Regional Transportation (SMART), Ann Arbor Transportation Authority (AATA), University of Michigan (U-M), Detroit People Mover (DPM), Blue Water Area Transit (BWAT), and Lake Erie Transit (LET). All transit operators contributed in-kind services and technical advice for the project. Without their assistance, this survey would not have been possible, or as successful.

The main purpose of this on-board transit survey is to update SEMCOG's Travel Demand Forecasting Model—specifically, to enhance the travel model's transit and mode-choice components. In addition, the collected data provided valuable, current information on travel patterns and demographics of transit riders, as well as transit service characteristics. This information is extremely important for regional transit planning and the transportation investment decision-making process.

Collecting new transit survey data was important for several reasons. The last regional on-board transit survey prior to this one was conducted in 2001, and since then, significant economic and transit system changes have occurred in the region. Furthermore, the region has a renewed interest in expanding rapid transit services, making the need for more up-to-date transit travel information even more critical. As an example, the Federal Transit Administration's (FTA's) New Starts and Small Starts program requires that recent travel information be used in alternatives analyses in order for proposed transit projects to qualify for federal funding.

Survey tasks involved developing a sampling plan; designing the survey instruments; conducting a pilot test; collecting, processing, and geocoding the collected questionnaires; weighting and expanding the data; analyzing the data; and reporting the results. For all systems except the DPM system, a self-administered survey was used. Data collection was performed from September 22, 2010, through March 30, 2011. A total of 18,495 completed questionnaires were collected.

Survey Design and Administration

Survey design and administration consisted of all tasks leading up to full-scale data collection, plus the actual data collection itself.

SEMCOG, FTA, NuStats, and CDM Smith had significant discussions about how to develop a sampling plan that would ensure the collection of adequate numbers of samples for analysis. A total of 18,000 samples were budgeted for the survey, and these samples had to be allocated among the region's transit systems. Under the guidance of FTA, a simple sample rate of eight percent was adopted instead of a traditional, statistics-based method for calculating sample rate. This eight-percent rate was further adjusted for each system based on variation of trip characteristics and each system's geographic spread.

Development of the initial survey questionnaire was based on SEMCOG's 2001 on-board survey, the contractor's knowledge and experience with other surveys, current regional conditions, and SEMCOG's modeling needs. Subsequently, the questionnaire was refined and finalized using findings from the pilot study and cognitive interviews. The questionnaire was also reviewed by FTA.

A pilot study was conducted to test the effectiveness of survey administration (e.g., survey form design, the distribution and collection process, response rates, etc.). The pilot was a full dress rehearsal intended to mimic the full-scale data collection and processing to follow. Two data collection modes were tested—a self-administered paper questionnaire and web-based data collection tool. The test results showed that the self-administered option was more productive, and thus was the primary method used for the survey, although the Web-based survey was still made available to riders. The pilot study results were also used to develop and finalize the data quality assurance and control (QA/QC) plan.

Following the pilot study, cognitive interviews were conducted with DDOT riders recruited at the Rosa Parks Transit Center to gauge their thoughts and motivations for participating in a survey of this type. DDOT riders were selected because they have historically lower response rates than riders from other systems, and because the DDOT system carries more than half of the region's transit ridership. These interviews were used to refine the marketing of the full-scale study and the questionnaire layout.

Prior to the full data collection, SEMCOG launched a public outreach effort via its website. A link to Frequently Asked Questions (FAQs) was placed on SEMCOG's website, addressing questions likely to be asked by the riding public. In addition, each transit agency provided support in promoting the survey and encouraged riders to participate. For example, posters were placed on buses and at the Rosa Parks Transit Center. Primary data collection occurred from late September 2010 through early December 2010. In March 2011, a second collection effort was undertaken on DDOT routes that had less desirable survey return rates during the first phase. Free ride tickets were distributed as an incentive in the second phase and the survey return rate was dramatically improved.

Quality Control and Data Processing

Quality control and data processing tasks included the entire QA/QC process, as well as tasks related to sample weighting and expansion.

The QA/QC process was an intensive effort performed jointly by NuStats and SEMCOG through all phases of the survey. Records were thoroughly examined for validity, with checks executed to search for and correct a large number of potential logical, data entry, geocoding, duplication, and other errors. Minimizing data errors and increasing accuracy helped maximize the number of usable survey samples that were collected. The survey records were scanned and verified by NuStats using specially designed software. In addition, SEMCOG performed an in-depth review of the data based on knowledge of the region and understanding of local travel.

Statistically sound methods for weighting and expanding the sample data were necessary to account for biases in the survey and to ensure that the data accurately reflected the transit riding population when used for analysis. Traditionally, the weighting and expansion process has involved calculating control totals based on the number of people boarding at stop locations (among other factors). However, the past few years have seen the emergence of a more sophisticated procedure that takes into account both boarding and alighting activity at transit stops. This emerging approach was strongly recommended by FTA for this on-board survey and was subsequently adopted by SEMCOG. At the heart of the adopted approach is a process known as iterative proportional fitting (IPF), which was used to calculate initial sample weights.

Due to transit system characteristics in Southeast Michigan, for the IPF process to work properly for this survey, stop-level IPF input data needed to be aggregated. There were more than 14,000 stop locations where boarding and alighting activity was observed during the survey, and about 9,000 of these locations had survey samples associated with them, spreading the available data very thin. To compensate, SEMCOG and NuStats developed and implemented a logical and practical stop aggregation approach that accommodated the IPF process and preserved observed travel patterns. A key feature of this aggregation approach was the consideration of land-use characteristics surrounding transit stops.

Along with the IPF weighting factor, the final expansion weight was also a function of factors that accounted for rider non-response on sampled trips and for trips that were not sampled. In addition, the final expansion weight took into account expected ridership calculated at both the route level and time-of-day level. The survey team believes that this unique approach maximized data efficiency by considering not only spatial travel patterns and land-use characteristics, but also the temporal distribution of trips.

Survey Results

NuStats created sets of statistics at both the regional level and the individual transit system level. These statistics focused on passengers' attitude towards the transit services, transit traveler's demographics, transit travel patterns, trip purposes, and service coverage and quality.

Statistics are also provided for certain corridors and in comparison to other regions across the country. Since the region is considering higher level rapid transit service to complement current bus service in major road corridors, additional survey statistics was developed for the Woodward, Gratiot and Michigan Avenue corridors. A set of general survey statistics was also created to compare Southeast Michigan with other major metropolitan areas where NuStats has conducted on-board surveys.

The region's bus system serves about 222,000 boardings every day, with about half of transit usage occurring on 10% of the routes in the system. There were 14,000 active bus stops surveyed, and 800 of them (6%) carried 50% of the daily regional ridership. The majority of riders (52%) reported not making any transfers to complete their trips, while 36% made one transfer.

The majority of transit riders in the SEMCOG region are transit dependent. Nearly 52% of riders do not have access to a vehicle on the survey day, and this number increased to 60% in the DDOT service area, which accounts for more than half of regional daily ridership. Another 46% of riders in the SEMCOG region do not have a valid driver's license, and furthermore, 20% of riders surveyed in the region were unemployed.

The SEMCOG region's transit systems primarily serve people with lower incomes. While each system varies, the survey found that on average 86% of riders were from households with an annual income of $50,000 or less, and that 40% of riders were from households making less than $10,000 annually. This ranges from 17% to 53%, depending upon the system. Since the majority of the region's transit riders are captive riders, improving existing service is a primary concern. Nearly 40% of those surveyed would like service to be more frequent; one-third would like extended service hours (earlier start/later end). Regional transit service is essential to riders, as 26% of the riders surveyed would not be able to make the trip if the service were not available.

For travel characteristics, the vast majority of transit trips made by riders either begin or end at home (84%), and 54% of riders used transit for work/university-related purposes. To access buses, 84% of respondents walked to/from the stop. Finally, 75% of riders were frequent riders (3-7 days per week).

Summary and Recommendations

Two areas of improvement are recommended for future on-board surveys. The first recommendation is that, if resources allow, a system-wide boarding/alighting count survey should be conducted prior to the next on-board survey. The collection effort should cover all routes, or at least major routes, by stop, direction, and time of day. This would allow for a more detailed understanding of passenger travel patterns, and could be used to develop a sample plan based on a more disaggregate approach, with sample collection goals established and weights calculated for cells stratified by system, route, direction, time of day, bus-stop "on" segment, and bus-stop "off" segment.

In recent years, FTA has emphasized using more accurate weighting and expansion methodologies, so that the results of origin-destination transit surveys more closely reflect actual travel characteristics both spatially and temporally. A stop-level count survey is one such solution promoted by FTA. While such system-wide boarding and alighting counts were not available as control totals for this survey, SEMCOG and the rest of the survey team worked hard to minimize the impact, making good use of the limited ridership data that was available and supplementing it with boarding and alighting activity observed during the survey.

The second recommendation relates to the way that on-board survey data is collected, given the known issues and biases with self-administered questionnaires. While this type of survey has been the industry standard for the past two decades, the trend, based on FTA direction and client/consultant data needs, is to move to an intercept interview (similar to what was conducted on the Detroit People Mover for this survey). Such interviews are administered by a surveyor on a transit vehicle who, in a random fashion, asks riders to participate in the survey. In addition, and most importantly, intercept interviews allow for an increased response rate and a minimization of potential biases introduced into the dataset.

In summary, the SEMCOG on-board transit survey provides valuable information for transit planning, policy decision making, and resource allocation. The survey team developed various innovative methods in survey design, sample expansion, and data quality control, all leading to a more robust and much improved dataset, both spatially and temporally, compared to previous regional surveys.

1. Introduction

Accurate and valid transit usage forecasts grounded on regional travel demand models are extremely important for local planning and investment, and are required for the increasingly stringent funding process of the Federal Transit Administration (FTA). To support the demand models' data requirements, up-to-date on-board transit surveys that are fully compliant with FTA guidelines are needed. Therefore, the Southeast Michigan Council of Governments (SEMCOG), working with transit research consultant NuStats, conducted a regional on-board survey for the riders on line-haul fixed bus routes operated by the Detroit Department of Transportation (DDOT), Suburban Mobility Authority for Regional Transportation (SMART), Ann Arbor Transportation Authority (AATA), University of Michigan Transit Service, Blue Water Area Transportation Commission (BWATC), Lake Erie Transit Commission (LETC), and the Detroit People Mover (DPM).

Since the previous O/D survey conducted in 2001, service changes and significant economic impact has reshaped much of the region. In addition, with new regional plans for expanded rapid transit services, FTA New Starts/Small Starts grant requirements are a focus. The purpose for conducting this large study is to update the SEMCOG Travel Demand Forecast Model (TDFM) and to enhance the transit and mode choice component based on the previously noted changes. The data collected should provide rich information on the current transit rider travel pattern, demographic information and transit service characteristics. The information is extremely valuable to the region's transit planning and transportation investment decision making process.

SEMCOG defined a set of criteria for a successful survey that includes the following:

- Proper coverage and representativeness across the full universe of transit users and all regional transit service providers;
- Sampling plan and data collection methodology focusing on trip purposes and transit access/egress mode;
- Completeness of detail in the trip O/D records collected, including accurate geocoding;
- Establishing baseline information for boarding/alighting and transfer rates; and
- Comprehensive and transparent documentation of all methods, procedures, and outcomes in the survey.

First, federal air quality standards have tightened, likely placing Southeast Michigan in a non-attainment category for ozone. In addition, the regulatory environment for air quality standards may soon include greenhouse gas emissions, placing further demands on metropolitan regions to slow or reduce transportation sector contributions to airborne pollutants, including carbon dioxide. Expanded transit service will be part of the mix of efforts to achieve and sustain air quality compliance.

Second, Southeast Michigan's economy for many decades was characterized by a strong automobile manufacturing sector. This portion of the region's economy has undergone dramatic structural shifts that necessitate the pursuit of a more diversified economy. With travel patterns changing, it is important to establish a new baseline of transit usage for our current system.

Third, SEMCOG's adopted regional transit plan calls for expansion of rapid transit service in both urban and suburban areas. Currently, with no existing rapid transit lines in operation, line-haul bus usage is an essential piece of estimating the transit market, especially for suburban areas, and planning the most appropriate mode for phased implementation of improved regional transit service. This study provides a database of all operational transit routes, origin and destination trip patterns, transit system utilization characteristics, and rider characteristics data.

SEMCOG and transit providers will use these data to characterize and predict travel patterns of customers traveling on transit systems in Southeast Michigan. The collected data will also be essential for the enhancement of the mode choice component of SEMCOG's TDFM and for producing model output that follows the recommendations of federal funding programs, specifically the Federal Transit Administration's New Starts/Small Starts grant program. Anticipated applications of these survey data include:

- Enhancement of the transit and mode-choice components of the SEMCOG Regional TDFM,
- Compliance with the travel model recommendations and guidelines for FTA New Starts/Small Starts applications,
- Identification of current levels of service,
- Establishing baseline information for boardings/alightings and transfer rates, and
- Identification of ridership patterns on local and express services.

The O/D survey was conducted among riders of fixed route bus services for all SEMCOG systems using self-administered questionnaires (or intercept interviews for DPM). Data collection was conducted on weekdays (Monday through Friday) from September 22 through December 3, 2010 and March 21 through 30, 2011. A total of 18,495 usable questionnaires, as included in the final data files, were collected for the O/D study out of a total of 114,901 eligible boardings. Eligible boardings include all passengers, aged 16 and older, on all surveyed trips where a completed questionnaire was collected (eligible boarding are described as those riders who are aged 16 and older by visual estimation), which is a response rate of 16.1 percent.

This report summarizes the survey methods and findings. Chapter 2 provides a description of the sampling approach, survey instrument and procedures, project challenges and solutions, and weighting and expansion methodology. Chapter 3 provides detailed information for the variables collected during the O/D study and summarizes the data at the regional, corridor and comparative levels. Chapter 4 summarizes the major findings and recommendations based on the analysis. Appendix A contains the SEMCOG Pilot Study Analysis. Appendix B contains the SEMCOG Cognitive Interviews Report. Appendix C contains the SEMCOG Pilot Findings and Full-Scale Data Collection Preparations. Appendix D contains the SEMCOG FAQs for the public outreach. Appendix E contains the SEMCOG questionnaires. Appendix F contains the Weighting and Expansion Memo. Appendix G contains the SEMCOG Assignment Report.

2. Survey Methods

2.1 Sampling Plan

In order to account for all the various systems and their ridership in the SEMCOG region, the sample plan was developed prior to the data collection with collaboration between SEMCOG, NuStats (and also subcontractor CDM Smith), and FTA for the most appropriate sample distribution.

The proposed sample plan was based on three main factors:

- First, the plan ensured that the sample adequately met data needs at the regional level.
- Second, the plan ensured the collection of adequate samples at the various times of day. Times of day (TOD) are defined by the individual transit agencies according to their needs. For the Detroit People Mover (DPM), this is defined as Peak for trips that start between 11:00 and 14:00, and Non-Peak for other trips. For systems other than DPM, the time periods are defined as AM Peak, Mid-day, PM Peak, and Evening/Early Morning time periods. Each system has different specifications for time of day definitions and is contained for each system in Appendix F, Weighting and Expansion.
- Third, the plan ensured that SEMCOG staff would have the ability to segment the sample on key variables, such as route, day of the week, time of day, and direction.

The population ridership figures were gathered by each agency from periods meant to best approximate the expected ridership to be encountered during the field data collection. Based on previous discussions with FTA regarding best current practices, NuStats suggested a 10 percent sample proportional to population ridership as a starting point in the sample design. This 10 percent figure needed to be balanced against project resources, which allowed for 18,000 sample pieces to be collected during the data collection. Therefore, concessions were made for systems, as well as for routes within systems. The population ridership figures and base 10 percent sample figures are contained in Table 2.1.

Table 2.1: Year 2010 Average Ridership by System and 10% Sample Allocation

System	Average Weekday Ridership	Samples Goal at 10%
DDOT	124,514	12,500
SMART	34,301	3,400
AATA	22,010	2,200
DPM	4,011	400
BWATC	2,491	250
LETC	877	100
UM	34,501	3,450
Total	222,705	22,300

Based on the needs of the study, systems and routes within the systems were scrutinized for similar travel characteristics to determine which systems/routes were candidates for sample size reduction and increase.

2.1.1 DDOT and SMART Overview

The two main systems of the Detroit area, DDOT and SMART, exhibited characteristics that make them prone to sample size reduction and increase, respectively. Meetings were held with each agency to better describe each route and route type to get a better understanding of the trip characteristics. DDOT serves only the city of Detroit and therefore has less variety in trip types than SMART, which serves a larger square mileage area encompassing more varied cities and jurisdictions. Because of this, the DDOT ridership could be accurately described with a sample size smaller than 10 percent, while SMART would need a proportionately larger sample. In the previously conducted 2001 SEMCOG OB study, the sample was drawn at a higher percent for SMART than for DDOT for this same reason. (Note, a different sample design was used for this study using confidence intervals sample allocation rather than a proportional sample draw.) The sample design and actual records collected from the 2001 study are presented in Table 2.2.

Table 2.2: 2001 SEMCOG On-Board Sample Goals and Samples Collected

System	Average Daily Ridership	Sample Goal	Sample Actual (O/D Pair)	% Sample Goal Actual
DDOT	158,215	6,448	5,624	3.6%
SMART	31,749	8,796	4,996	15.7%
AATA	15,229	2,249	1,922	12.6%
BWATC	1,421	683	107	7.5%
LETC	981	675	80	8.2%
Total	207,595	18,851	12,729	6.1%

Entering into this survey effort and based on the trip characteristics of these two systems, it was decided to set the sample goal of DDOT to 7.5 percent and set the sample goal of SMART to 13 percent. These two sample rates were determined as a trade-off between perceived sample needs and resources available with a higher percentage of riders required for SMART as compared to DDOT.

2.1.1.1 DDOT

A 7.5 percent sample produced a sample goal of 9,339, which is significantly higher than the 5,624 samples collected during the 2001 study. DDOT classifies routes by two categories, direction and size. For direction, there are three types of routes:

- All direction routes that serve downtown
- East/West routes that do not serve downtown
- North/South routes that do not serve downtown

For size they are classified as:

- Small – ridership less than 1,000
- Medium – ridership from 1,000–3,000
- Large – ridership greater than 3,000

NuStats implemented a 7.5 percent sample goal for DDOT routes, as shown in Table 2.3.

Table 2.3: DDOT 7.5 Percent Route Sample Goals

Route	Name	Average Daily Weekday Ridership	7.5% Ridership Sample
7	Cadillac-Harper	2,739	205
8	Caniff	489	37
9	Chalmers	1,478	111
10	Chene	1,778	133
11	Clairmount	1,208	91
12	Conant	932	70
13	Conner	1,392	104
14	Cross-town	6,330	475
15	Chicago-Davison	2,582	194
16	Dexter	10,040	753
17	Eight Mile	5,306	398
18	Fenkell	4,143	311
19	Fort	1,498	112
21	Grand River	8,939	670
22	Greenfield	5,226	392
23	Hamilton	2,945	221
25	Jefferson	3,112	233
27	Joy	2,932	220
29	Linwood	2,089	157
30	Livernois	1,962	147
31	Mack	3,971	298
32	McNichols	5,492	412
34	Gratiot	6,461	485
36	Oakland	456	34
37	Michigan	1,250	94
38	Plymouth	1,886	141
39	Puritan	913	68
40	Russell	645	48
41	Schaefer	1,933	145
43	Schoolcraft	1,093	82
45	Seven Mile	7,944	596
46	Southfield	1,379	103
47	Tireman	1,414	106

Route	Name	Average Daily Weekday Ridership	7.5% Ridership Sample
48	Van Dyke-Lafayette	4,208	316
49	Vernor	1,172	88
53	Woodward	12,466	935
54	Wyoming	1,271	95
60	Evergreen	2,866	215
76	Hayes Limited	176	13
78	Imperial Limited	398	30
Total for DDOT		124,514	9,339

2.1.1.2 SMART

A 13 percent sample rate produced a sample goal of 4,419, which is roughly in line with the 2001 study. In general, SMART categorizes their routes into four different route types:

1) Main Corridor – Serves major arterials and Detroit's central business district (CBD)
2) Community – Circulates within a community but does not cross jurisdictions, connect to Main Corridor routes
3) Cross-town – Like Community, but crosses jurisdictions
4) Commuter – Serves both formal and informal park and ride lots, express and limited stops to CBD

NuStats implemented a 13 percent sample goal for SMART routes, as shown in Table 2.4.

Table 2.4: SMART 13 Percent Route Sample Goals

Route	Name	Average Daily Weekday Ridership	13% Ridership Sample
125	Fort Street- Eureka Road	1,871	243
135	Southshore Express	35	5
140	Southshore	314	41
145	Carlysle	63	8
150	Taylor- Detroit	56	7
160	Downriver	189	25
190	Taylor Flyer	11	1
200	Michigan Avenue	2,531	329
202	Romulus	37	5
245	Cherry Hill	217	28
250	Ford Road	307	40
255	Ford Road Express	225	29
265	Warren Road	262	34
275	Telegraph	1,234	160
280	Middlebelt South	277	36
305-330	Grand River Beech Daily	385	50
385	Orchard Lake	97	13
400	Southfield- Orchard Ridge	288	37
405	Northwestern Highway	513	67
415-420	Greenfield- Southfield	1,304	170
430	Main Street- Big Beaver	97	13
445-475	Maple/Telegraph-Troy Limited	287	37
450-460	Pontiac- Somerset	4,259	554
465	Northfield Hills/Auburn Hills Limited	242	31
494	Dequindre	562	73
495	John R	1,791	233
510-515	Van Dyke	2,599	338
525	Groesbeck Helper	15	2
530	Schoenherr	143	19
550	Garfield	307	40
559	Auburn Hills	24	3
560-565	Gratiot	5,512	717
580	Harper	114	15

Route	Name	Average Daily Weekday Ridership	13% Ridership Sample
610-615	Kercheval	1,046	136
620	Charlevoix	89	12
635	Jefferson Express	102	13
710	Nine Mile	1,846	240
730	Ten Mile	636	83
740	Twelve Mile	1,384	180
752	North Hill Farms	201	26
753	Baldwin Road	256	33
756	Perry- Opdyke	204	27
760	Thirteen Mile- Fourteen Mile	401	52
780	Fifteen Mile	786	102
805	Grand River Park and Ride	297	39
830	Downriver Park and Ride	268	35
851	OCC- Northland Park and Ride	306	40
Total for SMART		33,990	4,419

2.1.2 University of Michigan (U-M)

The system displaying the greatest level of similarity for travel characteristics is U-M, making it the most appropriate system (and routes within a system) on which to reduce the sample size. Through discussions with UM, each route was characterized based on the trip types made. Routes were classified into three groups as follows:

1) Housing area to campus with trips made primarily by students,
2) Campus to campus with trips made primarily by students (also faculty and staff), and
3) Park and ride lots to campus/hospital trips made using drive access made by students, staff, medical school students, and patients.

For classifications 1 and 2, a lower percentage of trips were recommended based on trips overwhelmingly beginning and ending on campus with common trip types. A higher percentage was recommended for 3, as compared to 1 and 2 because of the drive access/egress at the park and ride lots, which allowed trips to either begin or end off of campus, thus showing a larger variation in trip type (note that these are still relatively homogeneous trips when compared to "standard" transit systems).

Using these classifications, NuStats suggested a 2 percent sample for classifications 1 and 2 (intra-campus trips) and a 5 percent sample for classification 3 (trips where either the origin or destination is off campus), as shown in table 2.5, below.

Table 2.5: U-M 5 Percent and 2 Percent Route Sample Goals

Route	Classification	Average Daily Weekday Ridership	2% Route Sample Goal (Non-drive), 5% (Drive)
Bursley-Baits	1	13,484	270
Northwood Express	1	702	14
North Campus	1	201	4
Oxford Shuttle	1	512	10
Oxford	1	1,021	20
Northwood	2	4,961	99
Diag to Diag	2	1,833	37
Commuter	3	9,450	473
Mitchell Glazier	3	1,547	77
Mitchell Express	3	274	14
Inter Campus	3	516	26
Total for U-M		34,501	1,044

2.1.3 Other Transit Providers

It was determined to hold the sample sizes for DPM, BWATC, LETC, and AATA at the 10 percent level. The route by route goals for each of these systems are provided in the following tables. Table 2.6 shows the route level sample sizes for AATA.

Table 2.6: AATA 10 Percent Route Sample Goals

Route	Name	Average Daily Weekday Ridership	10% Ridership Sample
1	Pontiac - Dhu Varren	810	81
2	Plymouth	2,463	246
3	Huron River	1,337	134
4	Washtenaw	3,052	305
5	Packard	2,319	232
6	Ellsworth	2,248	225
7	S. Main- East	1,241	124
8	Pauline	746	75
9	Jackson- Dexter	708	71
10	Ypsilanti- Northeast	506	51
11	Ypsilanti- South	243	24
12	Miller- Liberty	834	83
13	Newport	211	21
14	Geddes- E. Stadium	196	20
15	Scio Church- W. Stadium	253	25
16	Ann Arbor- Saline Rd	468	47
17	Amtrak- Depot St	46	5
18	Miller- University	435	44
20	Ypsilanti: Grove- Ecorse	544	54
22	North/South Connector	772	77
33	EMU Shuttle	611	61
36	Wolverine Tower Shuttle	1,697	170
601	Pontiac- University	128	13
609	Dexter -University	142	14
Total for AATA		22,010	2,201

The route level sample sizes for BWATC are presented in Table 2.7, while the route level sample sizes for LET are presented in Table 2.8.

Table 2.7: BWATC 10 Percent Route Sample Goals

Route	Average Daily Weekday Ridership	10% Ridership Sample
Route 1	369	37
Route 2	300	30
Route 3	402	40
Route 4	151	15
Route 5	369	37
Route 6	441	44
Route 9	383	38
94 Express	9	1
M29 S / 94 Express N	34	3
94 Express S / M29 N	33	3
Total for BWATC	2,491	249

Table 2.8: LET 10 Percent Route Sample Goals

Route	Name	Average Daily Weekday Ridership	10% Ridership Sample
2	Elm St	93	9
3	Southeast	84	8
4	Seventh St	106	11
5	Telegraph	131	13
6	McComb	103	10
7	S. Monroe	87	9
8	N. Monroe	173	17
9	S. Custer	100	10
Total for LET		877	88

For DPM, an alternative sample allocation and data collection method were designed based on the results of the pilot study. Rather than using a bus as the sampling unit for a route, the individual stations were used to produce station level goals. Surveyors conducted intercept interviews,[1] with passengers on the DPM, and these interviews were allocated to the individual stations where the passengers boarded the vehicles. The station level goals are provided in Table 2.9.

Table 2.9: DPM 10 Percent Station Sample Goals

Station	Average Daily Weekday Boardings*	10% Ridership Sample
Times Square	392	39
Michigan Ave.	279	28
Fort / Cass	150	15
Cobo Center	73	7
Joe Louis Arena	109	11
Financial District	182	18
Millender	202	20
Renaissance Center	630	63
Bricktown	125	13
Greektown	898	90
Cadillac	693	69
Broadway	115	12
Grand Circus	163	16
Total	4,011	401

**Special event days, including baseball and hockey game days as well as other special events, were removed from the average daily weekday ridership.*

[1] In an intercept interview, the surveyor asks the rider questions rather than having the rider self-complete the questionnaire.

2.1.4 Major Corridors in the Detroit Area

Knowing that there will be upcoming improvements in a few corridors in the Detroit area, special consideration was shown for some of the most prominent corridors. Using the system percentages previously mentioned for SMART (13 percent) and DDOT (7.5 percent), each corridor would have the total number of samples collected as listed in Table 2.10.

Table 2.10: Corridor Sampling Allocation for SMART and DDOT

Corridor	SMART Routes	SMART Samples (13%)	DDOT Routes	DDOT Samples (7.5%)	Total Samples
Michigan Ave.	200, 250, 255, 280	434	37	94	528
Gratiot Ave.	510, 515, 530, 560, 565, 580	1,088	34, 76	498	1,586
Woodward Ave.	415/420, 445/475, 450/460, 465, 495	1,025	53	935	1,960
M-59	510, 559	341			341

2.1.5 Sample Summary for SEMCOG On-Board Survey

Aggregating each system's route ridership produces the following sample sizes as shown in Table 2.11. (Within each system, individual route sample sizes were be allocated in the same proportion format as was constructed for the system sizes.)

Table 2.11: Sample Allocation at the System Level

System	AWR 2010	Samples Sizes	Sample Percentage
DDOT	124,514	9,339	7.5%
SMART	34,301	4,419	13.0%
AATA	22,010	2,201	10.0%
DPM	4,011	401	10.0%
BWATC	2,491	249	10.0%
LET	877	88	10.0%
UM	34,501	1,044	2% intra-campus / 5% PnR
Total	222,705	17,741	8.0%

While goals were created at the route level, the sample plan also elaborated on the data collection methodologies—a self-administered questionnaire with a mail-back option or an intercept interview for the DPM, as previously described, specifically with regard to trip selection. All surveyed trips were selected at the block level,[2] therefore providing de facto stratification to ensure the distribution of sampled trips was representative of the population of trips with regards to time of day and direction distributions.

The survey used a standard two-stage sampling approach that consisted of sampling bus trips and then sampling passengers. Every passenger over the age of 16 (determined by visual estimation) who boarded the sampled bus received a questionnaire. If the surveyor was not able to determine whether a rider's age was over 16 by direct observation (which is the standard procedure), the surveyor asked the boarding passenger if they were over 16 years old.

For a questionnaire to be considered a complete, the following information must have been captured: geocodable origin and destination addresses, boarding and alighting locations, origin and destination purposes, access and egress modes, and a one-way trip route sequence. (Boarding and alighting locations were captured by PDA and imputation, respectively, though alighting location was asked explicitly well.)

2.2 Pilot Test

2.2.1 Background

A pilot test was conducted prior to the full-scale data collection. The pilot test was a full dress rehearsal of the data collection processes for the O/D study and tested multiple methods of data collection (self-administered questionnaire versus a card inviting passengers to complete the questionnaire online) and included the following activities: questionnaire design, assignment generation, conducting of assignments, data processing, and final data file submittal and scrutiny to prepare for the upcoming full-scale data collection.

NuStats conducted a pilot survey from April 19 to April 28, 2010. The purpose of the pilot was to gain insight on potential issues that could arise during the full-scale data collection and limit these occurrences. The pilot test was designed to see what works well and what does not. For those items that do not work well, the pilot data and anecdotal information from the pilot were used to make improvements upon the methods used for the full-scale data collection. In addition, the pilot allowed staff the opportunity to work with the individual transit providers to ensure that the full-scale data collection logistical issues were minimized and to familiarize themselves with transit facilities, people, and bus schedules.

During the pilot, routes from DDOT, SMART, DPM, AATA, and UM were surveyed. For the purposes of the pilot, BWATC and LET were not studied because of their relatively low ridership figures in comparison to the other systems.

For DDOT and SMART, NuStats tested a self-administered questionnaire. For these two systems, the focus was on the survey and system logistics. For the DPM, AATA, and UM, one focus for the pilot was survey and system logistics, but NuStats also tested a second data collection methodology—Web-based data collection—based on the nature of the system trips. The rationale behind using this second method of collection relates to the relatively short trip times for passengers,

---

[2] "Blocks" represent the series of trips the bus is going to make and usually mimics the driver shifts. By boarding a single block and continuing to survey, there is no need for surveyors to transfer between buses over the course of a surveyor assignment.

especially for DPM and UM. For the routes in these three systems, half of the surveyed trips distributed the self-administered questionnaire, and the other half distributed survey cards inviting passengers to take the survey online.

2.2.2 Results

The results of the pilot test were mixed. When looking at the different data collection methods, the hopes that using a survey card asking the passenger to complete the questionnaire online did not produce better response rates than the traditional self-administered method. Therefore, for the full-scale data collection, it was determined to issue the questionnaire primarily as a self-administered instrument but give respondents the option to complete the instrument online. In fact, out of over 18,000 questionnaires completed for the full-scale data collection, less than 100 were completed online.

Overall, the response rates from the DDOT routes were lower than anticipated. Therefore, the cognitive interviews, which were scheduled to take place following the pilot test (discussed in the next section), focused only on DDOT riders. Furthermore, the public outreach effort through SEMCOG focused on DDOT passengers.

For SMART, AATA, and UM services, the pilot test performed well. Therefore, it was determined for the full-scale data collection that no substantial changes were needed to the instrument or method employed. The one minor change involved combining the self-administered/online option into the same questionnaire allowing a respondent to participate by either method.

For DPM, it was determined that self-administered questionnaires took too long to complete on the short trips; therefore, intercept interviews were conducted. The full questionnaire, or at least the vast majority of the questionnaire, could be completed from the time the passenger entered a station until they left the system at their alighting location.

A full description of the pilot test, including the questionnaire and results, is contained in Appendix A: SEMCOG Pilot Study Analysis. A couple of tables presented in the appendix are shown in Figures 2.1 and 2.2.

Figure 2.1: Web VS. Paper Option Response Rates

System	Web		Paper		Total	
	Complete	Response Rate	Complete	Response Rate	Complete	Response Rate
AATA	78	6.9%*	186	32.3%	264	15.5%
People Mover	32	3.8%	23	4.2%	55	4.0%
U of M	29	3.4%*	148	10.3%	174	7.9%
Total	139	5.1%	354	13.9%	493	9.3%

Figure 2.2: System-wide Response Rates

Transit System	Completed	Collected	Boardings	Response Rate
AATA	264	312	1,703	15.5%**
DDOT	226	324	1,324	17.1%
People Mover	55	77	1,378	4.0%
SMART	158	200	577	27.4%
U of M	174	202	2,200	7.9%*
Total	877	1,132*	7,182	12.2%

2.3 Cognitive Interviews

2.3.1 Background

Based on the experience from the pilot test, cognitive interviews were also conducted prior to the full-scale study in an attempt to optimize the full-scale data collection, especially with the DDOT passengers. NuStats conducted a series of in-person cognitive interviews with DDOT transit riders in advance of conducting the full-scale survey in the fall of 2010.

The goals of the cognitive testing were to:

1) Gather feedback on the survey instrument and suggestions for improving the instrument, thus improving data quality,
2) Learn what messages would be motivating and encouraging for participation,
3) Determine the appropriate media outlets for communication, and
4) Provide insight on the need for, preferred amounts of, and types of incentives.

Based on the goals above, a cognitive interview guide was developed to ask questions that would yield the needed information. For Goal 1, a section of questions assessed how easy or difficult it would be for respondents to complete a self-administered questionnaire asking about transit trips, exact addresses, and other information. Respondents completed the questionnaire while interviewers observed. Interviewers then went through the instrument, question by question, to assess areas of confusion or need for clarity, along with any suggestions for making the questionnaire easier to read and complete.

To accomplish Goal 2, interviewers asked questions about reasons for participation and ways to communicate the importance of the survey to riders. Questions about preferred media sources and the most useful methods for communicating the upcoming study to a broad range of riders helped to achieve Goal 3. For Goal 4, questions regarding how important an incentive was for encouraging participation were asked, as well as questions about whether people would be more likely to want bus passes, cash, or a big-ticket item (such as a flat screen TV or a computer).

2.3.2 Methods

NuStats recruited riders at the Rosa Parks Transit Center in downtown Detroit on June 7 and June 8, 2010. The recruiter approached a variety of riders in an attempt to gain cooperation from different types of transit users, taking into consideration gender, race, age, and perceived background. Each person was presented with an introductory script and was told about the incentive for participating in the cognitive research.

Recruitment for the interviews took place over a couple of days, throughout the day, prior to the interviews. NuStats slightly over-recruited the number of interviews needed to allow for those who did not show. On June 9 and 10, 2010, 23 interviews were conducted at the Renaissance Conference Center in downtown Detroit. Interviews lasted approximately 60 minutes, and each person received $75 for his or her time.

The basic structure for the interviews was to provide an introduction and goal of the interview, then present the survey instrument and ask the respondent to complete it. During the introduction, it was important to try to get the respondent to think back to his or her most recent and preferably "regular" bus trip, which for many respondents was the home-to-work commute. Previous testing has indicated that asking about their trip to the interview leads to confusion, as some people mix the two trips (their usual trip and the trip to the interview).

While testing of this nature introduces an artificial element—since participants are not on a bus, and they agreed to attend the interview knowing they would receive $75 as compensation for their time—the real value in conducting in-person interviews with actual transit riders, even under contrived conditions, is the ability to probe on the various issues encountered or points of confusion. More importantly, it is an opportunity to learn about ideas for how to improve the survey instrument from the perspective of the respondent, thus increasing participation rates as well as the quality of the data. Interviewers concentrated their efforts on gaining insight about motivations for taking the survey, messaging strategies that may help increase participation, and preferred media outreach venues. A relatively significant section of the interview concentrated on the incentive structure.

2.3.3 Findings

As expected, there were certain elements of the questionnaire that were problematic for participants, including not knowing the physical address of either the origin or destination, not understanding the origin and destination place name request, poor comprehension of the "one-way trip" concept, not understanding the example of one-way trip provided on the questionnaire, not knowing the route number (only the route name), and misinterpreting the prize drawing. All comments were considered when the full-scale data collection questionnaire was developed, such as adding the route name, in addition to route number, when trying to capture the route sequence for the one-way trip. In addition, the place name was placed before the address for both the origin and destination questions, and instructions were given to provide "Home" for the place name if that was the case.

Motivations and incentives for participation, as well as the appropriate media outlets, were also discussed with the participants. In general, these discussions produced expected results with a mix of responses where one participant contradicted another participant, i.e., some thought more incentives of a lower value were more appealing than fewer incentives of a higher value.

A full description of the cognitive interviews is contained in Appendix B: SEMCOG Cognitive Interviews Report.

2.4 Pilot Test Finding Meeting

Following the pilot test and the cognitive interviews, a pilot findings and full-scale data collection preparations meeting was held on August 25, 2010. NuStats and SEMCOG made a presentation at SEMCOG's Technical Advisory Committee monthly meeting. This meeting was designed to communicate study activities that had transpired up to that point and to lay out a plan for the remainder of the project. The main tool for these communications was a PowerPoint presentation, produced by SEMCOG and NuStats, highlighting the major findings, including descriptions of the requirements for a successful study, upcoming planned project schedule, the pilot analysis and findings (including the pilot test and the cognitive interviews), and recommendations for the full-scale data collection and public outreach.

When describing the pilot test, the two different data collection methodologies (self-administered and online data collection) were presented, along with the results of each at the overall and system level. Different analyses were conducted, including trip length both in terms of minutes and miles, to determine its impact when scrutinizing the two methods. The overall response rate was presented by system, along with the item response rates for both required and optional questions. Next, the cognitive interview process and outcomes were described. Finally, suggestions were made for the full-scale data collection based on the prior findings presented on a system-by-system basis.

The slides used for the presentation are contained in Appendix C: SEMCOG Pilot Findings and Full-Scale Data Collection Preparations.

2.5 Public Outreach

Using findings from the cognitive interviews, experience from previous SEMCOG survey efforts, and discussions between SEMCOG and NuStats, a public outreach plan was developed to alert the public about the survey and then to later update the progress of the study. The primary means through which this was accomplished was the SEMCOG website. From the website, those interested were able to click on a static page that provide frequently asked questions (FAQs) to find out additional information about the upcoming study. As data collection progressed, updates were posted on the website to keep interested parties abreast of the efforts being made.

Another way the study was advertised was through posting at the Rosa Parks Transit Center. With the majority of DDOT routes converging on the transit center, this location was viewed as the optimal location to advertise the upcoming and then current study. Individual postings also occurred on individual buses. The posting directed interested parties to the SEMCOG website for further explanation of the study.

The FAQs used for the public outreach is contained in Appendix D: SEMCOG FAQs

2.6 Survey Instrument

The survey instrument design was a multiple-stage process comprising a pilot and full-scale questionnaire design phase. During this process, SEMCOG and NuStats collaborated (with inputs from CDM Smith, the local transit agencies, and FTA), to design the survey instrument. Taking into consideration the feedback from the pilot test and the cognitive interviews, modifications were made to the pilot instrument in order to arrive at the final instrument that best presented the required questions using wording and question grouping that most appealed to the Southeast Michigan area ridership.

The final survey instrument was designed as a self-completion questionnaire with 18 self-coded questions. The set of data items is presented in Table 2.12. Prior to data collection, returned questionnaires were defined as “complete” and “usable” if the following questions were answered: origin address, destination address, mode of access, mode of egress, trip purposes, and trip path.

Questionnaires were designed in a two-sided double letter-size format and printed on heavy card stock for easy distribution and completion. Each questionnaire contained a business reply mail permit for off-bus completion and mail-back. The form was pre-printed with a unique serial number and barcode, which linked each questionnaire to a specific trip and bus stop boarding location. Text on the questionnaire invited passengers to register to win a monetary prize, one of 10 $100 prizes, by providing their name and telephone number. The questionnaire was designed to obtain information in three major categories: O/D travel patterns, access and egress modes, and rider demographics. As noted in Table 2.12, some of the required data elements were captured by means other than a question on the questionnaire. This approach had multiple benefits: (1) the questionnaire was shorter to enhance response rates, and (2) data quality was improved by circumventing respondent-provided information. The questionnaire was available in English and Spanish.

Appendix E contains the English survey instruments for both the pilot and full-scale data collection efforts.

Table 2.12: Data Elements and Capture Method

Data Elements	Capture Method
Day of Travel	GPS-enhanced Palm device
Time of Travel	GPS-enhanced Palm device
Route	GPS-enhanced Palm device
Direction	GPS-enhanced Palm device
Questionnaire Language	Field Code by editor
Origin Address	Questionnaire
Origin Trip Purpose	Questionnaire
Destination Address	Questionnaire
Destination Trip Purpose	Questionnaire
Access Mode	Questionnaire
Egress Mode	Questionnaire
Total Buses and Trains	Questionnaire
Trip Path	Questionnaire
Alighting Location	Questionnaire
Bus Stop On	GPS-enhanced Palm device
Bus Stop Off	Imputed using information from other sources: Destination, Egress Mode, Distance, and GPS data on bus stops for the sampled trip
Service Improvements	Questionnaire
Fare Type	Questionnaire
Fare Level	Questionnaire
Fare Subsidy	Questionnaire
Trip Frequency	Questionnaire
Alternative Mode	Questionnaire
Valid Driver's License	Questionnaire
Passenger Age	Questionnaire
Employment/Student Status	Questionnaire
Household Working Vehicles	Questionnaire
Household Size	Questionnaire
Household Workers	Questionnaire
Household Income	Questionnaire

2.7 Survey Administration

2.7.1 Labor recruitment and Training

NuStats subcontracted the survey labor (i.e., surveyors and counters) to two local employment agencies, Yoda Technologies and Scott Group. Employment criteria included the demonstration of: current or past residence in the service area, good work habits, people skills, honesty, maturity, possession of reliable personal transportation, and attention to details.

Training was conducted for the Detroit area collection (defined as DDOT, SMART, and DPM) on September the 20 and 21, 2010. A total of 48 surveyors where trained for the Detroit area collection. Full-scale data collection began on September 22 and ended on December 3 2010. Ann Arbor collection (defined as AATA and UM) began on October the 28, 2010 with training on October 27. A total of 22 surveyors where trained for Ann Arbor collection. Ann Arbor collection ended on November 19, 2010, with all initial goals met for the Ann Arbor Transportation Authority and all but one route reached goal for University of Michigan Transportation.

The training included a background of the survey project, an overview of the seven transit systems, safety and security training, and survey instruction, including one hour of role-playing and intensive tutoring. Surveyors received specific training in reading and comprehension of surveyor assignment sheets, basic survey procedures and etiquette, and survey subject approach techniques. Counter specific instruction included training in the use of the hand-held Palm computer devices, the ride count software program, counting techniques for the boarding and alighting passengers, as well as general on-board vehicle etiquette.

Following the completion of the initial data collection assignments, NuStats required the survey teams to return to the survey command center located at Rosa Parks Terminal (or location in Ann Arbor). Supervisors verified the accuracy of each survey team's work. Survey command center staff provided coaching and additional training when deemed necessary. Staff then distributed survey assignments for the next day.

2.7.2 Survey Administration

The full-scale survey was managed by an in-field survey team comprising 1) a field manager to oversee the entire field team, 2) a surveyor assistant to manage surveyors, and 3) a counter assistant to manage the counters and provide ridership count quality assurance for uploads/downloads to the Web-based field management system.

On-board data collection was conducted by teams comprising a surveyor and a counter. The surveyor handed out questionnaires, persuaded passengers to complete the questionnaires, assisted with questions, collected questionnaires, and scrutinized the returned questionnaires. The counter entered the questionnaire numbers into the Palm device to link questionnaires to a bus stop, counted the passengers boarding and alighting, ensured the unit had picked up accurate GPS location coordinates, collected questionnaires, and validated passenger loads after each stop. Daily surveyor assignments were distributed by the field manager or by the assistants. (For DPM, only surveyors were used.)

As assignments were handed out, information was updated in the Web-based field management system. When surveyors and counters returned from an assignment, the field manager or assistant checked the assignment results (i.e., quickly reviewed the questionnaires to spot any glaring performance issues) and downloaded the passenger count data from the Palm devices. Feedback and additional training were provided when errors were found in the data. If certain errors persisted, staff would be relieved of their services.

The field manager updated the assignment status in the Web-based field management system and then handed out the next assignment. Once the completed assignments were reviewed, the questionnaires went through the in-field editing process for inspection and coding prior to being sent to Austin, the location of NuStats' headquarters, for scanning and verification.

2.7.3 Surveyor Assignments

NuStats developed surveyor assignments by uploading the trip selection requirements to a Web-based field management system. Field managers printed the surveyor assignment sheets from the Web-based management system for the surveyor teams and included the assignment sheets with directions to/from the assignment starting/ending point. The assignment sheets additionally contained a barcode to link the assignment back to the field management system.

Surveyors distributed survey instruments to all boarding passengers over the age of 16, and counters tallied the number of passengers boarding and alighting the vehicle. The counters used a GPS-enhanced Palm device, pictured in the Figure 2.3.

Figure 2.3: GPS-Enhanced Palm Device for On-Board Counts

The GPS-enhanced Palm device recorded the location and time (arrival and departure) at each stop, while counters entered the number of passengers boarding and alighting. Counters also entered the number at the top of the surveyor instrument bundle prior to arrival at each stop. This process linked the sequence and range of survey instruments directly to a stop using the provided transit systems' digitized stop list file. The field manager uploaded the count data files to the Web-based field management system.

2.7.4 In-Field Survey Instrument Editing and Scanning

Following the surveyor check-in, completed questionnaires were presented to on-site data editors for editing and correction. Data editors were local residents who were familiar with the geography of the transit service area. Data editors reviewed each completed questionnaire and used geographic resources to complete or correct address information. Because the origin and destination questions are the most difficult to collect, using these geographic resources to "clean" addresses provided a means to "save/salvage" as many questionnaires as possible. After each questionnaire had been reviewed, the barcodes were scanned on the questionnaire using a procedure that identified the questionnaire as a "complete." This information was uploaded to the field management system as one data input for the status reports. "Complete" questionnaires were sent to Austin for scanning and verification. Data editors were also employed to call back riders who turned in questionnaires that were less than complete. The phone number came from the questionnaire and allowed for more partially filled out questionnaires to be converted to completed questionnaires.

After reviewing each survey instrument, the data editors scanned the barcodes on the survey instrument to identify the record as complete. Scanning the records simultaneously uploaded the data to the field management system as one data input for Web-based management status reports. The field editor would then scan the complete surveys, and the images would be sent electronically to staff at NuStats for verification.

2.7.5 Status Reporting

The Web-based field management system allowed the field manager to review surveyor assignments, provide progress reports and data summary tables, and monitor field staff performance. The field manager prepared status reports from the Web-based field management system. This automated Web application also allowed the field manager to conduct consistency checks, flag problem records, and clean and purge flagged records. The field manager reviewed the information for accuracy in the status, response, and performance reports to the Web-based field management system.

2.8 Data Quality Assurance

2.8.1 Data Verification

Respondents were given two options for completing the questionnaire: a hard-copy questionnaire or a Web-based option. For the hard-copy questionnaires, NuStats used ScanTron scanning technology to assist in data entry and minimize human error resulting from manual data entry methods. The scanning process involved electronically scanning batches of approximately 20 survey instruments to produce an image file of the documents. After scanning, the data results derived from the image files were individually reviewed and verified by comparing the scanned image to the data contained in the data file. Text data (primarily origin and destination address information) were reviewed for the purpose of correcting misspellings and verifying that the scanner correctly read numeric data. The raw data file output from scanned documents was maintained unaltered for comparison purposes, if necessary. The other option was to complete the questionnaire using a Web-based program, although less than 100 participants chose to participate online.

A data items matrix and data dictionary were developed based on the survey instruments and scanning programs using the following process:

- The data items matrix identified variable names, variable descriptions, data types, field widths, code sets, skips, and exact question wording as it appeared in the survey instruments.
- The data dictionary was based on variables listed in the data items matrix. The data dictionary consisted of variable names, data types, field widths, variable labels, and response labels. The labels were abbreviated as necessary to accommodate database field width restrictions.
- The data dictionary was checked to ensure agreement with the hard-copy survey instrument.
- The data structure was checked to ensure consistency for all data files created for the study.

Following the duplication of the original database, the data contained in the database copy were checked for data integrity. Various edit routines were programmed to check the consistency of data and to identify reporting, scanning, or entry errors. Data in the control file were then matched against survey data to ensure that all information was consistent between the two files.

Routine edit checks were conducted to examine survey instrument responses for reasonableness and consistency across items. Routine checks included:

- The total number of records in the data file was checked to determine if the amount was equal to the total number of scanned survey instruments.
- If duplicate records were identified, all duplicated data were checked against the original record. If all data were not identical, data were flagged for review. Otherwise, duplicates were corrected or removed (duplicate unique identifier).
- Records with multiple responses per question were reviewed for plausibility.
- Records with comments outside of the scanning area were either incorporated into the appropriate variable or into the open comments.
- Ten percent of data entries were re-verified.
- After the creation of the verified database, the paper records were merged with any Web-based records and reviewed for completeness, coherence, and consistency.

Response Checks

- Checking for proper data skips and patterns of answering questions consistent with prior answers.
- Checking for realistic responses (e.g., number of household workers is equal to or less than the number of household members).
- Responses of home to home trips were researched for correct one-way location types.
- Checking for high frequency of item non-response (missing data).
- Skip patterns were verified to be completed correctly.
- Non-required responses that were not valid skips were filled coded as "9".

Range Checks

- All categorical values were verified within the expected range.
- Outliers in continuous variables (variables that represent a continuum of values and do not have a code set) were flagged and reviewed.

Open-Ends Preparation (non-categorical, text variables)

- Routes were converted from an open response into standard route codes.
- Text variables associated with an "other" type category were reviewed. Text responses that belonged to one of the categories in the response list/code set were flagged and re-coded.
- All text responses were corrected for any spelling or typographical errors.
- All responses not marked "other" and included an open response were reviewed.

2.8.2 Geocoding

After the questionnaire data was compiled, the location data were split out and geocoded. NuStats sub-contracted the geocoding task to GeoStats of Atlanta, Georgia. The geocoding task included reviewing, cleaning, and geocoding the location data collected in the survey instrument and recorded in the Palm device or imputed from the survey instrument. The survey location data consisted of four location types: trip origin, boarding location, alighting location, and trip destination. The trip origin, destination, and alighting questions were explicitly asked on the survey instrument, while the boarding and alighting location data were automatically collected and recorded via the Palm device technology or imputed, respectively.

Trip Origin and Trip Destination

Geocoding of respondent-provided trip origin and trip destination addresses consisted of two stages. First, an automated batch run was first attempted to successfully geocode origin and destination addresses. The batch run attempted to match exact addresses or cross-streets obtained from respondents to a street coverage file provided by SEMCOG. Addresses or cross-streets matching the coverage file were assigned an X/Y coordinate and a value of “M” for matched, and placed in the “AV_STATUS” field. Addresses or cross-streets not matched during the batch run were flagged with an “AV_STATUS” value of “U” for unmatched, and passed to the next stage of geocoding.

During the next stage, addresses were researched using a series of resources, including Switchboard.com, Google.com (Internet search engines), and DeLorme Street Atlas USA (mapping software). Addresses that were matched to an exact address or cross-streets during this stage were assigned an X/Y coordinate and an “AV_STATUS” of “M”. Those remaining unmatched addresses were not assigned an X/Y coordinate and were given the “AV_STATUS” of “U”. Because origin and destination are required elements, unmatched records were removed from the final data file.

Boarding Location Assignment

GeoStats developed a technique to assign the boarding location of survey passengers using both the boarding information collected with the iPAQ devices along with the transit system route database. The boarding location was obtained directly from the passenger count data file using the survey instrument number and the ranges captured at each boarding location. Depending on the availability of GPS, one of the following two paths determined the location:

- If a GPS record was available, then it was used to select the nearest stop in the current sequence of stops (as determined by route/direction/pattern).
- If the record did not have a GPS record, but the counter selected a stop from the list, then the counter selected value was used.

Alighting and Transfer Imputation

The alighting imputation calculated the location where the passenger most likely exited the vehicle. The alighting imputation procedure used the survey instrument variables in conjunction with the assigned boarding information to determine if the passenger was surveyed during the final leg of the trip or if a transfer was performed at the end of the surveyed trip. The alighting stop selected in the imputation process when a future transfer did not occur was the closest stop to the destination in the route/direction/pattern list after the boarding stop. For the Blue Water system, only time points were available, so some short trips were imputed to have the same alighting as the boarding.

If it was determined that a future transfer occurred, the following imputation logic was applied to determine the transfer location and, therefore, the alighting location:

- The set of possible stop locations that the passenger could transfer to/from based on the reported sequence of routes and the current route were identified in order to determine the transfer location.
- The transfer location was then selected using a half-mile buffer, which included the stops closest to the destination where the two routes cross.

Geocoding Quality Control

Once geocoded, the records were subjected to a series of strict quality control checks. The quality control checks included:

- Running the unmatched locations through the geocoding process for a final geocoding attempt.
- Randomly selecting five percent of the geocoded address file to review in detail to ensure proper placement of the overall latitude/longitude points. The review process entailed mapping the geocoded points in ArcView and comparing the points with DeLorme street file.
- All cross-street points were queried and analyzed to ensure proper placement of the points (since a cross-street geocode does not reference a zone for zip code or city in ArcView; and the default placement of a geocoded cross-street in ArcView places the point in the southeast quadrant of that intersection.)
- A visual quality control check was first performed on each route. This check reviewed the geocoding and verified the accuracy of the location by route, and additionally analyzed the boarding and alighting locations relative to the each route. The visual check was conducted by querying boarding/alighting points according to each route. For example, all of the boarding/alighting matches for Route 5 were selected and displayed in the map view in ArcView. A visual check was conducted to make sure that most of these points were displayed on or within proximity of the route. Points that were not displayed on or near the route identified a respondent error.
- A visual quality control check was then performed by city. The geocoding was verified by querying the geocoded matches related to each city. These points were then displayed in the map view in ArcView and visually confirmed, and outlying locations were selected and confirmed to be correct.
- Records with the same origin and destination location were researched or possibly removed from the database if irresolvable.

2.8.3 Route Sequence and Locations Verification

At times, survey respondents were confused with the one-way trip nature of the questionnaire, and recorded information pertaining to roundtrip locations or routes, alternate routes, reversed locations, or routes used for a different trip. If uncorrected, this could result in incorrect transfer rates for a trip or inaccurate route usage. The geocoded locations and sequence of routes along the trip were reviewed for plausibility through the TrueRoute program. The heart of TrueRoute is the "PuT Passenger Surveys" module of VISUM (PuT stands for Public Transport). This is an add-on to VISUM that has the capability to calculate plausibility scores for on-board survey records. These scores are based on the records' O/D geocodes and the captured sequence of routes. As part of its processing, TrueRoute uses the provided input data to identify line routes (directional routes) and vehicle journeys (trips) in the VISUM network that match the provided inputs as closely as possible, and that allow the passenger to complete the trip between the provided origin and destination locations.

Plausibility scores were computed for the surveyed leg, the preceding leg (if reported), and the succeeding leg (if reported), and then for the overall path based on a .5-mile walk tolerance and 10-minute wait time tolerance. All records with an implausible leg were reviewed visually for possible errors in either the locations or the route sequence. Although not all public transit riders follow a logical path, this verified that the origin and destination locations were in the correct order given the surveyed route and that only routes that could have been used in the one-way O/D trip were recorded in the data set.

2.8.4 Additional Quality Assurance from SEMCOG

SEMCOG performed its own series of quality assurance and quality control checks on the pilot survey records as well as each draft of the official survey dataset. Early on, SEMCOG staff wrote up multiple pages of questions in MS Word documents, after which NuStats staff would respond with explanations and/or make corrections to the dataset and data dictionary as appropriate. As the official dataset neared finalization, SEMCOG staff had fewer questions, and exchanges of information were more frequently handled via email or over the phone. Even after the dataset was initially "finalized", more fields were later added, coinciding with the refinement of the survey expansion and weighting process. The final correction to the data was made on February 1, 2012, fixing errors identified in the boarding time variable.

Some checks involved simultaneously comparing the actual survey questionnaire, the data dictionary, and the dataset. These included ensuring that every field documented in the data dictionary appeared in the dataset and vice versa; checking that the field-type documentation in the data dictionary (integer, character, date/time, etc.) matched the actual field type in the database; and making sure that every response from the questionnaire was represented by one or more appropriate variables in the dataset and data dictionary.

For individual fields in the dataset, SEMCOG reviewed the clarity of each variable's definition, reviewed the descriptions of allowed responses (for those fields with lists of specific allowed responses), checked that only allowable responses appeared in each field, and questioned the processing of the data for cases in which certain values seemed over-represented or under-represented compared to expectations.

Reviewing fields one at a time was not always sufficient, however, since many variables were interrelated. To the extent that SEMCOG could determine the logic that governed the relationship among such fields, staff validated the data accordingly. For example:

- If the values in one field limited the range of acceptable values in other fields beyond what was written in the data dictionary, SEMCOG queried and validated all fields simultaneously;
- For questions with "Other" as a possible response (along with a place for the respondent to write their own open-ended comment on the form), SEMCOG checked that only those records with an "other" code had a comment (and vice versa); and
- For groups of fields corresponding to questions for which multiple responses were allowed, SEMCOG checked that coded values strictly increased from one field to the next, and that a valid value in one field was never preceded by a null value in a prior field;

Sections of the final dataset were devoted not only to responses culled directly from the survey forms, but also to geocoding and subsequent geospatial analysis, to describing boarding and alighting locations in detail, and to flagging potential errors and inconsistencies in the data. For each of these sections, SEMCOG corresponded with NuStats in order to learn the precise meaning and appropriate use of each variable before devising appropriate validation checks. Of particular importance was understanding the difference between the alighting locations as reported by the survey respondent and the alighting location as "imputed" by NuStats. The validation checks that

SEMCOG ultimately applied varied from section to section. A few of the more important ones included:

- Verifying that all origin, destination, boarding, and alighting locations were located within Southeast Michigan (as well as sorting out the subtleties of the many similar but different fields containing geocoding results);
- Checking that the appropriate linkage existed between the main dataset and the "master stop" table, which contains data for every potential bus stop on every bus route, direction, and pattern in the region; and
- Performing similar checks on the linkage between the main dataset and the "surveyed trips" table, which contains data about the number of boarding and alighting observations from the field for each surveyed bus trip.

In the checks described above, special attention was directed toward identifying data that should be the same in multiple tables, but differed for one reason or another.

In addition to the work described above, SEMCOG also undertook a significant QA/QC effort related to the survey weighting and expansion process. The weighting and expansion process was developed jointly by SEMCOG and NuStats, but as it evolved and new steps were appended to existing steps based on discussion among staff from both agencies, NuStats frequently initiated new calculations, including introducing factors to eliminate specific survey biases. As the process unfolded, SEMCOG made sure to not only fully understand the reasons behind each mathematical operation performed by NuStats, but also to replicate calculations whenever feasible.

Finally, it should be noted that SEMCOG followed through with questions about the data that were raised by the consultant (Cambridge Systematics) that used the on-board survey data to enhance SEMCOG's mode-choice and transit models. In fact, the final correction to the data (the boarding time issue mentioned in the first paragraph of this section) resulted from analysis performed by Cambridge Systematics that was subsequently reviewed by SEMCOG, NuStats and CDM Smith.

2.9 Survey Weighting and Expansion

2.9.1 Background

A logical weighting and expansion procedure is critical to account and adjust for biases in the survey data, which exist because not all trips in the transit system can be sampled and not all riders on sampled trips respond to the survey. In the absence of such a procedure, data users cannot confidently draw conclusions about the characteristics of the transit-riding population.

During the planning stages of SEMCOG's on-board transit survey, the original thought was to stratify samples by route, direction, time of day, and boarding location in order to calculate weighting and expansion factors—a traditional boarding based approach. However, SEMCOG and NuStats, under guidance from the Federal Transit Administration, collaboratively developed a new, thorough weighting and expansion methodology which goes beyond the traditional approach. Key components of the methodology include the following:

- Incorporating both boarding and alighting locations into the weighting process (using a method known as Iterative Proportional Fitting, IPF) to account for the fact that characteristics of transit riders can vary based on where they alight as well as where they board, and to better account for stop-level variation in response rates;

- Stratifying samples by pattern (a *pattern* is a unique trip path with a unique sequence of stops) because in Southeast Michigan, different patterns associated with a particular route often serve dissimilar geographic areas and function more like different routes rather than subtle route variations;
- Overcoming limitations in available target ridership data by developing an innovative method for calculating expected weekday transit ridership at the stop level using the observed boarding and alighting data that was collected during the survey;
- Developing useful software tools and logical methods for aggregating stops along patterns, with stop aggregation a necessity both technically (as a precursor to running the Iterative Proportional Fitting weighting procedure) and philosophically (to align the data with the observed spatial variation in ridership as much as possible), all while considering important factors such as land-use characteristics and the region's uneven distribution of boarding and alighting activity from pattern to pattern, and from stop to stop; and
- Incorporating a special time-of-day factor to ensure that the overall methodology is balanced, preserving temporal variation in ridership as well as spatial variation.

The SEMCOG and NuStats methodology stratifies the survey data in different ways for different calculations: the weighting component of the process stratifies the data by pattern, boarding location, and alighting location, while the expansion component stratifies the data by pattern and time of day. It was not plausible to simultaneously weight and expand by pattern, boarding and alighting stop location, and time of day throughout the entire process—the survey data would have been stretched too thin, and meaningful variations in ridership characteristics would have been lost.

2.9.2 Methodology

For each of the 18,495 records in the survey dataset, the process developed by SEMCOG and NuStats ultimately assigns a final expansion weight that is a function of five key factors: the sample factor, vehicle factor, response factor, expansion factor, and time-of-day factor.

Sample Factor

The first part of the process involved calculating expected weekday ridership estimates by pattern and stop. With such disaggregate data not available from the transit agencies participating in the survey, observed boarding and alighting activity collected during sampled trips was used to calculate the necessary estimates.

First, raw boarding and alighting observations were adjusted in such as way so that the total number of boardings and alightings precisely matched one another for each sampled trip—a necessity for future calculations. Next, these stop-level observations were further adjusted using the sample factor, a value that accounts for the fact that although most sampled bus trips were surveyed only once, some were surveyed two or three times. The sample factor ensures that observations on trips surveyed more than once were not overrepresented in calculations.

Simply stated, the sample factor is equal to the inverse of the number of times a particular scheduled bus trip was surveyed. For example, if the 9:08 AM bus trip on Pattern 1 was surveyed on both Tuesday and Wednesday, then two sets of stop-level boarding and alighting observations would exist for that trip, and the data for all of these records would need to be multiplied by a sample factor of 0.50.

Sample Factor = 1 / Number of times a trip was surveyed

Vehicle Factor

After the sample factor was calculated, the vehicle factor was then tabulated for each pattern and time-of-day combination (also known as *cell*). The vehicle factor serves as a basic expansion factor in the process, accounting for the fact that not every trip in the transit system was sampled. The vehicle factor was simply calculated by dividing the total number of scheduled trips by the number which were sampled. For example, if Pattern 1 had a total of 10 bus trips scheduled during the AM-peak period, but only 2 of these trips were sampled, then the vehicle factor for Pattern 1 during that time period would equal 10 divided by 2, or 5.

Vehicle Factor = Total trips per cell / Sampled trips per cell

Multiplying the balanced trip-level boarding and alighting observations by the sample factor and the vehicle factor, then aggregating the results by pattern and stop, yielded expected weekday boarding and alighting values at each stop for each surveyed pattern, which was necessary input for the next part of the process.

Response Factor

The next part of the process revolved around using the IPF procedure to incorporate both boarding and alighting stop locations into the calculation of response factors, which are weighting factors that account for riders on sampled trips who did not return complete, useable surveys.

In basic terms, IPF aligns the spatial distribution of completed survey records with the distribution of the transit-riding population at the pattern and stop level, where the population distribution is represented by the expected weekday boarding and alighting values calculated as described in the previous section. In other words, expected weekday ridership is a necessary input for the IPF procedure. For IPF to work, it was also necessary to aggregate individual stop locations into groups due to the unique nature of Southeast Michigan's transit system, which includes characteristics such as a large number of bus stops per pattern, uneven distribution of ridership from pattern to pattern, and uneven distribution of boarding and alighting activity from stop to stop along patterns.

To efficiently yet effectively aggregate stops, SEMCOG and NuStats separated patterns into three distinct tiers based on ridership level. Automated procedures were developed for grouping stops on patterns with "low" and "moderate" ridership, while many more resources were allocated toward aggregating stops on the 96 "high" ridership patterns, which together account for about 75% of the region's typical weekday ridership. A great deal of time and effort was spent aggregating stops for these high ridership patterns so that spatial variations in rider characteristics could be captured as accurately as possible via the IPF procedure.

The aggregation approach for high ridership patterns consisted of 1) using GIS functionality to automatically collapse data from the stop level to the node level (using nodes from the travel-demand model's highway network), and 2) manually combining nodes into groups. To assist in the manual effort, SEMCOG developed helpful MS Excel and ArcMap applications for use in conjunction with one another. SEMCOG also developed logical criteria for manually combining nodes, based on a variety of information displayed in the two applications, such as the land-use characteristics associated with each node.

With estimates of stop-level ridership available and with stops logically aggregated into groups, NuStats performed the final, official IPF runs for each of the 355 surveyed and weighted patterns in Southeast Michigan. When IPF runs failed to properly converge, adjustments were made to the grouping of stops as needed. Adjustments were also made to prevent the calculated response factors from being unreasonably large or small.

Expansion Factor

After the IPF procedure was run, the sum of the response factors for all patterns in the region totaled 217,181 riders per weekday. As expected, this value was close to but not precisely equal to the ultimate expansion target: 218,129, the average weekday transit ridership in Southeast Michigan (excluding the Detroit People Mover). Therefore, another adjustment was needed to proportionally align the IPF response weights to match average weekday ridership. The expansion factor fulfills this role.

Expansion factors were calculated at the line level, where *lines* are groups of routes for which average weekday ridership was uniformly available across the region. The calculation itself was a simple one: the average weekday ridership for the line divided by the weighted expected weekday ridership—that is, average weekday ridership divided by the sum of the response factors for every survey associated with the line.

For example, assume that Pattern 1 and Pattern 2 were the only two patterns associated with Line 1. Further assume that the sum of all the response factors calculated for Pattern 1 and Pattern 2 was 7,270, and that the line's average weekday ridership was 7,742 (as provided by the transit agency operating the line). In this case, the expansion factor for Line 1 would be 7,742 / 7,270 = 1.06, and this factor would be applied to every survey associated with that line.

Expansion Factor = Average Weekday Ridership / Weighted Expected Weekday Ridership

Time-of-Day Factor

The product of the response factor and the expansion factor equals a value known as the *initial expansion weight*. The initial expansion weight aligns the survey data with the spatial distribution of ridership observed during the survey. Because of the inclusion of the vehicle factor in IPF input calculations, it also partially accounts for distribution of ridership by time of day. Nevertheless, to better align the survey data with observed variation in ridership by time period, one final adjustment was made using the time-of-day factor.

Like the vehicle factor, the time-of-day factor was calculated at the cell level (that is, by pattern and time of day). The factor was tabulated so that the final expansion weights would match new cell-level expansion targets called *expected riders*. Expected riders, in turn, were calculated by multiplying balanced boarding and alighting observations by the sample factor, the vehicle factor, and the expansion factor; aggregating by pattern and time of day; then making a handful of adjustments, primarily to reallocate data from cells that were sampled but not associated with any completed surveys.

Once expected riders per cell were calculated, it was straightforward to calculate the time-of-day factor itself:

Time-of-day factor = Expected riders per cell / Initial expansion weight

For example, if there were 50 riders expected to board Pattern 1 during the midday period, but 40 was the sum of the initial expansion weights for surveys associated with the midday cell for Pattern 1, then the time-of-day factor for each of these surveys would be 50 / 40 = 1.25. In this example, surveys from the midday time period were underrepresented in the IPF weighting process relative to surveys for the pattern from other time periods, and the time-of-day factor was necessary to bring the sample data back in line with period-level observations.

As alluded to in the background section, in theory samples could have been stratified by time of day throughout the entire weighting and expansion process. However, SEMCOG and NuStats chose instead to follow the steps described above—in other words, to adjust the survey data with a time-of-day factor after weighting the data by pattern. Several related reasons factored into this decision.

First of all, if samples were stratified by time of day prior to IPF, then an IPF run would have been required for each pattern and time period combination, rather than just one run per pattern. In all, well over 1,000 IPF runs would have been needed, stretching the survey data very thin. With many fewer surveys relative to the number of IPF runs, much more stop aggregation would have been required in order for runs to converge, and precision in capturing spatial variations in ridership characteristics would have been lost.

Furthermore, as already noted, due to the lack of available disaggregate boarding and alighting data from the transit agencies participating in the survey, stop-level ridership estimates were calculated based on the boarding and alighting activity observed in the field during the survey. Since not all trips were surveyed, those observations were necessarily limited, and daily estimates were likely to be more accurate than estimates by time period. Accordingly, SEMCOG and NuStats felt more confident using daily estimates for the IPF weighting process rather than estimates by time period.

Again, the goal was to avoid compromising the accuracy of the response factors and to preserve observed spatial ridership distributions as much as possible, but to still—in the end—align the sample data with targets based on period-level observations.

Final Expansion Weight

The calculation of the final expansion weight for each survey record was simply an adjustment of the initial expansion weight using the time-of-day factor:

Final expansion weight = Initial expansion weight * Time-of-day factor

A full description of the weighting and expansion methodology is contained in Appendix F, SEMCOG Weighting and Expansion.

2.10 Survey Data Check & Transit Assignment

This section documents the on-board transit survey data-based transit assignment investigation performed by CDM Smith for the Southeast Michigan Council of Governments (SEMCOG). CDM Smith was tasked to compare the differences and similarities between the on-board survey results and the application of those results to the region's travel demand model.

The purpose of this investigation is to identify, explore and explain the difference between the SEMCOG model results and on-board transit surveys. CDM Smith's investigation focused in four areas:

- Transfer rates,
- Walk access time
- Transfer penalty, and
- Overall assessment on SEMCOG transit assignment

The investigation started by converting the survey data into origin-destination matrices, which serve as inputs to the transit assignment module in the SEMCOG model. After the initial comparison between survey data and model results, four avenues were explored to understand the differences found between the survey data and model results and these findings are detailed in Appendix G, On-Board Transit Survey Data Based Transit Assignment Investigation.

1. Comparison of the survey and model skims by alternative transit assignment algorithms in TransCAD – The normal assignment method for transit trips in TransCAD is known as Pathfinder. This method finds multiple likely transit paths based on the bus frequency at the first boarding stop and at potential transfer points. This method was compared to a single shortest (all-or-nothing) path method.
2. Adjustment of model assumptions such as transfer penalty time - The number of transfers at the system level was one of the main parameters examined. The difference in the number of transfers between the survey data and the SEMCOG model results could be a result of model assumptions such as transfer penalty time values. An adjustment in transfer penalty time was made to shed some light on the reasonableness of the current transfer penalty time.
3. Assignment of survey trips with no transfer. The consultant team assigned only trips with no transfers in the survey to understand and identify discrepancies between the SEMCOG model results and the survey data. Singling out these trips in the assignment excludes the impact of trips involving more complex routes and may reveal factors that could have been obscured in the overall assignment results.
4. Examination of selected survey records - Examination of selected survey records provided specific examples to guide the investigation and support the conclusions reached.

In addition, CDM Smith did an investigation on increasing the maximum access and egress time based on a separate review completed by the Cambridge Systematics (CS) in January 2012. The CS was under a contract with SEMCOG to update the regional travel demand forecasting model and this on-board transit survey records will be used as part of the transit model improvement.

In the *SEMCOG Checking Transit Networks and Path-Building Procedures* memo by Cambridge Systematics (CS) on January 4, 2012, the consultant recommended increasing the maximum walk access time allowed from 18 minutes to 36 minutes.

In this exercise, we tested this increase for the maximum access and egress times and used all the other original values in the SEMCOG transit assignment module. To be somewhat consistent with the CS memo, only the assigned walk access trips are reported in this investigation. A major difference between this and CS investigation is that 12,064 out of 13,336 original unweighted walk access survey trips are used in the CS investigation whereas weighted walk access survey (137,399) trips from the original 13,336 unweighted trips are used in this investigation.

The tables below show the results of increasing the maximum walk access time at the line-level comparison of assigned and survey trips. A comparison of Table 2.13 and Table 2.14 shows that the increase in maximum walk access time to 36 minutes improves the system level match of the assigned transit trips to the surveyed transit trips from a 2.0% difference to a 0.9% difference. At the route level, improvements are seen in BWAT, DPM, LETC, and SMART and UMI, while for AATA and DDOT, the match is slightly worse.

Table 2.13: Boarding from Original SEMCOG Model (Walk Access Trip Assignment with 18 min Walk Time)

Transit Operator	Period	Initial Model Run Boarding (A)	Survey Boarding (B)	Diff. A-B	% Diff. (A-B)/B
AATA	AM	4,111	4,047	64	1.60%
	MD	8,602	8,446	155	1.80%
	PM	5,132	4,506	626	13.90%
	OP	1,066	1,427	-362	-25.30%
	Total	18,911	18,427	484	2.60%
BWAT	AM	318	464	-146	-31.40%
	MD	1,125	1,570	-445	-28.30%
	PM	167	265	-98	-36.90%
	OP	-	-	0	N/A
	Total	1,610	2,298	-688	-29.90%
DDOT	AM	22,798	25,556	-2,757	-10.80%
	MD	32,412	33,580	-1,167	-3.50%
	PM	32,120	35,522	-3,402	-9.60%
	OP	13,521	12,960	561	4.30%
	Total	100,852	107,617	-6,766	-6.30%
DPM	AM	65	-	65	N/A
	MD	138	-	138	N/A
	PM	870	1,128	-258	-22.90%
	OP	779	1,603	-824	-51.40%
	Total	1,851	2,731	-880	-32.20%

Transit Operator	Period	Initial Model Run Boarding (A)	Survey Boarding (B)	Diff. A-B	% Diff. (A-B)/B
LETC	AM	47	111	-64	-58.00%
	MD	200	482	-282	-58.50%
	PM	57	42	14	33.90%
	OP	-	34	-34	-100.00%
	Total	303	669	-366	-54.70%
SMART	AM	9,289	6,916	2,373	34.30%
	MD	12,781	12,084	698	5.80%
	PM	10,572	5,860	4,712	80.40%
	OP	3,039	2,948	91	3.10%
	Total	35,681	27,807	7,874	28.30%
UMI	AM	3,122	1,243	1,879	151.20%
	MD	14,190	12,679	1,512	11.90%
	PM	12,539	11,646	893	7.70%
	OP	3,451	3,657	-206	-5.60%
	Total	33,303	29,225	4,078	14.00%
System Total	AM	39,750	38,336	1,414	3.70%
	MD	69,449	68,840	609	0.90%
	PM	61,457	58,970	2,487	4.20%
	OP	21,856	22,630	-774	-3.40%
	Total	192,512	188,775	3,736	2.00%

Table 2.14: Boarding from SEMCOG Model (Walk Access Trip Assignment with 36 min walk time)

Transit Operator	Period	Initial Model Run Boarding (A)	Survey Boarding (B)	Diff. A-B	% Diff. (A-B)/B
AATA	AM	4,141	4,047	94	2.30%
	MD	8,510	8,446	63	0.80%
	PM	5,239	4,506	733	16.30%
	OP	1,136	1,427	-291	-20.40%
	Total	19,026	18,427	599	3.30%
BWAT	AM	328	464	-136	-29.40%
	MD	1,094	1,570	-476	-30.30%
	PM	208	265	-57	-21.70%
	OP	-	-	0	N/A
	Total	1,629	2,298	-669	-29.10%
DDOT	AM	22,615	25,556	-2,941	-11.50%
	MD	32,068	33,580	-1,512	-4.50%
	PM	31,800	35,522	-3,722	-10.50%
	OP	13,362	12,960	402	3.10%
	Total	99,844	107,617	-7,773	-7.20%
DPM	AM	65	-	65	N/A
	MD	144	-	144	N/A
	PM	881	1,128	-247	-21.90%
	OP	781	1,603	-822	-51.30%
	Total	1,871	2,731	-860	-31.50%

Transit Operator	Period	Initial Model Run Boarding (A)	Survey Boarding (B)	Diff. A-B	% Diff. (A-B)/B
LETC	AM	47	111	-64	-58.00%
	MD	234	482	-248	-51.50%
	PM	42	42	0	0.00%
	OP	-	34	-34	-100.00%
	Total	323	669	-347	-51.80%
SMART	AM	9,071	6,916	2,155	31.20%
	MD	12,421	12,084	337	2.80%
	PM	10,110	5,860	4,250	72.50%
	OP	2,934	2,948	-14	-0.50%
	Total	34,536	27,807	6,729	24.20%
UMI	AM	3,147	1,243	1,904	153.20%
	MD	14,220	12,679	1,541	12.20%
	PM	12,578	11,646	932	8.00%
	OP	3,327	3,657	-331	-9.00%
	Total	33,271	29,225	4,046	13.80%
System Total	AM	39,412	38,336	1,077	2.80%
	MD	68,690	68,840	-150	-0.20%
	PM	60,858	58,970	1888	3.20%
	OP	21,540	22,630	-1,090	-4.80%
	Total	190,500	188,775	1724	0.90%

Tables 2.15 through 2.17 present the results of a comparison of model transfers vs. survey transfers using various transfer penalty and walk access thresholds in prediction success tables. The baseline prediction success rate is 57.7%, using a 6 minute transfer penalty and an 18 minute walk access limit (Table 2.15). When the transfer penalty is increased to 8 minutes, the prediction success rate increases to 59.3% as the better match to 0 and 1 transfer trips more than offsets the slightly worse results for 2 transfer trips (Table 2.16). However, increasing the transfer penalty (to 8 minutes) and increasing the walk access threshold to 36 minutes reduces the match to the number of transfers in the original unweighted survey database to 45.0%. From this result, we conclude that either the transfer penalty or the max access walk time should be adjusted, but not both parameters. The CS and CDM Smith analyses appear to be consistent in this conclusion.

Table 2.15: Unexpanded Survey Walk Access Trip Prediction Success Table (6 min Transfer Penalty and 18 min Walk Access)

Survey Transfers / Model Transfers	0	1	2	3	Total	Prediction Success Rate
0	4,940	2,103	217	4	7,264	57.7%
1	1,903	3,672	580	10	6,165	
2	561	870	331	14	1,776	
3	110	150	36	0	296	
Total	7,514	6,795	1,164	28	15,501	

Table 2.16: Unexpanded Survey Walk Access Trip Prediction Success Table (8 min Transfer Penalty and 18 min Walk Access)

Survey Transfers / Model Transfers	0	1	2	3	Total	Prediction Success Rate
0	4,940	2,103	217	4	7,264	57.7%
1	1,903	3,672	580	10	6,165	
2	561	870	331	14	1,776	
3	110	150	36	0	296	
Total	7,514	6,795	1,164	28	15,501	

Table 2.17: Unexpanded Survey Walk Access Trip Prediction Success Table (8 min Transfer Penalty and 36 min Walk Access)

Survey Transfers / Model Transfers	0	1	2	3	Total	Prediction Success Rate
0	3,530	3,429	300	5	7,264	45.0%
1	2,496	3,297	367	5	6,165	
2	729	896	148	3	1,776	
3	122	147	27	0	2,96	
Total	6,877	7,769	842	13	15,501	

3. Survey Results

The fully weighted and expanded SEMCOG data were used to create the following analyses, displayed in three separate sections. The first section displays system level frequencies of the survey questions, while the second section focuses on the Woodward, Gratiot, and Michigan corridors. For the third section, the SEMCOG and Greater Detroit area were compared against other major metropolitan areas where similar studies have occurred in the last five years.

3.1 Regional Data Summary and Analysis

A total of 18,495 questionnaires from seven transit agencies were completed for the survey. The response rates varied from a low of 10.8 percent on the University of Michigan routes to a high of 24.7 percent on the Blue Water routes. Overall, 16.1 percent of eligible riders on the observed trips completed a survey, with an eligible respondent being defined as a rider over 16 years old, based on visual inspection.

Table 3.1: SEMCOG Overall Survey Results by Transit Agency

Transit Agency	Average Daily Ridership	Sample Goal	Total Completes	Survey-Eligible Boardings	Total Observed Boardings	Eligible Response Rate	Observed Response Rate
Ann Arbor Transit (AATA)	22,010	2,532	2,557	12,484	12,758	20.5%	20.0%
Blue Water Area Transit (BWAT)	2,625	280	286	1,160	1,267	24.7%	22.6%
Detroit Department of Transportation (DDOT)	124,514	9,688	9,327	70,183	76,817	13.3%	12.1%
Detroit People Mover (DPM)	4,011	400	396				
Lake Erie Transit (LET)	877	110	98	495	511	19.8%	19.2%
Suburban Mobility Authority for Regional Transportation (SMART)	33,876	4,574	4,538	18,653	20,484	24.3%	22.2%
University of Michigan (UM)	34,227	1,293	1,293	11,926	11,931	10.8%	10.8%
Total	222,140	18,877	18,495	114,901	123,768	16.1%	14.9%

The tables in the following section display the weighted frequency of responses to the survey in the order of the questionnaire. The first question presented to respondents on the questionnaire was improvements needed on the transit system. The most common request was to increase the frequency of service, with over half of riders requesting this. Riders on the People Mover were much less likely to request this though, and typically responded that there were no improvements needed, other than to increase pedestrian access. Another important aspect of the survey was fare payment. A little under half of riders pay with cash or day pass, while just over 8 percent of riders have their fare paid by their employer. This is much more common for Ann Arbor riders, where 30 percent of riders have their fare paid by their employer.

Across the region, about half of the surveyed trips are made 3-5 days per week, with over a quarter being made more often. If the transit system was not available, 30 percent of riders would try to get a ride with someone else, 30 percent would forego the trip, and only 16 percent would drive. Like most transit systems, about half of the SEMCOG transit ridership population does not have a valid driver's license, although this is not the case for the People Mover, U of M and Ann Arbor systems, where riders are much more likely to have one. Similar to driver's licenses, about half of SEMCOG riders are from zero-vehicle households, although again, this is not true for the People Mover, U of M or Ann Arbor systems. Over half of SEMCOG riders come from households making less than $20,000 per year.

3.1.1 Survey Results by System

Across the region, the most common service improvement suggestion was to increase frequency of service, although it was not the most common for each transit agency. The most common suggestion for AATA riders and LET riders was to end service later, at 43.8 percent and 66.0 percent, respectively. The most common improvement suggestion for DPM riders was to improve pedestrian access, at 15.6 percent, although 52.0 percent of DPM riders reported that no service improvements were needed.

Table 3.2: Transit Improvements by System*

System	Transit Improvements								
	Make transfers easier	Start service earlier	End service later	Add new route	Improve pedestrian access	Increase frequency of service	No service improvements needed	Other, specify	Total
AATA	6.6%	10.4%	43.8%	6.7%	4.2%	36.7%	10.0%	9.3%	100.0%
BWAT	8.8%	20.3%	23.9%	5.5%	4.3%	26.4%	21.7%	11.8%	100.0%
DDOT	17.1%	21.6%	27.5%	3.3%	8.1%	56.8%	6.4%	4.2%	100.0%
DPM	7.9%	7.3%	14.3%	3.7%	15.6%	14.5%	52.0%	2.3%	100.0%
LET	8.2%	14.6%	66.0%	11.9%	2.7%	28.1%	9.0%	9.2%	100.0%
SMART	14.0%	15.7%	31.2%	8.6%	5.6%	38.1%	13.8%	6.1%	100.0%
UM	6.6%	7.6%	20.4%	7.1%	3.7%	62.2%	8.2%	6.6%	100.0%
Total	13.6%	17.0%	28.6%	5.1%	6.7%	51.5%	9.1%	5.5%	100.0%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

Riders on UM routes and DPM routes are most likely to get to their destination with one bus/train only, at 90 percent and 84 percent, respectively. BWAT riders are least likely to get their destination with one bus, at 35.3 percent, while the number for DDOT and LET riders is slightly higher at 39.9 percent and 39.1 percent, respectively.

Table 3.3: Total Buses by System

System	Total Buses (Imputed)				
	One	Two	Three	Four or more	Total
AATA	65.2%	31.2%	3.2%	0.4%	100.0%
BWAT	35.3%	52.2%	11.4%	1.1%	100.0%
DDOT	39.9%	44.0%	13.5%	2.6%	100.0%
DPM	84.0%	11.0%	4.9%	0%	100.0%
LET	39.1%	44.4%	16.5%	0%	100.0%
SMART	49.2%	37.9%	11.5%	1.4%	100.0%
UM	90.0%	9.7%	0.3%	0.1%	100.0%
Total	52.3%	36.0%	10.0%	1.7%	100.0%

Most trips are home-based for all transit systems (53.3 percent) except UM, where they are mostly university-based (55.3 percent). DPM has the highest percentage of both social-based trips and work-based trips, at 26.4 percent and 23.8 percent, respectively.

Table 3.4: Origin Purpose by System

System	Origin Trip Purpose								
	Home	University/ College	Shopping	Social, Eat Out, etc.	Work or Work-Related	High School/ Middle School	Medical Services	Other	Total
AATA	55.2%	16.9%	2.1%	4.2%	17.9%	1.4%	2.1%	0.1%	100.0%
BWAT	72.9%	2.0%	5.0%	5.4%	9.6%	3.0%	1.2%	1.0%	100.0%
DDOT	56.7%	5.8%	3.1%	8.5%	18.3%	4.3%	3.2%	0.1%	100.0%
DPM	42.6%	1.2%	2.5%	26.4%	23.8%	0.7%	2.2%	0.6%	100.0%
LET	66.5%	3.8%	8.7%	9.2%	8.5%	3.2%	0.2%	0%	100.0%
SMART	59.1%	5.2%	3.2%	6.3%	21.9%	2.1%	2.1%	0.2%	100.0%
UM	33.1%	55.3%	0.2%	1.4%	9.0%	0.1%	0.8%	0%	100.0%
Total	53.3%	14.3%	2.6%	6.9%	17.3%	2.9%	2.5%	0.1%	100.0%

Most riders walk to access the various transit systems (87.5 percent overall), although getting dropped off is common for both DPM and LET riders, at 19.5 percent and 18.3 percent, respectively. Park and rides are used for 8.3 percent of DPM trips, 7 percent of UM trips, and 6.7 percent of AATA trips. Carpool users constitute less than one percent of the ridership for each system.

Table 3.5: Access by System

System	Access						
	Walked/ Wheelchair	Dropped off	Drove alone	Carpool	Bicycled	Taxi	Total
AATA	86.6%	4.4%	6.7%	0.6%	1.6%	0%	100.0%
BWAT	86.9%	9.3%	1.2%	0.7%	1.7%	0.2%	100.0%
DDOT	89.1%	9.4%	0.5%	0.3%	0.3%	0.5%	100.0%
DPM	70.6%	19.5%	8.3%	0.8%	0.2%	0.5%	100.0%
LET	80.6%	18.3%	1.1%	0%	0%	0%	100.0%
SMART	80.9%	13.0%	2.7%	0.4%	2.5%	0.4%	100.0%
UM	90.9%	1.5%	7.0%	0.3%	0.2%	0.1%	100.0%
Total	87.5%	8.5%	2.6%	0.4%	0.8%	0.4%	100.0%

DPM riders are the most likely to pay with cash or a day pass, at 90.3 percent; while SMART, DDOT, and BWAT riders are the most likely (23 percent each) to use a different type of pass.

Table 3.6: Fare Paid by System*

System	Fare Paid					
	Cash/ Day Pass	Other Pass	Transfer	University of Michigan (Free)	Other, Specify	Total
AATA	22.8%	15.1%	4.6%	53.6%	5.1%	100.0%
BWAT	63.5%	23.0%	13.6%	0%	1.7%	100.0%
DDOT	60.9%	23.4%	15.9%	0.2%	1.4%	100.0%
DPM	90.3%	7.6%	0.9%	0%	1.8%	100.0%
LET	66.2%	13.1%	23.9%	3.3%	3.8%	100.0%
SMART	63.3%	23.5%	14.9%	0%	1.0%	100.0%
UM	1.3%	1.3%	0%	96.9%	0.5%	100.0%
Total	48.4%	18.7%	11.8%	21.0%	1.6%	100.0%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

The AATA region has the highest percentage of riders using a student (K-12) fare, at 10.7 percent, while none of the DPM riders used this fare type. Percentage-wise, AATA, LET, and BWAT, had the most senior/disabled riders, at 18.6 percent, 15.7 percent, and 14 percent, respectively. (This question was not asked on the University of Michigan routes, which do not charge a fare.)

Table 3.7: Fare Type by System

System	Fare Type			
	Regular Fare	Student (K-12)	Senior/ Disabled	Total
AATA	70.7%	10.7%	18.6%	100.0%
BWAT	80.4%	5.6%	14.0%	100.0%
DDOT	87.7%	6.5%	5.8%	100.0%
DPM	96.0%	0%	4.0%	100.0%
LET	81.4%	2.9%	15.7%	100.0%
SMART	84.4%	6.0%	9.6%	100.0%
Total	86.0%	6.5%	7.5%	100.0%

Overall, 8.2 percent of regional employers paid for the entire transit fare. AATA riders had the highest percentage (30 percent) of employer/school-assisted transit passes. LET and BWAT riders also had high percentages, at 11.8 percent and 10.4 percent respectively, and BWAT riders had the highest percentage of employers subsidizing only part of the fare, at 5.1 percent. (This question was not asked on the University of Michigan routes, which do not charge a fare.)

Table 3.8: Fare Subsidy by System

System	Employer Pay			
	Yes, entire fare	Yes, some of fare	No	Total
AATA	30.0%	3.4%	66.7%	100.0%
BWAT	10.4%	5.1%	84.5%	100.0%
DDOT	6.4%	1.7%	91.9%	100.0%
DPM	7.0%	0.3%	92.7%	100.0%
LET	11.8%	1.2%	87.0%	100.0%
SMART	7.1%	2.0%	90.9%	100.0%
Total	8.2%	1.9%	89.9%	100.0%

The dominant destination purposes were to get home or to work, with the exception of the UM and DPM buses, where they are used predominantly to either get to the university (for UM) or to get to social places (for DPM). AATA routes also had a high percentage of university-bound trips at 25.1 percent, versus 23.1 percent of trips that were toward a Work or Work-Related location.

Table 3.9: Destination Purpose by System

System	Destination Trip Purpose								
	Home	University/ College	Shopping	Social, Eat Out, etc.	Work or Work-Related	High School/ Middle School	Medical Services	Other	Total
AATA	32.6%	25.1%	6.2%	8.6%	23.1%	1.2%	3.0%	0.3%	100.0%
BWAT	21.6%	9.0%	14.6%	15.7%	31.5%	2.7%	4.9%	0%	100.0%
DDOT	33.2%	7.1%	4.6%	17.8%	26.2%	5.6%	5.4%	0.2%	100.0%
DPM	18.2%	2.7%	3.3%	39.1%	33.4%	0.8%	2.5%	0%	100.0%
LET	25.5%	8.4%	19.5%	20.9%	12.9%	0.9%	11.9%	0%	100.0%
SMART	32.4%	6.4%	6.1%	12.3%	35.6%	2.7%	4.3%	0.3%	100.0%
UM	26.8%	60.0%	0%	3.9%	8.2%	0%	1.0%	0%	100.0%
Total	31.6%	16.9%	4.4%	14.3%	24.7%	3.7%	4.3%	0.2%	100.0%

While most riders in the region were able to egress the transit system without using a vehicle, LET riders had a high percentage of riders (12.8 percent) who were picked up at their final stop. SMART also had a comparatively large number of riders with an auto egress, with 6.7 percent getting picked up, and 3 percent with other types of auto egress. SMART riders also had the highest percentage of bicycle riders in the region, at 1.8 percent of riders.

Table 3.10: Egress by System

System	Egress						
	Walk/ Wheelchair	Picked up	Drive alone	Carpool	Bicycle	Taxi	Total
AATA	91.0%	1.6%	5.5%	0.3%	1.6%	0%	100.0%
BWAT	93.9%	1.7%	2.3%	0.5%	1.4%	0.2%	100.0%
DDOT	93.9%	4.9%	0.4%	0.4%	0.2%	0.4%	100.0%
DPM	91.0%	3.8%	3.6%	0.4%	0.2%	0.9%	100.0%
LET	86.9%	12.8%	0%	0%	0.4%	0%	100.0%
SMART	88.5%	6.7%	2.2%	0.5%	1.8%	0.3%	100.0%
UM	92.9%	0.3%	5.7%	0%	1.1%	0%	100.0%
Total	92.5%	4.1%	2.0%	0.3%	0.7%	0.3%	100.0%

Across the region, about half of transit riders made the surveyed trips 3–5 days per week, with about one-quarter making the trips more often and one-quarter less often. The exception to this is LET, where only 5.5 percent of the trips were reported to be made more than five days per week and nearly half of the trips made less than three days per week.

Table 3.11: Trip Frequency by System

System	Trip Frequency						
	6–7 days per week	3–5 days per week	1–2 days per week	1–3 days per month	Less than 1 day per month	First time to make this trip	Total
AATA	20.7%	56.7%	11.7%	5.9%	2.1%	2.9%	100.0%
BWAT	24.9%	41.9%	20.0%	8.6%	2.0%	2.6%	100.0%
DDOT	26.6%	47.5%	11.0%	8.2%	3.0%	3.7%	100.0%
DPM	16.7%	29.0%	30.4%	7.7%	8.5%	7.6%	100.0%
LET	5.5%	48.7%	20.7%	20.0%	4.0%	1.1%	100.0%
SMART	24.3%	52.4%	10.6%	7.0%	2.8%	2.9%	100.0%
UM	46.4%	42.2%	7.1%	2.3%	1.2%	.8%	100.0%
Total	29.5%	47.6%	10.7%	6.5%	2.6%	3.0%	100.0%

Overall, the most common alternative travel mode if transit was not available would be to ride with someone else, at 30.7 percent. A very close second alternative, at 30.5 percent, was to not make the trip altogether. A notable exception is for DPM, where only 8.7 percent would forgo the trip, and 48.1 percent would have walked for the entire trip due to the nature of the service. BWAT had the highest percentage that would take a taxi, at 20.5 percent and the lowest percentage that would drive, at 4 percent. UM riders were the most likely to use a bicycle if transit was not available, at 19.3 percent.

Table 3.12: Alternative Travel Mode by System*

System	Alternative Travel Mode						
	Walk/ Wheelchair	Drive	Ride with someone else	Taxi	Bicycle	Would not make this trip	Total
AATA	20.4%	34.9%	21.0%	10.9%	13.8%	18.6%	100.0%
BWAT	37.2%	4.0%	26.8%	20.5%	14.8%	27.8%	100.0%
DDOT	18.1%	10.2%	37.4%	15.7%	4.7%	33.5%	100.0%
DPM	48.1%	13.7%	25.4%	10.7%	2.2%	8.7%	100.0%
LET	37.6%	10.5%	38.0%	4.7%	10.3%	34.8%	100.0%
SMART	12.1%	16.3%	31.9%	10.7%	6.5%	39.2%	100.0%
UM	25.7%	26.3%	14.0%	7.4%	19.3%	22.4%	100.0%
Total	19.4%	16.3%	30.7%	13.0%	8.4%	30.5%	100.0%

**Select all that apply*
Figures may not add up to totals due to multiple responses per record.

Across the region, transit riders are nearly evenly split between having and not having a valid driver's license. The areas with the highest concentration of riders with driver's licenses are served by UM, DPM, and AATA, at 91.3 percent, 79.4 percent, and 72.5 percent, respectively. BWAT and DDOT had the lowest concentration of riders with driver's licenses at 36.1 percent and 39 percent, respectively.

Table 3.13: Valid Driver's License by System

System	License		
	Yes	No	Total
AATA	72.5%	27.5%	100.0%
BWAT	36.1%	63.9%	100.0%
DDOT	39.0%	61.0%	100.0%
DPM	79.4%	20.6%	100.0%
LET	40.5%	59.5%	100.0%
SMART	51.0%	49.0%	100.0%
UM	91.3%	8.7%	100.0%
Total	53.7%	46.3%	100.0%

UM and AATA routes have the highest percentage of younger riders. Over 80 percent of UM riders are 18–25, while 63.8 percent of AATA riders are younger than 35. DPM riders have a larger percentage of riders in the 35–54 and 55–64 age ranges than any other system at 49.5 percent and 15.6 percent, while LET riders have the highest percent in the 65 plus range, at 6 percent.

Table 3.14: Age by System

System	Age						
	Under 18	18–25	26–34	35–54	55–64	65 + years of age	Total
AATA	3.3%	37.2%	23.3%	22.7%	10.1%	3.4%	100.0%
BWAT	7.3%	27.7%	19.8%	32.0%	10.0%	3.3%	100.0%
DDOT	7.8%	24.7%	16.5%	37.4%	11.8%	1.8%	100.0%
DPM	0.4%	11.9%	18.5%	49.5%	15.6%	4.1%	100.0%
LET	0.6%	29.1%	15.6%	41.1%	7.6%	6.0%	100.0%
SMART	3.6%	21.7%	19.5%	40.7%	12.0%	2.6%	100.0%
UM	2.1%	80.9%	7.3%	6.3%	3.4%	0%	100.0%
Total	5.6%	34.9%	16.2%	31.2%	10.2%	1.8%	100.0%

DPM has the largest percentage of full-time workers, at 66 percent of riders, while LET has the largest percentage of unemployed, seeking work riders, at 27.8 percent. UM has the highest percentage of university student riders, at 87.6 percent, while DDOT has the highest percentage of primary education student riders, due to its specific functioning in this capacity, at 8.9 percent.

Table 3.15: Worker/Student Status by System*

System	Employment Status									
	Full-time Worker	Part-time Worker	Home-maker	University/ College Student	Middle/ High School Student	Other Student	Unemployed, seeking work	Unemployed, not seeking work	Retired	Total
AATA	34.9%	22.2%	2.4%	45.4%	2.9%	2.5%	6.7%	1.9%	4.9%	100.0%
BWAT	21.6%	31.9%	6.0%	17.8%	6.3%	2.7%	21.5%	4.8%	9.7%	100.0%
DDOT	34.0%	20.9%	6.3%	18.2%	8.9%	4.6%	18.9%	2.5%	5.6%	100.0%
DPM	66.0%	10.5%	2.3%	5.6%	0.4%	0.4%	8.8%	4.0%	5.7%	100.0%
LET	21.1%	26.1%	13.1%	15.0%	0.6%	3.4%	27.8%	11.2%	8.5%	100.0%
SMART	46.3%	21.0%	4.5%	18.1%	3.5%	2.6%	14.0%	2.5%	5.5%	100.0%
UM	10.1%	19.8%	0.4%	87.6%	0.1%	0.4%	1.4%	1.8%	0.2%	100.0%
Total	32.3%	20.9%	4.6%	32.4%	5.8%	3.3%	13.9%	2.4%	4.7%	100.0%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

Overall, about half of transit riders in the region had one or more household vehicles, but this varies across systems. Within the Detroit area, DPM riders are much more likely to have at least one vehicle, at 76.3 percent, while DDOT riders are much less likely, at 39.6 percent. BWAT riders are the most dependent on transit, where 73.1 percent of riders do not have a household vehicle.

Table 3.16: Household Vehicles by System

System	Household Vehicle				
	None	One	Two	Three or more	Total
AATA	37.8%	35.6%	19.6%	7.1%	100.0%
BWAT	73.1%	18.8%	7.4%	0.7%	100.0%
DDOT	60.4%	27.5%	9.2%	2.9%	100.0%
DPM	23.7%	38.7%	27.3%	10.3%	100.0%
LET	64.0%	26.7%	3.3%	6.1%	100.0%
SMART	51.7%	32.1%	12.0%	4.2%	100.0%
UM	33.8%	28.9%	18.0%	19.4%	100.0%
Total	51.8%	29.3%	12.4%	6.4%	100.0%

AATA and LET riders typically come from smaller households, where one and two-person households make up 66 percent of the ridership for AATA, and 63.5 percent of the ridership for LET. UM ridership includes the highest percentage of four-plus-person households at 38.2 percent, although most of those household are likely non-related households. DDOT has the second highest percentage (36.4 percent) of four-plus-person households.

Table 3.17: Household Size by System

System	Household Size				
	One	Two	Three	Four or more	Total
AATA	28.4%	37.6%	14.6%	19.4%	100.0%
BWAT	25.5%	26.6%	18.8%	29.1%	100.0%
DDOT	18.1%	24.0%	21.5%	36.4%	100.0%
DPM	26.7%	29.3%	22.0%	22.0%	100.0%
LET	32.0%	31.5%	19.1%	17.3%	100.0%
SMART	20.4%	26.9%	23.7%	29.0%	100.0%
UM	16.3%	29.1%	16.4%	38.2%	100.0%
Total	19.5%	26.9%	20.2%	33.4%	100.0%

Region-wide over 70 percent of transit-riding households have at least one worker in the household. However, LET and BWAT riders are more likely to be unemployed, where 43.3 percent and 37.2 percent of riders, respectively, come from zero-worker households. DPM has the lowest percentage of zero-worker households and the highest percentage of one-worker and two-worker households.

Table 3.18: Household Employees by System

System	Household Workers				
	None	One	Two	Three or more	Total
AATA	21.1%	40.0%	29.8%	9.0%	100.0%
BWAT	37.2%	30.5%	21.1%	11.2%	100.0%
DDOT	29.0%	40.1%	22.6%	8.3%	100.0%
DPM	15.6%	43.8%	33.5%	7.1%	100.0%
LET	43.3%	33.6%	16.0%	7.2%	100.0%
SMART	23.8%	39.2%	27.4%	9.5%	100.0%
UM	30.2%	28.8%	27.5%	13.6%	100.0%
Total	27.6%	37.9%	25.0%	9.5%	100.0%

LET has the highest percentage of riders in the lowest income bracket of less than $10,000 per year, at 52.8 percent, while BWAT follows closely with 47.3 percent making less than $10,000 per year. DPM riders also had a small percentage of passengers in this category, at 16.7 percent, although most riders were in the higher income categories.

Table 3.19: Income by System

System	Income								
	Less than $10,000	$10,000-$19,999	$20,000-$29,999	$30,000-$39,999	$40,000-$49,999	$50,000-$59,999	$60,000-$74,999	$75,000 or more	Total
AATA	30.1%	11.7%	11.2%	12.9%	10.9%	11.8%	6.6%	4.8%	100.0%
BWAT	47.3%	22.0%	18.4%	5.1%	5.3%	0.5%	1.5%	0%	100.0%
DDOT	43.8%	16.6%	14.8%	11.2%	7.4%	4.8%	0.8%	0.7%	100.0%
DPM	16.7%	7.5%	4.6%	14.3%	18.4%	19.4%	8.4%	10.7%	100.0%
LET	52.8%	21.1%	13.6%	7.3%	2.7%	0.6%	1.9%	0%	100.0%
SMART	30.2%	17.5%	13.9%	13.7%	10.0%	8.6%	3.0%	3.1%	100.0%
UM	43.4%	5.0%	4.6%	7.0%	5.9%	8.1%	7.9%	18.2%	100.0%
Total	39.9%	14.4%	12.6%	11.1%	8.0%	6.7%	3.0%	4.4%	100.0%

3.1.2 Trip Purpose and Access/Egress Mode

Across all transit systems and time periods, the most common trip is from home to work, at 21.9 percent of trips, while the second most common is the reverse commute from work to home, at 13.7 percent. Home to University/College constitute 10.7 percent of trips, while University/College to Home make up 7.5 percent of trips.

Table 3.20: Regional Distribution of Origin Purpose by Destination Purpose

Origin Trip Purpose	Destination Trip Purpose								
	Home	University/ College	Shopping	Social, Eat Out, etc.	Work or Work-Related	High School/ Middle School	Medical Services	Other	Total
Home	0%	10.7%	3.4%	10.4%	21.9%	3.4%	3.5%	0%	53.3%
University/College	7.5%	5.4%	0.2%	0.6%	0.5%	0%	0.1%	0%	14.3%
Shopping	2.0%	0.1%	0.2%	0.2%	0.1%	0%	0.1%	0%	2.6%
Social, Eat Out, Recreational, Religious, Community, Personal Business	4.2%	0.2%	0.2%	1.5%	0.5%	0.1%	0.2%	0%	6.9%
Work or Work-Related	13.7%	0.4%	0.2%	1.2%	1.6%	.0%	0.2%	0%	17.3%
High School/ Middle School	2.4%	0%	0%	0.2%	0.1%	0.2%	.0%	0%	2.9%
Medical Services	1.8%	0%	0.1%	0.2%	0.1%	0%	0.2%	0%	2.5%
Other	0%	0%	0%	.0%	0%	.0%	.0%	0%	0.1%
Total	31.6%	16.9%	4.4%	14.3%	24.7%	3.7%	4.3%	0.1%	100.0%

Over 83 percent of riders walked to both access and egress the transit system. Six percent of riders were dropped off at the bus stop, and then walked to egress; while 1.8 percent of riders used a park and ride and then walked to egress.

Table 3.21: Distribution of Access Mode by Egress Mode

Access	Egress						
	Walked/ Wheelchair	Picked up	Drive alone	Carpool	Bicycled	Taxi	Total
Walked/Wheelchair	83.8%	1.9%	1.2%	0.2%	0.2%	0.1%	87.5%
Dropped off	6.2%	2.0%	0.1%	0.1%	0%	0%	8.5%
Drove alone	1.8%	0.1%	0.7%	0%	0%	0%	2.6%
Carpool	0.3%	0%	0%	0.1%	0%	0%	0.4%
Bicycle	0.3%	0%	0%	0%	0.5%	0%	0.8%
Taxi	0.2%	0%	0%	0%	0%	0.1%	0.4%
Total	92.5%	4.1%	2.0%	0.3%	0.7%	0.3%	100.0%

Across the region, 35.5 percent of riders were on home-based work trips, and nearly half of riders were on home-based non-work trips. DDOT was close to the region averages for both of these trip types, with 38.4 percent of riders on home-based work trips, and close to half of riders on home-based non-work trips, while the percentages were reversed for SMART, with 39.7 percent on home-based non-work trips, and close to half on home-based work trips. As would be expected, U-M had the highest percentage of non-home non-work trips, at 35.3 percent, and the Detroit People Mover had the highest percentage of non-home-based work related trips.

Table 3.22: Distribution of Trip Types

System	Trip Type				
	Home-based Non-Work	Home-based Work	Non-Home Non-Work	Non-Home-Based Work	Total
AATA	52.8%	35.1%	7.3%	4.9%	100.0%
BWAT	58.0%	36.5%	2.9%	2.7%	100.0%
DDOT	51.5%	38.4%	5.6%	4.5%	100.0%
DPM	34.1%	26.6%	11.8%	27.4%	100.0%
LET	71.0%	21.0%	7.7%	0.4%	100.0%
SMART	39.7%	51.8%	4.4%	4.1%	100.0%
UM	49.3%	10.7%	35.3%	4.7%	100.0%
Total	49.3%	35.5%	10.2%	4.9%	100.0%

3.2 Corridor Data Specific Analysis

Three of the most important corridors were also analyzed to show how these areas compare to the overall data. The three corridors selected were Woodward Avenue, Michigan Avenue, and Gratiot Avenue, which all lead into downtown from the Northwest, West, and Northeast, respectively. To analyze the ridership characteristics along these corridors, records were selected that used these corridors on their trip, defined as the following:

- The Woodward Avenue corridor is defined as riders who traveled on Woodward Avenue while being surveyed on the SMART 420, 450, 465, 475, 495, or 610, or DDOT 53 between Pontiac, MI and downtown.
- The Michigan Avenue corridor is defined as riders who used the SMART 200 or DDOT 37, covering Michigan Avenue between John Hix Road and downtown.
- The Gratiot Avenue corridor is defined as riders who traveled on Gratiot Avenue while being surveyed on the SMART 510, 515, 530, 560, 565, or 580, or DDOT 34 or 76 between 23 Mile Road and downtown.

Like most riders across the region, the most common service improvement suggestion was to increase frequency of service. Riders along the Woodward corridor were the least likely to suggest a new route.

Table 3.23: Transit Improvements by Corridor*

Corridor	Transit Improvements								
	Make transfers easier	Start service earlier	End service later	Add new route	Improve pedestrian access	Increase frequency of service	No service improvements needed	Other, specify	Total
WOODWARD	17.1%	16.2%	24.3%	4.1%	9.3%	52.9%	10.3%	4.2%	100.0%
MICHIGAN	13.7%	17.9%	22.8%	7.2%	5.9%	49.9%	8.9%	4.4%	100.0%
GRATIOT	16.7%	17.1%	24.3%	7.9%	9.0%	48.6%	11.8%	4.0%	100.0%
SEMCOG	13.6%	17.0%	28.6%	5.1%	6.7%	51.5%	9.1%	5.5%	100.0%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

Compared to the overall survey results, riders on the corridors were less likely to be able to get to their final destination with only one bus, particularly along the Woodward corridor where only 40.8 percent of riders reached their final destination without a transfer.

Table 3.24: Total Buses by Corridor

Corridor	Total Buses (Imputed)				
	One	Two	Three	Four or more	Total
WOODWARD	40.8%	43.1%	13.3%	2.8%	100.0%
MICHIGAN	41.2%	46.9%	10.0%	1.9%	100.0%
GRATIOT	47.4%	38.8%	11.0%	2.8%	100.0%
SEMCOG	52.3%	36.0%	10.0%	1.7%	100.0%

Origin purposes along the corridors closely mirror other trips in the SEMCOG region. The Gratiot corridor had a notably higher number of shopping-based trips at 5.9 percent versus the regional percentage of 2.6 percent, while the Woodward corridor had a larger number of medical-based trips, at 5.6 percent, versus the 2.5 percent across the region.

Table 3.25: Origin Purpose by Corridor

Corridor	Origin Trip Purpose								
	Home	University/ College	Shopping	Social, Eat Out, etc.	Work or Work-Related	High School/ Middle School	Medical Services	Other	Total
WOODWARD	56.0%	5.3%	2.7%	6.6%	20.3%	3.4%	5.6%	0.1%	100.0%
MICHIGAN	57.5%	9.3%	1.6%	9.7%	17.8%	3.4%	0.8%	0%	100.0%
GRATIOT	58.0%	4.2%	5.9%	7.8%	19.8%	2.2%	2.0%	0%	100.0%
SEMCOG	53.3%	14.3%	2.6%	6.9%	17.3%	2.9%	2.5%	0.1%	100.0%

Woodward corridor riders walked to access the transit system in the same proportion as the regional average, but the riders on the Michigan and Gratiot corridors were slightly more likely to get dropped off at their boarding stop rather than walk.

Table 3.26: Access by Corridor

Corridor	Access						
	Walked/ Wheelchair	Dropped off	Drove alone	Carpool	Bicycled	Taxi	Total
WOODWARD	86.7%	10.2%	0.6%	0.5%	1.8%	0.3%	100.0%
MICHIGAN	83.2%	14.9%	0.6%	0%	0.7%	0.6%	100.0%
GRATIOT	83.4%	12.9%	1.2%	0.4%	1.1%	1.0%	100.0%
SEMCOG	87.5%	8.5%	2.6%	0.4%	0.8%	0.4%	100.0%

Riders along the Michigan corridor were the most likely to use cash or a day pass, at 62.2 percent. Riders along the Woodward corridor were the most likely to use some other pass, at 26.9 percent.

Table 3.27: Fare Paid by Corridor*

Corridor	Fare Paid				
	Cash/ Day Pass	Other Pass	Transfer	Other	Total
WOODWARD	57.8%	26.9%	14.7%	2.2%	100.0%
MICHIGAN	62.2%	24.1%	14.0%	0.6%	100.0%
GRATIOT	61.5%	22.1%	17.6%	0.7%	100.0%
SEMCOG	48.4%	18.7%	11.8%	22.6%	100.0%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

The Woodward corridor had the highest percentage of riders using a senior/disabled fare type, at 10.1 percent, which is slightly higher than the regional average of 7.5 percent. In addition, the Michigan corridor had the highest percentage of student (K-12) fares, at 8.4 percent, versus the regional average of 6.5 percent.

Table 3.28: Fare Type by Corridor

Corridor	Fare Type			
	Regular Fare	Student (K-12)	Senior/ Disabled	Total
WOODWARD	84.3%	5.6%	10.1%	100.0%
MICHIGAN	85.4%	8.4%	6.2%	100.0%
GRATIOT	85.8%	6.6%	7.6%	100.0%
SEMCOG	86.0%	6.5%	7.5%	100.0%

Riders using the Michigan corridor were more likely to have their fare paid for by their employer, at 8.7 percent, while riders along the Woodward and Gratiot corridors were less likely than the regional average, at 7.1 percent and 6.2 percent, respectively.

Table 3.29: Fare Subsidy by Corridor

Corridor	Employer Pay			
	Yes, entire fare	Yes, some of fare	No	Total
WOODWARD	7.1%	2.3%	90.6%	100.0%
MICHIGAN	8.7%	2.4%	88.9%	100.0%
GRATIOT	6.2%	2.2%	91.6%	100.0%
SEMCOG	8.2%	1.9%	89.9%	100.0%

The Woodward corridor had the highest percentage of trips made for medical services than the other corridors, at 8.7 percent. Compared to the regional average, the Michigan corridor had more work-bound trips, at 34.6 percent versus the regional average of 24.7 percent. The Michigan corridor also had less social-bound trips, at 9.9 percent, versus the regional number of 14.3 percent.

Table 3.30: Destination Purpose by Corridor

Corridor	Destination Trip Purpose								
	Home	University/ College	Shopping	Social, Eat Out, etc.	Work or Work-Related	High School/ Middle School	Medical Services	Other	Total
WOODWARD	30.8%	7.7%	5.6%	17.5%	26.5%	2.8%	8.7%	0.4%	100.0%
MICHIGAN	30.8%	8.6%	8.4%	9.9%	34.6%	2.7%	4.3%	0.7%	100.0%
GRATIOT	32.3%	6.0%	7.2%	16.6%	27.9%	6.5%	3.3%	0.1%	100.0%
SEMCOG	31.6%	16.9%	4.4%	14.3%	24.7%	3.7%	4.3%	0.2%	100.0%

Similar to access modes, the three corridors have higher percentages of vehicle egress modes. All three corridors have higher percentages of passengers getting picked up at their final stop, with the highest being the Michigan corridor, at 7.9 percent versus the regional percentage of 4.1 percent egressing likewise.

Table 3.31: Egress by Corridor

Corridor	Egress						
	Walk/ Wheelchair	Picked up	Drive alone	Carpool	Bicycle	Taxi	Total
WOODWARD	92.1%	6.2%	0.6%	0.1%	0.8%	0.1%	100.0%
MICHIGAN	90.0%	7.9%	0.6%	0.6%	0.3%	0.5%	100.0%
GRATIOT	91.2%	5.6%	1.9%	0.2%	0.8%	0.2%	100.0%
SEMCOG	92.5%	4.1%	2.0%	0.3%	0.7%	0.3%	100.0%

The Gratiot corridor has the highest percentage of riders who made their trip more than five days per week, at 28 percent. The Michigan corridor had the highest percentage of riders who made their trip 3–5 days per week, at 55.9 percent.

Table 3.32: Trip Frequency by Corridor

Corridor	Trip Frequency						
	6–7 days per week	3–5 days per week	1–2 days per week	1–3 days per month	Less than 1 day per month	First time to make this trip	Total
WOODWARD	21.9%	49.1%	12.5%	9.3%	3.5%	3.8%	100.0%
MICHIGAN	19.7%	55.9%	10.8%	8.0%	3.6%	1.9%	100.0%
GRATIOT	28.0%	45.2%	13.5%	6.2%	3.9%	3.2%	100.0%
SEMCOG	25.2%	49.7%	11.2%	7.7%	2.9%	3.5%	100.0%

If bus service was not available, riders along the Michigan and Gratiot corridors were less likely to report that they would walk to make their trip, compared to the riders overall in the region. They were also more likely to forego the trip.

Table 3.33: Alternative Travel Mode by Corridor*

Corridor	Alternative Travel Mode						
	Walk/ Wheelchair	Drive	Ride with someone else	Taxi	Bicycle	Would not make this trip	Total
WOODWARD	18.3%	13.1%	36.6%	13.8%	4.7%	31.8%	100.0%
MICHIGAN	13.3%	11.4%	34.5%	14.7%	6.0%	40.2%	100.0%
GRATIOT	14.1%	12.2%	35.5%	11.9%	4.8%	38.0%	100.0%
SEMCOG	19.4%	16.3%	30.7%	13.0%	8.4%	30.5%	100.0%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

Riders along all three of the corridors were less likely to have a driver's license, compared to the region overall, especially along the Michigan corridor, where 63.8 percent of riders surveyed have no valid driver's license, compared to only 46.3 percent of SEMCOG riders overall.

Table 3.34: Valid Driver's License by Corridor

Corridor	License		
	Yes	No	Total
WOODWARD	46.2%	53.8%	100.0%
MICHIGAN	36.2%	63.8%	100.0%
GRATIOT	40.8%	59.2%	100.0%
SEMCOG	53.7%	46.3%	100.0%

Compared to the overall region, the Woodward corridor has the largest percentage of older riders. All three corridors have fewer riders in the 18–25 range when compared to the region as a whole.

Table 3.35: Age by Corridor

Corridor	Age						
	Under 18	18–25	26–34	35–54	55–64	65 + years of age	Total
WOODWARD	4.9%	15.0%	14.5%	43.8%	18.1%	3.8%	100.0%
MICHIGAN	7.1%	20.8%	17.6%	43.6%	10.2%	0.7%	100.0%
GRATIOT	5.1%	24.7%	19.3%	37.6%	11.2%	2.1%	100.0%
SEMCOG	5.6%	34.9%	16.2%	31.2%	10.2%	1.8%	100.0%

The Woodward corridor has a higher percentage of both retired riders, at 10.9 percent versus the regional average of 4.7 percent, and "unemployed – not seeking work" riders, at 4.5 percent in the corridor versus 2.4 percent in the region overall.

Table 3.36: Worker/Student Status by Corridor*

Corridor	Employment Status									
	Full-time Worker	Part-time Worker	Home-maker	University/ College Student	Middle/ High School Student	Other Student	Unemployed-seeking work	Unemployed-not seeking work	Retired	Total
WOODWARD	34.7%	20.3%	5.5%	16.8%	5.3%	3.0%	18.3%	4.5%	10.9%	100.0%
MICHIGAN	33.5%	23.5%	8.4%	21.1%	6.1%	3.1%	21.5%	2.3%	4.5%	100.0%
GRATIOT	38.9%	20.4%	6.4%	17.3%	7.2%	5.7%	16.2%	1.6%	4.6%	100.0%
SEMCOG	32.3%	20.9%	4.6%	32.4%	5.8%	3.3%	13.9%	2.4%	4.7%	100.0%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

All three corridors have a higher percentage of riders without a household vehicle, with the largest difference along the Woodward corridor where 61.7 percent of riders do not have a household vehicle, compared to 51.8 percent of overall riders in the SEMCOG region.

Table 3.37: Household Vehicles by Corridor

Corridor	Household Vehicle				
	None	One	Two	Three or more	Total
WOODWARD	61.7%	26.5%	9.4%	2.4%	100.0%
MICHIGAN	54.4%	31.3%	11.8%	2.4%	100.0%
GRATIOT	58.2%	27.5%	10.1%	4.2%	100.0%
SEMCOG	51.8%	29.3%	12.4%	6.4%	100.0%

The Woodward corridor has a larger percentage of riders from one-person households, at 28.1 percent versus the regional average of 19.5 percent. Similarly, the Woodward corridor also has a smaller percentage of four or more person households, at 25.7 percent versus the regional average of 33.4 percent.

Table 3.38: Household Size by Corridor

Corridor	Household Size				
	One	Two	Three	Four or more	Total
WOODWARD	28.1%	26.7%	19.5%	25.7%	100.0%
MICHIGAN	15.9%	22.9%	28.6%	32.6%	100.0%
GRATIOT	18.6%	26.5%	20.2%	34.7%	100.0%
SEMCOG	19.5%	26.9%	20.2%	33.4%	100.0%

The Woodward corridor has a higher percentage of zero-worker and one-worker households as compared to the region overall. The other two corridors were very close to the regional distribution.

Table 3.39: Household Employees by Corridor

Corridor	Household Workers				
	None	One	Two	Three or more	Total
WOODWARD	32.9%	42.0%	17.7%	7.4%	100.0%
MICHIGAN	26.3%	38.3%	23.0%	12.3%	100.0%
GRATIOT	26.1%	41.6%	24.6%	7.7%	100.0%
SEMCOG	27.6%	37.9%	25.0%	9.5%	100.0%

Riders along the Michigan corridor are less likely to have an annual household income of less than $10,000, at 30.2 percent versus the regional average of 39.9 percent, and more likely to have an annual household income between $10,000 and $29,999.

Table 3.40: Income by Corridor

Corridor	Income								
	Less than $10,000	$10,000–$19,999	$20,000–$29,999	$30,000–$39,999	$40,000–$49,999	$50,000–$59,999	$60,000–$74,999	$75,000 or more	Total
WOODWARD	41.6%	15.7%	11.9%	14.0%	8.6%	5.2%	1.7%	1.3%	100.0%
MICHIGAN	30.2%	20.0%	16.0%	10.4%	14.0%	6.2%	2.2%	0.9%	100.0%
GRATIOT	40.2%	18.1%	13.0%	11.5%	8.6%	6.4%	1.3%	0.9%	100.0%
SEMCOG	39.9%	14.4%	12.6%	11.1%	8.0%	6.7%	3.0%	4.4%	100.0%

3.3 Survey Results Compared to Other Recent Studies

The following analysis compares the ridership of SEMCOG regional transit agencies to other regions based on similar transit O/D studies conducted in the last five years, as well as distinguishing the ridership of SEMCOG overall from the ridership of Greater Detroit (defined as DDOT, SMART, and the DPM). Studies without a given question are left out of the respective table. Regions included for comparison are the following:

- Phoenix (Valley Metro) 2007
- Dallas (DART bus and rail) 2007
- Baltimore (MTA commuter rail, subway, light rail and bus) 2008
- Columbus (COTA and DATA) 2008
- Denver (RTD) 2008
- Charlotte (CATS) 2009
- Broward County (BCT) 2010

Riders in the Greater Detroit transit systems were more likely to suggest making transfers easier or starting service earlier, as compared the overall region.

Table 3.41: System Improvements by Region*

Region	Make transfers easier	Start service earlier	End service later	Add new route	Improve pedestrian access	Increase frequency of service	No service improvements needed	Other, specify	Total
SEMCOG	13.6%	17.0%	28.6%	5.1%	6.7%	51.5%	9.1%	5.5%	100%
Greater Detroit	16.2%	20.0%	28.0%	4.4%	7.7%	51.9%	9.0%	4.6%	100%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

Riders in Columbus were the most likely to not need to transfer, at 64 percent; SEMCOG has about 53% of the riders surveyed with no transfer during their journeys. This is the second highest in our comparison.

Table 3.42: Total Buses by Study

Region	One this bus only	Two	Three or More	Totals
SEMCOG	52.3%	36.0%	11.7%	100%
Greater Detroit	42.9%	41.9%	15.1%	100%
Columbus	64.0%	33.0%	3.0%	100%
Charlotte	39.0%	42.0%	19.0%	100%
Phoenix	37.0%	44.0%	19.0%	100%
Broward County	35.6%	39.7%	24.7%	100%
Dallas	46.0%	35.0%	19.0%	100%

**Some systems operated more than just bus services, as shown in the introduction.*

Riders in both Greater Detroit and SEMCOG overall have a much higher rate of Home-Based Non-Work trips, which make up nearly half of the trips for both groups. Denver has the lowest rate of these trips, at 17 percent, while the other regions range from 30–37 percent.

Table 3.43: Trip Types by Study

Region	Home-based Non-Work	Home-based Work	Non-Home Non-Work	Non-Home-Based-Work	Total
SEMCOG	49.3%	35.5%	10.3%	4.9%	100%
Greater Detroit	48.6%	40.9%	5.5%	5.0%	100%
Denver	17.0%	50.0%	29.0%	4.0%	100%
Columbus	34.0%	50.0%	7.0%	9.0%	100%
Charlotte	37.0%	48.0%	8.0%	7.0%	100%
Baltimore	30.0%	62.0%	3.0%	5.0%	100%
Phoenix	33.0%	44.0%	18.0%	5.0%	100%
Broward County	36.6%	49.3%	6.2%	7.9%	100%
Dallas	37.0%	53.0%	4.0%	6.0%	100%

Compared to Charlotte and Baltimore, SEMCOG riders are more likely to access and egress the transit system without a vehicle, although they are less likely than Columbus, Phoenix, and Broward County riders.

Table 3.44: Access Mode by Egress Mode by Study

Region	Walk-to-Walk (Wheelchair/ Bicycle)	Walk-to-Auto (Drive, Carpool, Taxi, Dropped off)	Auto-to-Walk (Drive, Carpool, Taxi, Dropped off)	Auto-to-Auto (Drive, Carpool, Taxi, Dropped off)	Total
SEMCOG	84.0%	3.0%	8.0%	5.0%	100%
Greater Detroit	84.3%	3.4%	8.9%	3.4%	100%
Columbus	88.0%	6.0%	5.0%	3.0%	100%
Charlotte	75.0%	9.0%	12.0%	4.0%	100%
Baltimore	76.0%	12.0%	9.0%	3.0%	100%
Phoenix	85.0%	4.0%	9.0%	2.0%	100%
Broward County	90.5%	2.3%	4.6%	2.6%	100%

SEMCOG riders are less likely to use a pass, other than a day pass. Within SEMCOG area, 18.7 percent of riders use a pass, while 23 percent of riders surveyed in the City of Detroit use a pass.

Table 3.45: Fare Paid by Study*

Region	Cash/ Day Pass	Other Pass	Transfer	University	Other	Totals
SEMCOG	48.4%	18.7%	11.8%	21.0%	1.6%	100%
Greater Detroit	62.1%	23.0%	15.3%	0.2%	1.3%	100%
Denver	35.0%	52.0%	4.0%	NA	9%	100%
Columbus	49.0%	32.1%	2.7%	8.2%	8.0%	100%
Baltimore	36.0%	26.0%	NA	2.0%	36.0%	100%
Dallas	25.0%	57.0%	NA	NA	18.0%	100%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

Denver has the highest percentage of student (K-12) fares, at 15 percent, while SEMCOG has the highest percentage of senior/disables fares, at 7.5 percent.

Table 3.46: Fare Type by Study

Region	Regular	Student(K-12)*	Senior/Disabled	Other	Totals
SEMCOG	86.0%	6.5%	7.5%	NA	100%
Greater Detroit	87.2%	6.2%	6.6%	NA	100%
Denver	81.0%	15.0%	NA	4.0%	100%
Columbus	78.0%	4.0%	5.0%	13.0%	100%
Baltimore	89.0%	4.0%	5.0%	2.0%	100%

**Information obtained from riders deemed surveyable. For SEMCOG, that was 16 and up.*
Other regions may have included younger respondents.

Compared to Denver and Baltimore, SEMCOG has a lower rate of employer-provided transit fare subsidies.

Table 3.47: Fare Subsidy (Employer Pay) by Study

Region	Yes Entire Fare	Yes Part of Fare	No	Totals
SEMCOG	8.2%	1.9%	89.9%	100%
Greater Detroit	6.5%	1.7%	91.7%	100%
Denver	13.0%	8.0%	79.0%	100%
Baltimore	7.0%	9.0%	84.0%	100%

Compared to Columbus, Baltimore, and Dallas, SEMCOG and Greater Detroit have higher rates of trips that are made more than five days per week, at 25 percent and 26.1 percent. SEMCOG also has a high rate of trips that are made less than four days per month, but this does not apply to Greater Detroit.

Table 3.48: Trip Frequency by Study

Region	6-7 days per week	1-5 days per week	Less than 4 days per month	Totals
SEMCOG	25.0%	60.9%	14.1%	100%
Greater Detroit	26.1%	59.5%	10.9%	100%
Columbus	20.0%	70.0%	10.0%	100%
Baltimore	14.0%	75.0%	11.0%	100%
Dallas	16.0%	69.0%	15.0%	100%

Baltimore and Dallas have the highest rates of transit riders who would use a vehicle if transit were not available, at 63 percent and 61 percent, respectively. Forty-seven percent of SEMCOG riders reported they would use a vehicle, which is similar to percentages in both Denver and Charlotte.

Table 3.49: Alternative Travel Mode by Study

Region	Walk/Wheelchair Bicycle	Auto (drive, carpool, dropped off)	Taxi/Other	Would not make this trip	Totals
SEMCOG	27.8%	47.0%	13.0%	30.5%	100%
Greater Detroit	22.6%	47.5%	14.5%	34.1%	100%
Denver	33.0%	47.0%	8.0%	22.0%	100%
Charlotte	16.0%	49.0%	7.0%	28.0%	100%
Baltimore	4.0%	63.0%	15.0%	18.0%	100%
Phoenix	25.0%	34.0%	11.0%	30.0%	100%
Broward County	25.3%	40.5%	13.0%	33.7%	100%
Dallas	16.0%	61.0%	2.0%	21.0%	100%

Typically, close to half of transit riders do not have a valid driver's license. Riders on the Greater Detroit systems are even more likely not to have a license, at 57.7 percent.

Table 3.50: Valid Driver's License by Study

Region	Yes	No	Totals
SEMCOG	53.7%	46.3%	100%
Greater Detroit	42.3%	57.7%	100%
Denver	60.0%	40.0%	100%
Columbus	48.0%	52.0%	100%
Charlotte	50.0%	50.0%	100%
Baltimore	51.0%	49.0%	100%
Phoenix	48.0%	52.0%	100%
Broward County	54.0%	46.0%	100%

SEMCOG has the lowest rate of riders 65 years of age and older, at 1.9 percent, while Broward County has the highest percentage of riders in this age group, at 7.7 percent.

Table 3.51: Age by Study

Region	Under 18	18-64	65 +	Totals
SEMCOG (16 and older)	5.6%	92.5%	1.9%	100%
Greater Detroit (16 and older)	6.7%	91.3%	2.0%	100%
Columbus (11 and older)	5.0%	93.0%	2.0%	100%
Baltimore (11 and older)	5.0%	92.0%	3.0%	100%
Phoenix (16 and older)	10.0%	88.0%	2.0%	100%
Broward County (16 and older)	2.1%	90.2%	7.7%	100%

SEMCOG has the lowest rate of employed riders, at 53.2 percent. This number is slightly higher for Greater Detroit, at 58 percent, but still lower than Columbus, Charlotte, Phoenix, Broward County, and Dallas.

Table 3.52: Worker/Student Status by Study*

Region	Full time/Part time worker	Student	Un-employed	Other	Totals
SEMCOG	53.2%	38.2%	16.3%	12.6%	100%
Greater Detroit	58.0%	29.6%	20.2%	11.4%	100%
Columbus	71.0%	19.0%	17.0%	9.0%	100%
Charlotte	71.0%	21.0%	26.0%	3.0%	100%
Phoenix	71.0%	27.0%	25.0%	4.0%	100%
Broward County	70.3%	6.5%	17.8%	12.3%	100%
Dallas	69.0%	16.0%	9.0%	6.0%	100%

**Select all that apply.*
Figures may not add up to totals due to multiple responses per record.

Typically, about half of transit riders do not own a household vehicle. However riders in Greater Detroit have the highest percentage of riders without a vehicle, at 57.9 percent.

Table 3.53: Household Vehicles by Study

Region	None	One	Two	Three or more	Totals
SEMCOG	51.9%	29.3%	12.4%	6.4%	100%
Greater Detroit	57.9%	28.7%	10.1%	3.3%	100%
Denver	29.0%	29.0%	26.0%	16.0%	100%
Columbus	48.0%	31.0%	14.0%	7.0%	100%
Charlotte	49.0%	28.0%	17.0%	6.0%	100%
Baltimore	49.0%	26.0%	17.0%	8.0%	100%
Phoenix	51.0%	27.0%	15.0%	7.0%	100%
Broward County	51.4%	31.8%	12.5%	4.3%	100%
Dallas	40.0%	31.0%	21.0%	8.0%	100%

Columbus riders are the most likely to be from one-person households.

Table 3.54: Household Size by Study

Region	One	Two	Three	Four or more	Totals
SEMCOG	19.5%	26.9%	20.2%	33.4%	100%
Greater Detroit	18.7%	24.7%	22.0%	34.6%	100%
Denver	21.0%	33.0%	18.0%	28.0%	100%
Columbus	26.0%	30.0%	18.0%	26.0%	100%
Charlotte	19.0%	28.0%	22.0%	31.0%	100%
Phoenix	17.0%	26.0%	20.0%	37.0%	100%
Broward County	18.0%	26.0%	18.0%	37.0%	100%
Dallas	20.0%	27.0%	19.0%	34.0%	100%

SEMCOG riders are the most likely to be from zero-worker households, at 27.6 percent, while the other regions range from 8–16.9 percent.

Table 3.55: Household Employees by Study

Region	None	One	Two	Three or More	Totals
SEMCOG	27.6%	37.9%	25.0%	9.5%	100%
Greater Detroit	27.7%	40.0%	23.8%	8.6%	100%
Denver	8.0%	38.0%	38.0%	16.0%	100%
Columbus	12.0%	47.0%	28.0%	13.0%	100%
Baltimore	14.0%	33.0%	34.0%	19.0%	100%
Broward County	16.9%	36.0%	27.8%	19.3%	100%

SEMCOG riders are more likely to be in the lowest income range than any other region in the comparison.

Table 3.56: Income by Study

Region	Less than $19,999	$20,000–$49,999	$50,000–$74,999	More than $75,000	Totals
SEMCOG	54.3%	31.7%	9.7%	4.3%	100%
Greater Detroit	57.1%	34.4%	7.2%	1.3%	100%
Denver	32.0%	32.0%	17.0%	19.0%	100%
Columbus	49.0%	32.0%	12.0%	7.0%	100%
Charlotte	44.0%	36.0%	8.0%	12.0%	100%
Baltimore	35.0%	38.0%	13.0%	14.0%	100%
Phoenix	47.0%	39.0%	7.0%	7.0%	100%
Dallas	42.0%	37.0%	10.0%	11.0%	100%

4. Major Findings and Recommendations

The SEMCOG On-board Survey is a large scale survey and data collection project. Section 2 presented survey procedures and the efforts in order to ensure a successful data collection. This section summarizes major achievements and findings from the survey as well as recommendations for the future.

4.1 Major Achievements and Innovations

Most of the steps in the On-Board Survey data collection follow traditional procedures that were similar to the SEMCOG 2001 on-board transit survey. Meanwhile, for an optimal final product, innovations were also made to improve survey collection efficiency and data quality.

Web Based Survey vs. Self-Administrated Survey

At the pilot test stage, NuStats tested a method of web-based data collection to mitigate a common problem in surveys of this type: riders who make short trips are not able to complete a questionnaire during their short trip. Because students spend more time online and tend to be more Internet savvy, this method was tested on riders of the UM shuttle system on campus. Unfortunately, this method did not provide positive results; in fact, the self-administered version performed better on UM routes than the Web-based version. Going forward, Web-based surveys should only be employed if the invitation is issued as an email blast to the respondents. Asking a rider to visit a Web site at a later point is not a suitable method to conduct this type of data collection.

Cognitive Interview

Also at the pilot test stage, cognitive interviews were conducted to better assess the motivations for riders to participate in a study of this type. The topics discussed included incentives, appropriate introductions to the potential participant, and different media sources used for information about the transit system. Another important part of the interviews was to determine the most comprehensible way to display the questions within the instrument so that it made the most sense to passengers. To accomplish this, the recruited transit riders were asked to complete the questionnaire, after which the moderator went through the questionnaire to determine the thought process behind each response. Using responses and finding from the interviewers, the final questionnaire was revised so as to appeal to the majority of the transit users.

Sample Weighting Procedure

The most impressive achievement during the survey was the development of a sample weighting procedure that was uniquely constructed based on the data available. The original plan was to develop weighting factors for each trip based on boarding location only approach, however, in recent years, the sample weighting process has been gradually transitioning from this simpler calculation to a more sophisticated procedure that considers both boarding and the alighting locations. FTA strongly recommended this emerging approach as it provides a much better way to reserve transit travel patterns.

An iterative proportional fitting (IPF) method was used in the SEMCOG on-board transit survey to calculate weighting factors that count for both boarding and alighting activities. In SEMCOG region, the transit service covers a very large area, and the travel pattern, service focus, route density, land use characteristics and stop level activities along the transit lines vary widely. In order to effectively perform an IPF based sample weighting process, including implementing the time-of-day factors, the survey records have to be aggregated taking the direction and time-of-day factors into consideration.

Sample Aggregation Process

The input to the IPF process has to meet certain requirement, such as number of boardings and alightings per stop location, overall service level, etc. There are about 14,000 active stop locations in the survey records, and these numbers are far exceeded an IPF program can reasonable handle. SEMCOG and NuStats developed and implemented a practical approach so the original survey samples can be aggregated to the level that not only accommodated the IPF process but also optimally preserved travel patterns. The survey team believes this unique approach balanced the sample expansion needs for preserving the spatial travel pattern, land use characteristics, and the temporal distribution of the trips.

To implement the land use characteristics in the stop level aggregation, SEMCOG developed a unique Land Use Index (LUI) measure that reflected population, employment, and activity center locations along the transit lines. For activity centers, SEMCOG considered hospital/medical facilities, universities, and the size of retail businesses.

Depending on route level boarding activity levels, three aggregation schemes were defined and used. Along with other criteria, the stops with similar LUI were grouped together and the spatial distribution of travel patterns was better preserved later in the IPF process.

Use of Incentives

The data collection efforts overall were successful. One major issue was the low response rates for some of the DDOT high-volume routes. This is partially due to inconsistency of the DDOT service. A large number of assignments were missed and had to be sent out multiple times because some buses were unable to operate as scheduled. Meanwhile, the service areas have a high concentration of captive riders with very low income. In order to combat the low response rate encountered during the first portion of the data collection, free ride tickets were purchased for the seven DDOT routes as incentives for completing a questionnaire for the second portion of the data collection. The incentive use was very effective and more than doubled the response to the survey, increasing overall response on the seven routes from 10 percent to 23 percent, as shown in Table 4.1 below.

Table 4.1: Response Rates on Incentivized Routes

Route	No Incentive			Incentive			Response Rate Increase
	Eligible Boardings	Completed Records	Response	Eligible Boardings	Completed Records	Response	
DOT-016	4,583	478	10%	1,250	254	20%	95%
DOT-018	1,604	222	14%	198	87	44%	217%
DOT-032	2,164	275	13%	495	112	23%	78%
DOT-034	3,256	420	13%	340	67	20%	53%
DOT-041	1,085	107	10%	142	47	33%	236%
DOT-045	4,154	269	6%	1,020	225	22%	241%
DOT-053	7,010	693	10%	3,445	792	23%	133%
Total	23,856	2,464	10%	6,890	1,584	23%	123%

4.2 Data Analysis and Findings

The survey team conducted an initial analysis using the fully weighted and expanded data set. Four separate data summaries and analysis were developed to highlight the major travel characteristics in the region:

- Data summaries and analysis at SEMCOG regional level
- Data summaries and analysis for each of seven service provider,
- Data summaries and analysis for three major regional transit corridors (Woodward Avenue, Michigan Avenue, and Gratiot Avenue), and
- A regional comparison to other major metropolitan areas where NuStats has recently conducted surveys

The analysis focused in four areas: passenger attitudes toward the transit services, transit traveler's demographics, transit travel patterns and service coverage/quality.

For the service coverage, the region's bus system serves about 222,000 boardings every day, with about half of transit usage on the 10% of the routes in the system. In the region, there were 14,000 active bus stops surveyed and 800 of these (6 percent) carried 50% of daily regional ridership. 52% of riders reported making no transfers to complete their trips, while 36% made one transfer.

Majority of the transit riders in the SEMCOG region are transit dependent. Nearly 52% of riders did not have access to a vehicle on the survey day and this number increased to 60% in the DDOT area, which covers more than 50% of the regional daily riders. 46% percent of riders in SEMCOG region did not have a valid driver's license, and furthermore, 20% of riders surveyed in the region are unemployed.

The transit system primarily serves people with lower incomes. The survey found 86% of riders were from households with an annual income of $50,000 or less, and 40% of riders were from households making less than $10,000 annually.

Since majority of the transit riders are captive riders, improving existing service is their primary concern. Nearly 40 percent of those surveyed would like service to be more frequent; one-third would like extended service hours (earlier start/later end). The transit service is essential to riders, as 26% of the riders were not able to make the surveyed trip if the service were not available.

For travel characteristics, the vast majority of transit trips made by riders either begin or end at home (84 percent), and 54% of riders used transit for work/university-related purposes. 84% respondents walked to access or egress the bus. 75% of riders were frequent riders, defined as riding 3–7 days per week.

4.3 Recommendations and Future Improvement

Improving on-board transit survey methodology is always one of the FTA focuses. In recent years, a better approach to obtain a more precise spatial and temporal distribution from survey records has been FTA's focus. Under FTA's recommendations, the survey team implemented a new sample design approach, adopted an IPF based sample expansion method, and addressed time-of-day distribution of the survey records. The final expanded survey data set was a better balance in preserving both spatial and temporal characteristics of travel behavior. However, the expanded survey data set still has its limitation due to lacking of benchmark data.

Two areas of improvement are recommended for subsequent on-board surveys. The first addresses the need for better ridership figures when developing the sample plan and weighting procedure. The second relates to the way the on-board survey data are collected, given the know issues and biases with self-administered questionnaires.

4.3.1 System-wide Stop Level Boarding/Alighting Counts

Lacking of stop level ridership control totals for boarding and alighting are a major limitation in this survey, this is especially true for developing the sample plan and weighting the data. When this study was initially designed, response rates were going to be collected using only the boarding locations for each questionnaire returned on each trip. As the survey design evolved, it was determined that the alighting location needed to be included as well. This improved weighting style has recently been developed and is endorsed by FTA.

For future studies, if resources allow, a system-wide collection of boarding and alightings counts for all routes, or at least major routes, by stop, either as pure counts or the more desirably pairs or flows, collected prior to the on-board survey, will allow for a more detailed understanding of passenger travel patterns. These counts can be used to develop the sample plan using more disaggregate units when attaching survey goals. More specifically, goals can be developed with not only the system and route in mind, but also with regard to direction, time of day, and bus stop or bus stop segment. Constructing goals in this way will create a de facto weighting and expansion plan wherein each cell (system, route, direction, time of day, bus stop segment on, and bus stop segment off) would represent a goal for collection in the field and a weight for each record collected within each cell.

4.3.2 Intercept Interview

While self-administered questionnaires have been the industry standard for the past two decades, the trend, based on FTA direction and client/consultant data needs, is to move to an intercept interview (similar to what was conducted on the People Mover for this survey). An intercept interview is one in which a surveyor boards a transit vehicle and, in a random fashion, asks riders to participate in the survey. Once the rider has agreed to participate, the surveyor collects the O/D and demographic information, taking between 4–6 minutes to conduct the interview. While paper and pencil is an option to conduct these interviews, the more prudent means of obtaining these data is through a Tablet PC. Not only would the geographic data be collected such that the locations can be validated, it would allow for the scrutiny of route-to-route transfers based on routes intersecting or running in a close proximity. Furthermore, this method of collecting data allows for the dataset to be constructed in near real time by removing the need for scanning in paper questionnaires, geocoding the location data, and fully processing the data.

In addition, and most importantly, intercept interviews would allow for an increased response rate and a minimization of potential biases introduced into the dataset. Traditionally, riders who are unable to complete the questionnaire—whether because they are making very short trips or are illiterate or have low literacy skills, among other reasons—are traditionally underrepresented in studies of this kind. As intercept interviews are shorter than self-administered questionnaires and simply because they require a verbal exchange between the surveyor and respondent, they potentially allow for a more representative dataset.

Finally, when the boarding and alighting pair's distribution are collected and merged with the intercept interviews, weighting is accomplished because the data are collected with specific regard to the population distribution. Data collection would be able to be monitored in real time to ensure the distributions collected matched the distribution of the population of riders.

Appendices

Appendix A: Pilot Study Analysis

Appendix B: Cognitive Interviews Report

Appendix C: Pilot Findings and Full-Scale Data Collection Preparations

Appendix D: Frequently Asked Questions

Appendix E: Questionnaire

Appendix F: Data Weighting & Expansion

Appendix G: Assignment Report

Southeast Michigan Council of Governments

Regional On-Board Transit Survey

Pilot Test Analysis

August 16, 2010

NuStats a ptv group company

206 Wild Basin Rd., Suite A-300
Austin, Texas 78746
Contact: Fred G'sell, Project Manager
(512) 306-9065
Fax (512) 306-9077

www.nustats.com

Table of Contents

List of Tables and Figures

1.0 Introduction

Background

NuStats conducted a pilot study for the upcoming SEMCOG Regional On-Board Transit Survey from April 19–April 28, 2010. The purpose of the pilot was to gain insight on potential issues that could arise during the full-scale data collection and limit these occurrences. The pilot test is designed to see what works well and what does not work well. For those items that do not work well, the pilot data and anecdotal information from the pilot are used to make improvements upon the methods used for the full-scale data collection. In addition, the pilot also allows the opportunity to work with the individual transit providers to insure that the full-scale data collection logistical issues are minimized by getting familiar with their facilities, people and bus schedules.

The pilot effort consisted of surveying routes from Detroit Department of Transportation (DDOT), Suburban Mobility Authority for Regional Transportation (SMART), Detroit People Mover (People Mover), Ann Arbor Transportation Authority (AATA), and University of Michigan Transit (U of M). For the purposes of the pilot, Blue Water Area Transportation Commission (BWATC) and Lake Eric Transit Commission (LETC) were not studied because of their relatively low ridership figures in comparison to the other listed systems.

For DDOT and SMART, NuStats tested a self-administered questionnaire. For these two systems, the focus was on survey and system logistics. For the People Mover, AATA, and U of M, one focus for the pilot was survey and system logistics, but NuStats also tested a second methodology: Web-based data collection. The rationale behind using this second method of collection is because of the relatively short trip times for passengers, especially for People Mover and U of M. For the routes in these systems, half of the surveyed trips distributed the self-administered questionnaire from DDOT and SMART, and the other half distributed survey cards inviting passengers to take the survey online using their computer or smartphone.

2.0 Survey Administration

Below is a break down, system by system, anecdotally describing the pilot test and lessons learned for the full-scale study.

Detroit Department of Transportation (DDOT)

For the DDOT portion of the pilot data collection, the command center was based out of DDOT garages. This did not work particularly well for various reasons, including logistics and lack of efficiencies in surveying labor. Because many of the routes use the Rosa Parks Transit Center (RPTC), NuStats would like to set up its command center in or around the RPTC and will work with DDOT to secure a location for the full-scale data collection. The central location of this terminal would allow surveyors assignments to begin and end at the same location, thus limiting non-productive surveyor time spent riding on the vehicles during non-revenue generating time. Assignments on routes that do not access RPTC would have to begin at other locations, but this would still be more efficient than using a DDOT garage.

Overall, DDOT passengers did not participate at a high level; therefore, these passengers became a focus for the cognitive interviews, as DDOT is the largest system in the study and is thought to be difficult based on perceived ridership. In addition, this system will be the primary focus of the public outreach program.

Suburban Mobility Authority for Regional Transportation (SMART)

For the SMART portion of the pilot data collection, the command center was based out of the garages as well. Because of the nature of SMART service, highly interlined routes that often run only during peak times, the garage location makes sense for the full-scale data collection. There were issues with the "run numbers" during the pilot that will have to be accounted for during the full-scale data collection when cutting assignments beginning and ending at the garages. Using the garages to begin and end assignments will allow the surveyors to collect data over the course of the entire driver "run" and get the surveyors back to the original assignment beginning location.

Overall, SMART passengers participated in the study at a high level; as such, the same methodology will be employed for the full-scale data collection.

Detroit People Mover (People Mover)

For the People Mover portion of the pilot data collection, the command center was informally located near RPTC. The data collection for this system was very difficult for multiple reasons. First, and as expected, the lengths of trips made by passengers are extremely short and limit their ability to complete the questionnaire. Second, because the distance and time between stops is so small, it became difficult to accurately count the number of boarding and alighting passengers while accurately attributing the correct boarding location to the boarding rider's questionnaires. As a result, the majority of the counters' PDA files were incorrect. In addition, the day the People Mover was surveyed was "bring your kids to work day," which created higher ridership than would probably be encountered over a typical weekday.

Based on this anecdotal information and the response rates reported later in this report, alternate methodologies may be employed for the full-scale data collection, including platform-based surveying and using station counts rather than PDA counts.

Ann Arbor Transportation Authority (AATA)

For the AATA portion of the pilot data collection, the command center was based out of the garages. While this worked for the pilot, NuStats recommends, for surveying efficiency, using a more centralized location for the full-scale data collection effort. While there is not any space in the Blake Transit Center, space near this location would be ideal. NuStats is looking into a potential short-term lease in the area, but would appreciate and consider any other recommendations. Additionally, some routes had sub-route designations. For example, Route 4 had Sub-routes of 4a, 4b, and 4c. This type of information will be valuable for the full-scale data collection logistics planning.

The ridership on this system was high. Many of the passengers encountered were very happy with the bus system and appreciated the service provided by AATA. Therefore, no methodological changes are recommended for the full-scale data collection.

University of Michigan Transit (U of M)

For the U of M portion of the pilot data collection, the command center was informally based along the routes in the system. Ideally, this full-scale data collection will be paired with the AATA data collection for optimum command center efficiency. The U of M system was one of the more difficult for which to collect data because of relatively short trips (either inter or intra campus) and college-aged ridership with limited appreciation for the study purpose.

Overall, the ridership participation was poor. Many of the students did not want to fill out the questionnaires and threw the questionnaires on the floor or left them on the bus seat. The public outreach effort for the U of M routes will also be important for the full-scale data collection. In addition, a different methodology, similar to the People Mover distributing questionnaires at popular bus stops such as Bursley-Baits, may help increase student participation by giving them additional time to complete the questionnaire.

3.0 Analysis

Trips Surveyed Overview

Assignments were generated to cover specific times of the day on different systems and routes. The Mid-day and PM Peak periods were primarily surveyed because each system was only surveyed for a single day, the exception being DDOT. There was a concern that if surveying occurred in the early morning, there would be a potential for too many missed trips because of surveyors being unfamiliar with the various locations. The following tables summarize the number of trips sampled by a preliminary time of day (TOD) definition.

Table 3.1: Number of Trips Sampled by TOD

Time-of-Day	Trips Sampled	% by TOD
AM Peak	1	0.3%
Mid-day	173	52.3%
PM Peak	146	44.1%
Evening/Early Morning	11	3.3%
Total	331	100.0%

For the full-scale study, specific TOD definitions will be utilized from each individual system' definitions, and the surveying will occur over the course of the entire day. The following are the frequency of trips surveyed by system/route:

Table 3.2: Number of Trips Sampled by Route

Route	Trips Sampled	% by Route
AAT-..2	21	6.3%
AAT-..4	16	4.8%
AAT-..8	5	1.5%
AAT-..9	7	2.1%
AAT-.36	23	6.9%
DOT-021	9	2.7%
DOT-022	7	2.1%
DOT-025	10	3.0%
DOT-034	8	2.4%
DOT-036	10	3.0%
DOT-078	9	2.7%
DPM-DPM	123	37.2%
SMT-.112	1	0.3%
SMT-.125	5	1.5%
SMT-.200	7	2.1%
SMT-.245	1	0.3%
SMT-.265	1	0.3%
SMT-.280	2	0.6%
SMT-.805	3	0.9%
UMI-Bursley-Baits	26	7.9%
UMI-Commuter	26	7.9%
UMI-Mitchell-Glazier	11	3.3%
Total	331	100.0%

4.0 Quantitative Results

4.1 Questionnaire Response Rate

Overall, there was a lower-than-average level of participation. Of the total 7,182 passengers who boarded the sampled trips, 1,132 questionnaires were collected, and 877 questionnaires were found to be complete and usable for an overall response rate of 12.2 percent. Out of 1,132 total collected questionnaires, 171 (15.1 percent) were collected by Web, and 120 (10.5 percent) were collected by mail back.

The definition of a complete questionnaire for this analysis included the following variables:

- Origin and destination addresses matched
- Trip purposes
- Access and egress modes
- Route sequence
- Boarding/Alighting of the surveyed route

When determining if a questionnaire is complete, the most common reason for a questionnaire failing was an incomplete or missing destination address, with 10.4 percent of collected surveys failing for this reason. The table below shows the frequency of collected surveys with missing or incomplete key components.

Table 4.1: Non-Response of Key Variables (1,132 Returned Questionnaire)

Question	Count	Percent
Origin Incomplete or Missing (geocoding fail)	78	6.9%
Destination Incomplete or Missing(geocoding fail)	118	10.4%
Origin Purpose Missing	17	1.5%
Destination Purpose Missing	27	2.4%
Access Mode Missing	20	1.8%
Egress Mode Missing	67	5.9%
Route Sequence Invalid	2	0.2%
Control File Incomplete*	17	1.5%
Origin and Destination Same**	29	2.6%

**The boarding and alighting could not be determined for these records.*

***Respondents with the same origin and destination were missing the complete origin to destination trip.*

There was a large variability in response between different routes. The response rate is calculated as the number of returned surveys deemed to be complete and usable divided by the number of boardings. The SMART 805 had the highest response rate: 65.6 percent. Most SMART routes had a higher-than-average response rate, while the People Mover had a very low response rate at 4.1 percent. Table 4.2 shows the response rate of the surveyed routes.

Table 4.2: Number of Completes & Response Rates by Route

Route	Completed	Collected	Boardings	Response Rate**
AAT-..2	95	106	369	25.7%
AAT-..4	74	92	582	12.7%
AAT-..8	25	32	103	24.3%
AAT-..9	37	44	203	18.2%
AAT-.36	33	38	446	7.4%
DOT-021	44	61	355	12.4%
DOT-022	37	52	154	24.0%
DOT-025	70	105	349	20.1%
DOT-034	32	57	263	12.2%
DOT-036	34	40	118	28.8%
DOT-078	9	9	85	10.6%
DPM-DPM	55	77	1378	4.0%
SMT-.112	0	0	3	0.0%
SMT-.125	51	69	181	28.2%
SMT-.200	56	71	285	19.6%
SMT-.245	4	5	11	36.4%
SMT-.265	2	4	12	16.7%
SMT-.280	3	5	21	14.3%
SMT-.805	42	46	64	65.6%
UMI-Bursley-Baits	47	50	1014	4.6%
UMI-Commuter	97	107	1077	9.0%
UMI-Mitchell-Glazier	30	45	109	27.5%
Total	877	1132*	7182	12.2%

** ROUTE was not identified on 17 surveys.*

*** Response rate = # of completes/ # of boardings*

4.2 Individual Question Response Rate

Respondents who completed the required variables typically completed most of the other variables. Of the 877 questionnaires that were deemed to be complete, the highest frequencies of non-response were for the questions asking about system improvement suggestions and household income, at 10.7 percent (94 records) and 7.6 percent (67 records), respectively. Item non-response rates for the non-required questions are shown in Table 4.3.

Table 4.3: Non-Response for Individual Questions (Based on total completes N=877)

Question	Missing Count	Percent Non-Response
Total # of buses/trains	1	0.1%
Age	9	1.0%
Household Size	13	1.5%
Household Workers	17	1.9%
Employment/Student Status	44	5.0%
Household Vehicle	14	1.6%
Car Availability	19	2.2%
Trip Frequency	14	1.6%
Alternative Mode if Bus Unavailable	22	2.5%
Round Trip Purpose	1	0.1%
Household Income	67	7.6%
Improvement Suggestion	94	10.7%

4.3 System Comparison

NuStats surveyed five transit systems in the Detroit/Ann Arbor area. There was a large difference in the response rates between the different transit systems. SMART displayed the highest response rate at 27.4 percent, while the Detroit People Mover had the lowest at 4.0 percent. Table 4.4 shows the response rates for the five surveyed transit systems.

Table 4.4: Number of Completes and Response Rates by Transit System

Transit System	Completed	Collected	Boardings	Response Rate**
AATA	264	312	1,703	15.5%
DDOT	226	324	1,324	17.1%
People Mover	55	77	1,378	4.0%
SMART	158	200	577	27.4%
U of M	174	202	2,200	7.9%
Total	877	1,132*	7,182	12.2%

** ROUTE was not identified on 17 surveys.*

*** Response rate = # of completes/ # of boardings*

It is important to note that the 15.5 percent overall response rate for AATA was greatly influenced by the lower response rate it received from the Web-only option discussed in the later section. This is also the case for U of M routes, but with a smaller impact.

4.4 Survey Method (self-administered vs. Web-based) Comparison

Questionnaire Response Rate

The survey was conducted using two different survey methods—one with a paper-based, self-administered questionnaire and one with a Web-only questionnaire. For the self-administered questionnaire, surveyors offered all passengers who were at least 16 years old a full-length questionnaire. They were asked to complete the questionnaire while on the bus or, if needed, return it in the mail. For the Web-only questionnaire, surveyors distributed a card to passengers who were at least 16 years old and asked them to participate in the survey via an online website. The DDOT and SMART routes were conducted exclusively with paper-based, self-administered questionnaires, and the remaining three transit systems had a portion of trips exclusively surveyed by one or the other. Overall, there was a better response rate with the paper-based surveys, as shown in Table 4.5.

Table 4.5: Response Rate by Method

Method	Trips	Completed	Collected	Boardings	Percent Complete
Web	138	139	171	2,728	5.1%
Paper	193	738	961	4,454	16.6%
Total	331	877	1,132	7,182	12.2%

A common reason for failed records with either survey method was missing or incomplete addresses, but this error was less common in the Web-based version. However, non-response of some of other key variables was slightly higher in the Web-based version, as is shown in Table 4.6.

Table 4.6: Non-Response of Key Variables by Method

Question	Web (N=171)		Paper (N=961)		Total (N=1,132)	
	Missing Count	Percent	Missing Count	Percent	Missing Count	Percent
Origin Incomplete or Missing (geocoding fail)	10	5.8%	68	7.1%	78	6.9%
Destination Incomplete or Missing (geocoding fail)	12	7.0%	106	11.0%	118	10.4%
Origin Purpose Missing	7	4.1%	10	1.0%	17	1.5%
Destination Purpose Missing	10	5.8%	17	1.8%	27	2.4%
Access Mode Missing	6	3.5%	14	1.5%	20	1.8%
Egress Mode Missing	12	7.0%	55	5.7%	67	5.9%
Route Sequence Invalid	1	0.6%	1	0.1%	2	0.2%
Control File Incomplete*	5	2.9%	12	1.2%	17	1.5%
Origin and Destination Same**	5	2.9%	24	2.5%	29	2.6%

**The boarding and alighting could not be determined for these records.*

***Respondents with the same origin and destination were missing the complete origin to destination trip.*

Individual Question Response Rate

In both survey types, respondents who completed the required variables typically completed most of the other variables. However, the usable surveys that were collected on the Web were typically more complete, as shown in Table 4.7.

Table 4.7: Non-Response for Individual Questions by Method

Question	Web (N=139)		Paper (N=738)		Total (N=877)	
	Missing Count	Percent	Missing Count	Percent	Missing Count	Percent
Total # of buses/trains	0	.0%	1	.1%	1	.1%
Age	0	.0%	9	1.2%	9	1.0%
Household Size	0	.0%	13	1.8%	13	1.5%
Household Workers	1	.7%	16	2.2%	17	1.9%
Employment/Student Status	3	2.2%	41	5.6%	44	5.0%
Household Vehicle	1	.7%	13	1.7%	14	1.6%
Car Availability	1	.7%	18	2.4%	19	2.2%
Trip Frequency	1	.7%	13	1.8%	14	1.6%
Alternative Mode if Bus Unavailable	3	2.1%	19	2.6%	22	2.5%
Round Trip Purpose	0	.0%	1	.1%	1	.1%
HH Income	7	5.0%	60	8.1%	67	7.6%
Improvement Suggestion	11	7.9%	83	11.2%	94	10.7%

System Comparisons – Overall

The three systems that were sampled using multiple methods had varying results. AATA and U of M had a much higher response with the paper-based survey, with response rates 3–4 times higher than the Web-based survey. However, the People Mover had an insignificant difference between the two methods, as shown in Table 4.8.

Table 4.8: Response Rate by Method per Transit System

System	Web		Paper		Total	
	Completed	Response Rate	Completed	Response Rate	Completed	Response Rate
AATA	78	6.9%	186	32.3%	264	15.5%
People Mover	32	3.8%	23	4.2%	55	4.0%
U of M	29	3.4%	148	10.3%	174	7.9%
Total	139	5.1%	354	13.9%	493	9.3%

System Comparisons – Trip Length

The average trip length, as measured using straight-line distance between boarding and alighting locations, was 1.7 miles for the Web-based survey and 1.6 miles for the paper based survey. These figures were not statistically significant though.

Table 4.9: Short Trip Analysis (Paper vs. Web) – Distance between Boarding and Alighting Locations (air distance in miles)

Survey Method	N	Mean	Std. Error of Mean	Std. Deviation	Minimum
Web	139	1.7	.14441255	1.702	.10
Paper	354	1.6	.06587349	1.239	.09
Total	493	1.6	.06237724	1.385	.09

The trip duration was calculated using the time of the surveyed boarding stop and the time of the imputed alighting stop. There was a mean of 11.3 minutes for the Web survey and 12.1 for the paper survey, with a larger variance in the Web survey data, as is shown in Table 4.10. While the Web-based trip time is a little shorter, it is not statistically significant.

Table 4.10: Short Trip Analysis (Paper vs. Web) – Trip Duration between Boarding and Alighting Locations (in minutes)

Survey Method	Mean	N	Std. Error of Mean	Std. Deviation	Minimum
Web	11.3	102*	.86652250	8.751	1
Paper	12.1	263*	.43220637	7.009	1
Total	11.9	365*	.39434651	7.534	1

**Trip duration could not be calculated for all cases.*

When comparing the time passengers were on the bus, the Web survey captured a greater percentage of passengers who were on for five minutes or less, with 28.4 percent of the Web surveys falling into this category, and 14.8 percent for the paper survey, as is shown in Table 4.11.

Table 4.11: Short Trip Analysis (Paper vs. Web) – Percentage of Respondents on the bus for Five Minutes or Less

Survey Method	N	5 minutes or less		Greater than 5 minutes	
		Count	Percent	Count	Percent
Web	102	29	28.4%	73	71.6%
Paper	263	39	14.8%	224	85.2%
Total	365	68	13.6%	297	81.4%

**Trip duration could not be calculated for all cases.*

While the Web-based option did not show significant trip distance and time differences overall, it did capture a higher percentage of trips that lasted for five minutes or less; therefore, it can be used for the full-scale study to help combat short trip bias.

An additional question asked passengers who completed the questionnaire on the Web if they completed it using a smartphone or a traditional laptop/desktop computer. Of the 171 questionnaires collected on the Web, only seven passengers (4 percent) used a smartphone, while the other 164 used a more traditional laptop or desktop.

5.0 Final Recommendations for Pilot Study

NuStats and CDM Smith discussed the pilot data set and findings presented above for the purposes of project recommendations for the full-scale data collection. Discussions occurred at the aggregate level, but also for each individual system due to the large variability in the response rates achieved by system.

In general, the desired response rate for the full-scale data collection will be higher than what was achieved during the pilot data collection, though some systems performed well during the pilot. It is expected that the response rates will improve for the full-scale study as surveyors become more proficient with the survey process as they gain more experience over a longer collection period. This, in addition to the efforts that are currently being made to improve the instrument, data collection methods, and public outreach, should allow response rates to improve for the full-scale data collection.

Specifically, SMART and AATA (non-Web version) performed well with response rates of 27.4% and 32.3% respectively. While we hope to make some improvements in response rates for these systems through the previously mentioned changes (more experienced survey teams, improved instrument, and public outreach), the focus for improvements are for the other systems. Therefore, no significant changes are recommended for these systems.

DDOT achieved a 17.1% response rate during the pilot. Knowing that the majority of the data collection will come from DDOT routes, any improvements in response rate will go a long way in making the entire survey successful. In order to get a better grasp on the DDOT ridership perspectives, cognitive interviews were held with bus passenger to attempt to get into the heads of the average DDOT user and make the survey more relevant to each. The cognitive interviews were the biggest source of the changes from the pilot to the current survey instrument. The cognitive interview summary report, pilot and full-scale questionnaire versions are part of the final report and available for future reference. Other than the instrument, the public outreach for DDOT should allow the response rates to improve by making the passengers more comfortable with the survey purpose and promote a civic-minded approach to the study. Banners will posted at the Rosa Parks TC and signs are going to be put on the bus about the study. Information will also be posted on the DDOT/SEMCOG Websites, Facebook and Twitter. In addition, a transit blog and press release will occur in conjunction with the DDOT data collection.

University of Michigan performed below expectation, 10.3% response rate, for the non-Web version. Due to the unique nature of the trips on these systems, very short free trips, the response rate is mostly a product of insufficient time for the average passenger to complete a questionnaire while on their one-way trip. For the full-scale, we are going to offer the paper version, but also give passengers the option to complete the survey via the Web. While this is not going to greatly improve the response rate, it should help somewhat. While public outreach and an improved instrument will increase the response rate, this method may still have to be altered for the full-scale data collection. For these routes, it might be appropriate to interview passengers at popular locations to allow sufficient time to complete the questionnaire. This will be determined once the data collection commences.

The People Mover performed well below expectation as well with a 4.2% response rates for the non-Web version, the lowest among the various systems. Therefore, significant changes are planned for the People Mover full-scale data collection. Platform based intercept interview surveying is going to be utilized for the full-scale study. Surveyors will conduct interviews of passengers and ride with them while on the People Mover increasing the time available to collect complete information. In order to facilitate this type of surveying, non-trip based, weighting will occur using station level counts rather than the trip/bus stop level allowing the passenger to begin the interview prior to the vehicle arriving.

Appendix B: Cognitive Interviews Report

Southeast Michigan Council of Governments

Transit On-Board Study Cognitive Interviews

Summary Report

July 2010

NuStats a ptv group company

206 Wild Basin Rd., Suite A-300
Austin, Texas 78746
Contact: Kimberly Hilsenbeck, *Cognitive Task Leader*
(512) 306-9065
Fax (512) 306-9077

www.nustats.com

Table of Contents

Introduction

Background

NuStats conducted a series of in-person cognitive interviews with transit riders in the Detroit area in advance of conducting the full survey in the fall of 2010.

The main cognitive testing goals were to:

1) Gather feedback on the survey instrument and suggestions for improving the instrument, thus improving data quality,
2) Learn what messages would be motivating and encouraging for participation,
3) Determine the appropriate media outlets for communication, and
4) Provide insight on the need for, preferred amounts of, and types of incentives.

Based on the goals above, NuStats developed a cognitive interview guide to ask questions that would yield the needed information. For Goal 1, a section of questions assessed how easy or difficult it would be for respondents to complete a self-administered questionnaire asking about transit trips, exact addresses, and other information. Respondents completed the questionnaire while NuStats interviewers observed. Interviewers then went through the instrument, question by question, to assess areas of confusion or need for clarity, along with any suggestions for making the questionnaire easier to read and complete.

To accomplish Goal 2, NuStats included questions about reasons for participation and ways to communicate the importance of the survey to riders. Questions about preferred media sources and the most useful methods for communicating the upcoming study to a broad range of riders helped to achieve Goal 3. For Goal 4, NuStats asked questions regarding how important an incentive was for encouraging participation, as well as whether people would be more likely to want bus passes, cash, or a big-ticket item (such as a flat screen TV or a computer).

Methods

NuStats recruited riders at the Rosa Parks Transit Center in downtown Detroit on June 7 and June 8, 2010. The recruiter approached a variety of riders in an attempt to gain cooperation from different types of transit users: women and men, within a range of ages (including younger adults who do not often participate), and of varying backgrounds. Each person heard the introductory script and was notified of the incentive for participating in the cognitive research.

Recruitment for the interviews took place during all hours of the day and evening. NuStats slightly over-recruited the number of interviews needed to allow for those who did not show. A copy of the recruitment script is in Appendix A.

On June 9 and 10, 2010, NuStats conducted 23 interviews at the Renaissance Conference Center in downtown Detroit. Interviews lasted approximately 60 minutes, and each person received $75 for his or her time.

One of our senior research associates developed the interview guide, an English copy of which is in Appendix B. She and two other experienced interviewers from NuStats conducted the interviews using the approved interview guide.

Findings

The basic structure for the interviews was to provide an introduction and goal of the interview, then present the survey instrument and ask the respondent to complete it. During the introduction, it was important to try to get the respondent to think back to his or her most recent and preferably “regular” bus trip, which for many respondents was the home-to-work commute. Previous testing has indicated that asking about their trip to the interview leads to confusion, as some people mix the two trips (their usual trip and the trip to the interview).

NuStats understands that testing of this nature introduces an artificial element since participants are not on a bus, and they agreed to attend the interview knowing they would receive $75 as compensation for their time. However, the real value in conducting in-person interviews with actual transit riders, even under contrived conditions, is the ability to probe on the various issues encountered or points of confusion. More importantly, it is an opportunity to learn about ideas for how to improve the survey instrument from the perspective of the respondent, thus increasing participation rates as well as the quality of the data. Interviewers concentrated their efforts on gaining insight about motivations for taking the survey, messaging strategies that may help increase participation, and preferred media outreach venues. A relatively significant section of the interview concentrated on the incentive structure.

Feedback on Survey Instrument

Completing an on-board questionnaire can be confusing for some respondents, in part because they are asked to do so under less than optimal conditions—often while on a crowded, moving vehicle with low lighting and bumpy roads; some with limited cognitive comprehension, literacy challenges, or sight/vision impairments; some with children in tow; and some without enough time if on a short bus ride. Yet the data collected in such studies is critical for transit planners and modelers. NuStats' testing brought attention to several issues respondents had in completing the questionnaire, even under much more optimal surroundings. The problems they encountered included:

Did not know exact addresses – Some participants simply did not know exact addresses, even of habitual destinations; in some cases, they left this information blank. Most provided cross streets when they left the address blank.

Did not understand what to write for Place Name – This caused quite a bit of confusion for many respondents. They were unsure of what to write, in particular if they were coming from home as they had indicated in the previous question. Most respondents would leave the Place Name question unanswered or give incorrect information. Several people put the name of the place they were going, which in turn caused the rest of their answers to be out of order.

Concept of a one-way trip was confusing – For example, some people had questions about start and end points; the example provided may have contributed to confusion for several people as it began at work and ended at home. Several respondents felt the example should begin at home and end at work.[1] A few respondents attempted to provide round-trip information, though this was not prevalent among the majority of respondents. A female respondent started completing the instrument about one trip, then switched to another. Interestingly, when one of the interviewers had respondents draw/write out their one-way trip in order to visualize it, the concept of a one-way trip seemed to make more sense.

[1] Previous testing showed that when ‘home to work’ is used as an example of a one-way trip, respondents have a higher likelihood of trying to report on their home to work trip rather than the current one-way trip they are taking when handed the questionnaire.

Did not look at or did not understand visual examples – While some respondents said they found the visual example of a one-way trip helpful, several admitted to skipping over the example or not understanding it. The example of a one-way trip was confusing to some respondents.

Did not always know bus route numbers – Respondents are not always able to give a bus route number. They seem to be able to easily write in the bus route name, e.g., Mack, Gratiot, and Jefferson.

Concept of a prize drawing was mostly favorable, but some were leery – For the most part, respondents reacted positively to the idea of a prize drawing. A few felt it was or would be perceived as a scam, yet even those respondents liked the idea and felt it was important to offer an incentive.

Bus passes, cash and, big-ticket items are appealing – Many respondents felt other riders would need an incentive to participate. They suggested bus passes, cash, or a big-ticket item.

Motivation and Encouragement to Take the Survey

NuStats knows from experience, and from conducting the pilot test for SEMCOG, that some respondents simply will not complete an on-board survey questionnaire. Others will take the paper questionnaire but will never complete it. Still, others may take and complete the survey, but never return it to NuStats. Therefore, using the cognitive testing to probe deeper on attitudes, motivations, and barriers with the goal of uncovering useful insights that would help increase survey participation rates was important.

NuStats interviewers questioned respondents on possible message strategies for encouraging survey participation. The goal was to uncover reasons and motivations for accepting, completing, and returning an on-board survey questionnaire.

Nearly all of the respondents interviewed said they would complete a questionnaire on the bus (keeping in mind, of course, that respondents likely felt obligated to say this); though notably, a few said they would not (but they may take the questionnaire and mail it back later). Completing a questionnaire on the bus would furthermore depend on several other factors, such as how crowded the bus was at the time, the length of their trip, and even their mood at the time. In addition, the surveyor's presentation of the instrument was very important: while respondents did not suggest formal dress, they did prefer a casual yet neat dress for an adequate comfort level. Additionally, the attitude and communication skills of the surveyor were paramount in getting riders to participate by speaking non-slang English, being polite, and not being too pushy. A few were wary of providing personal information. Some respondents said there would be riders who will not take the survey "no matter what." even with the offer of an incentive. In large part, respondents felt those who would not participate did not care much about the bus system, possibly because they were not regular transit users. Respondents who participated in the testing consistently said they care about the bus system.

The interviewers asked what respondents recommended in terms of motivating riders to take the survey. Some ideas included:

1) Let riders know taking the survey will benefit them and the bus they ride; i.e., personalize the message. Surveyors on the buses should say, "Take this survey to help the 220 bus." Or, "Take this survey to make your bus ride better."
2) Provide information that helping with the survey will help make system-wide transit improvements. In this sense, NuStats could appeal to the altruistic side of riders by explaining how the survey will help transit planners and the bus agencies involved to improve transit.

Media Sources and Communication Outreach

NuStats interviewers asked respondents about their preferences for receiving news and information, and to share their ideas for the best outreach communication venues to promote the survey. NuStats found that most respondents receive their news and information from TV, newspapers, and the Web, as well as word of mouth from friends, relatives, and neighbors. Suggestions for spreading the word about the survey included putting ads inside buses, TV news, hanging signs/flyers at bus stations and the downtown transfer center, and using community/grassroots organizations to help with word of mouth messaging, like Goodwill Industries, Moses, and Transportation Riders United.

Incentive Preferences

The drawing for 10 $100 cash prizes appealed to many participants, although a few said it was not necessary to include; in other words, riders who care about the bus system would complete the survey even without an incentive. Some respondents felt it may be perceived as a fraud or that people had little chance of actually winning. A handful of respondents did not see the information about the drawing while they filled out the survey, which indicates it was not visually appealing enough to catch their attention. In a few cases, however, interviewers noticed that rather than attributing this to a lack of visual appeal, some respondents did not read the instruction section.

NuStats used the cognitive interviewing to test the concept of offering a non-monetary/cash prize, such as monthly bus passes or big-ticket items such as a flat screen TV or computer. This section of the summary provides details on the results of that approach.

Several people suggested using a monthly bus pass for the drawing prizes as opposed to money; one-day passes were less appealing. Respondents who preferred bus passes liked the appeal of this instant gratification rather than being entered into a drawing they likely would not win (though currently, instant gratifications are not going to be employed). For those who preferred a monetary prize to a pass, the chance for more people to win (even if it was a smaller cash amount) was appealing. Some respondents favored a big-ticket item over cash or bus passes. Items such as a flat screen TV or lap top computer have broad appeal to a large audience, though other may view this as a poor use of funds.

Recommended Edits to Survey Instrument

NuStats recommends changes to the survey instrument using the pilot instrument, based on the input of transit riders, as follows (*Note: not all suggested edits will be practical or able to be implemented*):

1) Design the instrument in color, or use fonts somehow throughout to call attention to particular areas.
2) Modify the one-way trip visual example; use generic terms such as Start and End, Origin and Destination, or similar text.
3) Q2. Total number of buses used was a little confusing, add additional detail.
4) Q3. Add, “List each route number/name” to the instrument.
5) Q6a. This caused confusion for several people. NuStats recommends the addition of a simple line of text: “If you are not coming from home, what is the name of this place?” (*Note: if respondents incorrectly answer Q6A, they also often miss Q6b.*)
6) Q8a. This issue is similar to Q6a above. Add text: “If you are not going home…”
7) Q9. This question may seem redundant to respondents who are getting off directly at or very near their destination. For example, if the trip ends at Main and Broad Streets and their destination is at that intersection, they write the same streets twice.
8) Q15. Add / change choices: “Unemployed, but seeking work,” “Unemployed, not seeing work,” and “Retired.”
9) Q20. This caused some confusion among respondents. Recommend deleting.
10) Q22. Move this to the front of the instrument; this shows riders the transit system cares enough to ask about service improvements.

Recommended Public Outreach

NuStats, CDM Smith, and SEMCOG will determine a specific public outreach plan based on the feedback from the cognitive interviews, outreach that has been conducted by SEMCOG and the various transit agencies previously, and public transit groups, i.e. Transportation United. A meeting of the Transportation Coalition on August 18th will be the starting point for coordination between the various agencies insuring a collective effort in the public outreach.

Appendix A: Recruitment Script

Hi, I am talking to folks who use the bus system to ask their opinions about an upcoming survey. It is a great way to share your input and earn $75 for your time. This Wednesday and Thursday, we are inviting people to participate in an hour-long interview at the Renaissance Center on Jefferson Street.

There are only a few interviews each day. What time would work best for you? SHOW THE SCHEDULE AND GET THEM TO DECIDE WHICH TIME SLOT IS BEST.

Appendix B: Interview Guide

Introduction 2 minutes

- Interviewer introduces self
- Statement of purpose / Study objectives: We want your opinions about the survey questionnaire.
- Ground rules – honesty, don't be afraid to ask questions or speak your mind.
- Audiotaping our conversation to write the report.

Past experience with transit surveys (ice breaker) 5 minutes

As a transit rider, can you recall a time when someone asked you to complete a survey while riding on the bus or train?
How did you react? Did you take the survey?
What do you recall about what kinds of questions were asked?

Completing the Questionnaire 20 minutes

Think about the most recent bus trip you took in the past few days or week.

I'd like you to pretend you're on the bus and someone just handed you this questionnaire. I'd like you to take a few minutes to fill out the survey, but as you're doing this, tell me out loud what you're thinking. I want to hear about what's going through your mind as you fill it out.

PAY ATTENTION TO THE FOLLOWING: DID THEY LOOK AT THE EXAMPLE? DID THEY READ THE INSTRUCTIONS? DID THEY HAVE A PUZZLED LOOK AT ANY POINT DURING THE TIME THEY FILLED IT OUT?

WHEN THEY'RE DONE ASK: Were the instructions clear? Was there anything confusing about the survey?

GO THROUGH THEIR ANSWERS WITH THEM, POINT OUT AREAS THAT ARE INCOMPLETE OR MISSIONG AND DISCUSS

Examples: I noticed this section is blank. Let's talk about this.
It seemed like you had a puzzled look when you completed this section. I'd like to hear your thoughts about it.

Now I have a second questionnaire. It has the same basic questions but looks a little different from the other one. I'd like you to please take a few minutes to complete this survey, and again, tell me out loud what's happening in your head as you read through and answer it.

AGAIN PROBE ON AREAS OF CONFUSION
Tell me about this questionnaire. Things you didn't like as much? Were the instructions clear?
GET SPEFIFICS: language choices, visual appeal, confusing terms, size, etc.

Motivation for Taking Survey 15 minutes

If someone handed you a survey while you were riding a bus or train, how might you react? Would you take the questionnaire? Would you fill it out?

WHAT ABOUT THE CHANCE TO WIN $100 IN A DRAWING? LIKELIHOOD OF WINNING IS 1 IN ABOUT 3,000 – WOULD THAT BE APPEALING?
WHAT OTHER IDEAS/SUGGESTIONS FOR INCENTIVES

If you wouldn't want to take the survey, help me understand what more that person could do to get your participation. PROBE: What would you want to know? Do you have any concerns? We're trying to get at whether or not people are opposed to surveys, or if it's more of a time factor, or something else.

Conclusion and Thank you. 5 minutes

Now after you've gone through all of this, why do you think the survey is being done? What do you think is going to be done with your responses? Does any of this affect the answers you give?
Last question…Do you recommend any changes to the survey? Graphics? Format? Text?

Thank respondent(s) for participating…have them sign for their incentive.

Appendix C: Pilot Findings and Full-Scale Data Collection Preparations

SEMCOG Pilot Findings and Full-Scale Data Collection Preparations

SEMCOG

NuStats Introduction

- Fred G'sell – Project Manager
- Brad Carlson – Data Collection Coordinator
- Paul Hershkowitz – Wilbur Smith

NuStats GeoStats

SEMCOG

Presentation Overview

- Preview
- Pilot Findings and Analysis
- Cognitive Interviews Findings
- Questionnaire Design
- Data Collection Schedule
- Public Outreach
- Questions and Answers

NuStats GeoStats

SEMCOG

Preview

SEMCOG

Preview

- Regional Transit Systems for SEMCOG Travel Demand Forecast Model
  - DDOT
  - SMART
  - Ann Arbor (AATA)
  - U of Michigan (UM)
  - Detroit People Mover (DPM)
  - Blue Water (BWATC) – not surveyed
  - Lake Erie (LETC) – not surveyed

SEMCOG

Preview

- Criteria for a Successful Study
  - Coverage of the study area / transit providers
  - Statistical significance for the data collected
  - Data Collection with trip purposes and modes
  - Completeness of records, geocoding H/O/D/B/A and transfer activity
  - Real time assessment of quality and quantity
  - Full documentation of processes

NuStats

GeoStats

SEMCOG

Preview

Task	Date(s)
Conduct Full-scale Data Collection	9/20/10 - 11/12/10
Process Full-scale Data	9/27/10 - 12/10/10
Deliver Complete Draft Data	12/20/10
SEMCOG Review of Draft Data	1/10/11
Produce Final Data	1/21/11
Wilbur Smith Review of the Data	2/4/11
Data Weighting	2/25/11
Data Analysis and Report	3/18/11
Final Presentation	3/29/11

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

- Survey Administration Methods
  - DDOT and SMART – self-administered
  - UM, AATA, DPM – self-administered & Web-based
  - TOD distribution

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

- Overall figures
  - 7,182 eligible passengers boarded
  - 1,132 collected back from passengers / Web
  - 877 deemed complete using trip characteristics requirements
  - 171 collected via Web
  - 120 collected via mail back
  - 12.2% overall response rate

NuStats GeoStats

SEMCOG

Pilot Findings and Analysis

n=1,132

Question	Count	Percent
Origin Incomplete or Missing (geocoding fail)	78	6.9%
Destination Incomplete or Missing(geocoding fail)	118	10.4%
Origin Purpose Missing	17	1.5%
Destination Purpose Missing	27	2.4%
Access Mode Missing	20	1.8%
Egress Mode Missing	67	5.9%
Route Sequence Invalid	2	0.2%
Control File Incomplete	17	1.5%
Origin and Destination Same	29	2.6%

NuStats GeoStats

SEMCOG

Pilot Findings and Analysis

n=877

Question	Missing Count	Percent Non-Response
Total # of buses/trains	1	0.1%
Age	9	1.0%
Household Size	13	1.5%
Household Workers	17	1.9%
Employment/Student Status	44	5.0%
Household Vehicle	14	1.6%
Car Availability	19	2.2%
Trip Frequency	14	1.6%
Alternative Mode if Bus Unavailable	22	2.5%
Round Trip Purpose	1	0.1%
Household Income	67	7.6%
Improvement Suggestion	94	10.7%

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

Overall System Comparisons

Transit System	Completed	Collected	Boardings	Response Rate
AATA	264	312	1,703	15.5%**
DDOT	226	324	1,324	17.1%
People Mover	55	77	1,378	4.0%
SMART	158	200	577	27.4%
U of M	174	202	2,200	7.9%*
Total	877	1,132*	7,182	12.2%

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

Overall Method Comparison

Method	Trips	Completed	Collected	Boardings	Response Rate
Web	138	139	171	2,728	5.1%
Paper	193	738	961	4,454	16.7%
Total	331	877	1,132	7,182	12.2%

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

System / Method Comparison

System	Web		Paper		Total	
	Complete	Response Rate	Complete	Response Rate	Complete	Response Rate
AATA	78	6.9%*	186	32.3%	264	15.5%
People Mover	32	3.8%	23	4.2%	55	4.0%
U of M	29	3.4%*	148	10.3%	174	7.9%
Total	139	5.1%	354	13.9%	493	9.3%

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

Trip Length Comparison – Distance, miles

Survey Method	N	Mean	Std. Error of Mean	Std. Deviation	Minimum
Web	139	1.7	.14441255	1.702	.10
Paper	354	1.6	.06587349	1.239	.09
Total	493	1.6	.06237724	1.385	.09

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

Trip Length Comparison – Time, minutes

Survey Method	Mean	N	Std. Error of Mean	Std. Deviation	Minimum
Web	11.3	102*	.86652250	8.751	1
Paper	12.1	263*	.43220637	7.009	1
Total	11.9	365*	.39434651	7.534	1

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

Trip Length Comparison - % Short Trip, < 6 min

Survey Method	N	5 minutes or less		Greater than 5 minutes	
		Count	Percent	Count	Percent
Web	102	29	28.4%	73	71.6%
Paper	263	39	14.8%	224	85.2%

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

- Recommendations - DDOT
  - Use Rosa Parks TC for data collection base, not garages
  - Significant public outreach
    - Media outlets
    - Major transfer centers
    - Buses
    - Website
    - Grassroots groups

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

- Recommendations - SMART
  - Operate out of garages for data collection base
  - Less significant public outreach

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

- Recommendations - AATA
  - Ideally use space in downtown AA or campus in conjunction with UM
  - Use self-administered instrument, still have Web access
  - Less significant public outreach

NuStats

GeoStats

SEMCOG

Pilot Findings and Analysis

- Recommendations – U of M
  - Ideally use space in downtown AA or campus in conjunction with AATA
  - Use self-administered instrument, still have Web access
  - Significant public outreach
  - Potentially change the data collection methodology to account for "short trippers"

NuStats GeoStats

SEMCOG

Pilot Findings and Analysis

- Recommendations – Detroit People Mover
  - Use Rosa Parks TC for data collection base, not garages
  - Significant public outreach
    - Major transfer centers
    - Buses
    - Website
    - Grassroots groups

NuStats GeoStats

SEMCOG

Pilot Findings and Analysis

- Recommendations – People Mover
  - Ideally use Rosa Parks TC in conjunction with DDOT
  - Use self-administered instrument, still have Web access
  - Significant public outreach
  - Potentially change the data collection methodology to account for “short trippers”

NuStats

GeoStats

SEMCOG

Cognitive Interview Findings

NuStats

GeoStats

SEMCOG

Cognitive Interview Findings
Purpose

- Gather feedback on the survey instrument and suggestions for improving the instrument, thus improving data quality,
- Learn what messages would be motivating and encouraging for participation,
- Determine the appropriate media outlets for communication, and
- Provide insight on the need for, preferred amounts of, and types of incentives.

NuStats

GeoStats

SEMCOG

Cognitive Interview Findings
Recruitment

- Rosa Parks TC - June 7^{th} & 8^{th}
- Through out the day between 8a - 7p
- Different genders, ages, ethnicities and various perceived backgrounds
- Explanation of CI purpose, compensation and location
- Interviews occurred June 9^{th} and 10^{th} at the Renaissance Conference Center in downtown Detroit

NuStats

GeoStats

SEMCOG

Cognitive Interview Findings
Survey Instrument

- Did not know exact addresses
- Did not understand how to answer “Place Name”
- Concept of a one-way trip was confusing
- Did not look at or did not understand “Examples”
- Did not always know bus route numbers (names)
- Concept of a prize drawing was mostly favorable, but some were leery
  - Bus passes,
  - cash and,
  - big-ticket items are appealing.

NuStats GeoStats

SEMCOG

Cognitive Interview Findings
Motivation to Participate

- Nearly all said they would be willing to take one, though some would not complete it on the vehicle
- Reasons why not?
  - Crowded bus
  - Length of trip
  - Mood
- Surveyor impact
  - Dress
  - Attitude
  - Communication ability

NuStats GeoStats

SEMCOG

Cognitive Interview Findings
Motivation to Participate cont.

- Respondent suggestions
  - Personalize the message
  - Survey is for route X
  - Survey is to help improve service for everyone

NuStats

GeoStats

SEMCOG

Cognitive Interview Findings

Media Sources and Communication Outreach

- Respondent typically get their news / information from traditional sources – TV, Radio, Newspaper Web, but also through word of mouth
- Best way for passengers to get transit information
  - On buses
  - In transfer centers
  - System website
  - Community grassroots organizations
    - Transportation United
    - Moses

NuStats

GeoStats

SEMCOG

Cognitive Interview Findings

Incentive Preferences

- Varying opinions on need for incentive
  - $100 prize appealed to many
  - Others didn't think it was needed
  - Still others didn't even notice the drawing portion
- Feelings on different types of incentives
  - As expected, instant gratification more appealing than drawings
  - As a prize drawing, monthly bus passes were appealing
  - For those who preferred cash, more preferred the idea of more smaller prizes as opposed to fewer larger prizes
  - Some favored a "big ticket" prize over cash or pass

NuStats GeoStats

SEMCOG

Questionnaire Design

Review Pilot and Current Questionnaires

NuStats GeoStats

SEMCOG

Data Collection Schedule

SEMCOG

Data Collection Schedule

- Surveyor training – September 20th / 21st & as needed after the initial trainings
- DDOT surveying begins - September 22nd
- People Mover – early / mid October
- SMART – late October
- AATA & UM – October 25th
- LET & BWATC - TBD

SEMCOG

Public Outreach Plan

SEMCOG

Public Outreach Plan

- Media outlets (TV, Radio, Newspaper)
- Transit agencies Website links to SEMCOG
- Postings at major Transit Centers
- Postings in transit vehicles
- Grassroots efforts, i.e. Transportation United

SEMCOG

Questions?

Appendix D: Frequently Asked Questions

SEMCOG Information

April 17, 2010

Frequently Asked Questions
SEMCOG Onboard Transit Survey

What is the purpose of the onboard transit survey and how is the data used?

The onboard transit survey is used to help Southeast Michigan plan for improved transit infrastructure and service. Your responses help SEMCOG and its partners better understand the travel patterns and choices of transit riders. This also allows us to make more effective improvements and better serve transit users.

What is SEMCOG?

SEMCOG, the Southeast Michigan Council of Governments, is a membership organization of local governments and the region's designated Metropolitan Planning Organization (MPO). As an MPO, SEMCOG is responsible for regional transportation planning, including the region's transit system. In its efforts to improve the mobility of all people, SEMCOG works to help public and private transportation providers meet the mobility needs of the people and communities they serve. For more information about SEMCOG, click here.

How often are these studies performed and why now?

Onboard transit studies are typically conducted every five years. The region's last survey was conducted in 2001. In the past nine years, our region has changed – gas prices have risen, the transit system and operations have been modified, and our economy is transitioning. As a result, people have changed where, when, and how they make their trips, including their use of transit.

Who is conducting the study?

This study is being conducted by SEMCOG in partnership with the area's transit providers:

- Detroit Department of Transportation (DDOT)
- Suburban Mobility Authority for Regional Transportation (SMART)
- The Detroit People Mover
- Ann Arbor Transportation Authority (AATA) and the University of Michigan System
- Lake Erie Transit (LET)
- Blue Water Area Transit (BWAT)

A contractor was hired by SEMCOG to conduct the onboard survey. The people you see on the bus, handing out surveys are contracted employees. The data, however, belongs to SEMCOG and its transit partners.

Will I have access to the data or the final findings?

SEMCOG will produce a summary report when the study is complete.

Is the data shared with or sold to anyone other than SEMCOG and the transit providers?

No. We want to ensure the anonymity of our survey participants. Data will be “scrubbed” to delete any personal information. Records are then aggregated in order to be useful for transit planning. This final database is what is used by SEMCOG and its partners.

What if I receive a survey on more than one route?
Since the survey is about your specific one-way trip, please fill out the survey again regarding your additional one-way trip.

Will my route be eliminated or changed based on my information?
This information will allow SEMCOG and its partners to implement enhancements to the system that make the system more effective, reliable, and efficient. The ultimate goal is to better meet travelers’ needs.

Why do you need to ask sensitive information such as income?
Studies have shown that certain household characteristics, such as income, are key indicators of travel behavior. By understanding these relationships, planners are better able to predict how recommended service improvements will perform. Please keep in mind that personal information is removed from all records and the results are aggregated – a measure that ensures your privacy and anonymity.

When will the drawing for the incentive take place?
The drawing will occur on or before December 31, 2010. The date is dependent upon the input of the final survey.

How long will it take to complete the survey?
Your individual survey may take between 3-10 minutes to complete.

When will the project be completed?
We expect to have all transit lines surveyed by late fall 2010.

Who is paying for the survey?
SEMCOG receives transportation planning funds from the federal government that are being used to pay for the survey.

I received a second mail-back survey after I completed my TravelCount’10 survey on a DDOT bus. What is this survey?
Since SEMCOG partnered with local transit providers like DDOT, we gave them the opportunity to distribute their own surveys along with ours, as a cost savings measure. This survey has no affiliation with the SEMCOG survey other than our contractor dispersing them.

Please visit the following site for more information on the DDOT survey.

Http://_______

Appendix E: Questionnaire

Southeast Michigan Council of Governments

Questionnaire

July 2012

206 Wild Basin Rd., Suite A-300
Austin, Texas 78746
Contact: Fred G'sell, Project Manager
(512) 306-9065
www.nustats.com

Figure E1: Pilot Survey English Survey Instrument

Please provide any additional comments.

Return the completed survey to the surveyor, or drop it in any mailbox *(no postage required)*.

Thank you!

If you have additional customer comments or questions about **SEMCOG** and its services, please e-mail InfoCenter@semcog.org or call 313-342-3362 or 800-961-3334.

If returning by mail, please close with tape.

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 5478 AUSTIN TX
POSTAGE WILL BE PAID BY ADDRESSEE

NUSTATS
206 WILD BASIN RD STE A300
AUSTIN TX 78746-9907

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

SEMCOG
TravelCount '10

1. **Register to win one of ten $100 cash prizes when you answer all questions!**
Print letters and numbers clearly in upper case: A B C 1 2 3

Name

( ) - Phone Number

All information is confidential and will not be shared or sold.

The following questions are about this ONE-WAY TRIP you are making now!

(START) WORK → WALK → BUS → TRANSFER → BUS → WALK → HOME (END)

Example

Your trip may be different from our example.

2. **Including this bus, how many total buses will you ride to make this one-way trip?**

- ◯ One, this bus only
- ◯ Two
- ◯ Three or more

3. **Which bus or buses?**

	1st: Bus →	2nd: Bus →	3rd: Bus →	4th: Bus
Example:	200	29	—	—

4. **How did you get to the first bus on this one-way trip?**

- ◯ Walked/Wheelchair → # blocks
- ◯ Dropped off
- ◯ Bicycled
- ◯ Taxi
- ◯ Drove alone
- ◯ Carpooled

→ Parking Lot Name/Cross Streets

A-2012-E-0-0101

For office use only

Continue Inside →

Where are you coming from now?

5. What kind of place?

- ⬭ Home
- ⬭ Work or Work-Related
- ⬭ University/College (Student only)
- ⬭ High School/Middle School (Student only)
- ⬭ Shopping
- ⬭ Medical Services
- ⬭ Social, Eat out, Recreational, Religious, Community or Personal Business
- ⬭ Other, specify: ______________

6a. What is the name of this place/building?

6b. What is the exact street address?
(Provide the nearest cross streets if you don't know the exact address)

Address

Cross Street #1

and Cross Street #2

City Zip

Where are you going to now?

7. What kind of place?

- ⬭ Home
- ⬭ Work or Work-Related
- ⬭ University/College (Student only)
- ⬭ High School/Middle School (Student only)
- ⬭ Shopping
- ⬭ Medical Services
- ⬭ Social, Eat out, Recreational, Religious, Community or Personal Business
- ⬭ Other, specify: ______________

8a. What is the name of this place/building?

8b. What is the exact street address?
(Provide the nearest cross streets if you don't know the exact address)

Address

Cross Street #1

and Cross Street #2

City Zip

9. Where will you get off this bus?

Name of Place *(including Park & Ride or other lots)*

Cross Street #1

and Cross Street #2

10. How will you get from the last bus to your final destination on this one-way trip?

- ⬭ Walk/Wheelchair → # blocks
- ⬭ Picked up
- ⬭ Bicycle
- ⬭ Taxi
- ⬭ Drive alone
- ⬭ Carpool

→ ______________ Parking Lot Name/Cross Streets

11. What is your age?

⬭ Under 16 ⬭ 16 to 17 ⬭ 18 to 34 ⬭ 35 to 54 ⬭ 55 or older

12. Do you have a valid driver's license?

⬭ Yes ⬭ No

13. Including yourself, how many people live in your house/apartment?

⬭ 1 ⬭ 2 ⬭ 3 ⬭ 4 or more

14. Including yourself, how many of the people in your household are employed full-time or part-time?

⬭ None ⬭ 1 ⬭ 2 ⬭ 3 or more

15. Are you...

- ⬭ Full-time Worker
- ⬭ University/College Student
- ⬭ Part-time Worker
- ⬭ Middle/High School Student
- ⬭ Homemaker
- ⬭ Other Student

16. How many working vehicles are available to your household?

⬭ None ⬭ 1 ⬭ 2 ⬭ 3 or more

17. Was a car (or other personal vehicle) available to make this trip?

⬭ Yes ⬭ No

18. How many days a week do you usually make this trip?

- ⬭ 6-7 days/week
- ⬭ 1-3 days/month
- ⬭ 3-5 days/week
- ⬭ Less frequent ride
- ⬭ 1-2 days/week

19. If bus service was not available, how would you make this trip?

- ⬭ Walk/Wheelchair
- ⬭ Taxi
- ⬭ Drive
- ⬭ Bicycle
- ⬭ Ride with someone else
- ⬭ Would not make this trip

20. On this ROUND TRIP (between the time you left home and will return home) will you... *(check all that apply)*

- ⬭ Go to work
- ⬭ Go to school
- ⬭ Go shopping
- ⬭ Visit friends/Attend a religious or social event
- ⬭ Buy a meal
- ⬭ Other, specify: ______________

21. What was your estimated total household income in 2009 before taxes?

- ⬭ Less than $10,000
- ⬭ $30,000 - $39,999
- ⬭ $60,000 - $74,999
- ⬭ $10,000 - $19,999
- ⬭ $40,000 - $49,999
- ⬭ $75,000 or more
- ⬭ $20,000 - $29,999
- ⬭ $50,000 - $59,999

22. What one service do you feel needs the most improvement?

- ⬭ Make transfers easier
- ⬭ Improve pedestrian access
- ⬭ Start service earlier
- ⬭ Increase frequency of service
- ⬭ End service later
- ⬭ Other, specify on back
- ⬭ Add new route from ______________ to ______________

Please continue on the back →

Figure E2: Full-Scale Data Collection English Survey Instrument

Please provide any additional comments.

Return the completed survey to the surveyor, or drop it in any mailbox *(no postage required)*.
If you would like to fill out this survey electronically, visit: www.surveys.nustats.com/semcog

Thank you!

If you have additional customer comments or questions about SEMCOG and its services, please visit the survey website www.semcog.org/On-boardTransitSurvey.aspx, email InfoCenter@semcog.org, or call 313-324-3362 or 800-961-3334.

If returning by mail, please close with tape.

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 5478 AUSTIN TX
POSTAGE WILL BE PAID BY ADDRESSEE

NUSTATS
206 WILD BASIN RD STE A300
AUSTIN TX 78746-9907

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

SEMCOG TravelCount '10

Hi, thank you for helping us improve **YOUR** transit system.
Complete this survey and become eligible to win one of ten **$100** cash prizes!

Please provide your contact information.
Fill in answers clearly as shown in this example: A B C 1 2 3

Name

Phone Number ( ) -

All information is confidential and will not be shared or sold.

1. What one service do you feel needs the most improvement?

- ○ Make transfers easier
- ○ Start service earlier
- ○ End service later
- ○ Add new route from, specify: ______ to ______
- ○ Improve pedestrian access
- ○ Increase frequency of service
- ○ Other, specify on back
- ○ No service improvements needed

Answer the rest of the survey questions about the ONE-WAY TRIP you are taking now!

2. Including the bus where you were given this survey, how many total buses will you take on this one-way trip? (Please include People Mover rides in your total.)

- ○ One, this bus only
- ○ Two
- ○ Three or more

3. List all the route numbers (or names) you will use for this one-way trip.

	1st Route →	2nd Route →	3rd Route →	4th Route
Example:	125	PEOPLE MOVER	29 OR LINWOOD	—

A - 26023 - D10201D

For office use only. a. b. c.

Continue Inside →

4. ORIGIN LOCATION: Where are you coming from?

a. What kind of place?

- ○ Home
- ○ Work or Work-Related
- ○ University/College (Student only)
- ○ High School/Middle School (Student only)
- ○ Shopping
- ○ Medical Services
- ○ Social, Eat Out, Recreational, Religious, Community or Personal Business
- ○ Other, specify: ____________

b. What is the exact street address?
(Provide the nearest cross streets if you don't know the exact address)

Address

Cross Street #1

and Cross Street #2

City

Zip

c. Please record the name of the place or building (Enter HOME if you are coming from home).

d. How did you get from your origin location to the first bus stop (or People Mover station) on your trip?

- ○ Walked/Wheelchair → # blocks (please answer 0 if less than 1 block)
- ○ Dropped off
- ○ Bicycled
- ○ Taxi
- ○ Drove alone
- ○ Carpooled

→ Parking Lot Name/Cross Streets

5. When you got on the bus where you were given this survey, how did you pay?

- ○ Cash/Day Pass
- ○ Transfer
- ○ Other pass (weekly, monthly, prepaid card, etc.)
- ○ University of Michigan, free → Skip to Q8
- ○ Other, specify: ____________

6. Was your fare...

- ○ Regular fare
- ○ Student (K-12)
- ○ Senior/Disabled

7. Did your employer or another organization pay for your fare?

- ○ Yes, entire fare
- ○ Yes, some of fare
- ○ No

8. DESTINATION LOCATION: Where are you going to?

a. What kind of place?

- ○ Home
- ○ Work or Work-Related
- ○ University/College (Student only)
- ○ High School/Middle School (Student only)
- ○ Shopping
- ○ Medical Services
- ○ Social, Eat Out, Recreational, Religious, Community or Personal Business
- ○ Other, specify: ____________

b. What is the exact street address?
(Provide the nearest cross streets if you don't know the exact address)

Address

Cross Street #1

and Cross Street #2

City

Zip

c. Please record the name of the place or building (Enter HOME if you are going home).

d. How will you get to your destination location from the last bus stop (or People Mover station) on your trip?

- ○ Walk/Wheelchair → # blocks (please answer 0 if less than 1 block)
- ○ Picked up
- ○ Bicycle
- ○ Taxi
- ○ Drive alone
- ○ Carpool

→ Parking Lot Name/Cross Streets

9. How many days a week do you usually make this trip?

- ○ 6-7 days per week
- ○ 1-3 days per month
- ○ 3-5 days per week
- ○ Less than 1 day per month
- ○ 1-2 days per week
- ○ First time to make this trip

10. If bus service was not available, how would you make this trip?

- ○ Walk/Wheelchair
- ○ Taxi
- ○ Drive
- ○ Bicycle
- ○ Ride with someone else
- ○ Would not make this trip

11. Where will you get off the bus on which you were given this survey?

Name of Place *(including Park & Ride or other lots)*

Cross Street #1

and Cross Street #2

12. Do you have a valid driver's license?

- ○ Yes
- ○ No

13. What is your age?

- ○ Under 18
- ○ 26 - 34
- ○ 55 - 64
- ○ 18 - 25
- ○ 35 - 54
- ○ 65 + years of age

14. Are you... *(fill in all that apply)*

- ○ Full-time Worker
- ○ University/ College Student
- ○ Unemployed, but seeking work
- ○ Part-time Worker
- ○ Middle/ High School Student
- ○ Unemployed, not seeking work
- ○ Homemaker
- ○ Other Student
- ○ Retired

15. How many working vehicles are available to your household?

- ○ None
- ○ 1
- ○ 2
- ○ 3 or more

16. Including yourself, how many people live in your household?

- ○ 1
- ○ 2
- ○ 3
- ○ 4 or more

17. Including yourself, how many of the people in your household are employed full-time or part-time?

- ○ None
- ○ 1
- ○ 2
- ○ 3 or more

18. What was your total household income in 2009 before taxes?

- ○ Less than $10,000
- ○ $25,000 - $34,999
- ○ $75,000 - $99,999
- ○ $10,000 - $14,999
- ○ $35,000 - $49,999
- ○ $100,000 or more
- ○ $15,000 - $24,999
- ○ $50,000 - $74,999

Please continue on the back →

Figure E3: Pilot Postcard

Appendix F: Data Weighting & Expansion

Southeast Michigan Council of Governments

Regional On-Board Transit Survey

Data Weighting and Expansion

July 2012

NuStats Research Solutions

206 Wild Basin Rd., Suite A-300

Austin, Texas 78746

Contact: Fred G'sell, Project Manager

(512) 306-9065

www.nustats.com

Table of Contents

1. Introduction

1.1 The Need for Weighting and Expansion

Data collected during an on-board transit survey is subject to various biases, which exist because not all trips in the transit system can be sampled, and not all riders on sampled trips respond to the survey. A logical weighting and expansion procedure must be implemented to account and adjust for these sample biases. In the absence of such a procedure, data users cannot confidently draw inferences about the characteristics of the transit-riding population.

As a simple example to illustrate this point, assume that one bus trip serves 100 passengers per day, while another serves just 10 passengers per day. If 5 surveys were collected for each of these trips, then the response rates would equal 5 percent and 50 percent, respectively. Without weighting (in other words, if each collected survey were assigned an equal weight), the survey data collected on the trip with fewer passengers would be overrepresented relative to the data collected on the trip with more passengers (due to the difference in response rates). A logical weighting and expansion procedure properly accounts for such differences and adjusts the data accordingly.

1.2 Traditional Weighting and Expansion

Strictly speaking, the *weighting* component of the weighting and expansion process refers to adjusting the data to account for rider non-response on sampled trips, whereas the *expansion* component refers to 1) expanding the data to account for trips that were not sampled (so that the data encompasses the entire universe of daily transit trips), and 2) adjusting the data so that the final expansion weights, when summed up, equal the overall expansion target—the average weekday ridership of the transit system. That being said, in this appendix the terms *weighting* and *expansion* are occasionally used interchangeably for the purpose of readability.

To maximize the accuracy of the final expansion weight assigned to each survey record, samples must be separated into cells (or strata) based on various data elements, with different weights and factors calculated for samples in different cells. Traditionally, samples are stratified by route, direction, time of day, and boarding stop location.

In general, the more that samples are stratified, the more accurate the final expansion weights will be. However, this is only true if sufficient numbers of samples exist to support the desired level of stratification. When samples are stratified based on multiple data elements, but there are too few collected surveys associated with some cells, those cells must be merged with others to allow the weighting process to proceed and to avoid the calculation of unreasonably large expansion factors.

1.3 The SEMCOG and NuStats Approach: Key Components

Due to the unique nature of the transit system in Southeast Michigan, data limitations encountered during the survey project, and other mitigating factors, SEMCOG and NuStats collaboratively developed a new, thorough weighting and expansion methodology specifically for the SEMCOG on-board transit survey. Underlying the new methodology are procedures which go beyond the traditional boarding-based expansion approach, and which were applied under the guidance of the Federal Transit Administration.

Key components of the expansion methodology developed by SEMCOG and NuStats include the following:

- Use of a weighting method known as Iterative Proportional Fitting (IPF) to incorporate both boarding and alighting data (rather than boarding data alone) into the process. Alighting locations were included to account for the fact that characteristics of transit riders can vary based on where they alight as well as where they board, and to better account for stop-level variation in response rates.
- Inclusion of *patterns* in the stratification of samples. In Southeast Michigan, bus trips often follow different patterns (unique paths with unique sequences of stops) during different times of the day. The paths these patterns take can vary so much that in reality, different patterns associated with a particular route often function more like different routes rather than subtle route variations. This fact (and others) led to the inclusion of patterns as a stratification element.
- Development of an innovative method for calculating expected weekday ridership at the stop level by using the observed boarding and alighting data that was collected during the survey. An innovative method was needed because for the most part, transit agencies in the region were able to provide average weekday ridership by route only—ridership by pattern, stop, or time of day was not available, nor were there any on–off studies available to aid in estimating ridership at the disaggregate stop level.
- Development of innovative tools and methods for aggregating stops. Careful and logical aggregation of stops into groups was a high priority during this project, not only because aggregation was a necessary precursor to the IPF weighting procedure, but also because of the desire to preserve meaningful spatial variation in ridership as much as possible. In Southeast Michigan, ridership on many routes is distributed quite unevenly, with most stops servicing few riders, and just a small number of stops handling the majority of boarding and alighting activity. The aggregation tools and methods that were developed helped maintain the balance between aggregating too little (thereby creating mathematical problems for IPF) and aggregating too much (thereby failing to capture spatial variation in ridership with sufficient precision). The methods employed also allowed staff to consider meaningful information such as land-use characteristics while aggregating stops.
- Incorporation of a special time-of-day factor. Used toward the end of the process, this factor cleverly ensures that the overall methodology is balanced, preserving both the temporal variation in ridership observed during the survey, as well as the spatial variation.

Every step of the weighting and expansion process developed by SEMCOG and NuStats is laid out in Section 4 through Section 8 of this appendix. Preceding these sections are Section 2, which defines some frequently used terms, and Section 3, which provides more detail about the data that was available for developing the process. The appendix concludes with a brief summary in Section 9.

2. Terminology

This section describes and clarifies differences in meaning for the following terms that are used frequently in this appendix: trip, route, pattern, line, stop, time of day, and cell.

Trip

A *trip* refers to a single, regularly scheduled journey made by a transit vehicle during a typical weekday. A trip departs from a specific start location at a specific time, and arrives at a specific end location at a later time. In this appendix, except as noted in Section 8.3 (which describes the difference between linked and unlinked person-trips), the term *trip* always refers to trips made by transit vehicles, not trips made by individuals.

Due to transit system logistics and limited project resources, not all trips in the transit system can be sampled. In addition, of the trips that are sampled, some may be sampled just one time, while others may be sampled more than once (on different days). The weighting and expansion process must account for these issues.

During this survey project, 3,791 regularly scheduled bus trips were sampled across Southeast Michigan. Of these 3,791 trips, 3,518 were sampled once, 266 were sampled twice, and 7 were sampled three times. These statistics include fixed-route bus trips only, not trips made by the Detroit People Mover. It should also be noted that while nearly 3,800 trips were sampled, useable surveys were not returned for every single trip.

Route

A *route* is a group of similar trips, with each trip in the group assigned the same route name and number. These names and numbers are the designations familiar to the public, appearing in published schedules and on the displays of transit vehicles—for example, DDOT Route 53 (Woodward). Although each route consists of multiple individual trips, some routes have only a few trips scheduled each weekday, while others have dozens of scheduled trips.

Another way that routes can be characterized is by the direction of travel associated with them. Some routes consist of trips that travel in only one direction. Others consist of trips that travel in a loop, starting and ending at the same stop. Most routes, however, consist of trips that travel in two opposite directions, such as eastbound and westbound, or inbound and outbound.

A total of 146 routes were surveyed during this project, not including the Detroit People Mover. Complete, useable surveys were returned for all but one of these routes (SMART Route 580), meaning that 145 routes were weighted using the methodology described in this appendix. These 145 routes were associated with 254 unique route/direction combinations.

It should be noted that in a few cases, different real-world routes were assigned the same route number in the project database. For example, all of the various trips associated with SMART's Route 415 (Greenfield) and Route 420 (Southfield) were assigned the same route number (420) in the database. Local and express routes in a corridor were also typically tagged with the same database route number, even though their route numbers and names differ in reality.

Pattern

All of the various trips associated with a particular route and direction do not necessarily follow the exact same path and sequence of stops throughout the entire day. Each unique path variation is called a *pattern*. This term also refers to the sequence of stops associated with the path—in other words, each pattern has its own unique sequence of stops. In the hierarchy of the transit system, each scheduled trip, by definition, is associated with only one pattern. Patterns, in turn, are associated with one or more trips, and each pattern is associated with one (and only one) route and route/direction combination.

As stated in Section 1, weighting and expansion processes typically stratify samples by route and direction. However, the methodology developed for the SEMCOG on-board survey stratifies samples by pattern instead. This is because SEMCOG analyzed the relationship between patterns and routes throughout the region and found that the start and end points of patterns associated with particular routes are sometimes located far apart. In addition, some patterns with identical start and end points traverse very different paths midway through their journeys. In short, different patterns belonging to the same route can serve riders with dissimilar origins and destinations. This discovery led to the decision to include patterns as a sample stratification element. The fact that it was more straightforward to aggregate stops at the pattern level than at the route level also entered into the decision (see Section 6).

A total of 372 patterns were surveyed during this project, not including the Detroit People Mover. Complete, useable surveys were returned for 355 of these patterns, meaning that 355 patterns were weighted using the methodology described in this appendix. Note that there are significantly more patterns (355) than route/direction combinations (254). In other words, to work at the pattern level is to work at a more disaggregate level than route and direction.

It should be noted that the coded features that TransCAD calls *routes* in SEMCOG's travel-demand forecasting model are, in reality, *patterns*. It should also be noted that since, by definition, a pattern is associated with exactly one route/direction combination, it is redundant to say that samples are stratified "by route, direction, and pattern"—to say "stratified by pattern" is sufficient.

Line

Average weekday ridership estimates provided by the transit agencies were necessary for calculating sample collection goals as well as weighting and expanding the survey. Most of the ridership estimates that were provided were route-level estimates, but in a few cases this level of detail was unavailable, and ridership estimates were provided for a collective group of routes instead. For example, AATA provided a combined ridership estimate for all of the routes (4A, 4B, and 4C) that service the Washtenaw Avenue corridor between Ann Arbor and Ypsilanti.

With ridership estimates not available for every route, it was necessary to use a different term to refer to the more aggregate level at which ridership data was uniformly available—the term *line* was introduced for this purpose. Put simply, a line refers to one or more routes, grouped in such a way that weekday ridership estimates are available for each line. Most lines consist of only one route, but some consist of more than one (as in the Washtenaw example).

A total of 139 lines were surveyed during this project, not including the Detroit People Mover. Complete, useable surveys were returned for all but one line (a line consisting of only one route, SMART Route 580), meaning that 138 lines were weighted using the methodology described in this appendix. With 138 lines associated with only 145 routes, this illustrates the point that most lines consist of just one route.

Stop

As described in Section 2.3, each pattern has its own unique sequence of stops. *Stops* in general are simply locations where people can board or alight a transit vehicle. What is important to note is that unless stated otherwise, the term *stops,* in the context of this appendix, always refers to *pattern-stops,* not *physical stops*. To clarify, physical stops are actual bus stops, the real-world physical locations used for boarding and alighting. Pattern-stops, on the other hand, are a statistical concept, resulting from the fact that any number of patterns may use a particular physical stop, but that in the project database, physical stops are assigned different, unique ID numbers for each pattern associated with them. This means that the relationship between pattern-stops and physical stops is a many-to-one relationship, and that there are more pattern-stops than physical stops in the region. Put another way, pattern-stops denote the number of places a transit vehicle can stop along a particular pattern, ignoring the fact that some of these places also serve as stops for transit vehicles traversing other patterns.

The patterns surveyed during this project consist of 24,039 total stops (pattern-stops), with boarding and/or alighting activity observed at 14,657 of these locations. Of course, useable surveys were not returned for each and every surveyed pattern. Constraining the statistics to include only those patterns with complete, useable surveys yields a total of 23,142 stops, with boarding and/or alighting activity observed at 14,501 of these locations. Also of note, 9,303 was the total number of stops used for boarding and/or alighting by the 18,099 bus-trip respondents whose surveys were included in the final dataset.

Again, it must be emphasized that all of the above statistics refer to pattern-stops, not physical stops. In fact, due to a lack of complete consistency in how transit agencies described stop locations on different patterns, the actual number of real-world, physical bus stops in Southeast Michigan is unknown, let alone how many were active during the survey. It should also be noted that the above statistics exclude the 396 survey responses and 13 stations associated with the Detroit People Mover.

Time of Day

The term *time of day* does not refer to a specific time (such as 9:08 AM), but rather to a time period defined by transit agencies to help categorize their service. While all transit agencies in Southeast Michigan use time periods called *AM-peak*, *midday*, and *PM-peak*, the precise times that mark the boundaries of these periods vary from agency to agency. In addition, different agencies use different terminology to describe the time prior to the AM-peak period and the time following the PM-peak period. When this appendix refers to stratifying samples “by pattern and time of day”, this means stratifying by pattern and by the time periods listed in Table 2.1.

Table 2.1: Time Period Definition by Transit Agency

Transit Agency	Time of Day					
	AM Peak	Mid-Day	PM Peak	Evening	Night	Other
Ann Arbor Transit (AATA)	6:00-9:00	9:00-15:00	15:00-18:00	--	18:00-23:00	--
Blue Water Area Transit (BWAT)	6:00-9:00	9:00-15:00	15:00-18:00	18:00-22:00	--	--
Detroit Department of Transportation (DDOT)	6:00-9:00	9:00-14:00	14:00-18:00	18:00-3:00	3:00-6:00	--
Lake Erie Transit (LET)	8:00-9:00	9:00-15:00	15:00-18:00	--	--	6:00-8:00
Suburban Mobility Authority for Regional Transportation (SMART)	6:00-9:00	9:00-15:00	15:00-18:00	18:00-21:00	21:00-6:00	--
University of Michigan (UM)	6:00-9:00	9:00-15:00	15:00-18:00	--	18:00-23:00	23:00-3:00

Cell

For the purposes of this survey, *cell* refers to the stratification of collected samples and scheduled trips by pattern and time of day. In other words, each unique combination of pattern and time of day forms a different cell. The term *strata* is sometimes used in place of *cell*.

A cell is ignored if it is empty, meaning that there are no scheduled trips associated with that cell's pattern during that cell's time period. The remaining cells contain at least one trip, and it is assumed that all of the trips in a cell are sufficiently similar to one another to be grouped together during certain steps of the weighting and expansion process. In particular, two important factors—the vehicle factor and the time-of-day factor—are calculated at the cell level (see Section 5 for more about the vehicle factor, and Section 8 for more about the time-of-day factor).

Summary

For convenience and reference, many of the statistics cited in the preceding subsections are collected in Table 2.2 below.

Table 2.2: Summary of Collection Statistics

Surveyed transit agencies with fixed-route bus service	6
Trips surveyed	3,791
Lines with completed surveys	138
Routes with completed surveys	145
Route/direction combinations with completed surveys	254
Patterns with completed surveys	355
Total stops (pattern-stops) on patterns with completed surveys	23,142
Stops with observed boarding and alighting activity	14,501
Stops used by survey respondents	9,303

3. The Data Available for Weighting and Expansion

The process that SEMCOG and NuStats developed for calculating weighting and expansion factors for the SEMCOG on-board study relies primarily on the datasets described in this section.

3.1 Average Weekday Ridership

For this survey, each of the six participating bus-operating transit agencies provided estimates of average ridership for a typical 2010 weekday. These estimates provided the basis for establishing sample collection goals in addition to serving as an important data source for the weighting and expansion process. As explained in Section 2.4, while most ridership estimates were provided at the route level, uniform estimates across the region were available only at the more aggregate line level.

With the exception of the Detroit People Mover, which provided boarding counts for each of its 13 stations, no agency provided data at a more disaggregate level than route level. This means that ridership was not available by pattern, time of day, or stop. As the SEMCOG survey progressed, SMART did begin an on–off study for its system, which could have provided more detailed data. However, the results of that study were not available in time to be considered for the weighting and expansion of SEMCOG's survey. Furthermore, no on–off data was available for any other agency in the region, and the survey team wished to be as consistent as possible, from agency to agency, in the methodology used to weight the survey.

Table 3.1 provides a breakdown of average weekday ridership by transit agency, based on the line-level data that was provided. This table shows that the typical weekday ridership for Southeast Michigan during the time of the survey was 222,140 boardings (or alightings) per day. It should be noted that this total excludes weekend ridership (by definition), but also excludes a small handful of special weekday routes—see Section 3.2 for details.

Table 3.1: Average Weekday Ridership by Transit Agency

System	Average Weekday Ridership
AATA	22,010
BWAT	2,625
DDOT	124,514
DPM	4,011
LET	877
SMART	33,876
U-M	34,227
Total	222,140

3.2 Transit Schedules, Stop Locations, and Other Trip-Related Data

Along with ridership estimates, the transit agencies provided NuStats with detailed information about the trips they operated, including schedules, stop locations, and other important data. Using this information, NuStats derived a hierarchy of relationships among the key elements of the transit system: the trips associated with each cell, the sequences of stops associated with each pattern, the patterns associated with each route and direction, the routes associated with each line, and the lines associated with each transit agency. An understanding of these relationships was necessary for developing a logical weighting and expansion process.

In addition, this information was the source for two important tables compiled by NuStats: the *run-cut file* and the *master stop table*. The run-cut file, which was the basis for selecting the trips to sample for the survey, lists all of the trips made by each transit agency during a typical weekday. Accordingly, weekend trips were not included in the table, nor were a handful of certain types of weekday trips, including special Friday night service trips, BWAT trolley trips, U-M trips operated only when peak capacity was overwhelmed, and the few school bus trips operated by SMART. Such trips did not fall within the scope of the project and were never considered for sampling. It should also be noted that in the LET system, there were slight deviations in trip departure times from one weekday to the next, but trips were sampled and weighted assuming that there were no significant differences in ridership from day to day due to these shifts.

The master stop table lists all locations where riders can potentially board or alight from transit vehicles during trips. Originally, the master stop table was simply a combination of all of the pattern-stops denoted by the transit agencies, grouped and sorted by pattern and stop sequence number. However, NuStats ultimately expanded the table to include additional, unlisted stop locations where boarding and alighting was observed during the actual survey. It should be noted, though, that a complete list of possible stops for the BWAT system could not be generated for this project—instead, what was available and ultimately used were the time points from BWAT's published schedules. Also of note, the last stop along each pattern in the AATA system was not provided to NuStats and did not end up in the master stop table.

3.3 Observed Boarding and Alighting Activity

As part of the data collected in the field for each sampled trip, NuStats recorded the number of people boarding and alighting at each stop. As explained in upcoming sections of this appendix, this observed activity at the trip/stop level became an essential component of the weighting and expansion process. More specifically, this data was used in combination with the other sources to develop necessary input for the Iterative Proportional Fitting (IPF) weighting procedure and to calculate the expansion targets needed to adjust the data using the time-of-day factor.

3.4 Other Data Sources

Naturally, data from the collected surveys themselves was used in the weighting and expansion methodology. NuStats collected a total of 18,495 useable records for the final dataset that was delivered to SEMCOG. Each record's pattern number, time of day, boarding stop location, and alighting stop location were all incorporated into the process.

One important note about the alighting location: For the small number of surveys turned in by riders who boarded a bus on one trip, but rode through the end of that trip and alighted the bus on a subsequent trip, the actual alighting location was not used in the weighting process. Instead, the alighting location was changed to the final stop on the boarding pattern. This was necessary because for a survey record to be included in the IPF weighting procedure, both the boarding and alighting stops needed to be located on the same pattern.

Beyond the data sources already listed, a few more were used in the process of aggregating stops into groups (see Section 6). These sources include highway network nodes from SEMCOG's travel-demand forecasting model, plus data used to calculate land-use characteristics.

4. Methodology Overview

4.1 Introduction

This section and the four that follow describe, step by step, the entire weighting and expansion methodology developed by SEMCOG and NuStats for the SEMCOG on-board transit survey. The main goal of the methodology was to weight and expand samples in a balanced way, capturing both spatial and temporal ridership distributions as accurately as possible given data limitations and the unique characteristics of the region's transit system. Spatial variation in ridership was captured by incorporating an approach known as *Iterative Proportional Fitting* (IPF) into the process. Temporal variation in ridership was captured through the vehicle factor and time-of-day factor, which factored in the trip sampling rates per time period and the expected ridership per time period.

At different points in the methodology, survey samples were stratified in different ways. More specifically, for running the IPF procedure, the data was stratified by pattern, boarding location, and alighting location; for calculating vehicle and time-of-day factors, the data was stratified by pattern and time of day. SEMCOG and NuStats seriously considered the possibility of simultaneously stratifying the data by pattern, boarding location, alighting location, and time of day throughout the entire process, but analysis indicated that it was not feasible to do this while still maintaining desired levels of precision and balance—there were simply not enough samples to warrant stretching the data that thin.

4.2 Iterative Proportional Fitting

Referenced above, Iterative Proportional Fitting, or IPF, occupies a central place in the weighting and expansion methodology. Because nearly all of the process steps described in Section 5 through Section 8 of this appendix relate in one way or another to the IPF procedure, and because of the significant time spent preparing the IPF input data, testing the procedure itself, and refining the IPF output, this overview section is an appropriate place to explain what Iterative Proportional Fitting is and how it was used in the context of this project.

In general terms, Iterative Proportional Fitting, also known as *raking*, is a mathematical process by which a two-dimensional table of data is systematically adjusted so that 1) the numbers in all of the rows, when added up, equal a set of pre-defined control totals; and 2) the numbers in all of the columns, when added up, equal another set of pre-defined control totals. The control totals are sometimes called *marginal totals* (or *marginal control totals*), and the original, unadjusted table values are sometimes called *seeds*. The process is iterative, and typically continues until differences in values calculated during successive iterations fall within a specified tolerance, at which point the process is said to have converged.

In the context of this project, Iterative Proportional Fitting was used to calculate the response factor, an important component of the final expansion weight. The purpose of the response factor is to properly account for riders on sampled trips who did not return complete, useable surveys. By using IPF, SEMCOG and NuStats were able to incorporate both boarding and alighting stop locations into this weighting process, rather than boarding locations alone. (For more about the response factor, see Section 7.)

The IPF procedure was run (separately) for each of the 355 bus patterns for which complete, useable surveys were returned. The procedure was designed so that:

1. Each row and column would represent the pattern's boarding and alighting stop locations, respectively;
2. The control totals (marginal totals) for the rows would equal typical weekday boarding totals for each stop location;
3. The control totals for the columns would equal typical weekday alighting totals for each stop location;
4. The seeds would be the number of survey records associated with each boarding and alighting combination;
5. The procedure would generate a matrix of adjusted values representing the number of riders expected to board and alight via each possible combination of stop locations; and
6. Response factors would be calculated by dividing each adjusted value by its original seed.

To illustrate by way of a simplistic example, assume that Pattern 1 had only four stops associated with it. Further assume the following:

- that 20 surveyed transit riders boarded at stop 1 and alighted at stop 2,
- that 28 riders boarded at stop 1 and alighted at stop 3,
- that 15 riders boarded at stop 1 and alighted at stop 4,
- that 22 riders boarded at stop 2 and alighted at stop 3,
- that 10 riders boarded at stop 2 and alighted at stop 4, and
- that 4 riders boarded at stop 3 alighted at stop 4.

Further assume that the typical weekday ridership for Pattern 1 was as follows:

- a total of 900 boardings at stop 1, a total of 680 at stop 2, and 120 at stop 3; and
- a total of 400 alightings at stop 2, a total of 750 at stop 3, and 550 at stop 4.

The above numbers comprise the input data needed to run the IPF process for the pattern. Represented in matrix form, the seeds and control totals would look like this:

Table 4.1: Example Initial Input for Iterative Proportional Fitting (IPF)

Stops	1	2	3	4	Expected Ridership
1	0	20	28	15	900
2	0	0	22	10	680
3	0	0	0	4	120
4	0	0	0	0	0
Expected Ridership	0	400	750	550	1700

After Iterative Proportional Fitting, the adjusted data would look like this:

Table 4.2: Example IPF Results

Stops	1	2	3	4	Expected Ridership
1	0	400.0	306.8	193.2	900
2	0	0	443.2	236.8	680
3	0	0	0	120.0	120
4	0	0	0	0	0
Expected Ridership	0	400	750	550	1700

In words, on a typical weekday about 400 transit riders would be expected to board at stop 1 and alight at stop 2, about 307 riders would be expected to board at stop 1 and alight at stop 3, and so on.

Dividing the IPF output by the original seeds (numbers of surveys) yields these response factors:

Table 4.3: Example Weighting and Expansion Factors from IPF

Stops	1	2	3	4
1	N/A	20.00	10.96	12.88
2	N/A	N/A	20.15	23.68
3	N/A	N/A	N/A	30.00
4	N/A	N/A	N/A	N/A

In words, all of the samples corresponding to riders that boarded at stop 1 and alighted at stop 2 would be assigned a response factor of 20.00, all that boarded at stop 1 and alighted at stop 3 would be assigned a response factor of 10.96, and so forth.

Although designing an IPF procedure to calculate response factors was straightforward conceptually, preparing the necessary input data was not as straightforward. For one thing, typical weekday ridership at the pattern and stop level was not readily available. Therefore, an innovative method was needed for calculating IPF control totals—this method is the subject of Section 5. In addition, for most patterns, it was mathematically impossible to run the IPF procedure using individual stops as rows and columns. Instead, for the IPF procedure to properly converge and produce logical results, stops first needed to be aggregated into groups—the methods developed for doing this are the subject of Section 6.

Even after disaggregate stop-level ridership estimates were calculated and stops were aggregated, running the IPF procedure still presented several challenges—these are discussed in Section 7. And finally, as alluded to earlier in this section, the response factors that emerged from the IPF procedure were only one component of the final expansion weights that were calculated—the remaining calculations are the subject of Section 8.

4.3 Weighting and Expansion for the Detroit People Mover

It must be emphasized that all of the steps of the methodology as described in Section 5 through Section 8, plus the use of IPF as described in Section 4.2, apply only to samples collected for transit agencies operating fixed-route bus service: namely, the Detroit Department of Transportation (DDOT), the Suburban Mobility Authority for Regional Transportation (SMART), the Ann Arbor Transportation Authority (AATA), the University of Michigan's transit service (U-M), Blue Water Area Transit (BWAT), and Lake Erie Transit (LET). Samples collected on the Detroit People Mover (DPM), on the other hand, were weighted and expanded in a different way.

The Detroit People Mover (DPM) is an elevated fixed-guideway transit system that circles downtown Detroit. Because of this, trips on the DPM are characteristically different from bus trips. Passengers on the People Mover are on the vehicle for much shorter time periods, and the time between stops are very short. Both these things made the onboard data collection methodology used on the other systems impractical. Instead, surveyors on the People Mover conducted intercept interviews,[1] which could be done quicker, although didn't allow for a systematic count of passengers. The results were instead expanded to the average daily boardings per station, which were collected by DPM prior to the study.

5. Calculating IPF Control Totals

5.1 Introduction

As noted in Section 4, to populate marginal control totals and run the IPF process, SEMCOG and NuStats required estimates of typical weekday boarding and alighting activity—at the stop level—for each of the 355 bus patterns for which complete, useable surveys were returned. However, as noted in Section 2 and Section 3, Southeast Michigan transit agencies were only able to provide average weekday ridership at the line level, not by pattern and stop.

With limited data availability, SEMCOG and NuStats needed to develop an innovative method for calculating the necessary IPF input. Ultimately, a method was developed to take advantage of the fact that NuStats recorded the number of people observed getting on and off at each stop for each surveyed bus trip. As shown in the steps that follow, based on these observations, it was possible to calculate logical stop-level ridership estimates, referred to as *expected weekday ridership*, or *expected weekday boarding and alighting values*.

5.2 Balancing Boarding and Alighting Observations

As a prerequisite for performing an iterative proportional fitting, the control totals for each category must be equal. The IPF would fail (more specifically, the process would not converge) if boardings and alightings were not balanced. An analysis of the raw observations revealed that the total number of observed boardings for a given surveyed trip did not always equal the total number of observed alightings, although the totals never differed by much.

Such minor differences were expected, due to the impossibility of precisely accounting for every person's movement on a bus trip, especially on trips with very high ridership. However, measures had to be taken to ensure that the boarding and alighting counts were equal for each trip. For trips with more boardings than alightings, alightings were added to the last stop of the trip. For the reverse case, where trips had more alighting observations than boardings, boardings were added to the first stop of the trip. This ensured that boardings and alightings were balanced at every level from the trip up. This had the effect of adding uncounted activity to the trip in a logical way.

After the raw, trip-level boarding and alighting observations were balanced as described above, the total number of observed boardings and alightings for the region equaled 126,481 each. (This statistic includes only the observations made on the 355 patterns with complete, useable surveys—in other words, those patterns involved in the IPF weighting process.) With balancing complete, the next step toward calculating expected weekday ridership was to introduce two factors into the process: the *sample factor* and the *vehicle factor*.

[1] In an intercept interview, the surveyor asks the rider questions rather than having the rider self-complete the questionnaire.

5.3 Calculating the Sample Factor

The sample factor was calculated at the trip level. It accounts for the fact that most bus trips were surveyed only once, but some were surveyed two or three times, and it ensures that observations on trips surveyed more than once were not overrepresented in calculations.

Simply stated, the sample factor is equal to the inverse of the number of times a particular bus trip was surveyed. Since no trip was surveyed more than three times, there are only three possible values for the sample factor: 1.00 for trips surveyed once, 0.50 for trips surveyed twice, and 0.33 for trips surveyed three times. For example, if the 9:08 AM bus trip on Pattern 1 was surveyed on both Tuesday and Wednesday, then two sets of stop-level boarding and alighting observations would exist for that trip, and the data for all of these records would need to be multiplied by a sample factor of 0.50 before being combined.

Sample Factor = 1 / Number of times a trip was surveyed

5.4 Calculating the Vehicle Factor

The vehicle factor was calculated at the cell level (that is, by pattern and time-of-day strata). The vehicle factor is an expansion component in the weighting and expansion process. It accounts for the fact that not every trip in the transit system was surveyed. The vehicle factor was simply calculated by counting up the number of scheduled bus trips associated with each cell, identifying which trips were sampled (at least once) and which trips were not sampled at all, then dividing the total number of trips in the cell by the number that were sampled. For example, if Pattern 1 had a total of 10 bus trips scheduled during the AM-peak period, but only 2 of these trips were sampled, then the vehicle factor for the Pattern 1, AM-peak cell would equal 10 divided by 2, or 5. This calculation is based on the assumption that all trips associated with a particular cell should have similar ridership characteristics.

Vehicle Factor = Total trips per cell / Sampled trips per cell

5.5 Calculating Expected Weekday Ridership

With the observed boarding and alighting observations balanced, the sample factor calculated for each trip, the vehicle factor calculated for each cell, and the appropriate trips associated with each cell, it was straightforward to calculate the expected weekday boarding and alighting values for each stop on each pattern.

First, a table was compiled containing the balanced boarding and alighting observations for each stop on each surveyed trip. This table contained multiple sets of records for those trips surveyed more than once. In this table, the following calculations were made for each record:

B = Balanced observed boardings at stop * Sample factor for trip * Vehicle factor for cell

A = Balanced observed alightings at stop * Sample factor for trip * Vehicle factor for cell

Next, the “B” and “A” values calculated above were aggregated by pattern and stop, resulting in expected weekday boarding and alighting values for each stop on each pattern—the exact data needed to populate the marginal control totals for the IPF procedure:

Expected weekday boarding values by pattern and stop = ΣB by pattern and stop

Expected weekday alighting values by pattern and stop = ΣA by pattern and stop

Note that even though the calculated expected values were daily values (in other words, data was not stratified by time of day for Iterative Proportional Fitting), time-of-day biases were partially accounted for by incorporating the vehicle factor into the calculation. These biases were dealt with even more directly by incorporating the time-of-day factor into the process (see Section 8).

5.6 Example Calculation

This subsection provides an example of how the sample factor and vehicle factor were used to calculate expected weekday ridership. Assume that Pattern 2 had six trips associated with it, all scheduled during the PM-peak period—in other words, Pattern 2 was associated with six trips, but only one cell. Further assume that only two of these scheduled trips were sampled: the 4:00 PM trip and the 5:00 PM trip. Next, assume that the 4:00 PM trip was sampled on Wednesday only, but that the 5:00 PM trip was sampled on both Wednesday on Thursday. Finally, assume that at Stop 1 on Pattern 2, the following (balanced) boarding observations were made:

- 20 boardings on Wednesday during the 4:00 PM trip,
- 30 boardings on Wednesday during the 5:00 PM trip, and
- 24 boardings on Thursday during the 5:00 PM trip.

If this were the case, then the following would be true:

- the sample factor for the 4:00 PM trip would equal 1.00, as it was sampled only once;
- the sample factor for the 5:00 PM trip would equal 0.50, as it was sampled twice;
- the vehicle factor for the cell would be 6 / 2 = 3, since only two of the six trips were sampled;
- three boarding observation records would exist for Stop 1 on Pattern 2;
- the “B” values for each of these three records would be calculated as shown below:
  - $B_1 = 20 * 1.00 * 3 = 60$
  - $B_2 = 30 * 0.50 * 3 = 45$
  - $B_3 = 24 * 0.50 * 3 = 36$; and
- the expected weekday boarding value for Stop 1 on Pattern 2 would equal the sum of the three “B” values: 60 + 45 + 36 = 141.

5.7 Expected Weekday Ridership Versus Average Weekday Ridership

Expected weekday ridership, the subject of this section, should not be confused with the average weekday ridership estimates provided by the transit agencies participating in the study. These measures are two different entities. Expected weekday ridership was calculated only because average weekday ridership was not available at the disaggregate level (pattern and stop) needed for generating Iterative Proportional Fitting control totals.

It was anticipated that these two different measures, at the regional level, would be fairly close to one another but not precisely equal, and that assumption turned out to be true. The sum of all expected weekday boarding or alighting values was 217,173. The average weekday ridership for the region was 218,129 (this number excludes the Detroit People Mover to make the comparison valid).

The reason it was anticipated that expected weekday ridership would be close to average weekday ridership was confidence in the logic behind the sampling process, and confidence in the logic used to calculate expected ridership (as described in this section). In fact, it was a necessity that these values be fairly close to one another. Otherwise, using expected weekday ridership as a stand-in for average weekday ridership in the IPF procedure would not make sense.

The reason it was anticipated that the values would not precisely equal one another is that they come from different sources. Expected weekday ridership was calculated based on survey observations from a limited sample of trips, while average weekday ridership was tabulated by the transit agencies prior to the survey using their own data resources.

6. Aggregating Stops

6.1 The Need for Aggregation

The reason for calculating expected weekday ridership by pattern and stop, using the steps described in Section 5, was to prepare marginal control totals for the Iterative Proportional Fitting procedure. It would have been ideal if SEMCOG and NuStats could have used the calculated values directly—in other words, if each pattern's individual stops could have directly formed the rows and columns of the IPF tables. Unfortunately, this was not the case. It was realized early on that in order for the IPF procedure to converge, rows and columns would need to be comprised of groups of stops, rather than individual stop locations. In other words, without aggregation, data would have been stretched too thin in too many places to allow IPF to be used to calculate response factors.

With 18,099 surveys returned for the 355 patterns that were weighted via IPF (approximately 50 samples per pattern), the overall number of samples was not really the problem. Rather, the problem was due to three characteristics typical of the Southeast Michigan transit system: the large numbers of stops on patterns, the uneven ridership distribution from pattern to pattern, and the uneven distribution of boarding and alighting activity from stop to stop. Taken collectively, these three issues pointed to the need to aggregate stops on patterns.

Large Numbers of Stops

In Southeast Michigan, bus patterns are typically comprised of many stop locations. There is an average of 65 stops per weighted pattern (23,142 stops on 355 patterns). Even by ignoring the stops with no activity and considering just those locations at which boarding or alighting activity was observed during the survey, calculations yield an average of 41 stops per pattern (14,501 stops on 355 patterns). Another related statistic is the fact that 77 of the 355 patterns involved in the IPF process (22 percent) have 100 or more stops associated with them. Some patterns have nearly 200 stops. The presence of so many stops immediately signaled to SEMCOG and NuStats that some stop locations would likely need to be aggregated.

Uneven Pattern-Level Ridership Distribution

In Southeast Michigan, a relatively small number of patterns carry a relatively large share of regional ridership. As shown in Table 6.1 below, just 10 percent of patterns account for more than half of regional ridership, and just 20 percent of patterns account for more than 70 percent of ridership. Accordingly, it is also true that a relatively large number of patterns carry relatively little ridership. In fact, calculations reveal that 107 of the 355 weighted patterns (30 percent) have an expected weekday ridership of just 100 or less, and that 173 patterns (49 percent) have an expected weekday ridership of just 200 or less.

Table 6.1: Distribution of Expected Weekday Ridership by Weighted Pattern

Ridership/Pattern	Number of Patterns	Total Ridership	% Ridership	Cumulative %
>1540	35	112,985	52.2%	52.2%
801 - 1540	36	39,685	18.3%	70.5%
521 - 800	35	21,653	10.0%	80.5%
371 - 520	36	16,617	7.7%	88.2%
221 - 370	36	9,993	4.6%	92.8%
151 - 220	35	6,293	2.9%	95.7%
102 - 150	35	4,355	2.0%	97.7%
63 - 101	35	2,888	1.3%	99.1%
28 - 62	36	1,544	0.7%	99.8%
<28	36	492	0.2%	100.0%
Total	355	216,504	100%	

Naturally, low ridership implies a correspondingly low number of completed surveys. The data bears this out, with 107 patterns (30 percent) having only 10 or fewer completed surveys associated with them, and 156 patterns (44 percent) having only 20 or fewer completed surveys associated with them. The lack of a significant number of surveys on so many patterns was another indication that stop locations would likely need to be aggregated for IPF to work.

Uneven Stop-Level Activity Distribution

In Southeast Michigan, not only are there major variations in ridership from pattern to pattern, but also major variations in boarding and alighting activity from stop to stop along individual patterns. Most stops service few riders, with just a small number handling the majority of boarding and alighting activity. Table 6.2, which provides a distribution of expected weekday boarding and alighting values at the stop level, reveals this. Only about 5 percent of the region's stops are associated with half of all boarding and alighting activity. The lack of boarding and alighting activity at many stops, together with the lack of ridership and collected surveys on many patterns, plus the sheer number of stops on many patterns in the region, all indicated the need to aggregate stops into groups.

Table 6.2: Distribution of Expected Weekday Boarding and Alighting Values by Stop (On Weighted Patterns)

B+A Activity Per Stop	Number of Stops	Total B+A Activity	% Stops	% Activity
>0 and <=20	10,400	69,922	71.7%	16.1%
>20 and <=40	1,919	54,234	13.2%	12.5%
>40 and <=60	772	37,506	5.3%	8.7%
>60 and <=80	404	28,019	2.8%	6.5%
>80 and <=100	251	22,378	1.7%	5.2%
>100	755	220,948	5.2%	51.0%
Total	14,501	433,008	100.0%	100.0%

6.2 Deciding Not to Aggregate Patterns

Based on statistics such as those given in Section 6.1, the decision was made to combine stops into groups for the purpose of running the IPF procedure. However, it should be noted that SEMCOG and NuStats also considered grouping together the patterns associated with each route/direction combination. Major reasons that this idea was rejected included the following:

- Stop sequence information, as listed in the master stop table, was only available (and in fact, only has logical meaning) at the pattern level. If patterns were aggregated, the ability to align stops in sequence would have been lost, and without stop sequence, it would have been difficult to logically group stop locations for IPF purposes.
- Different patterns running in the same direction for the same bus route frequently traverse significantly different geographic areas and therefore serve riders with potentially dissimilar characteristics. Combining such patterns into groups would have needlessly compromised the precision of capturing these differences in the weighting process.
- As mentioned in Section 2.3, the TransCAD "routes" coded in SEMCOG's travel-demand forecasting model are patterns in reality. Therefore, to facilitate the comparison of survey data with transit-system input/output from the travel model, it made sense to stratify the survey data at the pattern level.

6.3 Tiers of Ridership

With the decision made to aggregate stops, the next step was to determine what method (or methods) to use to efficiently yet effectively perform this task. Noting the major discrepancies in ridership by pattern, SEMCOG decided that a one-size-fits-all approach for every pattern in the region would be unwise. Instead SEMCOG committed to investing a considerable amount of resources toward aggregating stops on patterns with higher ridership, and developing logical but less time-intensive procedures for aggregating stops on patterns with lower ridership.

Ultimately, patterns were separated into three different tiers, with different stop aggregation methods developed for each tier:

- *Tier 1*, the *high ridership* tier, consisted of 96 patterns responsible for 75 percent of the region's ridership. Nearly all of these patterns had an expected weekday ridership of 600 or more.
- *Tier 2*, the *moderate ridership* tier, consisted of 67 patterns. The expected weekday ridership for the patterns in this tier was generally in the 150 to 600 range.
- *Tier 3*, the *low ridership* tier, consisted of the remaining 192 patterns.

6.4 Aggregating Stops on Patterns with High Ridership

With the goal of capturing spatial variation in ridership characteristics as accurately as possible via the IPF procedure, SEMCOG developed an extremely detailed and thorough approach for aggregating stops on high ridership patterns. This approach consisted of two parts: 1) automatically aggregating data from the pattern/stop level up to the pattern/node level using nodes from the travel-demand model's highway network, and 2) manually combining nodes into the groups of stops ultimately used to form rows and columns for IPF tables.

Assigning Stops to Nodes

As an initial, automated aggregation step, data at the pattern and stop level was aggregated up to the pattern and node level using nodes from the HN10 version of SEMCOG's travel-demand forecasting model. Each stop was plotted in ArcMap (latitude and longitude values were available from the master stop table), then assigned the ID number of the model node nearest the stop (as the crow flies) via geoprocessing. Based on this newly created stop/node relationship, expected weekday boarding and alighting values, as well as the number of collected surveys associated with each boarding and alighting stop, were recalculated at the pattern/node level. In addition, for each pattern and node, the minimum stop sequence number (from among all of the stops associated with that pattern and node) was retained so that the correct travel path along the pattern was still identifiable.

The decision to aggregate data at the pattern and node level was made for two main reasons. First, this step reduced the number of locations that needed to be manually reviewed in the next part of the process: the original 14,501 pattern/stop locations were reduced to about 11,500 pattern/node locations—a modest reduction, but one that increased efficiency without losing much detail. Second, analyzing data at the pattern/node level allowed the total amount of boarding and alighting activity to be calculated for each location (node). Calculating which locations had relatively low, medium, and high levels of transit activity proved to be valuable during the manual grouping stage, and it was not possible to do this calculation with the data organized at the original pattern/stop level. (This is because each node has the same ID number regardless of the number of patterns associated with it, facilitating the summation of boarding and alighting data, whereas stops have different ID numbers on different patterns, even if they are located at the same physical location. Put another way, nodes are a surrogate for closely spaced real-world physical bus stops that serve the same geographic location.)

After the total amount of boarding and alighting activity at each node was calculated, locations were divided into three groups (not to be confused with the three tiers of ridership used to categorize patterns):

- The 5 percent of nodes with the most activity (each with greater than 376 expected boardings and alightings) were labeled as *high activity nodes*;
- The next 15 percent (those with greater than 93 expected boardings and alightings but fewer than 376) were labeled as *medium activity nodes*; and
- The remaining 80 percent of nodes were labeled as *low activity nodes*.

It should be noted that this GIS-based, automated aggregation step did have two minor drawbacks. First, due to a variety of factors, the node to which a stop was assigned was not guaranteed to actually be part of that stop's pattern —this was annoying but did not occur frequently and did prevent such nodes from being manually grouped later on. Second, for patterns containing sections of road that crossed one another, stops that were relatively far apart in sequence were sometimes assigned to the same node, which made grouping nodes at such locations more of a challenge—fortunately, patterns with this issue were few in the region.

Manually Grouping Nodes

The automated procedure described above was the first part of the approach SEMCOG developed to aggregate stops on high ridership patterns. The second (and far more time consuming) part was to manually decide, pattern by pattern and node by node, which nodes should logically be combined into groups. SEMCOG prepared two software tools to help accomplish this task and make the process flow relatively smoothly.

The first tool was a series of MS Excel spreadsheets listing every pattern/node combination, the expected weekday boarding and alighting data tabulated for that combination, the tabulated number of sample boardings and alightings for that combination, and the total amount of transit activity calculated for the node (as contributed by all patterns associated with the node). High activity and medium activity nodes were highlighted in blue and red, respectively, using conditional formatting. Conditional formatting was also used to highlight records with unusual data that required extra scrutiny, such as records that were assigned more sample alightings than expected alightings. Spreadsheet columns were included for typing in the group number to be associated with each pattern and node. Other columns were formulated to calculate, on the fly, the new group-level boarding and alighting totals as each new group number was entered. These measures all helped SEMCOG determine logical break points for the groups of nodes. They also helped avoid, for the most part, creating groups containing data that would later cause a pattern's IPF run to fail (not converge).

The second tool was an extremely helpful ArcMap application, developed for use in conjunction with the MS Excel spreadsheets as group numbers were recorded. In fact, SEMCOG staff typically worked in pairs, utilizing two large-screen monitors—one displaying the MS Excel application and the other the ArcMap application. The ArcMap tool displayed a map of any selected pattern—and the pattern's original stops and assigned nodes—together with aerial photography and a layer with roads and street names, providing an irreplaceable birds-eye view of the areas and neighborhoods traversed by the pattern, as well as the exact geographic alignment of the pattern's path. The map was labeled with the same boarding and alighting data tabulated in the MS Excel spreadsheets, and high-activity and medium-activity nodes on the map were color-coded to match the colors used in the spreadsheets.

Criteria Used to Group Nodes

Judgment, experience, and knowledge of the region were all key factors in the manual process of deciding which nodes should be grouped together for each pattern. As work progressed, many guiding principles evolved and were put into effect. These principles included the following:

- High activity nodes (such as nodes near transit centers and other major transfer points) were typically left alone and ungrouped;
- Low activity nodes were either combined with nearby higher activity nodes, or else strung together to form their own group, depending on other criteria;
- Boundaries between different neighborhoods or between residential and commercial or industrial areas (or any other logical visible boundaries) were frequently used to separate one group from another;
- Nodes in close proximity to one another were much more likely to be grouped than those with a significant gap between them;
- As already mentioned, groups were formed to try to avoid creating mathematical problems that would later cause the IPF process for a pattern to fail; and
- Nodes were assigned a land-use index (as explained in detail in the following paragraphs), and nodes with identical or similar indices were grouped together where feasible.

Land-Use Index

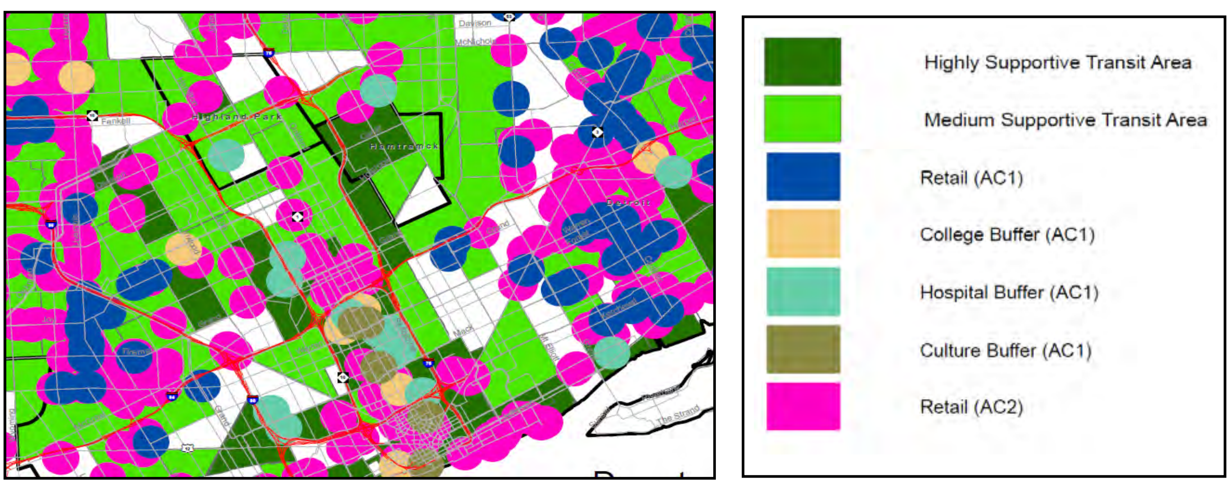
SEMCOG developed a tabulated value called the land-use index, which was instrumental in determining which nodes were similar to one another from a land-use characteristic perspective, and therefore were logical candidates for being grouped together.

The land-use index was assigned to each node, based on whether the node was located within a transit supportive area and/or located near a major activity center. For the purpose of tabulating this score, two categories of transit supportive areas were defined at the Traffic Analysis Zone (TAZ) level based on household and employment densities, and two categories of major activity centers were defined based on the type and size of the activity center, as shown below:

- Highly Supportive TAZ—Greater than 7 jobs or households per acre
- Medium Supportive TAZ—Between 3 and 7 jobs or households per acre
- Activity Center #1—Hospital, culture, university, or retail with >500,000 square feet (with buffer)
- Activity Center #2—Retail with 50,000 – 500,000 square feet (with buffer)

In ArcGIS, buffer polygons were drawn around each major activity center to facilitate overlays with the designated transit supportive TAZs. Figure 6.1 shows how the buffered activity centers and transit supportive zones can overlap one another.

Figure 6.1: Overlapping Transit Supportive Areas and Major Activity Center Buffers



Tabulated land-use indices ranged from 0 to 4, with larger values assigned to nodes with higher population and employment densities, and to nodes close to major activity centers. Table 7 provides the exact criteria used to tabulate scores.

- 0—No buffer areas at node
- 1—Medium supportive TAZ or Activity Center #2
- 2—Medium supportive TAZ and Activity Center #2, or Activity Center #1
- 3—Highly Supportive TAZ without Activity Centers or Medium Supportive TAZ with Activity Center #1
- 4—Highly Supportive TAZ with one or more Activity Centers

In words, the steps used to calculate the land-use index are as listed below:

1. The number of jobs and households per traffic analysis zone (TAZ) were tabulated, and GIS buffer areas were established around each major activity center (hospital, cultural center, university, and retail center) in the region.
2. TAZ's with greater than 7 jobs or households per acre were designated as "highly transit supportive" zones; TAZ's with between 3 and 7 jobs or households per acre were designated as "medium transit supportive" zones.
3. All major activity centers in the region, with the exception of retail centers less than 500,000 square feet in area, were designated as "level-1 activity centers"; retail centers from 50,000 to 500,000 square feet in area were designated as "level-2 activity centers".
4. A model node received a land-use index of **4** if it was located in a highly transit supportive zone *and* in the buffered area of any type of major activity center.
5. A model node received a land-use index of **3** if 1) it was located in a highly transit supportive zone but *not* in the buffered area of any type of major activity center, or 2) it was located in a medium transit supportive zone *and* in the buffered area of a level-1 activity center.
6. A model node received a land-use index of **2** if 1) it was located in a medium transit supportive zone *and* in the buffered area of a level-2 activity center, or 2) it was located in the buffered area of a level-1 activity center but *not* in a transit supportive TAZ.
7. A model node received a land-use index of **1** if 1) it was located in a medium transit supportive zone but *not* in the buffered area of any type of major activity center, or 2) it was located in the buffered area of a level-2 activity center but *not* in a transit supportive TAZ.
8. Every other model node received a land-use index of **0**. These nodes were not located in a transit supportive TAZ, nor were they located in the buffered area of any type of major activity center.

As mentioned in previous sections, the land-use index was added to both the MS Excel and ArcMap applications for reference, and nodes with similar index values were grouped together as much as feasible.

6.5 Aggregating Stops on Patterns with Moderate Ridership

Even though SEMCOG designed useful software tools to assist with aggregating stops on high ridership patterns, the Tier 1 process was so detailed and time consuming that it was neither practical nor desirable to use this process for all 355 patterns.

For patterns with moderate ridership, SEMCOG wrote a computer program to automatically assemble stops into groups for the purposes of IPF. A complete description of the program logic, which was based in part on the experience gained while manually combining nodes on high ridership patterns, is beyond the scope of this appendix. However, it should be noted that four main criteria were used for separating one group of stops from another:

1. Groups could only be comprised of consecutive stops along a pattern;
2. The total expected weekday ridership (in other words, the number of expected boardings or alightings) for all of the stops in the group needed to be greater than 100;
3. The number of survey samples associated with the group needed to be less than 50 percent of the total number of survey samples associated with the pattern; and
4. Consecutive stops could not be assigned to different groups if they were each associated with the same model node.

The motivation for developing this program was to combine stops into groups for as many additional patterns as possible while still working efficiently and saving staff time. The low ridership patterns that could not be processed using the program were handled as described in next subsection.

6.6 Aggregating Stops on Patterns with Low Ridership

NuStats aggregated stops for patterns with the lowest ridership, using logic somewhat similar to that employed in SEMCOG's aggregation program for moderate ridership patterns, but with the restrictions on nodes and expected ridership relaxed (among other changes). The essence of the process was to repetitively run patterns through the IPF procedure, each time automatically adjusting the grouping of stops until reasonable response factors were calculated.

The stops were first grouped at the most disaggregate level that could allow weights to be calculated, with boarding and alighting groups assembled independently then run through the IPF procedure to test for convergence. For patterns with groups that caused convergence problems, the groups were collapsed one at a time until convergence was achieved. Starting with the boarding groups, one target boarding group was chosen and then combined with either the previous boarding group or the subsequent boarding group. The target boarding group was the one with the greatest difference between its response rate and the average weighted response rate of its riders' alighting groups, since the difference in response rates is what typically caused the IPF procedure to fail.

After collapsing one boarding group per pattern, problem patterns were again run through the IPF procedure to test for convergence. For those patterns that still did not converge, the same aggregation process was applied, but this time on the alighting side. IPF runs and stop aggregation continued in this manner, alternately collapsing boarding and alighting groups, until all patterns converged. This sometimes resulted in a single expansion factor for all of the records associated with a pattern.

7. Calculating Response Factors Using IPF

7.1 The Response Factor

The weighting and expansion methodology developed by SEMCOG and NuStats for the SEMCOG on-board transit survey ultimately assigns a final, comprehensive expansion weight to each of the 18,495 survey records in the dataset. For surveys completed on fixed-route bus transit trips, the value of the final expansion weight is a function of the values of five key factors: the sample factor and vehicle factor (each described in Section 5), the expansion factor and time-of-day factor (each described in Section 8), and the response factor—the subject of this section.

Introduced in Section 4, the response factor is a weighting factor used to properly account for the fact that not all riders on sampled trips return complete, useable surveys, and moreover, that survey response rates can vary along the paths of sampled trips. For example, riders that board at one of the first stops along a bus trip and do not alight until reaching one of the last stops may be more likely to fill out a survey than riders who travel on the bus for short distances and spend only a small amount of time in the transit vehicle.

SEMCOG and NuStats, reasoning that variations in response rates are a function of both boarding and alighting location, and further reasoning that the attributes of transit riders themselves vary by both boarding and alighting location (especially for the many patterns in Southeast Michigan that have large numbers of stops and cover long distances through areas with diverse characteristics), decided to stratify the survey data by pattern, boarding location, and alighting location for the purpose of calculating response factors, rather than stratifying by boarding location alone as is often done in on-board transit surveys. To calculate response factors by incorporating boarding and alighting location simultaneously, SEMCOG and NuStats made use of the Iterative Proportional Fitting (IPF) procedure, the mathematics of which was explained in Section 4.2.

7.2 Running the IPF Procedure

The IPF procedure was run (separately) for each of the 355 bus patterns for which complete, useable surveys were received. The six statements listed in Section 4.2 about how the IPF tables were designed for this project are restated here, but reworded to take into account the information provided in Section 5 and Section 6:

1. IPF rows and columns represented groups of boarding and alighting stops, respectively. The stops associated with each group were aggregated as described in Section 6;
2. Each row's marginal control total equaled the sum of the expected weekday boarding values for all of the stops associated with the group represented by the row. Stop-level expected weekday boarding values were calculated as described in Section 5, then summed up for the group;
3. Each column's marginal control total equaled the sum of the expected weekday alighting values for all of the stops associated with the group represented by the column. Stop-level expected weekday alighting values were calculated as described in Section 5, then summed up for the group;
4. Based on the relationships established between stops and groups for each pattern, the number of survey records associated with each "group on" and "group off" combination were tabulated. These values became the seeds of the IPF tables;

5. The procedure generated a matrix of adjusted values representing the number of riders expected to board and alight via each possible group of boarding stops and each possible group of alighting stops, respectively; and
6. Response factors were calculated by dividing each adjusted value by its original seed.

NuStats performed the final, official IPF runs for every one of the 355 weighted patterns. During an IPF run, for each row and column, the difference between the control total and the sum of the adjusted values was continuously recalculated from iteration to iteration. When that difference dropped below 1.0 for each row and column in the entire IPF matrix, the program terminated, indicating that convergence was reached.

7.3 IPF Convergence Issues

As alluded to throughout this appendix, an IPF run is not guaranteed to converge given any arbitrary set of seeds and control totals. In the context of this project, the general reason that IPF runs failed was that the data was spread too thin—that is, if there were not enough surveys in parts of the matrix, relative to the number of groups assembled for a given pattern.

More specific reasons for lack of convergence included the following:

- Groups with expected boarding or alighting values equaling zero, but with a non-zero number of associated surveys;
- Groups with no associated surveys, but with non-zero expected boarding or alighting values; and
- Groups with such a small number of surveys in a row or column that when seeds were adjusted, there was no possible mathematical manipulation that could keep the sum of the adjusted values from exceeding the row control totals when columns were balanced, or vice versa.

As mentioned in Section 6.4, SEMCOG calculated running totals of group-level boarding and alighting statistics in its MS Excel application to maximize the chances that convergence would occur when the IPF procedure was run for high ridership patterns. Nevertheless, for both high ridership and moderate ridership patterns, some patterns failed to initially converge, and aggregation adjustments were required. SEMCOG made some of these adjustments, reviewing the problematic groups and re-aggregating nodes and stops as necessary. NuStats also assisted with adjustments, primarily by relaxing the criteria (initially used by SEMCOG) that groups must be comprised of identical sets of stops for both boarding and alighting. For low ridership patterns, convergence wasn't an issue in the same way it was for other patterns, because the procedure used to automatically generate groups of stops for low ridership patterns required convergence by definition (see Section 6.6).

Even when the IPF procedure successfully converged for each pattern, aggregation adjustments were sometimes made to tweak the response factors so that, as much as feasible, individual factors were neither too high nor too low. This was done to keep individual surveys from representing too many or too few expected riders, respectively. In the end, this effort was successful: the response factor was less than 60 for more than 99 percent of the samples, and only a handful of response factors were initially calculated to be less than 1.0. Since fractional response factors are not logical, such values were rounded up to equal exactly 1.0.

8. Final Calculations

After the IPF procedure was run, two important factors still needed to be calculated before final expansion weights could be assigned to each sample record in the on-board survey dataset. These two factors were the *expansion factor* and the *time-of-day factor*.

8.1 Calculating the Expansion Factor

As stated in Section 5.7, the sum of expected weekday boarding values (or alighting values, which equaled boarding values because of balancing) was 217,173. Since the IPF procedure used expected ridership for control totals, and since the purpose of the procedure was to calculate adjusted numbers that summed up to equal the control totals, it was expected that the sum of the response factors (the output of the IPF procedure) would also equal 217,173. It turned out that the sum of the response factors was a slightly larger value of 217,181 instead. Two possible reasons for the tiny difference between these numbers include the following: 1) convergence terminated the IPF process when the sums of adjusted values were close in value to, but not necessarily equal to, the marginal control totals; and 2) as stated in Section 7.3, response factors that would have been less than 1.0 if left alone were rounded up to equal 1.0.

With the sum of the response factors equaling a value of 217,181 riders per weekday, but with average weekday ridership for the region equaling a value of 218,129 (excluding the Detroit People Mover), and with average weekday ridership serving as the ultimate expansion target, a factor needed to be incorporated into the methodology to proportionally align the IPF weights to average weekday ridership. The *expansion factor* serves this purpose.

As noted in prior sections in this appendix, the most disaggregate level at which average weekday ridership was uniformly available across the region was the line level. Therefore, by necessity, the expansion factor was calculated at the line level—in other words, a separate expansion factor was calculated for each line. The calculation itself was a simple one: the average weekday ridership for the line divided by the weighted expected weekday ridership—that is, divided by the sum of the response factors for every survey associated with the line.

For example, assume that Pattern 1 and Pattern 2 were the only two patterns associated with Route 1, and that Route 1 was the only route associated with Line 1. Further assume that the sum of all the response factors calculated for Pattern 1 and Pattern 2 was 7,270, and that the line's average weekday ridership was 7,742 (as provided by the transit agency operating the line). In this case, the expansion factor for Line 1 would be 7,742 / 7,270 = 1.06, and this factor would be applied to every survey associated with that line.

Expansion Factor = Average Weekday Ridership / Weighted Expected Weekday Ridership

8.2 Calculating the Time-Of-Day Factor

<u>Initial Expansion Weight</u>

The calculation of the response factor—for which data was stratified by pattern, boarding location, and alighting location—together with the calculation of the expansion factor, allowed a value referred to as the *initial expansion weight* to be calculated for each survey record in the dataset:

Initial Expansion Weight = Response Factor * Expansion Factor

With the inclusion of the response factor, the initial expansion weight aligned the sample data with the spatial distribution of ridership observed during the survey. Because of the inclusion of the vehicle factor in IPF input calculations, the initial expansion weight also took into account the distribution of ridership by time of day. However, since weighting via IPF was carried out for the day as a whole, and since patterns may be associated with more than one time period during the day, the survey team ultimately decided to make one further adjustment to better align the sample data with observed variation in ridership by time period. This adjustment was accomplished via the *time-of-day factor*.

The time-of-day factor was calculated at the cell level (that is, by pattern and time of day) based on the time periods established by each transit agency (see Section 2.6), and was tabulated so that final expansion weights would match new cell-level expansion targets called *expected riders*. It should be noted that *expected riders* is a different variable than the similar sounding *expected weekday ridership* first introduced in Section 5 (and also known as *expected weekday boarding and alighting values*).

<u>Inflated Observations</u>

The first step in calculating expected riders was to generate a new variable called *inflated observations*. Recall from Section 5.5 that a table was created containing the balanced boarding and alighting observations for each stop on each surveyed trip, and that in the table, a variable called *B* was calculated as follows for each record:

B = Balanced observed boardings at stop * Sample factor for trip * Vehicle factor for cell

For the purpose of deriving inflated observations, a related variable called B_{exp} was calculated by multiplying "B" from above by the expansion factor described in Section 8.1:

$$B_{exp} = B * \text{Expansion factor for line}$$

Next, inflated observations for each cell were calculated by aggregating the stop-level B_{exp} values by pattern and time of day:

$$\text{Inflated observations by cell} = \Sigma B_{exp} \text{ by pattern and time of day}$$

<u>Expected Riders</u>

To transform *inflated observations* into the desired expansion target, *expected riders*, a complicated but logical series of computations were made. It is beyond the scope of this appendix to explain these computations in full, but it should be noted that they served the following three purposes: 1) reallocating data from cells that were sampled but not associated with any completed surveys; 2) accounting for the slight difference (mentioned in Section 8.1) between expected weekday ridership and the sum of post-IPF response factors; and 3) merging cells together as necessary.

It should also be noted that because of the third point above (the merging of cells, which occurred only a handful of times), most time-of-day factors were calculated at the cell level, but a few were calculated based on a combination of cells. SEMCOG and NuStats did not introduce new terminology to refer to cells grouped for this purpose—it was decided that it was accurate enough to state that the time-of-day factor was calculated at the cell level.

Factor Calculation

Once expected riders per cell were calculated, it was straightforward to calculate the time-of-day factor itself:

Time-of-day factor = Expected riders per cell / Initial expansion weight

For example, if there were 50 riders expected to board Pattern 1 during the midday period, but 40 was the sum of the initial expansion weights for surveys associated with the midday cell for Pattern 1, then the time-of-day factor for each of these surveys would be 50 / 40 = 1.25. In this example, surveys from the midday time period were underrepresented in the IPF weighting process relative to surveys for the pattern from other time periods, and the time-of-day factor was necessary to bring the sample data back in line with period-level observations.

Comments About the Time-of-Day Factor Calculation

As alluded to previously, in theory samples could have been stratified by time of day throughout the entire weighting and expansion process. However, SEMCOG and NuStats chose instead to follow the steps described above—in other words, to adjust the survey data with a time-of-day factor after weighting the data by pattern. Several related reasons factored into this decision.

First of all, if samples were stratified by time of day prior to Iterative Proportional Fitting, then an IPF run would have been required for each pattern and time period combination, rather than just one run per pattern. In all, well over 1,000 IPF runs would have been needed, stretching the survey data very thin. With many fewer surveys relative to the number of IPF runs, much more stop aggregation would have been required in order for runs to converge, and precision in capturing spatial variations in ridership characteristics would have been compromised.

Furthermore, as noted elsewhere, due to the lack of available disaggregate boarding and alighting data from the transit agencies participating in the survey, stop-level ridership estimates were calculated based on the boarding and alighting activity observed in the field during the survey. Since not all trips were surveyed, those observations were necessarily limited, and daily estimates were likely to be more accurate than estimates by time period. Accordingly, SEMCOG and NuStats felt more confident using daily estimates for the IPF weighting process rather than estimates by time period. Again, the goal was to avoid compromising the accuracy of the response factors and to preserve observed spatial ridership distributions as much as possible, but to still—in the end—align the sample data with targets based on period-level observations.

8.3 Calculating Final Expansion Weights

The calculation of the final expansion weight for each survey record was simply an adjustment of the initial expansion weight using the time-of-day factor:

Final expansion weight = Initial expansion weight * Time-of-day factor

It should be noted that with the calculation of final expansion weights, the weighted and expanded survey dataset represents typical weekday unlinked boarding/alighting trips, where in this paragraph only, the term *trips* refers to *person-trips*, not to scheduled vehicle trips (as defined in Section 2.1). Of course, in reality, passengers often ride on more than one transit vehicle as they travel from their origins to their destinations. When the survey dataset must be used to perform calculations involving individuals' origins and destinations (rather than boarding and alighting locations), the unlinked trip weights (in other words, the final expansion weights) must be converted into linked trip weights.

This is accomplished by simply dividing the final expansion weight by the total number of transit vehicles used by the survey respondent.

9. Summary

For the SEMCOG on-board transit survey, SEMCOG and NuStats collaboratively developed a new, detailed weighting and expansion approach that goes beyond traditional boarding-based methods. The approach centers around the Iterative Proportional Fitting (IPF) procedure, incorporating both boarding and alighting locations into the weighting process. The Federal Transit Administration strongly recommended using this emerging approach, as it better captures the spatial distribution of rider characteristics and better accounts for stop-level variation in survey response rates.

To prepare the input needed to run the IPF process, SEMCOG and NuStats had to overcome limitations in the data provided by the transit agencies participating in the survey, doing so by developing innovative methods for calculating ridership targets at the desired disaggregate level (by pattern and stop). These methods made use of the boarding and alighting observations that were collected at each stop along each sampled trip, transforming these measures into estimates of typical weekday ridership. Calculations at this stage incorporated two important factors to help control survey biases: the sample factor (which accounts for trips sampled more than once) and the vehicle factor (which accounts for the fact that only a subset of all typical weekday trips were sampled).

For the IPF procedure to work, the survey team also needed to aggregate boarding and alighting stop locations into groups. SEMCOG and NuStats developed and implemented an efficient, practical approach for accomplishing this task, accommodating not only the technical requirements of the IPF procedure, but also preserving observed regional travel patterns as much as feasible. In the SEMCOG region, transit services cover a large and diverse area, consisting of many long patterns with many stops—patterns with travel characteristics such as total ridership, service focus, route density, and land-use type, which vary widely from location to location. Capturing this variation was an important goal.

Patterns were divided into three different categories based on ridership, with different aggregation approaches used for different categories. The vast majority of resources were allocated to patterns with the highest ridership, ultimately encompassing patterns accounting for 75 percent of the region's average weekday ridership. For these patterns, both GIS-based automated procedures and detailed manual work were involved. Staff adopted criteria for logically aggregating stops into groups, and did the manual portion of the work with the help of in-house software applications that summarized and displayed many helpful measures, such as stop boarding and alighting activity levels and the land-use characteristics along each pattern.

With estimates of stop-level ridership available and with stops logically aggregated into groups, NuStats performed the official IPF runs for each of the 355 surveyed and weighted patterns in Southeast Michigan, resulting in the calculation of response factors. When IPF runs failed to properly converge, adjustments were made to the grouping of stops as needed. Adjustments were also made to prevent the calculated response factors from being unreasonably large or small.

To complete the calculation of final expansion weights after the IPF procedure was run, the survey team further adjusted the data using two additional factors: the expansion factor and the time-of-day factor. The purpose of the expansion factor was to proportionally align the IPF response weights to match known average weekday ridership values. The purpose of the time-of-day factor was to better align the survey data with observed variation in ridership by time period.

Calculating the time-of-day factor involved calculating final expansion targets for each pattern and time-period combination. These target values, known as *expected riders*, were a function of observed boarding and alighting activity, the sample factor, the vehicle factor, and the expansion factor. The final expansion weight allocated to each of the records in the survey dataset was a product of the time-of-day factor and an initial expansion weight based on the IPF response factor.

As a final note, the survey team believes that the approach described in this appendix makes the best use of available data by balancing the need to incorporate spatial distribution in ridership and consideration of land-use characteristics with the need to preserve the observed temporal distribution of sampled trips.

Appendix G: Assignment Report

ON-BOARD TRANSIT SURVEY DATA BASED TRANSIT ASSIGNMENT INVESTIGATION

Prepared for

Southeast Michigan Council of Governments

Report
November 2011

By

Table of Contents

List of Figures

List of Tables

1 Introduction

This report documents the on-board transit survey data-based transit assignment investigation performed by CDM Smith in October and November, for the Southeast Michigan Council of Governments (SEMCOG).

1.1 Project Description and Purpose

From April, 2010 to April, 2011, SEMCOG conducted an on-board transit survey, during which 18,495 surveys were collected. The original survey data have been post-processed and expanded to represent total daily transit boardings within the SEMCOG area. CDM Smith was tasked to compare the differences and similarities between the on-board survey results and the application of those results to the region's travel demand model.

The purpose of this investigation is to identify, explore and explain the difference between the SEMCOG model results and on-board transit surveys. Based on the findings, recommendations will be provided to guide future efforts to improve the regional model's performance as well as the quality of on-board transit surveys.

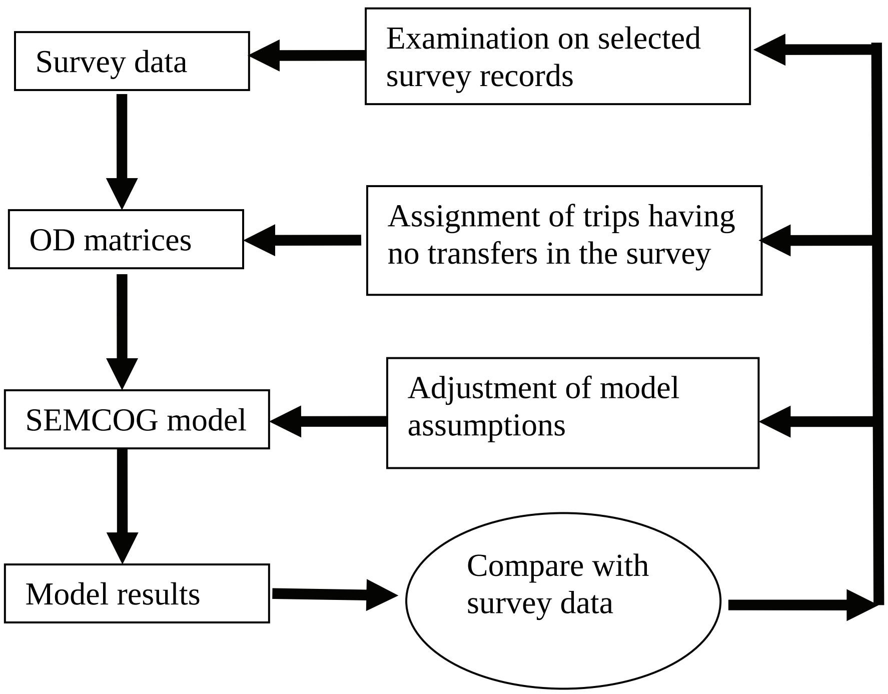
1.2 Investigation Flow Chart

The flow chart shown in **Figure 1** provides an overview of the investigation. The investigation started by converting the survey data into origin-destination matrices, which serve as inputs to the transit assignment module in the SEMCOG model. After the initial comparison between survey data and model results, four avenues were explored to understand the differences found between the survey data and model results:

1. Comparison of the survey and model skims by alternative transit assignment algorithms in TransCAD – The normal assignment method for transit trips in TransCAD is known as Pathfinder. This method finds multiple likely transit paths based on the bus frequency at the first boarding stop and at potential transfer points. This method was compared to a single shortest (all-or-nothing) path method. The results are discussed in Section 5.
2. Adjustment of model assumptions such as transfer penalty time - The number of transfers at the system level was one of the main parameters examined. The difference in the number of transfers between the survey data and the SEMCOG model results could be a result of model assumptions such as transfer penalty time values. An adjustment in transfer penalty time was made to shed some light on the reasonableness of the current transfer penalty time. The results are discussed in Section 6.2.
3. Assignment of survey trips with no transfer. The consultant team assigned only trips with no transfers in the survey to understand and identify discrepancies between the SEMCOG model results and the survey data. Singling out these trips in the assignment excludes the impact of trips involving more complex routes and may reveal factors that could have been obscured in the overall assignment results. The results are discussed in Section 6.3.

4. Examination of selected survey records - Examination of selected survey records provided specific examples to guide the investigation and support the conclusions reached. The results are discussed in Section 7.

Figure 1: Investigation Flow Chart



2 Survey Data Description

Two files were contained in the on-board transit survey data package from SEMCOG: *SEMCOG_DELIV_083111.mdb* and *SEMCOG_OB_DataMatrix_083111.xls*. The former contains 18,495 records of surveyed and expanded transit data. The latter serves as a dictionary explaining the meaning of each field in the *mdb* file. The following fields in the *mdb* file were used to convert the on-board transit survey data into OD matrices, which are needed in the transit assignment module: GETTO, IMP_TOBUS, GFROM, OGEOTAZ, DGEOTAZ, BTAZ, FTAZ, ROUTE, TOD and ExpWgt. **Table 1** documents the description of each field and its set of values.

Table 1: Fields Used in the Survey

Fields	Description	Values
GETTO	Access Mode	1=Walked/Wheelchair; 2=Dropped off; 3=Drove alone; 4=Carpooled; 5=Bicycled; 6=Taxi
IMP_TOBUS	Total Buses Taken	1=One;2=Two;3=Three;4=Four
GFROM	Egress Mode	1=Walk/Wheelchair; 2=Picked up; 3=Drive alone; 4=Carpool; 5=Bicycle; 6=Taxi
OGEOTAZ	Origin TAZ	N/A
DGEOTAZ	Destination TAZ	N/A
BTAZ	Boarding TAZ	N/A
FTAZ	Alighting TAZ	N/A
ROUTE	Route name	N/A
TOD	Time of day	N/A
Expwgt	Total Expansion Weight	N/A

3 OD Matrix Conversion Assumptions

The survey data were converted into OD matrices in the format that SEMCOG model uses as inputs. To avoid double-counting passengers who transfer and use more than one bus during a one-way trip, the data value in the field "ExpWgt" in each on-board transit survey record was adjusted – we divided it by the value in the field "IMP_TOBUS". Then the adjusted survey data were separated into twelve different groups based on different time periods and access modes, which are shown in Table **2**.

Table 2: Time Periods and Access Modes

Time Period		Access Mode
AM		Walk
MD		Auto (PrB)
PM		Auto (AtB)
OP		

The access mode was determined from data in the Access and Egress Mode fields, as shown in **Table 3**.

Table 3: Access Mode Look-Up Table

Access Mode	*Field "GETTO"*	*Field "GFROM"*
Walk	Walk/Bicycle	Walk/Bicycle
Auto (PrB)	Dropped off/Drove alone/Carpooled/Taxi	Walk/Bicycle
Auto (AtB)	Walk/Bicycle	Dropped off/Drove alone/Carpooled/Taxi
Auto (PrB)	Dropped off/Drove alone/Carpooled/Taxi	Dropped off/Drove alone/Carpooled/Taxi

For access walk trips, the origin TAZ and destination TAZ are the natural starting and ending points for the assignment of transit trips. For auto access trips, if the origin TAZ and destination TAZ were used as inputs to the travel demand model, the model would not be able to differentiate the non-transit segment of the trip from the transit segment of the trip in the assignment process. Therefore, "true" origins and destinations satisfying the transit assignment module in the SEMCOG model were defined and used in the process of developing transit trip O/D matrices, as shown in **Table 4**.

Table 4: "True" Origins and Destinations for Transit Trip Assignment

Access Mode	*True Origins*	*True Destinations*
Walk	Origin TAZ	Destination TAZ
Auto (PrB)	Boarding TAZ	Destination TAZ
Auto (AtB)	Origin TAZ	Alighting TAZ
Auto (PrB)	Boarding TAZ	Alighting TAZ

The resulting OD tables were imported into four matrices, *AM_OD.mtx*, *MD_OD.mtx*, *PM_OD.mtx and OP_OD.mtx*, which are the inputs for the transit assignment module in the SEMCOG model. The total number of trips in the OD matrices is summarized in **Table 5**.

Table 5: OD Trip Summary for the Model

Survey OD	AM	MD	PM	OP	Total
Walk	26,754	50,855	43,785	16,004	137,399
Auto Access	5,772	7,229	9,875	4,200	27,076
Sum	**32,526**	**58,085**	**53,660**	**20,204**	**164,475**

4 SEMCOG Model Description

The SEMCOG travel demand model was obtained from SEMCOG via the FTP link (ftp://ftp.semcog.org/outgoing/PlanPolicyDevelopment/GJ10_Chen_20110923/) on October 5, 2011. The model comprises eight modules: *Initialization, Trip Generation, Network Skimming, Trips Distribution, PA to OD, Mode Split, Assignment and Transit Assignment*. The Transit Assignment module is the most relevant to this investigation.

Not all the modules in the SEMCOG model were run. Other necessary inputs files for the transit assignment module generated by other preceding modules, such as transit skims, were taken directly from the output files of a completed run by SEMCOG, which were sent to CDM Smith via the same FTP link mentioned above.

5 Comparison between Survey and Model Skims with Alternative Assignment Algorithms

The SEMCOG model applies TransCAD's transit Pathfinder to perform the transit assignment. Before comparing transit assignment results, skimming was performed on the transit network to obtain 'best paths' between each of the survey OD pairs. The main purpose of this exercise was to examine how well the model's transfer rates compared to reported transfers in the survey. The numbers of transit trips involving transfers are shown in **Table 6** by time-of-day.

Table 6: Number of Transfers from Expanded Survey Data

Period	Number of Transfers				
	0	1	2	3+	Total
AM	21,632	18,165	4,823	815	45,435
MD	41,204	28,055	7,770	1,053	78,082
PM	14,247	9,980	2,486	555	27,268
OP	39,027	23,851	7,062	1,415	71,355
Total	116,110	80,052	22,143	3,838	222,143
Percent	52.3%	36.0%	10.0%	1.7%	100.0%

In the Pathfinder method, the best path is a set of links known as hyperpaths, which consist of a set of possible paths, each with different number of transfers. The skim output contains a weighted average of the number of transfers for each interchange, which often produces a decimal value. For example, a zone pair can have a 1.5 transfer in the skim output meaning that there could be any number of paths with 0,1 or 2 (maybe even 3) transfers possible leading to a weighted average (based on frequency) of 1.5. The transfer categories from Pathfinder results are grouped as follows:

- Zero transfer
- Between 0 and 1 transfer
- 1 transfer
- Between 1 and 2 transfers
- 2 transfers
- More than 2 transfers

This complicates the comparison of survey results to the transit path finder results. Hence, a simple shortest path skimming procedure was also performed. Although travelers do not always choose the shortest path, this method permits a more direct comparison to the survey transfers. This section contains comparisons between the reported transfers with Pathfinder and the shortest path algorithm. The comparison is made by time of day and, in certain cases, at each operator level. Results are presented for the following combinations of time-of-day and transit operators:

- Pathfinder Transit Assignment Method
  - Transit Survey vs. Pathfinder Model Skimming (AM)
  - Transit Survey vs. Pathfinder Model Skimming (MD)
  - Transit Survey vs. Pathfinder Model Skimming (AM, SMART)
  - Transit Survey vs. Pathfinder Model Skimming (AM, DDOT)
  - Transit Survey vs. Pathfinder Model Skimming (AM, UMI)
- Shortest Path Transit Assignment Method
  - Transit Survey vs. Shortest Path Model Skimming (AM)
  - Transit Survey vs. Shortest Path Model Skimming (MD)
  - Transit Survey vs. Shortest Path Model Skimming (AM, SMART)
  - Transit Survey vs. Shortest Path Model Skimming (AM, DDOT)
  - Transit Survey vs. Shortest Path Model Skimming (AM, UMI)

5.1 Pathfinder Transit Assignment Method

The following observations can be made regarding the comparison between transfers reported in the survey and the Pathfinder assignment results. The default path-building parameters for the SEMCOG model are specified in a file named “transitPara_E5.bin”.

In **Figure 2**, for all the survey records that reported 0 transfers in the AM peak period, the model predicts that nearly 70% are indeed 0 transfers but the remaining 30% have transfers in the set of best paths possible. The Pathfinder model also predicts that 30% of the 1 transfer reported trips have, in fact, best paths with 0 transfers. The trends are almost the same for the midday period as shown in **Figure 3**. **Figure 4**, **Figure 5**, and **Figure 6** show similar comparisons at the operator level for the three biggest operators - SMART, DDOT and UMI respectively. For SMART and DDOT operators, the model seems to predict many 0 transfer best paths, even when the survey reports more than 1 transfer. As noted below, the survey and model compare well for 0 transfer survey records for UMI. However, the model predicts more 1+ transfer trips than the survey produces.

5.2 Shortest Path Transit Assignment Method

To make a clearer one-to-one comparison between model and survey transfers, the shortest path algorithm was used. From **Figure 7**, it can be noted that the shortest path algorithm chooses a path with 0 transfers for 80% of the trips. There are around 18% of trips that need 1 transfer, according to the shortest path algorithm. The difference between the model and the survey is greater when we look at the survey records that reported 1 transfer trips. In this case, the shortest path algorithm found around 40% of these trips need no transfer. Similar results can be seen for the MD peak as shown in **Figure 8**. A similar trend continues for **Figure 9** and **Figure 10** that cover operators SMART and DDOT in the AM peak. For DDOT, the model seems to assign a lot (about 90 percent) of 0 and 1 transfer trips for survey records that reported 2 transfers. **Figure 11** shows the comparison for UMI, which compares really well for the 0 transfer trips but not the 1 or 2 transfer survey records.

Figure 2: Transit Survey vs. Pathfinder Model Skimming (AM)

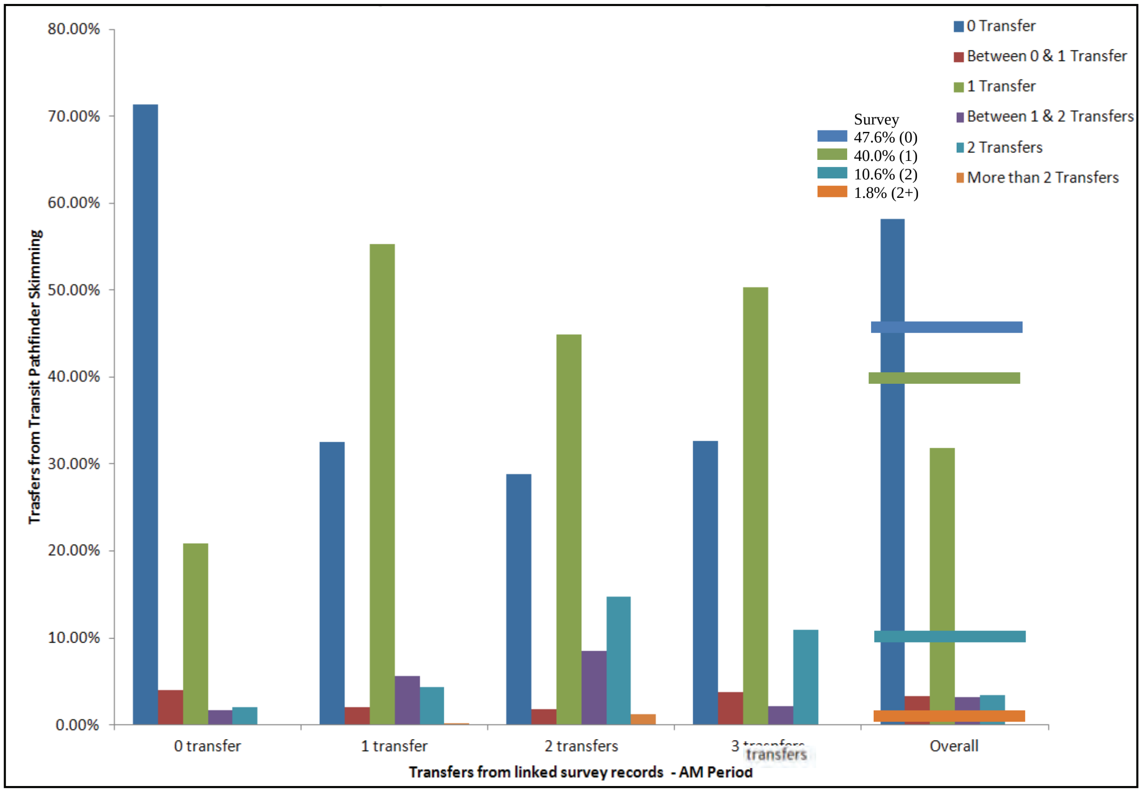


Figure 3: Transit Survey vs. Pathfinder Model Skimming (MD)

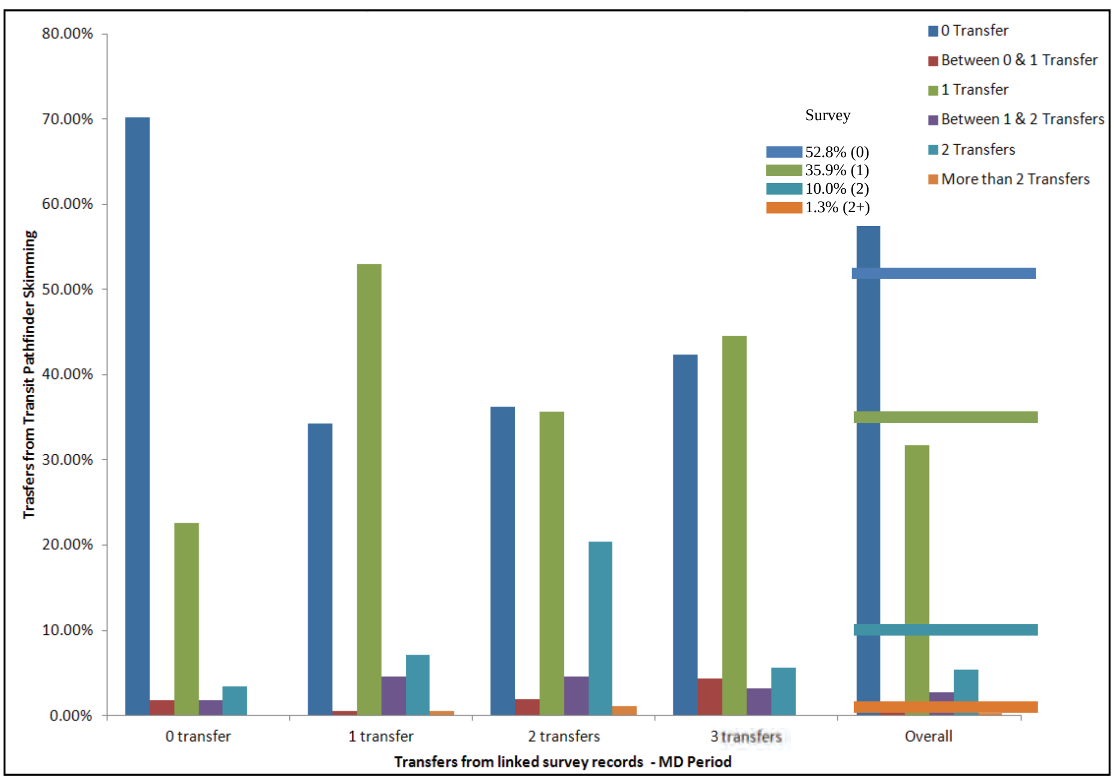


Figure 4: Transit Survey vs. Pathfinder Model Skimming (AM, SMART)

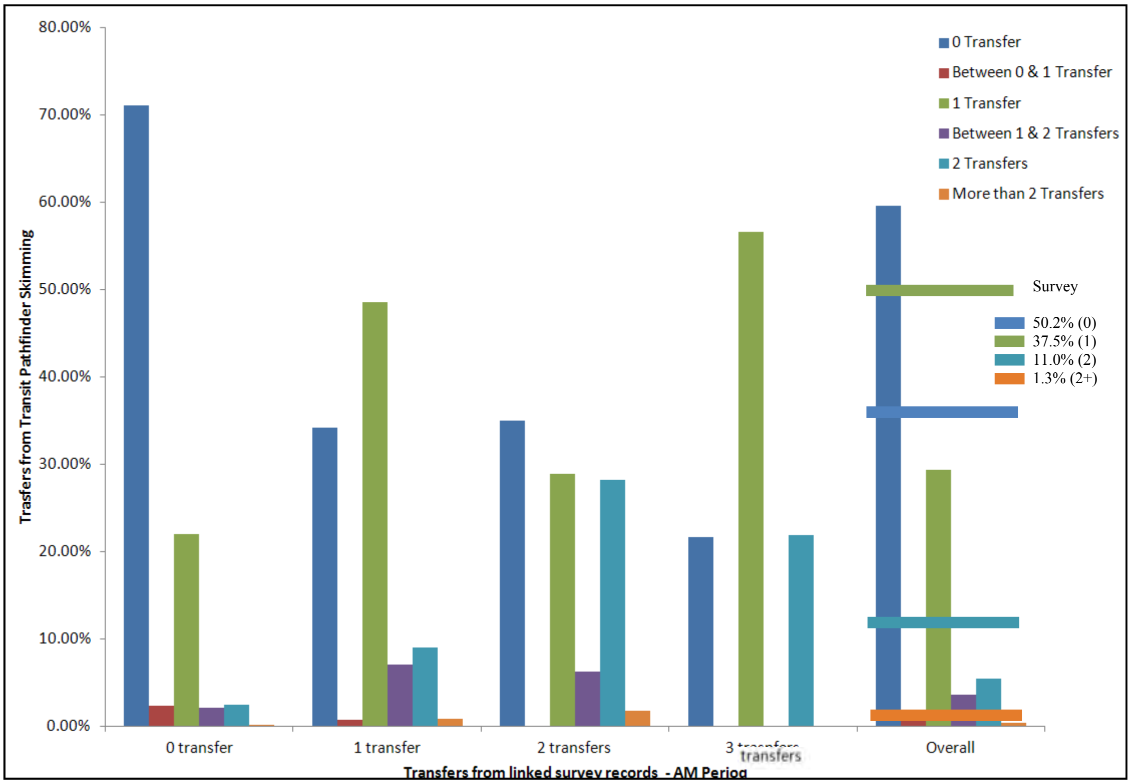


Figure 5: Transit Survey vs. Pathfinder Model Skimming (AM, DDOT)

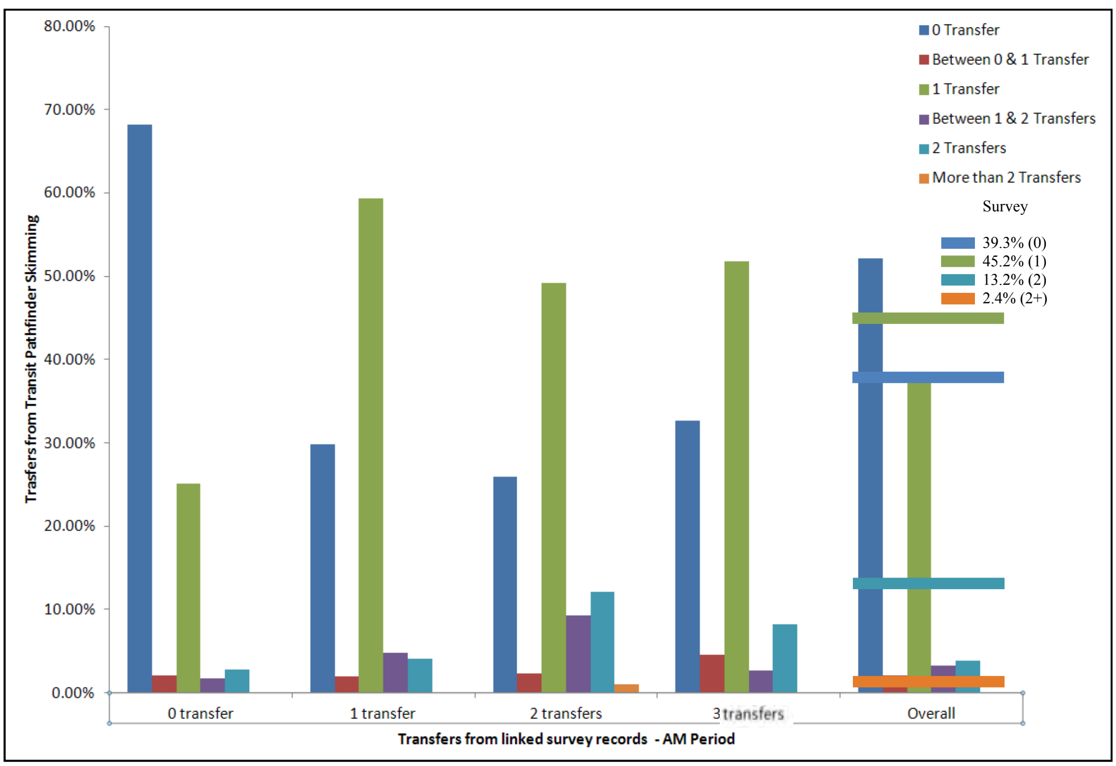


Figure 6: Transit Survey vs. Pathfinder Model Skimming (AM, UMI)

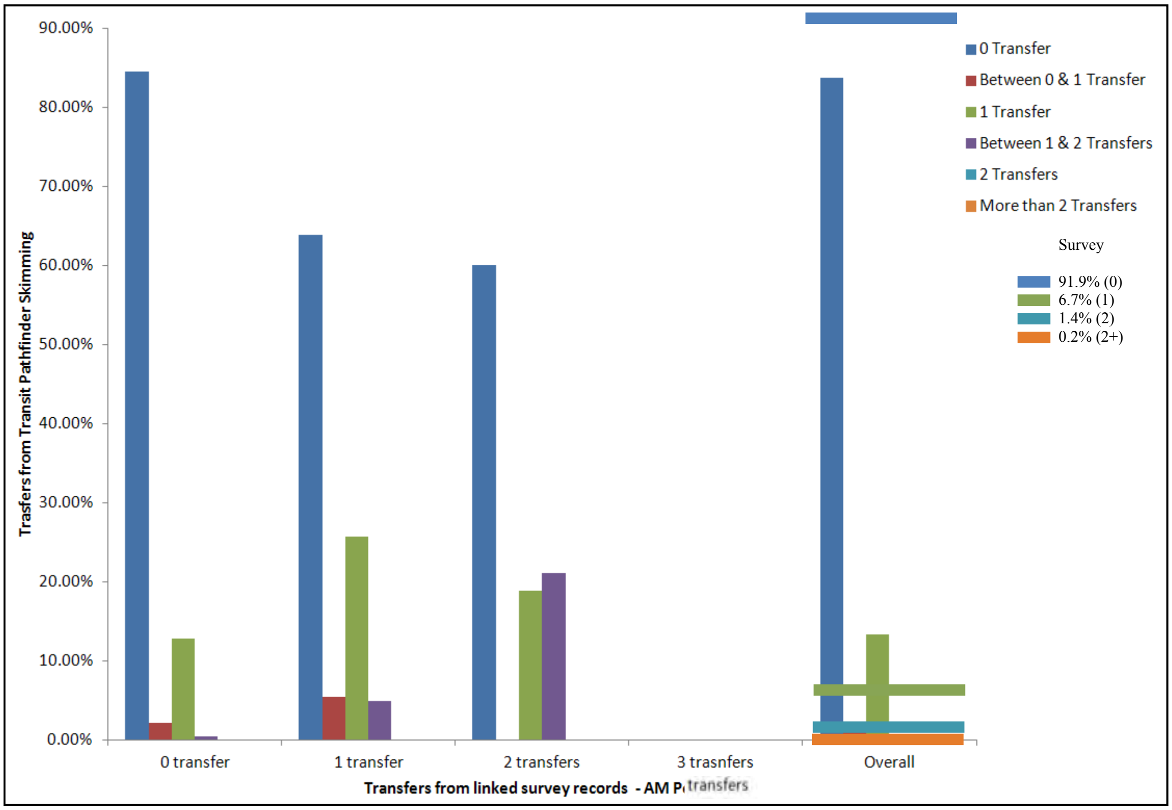


Figure 7: Transit Survey vs. Shortest Path Model Skimming (AM)

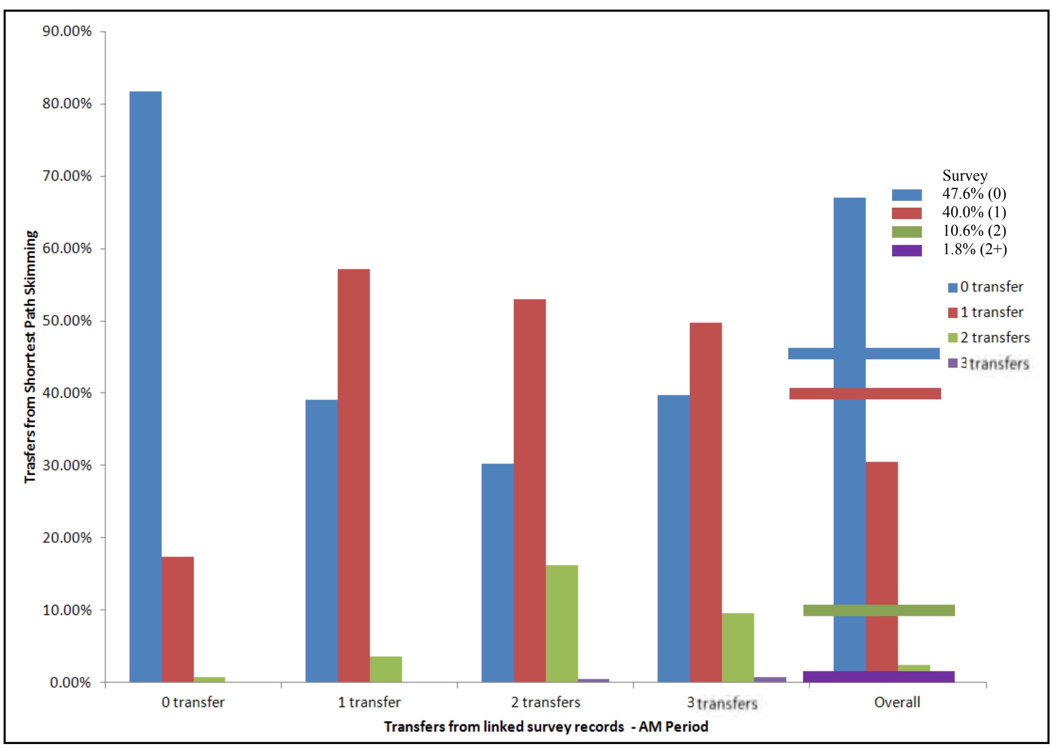


Figure 8: Transit Survey vs. Shortest Path Model Skimming (MD)

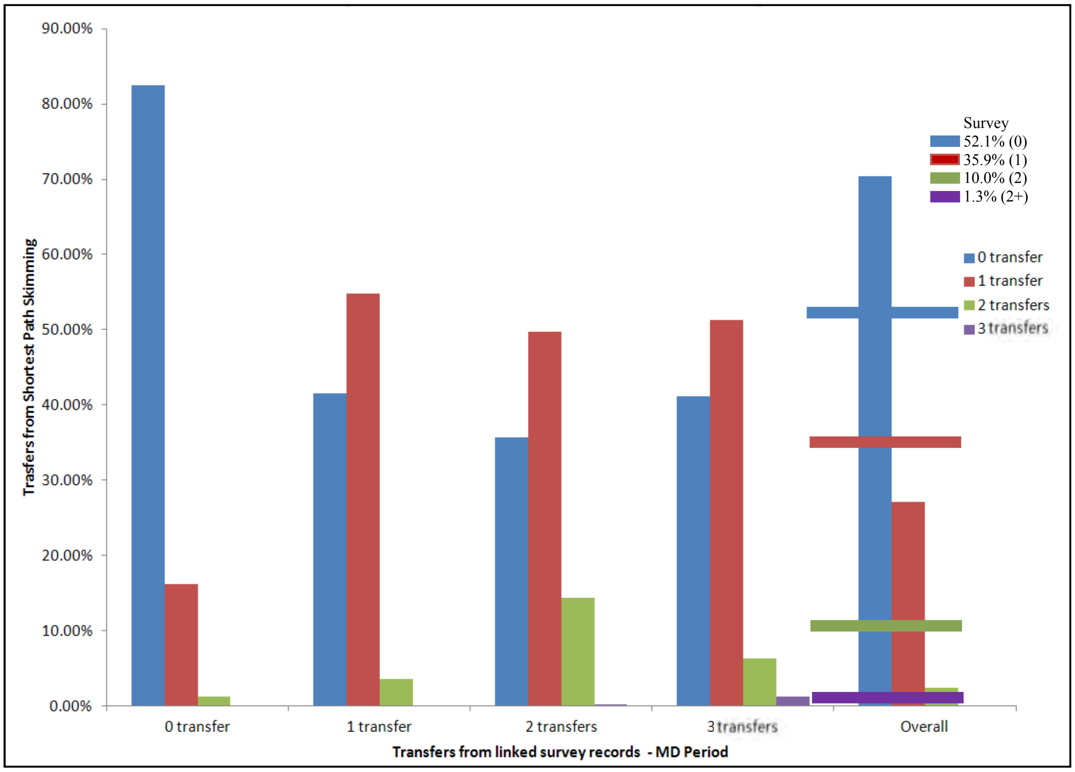


Figure 9: Transit Survey vs. Shortest Path Model Skimming (AM, SMART)

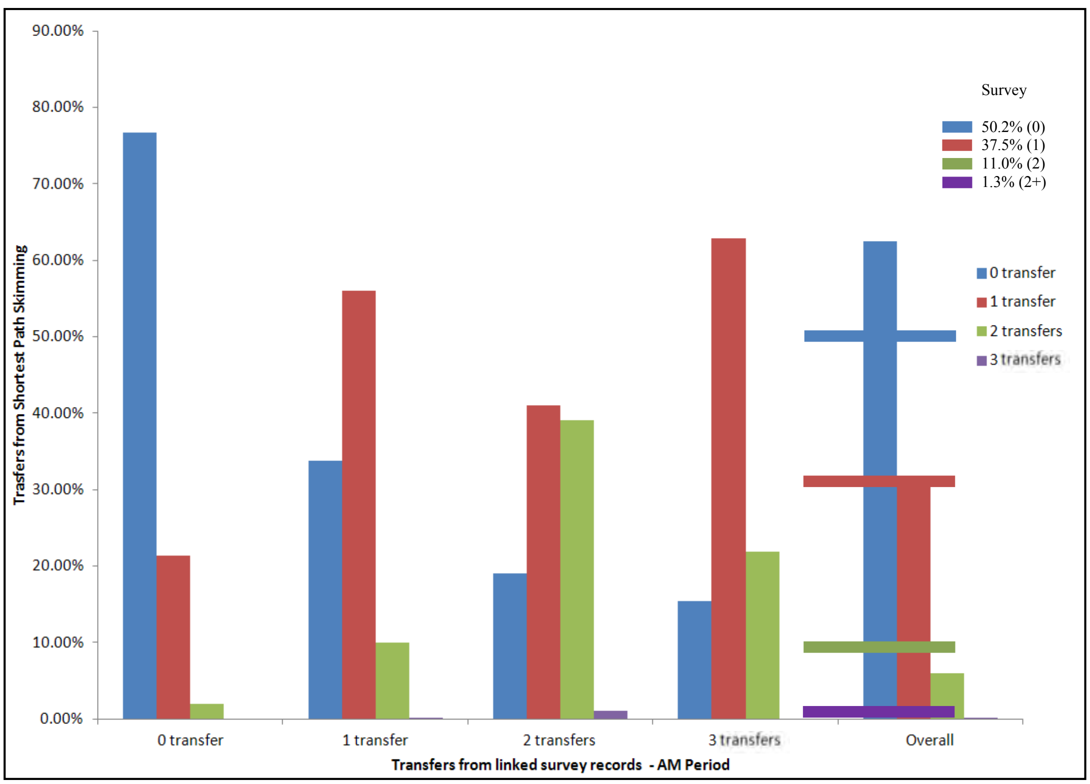


Figure 10: Transit Survey vs. Shortest Path Model Skimming (AM, DDOT)

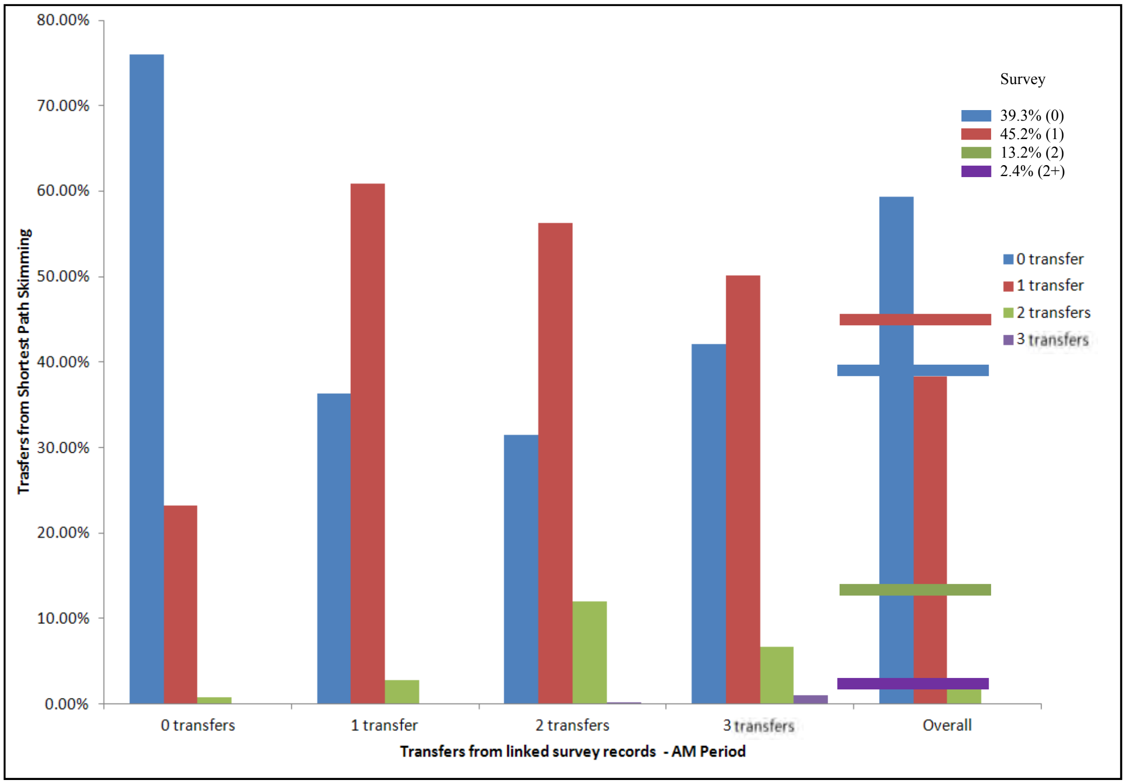
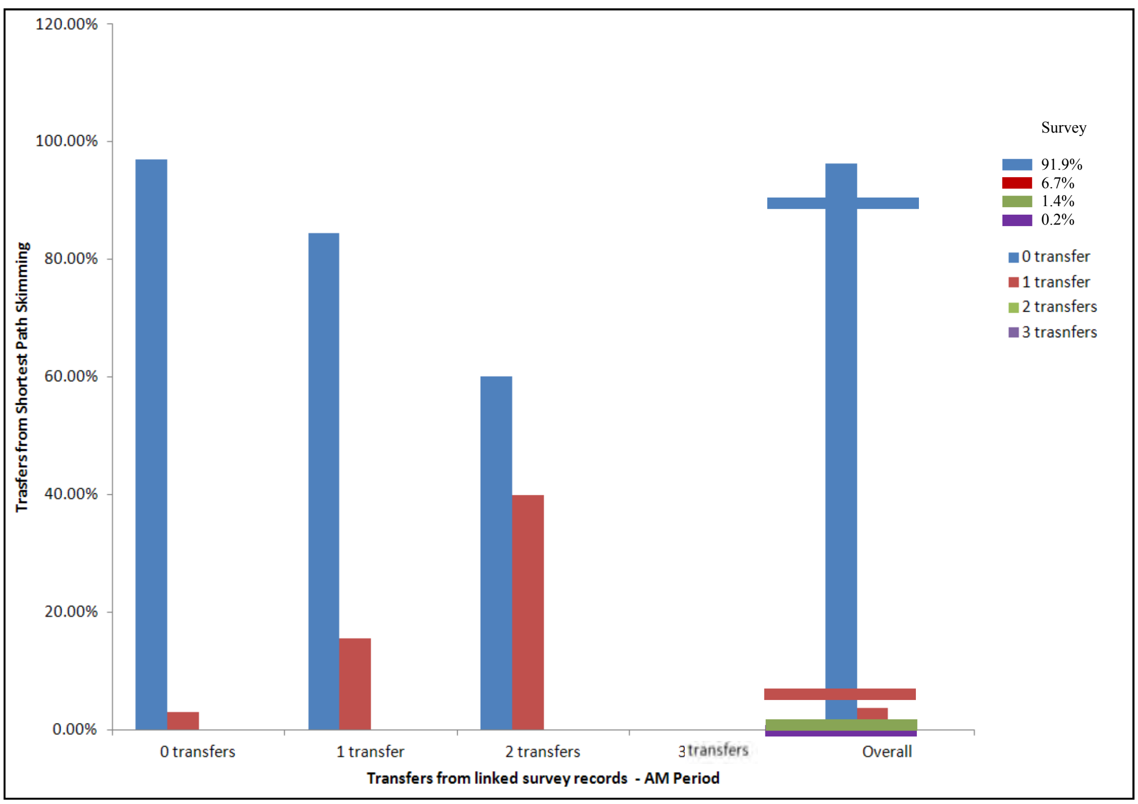


Figure 11: Transit Survey vs. Shortest Path Model Skimming (AM, UMI)



5.3 Walk Access Distances with the Pathfinder Assignment Method

The transit Pathfinder results were further analyzed to examine the access and egress walk distances from the skims. As mentioned earlier, all OD pairs were converted to 'true' origins and destinations. **Table 7** and **Table 8** show the distribution of access and egress walk distances by the transfers reported in the survey. The table also provides the split by selected operator. It can be seen that for the trips where survey reported 0 transfers, Pathfinder estimated 86% to walk less than half a mile. The other 14% would walk more than half a mile but within the one mile limit defined by the access walk parameter. **Table 8** reports the egress walk distances split by the same categories. The two tables were intended to serve as a quick check on the model's transit network so that obvious modeling errors (if any) could be identified. No transit network coding errors have been found based on our review.

Table 7: Access Distance in Pathfinder Assignment (AM)

Transfers in survey	Access distance	AATA	DDOT	SMART	UMICH	Total
0 Transfer	< 0.5 mile	87%	88%	77%	92%	86%
	=> 0.5 mile	13%	12%	23%	8%	14%
1 Transfer	< 0.5 mile	91%	88%	80%	91%	87%
	=> 0.5 mile	9%	12%	20%	9%	13%
2 Transfers	< 0.5 mile	96%	85%	89%	100%	86%
	=> 0.5 mile	4%	15%	11%	0%	14%
3 Transfers	< 0.5 mile	100%	94%	89%	N/A	94%
	=> 0.5 mile	0%	6%	11%	N/A	6%

Table 8: Egress Distance in Pathfinder Distribution (AM)

Transfers in survey	Egress distance	AATA	DDOT	SMART	UMICH	Total
0 Transfer	< 0.5 mile	92%	87%	85%	100%	90%
	=> 0.5 mile	8%	13%	15%	0%	10%
1 Transfer	< 0.5 mile	87%	85%	81%	100%	85%
	=> 0.5 mile	13%	15%	19%	0%	15%
2 Transfers	< 0.5 mile	77%	86%	77%	100%	84%
	=> 0.5 mile	23%	14%	23%	0%	16%
3Transfers	< 0.5 mile	100%	93%	85%	N/A	92%
	=> 0.5 mile	0%	7%	15%	N/A	8%

6 Comparison between Model Results and Survey Data

6.1 Initial Run

Based on the results from an initial run of the transit assignment module in the SEMCOG model, system-wide transfer rates, boardings at the transit operator's level, and boardings at the route level have been checked. The system-wide transfer rate and operator-level boarding comparison is shown in **Table 9** and **Table 10**, respectively. Boardings at the route level are not the priority at this stage and are listed in the appendix. All the model runs in Section 6 used Pathfinder.

Table 9: System-wide Transfer Rate Comparison

	AM	MD	PM	OP	Total	Inferred Transfer Rate
Initial Model Run Total Boardings	46,395	77,597	74,139	27,383	225,513	37.1%
Survey Total Boardings	45,436	78,082	71,355	27,267	222,140	35.1%
Survey Total OD trips	32,526	58,085	53,660	20,204	164,475	

Table 10: Boardings at the Transit Operator's Level in the Initial Model Run

Transit Operator	Period	Initial Model Run Boarding (A)	Survey Boarding (B)	Diff. A-B	% Diff. (A-B)/B
AATA	AM	5,295	5,047	249	4.9%
	MD	10,160	9,690	470	4.8%
	PM	6,701	5,588	1,114	19.9%
	OP	1,538	1,686	-147	-8.7%
	Total	23,695	22,010	1,685	7.7%
BWAT	AM	355	516	-161	-31.2%
	MD	1,125	1,821	-696	-38.2%
	PM	223	288	-66	-22.8%
	OP	-	-	0	N/A
	Total	1,702	2,625	-923	-35.1%
DDOT	AM	25,550	28,509	-2,960	-10.4%
	MD	36,532	37,838	-1,306	-3.5%
	PM	38,366	42,302	-3,936	-9.3%
	OP	17,027	15,865	1,162	7.3%
	Total	117,475	124,514	-7,039	-5.7%
DPM	AM	68	-	68	N/A
	MD	138	-	138	N/A
	PM	870	1,545	-676	-43.7%
	OP	920	2,466	-1,545	-62.7%
	Total	1,996	4,011	-2,015	-50.2%
LETC	AM	47	111	-64	-58.0%
	MD	200	631	-431	-68.3%
	PM	57	101	-44	-43.9%
	OP	-	34	-34	-100.0%
	Total	303	877	-574	-65.4%
SMART	AM	11,298	7,684	3,614	47.0%
	MD	14,737	14,387	350	2.4%
	PM	13,915	7,989	5,926	74.2%
	OP	4,286	3,817	470	12.3%
	Total	44,236	33,876	10,360	30.6%
UMI	AM	3,782	3,569	213	6.0%
	MD	14,705	13,716	989	7.2%
	PM	14,008	13,542	465	3.4%
	OP	3,611	3,400	211	6.2%
	Total	36,105	34,227	1,878	5.5%
System Total	AM	46,395	45,436	959	2.1%
	MD	77,597	78,082	-485	-0.6%
	PM	74,139	71,355	2,784	3.9%
	OP	27,383	27,267	116	0.4%
	Total	225,513	222,140	3,374	1.5%

The initial model run indicates a slightly higher system-wide transfer rate, 37.1%, than the transfer rate suggested by the survey data, 35.1%. This implies that increasing the transfer penalty is a plausible option to improve the model performance.

The boarding results at the transit operator level reveal that the SEMCOG model is producing boardings consistent with the survey data. The AM and PM peak periods have higher discrepancies than the MD and OP periods. The boardings of SMART routes suggested by the model differ from the survey the most.

6.2 Run with Adjustment of Model Assumptions

A key aspect of the transit assignment module in the SEMCOG model is its assumptions on parameters such as transfer penalty time, value of time and maximum number of transfers. Based on the initial model run results, the transfer penalty time was increased from six minutes to twelve minutes to test the sensitivity of the system-wide transfer rate to the transfer penalty time value.

The system-wide transfer rate fell to less than 35.1% when the transfer penalty time was set to twelve minutes. Therefore, the transfer penalty time was adjusted again to eight minutes, which produced a system-wide transfer rate of 34.9%, which is very close to the survey value of 35.1 percent.

Based on this sensitivity test and the results shown in **Table 11** and **Table 12**, it is recommended that consideration be given to increasing the transfer penalty by 33% in future model development.

Table 11: System-wide Transfer Rate with Increased Transfer Penalty

	AM	MD	PM	OP	Total	Inferred Transfer Rate
Model Run with 12-min Transfer Penalty Time Total Boarding	44,519	74,802	70,938	26,764	217,023	31.9%
Model Run with 8-minTransfer Penalty Time Total Boarding	45,691	76,491	72,531	27,146	221,858	34.9%
Survey Total Boarding	45,436	78,082	71,355	27,267	222,140	35.1%
Survey Total OD trips	32,526	58,085	53,660	20,204	164,475	

Table 12: Boardings at the Transit Operator's Level with 8-min Transfer Penalty Time

Transit Operator	Period	Initial Model Run Boarding (A)	Survey Boarding (B)	Diff. A-B	% Diff. (A-B)/B
AATA	AM	5,151	5,047	104	2.1%
	MD	9,997	9,690	307	3.2%
	PM	6,458	5,588	870	15.6%
	OP	1,530	1,686	-156	-9.3%
	Total	23,135	22,010	1,125	5.1%
BWAT	AM	340	516	-176	-34.1%
	MD	1,117	1,821	-703	-38.6%
	PM	218	288	-71	-24.5%
	OP	-	-	0	N/A
	Total	1,675	2,625	-950	-36.2%
DDOT	AM	25,205	28,509	-3,304	-11.6%
	MD	36,131	37,838	-1,707	-4.5%
	PM	37,746	42,302	-4,556	-10.8%
	OP	16,953	15,865	1,088	6.9%
	Total	116,034	124,514	-8,480	-6.8%
DPM	AM	53	-	53	N/A
	MD	137	-	137	N/A
	PM	823	1,545	-723	-46.8%
	OP	903	2,466	-1,562	-63.4%
	Total	1,917	4,011	-2,094	-52.2%
LETC	AM	47	111	-64	-58.0%
	MD	200	631	-431	-68.3%
	PM	57	101	-44	-43.9%
	OP	-	34	-34	-100.0%
	Total	303	877	-574	-65.4%
SMART	AM	11,194	7,684	3,510	45.7%
	MD	14,526	14,387	140	1.0%
	PM	13,599	7,989	5,610	70.2%
	OP	4,166	3,817	350	9.2%
	Total	43,486	33,876	9,610	28.4%
UMI	AM	3,702	3,569	132	3.7%
	MD	14,382	13,716	666	4.9%
	PM	13,631	13,542	89	0.7%
	OP	3,594	3,400	194	5.7%
	Total	35,308	34,227	1,081	3.2%
System Total	AM	45,691	45,436	255	0.6%
	MD	76,491	78,082	-1,591	-2.0%
	PM	72,531	71,355	1,176	1.6%
	OP	27,146	27,267	-121	-0.4%
	Total	221,858	222,140	-282	-0.1%

6.3 Run with No-Transfer Survey Trips

Assigning only trips with no transfers in the survey was undertaken to understand and identify discrepancies between the SEMCOG model results and survey data. Singling out this type of trip in the assignment excludes the impact of trips involving more complex routes and may reveal factors that could have been obscured in the overall assignment results.

Survey trips with no-transfers were identified and assigned in the SEMCOG model using the Pathfinder algorithm with six-minute and twelve-minute transfer penalties. **Table 13** indicates that the model results have much higher transfer rates than the survey for these comparatively straight-forward transit trips even when the transfer penalty is increased to twelve minutes. Boardings by operator for the no transfer survey trips are shown in **Table 14**.

Table 13: System-wide Transfer Rate for Assignment of No-Transfer Trips

	AM	MD	PM	OP	Total	Inferred Transfer Rate
No-Transfer Trip Model Run Total Boarding (6-min penalty)	27,650	50,064	49,594	18,169	145,477	25.3%
No-Transfer Trip Model Run Total Boarding (12-min penalty)	26,555	48,254	47,550	17,730	140,090	20.7%
No-Transfer Trip Survey Total Boarding	21,632	41,204	39,027	14,247	116,110	0.0%
No-Transfer Trip Survey Total OD trips	21,632	41,204	39,027	14,247	116,110	

The results led to discussions with Caliper, the developer of the TransCAD, regarding the unrealistically high transfer rate. The allocation of transit ridership is based on the bus frequency at each joint point (potential bus transfer points) of the set of paths. Even when having only survey trips with no transfers, the algorithm still allocates transit riders to the same set of paths at each joint point.

The Pathfinder algorithm is likely to assign zero-transfer survey trips to paths that correspond to multiple transfer survey records and multiple-transfer survey trips to paths with no transfer. Therefore, it is necessary to compare the number of transfers shown in the model with the survey records, which have been discussed in Section 5. A table documenting the number of transfers in the model and survey records has been generated and discussed in Section 5. The next step is to examine those records with different numbers of transfers in the model and survey records, and make necessary adjustments according to the findings. The table is not included in this report due to its size.

Table 14: Boardings at the Transit Operator's Level from Assignment of No-Transfer Trips (12-min penalty)

Transit Operator	Period	Initial Model Run Boarding (A)	Survey Boarding (B)	Diff. A-B	% Diff. (A-B)/B
AATA	AM	3,704	3,178	525	16.5%
	MD	7,423	6,524	899	13.8%
	PM	5,077	3,879	1,198	30.9%
	OP	915	768	147	19.1%
	Total	17,118	14,349	2,769	19.3%
BWAT	AM	148	114	34	29.4%
	MD	564	699	-135	-19.3%
	PM	138	113	25	22.5%
	OP	-	-	0	N/A
	Total	851	927	-76	-8.2%
DDOT	AM	13,163	11,199	1,964	17.5%
	MD	18,851	14,644	4,208	28.7%
	PM	20,736	16,878	3,858	22.9%
	OP	10,181	6,925	3,256	47.0%
	Total	62,931	49,647	13,285	26.8%
DPM	AM	7	-	7	N/A
	MD	69	-	69	N/A
	PM	757	1,331	-574	-43.1%
	OP	840	2,038	-1,198	-58.8%
	Total	1,672	3,370	-1,697	-50.4%
LETC	AM	-	4	-4	-100.0%
	MD	76	246	-170	-69.3%
	PM	57	90	-33	-36.6%
	OP	-	4	-4	-100.0%
	Total	132	343	-210	-61.4%
SMART	AM	6,223	3,856	2,367	61.4%
	MD	8,278	6,802	1,476	21.7%
	PM	7,808	4,052	3,757	92.7%
	OP	2,857	1,970	887	45.0%
	Total	25,166	16,680	8,486	50.9%
UMI	AM	3,310	3,281	29	0.9%
	MD	12,993	12,289	705	5.7%
	PM	12,977	12,685	292	2.3%
	OP	2,938	2,542	396	15.6%
	Total	32,219	30,796	1,423	4.6%
System Total	AM	26,555	21,632	4,923	22.8%
	MD	48,254	41,204	7,050	17.1%
	PM	47,550	39,027	8,523	21.8%
	OP	17,730	14,247	3,484	24.5%
	Total	140,090	116,110	23,980	20.7%

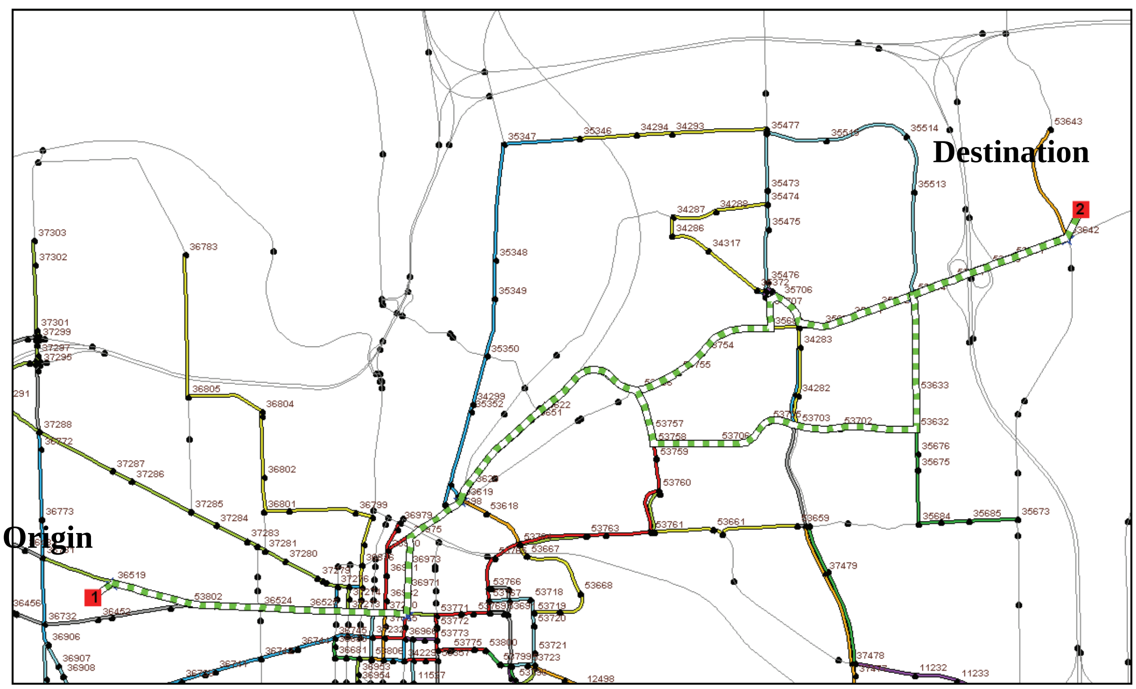
7 Examination of Selected Survey Records

With all the aggregated results shown in previous sections, we feel it is still important to review selected origin-destination trip-level results. Three OD pairs from the survey have been selected to compare the route suggested by the survey and Pathfinder in the SEMCOG model. Please note that this section is not intended to provide an exhaustive investigation, but rather to look into the results from another perspective.

(1) **Survey Sample #107152 from TAZ #2212 to TAZ#2335**
The survey is showing that no transfer (only AATA 2) is needed to get from TAZ #2212 to #2335 while the model suggests 2 transfers. More specifically, the model provides the following two routes:

a. AATA 609 EB → AATA 001 EB → UMI Intercampus NB
b. AATA 609 EB → AATA 02C EB → UMI Intercampus NB

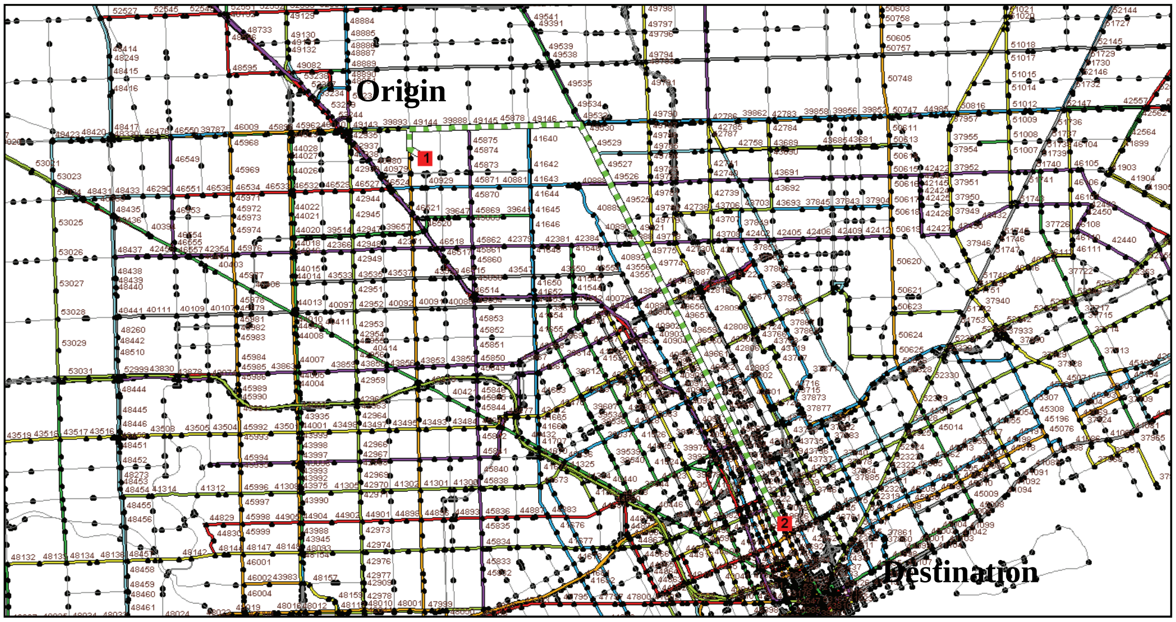
Figure 12: Two Paths from TAZ #2212 to #2335 by Pathfinder in the Model (AM)



There is no AATA 2 bus stop within one mile of the origin centroid in the model and the access mode at the origin is walk in the survey. This survey data record may need further scrutiny.

(2) **Survey Sample #20927 from TAZ #299 to TAZ#166**
The survey shows that one transfer is needed to get from TAZ #299 to #166 while the model suggests zero transfers. The model suggests two zero-transfer routes: SMART 415 and SMART 420, which overlap between TAZ#299 and #166. According to the survey, this transit rider boarded DDOT bus 53.

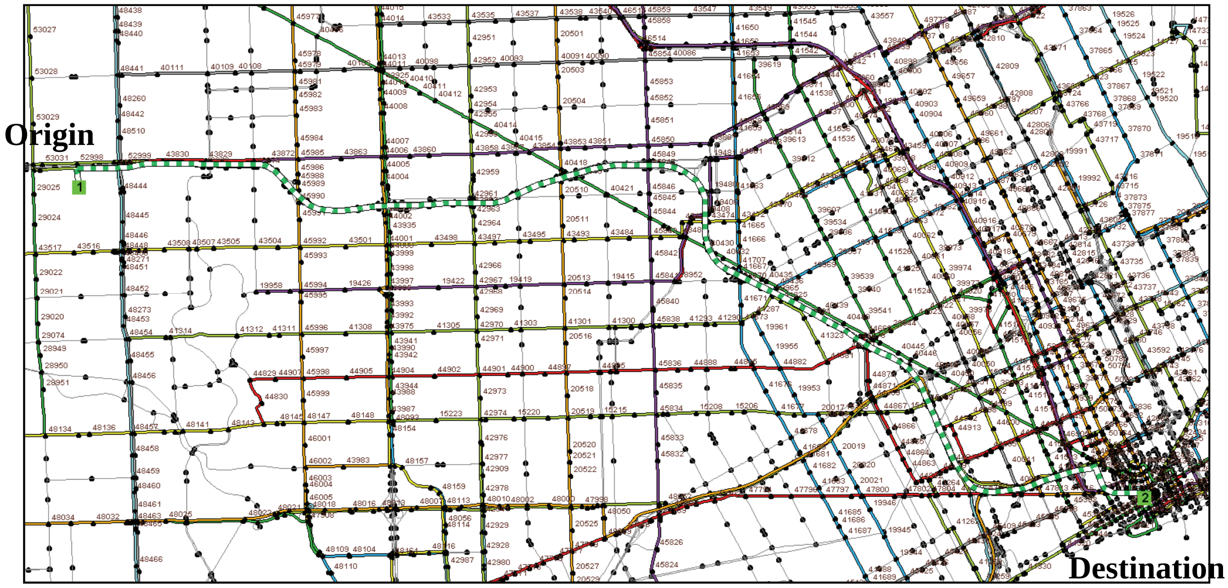
Figure 13: The Path from TAZ #299 to #166 by Pathfinder in the Model (AM)



The survey indicates that the first route taken is DDOT-17 and that the second is DDOT-53. The headways of DDOT-17 are less than what the model indicates for SMART 415 and 420. It is likely that the survey respondents preferred to get on this bus quicker despite the fact that it would incur one more transfer. This is an example showing that the Pathfinder algorithm is likely not capable of capturing certain rider preferences.

(3) **Survey Sample #56245 from TAZ #789 to TAZ#203**
Both the survey and model show that a transfer is needed to get from TAZ #789 to #203. The route, SMART 805, is the same in both the survey data and model results. This is an example showing that the survey record matches the model results.

Figure 13: The Path from TAZ #789 to #203 by Pathfinder in the Model (AM)



As shown in the above comparison with selected OD pairs, the model could suggest different routes with a different number of transfers in contrast with the survey data. This confirms the findings discussed in Section 5 on the initial comparison between the survey and model skims. Even though the system-wide transfer rate from the model is consistent with what the survey data suggest, many OD pairs will likely be taking very different routes in the model from those recorded in the survey.

8 Summary and Recommendation

In summary, this on-board transit survey based transit assignment in SEMCOG model has conducted the following tasks:

- Converted the survey data into OD matrices
- Compared the number of transfers and walk distance in the survey with those from the model skims using both Pathfinder and Shortest Path algorithms
- Tested and analyzed several model runs with different assumptions and inputs
- Examined the actual routes for selected survey OD pairs and compared them to model results

The consultant team found that the access time, transfer penalty and the Combination Factor used in the transit path-finder algorithm are important factors that influence the outcome of the transit assignment. Based on the findings, recommendations include:

- Examine those records with different number of transfers in the model and survey, and make adjustments to the survey records and/or model settings accordingly
- Explore more on Combination Factor setting in the Path-finder algorithm, so the assignment could better replicate the travel pattern shown in the survey.

APPENDIX

Table A-1 Route Level Comparison with 8-min Transfer Penalty (AM)

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
AATA Route 1	318	264	54	20.4%
AATA Route 10	101	90	11	12.7%
AATA Route 11	40	27	13	45.7%
AATA Route 12A	55	76	-21	-27.7%
AATA Route 12B	34	46	-12	-25.9%
AATA Route 13	45	65	-20	-30.5%
AATA Route 14	30	152	-122	-80.5%
AATA Route 15	45	103	-58	-56.3%
AATA Route 16	319	63	256	403.9%
AATA Route 17	4	46	-42	-92.2%
AATA Route 18	130	0	130	0.0%
AATA Route 1U	82	128	-46	-35.9%
AATA Route 2	369	520	-151	-29.1%
AATA Route 20	60	97	-37	-38.1%
AATA Route 22	157	157	0	-0.2%
AATA Route 3	239	391	-152	-38.9%
AATA Route 33	48	30	18	60.9%
AATA Route 36	438	445	-8	-1.7%
AATA Route 4	1,084	710	374	52.8%
AATA Route 5	440	591	-151	-25.5%
AATA Route 6	326	397	-70	-17.7%
AATA Route 609	120	80	40	50.4%
AATA Route 7	372	204	169	82.8%
AATA Route 8	235	234	2	0.7%
AATA Route 9	62	132	-71	-53.4%
BWAT	45	0	45	0.0%
BWAT Route 1	39	0	39	0.0%
BWAT Route 2	27	115	-89	-76.9%
BWAT Route 3	44	0	44	0.0%
BWAT Route 4	27	81	-54	-66.2%
BWAT Route 5	30	110	-80	-72.7%
BWAT Route 6	53	77	-25	-31.9%
BWAT Route 9	74	123	-48	-39.3%
DDOT Route 10	280	348	-67	-19.4%
DDOT Route 11	414	228	187	82.0%
DDOT Route 12	116	261	-145	-55.6%
DDOT Route 13	184	305	-121	-39.6%
DDOT Route 14	1,727	1,269	458	36.1%
DDOT Route 15	737	810	-74	-9.1%
DDOT Route 16	2,870	2,316	554	23.9%
DDOT Route 17	1,338	1,260	78	6.2%
DDOT Route 18	916	1,076	-160	-14.8%
DDOT Route 19	98	236	-138	-58.4%
DDOT Route 21	1,295	1,797	-501	-27.9%
DDOT Route 22	1,255	924	330	35.7%
DDOT Route 23	614	479	135	28.2%
DDOT Route 25	499	513	-14	-2.8%
DDOT Route 27	397	777	-380	-48.9%
DDOT Route 29	418	473	-55	-11.7%

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
DDOT Route 30	409	557	-148	-26.5%
DDOT Route 31	580	1,216	-635	-52.2%
DDOT Route 32	1,464	1,017	447	44.0%
DDOT Route 34	892	1,267	-375	-29.6%
DDOT Route 36	36	97	-60	-62.6%
DDOT Route 37	50	230	-180	-78.4%
DDOT Route 38	441	434	8	1.8%
DDOT Route 38/8	0	0	0	0.0%
DDOT Route 39	110	212	-101	-47.9%
DDOT Route 40	76	135	-58	-43.4%
DDOT Route 41	270	387	-117	-30.3%
DDOT Route 43	398	228	171	75.0%
DDOT Route 45	1,772	1,651	121	7.3%
DDOT Route 46	91	343	-252	-73.5%
DDOT Route 47	244	336	-92	-27.4%
DDOT Route 48	1,063	1,679	-616	-36.7%
DDOT Route 49	533	337	196	58.0%
DDOT Route 53	1,046	2,718	-1,672	-61.5%
DDOT Route 54	343	408	-65	-15.8%
DDOT Route 60	705	777	-72	-9.3%
DDOT Route 7	672	667	5	0.7%
DDOT Route 76	35	54	-19	-35.3%
DDOT Route 78	513	172	341	198.8%
DDOT Route 8	49	177	-128	-72.3%
DDOT Route 9	255	342	-88	-25.6%
DPM	53	0	53	0.0%
LETC Route 2	23	93	-70	-75.0%
LETC Route 3	23	0	23	0.0%
LETC Route 4	0	0	0	0.0%
LETC Route 5	0	0	0	0.0%
LETC Route 6	0	11	-11	-100.0%
LETC Route 7	0	0	0	0.0%
LETC Route 8	0	0	0	0.0%
LETC Route 9	0	7	-7	-100.0%
SMART Route 125	475	429	46	10.6%
SMART Route 135	109	0	109	0.0%
SMART Route 140	108	82	26	31.2%
SMART Route 145	35	0	35	0.0%
SMART Route 150	56	35	21	59.0%
SMART Route 160	63	46	17	37.6%
SMART Route 190	3	0	3	0.0%
SMART Route 200	990	642	348	54.1%
SMART Route 202	28	23	4	19.0%
SMART Route 245	49	39	10	27.1%
SMART Route 250	34	38	-4	-10.8%
SMART Route 255	254	105	149	141.9%
SMART Route 265	72	50	22	43.7%
SMART Route 275	362	276	86	31.2%
SMART Route 280	68	109	-41	-37.6%
SMART Route 305	41	0	41	0.0%
SMART Route 330	60	99	-38	-39.0%

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
SMART Route 385	16	10	6	57.6%
SMART Route 400	78	91	-12	-13.7%
SMART Route 405	132	0	132	0.0%
SMART Route 415	219	0	219	0.0%
SMART Route 420	143	17	126	737.2%
SMART Route 430	57	85	-28	-33.2%
SMART Route 445	86	0	86	0.0%
SMART Route 450	566	937	-371	-39.6%
SMART Route 460	418	0	418	0.0%
SMART Route 465	52	0	52	0.0%
SMART Route 475	133	0	133	0.0%
SMART Route 494	51	107	-56	-52.6%
SMART Route 495	752	375	378	100.9%
SMART Route 510	673	435	238	54.7%
SMART Route 515	4	0	4	0.0%
SMART Route 525	23	0	23	0.0%
SMART Route 530	139	143	-4	-2.7%
SMART Route 550	62	58	4	7.1%
SMART Route 559	5	24	-19	-80.2%
SMART Route 560	1,887	1,168	719	61.5%
SMART Route 565	56	0	56	0.0%
SMART Route 580	87	0	87	0.0%
SMART Route 610	54	298	-244	-81.9%
SMART Route 615	68	0	68	0.0%
SMART Route 620	157	0	157	0.0%
SMART Route 635	23	0	23	0.0%
SMART Route 710	359	436	-77	-17.8%
SMART Route 730	129	158	-30	-18.8%
SMART Route 740	319	266	53	19.9%
SMART Route 752	14	43	-29	-67.6%
SMART Route 753	31	24	7	29.6%
SMART Route 756	17	32	-15	-47.7%
SMART Route 760	70	59	10	17.7%
SMART Route 780	290	216	74	34.2%
SMART Route 805	250	142	108	76.3%
SMART Route 830	383	268	115	43.0%
SMART Route 851	589	129	460	355.7%
UMI	3,702	3,569	132	3.7%
SUM	45,691	45,236	454	1.0%

Table A-2 Route Level Comparison with 8-min Transfer Penalty (MD)

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
AATA Route 1	452	546	-94	-17.3%
AATA Route 10	161	229	-68	-29.6%
AATA Route 11	125	152	-27	-18.0%
AATA Route 12A	168	240	-72	-30.0%
AATA Route 12B	233	160	72	45.1%
AATA Route 13	81	56	25	44.1%
AATA Route 14	0	0	0	0.0%
AATA Route 15	90	96	-6	-6.0%
AATA Route 16	550	190	360	189.7%
AATA Route 17	33	0	33	0.0%
AATA Route 18	251	186	65	35.0%
AATA Route 1U	0	0	0	0.0%
AATA Route 2	674	1,095	-420	-38.4%
AATA Route 20	191	225	-34	-15.2%
AATA Route 22	275	370	-95	-25.6%
AATA Route 3	593	685	-92	-13.5%
AATA Route 33	411	534	-124	-23.1%
AATA Route 36	1,115	837	278	33.2%
AATA Route 4	1,072	1,036	36	3.5%
AATA Route 5	1,202	871	332	38.1%
AATA Route 6	974	1,108	-134	-12.1%
AATA Route 609	0	0	0	0.0%
AATA Route 7	809	493	316	64.2%
AATA Route 8	282	280	2	0.7%
AATA Route 9	257	303	-46	-15.2%
BWAT	4	168	-164	-97.4%
BWAT Route 1	143	261	-118	-45.1%
BWAT Route 2	73	156	-83	-53.2%
BWAT Route 3	293	402	-109	-27.1%
BWAT Route 4	95	70	25	36.1%
BWAT Route 5	94	259	-165	-63.6%
BWAT Route 6	167	245	-78	-31.8%
BWAT Route 9	248	260	-13	-4.8%
DDOT Route 10	400	373	27	7.2%
DDOT Route 11	570	305	264	86.6%
DDOT Route 12	77	175	-97	-55.7%
DDOT Route 13	284	347	-63	-18.1%
DDOT Route 14	2,438	1,823	615	33.7%
DDOT Route 15	708	643	65	10.1%
DDOT Route 16	4,498	3,658	840	23.0%
DDOT Route 17	1,744	1,708	36	2.1%
DDOT Route 18	1,454	1,189	265	22.3%
DDOT Route 19	293	678	-384	-56.7%
DDOT Route 21	2,216	2,652	-435	-16.4%
DDOT Route 22	1,655	1,334	321	24.1%
DDOT Route 23	855	834	22	2.6%
DDOT Route 25	911	1,016	-105	-10.4%
DDOT Route 27	1,027	1,005	22	2.2%
DDOT Route 29	465	544	-79	-14.5%

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
DDOT Route 30	442	764	-323	-42.2%
DDOT Route 31	797	1,214	-417	-34.3%
DDOT Route 32	1,894	1,618	277	17.1%
DDOT Route 34	1,698	2,084	-386	-18.5%
DDOT Route 36	23	113	-90	-79.6%
DDOT Route 37	74	520	-446	-85.8%
DDOT Route 38	330	537	-207	-38.6%
DDOT Route 38/8	213	0	213	0.0%
DDOT Route 39	215	325	-111	-34.0%
DDOT Route 40	184	235	-51	-21.7%
DDOT Route 41	702	604	98	16.3%
DDOT Route 43	350	381	-31	-8.1%
DDOT Route 45	2,035	2,085	-50	-2.4%
DDOT Route 46	141	395	-255	-64.4%
DDOT Route 47	78	291	-214	-73.4%
DDOT Route 48	1,476	1,010	465	46.1%
DDOT Route 49	319	270	49	18.3%
DDOT Route 53	2,883	4,577	-1,694	-37.0%
DDOT Route 54	225	373	-148	-39.7%
DDOT Route 60	773	833	-60	-7.2%
DDOT Route 7	636	711	-75	-10.6%
DDOT Route 76	0	0	0	0.0%
DDOT Route 78	437	0	437	0.0%
DDOT Route 8	62	189	-127	-67.2%
DDOT Route 9	551	428	123	28.8%
DPM	137	0	137	0.0%
LETC Route 2	22	0	22	0.0%
LETC Route 3	27	47	-20	-42.0%
LETC Route 4	21	65	-44	-68.2%
LETC Route 5	22	131	-109	-83.0%
LETC Route 6	34	85	-52	-60.5%
LETC Route 7	34	87	-53	-60.6%
LETC Route 8	8	146	-138	-94.8%
LETC Route 9	33	70	-37	-53.0%
SMART Route 125	907	865	42	4.9%
SMART Route 135	0	0	0	0.0%
SMART Route 140	172	144	27	19.0%
SMART Route 145	31	36	-4	-12.4%
SMART Route 150	0	0	0	0.0%
SMART Route 160	40	143	-104	-72.1%
SMART Route 190	0	11	-11	-100.0%
SMART Route 200	1,455	1,111	344	30.9%
SMART Route 202	0	0	0	0.0%
SMART Route 245	86	120	-34	-28.3%
SMART Route 250	169	163	6	3.8%
SMART Route 255	0	0	0	0.0%
SMART Route 265	119	96	23	23.8%
SMART Route 275	684	489	195	39.9%
SMART Route 280	100	168	-68	-40.7%
SMART Route 305	0	0	0	0.0%
SMART Route 330	127	155	-27	-17.8%

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
SMART Route 385	47	51	-3	-6.8%
SMART Route 400	100	93	7	7.1%
SMART Route 405	200	280	-80	-28.5%
SMART Route 415	367	0	367	0.0%
SMART Route 420	240	686	-446	-65.0%
SMART Route 430	31	12	19	153.4%
SMART Route 445	0	0	0	0.0%
SMART Route 450	911	1,926	-1,015	-52.7%
SMART Route 460	872	0	872	0.0%
SMART Route 465	0	43	-43	-100.0%
SMART Route 475	0	0	0	0.0%
SMART Route 494	205	223	-18	-8.0%
SMART Route 495	725	747	-22	-3.0%
SMART Route 510	1,048	1,159	-111	-9.6%
SMART Route 515	0	0	0	0.0%
SMART Route 525	0	0	0	0.0%
SMART Route 530	0	0	0	0.0%
SMART Route 550	116	101	15	14.9%
SMART Route 559	0	0	0	0.0%
SMART Route 560	2,805	2,479	325	13.1%
SMART Route 565	0	0	0	0.0%
SMART Route 580	0	0	0	0.0%
SMART Route 610	218	442	-224	-50.7%
SMART Route 615	171	0	171	0.0%
SMART Route 620	0	0	0	0.0%
SMART Route 635	0	0	0	0.0%
SMART Route 710	578	707	-130	-18.3%
SMART Route 730	295	306	-11	-3.6%
SMART Route 740	747	784	-37	-4.7%
SMART Route 752	55	119	-65	-54.1%
SMART Route 753	71	104	-33	-31.5%
SMART Route 756	143	126	17	13.9%
SMART Route 760	330	229	101	44.1%
SMART Route 780	360	267	93	34.8%
SMART Route 805	0	0	0	0.0%
SMART Route 830	0	0	0	0.0%
SMART Route 851	0	0	0	0.0%
UMI	14,382	13,716	666	4.9%
SUM	76,491	78,082	-1,591	-2.0%

Table A-3 Route Level Comparison with 8-min Transfer Penalty (PM)

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
AATA Route 1	101	0	101	0.0%
AATA Route 10	94	108	-14	-12.7%
AATA Route 11	46	63	-18	-27.8%
AATA Route 12A	73	143	-71	-49.3%
AATA Route 12B	63	168	-105	-62.3%
AATA Route 13	78	90	-12	-13.3%
AATA Route 14	2	44	-42	-94.4%
AATA Route 15	29	55	-26	-47.4%
AATA Route 16	365	116	249	214.1%
AATA Route 17	5	0	5	0.0%
AATA Route 18	238	249	-11	-4.3%
AATA Route 1U	34	0	34	0.0%
AATA Route 2	531	629	-98	-15.6%
AATA Route 20	108	158	-50	-31.6%
AATA Route 22	98	245	-148	-60.2%
AATA Route 3	224	258	-35	-13.4%
AATA Route 33	18	47	-29	-61.5%
AATA Route 36	1,232	414	818	197.3%
AATA Route 4	1,150	631	519	82.2%
AATA Route 5	667	685	-19	-2.7%
AATA Route 6	420	591	-172	-29.0%
AATA Route 609	149	62	87	139.4%
AATA Route 7	408	342	66	19.3%
AATA Route 8	229	233	-3	-1.5%
AATA Route 9	98	173	-75	-43.2%
BWAT	52	0	52	0.0%
BWAT Route 1	22	108	-86	-79.8%
BWAT Route 2	13	29	-16	-54.4%
BWAT Route 3	17	0	17	0.0%
BWAT Route 4	17	0	17	0.0%
BWAT Route 5	26	0	26	0.0%
BWAT Route 6	43	119	-76	-64.0%
BWAT Route 9	28	0	28	0.0%
DDOT Route 10	569	644	-75	-11.7%
DDOT Route 11	617	577	40	7.0%
DDOT Route 12	239	364	-125	-34.4%
DDOT Route 13	271	478	-206	-43.2%
DDOT Route 14	2,211	2,178	34	1.5%
DDOT Route 15	835	1,044	-208	-20.0%
DDOT Route 16	3,824	3,157	667	21.1%
DDOT Route 17	1,510	1,587	-77	-4.8%
DDOT Route 18	1,395	1,694	-298	-17.6%
DDOT Route 19	216	565	-349	-61.8%
DDOT Route 21	2,212	2,513	-302	-12.0%
DDOT Route 22	2,030	1,957	73	3.7%
DDOT Route 23	1,890	1,175	715	60.9%
DDOT Route 25	493	880	-388	-44.1%
DDOT Route 27	477	1,150	-672	-58.5%
DDOT Route 29	615	709	-94	-13.3%

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
DDOT Route 30	653	640	12	1.9%
DDOT Route 31	1,417	1,058	359	33.9%
DDOT Route 32	2,098	1,898	200	10.5%
DDOT Route 34	2,643	2,462	181	7.3%
DDOT Route 36	73	175	-101	-58.0%
DDOT Route 37	35	488	-453	-92.8%
DDOT Route 38	591	416	175	42.2%
DDOT Route 38/8	0	0	0	0.0%
DDOT Route 39	105	235	-130	-55.4%
DDOT Route 40	137	228	-91	-39.8%
DDOT Route 41	509	846	-337	-39.9%
DDOT Route 43	254	358	-104	-29.1%
DDOT Route 45	2,138	3,469	-1,331	-38.4%
DDOT Route 46	335	433	-98	-22.5%
DDOT Route 47	106	395	-289	-73.2%
DDOT Route 48	1,129	1,145	-17	-1.5%
DDOT Route 49	513	322	191	59.2%
DDOT Route 53	1,674	3,667	-1,993	-54.3%
DDOT Route 54	546	464	82	17.7%
DDOT Route 60	924	1,016	-92	-9.1%
DDOT Route 7	1,093	950	143	15.1%
DDOT Route 76	14	122	-108	-88.6%
DDOT Route 78	796	199	598	301.0%
DDOT Route 8	67	124	-56	-45.6%
DDOT Route 9	491	521	-31	-5.9%
DPM	823	1,545	-723	-46.8%
LETC Route 2	8	0	8	0.0%
LETC Route 3	14	37	-22	-61.0%
LETC Route 4	8	41	-33	-80.0%
LETC Route 5	0	0	0	0.0%
LETC Route 6	0	0	0	0.0%
LETC Route 7	14	0	14	0.0%
LETC Route 8	0	0	0	0.0%
LETC Route 9	12	23	-12	-50.0%
SMART Route 125	650	366	284	77.7%
SMART Route 135	51	35	16	46.1%
SMART Route 140	57	88	-31	-34.9%
SMART Route 145	30	27	3	10.4%
SMART Route 150	120	21	99	469.5%
SMART Route 160	40	0	40	0.0%
SMART Route 190	20	0	20	0.0%
SMART Route 200	1,511	668	843	126.3%
SMART Route 202	2	0	2	0.0%
SMART Route 245	86	59	27	46.0%
SMART Route 250	59	106	-47	-44.0%
SMART Route 255	420	112	308	275.8%
SMART Route 265	35	109	-75	-68.2%
SMART Route 275	413	139	274	197.0%
SMART Route 280	71	0	71	0.0%
SMART Route 305	0	0	0	0.0%
SMART Route 330	72	107	-35	-32.6%

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
SMART Route 385	20	25	-5	-19.7%
SMART Route 400	109	91	18	19.2%
SMART Route 405	197	207	-10	-4.8%
SMART Route 415	485	0	485	0.0%
SMART Route 420	454	456	-2	-0.4%
SMART Route 430	38	0	38	0.0%
SMART Route 445	156	0	156	0.0%
SMART Route 450	459	756	-297	-39.3%
SMART Route 460	457	0	457	0.0%
SMART Route 465	144	155	-11	-7.2%
SMART Route 475	133	0	133	0.0%
SMART Route 494	75	160	-85	-53.2%
SMART Route 495	948	475	473	99.7%
SMART Route 510	693	642	52	8.0%
SMART Route 515	3	0	3	0.0%
SMART Route 525	0	0	0	0.0%
SMART Route 530	128	0	128	0.0%
SMART Route 550	50	106	-56	-52.8%
SMART Route 559	6	0	6	0.0%
SMART Route 560	1,738	1,070	668	62.4%
SMART Route 565	115	0	115	0.0%
SMART Route 580	148	0	148	0.0%
SMART Route 610	104	187	-83	-44.4%
SMART Route 615	134	0	134	0.0%
SMART Route 620	160	89	71	80.3%
SMART Route 635	190	102	88	86.3%
SMART Route 710	545	468	76	16.3%
SMART Route 730	151	172	-21	-12.4%
SMART Route 740	264	213	51	23.7%
SMART Route 752	25	35	-10	-27.5%
SMART Route 753	35	72	-36	-50.7%
SMART Route 756	36	46	-10	-21.8%
SMART Route 760	134	90	44	48.2%
SMART Route 780	273	168	104	61.8%
SMART Route 805	304	145	160	110.4%
SMART Route 830	443	0	443	0.0%
SMART Route 851	610	152	458	301.1%
UMI	13,631	13,542	89	0.7%
SUM	72,531	71,171	1,360	1.9%

Table A-4 Route Level Comparison with 8-min Transfer Penalty (OP)

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
AATA Route 1	68	0	68	0.0%
AATA Route 10	48	79	-31	-39.5%
AATA Route 11	0	0	0	0.0%
AATA Route 12A	19	0	19	0.0%
AATA Route 12B	38	0	38	0.0%
AATA Route 13	0	0	0	0.0%
AATA Route 14	0	0	0	0.0%
AATA Route 15	3	0	3	0.0%
AATA Route 16	54	99	-44	-45.0%
AATA Route 17	0	0	0	0.0%
AATA Route 18	0	0	0	0.0%
AATA Route 1U	0	0	0	0.0%
AATA Route 2	44	138	-94	-67.9%
AATA Route 20	0	64	-64	-100.0%
AATA Route 22	39	0	39	0.0%
AATA Route 3	0	3	-3	-100.0%
AATA Route 33	0	0	0	0.0%
AATA Route 36	241	0	241	0.0%
AATA Route 4	440	676	-236	-34.9%
AATA Route 5	143	172	-29	-17.0%
AATA Route 6	182	153	30	19.4%
AATA Route 609	0	0	0	0.0%
AATA Route 7	138	203	-64	-31.8%
AATA Route 8	24	0	24	0.0%
AATA Route 9	48	100	-52	-51.7%
BWAT	0	0	0	0.0%
BWAT Route 1	0	0	0	0.0%
BWAT Route 2	0	0	0	0.0%
BWAT Route 3	0	0	0	0.0%
BWAT Route 4	0	0	0	0.0%
BWAT Route 5	0	0	0	0.0%
BWAT Route 6	0	0	0	0.0%
BWAT Route 9	0	0	0	0.0%
DDOT Route 10	391	413	-22	-5.3%
DDOT Route 11	176	98	78	79.1%
DDOT Route 12	89	132	-44	-33.0%
DDOT Route 13	346	263	83	31.5%
DDOT Route 14	943	1,061	-118	-11.1%
DDOT Route 15	159	85	74	86.7%
DDOT Route 16	785	909	-125	-13.7%
DDOT Route 17	1,291	751	540	71.9%
DDOT Route 18	532	184	348	189.2%
DDOT Route 19	76	20	56	286.1%
DDOT Route 21	847	1,977	-1,131	-57.2%
DDOT Route 22	1,003	1,011	-8	-0.8%
DDOT Route 23	1,193	458	736	160.8%
DDOT Route 25	397	702	-306	-43.5%
DDOT Route 27	322	0	322	0.0%
DDOT Route 29	565	363	203	55.9%

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
DDOT Route 30	49	0	49	0.0%
DDOT Route 31	690	483	207	42.9%
DDOT Route 32	696	959	-263	-27.4%
DDOT Route 34	821	648	173	26.6%
DDOT Route 36	0	72	-72	-100.0%
DDOT Route 37	362	12	350	2840.3%
DDOT Route 38	207	500	-293	-58.6%
DDOT Route 38/8	116	0	116	0.0%
DDOT Route 39	181	141	40	28.4%
DDOT Route 40	55	48	6	13.1%
DDOT Route 41	112	96	15	16.1%
DDOT Route 43	405	126	278	220.7%
DDOT Route 45	615	739	-124	-16.8%
DDOT Route 46	154	208	-53	-25.7%
DDOT Route 47	47	392	-345	-88.1%
DDOT Route 48	650	373	277	74.2%
DDOT Route 49	121	243	-121	-50.0%
DDOT Route 53	1,741	1,504	237	15.7%
DDOT Route 54	100	27	73	268.1%
DDOT Route 60	223	240	-17	-7.2%
DDOT Route 7	332	411	-79	-19.2%
DDOT Route 76	0	0	0	0.0%
DDOT Route 78	0	28	-28	-100.0%
DDOT Route 8	0	0	0	0.0%
DDOT Route 9	163	187	-24	-12.6%
DPM	903	2,466	-1,562	-63.4%
LETC Route 2	0	0	0	0.0%
LETC Route 3	0	0	0	0.0%
LETC Route 4	0	0	0	0.0%
LETC Route 5	0	0	0	0.0%
LETC Route 6	0	7	-7	-100.0%
LETC Route 7	0	0	0	0.0%
LETC Route 8	0	27	-27	-100.0%
LETC Route 9	0	0	0	0.0%
SMART Route 125	185	211	-26	-12.3%
SMART Route 135	0	0	0	0.0%
SMART Route 140	3	0	3	0.0%
SMART Route 145	0	0	0	0.0%
SMART Route 150	0	0	0	0.0%
SMART Route 160	0	0	0	0.0%
SMART Route 190	0	0	0	0.0%
SMART Route 200	479	110	369	335.0%
SMART Route 202	0	14	-14	-100.0%
SMART Route 245	0	0	0	0.0%
SMART Route 250	26	0	26	0.0%
SMART Route 255	0	8	-8	-100.0%
SMART Route 265	0	6	-6	-100.0%
SMART Route 275	255	330	-75	-22.7%
SMART Route 280	15	0	15	0.0%
SMART Route 305	0	0	0	0.0%
SMART Route 330	54	25	28	113.4%

Route name	Model Boarding (A)	SurveyBoarding(B)	Difference (A-B)	% (A-B)/B
SMART Route 385	24	11	12	110.5%
SMART Route 400	0	13	-13	-100.0%
SMART Route 405	0	26	-26	-100.0%
SMART Route 415	186	0	186	0.0%
SMART Route 420	127	145	-18	-12.5%
SMART Route 430	0	0	0	0.0%
SMART Route 445	0	0	0	0.0%
SMART Route 450	324	640	-317	-49.5%
SMART Route 460	309	0	309	0.0%
SMART Route 465	0	43	-43	-100.0%
SMART Route 475	0	0	0	0.0%
SMART Route 494	54	72	-18	-25.2%
SMART Route 495	126	194	-68	-34.9%
SMART Route 510	281	363	-82	-22.7%
SMART Route 515	0	0	0	0.0%
SMART Route 525	0	15	-15	-100.0%
SMART Route 530	0	0	0	0.0%
SMART Route 550	22	42	-20	-47.8%
SMART Route 559	0	0	0	0.0%
SMART Route 560	944	794	150	18.9%
SMART Route 565	0	0	0	0.0%
SMART Route 580	0	0	0	0.0%
SMART Route 610	103	119	-16	-13.7%
SMART Route 615	0	0	0	0.0%
SMART Route 620	0	0	0	0.0%
SMART Route 635	0	0	0	0.0%
SMART Route 710	184	235	-51	-21.6%
SMART Route 730	0	0	0	0.0%
SMART Route 740	156	121	35	29.1%
SMART Route 752	16	4	12	321.4%
SMART Route 753	39	56	-17	-30.1%
SMART Route 756	0	0	0	0.0%
SMART Route 760	0	22	-22	-100.0%
SMART Route 780	256	134	121	90.5%
SMART Route 805	0	11	-11	-100.0%
SMART Route 830	0	0	0	0.0%
SMART Route 851	0	25	-25	-100.0%
UMI	3,594	3,400	194	5.7%
SUM	27,146	27,240	-94	-0.3%