

Household Survey Processing

KATS Travel Model Update

Technical

Memorandum 5

prepared for

Kalamazoo Area Transportation Study

prepared by

Cambridge Systematics, Inc.

technical memorandum 5

Household Survey Processing

KATS Travel Model Update

prepared for

Kalamazoo Area Transportation Study

prepared by

Cambridge Systematics, Inc.
999 18th Street, Suite 3000
Denver, CO 80202

date

October 24, 2014

Table of Contents

1.0	Introduction	1-1
2.0	Survey Weighting and Expansion	2-1
2.1	Comparison to Observed Distributions	2-2
2.2	Expansion Factor Development.....	2-3
2.3	Use of Additional Records	2-5
3.0	Trip Purposes.....	3-6
4.0	TAZ Identification.....	4-7

List of Tables

Table 2.1	Survey and ACS Comparison: Household Income	2-2
Table 2.2	Survey and Census Comparison: Household Size	2-3
Table 2.3	Survey and ACS Comparison: Household Vehicles	2-3
Table 2.4	Survey and ACS Comparison: Household Workers	2-3
Table 2.5	Expansion Summary: Household Income	2-4
Table 2.6	Expansion Summary: Household Size	2-4
Table 2.7	Expansion Summary: Household Vehicles	2-4
Table 2.8	Expansion Summary: Household Workers	2-5
Table 3.1	Trip Purpose Identification.....	3-6
Table 3.2	Number of Trips by Purpose.....	3-7

1.0 Introduction

In 2004 and 2005, MDOT and its consultants conducted the MI Travel Counts (MITC) project, a comprehensive household travel data collection program. This was a household-based travel diary survey that recorded trip-making for approximately 14,000 households statewide, including 291 households within Kalamazoo County and 33 in Van Buren County. This was a 48-hour survey in which trip-making was recorded for each participating household over a 2-day period.

The MITC serves as the foundation for estimation of the trip generation model, development of trip distribution and mode choice calibration targets, and estimation of time of day model parameters. The survey includes information collected from participating households including demographic information such as household income and size, as well as detailed diaries of all travel conducted by participating household members during the survey period.

Prior to delivery for use in the KATS model update, MDOT processed the household survey dataset and performed significant quality control review of the contents. MDOT also geocoded each location in the survey, producing geographic coordinates for each trip end for all recorded trips. The survey process and results are documented in the document 2004-2005 Comprehensive Household Travel Data Collection Program, MI Travel Counts, Final Report (August 31, 2005).

2.0 Survey Weighting and Expansion

Prior to using the household travel survey to develop model parameters, it was necessary to compare the demographic and geographic distribution of survey records to observed distributions obtained from American Community Survey (ACS) data and the Census. Ideally, distributions of surveyed households will match the distributions obtained from ACS and the Census. However, varying response and participation by different groups often lead to differences between surveyed and observed distributions. Such differences, in turn, could impact the calculation of key metrics such as total trips, trips by purpose, trip distribution, time-of-day of travel, and mode splits.

Therefore, it is essential that a set of weighting and expansion factors be created to reconcile survey records with observed distributions. These factors

simultaneously *weight* the survey records so that distributions match observed data, and *expand* the survey records to represent the total population of the KATS region. Expansion factors were only developed for households located within the KATS modeling area. Development of these weighting and expansion factors is described below.

2.1 COMPARISON TO OBSERVED DISTRIBUTIONS

The Census bureau provides household and person tabulations in the ACS data and Census summaries. The 2008 - 2012 5-year dataset and 2010 Census, were used to expand the survey data, since this range is centered around the KATS model's 2010 base year. Because the Census is a full count, the ACS totals were all matched to the Census totals. The ACS was used for distributions which were not included in the available Census summaries.

Distribution of survey records and Census/ACS data were compared for each of the following variables:

- Household Income (3 variables, ACS),
- Household Size (5 variables, Census),
- Number of Workers (4 variables, ACS), and
- Number of Vehicles (4 variables, ACS).

Household survey and Census/ACS distributions are shown in Table 1.1 through Table 1.4. Each of these tables also includes the expanded totals. A normalized sampling factor greater than 1 indicates an under-sampled category, while a factor less than 1 indicates an over-sample. As expected, the household survey captured sufficient records in most of the reviewed categories. Two categories (0-vehicle households and 3+ worker households) have relatively small samples, but represent a small portion of the regional population. This suggests that caution should be used when evaluating results for these categories.

Table 2.1 Survey and ACS Comparison: Household Income

Income Group	Household Survey Records	Household Survey Records (%)	ACS Records	ACS Records (%)	Normalized Sampling Factor
Low (\$0 - 20k)	55	18%	23,831	22%	1.20
Middle (\$20 - 60k)	129	42%	43,781	40%	0.94
High (\$60k +)	123	40%	43,148	39%	0.97
Total	307	100%	110,760	100%	1.00

Note: Household records with missing income data are excluded from this table.

Table 2.2 Survey and Census Comparison: Household Size

Household Size	Household Survey Records	Household Survey Records (%)	Census Records	Census Records (%)	Normalized Sampling Factor
1	90	28%	31,982	29%	1.04
2	117	36%	38,508	35%	0.96
3	56	17%	16,788	15%	0.88
4	36	11%	13,976	13%	1.14
5+	25	8%	9,506	9%	1.11
Total	324	100%	110,760	100%	1.00

Table 2.3 Survey and ACS Comparison: Household Vehicles

Number of Autos	Household Survey Records	Household Survey Records (%)	ACS Records	ACS Records (%)	Normalized Sampling Factor
0	20	6%	8,130	7%	1.19
1	109	34%	39,529	36%	1.06
2	134	41%	43,589	39%	0.95
3+	61	19%	19,511	18%	0.94
Total	324	100%	110,760	100%	1.00

Table 2.4 Survey and ACS Comparison: Household Workers

Number of Workers	Household Survey Records	Household Survey Records (%)	ACS Records	ACS Records (%)	Normalized Sampling Factor
0	81	25%	32,192	29%	1.15
1	132	41%	42,533	38%	0.94
2	95	30%	30,690	28%	0.92
3+	13	4%	5,345	5%	1.21
Total	321	100%	110,760	100%	1.00

2.2 EXPANSION FACTOR DEVELOPMENT

The variables defined above result in 204 unique combinations of household categories (e.g., 4+ size, 2+ workers, 2 vehicles, and medium income). Sufficient data are not available in the household survey data to produce expansion factors at the individual 204-cell level. Consequently, a raking procedure or iterative proportional fitting (IPF) process was implemented to establish expansion weights in an iterative fashion. Resulting factors produce household totals within 1% of the target ACS values.

The ACS data were used to establish univariate, or marginal, totals for control variable categories. The household survey data were then tabulated by the

control variable categories to provide an initial estimate of the joint distribution. The joint distribution cells were factored to match marginal totals sequentially for each of the 4 control variables (i.e., income, household size, number of vehicles, and number of workers). Factors were applied iteratively until expanded marginal totals matched the target values developed from ACS data.

Expansion factors were developed so that when applied, the sum of expansion factors for all household records in a category will sum to the number of observed households. These household expansion factors are applicable to every person and trip record for use in survey analysis. Weighted household survey summaries and ACS distributions are shown in Tables 2.5 through 2.8.

Table 2.5 Expansion Summary: Household Income

Income Group	Household Survey Records	Household Survey Records (%)	ACS Records	ACS Records (%)
Low (\$0 - 20k)	22,236	21%	23,831	22%
Middle (\$20 - 60k)	40,925	40%	43,781	40%
High (\$60k +)	40,391	39%	43,148	39%
Total	103,552	100%	110,760	100%

Note: Household records with missing income data are excluded from this table.

Table 2.6 Expansion Summary: Household Size

Household Size	Household Survey Records	Household Survey Records (%)	Census Records	Census Records (%)
1	32,058	29%	31,982	29%
2	38,515	35%	38,508	35%
3	16,759	15%	16,788	15%
4	13,940	13%	13,976	13%
5+	9,489	9%	9,506	9%
Total	110,760	100%	110,760	100%

Table 2.7 Expansion Summary: Household Vehicles

Number of Autos	Household Survey Records	Household Survey Records (%)	ACS Records	ACS Records (%)
0	8,162	7%	8,130	7%
1	39,600	36%	39,529	36%
2	43,525	39%	43,589	39%
3+	19,472	18%	19,511	18%
Total	110,760	100%	110,760	100%

Table 2.8 Expansion Summary: Household Workers

Number of Workers	Household Survey Records	Household Survey Records (%)	ACS Records	ACS Records (%)
0	32,192	29%	32,192	29%
1	42,533	38%	42,533	38%
2	30,690	28%	30,690	28%
3+	5,345	5%	5,345	5%
Total	321	100%	110,760	100%

If a single expansion factor were applied to all records, each survey record would reflect approximately 342 households. Weighted expansion factors range from 182.7 to 544.5. These expansion factors are reasonable, with normalized expansion factors ranging from 0.53 to 1.59.

2.3 USE OF ADDITIONAL RECORDS

The MI Travel Counts survey dataset contains 324 household records within the KATS modeling area, for an equivalent of 3,152 un-expanded trip records made by study area households. This information may be sufficient for use in some analysis such as estimation of trip attraction rates, development of trip distribution calibration targets, and creation of model choice calibration targets. However, the 324 sampled households do not provide a sufficient sample size for estimation of cross-classified trip production rates.

Production rate cross classification analysis requires analysis for 67 individual categories (6 trip purposes along with 4, 9, or 15 household types depending on trip purpose). Household samples from outside of the KATS modeling area were added to the dataset for analysis tasks requiring additional data. Discussions with KATS and MDOT staff identified several areas similar in size and character to the Kalamazoo area. Areas included in the borrowed dataset include:

- Small Urban Modeled Areas (SUMAs, region code 7);
- Transportation Management Areas (TMAs, region code 6); and
- Washtenaw County.

These additional records are only used where records within the KATS area are insufficient for analysis. The larger dataset includes 4,336 households.

3.0 Trip Purposes

After review of survey data, MDOT's Urban Model Improvement Program (UMIP) documents, and discussion with the project team, the trip purposes listed below were identified for inclusion in the updated KATS Model.

- Home-Based Work (HBW): Trips between a traveler's residence and workplace.
- Home-Based Shop (HBS): Shopping trips starting or ending at the traveler's residence.
- Home-Based School (HBSc): Trips made by students between home and a K-12 school.
- Home-Based University (HBU): Trips made by university students and visitors between home and a traditional 4-year university.
- Home-Based Other (HBO): All remaining trips starting or ending at the traveler's residence.
- Work-Based Other (WBO): Trips starting or ending at the workplace, but with neither end at the traveler's residence.
- Other-Based Other (OBO): Trips that do not start or end at the traveler's residence or workplace.

A trip purpose was assigned to each of the records present in the household survey dataset. Trip purpose is defined based on the origin and destination place type as shown in Table 3.1. The resulting numbers of trips by purpose are shown in Table 3.2.

During processing of survey data, each trip was identified as either a production to attraction trip, or an attraction to production trip. For home-based trip purposes, the trip production end was identified as the trip maker's home. For WBO trips, the trip production was identified as the trip maker's workplace, or the work-related location if the trip was not made to or from the primary workplace. For OBO trips, the trip production was identified as the trip origin.

Table 3.1 Trip Purpose Identification

Production Place Type	Attraction Place Type	Trip Purpose
Home	Workplace	HBW
Home	Shopping	HBS
Home	School (K-12)	HBSc
Home	University	HBU
Home	Other non-home	HBO

Workplace	Workplace	WBO
Workplace	Non-home and Non-work	WBO
Non-home and Non-work	Non-home and Non-work	OBO

Notes: Other non-home includes childcare, eat out, personal business, religious/community, social, recreation, accompany another person, and pick-up/drop-off passenger. Trips with an activity of "turn around" are excluded from analysis.

Table 3.2 Number of Trips by Purpose

Trip Purpose	Number of Trips (Study Area)	Percent of Trips (Study Area)	Number of Trips (with borrowed data)	Percent of Trips (with borrowed data)
HBW	424	13%	6,600	17%
HBS	224	7%	3,469	9%
HBS _c	227	7%	3,384	9%
HBU	22	1%	245	1%
HBO	848	27%	12,170	31%
WBO	295	9%	3,772	10%
OBO	1,112	35%	9,040	23%
Total	3,152	100%	38,435	100%

Note: Study area values are weighted and expanded. Values with borrowed data are not weighted or expanded.

4.0 TAZ Identification

The household travel survey contained latitude and longitude coordinates for each household and for each place visited. In some cases, the geocoding process used to convert address data provided by survey respondents into geographic coordinates either failed, or placed records at a central zip code location. In such cases, survey records can be used in analysis that does not require specific geographic location, but must be dropped from analysis that requires geolocation of trip-ends. All records with valid coordinates have been matched to a traffic analysis zone for further analysis.