Appendix A. Detailed Literature Review

A-1. A Basic Cost-Benefit Analysis for Univariate Statistics

Multi-day travel surveys offer a potential to reduce survey costs by sampling fewer households. This potential exists because a large portion of the survey cost is involved in recruiting respondents, so the cost of adding a second day of data collection is expected to be less than the cost of recruiting a second household. To understand whether that potential is realized, it is necessary to consider the value of the additional data versus its cost. Studies addressing this topic in the case of univariate statistics are discussed below.

Stopher, Kockelman, Greaves, and Clifford (2008) (called 'SKGC 2008' below) provide a framework for the evaluation of variance for multiday surveys in the GPS era. They develop a model for person kilometers of travel (PKT) per day as follows:

$$y_{jt} = \mu + \delta_j + \varepsilon_{jt} \quad \text{Eq } (2-1)$$

where y_{jt} is individual j's PKT on day t, μ is the overall mean PKT, δ_j is the difference between the overall mean and person j's mean daily travel, and ε_{jt} is the difference between the person-level mean $\mu + \delta_j$ and the particular day's travel.

Heteroscedasticity is allowed for by defining σ_j^2 as the variance of ε_{jt}, and allowing for differing σ_j^2 values across individuals j, but then quickly revert to a simpler variance structure assuming a common σ_ε^2 over all individuals j. With this assumption, their model is a simple special case of the Pas (1986) model as given above in Equation 1. They define a $\sigma_{j|u}^2$ which is the variance of δ_j, and define a constant $K = \sigma_{j|u}^2 + \sigma_\varepsilon^2$. Note that K is the same as σ^2 in the Pas (1986) model.

SKGC 2008 provides an empirical example from two waves of a GPS-based panel survey in South Australia. In the first wave, the ratio $\sigma_\varepsilon^2 / \sigma_{j|u}^2$ was 3.36. In the second wave, the ratio was 17.35. Defining a as the correlation coefficient $a = \frac{\sigma_{j|u}^2}{\sigma_{j|u}^2 + \sigma_\varepsilon^2}$ we have a =22.9%, and a =5.5% in the two waves respectively. Their design effect reduces after some algebra to Pas's design effect $\frac{1+a(T-1)}{T}$. The table below summarizes their conclusions. The design effects represent reductions in person-level sample sizes when the number of days T is increased to 7 or 15 (from a benchmark single-day survey).

Table A-1-1. Correlation coefficients and design effects for South Australia GPS Study.

Ratio $\sigma_\varepsilon^2/\sigma_{j\|u}^2$	Correlation coefficient a	Number of days T	Design effect— $\frac{1+a(T-1)}{T}$.
3.36	22.9%	7	33.9%
3.36	22.9%	15	28.1%
17.35	5.5%	7	19.0%
17.35	5.5%	15	11.8%

The second wave variance components were based on some outlying values (some persons who took very long trips on particular days, which increased the day-to-day variability considerably), and thus has greater variability. We believe the 5.5% value may be somewhat of an outlier, but such high within-person variabilities may occur as a matter of course in the presence of outlying infrequent, long trips.

SKGC 2008 also provides explicit cost functions, which is rare in the literature, based on their experience in the Australian context with both telephone recruitment and face-to-face recruitment. These are given in Table A-1-2. Two types of recruitment are compared: telephone recruitment and face-to-face recruitment. The cost per household includes the cost from the loss of nonrespondents and the cost of followup to complete the interviews among the respondents. The diary survey is assumed to cover one single travel day. The GPS survey is assumed to cover 15 travel days, and includes a component for processing the GPS data for each collected day. The diary survey covers 3,000 households (and 3,000 collected household-days) and the GPS study covers 825 households (12,750 collected household-days).

Table A-1-2. Costs for four types of surveys in the Australian context.

	Household sample size	Number of collected travel days	Total collected travel days	Cost per household	Total cost
Diary Survey (Telephone Recruit)	3,000	1	3,000	$175	$525,000
GPS Survey (Telephone Recruit)	850	15	12,750	$500	$425,000
Diary Survey (Face-to-face Recruit)	3,000	1	3,000	$350	$1,050,000
GPS Survey (Face-to-face Recruit)	850	15	12,750	$680	$578,000

The Pas (1986) paper (described in greater detail in Section 4 below) develops an explicit cost model for comparing single-day and multiday studies. The cost of collecting T days of travel behavior from a single individual is assumed to be $C = p + qT$ where q is the cost of each collected day, and p is an 'overhead' cost for recruiting the individual. If N_M and N_S are the person-level sample sizes for the putative multiday and single-day studies, then the costs of these surveys using this simple cost model are

$$C_M = (p + qT)N_M, \quad C_S = (p + q)N_S \qquad \text{Eq}(2-2)$$

Suppose C_S is the cost of a benchmark single stage study with sample size N_S that achieves set variance level V. Then $C_M = K_C C_S$ is the cost of a multiday study with T days that achieves the same variance level, with

$$K_C = \left(1 + \frac{q}{p}T\right)\left(1 + \frac{q}{p}\right)^{-1} \frac{1 + a(T-1)}{T} \qquad \text{Eq}(2-3)$$

Pas (1986) calls this a 'cost scale factor'. If K_C is greater than 1, then that means the single-day study that achieves the same precision is less expensive. If K_C is considerably smaller than 1, that means the multi-day study is less expensive. We can find the optimal T for given values of q, p, and a. Pas (1986) assumed fairly large values of q/p being in the pre-GPS environment, and the optimal number of days T was not always a large number. With GPS technology, q/p may be much smaller (assuming the simple cost model as given is still valid: this may not be the case). This will mean the optimal designs may correspond to much larger values of T.

For example, suppose we apply q/p values that are consistent with the SKGC 2008 paper, which come from a GPS study. These q/p values come from a GPS study where there is fairly extensive processing of the GPS data, resulting in higher q/p ratios. The cost assumptions from this paper are given in Table A-1-2 above. One assumption that can be made is to assume that the cost of recruitment into the diary study ($175 for telephone, $350 for face-to-face) is p, and then q can be computed by taking the cost of the 15-travel day GPS study ($500 for telephone, $680 for face-to-face) by subtracting the latter cost from the former cost, and dividing by 15. This will result in a q value of $21.75 for the telephone recruitment scenario and $22.00 for the face-to-face recruitment scenario. The q/p values are Tables A-1-3 and A-1-4 provide two scenarios from the SKGC 2008 paper to illustrate the calculation of the K_C factor based on Equation (2-3) and the calculation of the optimal number of days for each of the four scenarios. The Cost Ratio is the factor $\left(1 + \frac{q}{p}T\right)\left(1 + \frac{q}{p}\right)^{-1}$ from Equation 2-3, and the Design Effect is the factor $\frac{1+a(T-1)}{T}$.

Both tables include a correlation coefficient of 22.9% (that from PKT for the South Australia study, Wave 1). Table A-1-3 presents a scenario with a q/p of 0.124, and Table A-1-4 a q/p of 0.063.

Table A-1-3. Calculation of K_C factor and optimal T for q/p equal to 0.124, correlation 22.9%.

p-marginal cost per house-hold	q-cost per day	q/p	Corre-lation a	Number of days T	Cost Ratio	Design Effect	K_C factor
$175	$21.67	0.124	22.9%	1	1.000	1.00	100.0%
$175	$21.67	0.124	22.9%	2	1.110	0.61	68.2%
$175	$21.67	0.124	22.9%	3	1.220	0.49	59.3%
$175	$21.67	0.124	22.9%	4	1.331	0.42	56.2%
$175	**$21.67**	**0.124**	**22.9%**	**5**	**1.441**	**0.38**	**55.3%**
$175	$21.67	0.124	22.9%	6	1.551	0.36	55.5%
$175	$21.67	0.124	22.9%	7	1.661	0.34	56.4%
$175	$21.67	0.124	22.9%	10	1.992	0.31	61.0%
$175	$21.67	0.124	22.9%	15	2.542	0.28	71.4%

Table A-1-4. Calculation of K_C factor and optimal T for q/p equal to 0.063, correlation 22.9%.

p-marginal cost per house-hold	q-cost per day	q/p	Corre-lation a	Number of days T	Cost Ratio	Design Effect	K_C factor
$350	$22.00	0.063	22.9%	1	1.000	1.000	100.0%
$350	$22.00	0.063	22.9%	2	1.059	0.615	65.1%
$350	$22.00	0.063	22.9%	5	1.237	0.384	47.4%
$350	$22.00	0.063	22.9%	6	1.296	0.358	46.4%
$350	**$22.00**	**0.063**	**22.9%**	**7**	**1.355**	**0.339**	**46.0%**
$350	$22.00	0.063	22.9%	8	1.414	0.326	46.1%
$350	$22.00	0.063	22.9%	10	1.532	0.306	47.0%
$350	$22.00	0.063	22.9%	12	1.651	0.294	48.5%
$350	$22.00	0.063	22.9%	15	1.828	0.281	51.3%
$350	$22.00	0.063	22.9%	20	2.124	0.268	56.9%

The optimal number of days is 5 for the q/p of 0.124 and is 7 for the q/p of 0.063. The optimal K_C factor is 55% for q/p of 0.124 and 46% for q/p equal to 0.063. The larger 'upfront' cost pushes the optimal design towards more days, and a greater cost reduction can be achieved as well.

A-2. Cost Function Parameter Considerations

Another critical parameter (assuming the linear cost model from Pas (1986) is a reasonable approximation of the true cost structure) is the q/p ratio. In Section 2, a q/p ratio is derived from cost documentation provided from Stopher, Kockelman et al. (2008). In a GPS-based study, p will include all 'one-time costs': the cost of recruiting the household, carrying out interviews (both initial interviews and possibly follow-up interviews), and the cost of providing GPSs to the household. The q parameter includes costs for 'each extra day', which include any costs which are incurred for collecting each separate travel day per se. These costs include the costs of cleaning and analyzing the GPS data that comes back from the households, as the magnitude of these costs should generally be a linear function of the number of travel days. The status of the cost of any follow-up recall interview is somewhat problematic. It is a one-time event (following the return of the GPSs), but for example the monetary incentive to the household may need to be larger if there are a larger number of travel days and the interview is longer, and analyzing and cleaning this interview data may be proportional to the number of travel days recoded . Any costs need to be separated out as one-time and per-travel day to make the cost model relevant.

In the 2012 Northeast Ohio Regional Travel Survey (Wilhelm et al. (2013)), part of the survey was done with GPS only and part was done with GPS and a prompted recall interview. These parts were randomly assigned with a 2:1 ratio for GPS only to GPS with prompted recall[8]. The prompted recall presented the sampled persons with their GPS trips, and asked them questions about trip purpose and other trip details using a CATI or Web-based interview. The quality of the data was higher with the prompted recall followup, but collecting this data was a considerable extra expense.

The separation of q and p costs are heavily tied to the degree of processing and followup of the GPS data, and this in turn determines the quality of the data. These are design choices that need to be made carefully.

A-3. Proposed Formula for Sample Size Reduction

As noted in the text, Parsons Brinckerhoff, et al (2014) propose a formula for a multi-day survey versus its single day equivalent. That formula is discussed in further detail here. It is:

$$S_N = S_0 * \frac{R + D}{(R + 1) * D} \qquad \text{Eq } (3 - 1)$$

where

- S_N is the new (reduced) sample size,
- S_0 is the sample size for a one-day survey,
- R is the ratio of day-to-day (intra-person) variability σ_ε^2 to inter-person variability $\sigma_{j|u}^2$.

[8] Households with all members age 75 and over were not part of this randomization, and did not receive a GPS.

- D is the sample length in days.

This method is referred to in the report as the Vovsha method, after one of the reports authors, and we continue with that conventions. It should be noted that the design effect $\frac{R+D}{(R+1)*D}$ is equivalent to the Pas design effect $\frac{1+a(T-1)}{T}$ in Section 4. The correlation coefficient a is $\frac{\sigma_{j|u}^2}{\sigma_\varepsilon^2+\sigma_{j|u}^2}$, and $R = \frac{\sigma_\varepsilon^2}{\sigma_{j|u}^2}$, so that $\frac{1}{a} = R + 1$. After replacing T in the Pas Equation with D, we obtain $1 + a(D-1) = 1 + \frac{D-1}{R+1} = \frac{R+D}{R+1}$, so that $\frac{1+a(D-1)}{D} = \frac{R+D}{(R+1)*D}$. In this formula, as R approaches zero, S_N will equal S_0. Such a situation would apply if there is no variability in the data over time, such as with auto ownership, where the number of vehicles owned is the same on all travel days. With no intra-person variability, this formula indicates no value to collecting additional travel days. Conversely, as R approaches infinity, the new sample size is the one-day sample size divided by the number of days ($S_N = S_0/D$). In other words, this situation would imply that adding additional days is equivalent to randomly sampling additional households.

There are two important points to make about this derivation. First, it is specific to one particular component of travel, so will be different for a model like car ownership versus destination choice. Vovsha addresses this by considering the relative importance of several model components to achieve a weighted average of the equivalent sample size across all model components. Second, the value of R is central to the calculation, and in the analysis provided is taken as an assumption.

Table A-3-1 presents an example application of the Vovsha method, as reported in Parsons Brinckerhoff, et al (2014). In this example, the assumed R values vary from 0 for the first three long-term choices, to 4 for the daily pattern of trips. Using the assumed relative importance of each model component, they estimate that a sample of 3,000 households surveyed for three days is roughly equivalent to 5,000 households surveyed for 1 day.

Table A-3-1. Example Application of Vovsha Method

Sub-model / Travel dimension	S_N	D	R	S_N*[(1+R)*D / (R+D)]	Relative importance	S_0
Car ownership	3,000	3	0	3,000	0.1	300
Workplace choice	3,000	3	0	3,000	0.1	300
School choice	3,000	3	0	3,000	0.1	300
Daily pattern of trips	3,000	3	4	6,429	0.2	1,286
Non-work dest. choice	3,000	3	2	5,400	0.1	540
Time of day choice	3,000	3	3	6,000	0.1	600
Mode choice	3,000	3	2	5,400	0.2	1,080
VMT (Stopher)	3,000	3	3.36	6,170	0.1	617
Overall Assessment	3,000	3			1	5,023

It is worth noting that the R values in the Vovsha method have a similar, but inverse interpretation to the design effects at the heart of this study (and discussed further in the methodology report). In both cases, a value of 1 indicates an equivalency between adding households and adding days, but lower R values indicate less value to adding days, whereas higher design effects indicate less value to adding days.

A-4. Within and Between-Person Variance Shares from Other Univariate Studies

A number of studies provide variance components for a variety of univariate travel characteristics in the multiday context. As noted above, the relative values of the day-to-day versus person-to-person variance is important to understanding the value contributed by collecting additional days of travel data. In that context, we examine existing evidence for these values.

Table A-4-1 presents variance components from the various recent papers by Stopher and Pas. Included is the travel characteristic measured, the between person variability as a fraction of total variability (also equal to the correlation coefficient), and the within-person variability as a fraction of total variability (one minus the correlation coefficient). As can be seen these shares vary considerably across travel characteristics, studies, and types of study (GPS or Diary). The highest within-person variability is registered as 94.6%, but as indicated in SKGC 2008 this is somewhat driven by outlying observations. The correlation coefficients generally range from 20% to 80%.

Table A-4-1. Variance Components for Travel Characteristics from Recent Literature (I).

Travel characteristic	Between-person variability share	Within-person variability share	Study	Number of days	Type of study	Reference
Person kilometers per day (PKT)	22.9%	77.1%	South Australia Wave 1	15-30	GPS	SKGC 2008
Person kilometers per day (PKT)	5.5%	94.6%	South Australia Wave 2	15-30	GPS	SKGC 2008
Stops	50.5%	49.5%	Reading, UK	5	Diary	Pas (1986)
Tours	48.0%	52.0%	Reading, UK	5	Diary	Pas (1986)
Subsistence stops	72.8%	27.2%	Reading, UK	5	Diary	Pas (1986)
Maintenance stops	39.0%	61.0%	Reading, UK	5	Diary	Pas (1986)
Leisure stops	37.2%	62.8%	Reading, UK	5	Diary	Pas (1986)
Trips	62.0%	38.0%	Seattle, WA	3	Diary	Pas and Sundar (1995)

As one might expect, the 'subsistence stops' have the lower within-person variability, as these shouldn't be much different across days. On the other hand maintenance stops and leisure stops have much higher variability within persons across days. The lower correlation coefficients for these are consistent with lower variances one might expect for these travel characteristics for multiday studies.

These results are also replicated in Kang and Scott (2009). The empirical results in this paper are based on the Toronto Activity Panel Survey from Toronto, Ontario in Canada. This survey was a seven-day survey (seven consecutive days), and was diary-based. Table A-4-2 below is from Table 2 in Kang and Scott (2009). The results are roughly consistent with the Table A-4-1 results. The seven-day period covering weekdays and weekends has a much higher within-person variability than weekdays or weekends taken alone.

Table A-4-2. Variance Components for Travel Characteristics from Kang and Scott (2009).

Time Scale	Type of Trip	Between person variability share	Within person variability share
Across one week	Independent maintenance	29%	71%
Across one week	Independent discretionary	34%	66%
Across one week	Joint maintenance	18%	82%
Across one week	Joint discretionary	24%	76%
Weekdays	Independent maintenance	45%	55%
Weekdays	Independent discretionary	42%	58%
Weekdays	Joint maintenance	22%	78%
Weekdays	Joint discretionary	29%	71%
Weekend days	Independent maintenance	61%	39%
Weekend days	Independent discretionary	78%	22%
Weekend days	Joint maintenance	53%	47%
Weekend days	Joint discretionary	65%	35%

Pendyala (2014) provides considerable empirical data for these variance components based on a recent pilot experiment in the Lexington, KY area. In this experiment, one vehicle in each of 100 households was fit with a GPS unit and a hand-held computer was also provided (81 households provided usable data for at least three days). The three-weekday sample showed within-person variability shares of right around 50% for total trips, non-work trips, mid-day non-work trips, PDA[9] travel time, and GPS travel time. The corresponding within-person variability shares were higher for the 3-5 weekday sample and the 3-5 day sample, but these were based on smaller data sets. The within-person variability shares for GPS VMT[10], GIS VMT and first home departure time for the three-weekday sample were around 60%. Only final home arrival time and final work departure time were under 50%, with final work departure time around 30%. One would expect this to show the smallest day-to-day variability. The corresponding correlation coefficients are all then in the 30%-70% range as for Table A-4-1.

A-5. Multivariate Analysis of Multiday Studies

Moving beyond univariate measures, multi-variate analysis is also of interest. Pas (1986) and Koppelman and Pas (1984) were written during the pre-GPS era, but much of the theoretical development in these papers is still relevant in the current period (though the multiday 'cost-

[9] Personal Digital Assistant device.

[10] Vehicle Miles Traveled.

structure index' (q/p using Pas's notation) is much smaller in GPS-only surveys than in old-style diary surveys).

Pas (1986) and Koppelman and Pas (1984) start with a linear trip-generation model for daily travel as follows:

$$Y_{jt} = \mathbf{X}'_{jt}\boldsymbol{\beta} + \varepsilon_{jt}, \ \ j = 1, \dots, J, t = 1, \dots, T \qquad \text{Eq (A.8.1)}$$

where

- Y_{jt} is the number of trips by individual j on day t,
- $\mathbf{X}_{jt}$ is a $K \times 1$ vector of variables describing individual j and their environment on day t,
- $\boldsymbol{\beta}$ is a $K \times 1$ vector of parameters,
- ε_{jt} is an error term for individual j on day t.

The random error term has the following properties. The ε_{jt} are multivariate normal all with expectation 0 and

$$Cov(\varepsilon_{jt}, \varepsilon_{j't'}) = \begin{cases} \sigma^2 = \sigma_u^2 + \sigma_w^2 & j = j', t = t' \\ \sigma_u^2 & j = j', t \neq t' \\ 0 & j \neq j' \end{cases} \qquad \text{Eq(A.8.2)}$$

They call this the 'crossed-error structure' (terminology from Fuller and Battese (1974)). Define a as the correlation coefficient $= \frac{\sigma_u^2}{\sigma_u^2+\sigma_w^2}$, which represents the correlation between ε_{jt} and $\varepsilon_{jt'}$ within an individual.

For a one-day travel study (a travel study with a diary for each individual covering only one day) under this model $\boldsymbol{\beta}$ is estimated as $\widehat{\boldsymbol{\beta}}_s = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-\mathbf{1}}\left(\mathbf{X}^{\mathbf{T}}\mathbf{Y}\right)$, where $\mathbf{X}$ is a matrix of order $J \times K$, and $\mathbf{Y}$ is a vector of length J. The variance matrix for $\widehat{\boldsymbol{\beta}}_s$ under the model is

$$Var\left(\widehat{\boldsymbol{\beta}}_s\right) = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-\mathbf{1}}\sigma^2 \qquad \text{Eq(A.8.3)}$$

For multi-day travel (a travel study with a diary for each individual covering $T > 1$ days), the papers deal with the following special case. The vector $\mathbf{X}_{\boldsymbol{j}}$ is assumed to be fixed across the T days for each individual j, and an average is taken across the T days for the y-values for each individual, generating a $\overline{\mathbf{Y}}$ mean vector of length J. The regression parameter estimate here is $\widehat{\boldsymbol{\beta}}_M = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-\mathbf{1}}\left(\mathbf{X}^{\mathbf{T}}\overline{\mathbf{Y}}\right)$. Based on their assumed 'crossed-error structure' model, the variance of this parameter estimate is $Var\left(\widehat{\boldsymbol{\beta}}_M\right) = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-\mathbf{1}}\sigma^2\frac{1+a(T-1)}{T}$.

Under this model, a very simple relationship is derived between the two variance matrices:

$$Var\left(\widehat{\boldsymbol{\beta}}_M\right) = \frac{1 + a(T-1)}{T} Var\left(\widehat{\boldsymbol{\beta}}_s\right) \qquad \text{Eq(A.8.4)}$$

Even though the two matrices are $K \times K$, the ratio between each pair of corresponding variances and covariances (K variances and $K(K-1)/2$ unique covariances) is the constant value $\frac{1+a(T-1)}{T}$, which can be interpreted then as a design effect from multiday sampling as opposed to single-day sampling. Suppose N_S is the sample size for a benchmark single stage study that achieves variance level V. Then $N_M = \frac{1+a(T-1)}{T} N_S$ will achieve the same precision in a multiday study with T collected days rather than 1 collected day. As T becomes larger, the precision-matching N_M should be become progressively smaller than the benchmark N_S. The actual function depends on a. In the extreme case of $a = 0$ (days are independent within individuals, and each extra day provides as much information as the first day), the precision matching N_M will be equal to N_S/T, and will get very small as T increases. In the other extreme case of $a = 1$ (further days are all the same within individuals as the first day and provide no new information), the precision-matching N_M will be equal to the benchmark N_S no matter the value of T. Adding extra collection days does not allow any reduction in the person-level sample sizes.

Note: references for Appendix A are included with references for the main text.

Appendix B. Jackknife Variance Estimation

The version of the jackknife we will use here will follow the stratification structure for this travel survey, as outlined in Section 1.

The sample size within each stratum is $n_s,\ s = 1, \ldots, S$. We subscript sampled households within each stratum as $sh, s = 1, \ldots, S, h = 1, \ldots, n_s$. The sample weight for each household is w_{sh} (also called the 'full-sample weight' to distinguish it from the replicate weights. The three y-estimators can be rewritten as:

$$\bar{y}^{(1)} = \frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh} y_{sh1}}{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}} \quad \bar{y}^{(2)} = \frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}(y_{sh1}+y_{sh2})}{2*\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}} \quad \bar{y}^{(3)} = \frac{\sum_{s=1}^{S}\sum_{h=1}^{n_s} w_{sh}\sum_{d=1}^{D_{sh}} y_{shd}}{\sum_{s=1}^{S}\sum_{h=1}^{n_s} w_{sh}*D_{sh}}$$

We summarize these three expressions as $\bar{y}^{(d)}$, $d = 1,2,3$:

$$\bar{y}^{(d)} = \frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh} y_{sh}^{(d)}}{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}} \quad \text{with } y_{sh}^{(1)} = y_{sh1},\ y_{sh}^{(2)} = \frac{y_{sh1} + y_{sh2}}{2},\ y_{sh}^{(3)} = \frac{\sum_{d=1}^{D_{sh}} y_{shd}}{D_{sh}}$$

We created replicate groups with roughly 10 sampled households. Thus the total number of replicate groups m_s for each stratum is $[n_s/10]$ ($n_s/10$ rounded to the nearest integer). This resulted in 453 replicate weights.

We subscript the replicate groups as $st, s = 1, \ldots, S, t = 1, \ldots, m_s$. Write $S(s)$ as the set of households mapped to stratum s and $S(st)$ as the set of households mapped to replicate group st. The replicate weights are subscripted as $s't'$, with $s' = 1, \ldots, S,\ t' = 1, \ldots, m_s$, then the replicate weights $w_{sh}(s't')$ are defined as follows:

$$w_{sh}(s't') = \begin{cases} 0 & s = s', h \in S(s't') \\ \dfrac{m_s}{(m_s - 1)} w_{sh} & s = s', h \notin S(s't') \\ w_{sh} & s \neq s' \end{cases}$$

See for example Valliant et al. (2013), Section 15.4.1.

The strata are given in Table B-1-1 below. The strata are crossings of county (five digit FIPS code) and Sample Type[11]. Strata with less than 10 households were collapsed with other strata in the same county (e.g., 39055_67 collapses 39055_6 and 39055_7).

---

11 Sample Type is 1=Address-based matched Sample; 2=Address-based Unmatched Sample; 3=General Listed Sample; 4=Target Large HH (3+ persons); 5=Target one-peron household with income less than $25,000 annually; 6=Other low income household; 7=High probability zero-vehicle household; 8=General listed transit oversample. See Wilhelm et al. (2013) for details of sample stratification.

Table B-1-1. Stratification structure for jackknife replicate weights.

Variance Stratum	House-holds	Number of Variance Strata
39035_1	493	49
39035_2	686	69
39035_3	272	27
39035_4	393	39
39035_5	27	3
39035_6	157	16
39035_7	451	45
39035_8	534	53
39055_1	40	4
39055_2	32	3
39055_3	47	5
39055_4	33	3
39055_67	23	2
39085_1	97	10
39085_2	81	8
39085_3	138	14
39085_4	70	7
39085_56	24	2
39085_8	57	6
39093_1	90	9
39093_2	134	13
39093_3	133	13
39093_4	121	12
39093_57	22	2
39093_6	42	4
39093_8	29	3
39103_1	66	7
39103_2	77	8
39103_3	91	9
39103_45	50	5
39103_68	30	3

The jackknife replicates for each estimator and each jackknife weight are as follows:

$$\bar{y}^{(d)}(s't') = \frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}(s't')y_{sh}^{(d)}}{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}(s't')}$$

The jackknife variance estimators are:

$$v_J\left(\bar{y}^{(d)}\right) = \sum_{s'=1}^{S} \frac{(m_{s'}-1)}{m_{s'}} \sum_{t'=1}^{m_{s'}} \left(\bar{y}^{(d)}(s't') - \bar{y}^{(1)}\right)^2$$

B-2. Jackknife Variance Estimation for Differences

The jackknife variance estimator of the difference $\bar{y}^{(2)} - \bar{y}^{(1)}$ is computed as

$$v_J\left(\bar{y}^{(2)} - \bar{y}^{(1)}\right) = \sum_{s'=1}^{S} \frac{(m_{s'} - 1)}{m_{s'}} \sum_{t'=1}^{m_{s'}} \left(\bar{y}^{(2)}(s't') - \bar{y}^{(1)}(s't')\right)^2$$

The formula is similar for $\bar{y}^{(3)} - \bar{y}^{(1)}$. Doing the calculation in this way will correctly account for the covariance between the mean values. Assuming independence between $\bar{y}^{(1)}$ and $\bar{y}^{(2)}$, $\bar{y}^{(3)}$ would be a serious error.

B-3. Jackknife Variance Estimation for Weighted Model Parameter Estimates

Each of the three weighted model parameter estimates $\hat{\theta}^{(1)}$, $\hat{\theta}^{(2)}$, $\hat{\theta}^{(3)}$ are re-estimated using each set of replicate weights one by one. For the one-day file, this results in replicate-weighted jackknife parameter estimate sets $\hat{\theta}^{(1)}(s't'), s' = 1, \ldots, S, t' = 1, \ldots, m_{s'}$ and jackknife variance estimators:

$$v_J\left(\hat{\theta}^{(1)}\right) = \sum_{s'=1}^{S} \frac{(m_{s'} - 1)}{m_{s'}} \sum_{t'=1}^{m_{s'}} \left(\hat{\theta}^{(1)}(s't') - \hat{\theta}^{(1)}\right)^2$$

For the two-day file, we have replicate-weighted jackknife parameter estimate sets $\hat{\theta}^{(2)}(s't'), s' = 1, \ldots, S, t' = 1, \ldots, m_{s'}$ and jackknife variance estimators:

$$v_J\left(\hat{\theta}^{(2)}\right) = \sum_{s'=1}^{S} \frac{(m_{s'} - 1)}{m_{s'}} \sum_{t'=1}^{m_{s'}} \left(\hat{\theta}^{(2)}(s't') - \hat{\theta}^{(2)}\right)^2$$

For the full file, we have replicate-weighted jackknife parameter estimate sets $\hat{\theta}^{(3)}(s't'), s' = 1, \ldots, S, t' = 1, \ldots, m_{s'}$ and jackknife variance estimators:

$$v_J\left(\hat{\theta}^{(3)}\right) = \sum_{s'=1}^{S} \frac{(m_{s'} - 1)}{m_{s'}} \sum_{t'=1}^{m_{s'}} \left(\hat{\theta}^{(3)}(s't') - \hat{\theta}^{(3)}\right)^2$$

B-4. Jackknife Variance Estimation for Unweighted Model Parameter Estimates

In some cases, an unweighted model parameter estimate is preferred. The weights w_{sh} are replaced by unit weights (all equal to 1). The jackknife replicate weights are all computed in the same way except that the base weight is 1 rather than w_{sh}.

As above, we subscript the replicate groups as $st, s = 1, \ldots, S, t = 1, \ldots, m_s$. Write $S(s)$ as the set of households mapped to stratum s and $S(st)$ as the set of households mapped to replicate group st. The replicate weights are subscripted as $s't'$, with $s' = 1, \ldots, S,\ t' = 1, \ldots, m_s$, then the replicate weights $u_{sh}(s't')$ are defined as follows:

$$u_{sh}(s't') = \begin{cases} 0 & s = s', h \in S(s't') \\ \dfrac{m_s}{(m_s - 1)} & s = s', h \notin S(s't') \\ 1 & s \neq s' \end{cases}$$

B-5. Degrees of Freedom for Jackknife Variance Estimators

As the variance estimators are of direct interest in this study, it is important to generate confidence intervals for the variance estimators. This in turn requires a measure of the variance of the jackknife variance estimators. This section provides an approximate method for computing degrees of freedom, which determine then the variance of the variance estimators. Note that this 'degrees of freedom' terminology arises from the fact that the variance estimator should generally (under certain conditions) follow a Chi-Square distribution.

The following formula from Valliant and Rust (2010) defines the relationship between the degrees of freedom and the variance $Var(v(\hat{\theta}))$ of a variance estimator $v(\hat{\theta})$:

$$DF = \frac{2 * \{Var(\hat{\theta})\}^2}{Var(v(\hat{\theta}))}$$

A general rule of thumb (see for example Valliant and Rust (2010)) for jackknife variance estimators for stratified samples as we have in the 2012 Northeast Ohio Regional Travel Survey is that DF should be equal to the number of replicates R minus the number of strata H. In this case, R is 453 and H is 21, making DF equal to 432. We would expect though that variance estimates for eestimates for domains will have less degrees of freedom, as some of the replicate estimates will just be equal to the full-sample estimate. Thus 432 can be considered an upper bound.

The DF values are used to generate 95% confidence intervals for the variances (standard errors), based on the 0.025 and the 0.975 percentiles of the χ^2 distribution with degrees of freedom equal to the DF value (rounded to an integer). Generally we found that these 95% confidence intervals were very wide when the DF values were less than 30, which happened frequently.

We will begin by developing a degrees of freedom approximation for the jackknife estimator of variance for $\bar{y}^{(d)}$:

$$v_J\left(\bar{y}^{(d)}\right) = \sum_{s'=1}^{S} \frac{(m_{s'}-1)}{m_{s'}} \sum_{t'=1}^{m_{s'}} \left(\bar{y}^{(d)}(s't') - \bar{y}^{(1)}\right)^2$$

We have $\bar{y}^{(d)} = \frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh} y_{sh}^{(d)}}{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}}$, where which can be rewritten as

$$\bar{y}^{(d)} \;=\; \sum_{s=1}^{S} W_s \frac{1}{m_s} \sum_{t=1}^{m_s} w_{st}\bar{y}_{st}^{(d)} \;=\; \sum_{s=1}^{S} W_s \bar{y}_s^{(d)}$$

with

$$w_{st} = m_s \frac{\sum_{h\in S(st)} w_{sh}}{\sum_{t=1}^{m_s}\sum_{h\in S(st)} w_{sh}}$$

$$W_s = \frac{\sum_{t=1}^{m_s}\sum_{h\in S(st)} w_{sh}}{\sum_{s=1}^{S}\sum_{t=1}^{m_s}\sum_{h\in S(st)} w_{sh}}$$

$$\bar{y}_{st}^{(d)} = \frac{\sum_{h\in S(st)} w_{sh} y_{sh}^{(d)}}{\sum_{h\in S(st)} w_{sh}}$$

$$\bar{y}_s^{(d)} = \frac{1}{m_s}\sum_{t=1}^{m_s} w_{st}\bar{y}_{st}^{(d)}$$

Note that $\sum_{t=1}^{m_s} w_{st} = m_s$, $\sum_{s=1}^{S} W_s = 1$. The mean of y is being redefined as a weighted mean of $\bar{y}_{st}^{(d)}$ ‘extended sample unit’ values which correspond to the original sample units assigned to each replicate st.

Rewriting $\bar{y}^{(d)}$ in this way, an approximate variance can be computed for $\bar{y}^{(d)}$ as follows, assuming a with-replacement stratified simple random sample design (see for example Cochran 1977, Eq (5.12)),

$$Var\left(\bar{y}^{(d)}\right) \;=\; \sum_{s=1}^{S} W_s^2 \frac{1}{m_s(m_s-1)} \sum_{t=1}^{m_s} \left(w_{st}\bar{y}_{st}^{(d)} - \bar{y}_s^{(d)}\right)^2 \;=\; \sum_{s=1}^{S} W_s^2 \frac{S_s^2(d)}{m_s}$$

with $S_s^2(d) = \frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left(w_{st}\bar{y}_{st}^{(d)} - \bar{y}_s^{(d)}\right)^2$. Define $S_s^4(d)$ as $\{S_s^2(d)\}^2$. Define

$$S_s^{(4)}(d) = \frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left(w_{st}\bar{y}_{st}^{(d)} - \bar{y}_s^{(d)}\right)^4, \quad \hat{\beta}_s(d) = \frac{S_s^{(4)}(d)}{S_s^4(d)},$$

$$DF\left(\bar{y}^{(d)}\right) = \frac{2\left\{\sum_{s=1}^{S} W_s^2 \frac{S_s^2(d)}{m_s}\right\}^2}{\sum_{s=1}^{S} \frac{W_s^4}{m_s^3} S_s^4(d)(\beta_s(d) - 1)}$$

These definitions follow Valliant and Rust (2010), Equation (7), for a stratified simple random sample.

The $s't'^{\text{th}}$ replicate estimate for $\bar{y}^{(d)}$ can be rewritten as follows:

$$\bar{y}^{(d)}(s't') = \frac{\sum_{s'=1}^{S}\sum_{h=1}^{H} w_{sh}(s't') y_{sh}^{(d)}}{\sum_{s'=1}^{S}\sum_{h=1}^{H} w_{sh}(s't')}$$

We can rewrite $\bar{y}^{(d)}(s't')$ as follows:

$$\bar{y}^{(d)}(s't') \ = \ \sum_{s \neq s'}^{S} W_s \frac{1}{m_s} \sum_{t=1}^{m_s} w_{st} \bar{y}_{st}^{(d)} \ + \ \frac{W_{s'}}{(m_{s'} - 1)} \sum_{t \neq t'} w_{s't} \bar{y}_{s't}^{(d)}$$

In contrast,

$$\bar{y}^{(d)} = \ \sum_{s \neq s'}^{S} W_s \frac{1}{m_s} \sum_{t=1}^{m_s} w_{st} \bar{y}_{st}^{(d)} \ + \ \frac{W_{s'}}{m_{s'}} \sum_{t=1}^{m_s} w_{s't} \bar{y}_{s't}^{(d)}$$

And then,

$$\begin{aligned}
\left(\bar{y}^{(d)}(s't') - \bar{y}^{(d)}\right) \ &= \frac{W_{s'}}{(m_{s'} - 1)} \sum_{t \neq t'} w_{s't} \bar{y}_{s't}^{(d)} - \frac{W_{s'}}{m_{s'}} \sum_{t=1}^{m_s} w_{s't} \bar{y}_{s't}^{(d)} \ = \\
= \ &\frac{W_{s'}}{(m_{s'} - 1)} \left\{ \sum_{t=1}^{m_s} w_{s't} \bar{y}_{s't}^{(d)} - w_{s't'} \bar{y}_{s't'}^{(d)} \right\} - \frac{W_{s'}}{m_{s'}} \sum_{t=1}^{m_s} w_{s't} \bar{y}_{s't}^{(d)} \ = \\
= \ &\frac{W_{s'}}{(m_{s'} - 1)} \left(\bar{y}_s^{(d)} - w_{s't'} \bar{y}_{s't'}^{(d)} \right)
\end{aligned}$$

So that

$$\frac{(m_{s'} - 1)}{W_{s'}} \left(\bar{y}^{(d)}(s't') - \bar{y}^{(d)}\right) = \bar{y}_s^{(d)} - w_{s't'} \bar{y}_{s't'}^{(d)}$$

$$\left\{\frac{(m_{s'} - 1)}{W_{s'}} \left(\bar{y}^{(d)}(s't') - \bar{y}^{(d)}\right)\right\}^2 = \left(w_{s't'} \bar{y}_{s't'}^{(d)} - \bar{y}_s^{(d)}\right)^2$$

$$\left\{\frac{(m_{s'} - 1)}{W_{s'}} \left(\bar{y}^{(d)}(s't') - \bar{y}^{(d)}\right)\right\}^4 = \left(w_{s't'} \bar{y}_{s't'}^{(d)} - \bar{y}_s^{(d)}\right)^4$$

The adjusted squared differences between the jackknife estimators and the full-sample estimator can be used to construct estimators of $S_s^{(4)}(d)$, $\beta_s(d)$, and $DF\left(\bar{y}^{(d)}\right)$. For example, $S_s^{(4)}(d)$ can be calculated as:

$$S_s^{(4)}(d) = \frac{1}{(m_s - 1)} \sum_{t=1}^{m_s} \left\{ \frac{(m_s - 1)}{W_s} \left(\bar{y}^{(d)}(s't') - \bar{y}^{(d)} \right) \right\}^4$$

The estimator $\hat{\beta}_s(d)$ is inherently unstable as it is based on fourth moments, and these are inherently unstable especially with small sample sizes. To allow for this, for continuous variables such as trip distance and trip duration we generated an Empirical Bayes estimator $\tilde{\beta}_s(d)$ by shrinking $\hat{\beta}_s(d)$ back to 3.0 (which is the value of β under normality) as follows:

$$\tilde{\beta}_s(d) = \begin{cases} 3.0 & \hat{\beta}_s(d) < 3.0 \\ \left\{ \frac{30 * 3.0}{30 + m_s} \right\} + \left\{ \frac{m_s * \hat{\beta}_s(d)}{30 + m_s} \right\} & \hat{\beta}_s(d) \geq 3.0 \end{cases}$$

Estimates $\hat{\beta}_s(d)$ less than 3.0 are shrunk back completely to 3.0. Estimates larger than 3.0 are shrunk back to 3.0 by taking a weighted average of 3.0 (the prior mean for $\beta_s(d)$ and $\hat{\beta}_s\ (d)$ (the sample estimate of $\beta_s(d)$, based on a sample size of m_s). The weights in the weighted average of prior and estimate assume the precision of the prior value of 3.0 is equivalent to a sample size of 30, and the precision of the estimated value $\hat{\beta}_s(d)$ is equal to the sample size m_s. For mean values $\bar{y}^{(d)}$ based on dichotomous 0-1 variables, we did no such shrinkage. In this case, the $\hat{\beta}_s(d)$ are stable enough not to require shrinkage.

B-5-2. Degrees of Freedom for Jackknife Variance Estimators for Parameter Estimates

For parameter estimates $\hat{\theta}^{(d)}, d = 1,2,3$ we have a similar jackknife variance estimator:

$$v_J\left(\hat{\theta}^{(d)}\right) = \sum_{s=1}^{S} \frac{(m_s - 1)}{m_s} \sum_{t=1}^{m_s} \left(\hat{\theta}^{(d)}(st) - \hat{\theta}^{(d)} \right)^2$$

We compute a degrees of freedom for parameter estimates $\hat{\theta}^{(d)}$ as follows:

$$DF\left(\hat{\theta}^{(d)}\right) = \frac{2 * \left\{ Var\left(\hat{\theta}^{(d)}\right) \right\}^2}{Var\left(v(\hat{\theta}^{(d)})\right)}$$

This is estimated by:

$$DF\left(\hat{\theta}^{(d)}\right) = \frac{2 * \left\{v_J\left(\hat{\theta}^{(d)}\right)\right\}^2}{Var\left(v_J\left(\hat{\theta}^{(d)}\right)\right)}$$

We estimate $Var\left(v_J\left(\hat{\theta}^{(d)}\right)\right)$ as

$$Var\left(v_J\left(\hat{\theta}^{(d)}\right)\right) = \sum_{s=1}^{S} \frac{W_s^4}{m_s^3} S_s^4\left(\hat{\theta}^{(d)}\right)\left(\beta_s\left(\hat{\theta}^{(d)}\right) - 1\right)$$

with

$$S_s^2\left(\hat{\theta}^{(d)}\right) = \frac{1}{(m_s - 1)} \sum_{t=1}^{m_s} \left\{\frac{(m_s - 1)}{W_s}\left(\hat{\theta}^{(d)}(st) - \hat{\theta}^{(d)}\right)\right\}^2, \quad S_s^4\left(\hat{\theta}^{(d)}\right) = \left\{S_s^2\left(\hat{\theta}^{(d)}\right)\right\}^2$$

$$S_s^{(4)}\left(\hat{\theta}^{(d)}\right) = \frac{1}{(m_s - 1)} \sum_{t=1}^{m_s} \left\{\frac{(m_s - 1)}{W_s}\left(\hat{\theta}^{(d)}(st) - \hat{\theta}^{(d)}\right)\right\}^4, \quad \beta_s\left(\hat{\theta}^{(d)}\right) = \frac{S_s^{(4)}\left(\hat{\theta}^{(d)}\right)}{S_s^4\left(\hat{\theta}^{(d)}\right)}$$

This can be justified if $\hat{\theta}^{(d)}$ is a smooth function of a set of stratified means coming from the sample. Suppose for example that

$$\hat{\theta}^{(d)} = g\left(\bar{u}_1^{(d)}, \dots, \bar{u}_c^{(d)}, \dots, \bar{u}_C^{(d)}\right) = g\left(\bar{\mathbf{u}}^{(\mathbf{d})}\right), \; d = 1,2,3$$

Each $\bar{u}_c^{(d)}$ is

$$\bar{u}_c^{(d)} = \frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh} u_{sh,c}^{(d)}}{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}} \quad \text{with } u_{sh,c}^{(1)} = u_{sh,c}, \; u_{sh,c}^{(2)} = \frac{u_{sh,c}^{(1)} + u_{sh,c}^{(2)}}{2}, u_{sh,c}^{(3)} = \frac{\sum_{d=1}^{D_{sh}} u_{sh,c}^{(d)}}{D_{sh}}$$

where $g(\,)$ is a C by 1 continuous vector function of a vector of stratified weighted means based on the sample, and the $\bar{u}_c^{(d)}$ are stratified mean values which may include cross-products. For example, suppose we have a parameter vector estimate $\hat{\boldsymbol{\beta}}^{(d)} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\left(\mathbf{X}'\mathbf{W}\bar{\mathbf{y}}^{(d)}\right)$, with K elements k =1,…, K ($\mathbf{X}$ is an n by K matrix, and $\bar{\mathbf{y}}^{(d)}$ is an n by 1 vector, where n is the sample size). Each element of this parameter vector will be a smooth function of cross-product stratified mean values, with the (k_1, k_2) element of $\mathbf{X}'\mathbf{W}\mathbf{X}$ being

$$\frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh} x_{k_1,sh} x_{k_2,sh}}{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}}$$

and the k_1 element of $\mathbf{X}'\mathbf{W}\bar{\mathbf{y}}^{(d)}$ being

$$\frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh} x_{k_1,sh} y_{sh}^{(d)}}{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}}$$

There are a total of $C = 2 * K + \frac{K*(K-1)}{2}$ cross-products comprising the argument of the g-function of the parameter vector $\hat{\boldsymbol{\beta}}^{(1)}$ (the $g(\,)$ function in this case inverts the matrix of $\frac{K*(K+1)}{2}$ cross-products of $\mathbf{X'WX}$, and pre-multiplies this to the K -vector of cross-products $\mathbf{X'W\bar{y}}^{(d)}$: this is a smooth continuous, differentiable function (though not linear)).

Suppose $g\left(\bar{u}_1^{(d)}, \dots, \bar{u}_c^{(d)}, \dots, \bar{u}_C^{(d)}\right)$ has continuous partial derivatives $\frac{\partial \mathbf{g}}{\partial \mathbf{u}} = \left\{\frac{\partial g}{\partial u_1}, \dots . \frac{\partial g}{\partial u_C}\right\}$ in a neighborhood of the expected value $\mathbf{U}^{(d)} = E\left(\bar{\mathbf{u}}^{(\mathbf{d})}\right)$. Under appropriate regularity conditions on the moments of the distribution of $\bar{\mathbf{u}}^{(\mathbf{d})}$, we can approximate the variance of $\hat{\theta}^{(d)}$ as follows:

$$Var\left(\hat{\theta}^{(d)}\right) = \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T Var\left(\bar{\mathbf{u}}^{(\mathbf{d})}\right)\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\} + \quad o(n^{-1})$$

where the remainder term is a term of lower order (small compared to the first term as n gets large: note that the first term is O(n^{-1}) under appropriate assumptions on the relevant components[12]). See for example Wolter (2007) , Section 6.2.

The variance estimator then for $\hat{\theta}^{(d)}$ based on this approximation is:

$$v_{TS}\left(\hat{\theta}^{(d)}\right) = \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \mathbf{VC}\left(\bar{\mathbf{u}}^{(\mathbf{d})}\right)\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}$$

The C by C matrix $\mathbf{VC}\left(\bar{\mathbf{u}}^{(\mathbf{d})}\right)$ has as its (c_1, c_1) diagonal element

$$\left\{\mathbf{VC}\left(\bar{\mathbf{u}}^{(\mathbf{d})}\right)\right\}_{c_1,c_1} = \sum_{s=1}^{S} W_s^2 \frac{S_s^2(u_{c_1}^{(d)})}{m_s}, \quad \text{with } S_s^2\left(u_{c_1}^{(d)}\right) = \frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left(w_{st}\bar{u}_{c_1,st}^{(d)} - \bar{u}_{c_1,s}^{(d)}\right)^2$$

and $\bar{u}_{c_1,s}^{(d)} = \frac{1}{m_s}\sum_{t=1}^{m_s} w_{st}\bar{u}_{c_1,st}^{(d)}$. Write $\bar{U}_{c_1,s}^{(d)} = E\left(\bar{u}_{c_1,s}^{(d)}\right)$.

The off-diagonal (c_1, c_2) element of $\mathbf{VC}\left(\bar{\mathbf{u}}^{(\mathbf{d})}\right)$ is:

$$\left\{\mathbf{VC}\left(\bar{\mathbf{u}}^{(\mathbf{d})}\right)\right\}_{c_1,c_2} = \sum_{s=1}^{S} W_s^2 \frac{S_s^2\left(u_{c_1}^{(d)}, u_{c_2}^{(d)}\right)}{m_s},$$

$$\text{with } S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)}) = \frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left(w_{st}\bar{u}_{c_1,st}^{(d)} - \bar{u}_{c_1,s}^{(d)}\right)\left(w_{st}\bar{u}_{c_2,st}^{(d)} - \bar{u}_{c_2,s}^{(d)}\right)$$

Note that $\bar{u}_{c_2,s}^{(d)} = \frac{1}{m_s}\sum_{t=1}^{m_s} w_{st}\bar{u}_{c_2,st}^{(d)}$ and $\bar{U}_{c_2,s}^{(d)} = E\left(\bar{u}_{c_2,s}\right)$. Define a C by C matrix $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}^{(\boldsymbol{d})}$ with (c_1, c_1) diagonal element $S_s^2\left(u_{c_1}^{(d)}\right)$ and off-diagonal (c_1, c_2) element $S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)})$. Then

[12] In other words, n times the first term remains bounded below by a constant strictly larger than 0, and remains bounded above by another larger constant as n tends to infinity.

$$\mathbf{VC}\left(\overline{\mathbf{u}}^{(\mathbf{d})}\right) = \sum_{s=1}^{S} W_s^2 \frac{\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}}{m_s}$$

and

$$v_{TS}\left(\hat{\theta}^{(d)}\right) = \sum_{s=1}^{S} W_s^2 \frac{\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})} \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}}{m_s}$$

This is an estimator of the first term of the Taylor Series approximation of the true variance of $\hat{\theta}^{(d)}$:

$$Var_{TS}\left(\hat{\theta}^{(d)}\right) = \sum_{s=1}^{S} W_s^2 \frac{\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \boldsymbol{\Sigma}_{\boldsymbol{us}}^{(\boldsymbol{d})} \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}}{m_s}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{us}}^{(\boldsymbol{d})}$ is the variance-covariance matrix within stratum $\boldsymbol{s}$ of the $\boldsymbol{u}$-vector.

To compute the variance of $v_{TS}\left(\hat{\theta}^{(d)}\right)$ as an estimator of $Var_{TS}\left(\hat{\theta}^{(d)}\right)$, we use the fact of independence across strata: each term in $v_{TS}\left(\hat{\theta}^{(d)}\right)$ as a random variable is independent of each other term. Also, W_s^2, m_s, and $\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}$ are constants. Thus,

$$Var\left\{v_{TS}\left(\hat{\theta}^{(d)}\right)\right\} = \sum_{s=1}^{S} W_s^4 \frac{Var\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})} \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\right)}{m_s^2}$$

Following for example Harville (1997), Section 16.2, we can define the vec operator of an m_s x m_s matrix $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}$ as an m_s^2 x 1 long column vector with the columns of $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}$ lined up:

$$vec\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right) = vec\left(\begin{pmatrix} s_{11} & \cdots & s_{1m_s} \\ \vdots & \ddots & \vdots \\ s_{m_s1} & \cdots & s_{m_sm_s} \end{pmatrix}\right) =$$
$$= \left(s_{11}, \ldots\ldots, s_{m_s1}, \ldots\ldots, \ldots\ldots\ldots, s_{1m_s}, \ldots\ldots, s_{m_sm_s}\right)^T$$

From Theorem 16.2.1 in Harville (1997), we can compute the following (where ⊗ is a Kronecker product (see also Harville (1997)):

$$\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}^T \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})} \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\} \quad = \quad \left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}^T \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}^T\right) vec\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right)$$

Note that $\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}^T \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})} \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}$ is a scalar (a quadratic form), $\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}^T \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}^T\right)$ is a 1x m_s^2 vector with the cross-products of the partial derivative elements of the $\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}$ vector and

$vec\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right)$ is an m_s^2 x 1 vector with the elements of $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}$ lined up. Some of these elements (m_s elements in all) are the sample variances $S_s^2\left(u_{c_1}^{(d)}\right)$, and rest ($m_s * (m_s - 1)$ elements in all) are the sample covariances $S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)})$.

Thus we can write:

$$Var\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})} \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\right) =$$
$$\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\right) \left\{Var\left(vec\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right)\right)\right\} \left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\} \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\right)$$

Note that $\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\right)$ is a 1x m_s^2 vector of constants, $\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\} \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\right)$ is a m_s^2 x 1 of constants and $Var\left(vec\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right)\right)$ is a m_s^2 x m_s^2 variance-covariance matrix with the variances of the sample variances $S_s^2\left(u_{c_1}^{(d)}\right)$ and sample covariances $S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)})$ along the diagonal, and the covariances of the sample variances and sample covariances as the off-diagonal elements.

To compute the variances of the $S_s^2\left(u_{c_1}^{(d)}\right)$ and the $S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)})$ (the diagonal elements of $Var\left(vec\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right)\right)$), we simplify the notation to better match the development of this in Hansen, Hurwitz, and Madow (1953) (pp. 99-101) (called "HHM" below). Define $z_{1st} = w_{st}\bar{u}_{c_1,st}^{(d)} - \bar{U}_{c_1,s}^{(d)}$ and $\bar{z}_{1s} = \bar{u}_{c_1,s}^{(d)} - \bar{U}_{c_1,s}^{(d)}$. Define $z_{2st} = w_{st}\bar{u}_{c_2,st}^{(d)} - \bar{U}_{c_2,s}^{(d)}$ and $\bar{z}_{2s} = \bar{u}_{c_2,s}^{(d)} - \bar{U}_{c_2,s}^{(d)}$. Note that $E\{z_{1st}\} = E\{\bar{z}_{1s}\} = E\{z_{2st}\} = E\{\bar{z}_{2s}\} = 0$. Then

$$S_s^2\left(u_{c_1}^{(d)}\right) = \frac{1}{(m_s-1)}\sum_{t=1}^{m_s}(z_{1st} - \bar{z}_{1s})^2, \quad S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)}) = \frac{1}{(m_s-1)}\sum_{t=1}^{m_s}(z_{1st} - \bar{z}_{1s})(z_{2st} - \bar{z}_{2s})$$

Assume $m_s \to \infty$ and $\bar{z}_{1s} = O_p\left({m_s}^{-\frac{1}{2}}\right)$, $\bar{z}_{2s} = O_p\left({m_s}^{-\frac{1}{2}}\right)$.

Following HHM,

$$E\left\{S_s^2\left(u_{c_1}^{(d)}\right)\right\}^2 = \frac{1}{(m_s-1)^2}E\left\{\left(\sum_{t=1}^{m_s} z_{1st}^2\right)^2 - 2m_s\{\bar{z}_{1s}\}^2\sum_{t=1}^{m_s} z_{1st}^2 + m_s^2\{\bar{z}_{1s}\}^4\right\}$$

Following HHM, under appropriate regularity conditions, the second and third terms of this expansion are of lower order in m_s as $m_s \to \infty$, so that we have

$$E\left\{S_s^2\left(u_{c_1}^{(d)}\right)\right\}^2 = \frac{1}{(m_s-1)^2}E\left\{\left(\sum_{t=1}^{m_s} z_{1st}^2\right)^2\right\} + o({m_s}^{-1}) \ldots$$

We have

$$\left(\sum_{t=1}^{m_s} z_{1st}^2\right)^2 = \sum_{t=1}^{m_s} z_{1st}^4 + \sum_{t\neq t'}^{m_s} z_{1st}^2 z_{1st'}^2$$

Write $E(z_{1st}^4) = \mu_4(1s)$, $E(z_{1st}^2) = \sigma^2(1s)$, and assume the z_{1st} are independent. Then

$$E\left\{S_s^2\left(u_{c_1}^{(d)}\right)\right\}^2 = \frac{1}{(m_s-1)^2} E\{m_s\mu_4(1s) + m_s(m_s-1)\sigma^4(1s)\} + o(m_s{}^{-1})$$
$$= \frac{\mu_4(1s)}{m_s} + \sigma^4(1s) + \ o(m_s{}^{-1})$$

So

$$Var\{S_s^2(u_{c_1}^{(d)})\} = E\{S_s^2(u_{c_1})\}^2 - \left\{E\left(S_s^2(u_{c_1})\right)\right\}^2 = \ E\{S_s^2(u_{c_1})\}^2 - \sigma^4(1s)$$
$$= \frac{\mu_4(1s)}{m_s} + o(m_s{}^{-1})$$

For $S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)})$ we have

$$E\left\{S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)})\ \right\}^2 = \frac{1}{(m_s-1)^2} E\left\{\left(\sum_{t=1}^{m_s} z_{1st}^2\right)^2 - 2m_s\{\bar{z}_{1s}\}^2 \sum_{t=1}^{m_s} z_{1st}^2 + m_s^2\{\bar{z}_{1s}\}^4\right\}$$

As above, under appropriate regularity conditions, the second and third terms of this expansion are of lower order in m_s as $m_s \to \infty$, so that we have

$$E\left\{S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)})\ \right\}^2 = \frac{1}{(m_s-1)^2} E\left\{\left(\sum_{t=1}^{m_s} z_{1st} z_{2st}\right)^2\right\} \ + o(m_s{}^{-1}) \ \ldots$$

We have

$$\left(\sum_{t=1}^{m_s} z_{1st} z_{2st}\right)^2 = \sum_{t=1}^{m_s} z_{1st}^2 z_{2st}^2 + \sum_{t\neq t'}^{m_s} z_{1st} z_{2st} z_{1st'} z_{2st'}$$

Write $E(z_{1st}^2 z_{2st}^2) = \mu_4(12s)$, $E(z_{1st} z_{2st}) = \sigma^2(12s)$, and assume the z_{1st} are independent (across t). Then

$$E\left\{S_s^2(u_{c_1}^{(d)}, u_{c_2}^{(d)})\ \right\}^2 = \frac{1}{(m_s-1)^2} E\{m_s\mu_4(12s) + m_s(m_s-1)\sigma^4(12s)\} + o(m_s{}^{-1})$$
$$= \frac{\mu_4(12s)}{m_s} + \sigma^4(12s) + \ o(m_s{}^{-1})$$

So

$$Var\{S_s^2(u_{c_1}^{(d)},u_{c_2}^{(d)})\} = E\{S_s^2(u_{c_1}^{(d)},u_{c_2}^{(d)})\}^2 - \{S_s^2(u_{c_1}^{(d)},u_{c_2}^{(d)})\}^2 = E\{S_s^2(u_{c_1}^{(d)},u_{c_2}^{(d)})\}^2 - \sigma^4(12s)$$
$$= \frac{\mu_4(12s)}{m_s} + o(m_s{}^{-1})$$

The corresponding covariances between the sample variances and covariances can be shown to have a similar form. We can summarize all of this as:

$$\left\{Var\left(vec\left(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right)\right)\right\} = \frac{1}{m_s}\mathbf{M}_{\boldsymbol{us}}^{(\boldsymbol{d})} + o(m_s{}^{-1})$$

where $\mathbf{M}_{\boldsymbol{us}}^{(\boldsymbol{d})}$ contains the mixed fourth moments $\mu_4(1s)$, $\mu_4(12s)$, in the main diagonal, and corresponding covariance mixed fourth moments in the off-diagonal elements.

$$Var\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}^T \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{us}}^{(\boldsymbol{d})} \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}\right) \approx \left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}^T \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}^T\right)\left\{\frac{1}{m_s}\mathbf{M}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right\}\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\} \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}\right)$$

and

$$Var\{v_{TS}(\hat{\theta}^{(d)})\} \approx \sum_{s=1}^{S} W_s^4 \frac{\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}^T \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}^T\right)\left\{\mathbf{M}_{\boldsymbol{us}}^{(\boldsymbol{d})}\right\}\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\} \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}\right)}{m_s^3}$$

Under regularity conditions, the jackknife pseudo-value can also be expanded accordingly. For jackknife parameter estimates $\hat{\theta}^{(d)}(s't'), s' = 1, \dots, S, t' = 1, \dots, m_{s'}$, we have

$$\hat{\theta}^{(d)}(s't') - \hat{\theta}^{(1)} = \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(\boldsymbol{d})}}\right\}^T \{\bar{\mathbf{u}}^{(\mathbf{d})}(s't') - \bar{\mathbf{u}}^{(\mathbf{d})}\}$$

where $\bar{\mathbf{u}}^{(\mathbf{d})}$ consists of the C elements

$$\bar{u}_c^{(d)} = \frac{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh} u_{c,sh}^{(d)}}{\sum_{s=1}^{S}\sum_{h=1}^{H} w_{sh}}, \quad c = 1, \dots, C$$

and $\bar{\mathbf{u}}^{(\mathbf{d})}(s't')$ consists of the C elements (replacing w_{sh} with $w_{sh}(s't')$):

$$\bar{u}_c^{(d)}(s^{'}t^{'}) = \frac{\sum_{s^{'}=1}^{S}\sum_{h=1}^{H} w_{sh}(s^{'}t^{'}) u_{sh}^{(d)}}{\sum_{s^{'}=1}^{S}\sum_{h=1}^{H} w_{sh}(s^{'}t^{'})}, \quad c = 1, \dots, C$$

We can decompose $\bar{u}_c^{(d)}(s^{'}t^{'})$ as we did for $\bar{y}^{(d)}(s't')$:

$$\bar{u}_c^{(d)}(s^{'}t^{'}) = \sum_{s \neq s^{'}}^{S} W_s \frac{1}{m_s}\sum_{t=1}^{m_s} w_{st}\bar{u}_{c,st}^{(d)} + \frac{W_{s^{'}}}{(m_{s^{'}}-1)}\sum_{t \neq t^{'}} w_{s^{'}t}\bar{u}_{c,s^{'}t}^{(d)} \qquad c = 1, \dots, C$$

with

$$\bar{u}_{c,st}^{(d)} = \frac{\sum_{h\in S(st)} w_{sh} u_{sh,c}^{(d)}}{\sum_{h\in S(st)} w_{sh}} \qquad \bar{u}_{c,s}^{(d)} = \frac{1}{m_s}\sum_{t=1}^{m_s} w_{st}\bar{u}_{c,st}^{(d)}$$

This is all justifiable as the $u_{sh,c}^{(d)}$ values are just a particular y-value (a cross-product in this case) that are aggregated up using the original w_{sh} or the replicate $w_{sh}(s't')$ weights as any other y-values.

Thus the C-vector $\bar{\mathbf{u}}^{(\mathbf{d})}(s't') - \bar{\mathbf{u}}^{(\mathbf{d})}$ can be written as C elements

$$\left(\bar{u}_c^{(d)}(s't') - \bar{u}_c^{(d)}\right) = \frac{W_{s'}}{(m_{s'}-1)}\sum_{t\neq t'} w_{s't}\bar{u}_{c,s't}^{(d)} - \frac{W_{s'}}{m_{s'}}\sum_{t=1}^{m_s} w_{s't}\bar{u}_{c,s't}^{(d)} =$$

$$= \frac{W_{s'}}{(m_{s'}-1)}\left\{\sum_{t=1}^{m_s} w_{s't}\bar{u}_{c,s't}^{(d)} - w_{s't'}\bar{u}_{c,s't}^{(d)}\right\} - \frac{W_{s'}}{m_{s'}}\sum_{t=1}^{m_s} w_{s't}\bar{u}_{c,s't}^{(d)} =$$

$$= \frac{W_{s'}}{(m_{s'}-1)}\left(\bar{u}_{c,s'}^{(d)} - w_{s't'}\bar{u}_{c,s't'}^{(d)}\right) \qquad c = 1,\dots,C$$

And

$$\hat{\theta}^{(d)}(s't') - \hat{\theta}^{(d)} = \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \left\{\bar{\mathbf{u}}^{(\mathbf{d})}(s't') - \bar{\mathbf{u}}^{(\mathbf{d})}\right\} = \sum_{c=1}^{C}\left\{\frac{\partial g}{\partial u_c}|_{\mathbf{U}^{(d)}}\right\}\left(\bar{u}_c^{(d)}(s't') - \bar{u}_c^{(d)}\right) =$$

$$= \sum_{c=1}^{C}\left\{\frac{\partial g}{\partial u_c}|_{\mathbf{U}^{(d)}}\right\}\frac{W_{s'}}{(m_{s'}-1)}\left(\bar{u}_{c,s'}^{(d)} - w_{s't'}\bar{u}_{c,s't'}^{(d)}\right) =$$

$$= \frac{W_{s'}}{(m_{s'}-1)}\sum_{c=1}^{C}\left\{\frac{\partial g}{\partial u_c}|_{\mathbf{U}^{(d)}}\right\}\left(\bar{u}_{c,s'}^{(d)} - w_{s't'}\bar{u}_{c,s't'}^{(d)}\right)$$

So that

$$\frac{(m_{s'}-1)}{W_{s'}}\left(\hat{\theta}^{(d)}(s't') - \hat{\theta}^{(d)}\right) = \sum_{c=1}^{C}\left\{\frac{\partial g}{\partial u_c}|_{\mathbf{U}^{(d)}}\right\}\left(\bar{u}_{c,s'}^{(d)} - w_{s't'}\bar{u}_{c,s't'}^{(d)}\right) = -\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T \left\{\Delta\mathbf{u}_{st}^{(d)}\right\}$$

where $\Delta\mathbf{u}_{\mathbf{st}}$ is a C-vector with elements $\left\{w_{st}\bar{u}_{c,st}^{(d)} - \bar{u}_{c,s}^{(d)}\right\}_{c=1,\dots,C}$.

We have

$$\left(\frac{(m_{s'}-1)}{W_{s'}}\left(\hat{\theta}^{(d)}(s't') - \hat{\theta}^{(d)}\right)\right)^2 = \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}^T \left\{\Delta\mathbf{u}_{st}^{(d)}\right\}\left\{\Delta\mathbf{u}_{st}^{(d)}\right\}^T \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_U\right\}$$

and

$$\left(\frac{(m_{s'}-1)}{W_{s'}}\left(\hat{\theta}^{(d)}(s't')-\hat{\theta}^{(d)}\right)\right)^4$$
$$=\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}$$

Repeatedly using the identity $vec(\mathbf{ABC})=(\mathbf{C}'\otimes\boldsymbol{A})vec(\mathbf{B})$ (Theorem 16.2.1 from Harville (1997)), we can rearrange (note that for a scalar a, $vec(a)=a$):

$$\left(\frac{(m_{s'}-1)}{W_{s'}}\left(\hat{\theta}^{(d)}(s't')-\hat{\theta}^{(d)}\right)\right)^4$$
$$=\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\otimes\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\right)vec\left\{\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\right\}=$$
$$=\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\otimes\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\right)\left(\left\{\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\right\}^T\otimes\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\right)vec\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T$$

Now $\left\{\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\right\}^T=\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T$ and $vec\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T=\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\otimes\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\right)$ so that

$$\left(\frac{(m_{s'}-1)}{W_{s'}}\left(\hat{\theta}^{(d)}(s't')-\hat{\theta}^{(d)}\right)\right)^4=$$
$$=\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\otimes\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\right)\left(\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\otimes\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\right)\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\otimes\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\right)$$

So that

$$S_s^{(4)}\left(\hat{\theta}^{(d)}\right)=\frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left\{\frac{(m_s-1)}{W_s}\left(\hat{\theta}^{(d)}(st)-\hat{\theta}^{(d)}\right)\right\}^4=$$
$$\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\otimes\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}^T\right)\left\{\frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left(\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\otimes\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\right)\right\}\left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\otimes\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}|_{\mathbf{U}^{(d)}}\right\}\right)$$

The (1,1) element of the matrix $\frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left(\{\Delta\mathbf{u_{st}}\}\{\Delta\mathbf{u_{st}}\}^T\otimes\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\right)$ is for example $\frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left(w_{st}\bar{u}_{1,st}-\bar{u}_{1,s}\right)^4$, which has as its expectation the (1,1) element of $\mathbf{M}_{\boldsymbol{us}}^{(\boldsymbol{d})}$. Thus we can claim:

$$E\left\{\frac{1}{(m_s-1)}\sum_{t=1}^{m_s}\left(\{\Delta\mathbf{u_{st}}\}\{\Delta\mathbf{u_{st}}\}^T\otimes\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}\left\{\boldsymbol{\Delta}\mathbf{u}_{st}^{(d)}\right\}^T\right)\right\}=\mathbf{M}_{\boldsymbol{us}}^{(\boldsymbol{d})}$$

and finally

$$S_s^{(4)}\left(\hat{\theta}^{(d)}\right) = \left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}\Big|_{\mathbf{U}^{(d)}}\right\}^T \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}\Big|_{\mathbf{U}^{(d)}}\right\}^T\right) \left\{\mathbf{M}_{us}^{(d)}\right\} \left(\left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}\Big|_{\mathbf{U}^{(d)}}\right\} \otimes \left\{\frac{\partial \mathbf{g}}{\partial \mathbf{u}}\Big|_{\mathbf{U}^{(d)}}\right\}\right)$$

References

Cochran, W. G. (1977). *Sampling Techniques, 3rd Ed.* New York: John Wiley & Sons.

Hansen, M. H., Hurwitz, W. N., and Madow, W. G. (1953). *Sample Survey Methods and Theory Volume II*. New York: John Wiley & Sons.

Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective.* New York: Springer (www.springer.com)

Shao, J., and Tu, D. (1995). *The Jackknife and Bootstrap.* New York: Springer (www.springer.com)

Valliant, R., Dever, J. A., and Kreuter, F. (2013). *Practical Tools for Designing and Weighting Survey Samples.* New York: Springer (www.springer.com).

Valliant, R. and Rust, K. F. (2010). Degrees of freedom approximations and rules-of-thumb. *Journal of Official Statistics 26 (4)*, 585-602.

Wolter, K. M. (2007). *Introduction to Variance Estimation, 2nd ed.* New York: Springer (www.springer.com).

Appendix C. Results Comparing GPS-Only and GPS-With-Prompted-Recall Data

There were 2,775 households which had GPS tracking only, and 1,312 household which had GPS with prompted recall. The two sets were generated by randomized assignment, so there is no reason why the two sets of households should differ in their estimates or measured outcomes in any sense. Any significant difference must in fact be an artifact of data collection differences, as well as possibly data cleaning and processing differences. For example, the GPS-only data replaces the recall data with imputation. This imputation process is very difficult and is subject to considerable error.

Table C-1 presents the estimated mean number of trips per person for the two GPS strata with the jackknife standard errors for these estimates and 95% confidence intervals. The absolute value of the t-statistic for the difference is 1.8, which has a p-value of 0.073 for a two-sided test of the null hypothesis of no difference (with an alpha value of 0.05). The t-statistic for the difference was computed using a jackknife standard error for the difference per se, following the theory as given in Appendix B-2. Thus we register a marginally significant difference. The magnitude of this difference (4.23 vs. 4.03) is not very large.

Table C-1. Weighted mean number of trips per person comparing the GPS-with-prompted-recall households with GPS-only households.

GPS stratum	Mean trips per person	Std Err	Lower bound 95% CI	Upper bound 95% CI
GPS with prompted recall	4.231	0.090	4.051	4.411
GPS only	4.026	0.078	3.871	4.181

Table C-2 presents the estimated mean number of trips per sampled person for the two GPS strata restricted to domains defined by designated trip purpose. Note that in most of the domains there are large numbers of zero-trip persons, which pulls the mean value towards 0. We would expect measurement differences between the GPS only data and the GPS with prompted recall data, as in the GPS with prompted recall case the designated trip purpose comes from a recall interview, and in the GPS only case the designated trip purpose is imputed.

The table includes the mean values for the domain and GPS stratum, the jackknife standard errors, the difference between GPS strata for the domain mean values, the t-statistic for the difference, and the p-value for the difference[13]. The rows are ordered by the sign and magnitude of the difference GPS-only mean minus GPS prompted recall mean. As can be seen there was no significant difference between the GPS strata for the home-based university and home-based other domains. The home-based shopping, work, and school domains showed significantly higher

[13] The t-statistic is the difference divided by the jackknife standard error of the difference. The p-value is the two-sided p-value for the null hypothesis of zero difference.

mean estimated trips per person for the GPS prompted recall stratum. Both non-home-based domains and the home-based social/recreation domain showed significantly higher mean estimated trips per person for the GPS only stratum. There is no reason why the difference should be non-zero except due to imputation error, and the differences likely reflect imputation error.

Table C-2. Weighted mean number of trips per person comparing the GPS-with-prompted-recall households with GPS-only households, by trip purpose domain.

Trip purpose domain	Gflag Stratum	Estimated trips per person	Jackknife standard error	Difference	T-stat for difference	P-value for difference
Home-Based Other	GPS PrmpRec	1.095	0.039			
Home-Based Other	GPS Only	0.980	0.025	0.1146	2.49	0.0132
Home-Based School	GPS PrmpRec	0.214	0.021			
Home-Based School	GPS Only	0.133	0.009	0.0809	3.38	0.0008
Home-Based Shopping	GPS PrmpRec	0.412	0.023			
Home-Based Shopping	GPS Only	0.233	0.010	0.1799	7.08	< 0.0001
Home-Based Social/Recr	GPS PrmpRec	0.587	0.035			
Home-Based Social/Recr	GPS Only	0.697	0.021	-0.1107	-2.72	0.0067
Home-Based University	GPS PrmpRec	0.012	0.003			
Home-Based University	GPS Only	0.008	0.002	0.0042	1.11	0.2684
Home-Based Work	GPS PrmpRec	0.328	0.016			
Home-Based Work	GPS Only	0.221	0.009	0.1070	5.52	<0.0001
Non-Home-Based Other	GPS PrmpRec	1.429	0.056			
Non-Home-Based Other	GPS Only	1.542	0.049	-0.1130	-1.57	0.1180
Non-Home-Based Work	GPS PrmpRec	0.154	0.011			
Non-Home-Based Work	GPS Only	0.212	0.010	-0.0577	-3.94	<0.0001

Table C-3 (in eight parts) presents the difference between the GPS-Prompted-Recall and the GPS-Only data sets for mean number of tours by tour type for eight person types. As above, there should be no significant difference between these two data sets for mean number of tours for any of tour type or person type, as assignment to these data sets is random. Any differences are an artifact of data collection. The results in the eight tables can be summarized as follows:

- For full-time and part-time workers both, there are somewhat more tours per person/day for GPS-Prompted-Recall than for GPS-Only, but considerably more work tours per person/day. There are considerably fewer tours of other kinds per person/day among GPS-Prompted-Recall (compared with GPS-Only)..
- For university students, there is no significant difference in total tours per person-day between the two sets. There are significantly less social/recreational tours per person/day for GPS-Prompted-Recall, with more tours of other kinds (as compared to GPS-Only).
- For GPS Prompted-Recall opposed to GPS-Only for non-workers and retirees, there are significantly more trips per person/day, divided up between the three possible categories.

- For driving-age children, there were no significant differences between the two sets.
- For pre-driving age children and pre-school children, there were no significant differences in total trips per person/day between the two sets, but in both cases there were more social/recreational trips per person/day for the GPS-Only group (with fewer school trips and shopping trips for pre-driving age children).

Table C-3, Part 1. Average number of tours per person comparing the GPS-with-prompted-recall with GPS-only households for full-time workers, by tour type.

Tour Type	GPS Stratum	Average Number of Tours Per Person[1]	Jackknife Standard Error	Difference with GPS-with-prompted-recall	T-stat for difference with GPS-with-prompted-recall	P-value for No difference with GPS-with-prompted-recall
Work	GPS PrmpRec	1.309	0.038			
Work	GPS-Only	0.470	0.015	-0.8384	-20.64	<0.0001
Shop	GPS PrmpRec	0.028	0.007			
Shop	GPS-Only	0.113	0.009	0.0849	7.85	<0.0001
Social/Recreational	GPS PrmpRec	0.011	0.004			
Social/Recreational	GPS-Only	0.331	0.020	0.3205	16.00	<0.0001
Other	GPS PrmpRec	0.049	0.013			
Other	GPS-Only	0.221	0.014	0.1711	9.80	<0.0001
Work-Based Subtour	GPS PrmpRec	0.076	0.010			
Work-Based Subtour	GPS-Only	0.074	0.007	-0.0020	-0.16	0.8710
Total Tours	GPS PrmpRec	1.472	0.038			
Total Tours	GPS-Only	1.208	0.027	-0.2639	-6.07	<0.0001

Table C-3, Part 2. Average number of tours per person comparing the GPS-with-prompted-recall with GPS-only households for part-time workers, by tour type.

Tour Type	GPS Stratum	Average Number of Tours Per Person[1]	Jackknife Standard Error	Difference with GPS-with-prompted-recall	T-stat for Difference with GPS-with-prompted-recall	P-value for No Difference with GPS-with-prompted-recall
Work	GPS PrmpRec	1.336	0.070			
Work	GPS-Only	0.271	0.027	-1.0646	-14.27	<0.0001
Shop	GPS PrmpRec	0.026	0.010			
Shop	GPS-Only	0.279	0.029	0.2534	8.13	<0.0001
Social/Recreational	GPS PrmpRec	0.009	0.007			
Social/Recreational	GPS-Only	0.402	0.036	0.3925	10.64	<0.0001
Other	GPS PrmpRec	0.046	0.016			
Other	GPS-Only	0.354	0.034	0.3075	8.16	<0.0001
Work-Based Subtour	GPS PrmpRec	0.022	0.011			
Work-Based Subtour	GPS-Only	0.012	0.006	-0.0100	-0.83	0.4086
Total Tours	GPS PrmpRec	1.439	0.074			
Total Tours	GPS-Only	1.318	0.064	-0.1212	-1.20	0.2299

Table C-3, Part 3. Average number of tours per person comparing the GPS-with-prompted-recall with GPS-only households for university students, by tour type.

Tour Type	GPS Stratum	Average Number of Tours Per Person	Jackknife Standard Error	Difference with GPS-with-prompted-recall	T-stat for difference with GPS-with-prompted-recall	P-value for No difference with GPS-with-prompted-recall
School	GPS PrmpRec	0.090	0.034			
School	GPS-Only	0.060	0.019	-0.0294	-0.78	0.4354
University	GPS PrmpRec	0.206	0.094			
University	GPS-Only	0.101	0.031	-0.1048	-1.05	0.2927
Shop	GPS PrmpRec	0.356	0.075			
Shop	GPS-Only	0.232	0.047	-0.1238	-1.39	0.1642
Social/Recreational	GPS PrmpRec	0.153	0.057			
Social/Recreational	GPS-Only	0.344	0.057	0.1903	2.40	0.0170
Other	GPS PrmpRec	0.264	0.076			
Other	GPS-Only	0.180	0.038	-0.0834	-1.00	0.3192
Total Tours	GPS PrmpRec	1.069	0.113			
Total Tours	GPS-Only	0.918	0.101	-0.1512	-1.00	0.3191

Table C-3, Part 4. Average number of tours per person comparing the GPS-with-prompted-recall with GPS-only households for non-workers, by tour type.

Tour Type	GPS Stratum	Average Number of Tours Per Person	Jackknife Standard Error	Difference with GPS-with-prompted-recall	T-stat for Difference with GPS-with-prompted-recall	P-value for No Difference with GPS-with-prompted-recall
Work	GPS PrmpRec	0.012	0.008			
Work	GPS-Only	0	0	-0.0123	-1.50	0.1333
Shop	GPS PrmpRec	0.524	0.046			
Shop	GPS-Only	0.436	0.031	-0.0884	-1.61	0.1088
Social/Recreational	GPS PrmpRec	0.344	0.045			
Social/Recreational	GPS-Only	0.340	0.031	-0.0044	-0.08	0.9378
Other	GPS PrmpRec	0.472	0.072			
Other	GPS-Only	0.210	0.023	-0.2617	-3.50	0.0005
Total Tours	GPS PrmpRec	1.352	0.094			
Total Tours	GPS-Only	0.986	0.054	-0.3668	-3.30	0.0010

Table C-3, Part 5. Average number of tours per person comparing the GPS-with-prompted-recall with GPS-only households for retirees, by tour type.

Tour Type	GPS Stratum	Average Number of Tours Per Person[1]	Jackknife Standard Error	Difference with GPS-with-prompted-recall	T-stat for Difference with GPS-with-prompted-recall	P-value for No Difference with GPS-with-prompted-recall
Work	GPS PrmpRec	0.004	0.004			
Work	GPS-Only	0	0	-0.0039	-1.00	0.3188
Shop	GPS PrmpRec	0.514	0.039			
Shop	GPS-Only	0.370	0.027	-0.1441	-2.97	0.0031
Social/Recreational	GPS PrmpRec	0.412	0.046			
Social/Recreational	GPS-Only	0.371	0.025	-0.0402	-0.78	0.4386
Other	GPS PrmpRec	0.356	0.060			
Other	GPS-Only	0.252	0.024	-0.1037	-1.59	0.1126
Total Tours	GPS PrmpRec	1.286	0.082			
Total Tours	GPS-Only	0.994	0.048	-0.2919	-2.88	0.0042

Table C-3, Part 6. Average number of tours per person comparing the GPS-with-prompted-recall with GPS-only households for driving-age children, by tour type.

Tour Type	GPS Stratum	Average Number of Tours Per Person[1]	Jackknife Standard Error	Difference with GPS-with-prompted-recall	T-stat for Difference with GPS-with-prompted-recall	P-value for No Difference with GPS-with-prompted-recall
School	GPS PrmpRec	0.434	0.084			
School	GPS-Only	0.409	0.059	-0.0252	-0.25	0.8043
University	GPS PrmpRec	0.025	0.025			
University	GPS-Only	0.006	0.006	-0.0189	-0.73	0.4637
Shop	GPS PrmpRec	0.224	0.077			
Shop	GPS-Only	0.072	0.026	-0.1516	-1.86	0.0630
Social/Recreational	GPS PrmpRec	0.336	0.132			
Social/Recreational	GPS-Only	0.324	0.053	-0.0128	-0.09	0.9305
Other	GPS PrmpRec	0.191	0.079			
Other	GPS-Only	0.184	0.046	-0.0066	-0.07	0.9421
Total Tours	GPS PrmpRec	1.210	0.162			
Total Tours	GPS-Only	0.995	0.088	-0.2151	-1.14	0.2529

Table C-3, Part 7. Average number of tours per person comparing the GPS-with-prompted-recall with GPS-only households for pre-driving-age children, by tour type.

Tour Type	GPS Stratum	Average Number of Tours Per Person[1]	Jackknife Standard Error	Difference with GPS-with-prompted-recall	T-stat for Difference with GPS-with-prompted-recall	P-value for No Difference with GPS-with-prompted-recall
School	GPS PrmpRec	0.695	0.038			
School	GPS-Only	0.585	0.032	-0.1098	-2.15	0.0323
Shop	GPS PrmpRec	0.098	0.021			
Shop	GPS-Only	0.042	0.014	-0.0565	-2.23	0.0264
Social/Recreational	GPS PrmpRec	0.216	0.038			
Social/Recreational	GPS-Only	0.364	0.029	0.1481	3.07	0.0022
Other	GPS PrmpRec	0.097	0.019			
Other	GPS-Only	0.119	0.016	0.0219	0.83	0.4073
Total Tours	GPS PrmpRec	1.107	0.045			
Total Tours	GPS-Only	1.110	0.042	0.0037	0.06	0.9533

Table C-3, Part 8. Average number of tours per person comparing the GPS-with-prompted-recall with GPS-only households for preschool children, by tour type.

Tour Type	GPS Stratum	Average Number of Tours Per Person	Jackknife Standard Error	Difference with GPS-with-prompted-recall	T-stat for Difference with GPS-with-prompted-recall	P-value for No Difference with GPS-with-prompted-recall
School	GPS PrmpRec	0.268	0.048			
School	GPS-Only	0.254	0.035	-0.0138	-0.24	0.8142
Shop	GPS PrmpRec	0.185	0.063			
Shop	GPS-Only	0.103	0.035	-0.0819	-1.15	0.2512
Social/Recreational	GPS PrmpRec	0.301	0.060			
Social/Recreational	GPS-Only	0.558	0.071	0.2576	2.85	0.0046
Other	GPS PrmpRec	0.168	0.036			
Other	GPS-Only	0.168	0.036	-0.0002	0.00	0.9974
Total Tours	GPS PrmpRec	0.922	0.073			
Total Tours	GPS-Only	1.084	0.085	0.1617	1.42	0.1573

Appendix D. Results Comparing Collection Days

There were 5,708 persons in households in the GPS tracking only stratum with at least one trip in the three-day period (persons with no trips at all are excluded from the analysis). The designation of the first day, second day, third day were generated by randomized assignment, so there is no reason why the day results should differ in their estimates or measured outcomes in any sense. Any significant difference must in fact be an artifact of data collection differences.

Table D-1 presents the estimated mean number of trips per person for these sampled households for Day 1 (the first data collection day) and Day 2 (the second data collection day) with the jackknife standard errors for these estimates. These calculations include the full household data file. The absolute value of the t-statistic for the difference is 13.8, which has a p-value less than 0.0001 for a two-sided test of the null hypothesis of no difference (with an alpha value of 0.05). The t-statistic for the difference was computed using a jackknife standard error for the difference per se, following the theory as given in Appendix B-2. Thus we ***do*** register a significant difference that can't be explained by chance alone, indicating collection and processing differences. The magnitude of this difference (4.026 vs. 2.627) is fairly sizeable: the second day of data collection definitely has a smaller mean than the first day of data collection. Again because of the randomized nature of the day assignment, this has to be a species of measurement error.

Table D-1. Weighted mean number of trips per person comparing Day 2 of data collection for each household with Day 1 of data collection for each household for the full data set.

Collection Day	Mean trips per person	Jack-knife stan-dard error	Diffe-rence with Day 1	T-stat for diffe-rence with Day 1	P-value for no diffe-rence with Day 1
Day 1	4.026	0.078			
Day 2	2.627	0.069	-1.39884	-13.82	<0.0001

Table D-2 presents a similar comparison for Day 1 (the first data collection day), Day 2 (the second data collection day), and Day 3 (the third day of data collection), restricting the data set to households which are assigned Monday, Tuesday, or Wednesday as Day 1 (households with Thursday or Friday as Day 1 have only two data collection days), with the jackknife standard errors for these estimates and 95% confidence intervals. This is called the 'MTW data set' below. We see a significant difference between Day 1 and each of Day 2 and Day 3, but not between Day 2 and Day 3. There is a dropoff between Day 1 and the later data collection days, but there is not a further dropoff from Day 2 to Day 3.

Table D-2. Weighted mean number of trips per person comparing Day 1, Day 2, and Day 3 of data collection for each household with Monday, Tuesday, or Wednesday as Day 1.

Collection Day	Mean trips per person	Jack-knife stan-dard error	Diffe-rence with Day 1	T-stat for diffe-rence with Day 1	P-value for no diffe-rence with Day 1
Day 1	3.917	0.095			
Day 2	2.678	0.095	-1.239	-8.95	<0.0001
Day 3	2.640	0.101	-1.276	-10.98	<0.0001

Table D-3 presents the estimated mean number of trips per sampled person again for Day 1 (the first data collection day), Day 2 (the second data collection day), and Day 3 (the third day of data collection), restricting to the 'MTW' data set, in this case by domains defined by trip purpose. As in Table D-2 we see generally a significant difference between Day 1 and each of Day 2 and Day 3, but not between Day 2 and Day 3.

The table includes the mean trips per person for the domain and collection day, the jackknife standard errors, the difference between collection days for the domain mean values, the t-statistic for the difference between Days 2 and 3 and Day 1, and the p-value for this difference[14]. In all domains, Day 2 and Day 3 have significantly lower mean trips per person than Day 1. Day 2 and Day 3 in all cases do not significantly differ from each other.

[14] The t-statistic is the difference divided by the jackknife standard error of the difference. The p-value is the two-sided p-value for the null hypothesis of zero difference.

Table D-3. Weighted mean number of trips per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set, by trip purpose domain.

Trip purpose domain	Collection Day	Mean trips per person	Jack-knife stan-dard error	Diffe-rence with Day 1	T-stat for diffe-rence with Day 1	P-value for no diffe-rence with Day 1	Sig/ NonSig Day 2 to Day 3 diffe-rence
Home-Based Other	Day 1	0.960	0.031				
Home-Based Other	Day 2	0.639	0.022	0.3206	8.89	<0.0001	NS
Home-Based Other	Day 3	0.606	0.023	0.3535	10.56	<0.0001	NS
Home-Based School	Day 1	0.121	0.011				
Home-Based School	Day 2	0.030	0.005	0.0912	8.00	<0.0001	NS
Home-Based School	Day 3	0.022	0.004	0.0989	8.91	<0.0001	NS
Home-Based Shopping	Day 1	0.218	0.014				
Home-Based Shopping	Day 2	0.144	0.014	0.0742	3.83	0.0001	NS
Home-Based Shopping	Day 3	0.150	0.013	0.0687	3.71	0.0002	NS
Home-Based Social/Recr	Day 1	0.696	0.029				
Home-Based Social/Recr	Day 2	0.415	0.029	0.2811	7.61	<0.0001	NS
Home-Based Social/Recr	Day 3	0.415	0.021	0.2809	8.88	<0.0001	NS
Home-Based University	Day 1	0.010	0.002				
Home-Based University	Day 2	0.004	0.001	0.0056	2.04	0.0424	NS
Home-Based University	Day 3	0.002	0.001	0.0080	3.25	0.0013	NS
Home-Based Work	Day 1	0.225	0.013				
Home-Based Work	Day 2	0.131	0.009	0.0943	8.32	<0.0001	NS
Home-Based Work	Day 3	0.121	0.010	0.1034	9.02	<0.0001	NS
Non-Home-Based Other	Day 1	1.484	0.059				
Non-Home-Based Other	Day 2	1.162	0.057	0.3218	3.86	0.0001	NS
Non-Home-Based Other	Day 3	1.180	0.068	0.3034	3.95	<0.0001	NS
Non-Home-Based Work	Day 1	0.204	0.013				
Non-Home-Based Work	Day 2	0.154	0.012	0.0499	3.69	0.0002	NS
Non-Home-Based Work	Day 3	0.144	0.011	0.0595	4.51	<0.0001	NS

Table D-4 presents a similar calculation as Table D-3, but in this case it is in terms of percentages of trips (so that the eight categories add to 1 for each person-day with at least one trip). Again there should not be any difference between the days: the reason for example for the much lower percentage of school trips in collection days 2 and 3 is unknown. In general, the trip percentages by trip purpose show much less difference between Day 1 and Days 2 and 3 than the absolute numbers of trips. There has been a drop off in the number of trips in Days 2 and 3, but not much of a shift in the distribution of these trips.

Table D-4. Weighted percentages of trips per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set, by trip purpose domain (for person-days with at least one trip).

Trip purpose domain	Collection Day	Percent of trips	Jack-knife stan-dard error	Diffe-rence with Day 1	T-stat for diffe-rence with Day 1	P-value for no diffe-rence with Day 1	Sig/ NonSig Day 2 to Day 3 diffe-rence
Home-Based Other	Day 1	0.268	0.008				
Home-Based Other	Day 2	0.281	0.009	0.0126	1.05	0.2946	NS
Home-Based Other	Day 3	0.264	0.009	-0.0039	-0.34	0.7317	NS
Home-Based School	Day 1	0.058	0.005				
Home-Based School	Day 2	0.025	0.005	-0.0335	-4.73	<0.0001	NS
Home-Based School	Day 3	0.016	0.003	-0.0421	-7.37	<0.0001	NS
Home-Based Shopping	Day 1	0.057	0.004				
Home-Based Shopping	Day 2	0.055	0.005	-0.0024	-0.41	0.6830	NS
Home-Based Shopping	Day 3	0.057	0.004	-0.0010	-0.17	0.8657	NS
Home-Based Social/Recr	Day 1	0.193	0.008				
Home-Based Social/Recr	Day 2	0.163	0.008	-0.0300	-2.66	0.0082	S
Home-Based Social/Recr	Day 3	0.191	0.009	-0.0019	-0.18	0.8585	S
Home-Based University	Day 1	0.003	0.001				
Home-Based University	Day 2	0.003	0.001	-0.0004	-0.23	0.8156	NS
Home-Based University	Day 3	0.001	0.001	-0.0018	-1.21	0.2255	NS
Home-Based Work	Day 1	0.092	0.006				
Home-Based Work	Day 2	0.084	0.007	-0.0076	-1.13	0.2610	NS
Home-Based Work	Day 3	0.085	0.008	-0.0070	-0.86	0.3902	NS
Non-Home-Based Other	Day 1	0.276	0.007				
Non-Home-Based Other	Day 2	0.328	0.008	0.0524	4.75	<0.0001	NS
Non-Home-Based Other	Day 3	0.332	0.010	0.0565	5.44	<0.0001	NS
Non-Home-Based Work	Day 1	0.052	0.003				
Non-Home-Based Work	Day 2	0.061	0.004	0.0089	1.98	0.0479	NS
Non-Home-Based Work	Day 3	0.053	0.003	0.0010	0.25	0.8024	NS

We did similar calculations as those given in Tables D-1 through D-4 for trip length and trip duration. In this case, the estimates are ratio estimates: persons with no trips at all for a given day are excluded from the estimators all together (the variance calculations do not include the zeroes for no-trip-days). These results are given in Tables D-5 and D-6 below.

Table D-5. Weighted mean trip length and trip duration comparing Day 1 and Day 2 of data collection for each household for the full data set.

Type of Measurement	Collection day	Mean per person	Jackknife standard error	Differen ce with Day 1	T-stat for diffe-rence with Day 1	P-value for no diffe-rence with Day 1
Mean Distance of Trips	Day 1	7.116	1.180			
Mean Distance of Trips	Day 2	6.561	0.196	-0.555	-0.46	0.6450
Mean Duration of Trips	Day 1	15.005	0.977			
Mean Duration of Trips	Day 2	12.906	0.208	-2.099	-2.07	0.0387

Table D-6. Weighted mean trip length and trip duration comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set.

Type of Measurement	Collection Day	Mean per Person	Jackknife Standard Error	Diffe-rence with Day 1	T-statistic for Difference with Day 1	P-value for No Difference with Day 1
Mean Distance of Trips	Day 1	7.644	2.000			
Mean Distance of Trips	Day 2	6.096	0.198	-1.548	-0.77	0.4424
Mean Distance of Trips	Day 3	6.323	0.224	-1.322	-0.67	0.5049
Mean Duration of Trips	Day 1	15.468	1.671			
Mean Duration of Trips	Day 2	12.582	0.235	-2.886	-1.70	0.0901
Mean Duration of Trips	Day 3	12.628	0.256	-2.840	-1.72	0.0865

As can be seen, the mean trip distance for Day 1 is nominally larger than that of Day 2 in the full data set, and Day 2 or Day 3 in the MTW data set, but the differences are not significant. We would accept the null hypothesis of no difference between the days for mean distance. The mean trip duration for Day 1 is larger than that of Day 2 for the full data set, and larger than that for Day 2 and Day 3 in the MTW data set. In the MTW data set case the difference is only marginally significant (significant at the 10% level), but this is likely due to the smaller household sample size. There is evidence that Day 1 is different (larger) than Days 2 and 3 for trip duration. Days 2 and 3 are not significantly different from each other.

We did similar calculations for trip distance and trip duration by trip purpose domain, as was done for Table D-4 above. There were only a few significant differences between Day 1 and Days 2 and 3 per trip purpose domain for trip length and trip duration, and these were not consistent. The smaller sample sizes may not allow for clear differences as we can see for Table D-5.

Tables D-7 through D-9 provide similar calculations for percentage of trips by mode and auto sufficiency (no autos in household for Table D-7, more workers than autos in household for

Table D-8, and as many or more autos than workers in household for Table D-9). These calculations are only over person-days with at least one trip.

Table D-7. Weighted percentage of trips per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for persons in households with no autos, by mode domain (for person-days with at least one trip).

Mode	Collection Day	Percent of Trips	Jack-knife Standard Error	Difference with Day 1	T-stat for Difference with Day 1	P-value for No Difference with Day 1	Sig/NonSig Day 2 to Day 3 Difference
Drive Alone	Day 1	0.419	0.061				
Drive Alone	Day 2	0.630	0.065	0.2110	2.78	0.0057	NS
Drive Alone	Day 3	0.629	0.081	0.2097	2.98	0.0030	NS
Shared Ride 2	Day 1	0.001	0.001				
Shared Ride 2	Day 2	0	0	-0.0008	-0.99	0.3245	NS
Shared Ride 2	Day 3	0	0	-0.0008	-0.99	0.3245	NS
Shared Ride 3+	Day 1	0.002	0.002				
Shared Ride 3+	Day 2	0	0	-0.0020	-0.99	0.3245	NS
Shared Ride 3+	Day 3	0	0	-0.0020	-0.99	0.3245	NS
Walk	Day 1	0.570	0.061				
Walk	Day 2	0.342	0.055	-0.2281	-3.24	0.0013	NS
Walk	Day 3	0.362	0.080	-0.2084	-2.99	0.0029	NS
Walk to Local Bus	Day 1	0.008	0.004				
Walk to Local Bus	Day 2	0.028	0.028	0.0200	0.71	0.4770	NS
Walk to Local Bus	Day 3	0.009	0.010	0.0015	0.14	0.8861	NS
Other	Day 1	0	0				
Other	Day 2	0	0	0			NS
Other	Day 3	0	0	0			NS

There are too few trips really for shared rides and walks to local bus (or subway). For drive alone trips and walking trips, there is evidence of significant differences between Days 2 and 3 and Day 1. Days 2 and 3 have more driving trips and fewer walking trips than Day 1. The reasons for this must be in data collection. It should be noted that the drive-alone mode percentage is unlikely to be correct in these cases. Households with no autos should not generally have drive-alone trips. The GPS information is clearly not being correctly interpreted. For the narrow purpose of evaluating differences between days, this data can be accepted provisionally, but it can't be accepted as a true analysis of travel modes in households.

Table D-8. Weighted percentage of trips per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for persons in households where the number of autos is less than the number of workers, by mode domain (for person-days with at least one trip).

Mode	Collec-tion Day	Per-cent of Trips	Jack-knife Standard Error	Diffe-rence with Day 1	T-stat for Diffe-rence with Day 1	P-value for No Diffe-rence with Day 1	Sig/NonSig Day 2 to Day 3 Difference
Drive Alone	Day 1	0.787	0.032				
Drive Alone	Day 2	0.690	0.088	-0.0965	-1.07	0.2873	S
Drive Alone	Day 3	0.870	0.024	0.0828	2.43	0.0157	S
Shared Ride 2	Day 1	0.009	0.005				
Shared Ride 2	Day 2	0.032	0.015	0.0231	1.46	0.1445	NS
Shared Ride 2	Day 3	0.043	0.018	0.0343	1.77	0.0773	NS
Shared Ride 3+	Day 1	0.009	0.008				
Shared Ride 3+	Day 2	0.117	0.102	0.1080	1.06	0.2916	NS
Shared Ride 3+	Day 3	0.001	0.001	-0.0083	-0.99	0.3239	NS
Walk	Day 1	0.189	0.030				
Walk	Day 2	0.160	0.041	-0.0290	-0.64	0.5243	NS
Walk	Day 3	0.087	0.016	-0.1025	-3.48	0.0005	NS
Walk to Local Bus	Day 1	0.003	0.002				
Walk to Local Bus	Day 2	0.001	0.001	-0.0025	-1.26	0.2097	NS
Walk to Local Bus	Day 3	0	0	-0.0031	-1.65	0.1006	NS
Other	Day 1	0.003	0.003				
Other	Day 2	0	0	-0.0032	-1.05	0.2941	NS
Other	Day 3	0	0	-0.0032	-1.05	0.2941	NS

For households with at least one auto, but ‘insufficient autos’, there are more shared rides and many more ‘drive-alone’ trips than the no auto households. The differences between Day 1, Day 2, and Day 3 are not great. The only significant difference is a shift of trips from walking to drive-alone for Day 3 as opposed to Day 1 and Day 2.

Table D-9. Weighted percentage of trips per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for persons in households where the number of autos is greater than or equal to the number of workers, by mode domain (for person-days with at least one trip).

Mode	Collec-tion Day	Percent of Trips	Jack-knife Standard Error	Diffe-rence with Day 1	T-stat for Diffe-rence with Day 1	P-value for No Diffe-rence with Day 1	Sig/NonSig Day 2 to Day 3 Difference
Drive Alone	Day 1	0.757	0.017				
Drive Alone	Day 2	0.929	0.010	0.1717	9.57	<0.0001	NS
Drive Alone	Day 3	0.940	0.010	0.1828	10.75	<0.0001	NS
Shared Ride 2	Day 1	0.040	0.008				
Shared Ride 2	Day 2	0.019	0.007	-0.0215	-3.43	0.0007	NS
Shared Ride 2	Day 3	0.022	0.009	-0.0182	-1.81	0.0714	NS
Shared Ride 3+	Day 1	0.025	0.011				
Shared Ride 3+	Day 2	0.005	0.004	-0.0206	-1.76	0.0792	NS
Shared Ride 3+	Day 3	0.001	0.001	-0.0242	-2.16	0.0316	NS
Walk	Day 1	0.176	0.011				
Walk	Day 2	0.047	0.008	-0.1290	-10.30	<0.0001	NS
Walk	Day 3	0.037	0.005	-0.1395	-12.47	<0.0001	NS
Walk to Local Bus	Day 1	0.001	0.000				
Walk to Local Bus	Day 2	0	0	-0.0008	-1.72	0.0858	NS
Walk to Local Bus	Day 3	0	0	-0.0008	-1.72	0.0858	NS
Other	Day 1	0.000	0.000				
Other	Day 2	0.000	0.000	0.0002	0.59	0.5564	NS
Other	Day 3	0	0	-0.0001	-1.00	0.3182	NS

Unlike for households with insufficient autos, the households with sufficient autos show very strong differences between Day 1 and Days 2 and 3 for drive alone trips and walking trips. Days 2 and 3 have a higher percentage of driving trips than walking trips than Day 1. Day 1 has many walking trips, but Days 2 and 3 have a significantly lower percentage of walking trips. The reasons for this must be in data collection..

Table D-10, Part 1. Average number of tours per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for full-time workers, by tour type.

Tour Type	Collection Day	Average Number of Tours Per Person	Jack-knife Standard Error	Difference with Day 1	T-stat for Difference with Day 1	P-value for No Difference with Day 1	Sig/NonSig Day 2 to Day 3 Difference
Work	Day 1	0.456	0.020				
Work	Day 2	0.295	0.016	-0.1603	-8.19	<0.0001	NS
Work	Day 3	0.280	0.017	-0.1762	-9.08	<0.0001	NS
Shop	Day 1	0.113	0.013				
Shop	Day 2	0.089	0.010	-0.0240	-1.49	0.1374	NS
Shop	Day 3	0.076	0.009	-0.0365	-2.56	0.0108	NS
Social/Recreational	Day 1	0.334	0.029				
Social/Recreational	Day 2	0.255	0.022	-0.0791	-2.75	0.0062	NS
Social/Recreational	Day 3	0.231	0.017	-0.1036	-2.94	0.0034	NS
Other	Day 1	0.222	0.019				
Other	Day 2	0.167	0.016	-0.0556	-2.40	0.0168	NS
Other	Day 3	0.176	0.015	-0.0465	-1.98	0.0480	NS
Work-Based Subtour	Day 1	0.069	0.009				
Work-Based Subtour	Day 2	0.056	0.009	-0.0129	-1.18	0.2392	S
Work-Based Subtour	Day 3	0.034	0.006	-0.0344	-3.28	0.0011	S
Total Tours	Day 1	1.194	0.036				
Total Tours	Day 2	0.862	0.037	-0.3319	-8.20	<0.0001	NS
Total Tours	Day 3	0.796	0.036	-0.3973	-8.24	<0.0001	NS

There is a significant difference in total tours and work tours between Days 2 and 3 and Day 1 for full-time workers. Day 1 has significantly more tours and work tours than Days 2 and 3. There are also a significantly larger number of tours for the other tour types as well, though less pronounced than the work tours. Day 2 and Day 3 do not differ from each except possibly for work-based subtours, though the difference is limited. The real difference is between Day 1 and Days 2 and 3.

Table D-10, Part 2. Average number of tours per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for part-time workers, by tour type.

Tour Type	Collection Day	Average Number of Tours Per Person	Jack-knife Standard Error	Difference with Day 1	T-stat for Difference with Day 1	P-value for No Difference with Day 1	Sig/NonSig Day 2 to Day 3 Difference
Work	Day 1	0.324	0.041				
Work	Day 2	0.149	0.026	-0.1748	-4.13	<0.0001	NS
Work	Day 3	0.154	0.024	-0.1699	-4.14	<0.0001	NS
Shop	Day 1	0.240	0.031				
Shop	Day 2	0.180	0.025	-0.0597	-1.69	0.0922	NS
Shop	Day 3	0.183	0.030	-0.0566	-1.63	0.1046	NS
Social/Recreational	Day 1	0.337	0.036				
Social/Recreational	Day 2	0.255	0.037	-0.0821	-1.71	0.0888	NS
Social/Recreational	Day 3	0.270	0.040	-0.0673	-1.44	0.1500	NS
Other	Day 1	0.390	0.044				
Other	Day 2	0.292	0.037	-0.0979	-2.06	0.0395	NS
Other	Day 3	0.230	0.040	-0.1599	-3.03	0.0026	NS
Work-Based Subtour	Day 1	0.021	0.010				
Work-Based Subtour	Day 2	0.023	0.011	0.0020	0.14	0.8874	NS
Work-Based Subtour	Day 3	0.014	0.009	-0.0067	-0.92	0.3592	NS
Total Tours	Day 1	1.311	0.076				
Total Tours	Day 2	0.898	0.070	-0.4124	-4.07	<0.0001	NS
Total Tours	Day 3	0.850	0.071	-0.4603	-5.44	<0.0001	NS

There is a significant difference in total tours and work tours between Days 2 and 3 and Day 1 for part-time workers as there was for full-time workers. Day 1 has significantly more tours and work tours than Days 2 and 3. There are also a significantly larger number of tours for the other tour types as well, though less pronounced than the work tours, as for full-time workers. Day 2 and Day 3 do not differ from each anywhere for part-time workers. The real difference is between Day 1 and Days 2 and 3.

Table D-10, Part 3. Average number of tours per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for university students, by tour type.

Tour Type	Collection Day	Average Number of Tours Per Person	Jack-knife Standard Error	Difference with Day 1	T-stat for Difference with Day 1	P-value for No Difference with Day 1	Sig/NonSig Day 2 to Day 3 Difference
School	Day 1	0.038	0.017				
School	Day 2	0.020	0.013	-0.0185	-1.65	0.0991	NS
School	Day 3	0.011	0.007	-0.0276	-1.77	0.0772	NS
University	Day 1	0.127	0.047				
University	Day 2	0.081	0.031	-0.0459	-1.10	0.2721	NS
University	Day 3	0.040	0.026	-0.0874	-2.66	0.0080	NS
Shop	Day 1	0.244	0.064				
Shop	Day 2	0.343	0.156	0.0985	0.58	0.5644	NS
Shop	Day 3	0.072	0.025	-0.1721	-2.45	0.0149	NS
Social/Recreational	Day 1	0.318	0.071				
Social/Recreational	Day 2	0.391	0.095	0.0724	0.63	0.5279	NS
Social/Recreational	Day 3	0.197	0.057	-0.1215	-1.83	0.0683	NS
Other	Day 1	0.205	0.052				
Other	Day 2	0.145	0.044	-0.0602	-0.81	0.4174	NS
Other	Day 3	0.090	0.030	-0.1152	-1.75	0.0814	NS
Total Tours	Day 1	0.933	0.128				
Total Tours	Day 2	0.979	0.254	0.0464	0.15	0.8808	S
Total Tours	Day 3	0.409	0.077	-0.5238	-4.13	<0.0001	S

For university students, the breakdown is different than that for full-time and part-time workers. Day 2 does not differ significantly from Day 1, but Day 3 does differ significantly from both Day 1 and Day 2 (a significant dropoff in average number of total tours per person, concentrating then in university, shopping and other trips).

Table D-10, Part 4. Average number of tours per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for non-workers, by tour type.

Tour Type	Collection Day	Average Number of Tours Per Person	Jack-knife Standard Error	Difference with Day 1	T-stat for Difference with Day 1	P-value for No Difference with Day 1	Sig/NonSig Day 2 to Day 3 Difference
Shop	Day 1	0.414	0.034				
Shop	Day 2	0.260	0.036	-0.1537	-3.45	0.0006	NS
Shop	Day 3	0.366	0.054	-0.0483	-0.72	0.4724	NS
Social/Recreational	Day 1	0.346	0.040				
Social/Recreational	Day 2	0.195	0.028	-0.1512	-3.40	0.0007	NS
Social/Recreational	Day 3	0.232	0.039	-0.1141	-2.81	0.0052	NS
Other	Day 1	0.216	0.031				
Other	Day 2	0.120	0.018	-0.0958	-3.44	0.0006	NS
Other	Day 3	0.114	0.019	-0.1017	-3.15	0.0018	NS
Total Tours	Day 1	0.976	0.064				
Total Tours	Day 2	0.575	0.054	-0.4008	-5.62	<0.0001	NS
Total Tours	Day 3	0.712	0.073	-0.2641	-2.77	0.0058	NS

Non-workers are similar to workers in part-time workers in that there is a significant difference in total tours and work tours between Days 2 and 3 and Day 1. Day 1 has significantly more tours and work tours than Days 2 and 3. There are also a significantly larger number of tours for all three kinds of tours for non-workers (shopping, social/recreational, other),. Day 2 and Day 3 do not differ significantly from each other (though Day 2 is nominally lower).

Table D-10, Part 5. Average number of tours per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for retirees, by tour type.

Tour Type	Collection Day	Average Number of Tours Per Person	Jack-knife Standard Error	Difference with Day 1	T-stat for Difference with Day 1	P-value for No Difference with Day 1	Sig/NonSig Day 2 to Day 3 Difference
Shop	Day 1	0.334	0.032				
Shop	Day 2	0.252	0.033	-0.0819	-1.65	0.0998	NS
Shop	Day 3	0.229	0.026	-0.1050	-2.71	0.0070	NS
Social/Recreational	Day 1	0.385	0.038				
Social/Recreational	Day 2	0.263	0.029	-0.1228	-2.59	0.0098	NS
Social/Recreational	Day 3	0.305	0.035	-0.0808	-1.92	0.0557	NS
Other	Day 1	0.288	0.037				
Other	Day 2	0.149	0.022	-0.1386	-3.16	0.0017	NS
Other	Day 3	0.171	0.025	-0.1169	-2.80	0.0053	NS
Total Tours	Day 1	1.007	0.070				
Total Tours	Day 2	0.664	0.060	-0.3433	-3.32	0.0010	NS
Total Tours	Day 3	0.704	0.051	-0.3027	-4.07	<0.0001	NS

Retirees are very similar to non-workers in that there is a significant different in total tours and work tours between Days 2 and 3 and Day 1. Day 1 has significantly more tours and work tours than Days 2 and 3. There are also a significantly larger number of tours for all three kinds of tours for retirees (shopping, social/recreational, other): the same as for non-workers. Day 2 and Day 3 do not differ significantly from each other.

Table D-10, Part 6. Average number of tours per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for driving-age children, by tour type.

Tour Type	Collection Day	Average Number of Tours Per Person	Jack-knife Standard Error	Difference with Day 1	T-stat for Difference with Day 1	P-value for No Difference with Day 1	Sig/NonSig Day 2 to Day 3 Difference
School	Day 1	0.373	0.066				
School	Day 2	0.303	0.059	-0.0699	-0.83	0.4043	NS
School	Day 3	0.204	0.054	-0.1688	-2.49	0.0130	NS
Shop	Day 1	0.046	0.021				
Shop	Day 2	0.035	0.027	-0.0109	-0.43	0.6642	NS
Shop	Day 3	0.012	0.012	-0.0345	-1.93	0.0538	NS
Social/Recreational	Day 1	0.272	0.066				
Social/Recreational	Day 2	0.252	0.118	-0.0194	-0.14	0.8900	NS
Social/Recreational	Day 3	0.218	0.091	-0.0534	-0.71	0.4773	NS
Other	Day 1	0.218	0.064				
Other	Day 2	0.077	0.042	-0.1408	-1.86	0.0636	NS
Other	Day 3	0.095	0.050	-0.1224	-1.64	0.1022	NS
Total Tours	Day 1	0.908	0.117				
Total Tours	Day 2	0.667	0.140	-0.2410	-1.30	0.1956	NS
Total Tours	Day 3	0.529	0.117	-0.3791	-3.57	0.0004	NS

Driving-age children show a weaker pattern of differences between Day 1 and Days 2 and 3. The Day 2 and 3 tours per person are lower, but are not generally significantly different (except for Day 3 for school tours and total tours). Day 2 and Day 3 are not significantly different. The smaller sample sizes may be causing the relative lack of significance.

Table D-10, Part 7. Average number of tours per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for pre-driving-age children, by tour type.

Tour Type	Collec-tion Day	Average Number of Tours Per Person	Jack-knife Standard Error	Diffe-rence with Day 1	T-stat for Diffe-rence with Day 1	P-value for No Diffe-rence with Day 1	Sig/NonSig Day 2 to Day 3 Difference
School	Day 1	0.530	0.040				
School	Day 2	0.092	0.017	-0.4382	-10.75	<0.0001	NS
School	Day 3	0.077	0.015	-0.4527	-11.36	<0.0001	NS
Shop	Day 1	0.037	0.018				
Shop	Day 2	0	0	-0.0372	-2.07	0.0388	NS
Shop	Day 3	0.007	0.006	-0.0299	-1.71	0.0875	NS
Social/Recreational	Day 1	0.394	0.041				
Social/Recreational	Day 2	0.077	0.022	-0.3175	-6.45	<0.0001	NS
Social/Recreational	Day 3	0.081	0.018	-0.3133	-7.18	<0.0001	NS
Other	Day 1	0.124	0.024				
Other	Day 2	0.063	0.016	-0.0615	-2.09	0.0374	NS
Other	Day 3	0.027	0.011	-0.0975	-4.14	<0.0001	NS
Total Tours	Day 1	1.086	0.055				
Total Tours	Day 2	0.231	0.037	-0.8544	-12.11	<0.0001	NS
Total Tours	Day 3	0.192	0.026	-0.8933	-14.50	<0.0001	NS

Pre-driving-age children show the strongest pattern of differences between Day 1 and Days 2 and 3 of all of the person types. The Day 2 and 3 tours per person are radically lower than Day 1, registering very strong significance levels across the board. We don’t know why this group is different from the others in this regard. Day 2 and 3 are not significantly different from each other.

Table D-10, Part 8. Average number of tours per person comparing Day 1, Day 2, and Day 3 of data collection for the MTW data set for preschool children, by tour type.

Tour Type	Collection Day	Average Number of Tours Per Person	Jackknife Standard Error	Difference with Day 1	T-stat for Difference with Day 1	P-value for No Difference with Day 1	Sig/NonSig Day 2 to Day 3 Difference
School	Day 1	0.258	0.048				
School	Day 2	0	0	-0.2584	-5.40	<0.0001	NS
School	Day 3	0	0	-0.2584	-5.40	<0.0001	NS
Shop	Day 1	0.073	0.030				
Shop	Day 2	0	0	-0.0728	-2.40	0.0169	NS
Shop	Day 3	0	0	-0.0728	-2.40	0.0169	NS
Social/Recreational	Day 1	0.547	0.100				
Social/Recreational	Day 2	0	0	-0.5466	-5.46	<0.0001	NS
Social/Recreational	Day 3	0	0	-0.5466	-5.46	<0.0001	NS
Other	Day 1	0.122	0.038				
Other	Day 2	0	0	-0.1221	-3.23	0.0013	NS
Other	Day 3	0	0	-0.1221	-3.23	0.0013	NS
Total Tours	Day 1	1.000	0.118				
Total Tours	Day 2	0	0	-0.9999	-8.45	<0.0001	NS
Total Tours	Day 3	0	0	-0.9999	-8.45	<0.0001	NS

Preschool children also show (as well as pre-driving age children) as strong a pattern of differences between Day 1 and Days 2 and 3 as could be possible. Day 1 has registered tours per person for many tour types, but Days 2 and 3 have absolutely no tours at all. Obviously this is an artifact of data collection.

Appendix E. Results for Tables

Appendix E-1. Results for Auto Ownership by County Tables

Table E-1-1 below presents weighted sample frequencies for County crossed with Number of Autos. The weights are normalized to add to the overall household sample size 4,540. Provided in the columns are the following values:

- Number of sampled households: the unweighted responding sample size of households in each cell;
- Weighted frequency of households: the total of the normalized weight in each cell;
- Weighted percentage of households: the percentage of each cell of the total weighted sample size of 4,540;
- Standard error of weighted percentage: jackknife standard error of the weighted percentage;
- Simple random sampling SRS standard error benchmark: the standard error for a simple random sample with the same percentage with a total sample size of 4,540;
- Design effect: the ratio of the jackknife standard error to the SRS standard error benchmark;
- Weight design effect: the design effect expected for a simple random sample using the design weights as 'haphazard' weights.

The SRS standard error benchmark represents the expected standard error for an unweighted percentage with the same population percentage and the same overall sample size. The design effect measures the degree to which the jackknife standard error is close to the SRS standard error. In this case, the true standard error should deviate from an SRS standard error from stratification and weighting differences. Stratification should lower variance, and unequal weights should increase variance. The Weight design effect is equal to $1+CV^2$, where CV is the coefficient of variation of the weights. This is the degree to which unequal weights should increase the variance when the weights are 'haphazard': they are not correlated to the measured estimand characteristic (see Kish (1992)).

Table E-1-1. Weighted Sample Frequencies for County crossed with Number of Autos in Households.

County	Number of autos in HH	Number of sampled house-holds	Weigh-ted fre-quency of house-holds	Wgtd pct of HHs	Stand-ard error of wgtd pct	SRS StdErr bench-mark	Design effect	Wgt Deff
Cuyahoga	0	551	393.8	8.67%	0.48%	0.42%	1.330	2.234
Cuyahoga	1	1,169	1,136.3	25.03%	0.87%	0.64%	1.832	1.825
Cuyahoga	2	947	971.0	21.39%	0.73%	0.61%	1.428	1.639
Cuyahoga	3	346	359.5	7.92%	0.53%	0.40%	1.781	1.755
Cuyahoga	Total	3,013	2,860.6	63.01%	0.49%	0.72%	0.464	1.815
Geauga	0	3	1.6	0.04%	0.00%	0.03%	0.016	1.374
Geauga	1	46	38.8	0.85%	0.12%	0.14%	0.732	1.127
Geauga	2	62	73.7	1.62%	0.17%	0.19%	0.787	1.304
Geauga	3	64	73.9	1.63%	0.22%	0.19%	1.336	1.133
Geauga	Total	175	187.9	4.14%	0.12%	0.30%	0.161	1.230
Lake	0	18	22.6	0.50%	0.19%	0.10%	3.322	2.019
Lake	1	173	192.5	4.24%	0.47%	0.30%	2.440	1.423
Lake	2	190	199.9	4.40%	0.26%	0.30%	0.701	1.118
Lake	3	86	94.7	2.09%	0.24%	0.21%	1.292	1.092
Lake	Total	467	509.7	11.23%	0.51%	0.47%	1.189	1.274
Lorain	0	29	28.8	0.63%	0.09%	0.12%	0.597	2.254
Lorain	1	179	205.6	4.53%	0.30%	0.31%	0.922	1.351
Lorain	2	237	268.1	5.91%	0.33%	0.35%	0.874	1.227
Lorain	3	126	125.6	2.77%	0.19%	0.24%	0.621	1.187
Lorain	Total	571	628.1	13.84%	0.35%	0.51%	0.457	1.306
Medina	0	13	11.5	0.25%	0.07%	0.07%	1.001	1.087
Medina	1	76	80.5	1.77%	0.19%	0.20%	0.965	1.266
Medina	2	147	171.3	3.77%	0.28%	0.28%	0.959	1.173
Medina	3	78	90.3	1.99%	0.23%	0.21%	1.192	1.198
Medina	Total	314	353.6	7.79%	0.22%	0.40%	0.310	1.199
5-County Total	0	614	458.3	10.09%	0.52%	0.45%	1.335	2.221
5-County Total	1	1,643	1,653.7	36.42%	0.98%	0.71%	1.873	1.672
5-County Total	2	1,583	1,684.0	37.09%	0.88%	0.72%	1.504	1.444
5-County Total	3	700	744.0	16.39%	0.68%	0.55%	1.548	1.440
5-County Total	Total	4,540	4,540.0	100.00%	0.00%	0.00%		1.000

As can be seen the jackknife design effects roughly align with the weight Deff for the cells, reflecting that the main influence on variance are the differential weights, though there are obviously correlations between the y-characteristic and the weights that are picked up by the jackknife. The jackknife design effects for the county totals are much smaller than the weight Deff for the county cells. This reflects the effect of stratification, which the jackknife variance estimator is designed to pick up.

Table E-1-2 presents row percentages: the percentages of weighted households in each number-of-autos cell within each county. These row percentages add to 100% for each county. The jackknife standard errors for these row percentages are provided, as well as SRS standard error benchmarks (these are based on taking a simple random sample within each county, with the row percentage as the population percentage and the county household sample size as the SRS sample size). A Deff is computed as the ratio of the jackknife standard error and the SRS benchmark.

Table E-1-2. Weighted Sample Frequencies for County crossed with Number of Autos in Households.

County	Number of autos in HH	Row pct	Row jack-knife std err	Row SRS std err bench-mark	Row deff
Cuyahoga	0	13.77%	0.76%	0.63%	1.485
Cuyahoga	1	39.72%	1.37%	0.89%	2.345
Cuyahoga	2	33.94%	1.11%	0.86%	1.663
Cuyahoga	3	12.57%	0.83%	0.60%	1.903
Cuyahoga	Total	100.00%	0.00%	0.00%	
Geauga	0	0.86%	0.09%	0.70%	0.017
Geauga	1	20.62%	2.92%	3.06%	0.914
Geauga	2	39.21%	4.21%	3.69%	1.299
Geauga	3	39.32%	4.59%	3.69%	1.549
Geauga	Total	100.00%	0.00%	0.00%	
Lake	0	4.43%	1.51%	0.95%	2.522
Lake	1	37.77%	2.86%	2.24%	1.620
Lake	2	39.22%	3.16%	2.26%	1.952
Lake	3	18.58%	2.36%	1.80%	1.721
Lake	Total	100.00%	0.00%	0.00%	
Lorain	0	4.58%	0.69%	0.87%	0.631
Lorain	1	32.73%	1.79%	1.96%	0.834
Lorain	2	42.69%	1.87%	2.07%	0.813
Lorain	3	20.00%	1.51%	1.67%	0.811
Lorain	Total	100.00%	0.00%	0.00%	
Medina	0	3.25%	0.98%	1.00%	0.956
Medina	1	22.76%	2.44%	2.37%	1.061
Medina	2	48.45%	3.26%	2.82%	1.333
Medina	3	25.54%	2.71%	2.46%	1.211
Medina	Total	100.00%	0.00%	0.00%	

There is no clear pattern distinguishing the jackknife standard errors from the standard errors assuming simple random sampling with differential weights in Table E-2. In this case, the design effects are not definitively different from 1 (sometimes they are larger than 1, sometimes smaller).

Appendix E-2. Results for Trip Distance and Trip Duration by Trip Purpose Domain

Table E-2-1 below present weighted sample means of trip distance and trip duration by trip purpose domain. These sample means are computed using the one-day file, the two-day file, and the full file. The jackknife standard errors are computed as well, as well as degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). We only included in this analysis estimates for which the degrees of freedom for the standard errors exceeded 30 for each file. The standard errors when the degrees of freedom are less than 30 have very wide confidence intervals, and including them in this analysis is not likely to contribute any scientifically meaningful information.

Table E-2-1. Trip Distance and Trip Duration Means and Standard Errors by File

Trip Purpose Domain	Variable	Data File	Total Trips	Domain Mean	Jack-knife Std Error	De-grees of Free-dom	CI for Std Err LB	CI for Std Err UB
Home-based School	Trip Distance	1-1dy	645	3.093	0.239	72	0.205	0.285
Home-based School	Trip Distance	2-2dy	788	3.352	0.252	64	0.215	0.305
Home-based School	Trip Distance	3-All	867	3.454	0.250	71	0.215	0.300
Home-based Shopping	Trip Distance	1-1dy	1,601	4.993	0.262	69	0.225	0.314
Home-based Shopping	Trip Distance	2-2dy	2,615	5.138	0.232	109	0.205	0.267
Home-based Shopping	Trip Distance	3-All	3,218	5.270	0.217	113	0.192	0.249
Home-based Work	Trip Distance	1-1dy	1,524	11.244	0.387	153	0.348	0.436
Home-based Work	Trip Distance	2-2dy	2,453	11.264	0.355	113	0.314	0.408
Home-based Work	Trip Distance	3-All	2,949	11.335	0.343	140	0.307	0.389
Non home-based Work	Trip Distance	1-1dy	1,442	7.235	0.311	102	0.273	0.360
Non home-based Work	Trip Distance	2-2dy	2,431	7.011	0.279	121	0.248	0.319
Non home-based Work	Trip Distance	3-All	2,979	6.937	0.249	133	0.223	0.283
Home-based School	Trip Duration	1-1dy	645	15.736	0.646	67	0.553	0.778
Home-based School	Trip Duration	2-2dy	788	15.010	0.545	129	0.486	0.620
Home-based School	Trip Duration	3-All	867	14.931	0.550	125	0.489	0.627
Home-based Work	Trip Duration	1-1dy	1,524	18.876	0.506	123	0.450	0.578
Home-based Work	Trip Duration	2-2dy	2,453	18.909	0.438	123	0.390	0.501
Home-based Work	Trip Duration	3-All	2,949	18.988	0.426	135	0.381	0.484
Non home-based Work	Trip Duration	1-1dy	1,442	13.613	0.423	147	0.380	0.478
Non home-based Work	Trip Duration	2-2dy	2,431	13.145	0.369	122	0.328	0.422
Non home-based Work	Trip Duration	3-All	2,979	12.992	0.324	144	0.291	0.366

Table E-2-2 provides estimates of the intra-person correlation for the two-day and full files by comparing the standard errors for the three files. Included are the following fields:

- Total trips $n^{(1)}$, $n^{(2)}$, $n^{(3)}$: the total number of trips that support the estimates (across persons and days) from the one-day, two-day, and full files respectively;
- Jackknife standard errors: the square roots of the jackknife variances $v(\bar{y}^{(1)})$, $v(\bar{y}^{(2)})$, $v(\bar{y}^{(3)})$.
- Inverse trip ratio: the ratio of the reciprocal of total trips for the two-day and full files ($1/n^{(2)}$ and $1/n^{(3)}$ respectively) to the reciprocal of total trips for the one-day file ($1/n^{(1)}$);
- Jackknife variance ratio: the ratio of the jackknife variance for the two-day and full files ($v(\bar{y}^{(2)})$ and $v(\bar{y}^{(3)})$ respectively) to the jackknife variance for the one-day file ($v(\bar{y}^{(1)})$);
- Roh calculation: the estimate of the within-person rate of homogeneity for the two-day and full files;
- Estimated within-person rate of homogeneity: equal to the mean of the two-day and full file $\boldsymbol{a}$ estimates;
- Pas design effect.

If the total trips were sampled in a simple random sample from some super-population of trips (i.e., with no clustering by persons or days), then there would be direct equality between the inverse trip ratios and the jackknife variance ratios (i.e., $v(\bar{y}^{(2)})/v(\bar{y}^{(1)}) = (1/n^{(2)})/(1/n^{(1)})$, $v(\bar{y}^{(3)})/v(\bar{y}^{(1)}) = (1/n^{(3)})/(1/n^{(1)})$). In all cases in Table E-2-2, the jackknife variance ratios are larger than the inverse trip ratios, which are consistent with a positive within-person $\boldsymbol{a}$. In the first case (home-based school domain—trip distance), the jackknife variances are actually larger for the two-day and full-files. In the case of 100% within-person correlation, the variances for the two-day and full files should be equal to the variance for the one-day file, rendering the extra trips in the second and third days for each person entirely superfluous. The estimated jackknife standard errors are consistent with this[15]. We estimate $\boldsymbol{a}$ as 100% in this case (though the estimates are also consistent with a large $\boldsymbol{a}$ less than 100%).

The $\boldsymbol{a}$ estimates for the two-day file are computed using the following formula:

$$roh^{(2)} = \frac{\{2 * v(\bar{y}^{(2)})/v(\bar{y}^{(1)})\} - 1}{2 - 1}$$

This is an inversion of the formula

$$\frac{v(\bar{y}^{(2)})}{v(\bar{y}^{(1)})} = \frac{1 + roh * (T - 1)}{T}$$

[15] The fact that they are actually slightly larger can be attributed to error in the variance estimates.

from Section 1 (with $T = 2$ for the two-day file). The a estimates for the full file are computed using the following formula:

$$roh^{(3)} = \frac{\{2.6 * v(\bar{y}^{(3)})/v(\bar{y}^{(1)})\} - 1}{2.6 - 1}$$

This is an inversion of the formula

$$\frac{v(\bar{y}^{(3)})}{v(\bar{y}^{(1)})} = \frac{1 + roh * (T - 1)}{T}$$

from Section 1 (with $T = 2.6$ for the full file[16]). The estimated within-person $\widehat{roh}$ is computed as the average of $roh^{(2)}$ and $roh^{(3)}$. Finally the Pas design effect is computed as:

$$\widehat{deff}\left(\bar{y}^{(T)}\right) = \frac{1 + \widehat{roh} * (T - 1)}{T}$$

[16] This T value is computed as T=3 for starting collection days Monday, Tuesday, and Wednesday, and T=2 for starting collection days Thursday and Friday, with each of the five days having an equal chance of being assigned to the household.

Table E-2-2. Trip Distance and Trip Duration Standard Errors and Intra-Person Correlation Calculations.

Trip Purpose Domain	Variable	Data File	Total trips	Jackknife Std Error	Inverse Trip Ratio	Jackknife variance ratio	α calculation	Estimated within-person α	Pas design effect
Home-basd School	Trip Dstnce	1-1dy	645	0.239	1.000	1.000		100%	1.00
Home-basd School	Trip Dstnce	2-2dy	788	0.252	0.819	1.117		100%	1.00
Home-basd School	Trip Dstnce	3-All	867	0.250	0.744	1.101		100%	1.00
Hom-bsd Shopping	Trip Dstnce	1-1dy	1,601	0.262	1.000	1.000		52.9%	1.00
Hom-bsd Shopping	Trip Dstnce	2-2dy	2,615	0.232	0.612	0.785	57.0%	52.9%	0.76
Hom-bsd Shopping	Trip Dstnce	3-All	3,218	0.217	0.498	0.685	48.7%	52.9%	0.71
Home-based Work	Trip Dstnce	1-1dy	1,524	0.387	1.000	1.000		66.6%	1.00
Home-based Work	Trip Dstnce	2-2dy	2,453	0.355	0.621	0.839	67.9%	66.6%	0.83
Home-based Work	Trip Dstnce	3-All	2,949	0.343	0.517	0.786	65.3%	66.6%	0.79
Non hm-bsd Work	Trip Dstnce	1-1dy	1,442	0.311	1.000	1.000		51.8%	1.00
Non hm-bsd Work	Trip Dstnce	2-2dy	2,431	0.279	0.593	0.807	61.5%	51.8%	0.76
Non hm-bsd Work	Trip Dstnce	3-All	2,979	0.249	0.484	0.644	42.2%	51.8%	0.70
Home-bsd School	Trip Duratn	1-1dy	645	0.646	1.000	1.000		48.5%	1.00
Home-bsd School	Trip Duratn	2-2dy	788	0.545	0.819	0.710	42.1%	48.5%	0.74
Home-bsd School	Trip Duratn	3-All	867	0.550	0.744	0.723	55.0%	48.5%	0.68
Home-based Work	Trip Duratn	1-1dy	1,524	0.506	1.000	1.000		51.3%	1.00
Home-based Work	Trip Duratn	2-2dy	2,453	0.438	0.621	0.751	50.1%	51.3%	0.76
Home-based Work	Trip Duratn	3-All	2,949	0.426	0.517	0.708	52.6%	51.3%	0.70
Non hm-bsd Work	Trip Duratn	1-1dy	1,442	0.423	1.000	1.000		42.6%	1.00
Non hm-bsd Work	Trip Duratn	2-2dy	2,431	0.369	0.593	0.762	52.5%	42.6%	0.71
Non hm-bsd Work	Trip Duratn	3-All	2,979	0.324	0.484	0.587	32.8%	42.6%	0.65

Any differences between the two-day file and the full-file α estimates would indicate a variance pattern more complicated than the Pas (1986) framework. In particular, the design effect would not be a direct function of a single α value and the number of days T, but a more complicated function. The apparent differences between the α estimates for each trip domain (e.g., 61.5% and 42.2% for trip distance for the non home-based work domain) can be explained from the noise in the jackknife standard errors. This justifies computing a final α estimate as an average of the two-day and full-file α estimates.

These final α estimates do differ across trip variable and domain, which can be allowed for in the Pas framework. They range from a low for 42.6% for non home-based work trip duration to 100% for home-based shopping trip distance. In general, one might expect a higher correlation for trip distance than for trip duration, as repeated trips (to work or school for example) might share a trip distance, but the trip duration may vary across days due to variable traffic. This is in fact what one can see from Table E-2-2, notwithstanding the noise in the jackknife variance estimates.

Appendix E-3. Results for Mean Tours per Person per Day by Tour Type and Person Type

Table E-3-1 (in six parts) presents weighted sample means of mean tours per person per day[17]. These sample means are computed using the one-day file, the two-day file, and the full file. The jackknife standard errors are computed as well, as well as degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). We only included in this analysis estimates for which the degrees of freedom for the standard errors exceeded 30 for each file. The standard errors when the degrees of freedom are less than 30 have very wide confidence intervals, and including them in this analysis is not likely to contribute any scientifically meaningful information.

Table E-3-1, Part 1. Mean Tours per Person per Day and Standard Errors for Full-Time Workers, by File

Tour Purpose	Data File	Total Tours	Average Number of Tours per Person	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
Work	1-1dy	2,150	0.470	0.015	228	0.014	0.017
Work	2-2dy	4,300	0.382	0.012	174	0.011	0.013
Work	3-All	5,574	0.359	0.011	154	0.010	0.013
Shop	1-1dy	2,150	0.113	0.009	33	0.007	0.011
Shop	2-2dy	4,300	0.100	0.006	135	0.005	0.006
Shop	3-All	5,574	0.095	0.005	132	0.004	0.006
Other	1-1dy	2,150	0.221	0.014	140	0.012	0.016
Other	2-2dy	4,300	0.200	0.010	208	0.009	0.011
Other	3-All	5,574	0.195	0.008	174	0.008	0.009
Work-Based Subtour	1-1dy	2,150	0.074	0.007	77	0.006	0.008
Work-Based Subtour	2-2dy	4,300	0.066	0.005	137	0.005	0.006
Work-Based Subtour	3-All	5,574	0.059	0.004	186	0.004	0.005
Total Tours	1-1dy	2,150	1.208	0.027	210	0.024	0.030
Total Tours	2-2dy	4,300	1.032	0.022	142	0.020	0.025
Total Tours	3-All	5,574	0.978	0.020	223	0.018	0.022

[17] Pre-driving children and pre-school children were excluded as their data was only collected through logs (not by GPS).

Table E-3-1, Part 2. Mean Tours per Person per Day and Standard Errors for Part-Time Workers, by File

Tour Purpose	Data File	Total Tours	Average Number of Tours per Person	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
Work	1-1dy	515	0.271	0.027	78	0.024	0.033
Work	2-2dy	1,030	0.204	0.019	150	0.017	0.021
Work	3-All	1,360	0.193	0.017	203	0.015	0.018
Shop	1-1dy	515	0.279	0.029	116	0.026	0.033
Shop	2-2dy	1,030	0.227	0.019	81	0.016	0.022
Shop	3-All	1,360	0.217	0.018	61	0.015	0.022
Social/Recreational	1-1dy	515	0.402	0.036	70	0.031	0.044
Social/Recreational	2-2dy	1,030	0.337	0.027	37	0.022	0.035
Social/Recreational	3-All	1,360	0.321	0.025	47	0.021	0.032
Other	1-1dy	515	0.354	0.034	120	0.030	0.039
Other	2-2dy	1,030	0.309	0.025	57	0.021	0.031
Other	3-All	1,360	0.291	0.022	85	0.019	0.026
Total Tours	1-1dy	515	1.318	0.064	87	0.055	0.075
Total Tours	2-2dy	1,030	1.090	0.048	35	0.039	0.062
Total Tours	3-All	1,360	1.036	0.043	61	0.037	0.053

Table E-3-1, Part 3. Mean Tours per Person per Day for University Students, by File

Tour Purpose	Data File	Total Tours	Average Number of Tours per Person	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
Social/Recreational	1-1dy	178	0.344	0.057	35	0.046	0.074
Social/Recreational	2-2dy	356	0.337	0.049	58	0.042	0.060
Social/Recreational	3-All	465	0.304	0.043	49	0.036	0.053
Other	1-1dy	178	0.180	0.038	30	0.031	0.051
Other	2-2dy	356	0.161	0.024	63	0.021	0.029
Other	3-All	465	0.144	0.018	74	0.016	0.022

Table E-3-1, Part 4. Mean Tours per Person per Day for Non-Workers, by File

Tour Purpose	Data File	Total Tours	Average Number of Tours per Person	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
Shop	1-1dy	809	0.436	0.031	93	0.027	0.037
Shop	2-2dy	1,618	0.361	0.026	49	0.021	0.032
Shop	3-All	2,121	0.362	0.022	83	0.019	0.026
Other	1-1dy	809	0.210	0.023	244	0.021	0.025
Other	2-2dy	1,618	0.161	0.016	148	0.014	0.018
Other	3-All	2,121	0.150	0.014	129	0.012	0.016
Total Tours	1-1dy	809	0.986	0.054	72	0.047	0.065
Total Tours	2-2dy	1,618	0.796	0.042	91	0.036	0.049
Total Tours	3-All	2,121	0.776	0.038	100	0.033	0.044

Table E-3-1, Part 5. Mean Tours per Person per Day and Standard Errors for Retirees, by File

Tour Purpose	Data File	Total Tours	Average Number of Tours per Person	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
Shop	1-1dy	925	0.370	0.027	173	0.024	0.030
Shop	2-2dy	1,850	0.314	0.017	153	0.016	0.020
Shop	3-All	2,400	0.295	0.015	125	0.014	0.017
Social/Recreational	1-1dy	925	0.371	0.025	41	0.021	0.032
Social/Recreational	2-2dy	1,850	0.316	0.020	68	0.017	0.024
Social/Recreational	3-All	2,400	0.314	0.020	46	0.017	0.025
Other	1-1dy	925	0.252	0.024	186	0.021	0.026
Other	2-2dy	1,850	0.194	0.014	164	0.013	0.016
Other	3-All	2,400	0.189	0.012	140	0.011	0.014
Total Tours	1-1dy	925	0.994	0.048	287	0.044	0.052
Total Tours	2-2dy	1,850	0.825	0.034	83	0.029	0.040
Total Tours	3-All	2,400	0.798	0.032	84	0.028	0.038

Table E-3-1, Part 6. Mean Tours per Person per Day for Driving-Age Children, by File

Tour Purpose	Data File	Total Tours	Average Number of Tours per Person	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
School	1-1dy	138	0.409	0.059	42	0.048	0.075
School	2-2dy	276	0.347	0.040	48	0.034	0.050
School	3-All	355	0.315	0.036	56	0.030	0.044
Total Tours	1-1dy	138	0.995	0.088	69	0.076	0.106
Total Tours	2-2dy	276	0.829	0.065	65	0.056	0.079
Total Tours	3-All	355	0.762	0.067	79	0.058	0.080

Table E-3-2 (in six parts) provides estimates of the intra-person correlation for the two-day and full files by comparing the standard errors for the three files, following the approach as given in Appendix E-2 for Table E-2-2.

Table E-3-2, Part 1. Mean Tours per Person per Day Standard Errors and Intra-Person Correlation Calculations for Full-Time Workers, by File.

Tour Purpose	Data File	Total Tours	Jackknife Standard Error	Inverse Tour Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
Work	1-1dy	2,150	0.015	1.00	1.00		28.06%	1.00
Work	2-2dy	4,300	0.012	0.50	0.63	26.47%	28.06%	0.64
Work	3-All	5,574	0.011	0.39	0.57	29.65%	28.06%	0.56
Shop	1-1dy	2,150	0.009	1.00	1.00		-13.38%	1.00
Shop	2-2dy	4,300	0.006	0.50	0.41	-17.81%	-13.38%	0.43
Shop	3-All	5,574	0.005	0.39	0.33	-8.95%	-13.38%	0.30
Other	1-1dy	2,150	0.014	1.00	1.00		-4.59%	1.00
Other	2-2dy	4,300	0.010	0.50	0.47	-5.91%	-4.59%	0.48
Other	3-All	5,574	0.008	0.39	0.36	-3.27%	-4.59%	0.36
Work-Based Subtour	1-1dy	2,150	0.007	1.00	1.00		22.09%	1.00
Work-Based Subtour	2-2dy	4,300	0.005	0.50	0.65	30.89%	22.09%	0.61
Work-Based Subtour	3-All	5,574	0.004	0.39	0.47	13.30%	22.09%	0.52
Total Tours	1-1dy	2,150	0.027	1.00	1.00		31.11%	1.00
Total Tours	2-2dy	4,300	0.022	0.50	0.68	35.23%	31.11%	0.66
Total Tours	3-All	5,574	0.020	0.39	0.55	26.99%	31.11%	0.58

For full-time workers, the estimated *α*'s are slightly negative for shopping and other tours, but are positive for work tours (25-30%), as one might expect, but the work tour correlation is not as high as one might expect given the consistency across days one might expect for work tours among full-time workers. The noise in the data may reduce what otherwise might be a larger *α* value.

Table E-3-2, Part 2. Mean Tours per Person per Day Standard Errors and Intra-Person Correlation Calculations for Part-Time Workers, by File.

Tour Purpose	Data File	Total Tours	Jackknife Standard Error	Inverse Tour Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
Work	1-1dy	515	0.027	1.00	1.00		-5.72%	1.00
Work	2-2dy	1,030	0.019	0.50	0.46	-8.68%	-5.72%	0.47
Work	3-All	1,360	0.017	0.38	0.37	-2.77%	-5.72%	0.35
Shop	1-1dy	515	0.029	1.00	1.00		-8.03%	1.00
Shop	2-2dy	1,030	0.019	0.50	0.42	-15.14%	-8.03%	0.46
Shop	3-All	1,360	0.018	0.38	0.38	-0.92%	-8.03%	0.34
Social/Recreational	1-1dy	515	0.036	1.00	1.00		14.78%	1.00
Social/Recreational	2-2dy	1,030	0.027	0.50	0.57	14.08%	14.78%	0.57
Social/Recreational	3-All	1,360	0.025	0.38	0.48	15.47%	14.78%	0.48
Other	1-1dy	515	0.034	1.00	1.00		8.22%	1.00
Other	2-2dy	1,030	0.025	0.50	0.54	8.66%	8.22%	0.54
Other	3-All	1,360	0.022	0.38	0.43	7.78%	8.22%	0.44
Total Tours	1-1dy	515	0.064	1.00	1.00		12.35%	1.00
Total Tours	2-2dy	1,030	0.048	0.50	0.56	12.24%	12.35%	0.56
Total Tours	3-All	1,360	0.043	0.38	0.46	12.47%	12.35%	0.46

For part-time workers, the estimated *α*'s are all slightly positive or slightly negative. One might expect less consistency for work tours for part-time workers.

Table E-3-2, Part 3. Mean Tours per Person per Day Standard Errors and Intra-Person Correlation Calculations for University Students, by File.

Tour Purpose	Data File	Total Tours	Jackknife Standard Error	Inverse Tour Ratio	Jackknife Variance Ratio	α Calcu-lation	Estimated Within-Person α	Pas Design Effect
Social/Recreational	1-1dy	178	0.057	1.00	1.00		39.60%	1.00
Social/Recreational	2-2dy	356	0.049	0.50	0.75	50.08%	39.60%	0.70
Social/Recreational	3-All	465	0.043	0.38	0.56	29.12%	39.60%	0.63
Other	1-1dy	178	0.038	1.00	1.00		-22.84%	1.00
Other	2-2dy	356	0.024	0.50	0.40	-19.71%	-22.84%	0.39
Other	3-All	465	0.018	0.38	0.22	-25.96%	-22.84%	0.24

Table E-3-2, Part 4. Mean Tours per Person per Day Standard Errors and Intra-Person Correlation Calculations for Non-Workers, by File.

Tour Purpose	Data File	Total Tours	Jackknife Standard Error	Inverse Tour Ratio	Jackknife Variance Ratio	α Calcu-lation	Estimated Within-Person α	Pas Design Effect
Shop	1-1dy	809	0.031	1.00	1.00		26.01%	1.00
Shop	2-2dy	1,618	0.026	0.50	0.67	34.76%	26.01%	0.63
Shop	3-All	2,121	0.022	0.38	0.49	17.27%	26.01%	0.54
Other	1-1dy	809	0.023	1.00	1.00		-2.40%	1.00
Other	2-2dy	1,618	0.016	0.50	0.49	-2.33%	-2.40%	0.49
Other	3-All	2,121	0.014	0.38	0.37	-2.48%	-2.40%	0.37
Total Tours	1-1dy	809	0.054	1.00	1.00		17.35%	1.00
Total Tours	2-2dy	1,618	0.042	0.50	0.59	18.20%	17.35%	0.59
Total Tours	3-All	2,121	0.038	0.38	0.49	16.49%	17.35%	0.49

Table E-3-2, Part 5. Mean Tours per Person per Day Standard Errors and Intra-Person Correlation Calculations for Retirees, by File.

Tour Purpose	Data File	Total Tours	Jackknife Standard Error	Inverse Tour Ratio	Jackknife Variance Ratio	α Calcu-lation	Estimated Within-Person α	Pas Design Effect
Shop	1-1dy	925	0.027	1.00	1.00		-11.94%	1.00
Shop	2-2dy	1,850	0.017	0.50	0.43	-14.28%	-11.94%	0.44
Shop	3-All	2,400	0.015	0.39	0.33	-9.59%	-11.94%	0.31
Social/Recreational	1-1dy	925	0.025	1.00	1.00		36.14%	1.00
Social/Recreational	2-2dy	1,850	0.020	0.50	0.65	30.11%	36.14%	0.68
Social/Recreational	3-All	2,400	0.020	0.39	0.64	42.16%	36.14%	0.61
Other	1-1dy	925	0.024	1.00	1.00		-24.49%	1.00
Other	2-2dy	1,850	0.014	0.50	0.35	-30.72%	-24.49%	0.38
Other	3-All	2,400	0.012	0.39	0.27	-18.26%	-24.49%	0.23
Total Tours	1-1dy	925	0.048	1.00	1.00		6.36%	1.00
Total Tours	2-2dy	1,850	0.034	0.50	0.51	1.83%	6.36%	0.53
Total Tours	3-All	2,400	0.032	0.39	0.45	10.89%	6.36%	0.42

Table E-3-2, Part 6. Mean Tours per Person per Day Standard Errors and Intra-Person Correlation Calculations for Driving-Age Children, by File.

Tour Purpose	Data File	Total Tours	Jackknife Standard Error	Inverse Tour Ratio	Jackknife Variance Ratio	α Calcu-lation	Estimated Within-Person α	Pas Design Effect
School	1-1dy	138	0.059	1.00	1.00		-3.87%	1.00
School	2-2dy	276	0.040	0.50	0.47	-6.43%	-3.87%	0.48
School	3-All	355	0.036	0.39	0.38	-1.30%	-3.87%	0.36
Total Tours	1-1dy	138	0.088	1.00	1.00		20.75%	1.00
Total Tours	2-2dy	276	0.065	0.50	0.55	9.79%	20.75%	0.60
Total Tours	3-All	355	0.067	0.39	0.58	31.72%	20.75%	0.51

For university students, non workers, retirees, and driving-age children, the results are ambiguous. The sample sizes may not be large enough to sustain reliable estimates and there may be issues with data quality.

Appendix E-4. Results for County to County Trip Percentages

Table E-4-1 below presents weighted percentages of trips according to the starting and ending counties of the trips. These sample percentages are computed using the one-day file, the two-day file, and the full file. The jackknife standard errors are computed as well, as well as degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). As in Appendix E-2, we only included in this analysis estimates for which the degrees of freedom for the standard errors exceeded 30 for each file.

Table E-4-2 then provides estimates of the intra-person correlation for the two-day and full files by comparing the standard errors for the three files, following the approach as given in Appendix E-2 for Table E-2-2.

Table E-4-1. Trip Weighted Percentages by Starting and Ending County (of Trip)and Standard Errors by File

Start County to End County	Data File	Total Trips	Percent of Trips	Jack-knife Std Error	De-grees of free-dom	CI for Std Err LB	CI for Std Err UB
Cuyahoga to Cuyahoga	1-1dy	26,408	60.44%	1.165%	323	1.082%	1.263%
Cuyahoga to Cuyahoga	2-2dy	44,048	59.76%	1.035%	372	0.965%	1.115%
Cuyahoga to Cuyahoga	3-All	54,421	59.90%	1.033%	309	0.957%	1.121%
Cuyahoga to Geauga	1-1dy	26,408	0.42%	0.085%	49	0.071%	0.105%
Cuyahoga to Geauga	2-2dy	44,048	0.47%	0.072%	88	0.063%	0.085%
Cuyahoga to Geauga	3-All	54,421	0.50%	0.082%	56	0.069%	0.100%
Cuyahoga to Lorain	1-1dy	26,408	1.44%	0.135%	69	0.116%	0.162%
Cuyahoga to Lorain	2-2dy	44,048	1.37%	0.112%	68	0.096%	0.135%
Cuyahoga to Lorain	3-All	54,421	1.34%	0.105%	74	0.091%	0.125%
Cuyahoga to Unknown	1-1dy	26,408	0.72%	0.078%	30	0.062%	0.104%
Cuyahoga to Unknown	2-2dy	44,048	0.77%	0.079%	31	0.064%	0.106%
Cuyahoga to Unknown	3-All	54,421	0.79%	0.073%	50	0.061%	0.091%
Geauga to Cuyahoga	1-1dy	26,408	0.47%	0.083%	45	0.069%	0.105%
Geauga to Cuyahoga	2-2dy	44,048	0.49%	0.073%	127	0.065%	0.083%
Geauga to Cuyahoga	3-All	54,421	0.53%	0.087%	115	0.077%	0.099%
Geauga to Geauga	1-1dy	26,408	2.66%	0.436%	153	0.392%	0.491%
Geauga to Geauga	2-2dy	44,048	2.59%	0.386%	213	0.353%	0.427%
Geauga to Geauga	3-All	54,421	2.45%	0.342%	141	0.306%	0.387%
Geauga to Unknown	1-1dy	26,408	0.13%	0.044%	55	0.037%	0.054%
Geauga to Unknown	2-2dy	44,048	0.13%	0.031%	47	0.025%	0.038%
Geauga to Unknown	3-All	54,421	0.11%	0.025%	104	0.022%	0.029%
Lake to Lake	1-1dy	26,408	8.11%	0.647%	92	0.566%	0.757%
Lake to Lake	2-2dy	44,048	8.54%	0.709%	34	0.574%	0.929%
Lake to Lake	3-All	54,421	8.39%	0.627%	36	0.510%	0.815%
Lorain to Cuyahoga	1-1dy	26,408	1.46%	0.132%	64	0.113%	0.160%
Lorain to Cuyahoga	2-2dy	44,048	1.40%	0.111%	68	0.095%	0.133%
Lorain to Cuyahoga	3-All	54,421	1.38%	0.104%	77	0.090%	0.123%
Lorain to Lorain	1-1dy	26,408	10.83%	0.632%	163	0.570%	0.709%
Lorain to Lorain	2-2dy	44,048	10.62%	0.643%	115	0.569%	0.738%
Lorain to Lorain	3-All	54,421	11.00%	0.722%	52	0.606%	0.894%
Medina to Medina	1-1dy	26,408	5.04%	0.443%	56	0.374%	0.544%
Medina to Medina	2-2dy	44,048	5.31%	0.431%	122	0.383%	0.493%
Medina to Medina	3-All	54,421	5.18%	0.369%	96	0.323%	0.429%
Unknown to Geauga	1-1dy	26,408	0.15%	0.048%	162	0.044%	0.054%
Unknown to Geauga	2-2dy	44,048	0.13%	0.030%	69	0.026%	0.036%
Unknown to Geauga	3-All	54,421	0.11%	0.025%	251	0.023%	0.027%
Unknown to Lake	1-1dy	26,408	0.08%	0.027%	33	0.022%	0.035%
Unknown to Lake	2-2dy	44,048	0.09%	0.022%	39	0.018%	0.028%
Unknown to Lake	3-All	54,421	0.09%	0.021%	34	0.017%	0.028%

Table E-4-2. County to County Trip Percentage Standard Errors and Intra-Person Correlation Calculations.

Start County to End County	Data File	Total Trips	Jack-knife Std Err	Inverse Trip Ratio	Jack-knife variance Ratio	α Calcu-lation	Esti-mated within-person α	Pas De-sign Ef-fect
Cuyahoga to Cuyahoga	1-1dy	26,408	1.165%	1.00	1.00		61.39%	1.00
Cuyahoga to Cuyahoga	2-2dy	44,048	1.035%	0.60	0.79	57.62%	61.39%	0.81
Cuyahoga to Cuyahoga	3-All	54,421	1.033%	0.49	0.79	65.16%	61.39%	0.76
Cuyahoga to Geauga	1-1dy	26,408	0.085%	1.00	1.00		67.51%	1.00
Cuyahoga to Geauga	2-2dy	44,048	0.072%	0.60	0.73	46.72%	67.51%	0.84
Cuyahoga to Geauga	3-All	54,421	0.082%	0.49	0.93	88.31%	67.51%	0.80
Cuyahoga to Lorain	1-1dy	26,408	0.135%	1.00	1.00		37.56%	1.00
Cuyahoga to Lorain	2-2dy	44,048	0.112%	0.60	0.69	38.86%	37.56%	0.69
Cuyahoga to Lorain	3-All	54,421	0.105%	0.49	0.61	36.27%	37.56%	0.62
Cuyahoga to Unknown	1-1dy	26,408	0.078%	1.00	1.00		90.28%	1.00
Cuyahoga to Unknown	2-2dy	44,048	0.079%	0.60	1.04	100.00%	90.28%	0.95
Cuyahoga to Unknown	3-All	54,421	0.073%	0.49	0.88	80.56%	90.28%	0.94
Geauga to Cuyahoga	1-1dy	26,408	0.083%	1.00	1.00		76.08%	1.00
Geauga to Cuyahoga	2-2dy	44,048	0.073%	0.60	0.76	52.16%	76.08%	0.88
Geauga to Cuyahoga	3-All	54,421	0.087%	0.49	1.08	100.00%	76.08%	0.85
Geauga to Geauga	1-1dy	26,408	0.436%	1.00	1.00		46.96%	1.00
Geauga to Geauga	2-2dy	44,048	0.386%	0.60	0.78	56.74%	46.96%	0.73
Geauga to Geauga	3-All	54,421	0.342%	0.49	0.61	37.17%	46.96%	0.67
Geauga to Unknown	1-1dy	26,408	0.044%	1.00	1.00		-8.16%	1.00
Geauga to Unknown	2-2dy	44,048	0.031%	0.60	0.47	-5.20%	-8.16%	0.46
Geauga to Unknown	3-All	54,421	0.025%	0.49	0.32	-11.12%	-8.16%	0.33
Lake to Lake	1-1dy	26,408	0.647%	1.00	1.00		95.09%	1.00
Lake to Lake	2-2dy	44,048	0.709%	0.60	1.20	100.00%	95.09%	0.98
Lake to Lake	3-All	54,421	0.627%	0.49	0.94	90.18%	95.09%	0.97
Lorain to Cuyahoga	1-1dy	26,408	0.132%	1.00	1.00		38.55%	1.00
Lorain to Cuyahoga	2-2dy	44,048	0.111%	0.60	0.70	39.38%	38.55%	0.69
Lorain to Cuyahoga	3-All	54,421	0.104%	0.49	0.62	37.71%	38.55%	0.62
Lorain to Lorain	1-1dy	26,408	0.632%	1.00	1.00		100.0%	1.00
Lorain to Lorain	2-2dy	44,048	0.643%	0.60	1.03	100.00%	100.0%	1.00
Lorain to Lorain	3-All	54,421	0.722%	0.49	1.31	100.00%	100.0%	1.00
Medina to Medina	1-1dy	26,408	0.443%	1.00	1.00		69.63%	1.00
Medina to Medina	2-2dy	44,048	0.431%	0.60	0.95	89.37%	69.63%	0.85
Medina to Medina	3-All	54,421	0.369%	0.49	0.69	49.90%	69.63%	0.81
Unknown to Geauga	1-1dy	26,408	0.048%	1.00	1.00		-21.24%	1.00
Unknown to Geauga	2-2dy	44,048	0.030%	0.60	0.38	-23.08%	-21.24%	0.39
Unknown to Geauga	3-All	54,421	0.025%	0.49	0.27	-19.39%	-21.24%	0.25
Unknown to Lake	1-1dy	26,408	0.027%	1.00	1.00		37.74%	1.00
Unknown to Lake	2-2dy	44,048	0.022%	0.60	0.68	35.34%	37.74%	0.69
Unknown to Lake	3-All	54,421	0.021%	0.49	0.63	40.14%	37.74%	0.62

As with Table E-2-2, any differences between the two-day file and the full-file α estimates would indicate a variance pattern more complicated than the Pas (1986) framework. In Table E-3-2, there are differences in the estimated α's between the two-day and full-file, but these differences are not systematic and can be explained by the noise in the standard errors. We can accept provisionally the Pas framework again and assume the underlying intra-person α values are the same across days.

These final α estimates do differ across county pairs, which can again be allowed for in the Pas framework. For percentage of trips within Lorain County for example, the variances for the two-day and full files are actually higher than the one-day file, indicating maximum intra-person correlation. This type of travel pattern does not vary much across days (as if persons either take all their trips within the county, or some other pattern). On the opposite side of the spectrum, percentage of trips from Geauga County to 'Unknown' (outside the five-county region) or from 'Unknown' to Geauga County, show a negative α, consistent with no correlation across days within persons. In this case, it is as if this trip incidence is a simple random sample from a infinite population of trips, with no clustering within persons. For other county pairs, the ranges are across the spectrum. There is noise in the variances, but it appears that there are in fact differences in within-person correlations across these county pairs.

Appendix E-5. Mode Choice by Auto Sufficiency Tables

Table E-5-1 below presents weighted percentages by mode choice (for trips), by household auto sufficiency domain (no autos, fewer autos than workers, as many or more autos than workers). These sample percentages are computed using the one-day file, the two-day file, and the full file. The jackknife standard errors are computed as well, as well as degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). As in Appendix E-2, we only included in this analysis estimates for which the degrees of freedom for the standard errors exceeded 30 for each file. There were only two mode choice percentages (drive alone and walking) for two domains (no autos, as many or more autos than workers) that had at least 30 degrees of freedom for each file.

Table E-5-2 then provides estimates of the intra-person correlation for the two-day and full files by comparing the standard errors for the three files, following the approach as given in Appendix E-2 for Table E-5-2.

Table E-5-1. Trip Weighted Percentages and Standard Errors by Mode Choice for Auto Sufficiency Domains by File.

Auto Sufficiency	Mode	Data File	Total Trips	Percent of Trips	Jack-knife Std Error	De-grees of free-dom	CI for Std Err LB	CI for Std Err UB
No Autos	Drive Alone	1-1dy	1,754	49.26%	4.642%	34	3.755%	6.082%
No Autos	Drive Alone	2-2dy	2,813	57.13%	4.908%	143	4.399%	5.551%
No Autos	Drive Alone	3-All	3,363	61.24%	4.605%	89	4.016%	5.397%
No Autos	Walk	1-1dy	1,754	47.00%	4.687%	42	3.864%	5.957%
No Autos	Walk	2-2dy	2,813	39.27%	4.632%	201	4.220%	5.133%
No Autos	Walk	3-All	3,363	35.71%	4.320%	119	3.834%	4.948%
Autos ≥ Workers	Drive Alone	1-1dy	22,027	81.54%	1.054%	41	0.868%	1.345%
Autos ≥ Workers	Drive Alone	2-2dy	36,740	86.36%	0.727%	32	0.585%	0.962%
Autos ≥ Workers	Drive Alone	3-All	45,469	88.04%	0.658%	37	0.537%	0.852%
Autos ≥ Workers	Walk	1-1dy	22,027	12.92%	0.684%	125	0.609%	0.781%
Autos ≥ Workers	Walk	2-2dy	36,740	9.25%	0.450%	141	0.403%	0.509%
Autos ≥ Workers	Walk	3-All	45,469	8.02%	0.388%	136	0.347%	0.440%

Table E-5-2. Mode Choice Percentages and Standard Errors and Intra-Person Correlation Calculations by Auto Sufficiency Domain.

Auto Sufficiency	Mode	Data File	Total Trips	Jack-knife Std Error	In-verse Trip Ratio	Jack-knife vari-ance ratio	α Calcu-lation	Within-Day α	Pas De-sign Effect
No Autos	Drv Aln	1-1dy	1,754	4.642%	1.00	1.00		98.70%	1.00
No Autos	Drv Aln	2-2dy	2,813	4.908%	0.62	1.12	100.0%	98.70%	0.99
No Autos	Drv Aln	3-All	3,363	4.605%	0.52	0.98	97.39%	98.70%	0.99
No Autos	Walk	1-1dy	1,754	4.687%	1.00	1.00		85.45%	1.00
No Autos	Walk	2-2dy	2,813	4.632%	0.62	0.98	95.34%	85.45%	0.93
No Autos	Walk	3-All	3,363	4.320%	0.52	0.85	75.56%	85.45%	0.91
Autos≥Wrkrs	Drv Aln	1-1dy	22,027	1.054%	1.00	1.00		-1.99%	1.00
Autos≥Wrkrs	Drv Aln	2-2dy	36,740	0.727%	0.60	0.48	-4.82%	-1.99%	0.49
Autos≥Wrkrs	Drv Aln	3-All	45,469	0.658%	0.48	0.39	0.83%	-1.99%	0.37
Autos≥Wrkrs	Walk	1-1dy	22,027	0.684%	1.00	1.00		-11.9%	1.00
Autos≥Wrkrs	Walk	2-2dy	36,740	0.450%	0.60	0.43	-13.53%	-11.9%	0.44
Autos≥Wrkrs	Walk	3-All	45,469	0.388%	0.48	0.32	-10.26%	-11.9%	0.31

It is clear that the drive-alone percentage for households with no autos is far too high and reflects defects in the GPS imputation without the recall feature. Setting this issue aside, the α values for the two mode percentage estimates for households with no autos are very high (99%

and 85%): later data collection days are very much like earlier days. On the other hand, the *α* values for the two mode percentage estimates for household with at least as many autos as households are slightly negative (-2% and -12%). Later data collection days are not at all like earlier data collection days: the data set is more reflective of collection days being independent. We are not sure if this result is just finally a function of noise in the data, given one might expect a certain consistency across days for travel mode.

Tables E-5-3 and E-5-5 further illustrate the sample size and GPS-imputation issues as they relate to trip modes. Table E-5-3 shows the number of observed trips by mode, for the GPS-with-recall segment and each day of the GPS-only segment. Table E-5-4 shows the weighted trips, and Table E-5-5 shows the mode shares with the same break-outs.

Table E-5-3. Number of trip observations by mode for each sample type and day number.

	GPS-with-Recall	GPS-Only		
Mode	Day 1	Day 1	Day 2	Day 3/4
Drive-Alone	6,503	21,864	13,431	12,475
Shared Ride 2	3,141	717	213	269
Shared Ride 3+	1,706	295	118	12
Walk	1,445	3,713	852	615
Bike	85	5	0	0
Local Bus	270	69	14	11
Express Bus	12	0	0	0
Rail	38	8	4	2
Other	337	9	1	0
Total	13,537	26,680	14,633	13,384

Table E-5-4. Number of weighted trips by mode for each sample type and day number.

	GPS-with-Recall	GPS-Only		
Mode	Day 1	Day 1	Day 2	Day 3/4
Drive-Alone	5,871	20,778	12,941	12,139
Shared Ride 2	2,918	1,066	231	302
Shared Ride 3+	2,142	366	270	8
Walk	1,281	3,928	814	568
Bike	80	3	0	0
Local Bus	207	53	26	9
Express Bus	15	0	0	0
Rail	45	6	6	1
Other	459	8	1	0
Total	13,017	26,209	14,289	13,027

Table E-5-4. Mode shares for each sample type and day number.

Mode	GPS-with-Recall Day 1	GPS-Only Day 1	Day 2	Day 3/4
Drive-Alone	45.10%	79.30%	90.60%	93.20%
Shared Ride 2	22.40%	4.10%	1.60%	2.30%
Shared Ride 3+	16.50%	1.40%	1.90%	0.10%
Walk	9.80%	15.00%	5.70%	4.40%
Bike	0.60%	0.00%	0.00%	0.00%
Local Bus	1.60%	0.20%	0.20%	0.10%
Express Bus	0.10%	0.00%	0.00%	0.00%
Rail	0.30%	0.00%	0.00%	0.00%
Other	3.50%	0.00%	0.00%	0.00%
Total	100.00%	100.00%	100.00%	100.00%

There are a few observations of note in these tables.

First, there are a small number of observations beyond the first four rows. These tables are not segmented by either trip purpose or auto sufficiency, so with those segmentations added, the data would be even thinner. In itself, this addresses one important issue—the sample size even with the full three day sample (and potentially with the GPS-with-recall sample included) is not sufficient to provide a trustworthy observation of the mode shares in the Cleveland region. This not unusual for household travel surveys, especially in a region with low transit mode shares. It serves to further illustrate the importance of collecting an onboard transit survey if understanding transit demand and ridership markets is a planning priority.

Second, the mode shares are very different across the samples. The GPS-with-recall has a drive-alone mode share of 45% across all purposes, compared to day one of the GPS-only sample which has a drive-alone mode share of 79%. Days two and three are even higher, over 90%. This is similar to the findings of Section 2 where we found a bias between the GPS-with-recall and day one of the GPS-only sample, and of Section 3 where we found a bias between the first and subsequent days within the GPS-only sample. We suspect that there is a limitation of the GPS mode imputation process where it does not pick up non-drive-alone modes very well.

Appendix F. Estimates and Design Effects for Model Estimation: Technical Details

For the estimated models, the output is a vector parameter estimate $\boldsymbol{\beta}$. We will compute three versions of each model's parameter based on the one-day, the two-day, and the full files: $\hat{\boldsymbol{\beta}}^{(1)}$, $\hat{\boldsymbol{\beta}}^{(2)}$, $\hat{\boldsymbol{\beta}}^{(3)}$. These are weighted estimates using the w_{sh} as weights. Jackknife variance estimators $v_J(\hat{\boldsymbol{\beta}}^{(1)})$, $v_J(\hat{\boldsymbol{\beta}}^{(2)})$, $v_J(\hat{\boldsymbol{\beta}}^{(3)})$ are computed as per the formulas in Appendix A.3. In some cases, it is considered appropriate to compute an unweighted estimate rather than a weighted estimate. In this case the weights w_{sh} are replaced with unit weights (all equal to 1), and the jackknife replicate weights are those as given in Appendix B-3. Otherwise, the unweighted analyses proceed in the same way as the weighted analyses (with formulas given in Appendix B-4).

In this case the desired parameter is a K-vector. Each of these vector elements will have separate variances, and each will also potentially separate design effects. In the simplest case, the design effects for the K parameter elements will all be equal to a common design effect $deff$. This common design effect $deff$ can be decomposed as $\frac{1+a(T-1)}{T}$, with a being a common rate of homogeneity.

It may be though that vector elements will have separate variances and separate design effects as well. In this case we have design effects $deff_1,\ldots, deff_k,\ldots\ deff_K$ for each vector element one by one, with corresponding separate rates of homogeneity $a_1,\ldots, a_k,\ldots\ a_K$. But this does not capture the effects of the sample design on the overall variance of the vector parameter $\hat{\boldsymbol{\beta}}^{(T)}$, as there are covariances as well as variances. The full variance in this case can be summarized in K 'eigenvalues', which are in fact variances for particular linear combinations of $\hat{\boldsymbol{\beta}}^{(T)}$. The largest eigenvalue corresponds to the particular linear combination of the vector $\hat{\boldsymbol{\beta}}^{(T)}$ that has the largest variance, the smallest eigenvalue to the particular linear combination with the smallest variance.

The design effects can also be analyzed in exactly the same way. The design effect in the univariate parameter case is the ratio of one variance to another variance. For vector parameters, the design effect becomes one variance matrix "divided" by another[18]. This variance matrix is itself a K by K symmetric matrix, which can be summarized by K eigenvalues. These eigenvalues represent the 'design effect' magnitude for particular linear combinations. Suppose we call these eigenvalues $gdeff_1,\ldots, gdeff_k,\ldots\ gdeff_K$ (the g prefix indicating 'generalized design effect'[19]). Each of these can be matched to a factors $a_1,\ldots, a_k,\ldots\ a_K$ based on the formula $\frac{1+a(T-1)}{T}$. If these generalized design effects are equal, then everything simplifies to a single design effect which can represent the full $\hat{\boldsymbol{\beta}}^{(T)}$ vector.

For the estimated models, the output is a vector parameter estimate $\boldsymbol{\beta}$. We will compute three versions of each model's parameter based on the one-day, the two-day, and the full files:

[18] One variance matrix multiplied to the inverse of another.

[19] See for example Skinner et al. (1989), p. 43.

$\widehat{\boldsymbol{\beta}}^{(1)}$, $\widehat{\boldsymbol{\beta}}^{(2)}$, $\widehat{\boldsymbol{\beta}}^{(3)}$. These are weighted estimates using the w_{sh} as weights. Jackknife variance estimators $v_J(\widehat{\boldsymbol{\beta}}^{(1)})$, $v_J(\widehat{\boldsymbol{\beta}}^{(2)})$, $v_J(\widehat{\boldsymbol{\beta}}^{(3)})$ are computed as per the formulas in Appendix A.3. In some cases, unweighted estimates of the vector parameters are preferred. The jackknife variance estimates $v_J(\widehat{\boldsymbol{\beta}}^{(1)})$, $v_J(\widehat{\boldsymbol{\beta}}^{(2)})$, $v_J(\widehat{\boldsymbol{\beta}}^{(3)})$ are computed using the replicate weights as given in Section A.4.

Design effects in this case are more complicated than the simple case of a mean value as discussed in Section 1. The simplest case is a linear regression model where the regression predictor variables are all defined at the household level (i.e., their values don't change across the days). For example, household size, household location, presence of workers, senior citizens, school children, etc. are all predictors which can be viewed as fixed at the household level across all days. Suppose $\mathbf{X}$ is an n by K matrix consisting of K n-vectors of fixed predictors $\mathbf{x}_k$ (for example $\mathbf{x}_1$ might be the number of adults, $\mathbf{x}_2$ the number of working adults, etc.). Suppose $\mathbf{W}$ is an n by n diagonal matrix with the sample weights of each household along the diagonal (if an unweighted estimate is being computed then W will be the n by n identity matrix). Suppose $\bar{\mathbf{y}}^{(1)}$ is a vector of outcome variables from the single day file, $\bar{\mathbf{y}}^{(2)}$ a vector of the mean of an outcome variable over two days from the two day file, $\bar{\mathbf{y}}^{(3)}$ a vector of the mean of an outcome variable over all days from the full file. Then

$$\widehat{\boldsymbol{\beta}}^{(1)} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\left(\mathbf{X}'\mathbf{W}\bar{\mathbf{y}}^{(1)}\right) \quad \widehat{\boldsymbol{\beta}}^{(2)} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\left(\mathbf{X}'\mathbf{W}\bar{\mathbf{y}}^{(2)}\right) \quad \widehat{\boldsymbol{\beta}}^{(3)} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\left(\mathbf{X}'\mathbf{W}\bar{\mathbf{y}}^{(3)}\right)$$

Write $\boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}$ as a K by K matrix with the true sampling variance of $\widehat{\boldsymbol{\beta}}^{(1)}$. This variance will include the effects of weights, stratification, and potential heteroscedasticity, as well as the effect of day clustering (though only one day is represented from each household). Then $v_J(\widehat{\boldsymbol{\beta}}^{(1)})$ will be a consistent estimator[20] of $\boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}$. Likewise $\boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{2})}$ is a k by k matrix with the true sampling variance of $\widehat{\boldsymbol{\beta}}^{(2)}$. In this case, $\boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{2})}$ has the same weighting, stratification, heteroscedasticity and day clustering effects, but it now represents a $\bar{\mathbf{y}}^{(2)}$ value that is a mean over two days. $v_J(\widehat{\boldsymbol{\beta}}^{(2)})$ will be a consistent estimator of $\boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{2})}$. Finally, $\boldsymbol{V}_{\boldsymbol{\beta}}^{(3)}$ as a k by k matrix with the true sampling variance of $\widehat{\boldsymbol{\beta}}^{(3)}$. $\boldsymbol{V}_{\boldsymbol{\beta}}^{(3)}$ represents a $\bar{\mathbf{y}}^{(2)}$ value that is a mean over the full set of days for each household (three or four). $v_J(\widehat{\boldsymbol{\beta}}^{(3)})$ will be a consistent estimator of $\boldsymbol{V}_{\boldsymbol{\beta}}^{(3)}$.

The design effect for a univariate estimator is the ratio of the variance under the sample design to some benchmark variance. The generalization of this is to define a generalized design effect matrix. One reference for this is Skinner, Holt, and Smith (1989), Section 2.11. The eigenvalues of this generalized design matrix become the 'generalized design effects'. In this case the design effect matrices are $\mathrm{vr}\left(\widehat{\boldsymbol{\beta}}^{(2)}, \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right) = \left\{\boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right\}^{-\mathbf{1}} \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{2})}$ and $\mathrm{vr}\left(\widehat{\boldsymbol{\beta}}^{(3)}, \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right) = \left\{\boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right\}^{-\mathbf{1}} \boldsymbol{V}_{\boldsymbol{\beta}}^{(3)}$.

---

[20] In general, the jackknife variance estimator when correctly defined should generate unbiased estimators of the variances of totals, and consistent estimators of the variances of means, as well as 'smooth functions' of means. Smoothness here means continuity and differentiability. Regression coefficients are in fact smooth functions of sample means of the cross products of the X predictor vectors and of the X predictor vectors and the y vector. The jackknife variance estimators should succeed in being a consistent estimator of the true sampling variance without actually disaggregating the components of this sampling variance. This is a powerful property.

These are estimated in turn consistently by $\widehat{\mathrm{vr}}\left(\widehat{\boldsymbol{\beta}}^{(2)}, \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right) = \left\{v_J\left(\widehat{\boldsymbol{\beta}}^{(1)}\right)\right\}^{-\mathbf{1}} v_J\left(\widehat{\boldsymbol{\beta}}^{(2)}\right)$ and $\widehat{\mathrm{vr}}\left(\widehat{\boldsymbol{\beta}}^{(3)}, \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right) = \left\{v_J\left(\widehat{\boldsymbol{\beta}}^{(1)}\right)\right\}^{-\mathbf{1}} v_J\left(\widehat{\boldsymbol{\beta}}^{(3)}\right)$.

Pas (1986) and Koppelman and Pas (1984) develops a similar framework under a simple model. They work with an unweighted regression parameter estimate $\overline{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}(\mathbf{X}'\overline{\mathbf{Y}})$, where $\mathbf{X}$ is a predictor matrix assumed constant over the observation period, and $\overline{\mathbf{Y}}$ is a vector of means over the observation period for the y-variable of interest. Based on their assumed 'crossed-error structure' model, the variance of this parameter estimate is $\mathrm{Var}\left(\boldsymbol{\beta}^{(\boldsymbol{T})}\right) = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\sigma^2 \frac{1+a(T-1)}{T}$, where a is a correlation across days within an individual, and T is the number of days. For a one-day file, this reduces to the simple $\left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-\mathbf{1}}\sigma^2$. If we compute the design effect matrix $\left\{\mathrm{Var}\left(\boldsymbol{\beta}^{(\mathbf{1})}\right)\right\}^{-1}\mathrm{Var}\left(\boldsymbol{\beta}^{(\boldsymbol{T})}\right)$ in this case we get $\frac{1+a(T-1)}{T}\mathbf{I}$**,** where **I** is the identity matrix (a diagonal matrix with the constant $\frac{1+a(T-1)}{T}$ along the diagonal).

Our anticipation is that the eigenvalues $\delta_1^{(T)}, \ldots \delta_K^{(T)}$ of $\mathrm{vr}\left(\widehat{\boldsymbol{\beta}}^{(T)}, \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right), T = 2{,}3$ can be written as $\delta_k^{(T)} = \frac{1+roh_k(T-1)}{T}$, where the roh_k values are homogeneity measures that are constant across the two-day and full files. As in the simple univariate case discussed in Section 1, these 'rate of homogeneity' values are analogous to correlation coefficients, but are not exactly the same as they include sampling effects as well as population relationships.

In the ideal case, all of the eigenvalues $\delta_1^{(T)}, \ldots \delta_K^{(T)}$ will be equal to a common value and the design effect matrix will be a constant times the K by K identity matrix $\mathbf{I}_{\boldsymbol{K}}$, with differing constants for the two-day and full files. This will certainly simplify the conclusions from the analysis. Our goal is to find simplifying patterns from the empirical jackknife variance matrices, if those patterns can be justified. The empirical jackknife variance matrices will be subject to sampling error themselves, as they are random variables, so the eigenvalues will need to be analyzed as to whether they are consistent with particular models (we will test for example whether the empirical eigenvalues are consistent with an underlying constant value or not, and whether eigenvalues from a particular analysis for a domain are the same as those from another domain, or not). We are hopeful that the degrees of freedom in the jackknife variance estimator should be sufficient to distinguish various hypotheses with some precision.

References.

Koppelman, F. S., and Pas, E. I. (1984). Estimation of disaggregate regression models of person trip generation with multiday data. *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory* (eds. J. Volmuller, R. Hamerslag), Utrecht, Netherlands: VNU Science Press, 513-531.

Pas, E. I. (1986). Multiday samples, parameter estimation precision, and data collection costs for least squares regression trip-generation models. *Environment and Planning A, 18*, 73-87.

Skinner, C. J., Holt, D., Smith, T. M. F. (1989). *Analysis of Complex Surveys.* New York: John Wiley and Sons.

Appendix G. Model Estimation Results: Parameter Estimates, Variances, and Design Effects

Appendix G provides estimation results, jackknife standard errors, and design effects in detail for the models fitted for the three data files (auto ownership: Appendix G-1; non-work tour generation: Appendix G-2; work tour generation: Appendix G-3; work tour mode choice models: Appendix G-4; social/recreational tour mode choice model: Appendix G-5).

Appendix G-1. Model Estimation Results for Auto Ownership

The auto ownership model predicts the probability of a household owning 0, 1, 2 or 3+ vehicles (cars or light trucks). It is a multinomial logit (MNL) model of the form:

$$\Pr(i) = \frac{\exp(U_i)}{\sum_{j \in J} \exp(U_j)}$$

where $\Pr(i)$ is the probability of alternative *i*, U_i is the utility of alternative *i*, and *J* is the set of all alternatives. The utility can be expressed as $U_i = \beta X_i$ where β is the vector of estimated model coefficients and X_i is the vector of predictors.

Table G-1-1 presents the parameter estimates for the Multinomial Auto Ownership Model. The parameters are relative factors for the probability that a household fell into the particular category. T-statistics are presented for the null hypothesis that the coefficients are zero. Coefficients that are insignificant or marginally significant, but still included in the model are highlighted. Both 'model-based' and jackknife t-statistics are presented, using model-based and jackknife standard errors in the t-statistic denominator respectively. The 'model-based' standard errors are those coming from the model fit assuming simple random sampling, but including differential weights. This again is the simple random sampling with 'haphazard weights' paradigm[21]. Note that a larger t-statistic reflects a smaller standard error. In general, the jackknife standard errors mirror the model-based standard errors. This reflects the lack of clustering at this household level, with no differences across days either (auto ownership is determined in this data file at the household level).

[21] The 'haphazard weights' paradigm is from Kish, L. (1992), "Unequal pi weighting", *Journal of Official Statistics 8 (2)*: 183-200. This refers to the increase in variance induced by the use of weights when there is no relationship between the weights and the underlying variance structure.

Table G-1-1. Parameter Estimates for Auto Ownership Model

Description	Alternative								
	0 Autos			2 Autos			3+ Autos		
	Coeff	t-stat		Coeff	t-stat		Coeff	t-stat	
		Model-based	Jack-knife		Model-based	Jack-knife		Model-based	Jack-knife
Household size 1									
Household size 2				2.258	18.02	17.63	1.905	9.03	9.00
Household size 3				2.264	13.04	12.77	2.613	10.47	10.59
Household size 4+	-0.406	-1.72	-1.64	2.381	13.66	12.97	2.546	10.16	10.27
0 workers in household									
1 worker in household	-0.783	-6.46	-6.49						
2 workers in household	-0.961	7.53	7.70	0.886	7.53	7.70	1.076	7.11	7.13
3+ workers in household							2.977	14.90	15.07
Income $0-9k									
Income $10-24k	-0.898	-7.20	-7.18						
Income $25-49k	-2.265	-13.36	-12.97	1.014	7.91	7.30	1.051	4.76	4.72
Income $50-99k	-3.841	-8.26	-7.90	2.105	15.66	15.17	2.381	11.14	10.91
Income $100k+	-2.962	-4.03	-3.01	2.742	12.94	15.72	3.280	12.04	13.43
Mixed density msr	0.097	1.70	1.78	-0.264	-5.03	-4.94	-0.449	-5.56	-5.20
Composite logsum msr	0.255	6.59	6.11	-0.034	-2.52	-2.55	-0.089	-5.32	-5.42
Alternative-Specific Cnst	1.184	5.51	5.47	-3.023	-16.08	-15.4	-4.685	-15.36	-15.67
Home TAZ is Missing	-0.286	-1.20	-1.14	-0.551	-2.23	-2.21	-0.292	-0.83	-0.87

The reference alternative in the model is owning 1 auto. The predictors for this model are described as follows:

- Households with a larger size tend to own more autos, although household size has little to no effect on the choice between owning 0 or 1 autos.
- Households with more workers tend to own more autos.
- Higher income households tend to ow more autos.
- The mixed density measure represents the mix of households and employment, as defined in Appendix I. Mixed use areas are correlated with owning fewer autos.
- The composite logsum measure represents the relative accessibility by transit versus by auto, as defined in Appendix I. In places where transit accessibility is high relative to auto accessibility, households are likely to own fewer autos.
- An alternative specific constant is used to match the aggregate shares.
- A separate parameter is estimated if the home TAZ is missing, to avoid biasing the density and composite logsum estimates.

Appendix G-2. Model Estimation Results for Non-worker Tour Generation

The tour generation model jointly predicts the number and purpose of tours made by individual. The models are specific to person type, and this section describes the models estimated for non-workers. Appendix I enumerates the possible alternatives, with non-workers restricted to the first eight alternatives:

0. Stay at home (H)
1. 1+ shopping tour (SH)
2. 1+ social/recreation tour (SR)
3. 1+ other tour (O)
4. 1+ shopping tour and 1+ social/recreation tour (SH-SR)
5. 1+ shopping tour and 1+ other tour (SH-O)
6. 1+ social/recreation tour and 1+ other tour (SR-O)
7. 1+ shopping tour, 1+ social/recreation tour and 1+ other tour (SH-SR-O)

The models are estimated from each of the three files using weighted maximum likelihood estimation, and the jackknife weights were utilized to generate jackknife estimates for each of the three files. The weights for the base estimation are all 1, with the weights varying for each jackknife replicate, but still 1 for most observations. A weighted MLE parameter vector estimate and jackknife variance matrices are generated for each file, and comparisons made of the variance matrices. Tables G-2-1, G-2-2, and G-2-3 provide the parameter estimates from the full data set, the two-day data set, and the one-day data set respectively.

In all cases, stay at home is the reference alternative. The models include a set of alternative specific constants to match the aggregate shares by alternative. There are a number of demographic variables that interact with specific tour purposes. For example, women are more likely to make shopping tours, but less likely to make social/recreation and other tours. Higher income travelers are more likely to make social/recreation and other tours, but income was not found to have a significant effect on the propensity to make shopping tours. 0 vehicle households are less likely to make certain types of tours. The presence of workers in the household, other non-working adults in the household, and children in the household affect the likelihood of making some tours. This is probably because the presence and type of other household members affects the allocation of household maintenance activities within the household. The highway logsum is a measure of overall highway accessibility at the home location. Non-workers living in more accessible locations are slightly, but insignificantly, less likely to make shopping tours, and more likely to make social/recreation tours. These results are generally logical, and provide an illustration of the types of detailed demographic variables that can be included in such models.

Table G-2-1. Parameter Estimates for Non-Worker Tour Generation Model: Full Data Set

Description	Alternative includes:	Coefficient parameter estimate	Jackknife T-statistic
Alternative-Specific Constant	SH	2.4491	1.08
	SR	-5.6098	-2.48
	O	-1.9129	-20.81
	SH-SR	-3.1755	-0.95
	SH-O	0.7434	0.33
	SR-O	-7.0853	-3.12
	SH-SR-O	-4.6009	-1.37
Person is female	Shopping tour	0.2041	2.87
	Social/recreation tour	-0.2756	-3.80
	Other tour	-0.1844	-1.95
Person is under age 35	Other tour	-0.5431	-2.71
Person is age 65+	Shopping tour	-0.2561	-3.34
Person has no driver's license	Shopping tour	-0.5145	-5.11
	Social/recreation tour	-0.4917	-4.30
Income $25-49k	Social/recreation tour	0.2150	1.93
	Other tour	0.1807	1.44
Income $50k+	Social/recreation tour	0.3435	2.99
	Other tour	0.3178	2.86
0 Vehicles	Shopping tour	-0.4969	-3.86
	Social/recreation tour	-0.3983	-2.81
Workers present in HH	Shopping tour	-0.3806	-4.31
	Social/recreation tour	-0.2141	-2.30
Other non-working adults present in HH	Shopping tour	-0.2774	-3.44
Children present in HH	Other tour	0.2991	2.17
Highway logsum	Shopping tour	-0.2241	-1.30
	Social/recreation tour	0.3529	2.07

Table G-2-2. Parameter Estimates for Non-Worker Tour Generation Model: Two-Day Data Set

Description	**Alternative includes:**	**Coefficient parameter estimate**	**Jackknife T-statistic**
Alternative-Specific Constant	SH	3.6527	1.57
	SR	-8.2389	-3.25
	O	-1.7299	-16.40
	SH-SR	-4.7961	-1.34
	SH-O	1.8400	0.79
	SR-O	-9.5747	-3.77
	SH-SR-O	-6.1287	-1.72
Person is female	Shopping tour	0.2098	2.68
	Social/recreation tour	-0.2294	-2.88
	Other tour	-0.1268	-1.24
Person is under age 35	Other tour	-0.4407	-2.15
Person is age 65+	Shopping tour	-0.2090	-2.43
Person has no driver's license	Shopping tour	-0.4765	-4.03
	Social/recreation tour	-0.4233	-3.18
Income $25-49k	Social/recreation tour	0.2813	2.30
	Other tour	0.0979	0.69
Income $50k+	Social/recreation tour	0.3796	2.97
	Other tour	0.2500	1.99
0 Vehicles	Shopping tour	-0.4215	-2.69
	Social/recreation tour	-0.3478	-2.13
Workers present in HH	Shopping tour	-0.3961	-4.07
	Social/recreation tour	-0.2371	-2.20
Other non-working adults present in HH	Shopping tour	-0.2627	-2.81
Children present in HH	Other tour	0.3038	1.99
Highway logsum	Shopping tour	-0.3031	-1.72
	Social/recreation tour	0.5555	2.92

Table G-2-3. Parameter Estimates for Non-Worker Tour Generation Model: One-Day Data Set

Description	Alternative includes:	Coefficient parameter estimate	Jackknife T-statistic
Alternative-Specific Constant	SH	2.4853	0.82
	SR	-8.1851	-2.38
	O	-1.1933	-8.99
	SH-SR	-6.2801	-1.18
	SH-O	0.6389	0.21
	SR-O	-9.4956	-2.76
	SH-SR-O	-7.7068	-1.46
Person is female	Shopping tour	0.0596	0.59
	Social/recreation tour	-0.3886	-3.80
	Other tour	-0.2169	-1.77
Person is under age 35	Other tour	-0.4636	-2.02
Person is age 65+	Shopping tour	-0.1683	-1.46
Person has no driver's license	Shopping tour	-0.5636	-3.85
	Social/recreation tour	-0.4640	-2.81
Income $25-49k	Social/recreation tour	0.2274	1.37
	Other tour	0.1288	0.79
Income $50k+	Social/recreation tour	0.4188	2.48
	Other tour	0.3304	2.11
0 Vehicles	Shopping tour	-0.5158	-2.78
	Social/recreation tour	-0.2983	-1.45
Workers present in HH	Shopping tour	-0.2795	-2.23
	Social/recreation tour	-0.2234	-1.70
Other non-working adults present in HH	Shopping tour	-0.1524	-1.28
Children present in HH	Other tour	0.3023	1.71
Highway logsum	Shopping tour	-0.1788	-0.78
	Social/recreation tour	0.5918	2.29

Table G-2-5 below present the parameter estimates as they are given in tables G-2-1, G-2-2, and G-2-3, but with the three day-file estimates together for each parameter. Also included are jackknife standard errors for the parameter estimates (Tables G-2-1 through G-2-3 present the t-statistics, which are the parameter estimates divided by the standard errors) Table G-2-5 presents degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). Note that in these tables we provided all estimates, even when the degrees of freedom were smaller than 30, to keep the parameter vector whole. But the variance calculations with low degrees of freedom should be treated with skepticism. Table G-2-4 presents a listing of short parameter names used in Tables G-2-5 and G-2-6 (to save space), linking back to the parameter descriptions in Tables G-2-1 through G-2-3.

Table G-2-4. Comparison of Full Parameter Names to Short Parameter Names for Non-Worker Tour Generation Model Parameters

Full Parameter Name		Short Parameter Name
Description	**Alternative includes:**	**Label**
Alternative-Specific Constant	SH	asc2
	SR	asc3
	O	asc4
	SH-SR	asc5
	SH-O	asc6
	SR-O	asc7
	SH-SR-O	asc8
Person is female	Shopping tour	female_sh
	Social/recreation tour	female_sr
	Other tour	female_o
Person is under age 35	Other tour	age35u_o
Person is age 65+	Shopping tour	age65p_sh
Person has no driver's license	Shopping tour	nolic_sh
	Social/recreation tour	nolic_sr
Income $25-49k	Social/recreation tour	inc3_sr
	Other tour	inc3_o
Income $50k+	Social/recreation tour	inc45_sr
	Other tour	inc45_o
0 Vehicles	Shopping tour	veh0_sh
	Social/recreation tour	veh0_sr
Workers present in HH	Shopping tour	wkrs_sh
	Social/recreation tour	wkrs_sr
Other non-working adults present in HH	Shopping tour	othnwk_sh
Children present in HH	Other tour	kids_o
Highway logsum	Shopping tour	hwylsum_sh
	Social/recreation tour	hwylsum_sr

Table G-2-5 (Part 1). Parameter Estimates for Non-Worker Tour Generation Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
age35u_o	1-1dy	1,734	-0.464	0.230	31	0.184	0.305
age35u_o	2-2dy	3,138	-0.441	0.205	18	0.155	0.303
age35u_o	3-All	4,521	-0.543	0.200	24	0.156	0.279
age65p_sh	1-1dy	1,734	-0.168	0.115	64	0.098	0.139
age65p_sh	2-2dy	3,138	-0.209	0.086	98	0.076	0.100
age65p_sh	3-All	4,521	-0.256	0.077	98	0.067	0.089
asc2	1-1dy	1,734	2.485	3.016	75	2.601	3.590
asc2	2-2dy	3,138	3.653	2.324	62	1.977	2.819
asc2	3-All	4,521	2.449	2.271	44	1.880	2.869
asc3	1-1dy	1,734	-8.185	3.433	51	2.877	4.257
asc3	2-2dy	3,138	-8.239	2.531	38	2.069	3.262
asc3	3-All	4,521	-5.610	2.263	66	1.934	2.728
asc4	1-1dy	1,734	-1.193	0.133	30	0.106	0.177
asc4	2-2dy	3,138	-1.730	0.105	22	0.082	0.149
asc4	3-All	4,521	-1.913	0.092	39	0.075	0.118
asc5	1-1dy	1,734	-6.280	5.326	46	4.426	6.689
asc5	2-2dy	3,138	-4.796	3.586	52	3.010	4.437
asc5	3-All	4,521	-3.176	3.357	57	2.838	4.110
asc6	1-1dy	1,734	0.639	3.018	78	2.610	3.579
asc6	2-2dy	3,138	1.840	2.317	62	1.972	2.811
asc6	3-All	4,521	0.743	2.272	45	1.885	2.862
asc7	1-1dy	1,734	-9.496	3.444	52	2.891	4.261
asc7	2-2dy	3,138	-9.575	2.540	39	2.081	3.262
asc7	3-All	4,521	-7.085	2.273	63	1.936	2.753
asc8	1-1dy	1,734	-7.707	5.292	47	4.406	6.629
asc8	2-2dy	3,138	-6.129	3.566	51	2.988	4.422
asc8	3-All	4,521	-4.601	3.350	57	2.832	4.101

Table G-2-5 (Part 2). Parameter Estimates for Non-Worker Tour Generation Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
female_o	1-1dy	1,734	-0.217	0.123	25	0.096	0.169
female_o	2-2dy	3,138	-0.127	0.103	32	0.082	0.136
female_o	3-All	4,521	-0.184	0.095	36	0.077	0.123
female_sh	1-1dy	1,734	0.060	0.101	44	0.083	0.127
female_sh	2-2dy	3,138	0.210	0.078	69	0.067	0.094
female_sh	3-All	4,521	0.204	0.071	71	0.061	0.085
female_sr	1-1dy	1,734	-0.389	0.102	181	0.093	0.114
female_sr	2-2dy	3,138	-0.229	0.080	159	0.072	0.090
female_sr	3-All	4,521	-0.276	0.073	117	0.064	0.083
hwylsum_sh	1-1dy	1,734	-0.179	0.229	77	0.197	0.271
hwylsum_sh	2-2dy	3,138	-0.303	0.177	62	0.150	0.214
hwylsum_sh	3-All	4,521	-0.224	0.173	44	0.143	0.218
hwylsum_sr	1-1dy	1,734	0.592	0.258	52	0.216	0.319
hwylsum_sr	2-2dy	3,138	0.556	0.190	40	0.156	0.244
hwylsum_sr	3-All	4,521	0.353	0.170	70	0.146	0.204
inc3_o	1-1dy	1,734	0.129	0.162	82	0.141	0.192
inc3_o	2-2dy	3,138	0.098	0.143	35	0.116	0.186
inc3_o	3-All	4,521	0.181	0.126	28	0.100	0.170
inc3_sr	1-1dy	1,734	0.227	0.167	97	0.146	0.194
inc3_sr	2-2dy	3,138	0.281	0.122	57	0.103	0.150
inc3_sr	3-All	4,521	0.215	0.111	24	0.087	0.155
inc45_o	1-1dy	1,734	0.330	0.156	69	0.134	0.188
inc45_o	2-2dy	3,138	0.250	0.126	108	0.111	0.145
inc45_o	3-All	4,521	0.318	0.111	86	0.097	0.131
inc45_sr	1-1dy	1,734	0.419	0.169	46	0.140	0.212
inc45_sr	2-2dy	3,138	0.380	0.128	26	0.101	0.175
inc45_sr	3-All	4,521	0.343	0.115	33	0.093	0.151

Table G-2-5 (Part 3). Parameter Estimates for Non-Worker Tour Generation Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
kids_o	1-1dy	1,734	0.302	0.177	88	0.154	0.208
kids_o	2-2dy	3,138	0.304	0.153	44	0.126	0.193
kids_o	3-All	4,521	0.299	0.138	41	0.113	0.175
nolic_sh	1-1dy	1,734	-0.564	0.146	63	0.125	0.177
nolic_sh	2-2dy	3,138	-0.476	0.118	37	0.096	0.153
nolic_sh	3-All	4,521	-0.515	0.101	36	0.082	0.131
nolic_sr	1-1dy	1,734	-0.464	0.165	54	0.139	0.203
nolic_sr	2-2dy	3,138	-0.423	0.133	35	0.108	0.174
nolic_sr	3-All	4,521	-0.492	0.114	39	0.094	0.147
othnwk_sh	1-1dy	1,734	-0.152	0.119	69	0.102	0.143
othnwk_sh	2-2dy	3,138	-0.263	0.093	64	0.080	0.113
othnwk_sh	3-All	4,521	-0.277	0.081	79	0.070	0.096
veh0_sh	1-1dy	1,734	-0.516	0.186	33	0.150	0.245
veh0_sh	2-2dy	3,138	-0.422	0.157	31	0.126	0.208
veh0_sh	3-All	4,521	-0.497	0.129	51	0.108	0.160
veh0_sr	1-1dy	1,734	-0.298	0.206	27	0.163	0.280
veh0_sr	2-2dy	3,138	-0.348	0.163	40	0.134	0.209
veh0_sr	3-All	4,521	-0.398	0.142	49	0.118	0.176
wkrs_sh	1-1dy	1,734	-0.280	0.125	63	0.107	0.152
wkrs_sh	2-2dy	3,138	-0.396	0.097	77	0.084	0.116
wkrs_sh	3-All	4,521	-0.381	0.088	108	0.078	0.102
wkrs_sr	1-1dy	1,734	-0.223	0.131	130	0.117	0.149
wkrs_sr	2-2dy	3,138	-0.237	0.108	38	0.088	0.139
wkrs_sr	3-All	4,521	-0.214	0.093	71	0.080	0.111

Table G-2-6 provides estimates of the intra-person correlation for the two-day and full files by comparing the standard errors for the three files. Included are the following fields:

- Total trips $n^{(1)}$, $n^{(2)}$, $n^{(3)}$: the total number of person-days that support the estimates from the one-day, two-day, and full files respectively;
- Jackknife standard errors: the square roots of the jackknife variances $v(\hat{\theta}^{(1)})$, $v(\hat{\theta}^{(2)})$, $v(\hat{\theta}^{(3)})$.
- Inverse person-days ratio: the ratio of the reciprocal of total person-days for the two-day and full files ($1/n^{(2)}$ and $1/n^{(3)}$ respectively) to the reciprocal of total person-days for the one-day file ($1/n^{(1)}$);

- Jackknife variance ratio: the ratio of the jackknife variance for the two-day and full files ($v(\hat{\theta}^{(2)})$ and $v(\hat{\theta}^{(3)})$ respectively) to the jackknife variance for the one-day file ($v(\hat{\theta}^{(1)})$);
- Roh calculation: the estimate of the within-person rate of homogeneity for the two-day and full files;
- Estimated within-person rate of homogeneity: equal to the mean of the two-day and full file a estimates;
- Pas design effect.

If the total trips were sampled in a simple random sample from some super-population of trips (i.e., with no clustering by persons or days), then there would be direct equality between the inverse person-days ratios and the jackknife variance ratios (i.e., $v(\hat{\theta}^{(2)})/v(\hat{\theta}^{(1)}) = \left(1/n^{(2)}\right)/\left(1/n^{(1)}\right), v(\hat{\theta}^{(3)})/v(\hat{\theta}^{(1)}) = \left(1/n^{(3)}\right)/\left(1/n^{(1)}\right)$).

The a estimates for the two-day file are computed using the following formula:

$$roh^{(2)} = \frac{\left\{2 * v(\hat{\theta}^{(2)})/v(\hat{\theta}^{(1)})\right\} - 1}{2 - 1}$$

This is an inversion of the formula

$$\frac{v(\hat{\theta}^{(2)})}{v(\hat{\theta}^{(1)})} = \frac{1 + roh * (T - 1)}{T}$$

from Section 1 (with $T = 2$ for the two-day file). The a estimates for the full file are computed using the following formula:

$$roh^{(3)} = \frac{\left\{2.6 * v(\hat{\theta}^{(3)})/v(\hat{\theta}^{(1)})\right\} - 1}{2.6 - 1}$$

This is an inversion of the formula

$$\frac{v(\hat{\theta}^{(3)})}{v(\hat{\theta}^{(1)})} = \frac{1 + roh * (T - 1)}{T}$$

from Section 1 (with $T = 2.6$ for the full file[22]). The estimated within-person $\widehat{roh}$ is computed as the average of $roh^{(2)}$ and $roh^{(3)}$. Finally the Pas design effect is computed as:

$$\widehat{deff}\left(\bar{y}^{(T)}\right) = \frac{1 + \widehat{roh} * (T - 1)}{T}$$

[22] This T value is computed as T=3 for starting collection days Monday, Tuesday, and Wednesday, and T=2 for starting collection days Thursday and Friday, with each of the five days having an equal chance of being assigned to the household.

Table G-2-6 (Part 1). Jackknife Standard Errors, *α* calculations, and design effects for Non-Worker Tour Generation Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
age35u_o	1-1dy	1,734	0.230	1.00	1.00		60.32%	1.00
age35u_o	2-2dy	3,138	0.205	0.55	0.80	59.33%	60.32%	0.80
age35u_o	3-All	4,521	0.200	0.38	0.76	61.30%	60.32%	0.76
age65p_sh	1-1dy	1,734	0.115	1.00	1.00		10.50%	1.00
age65p_sh	2-2dy	3,138	0.086	0.55	0.56	11.61%	10.50%	0.55
age65p_sh	3-All	4,521	0.077	0.38	0.44	9.39%	10.50%	0.45
asc2	1-1dy	1,734	3.016	1.00	1.00		24.16%	1.00
asc2	2-2dy	3,138	2.324	0.55	0.59	18.69%	24.16%	0.62
asc2	3-All	4,521	2.271	0.38	0.57	29.63%	24.16%	0.53
asc3	1-1dy	1,734	3.433	1.00	1.00		8.44%	1.00
asc3	2-2dy	3,138	2.531	0.55	0.54	8.76%	8.44%	0.54
asc3	3-All	4,521	2.263	0.38	0.43	8.12%	8.44%	0.44
asc4	1-1dy	1,734	0.133	1.00	1.00		20.84%	1.00
asc4	2-2dy	3,138	0.105	0.55	0.63	26.24%	20.84%	0.60
asc4	3-All	4,521	0.092	0.38	0.48	15.45%	20.84%	0.51
asc5	1-1dy	1,734	5.326	1.00	1.00		-3.64%	1.00
asc5	2-2dy	3,138	3.586	0.55	0.45	-9.35%	-3.64%	0.48
asc5	3-All	4,521	3.357	0.38	0.40	2.06%	-3.64%	0.36
asc6	1-1dy	1,734	3.018	1.00	1.00		23.75%	1.00
asc6	2-2dy	3,138	2.317	0.55	0.59	17.90%	23.75%	0.62
asc6	3-All	4,521	2.272	0.38	0.57	29.60%	23.75%	0.53
asc7	1-1dy	1,734	3.444	1.00	1.00		8.54%	1.00
asc7	2-2dy	3,138	2.540	0.55	0.54	8.80%	8.54%	0.54
asc7	3-All	4,521	2.273	0.38	0.44	8.27%	8.54%	0.44
asc8	1-1dy	1,734	5.292	1.00	1.00		-3.30%	1.00
asc8	2-2dy	3,138	3.566	0.55	0.45	-9.20%	-3.30%	0.48
asc8	3-All	4,521	3.350	0.38	0.40	2.60%	-3.30%	0.36

Table G-2-6 (Part 2). Jackknife Standard Errors, *α* calculations, and design effects for Non-Worker Tour Generation Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calculation	Estimated Within-Person *α*	Pas Design Effect
female_o	1-1dy	1,734	0.123	1.00	1.00		36.90%	1.00
female_o	2-2dy	3,138	0.103	0.55	0.70	39.55%	36.90%	0.68
female_o	3-All	4,521	0.095	0.38	0.60	34.26%	36.90%	0.61
female_sh	1-1dy	1,734	0.101	1.00	1.00		19.61%	1.00
female_sh	2-2dy	3,138	0.078	0.55	0.60	20.60%	19.61%	0.60
female_sh	3-All	4,521	0.071	0.38	0.50	18.62%	19.61%	0.51
female_sr	1-1dy	1,734	0.102	1.00	1.00		20.58%	1.00
female_sr	2-2dy	3,138	0.080	0.55	0.61	21.77%	20.58%	0.60
female_sr	3-All	4,521	0.073	0.38	0.50	19.39%	20.58%	0.51
hwylsum_sh	1-1dy	1,734	0.229	1.00	1.00		24.78%	1.00
hwylsum_sh	2-2dy	3,138	0.177	0.55	0.60	19.48%	24.78%	0.62
hwylsum_sh	3-All	4,521	0.173	0.38	0.57	30.09%	24.78%	0.54
hwylsum_sr	1-1dy	1,734	0.258	1.00	1.00		8.76%	1.00
hwylsum_sr	2-2dy	3,138	0.190	0.55	0.55	9.13%	8.76%	0.54
hwylsum_sr	3-All	4,521	0.170	0.38	0.44	8.39%	8.76%	0.44
inc3_o	1-1dy	1,734	0.162	1.00	1.00		44.80%	1.00
inc3_o	2-2dy	3,138	0.143	0.55	0.77	54.75%	44.80%	0.72
inc3_o	3-All	4,521	0.126	0.38	0.60	34.85%	44.80%	0.66
inc3_sr	1-1dy	1,734	0.167	1.00	1.00		8.95%	1.00
inc3_sr	2-2dy	3,138	0.122	0.55	0.54	7.69%	8.95%	0.54
inc3_sr	3-All	4,521	0.111	0.38	0.45	10.22%	8.95%	0.44
inc45_o	1-1dy	1,734	0.156	1.00	1.00		24.15%	1.00
inc45_o	2-2dy	3,138	0.126	0.55	0.64	28.95%	24.15%	0.62
inc45_o	3-All	4,521	0.111	0.38	0.50	19.36%	24.15%	0.53
inc45_sr	1-1dy	1,734	0.169	1.00	1.00		13.69%	1.00
inc45_sr	2-2dy	3,138	0.128	0.55	0.57	14.72%	13.69%	0.57
inc45_sr	3-All	4,521	0.115	0.38	0.46	12.66%	13.69%	0.47

Table G-2-6 (Part 3). Jackknife Standard Errors, *α* calculations, and design effects for Non-Worker Tour Generation Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
kids_o	1-1dy	1,734	0.177	1.00	1.00		42.01%	1.00
kids_o	2-2dy	3,138	0.153	0.55	0.74	48.40%	42.01%	0.71
kids_o	3-All	4,521	0.138	0.38	0.60	35.61%	42.01%	0.64
nolic_sh	1-1dy	1,734	0.146	1.00	1.00		22.19%	1.00
nolic_sh	2-2dy	3,138	0.118	0.55	0.65	30.07%	22.19%	0.61
nolic_sh	3-All	4,521	0.101	0.38	0.47	14.32%	22.19%	0.52
nolic_sr	1-1dy	1,734	0.165	1.00	1.00		22.66%	1.00
nolic_sr	2-2dy	3,138	0.133	0.55	0.65	29.96%	22.66%	0.61
nolic_sr	3-All	4,521	0.114	0.38	0.48	15.35%	22.66%	0.52
othnwk_sh	1-1dy	1,734	0.119	1.00	1.00		17.02%	1.00
othnwk_sh	2-2dy	3,138	0.093	0.55	0.61	22.43%	17.02%	0.59
othnwk_sh	3-All	4,521	0.081	0.38	0.46	11.61%	17.02%	0.49
veh0_sh	1-1dy	1,734	0.186	1.00	1.00		28.83%	1.00
veh0_sh	2-2dy	3,138	0.157	0.55	0.71	42.32%	28.83%	0.64
veh0_sh	3-All	4,521	0.129	0.38	0.48	15.35%	28.83%	0.56
veh0_sr	1-1dy	1,734	0.206	1.00	1.00		20.15%	1.00
veh0_sr	2-2dy	3,138	0.163	0.55	0.63	25.88%	20.15%	0.60
veh0_sr	3-All	4,521	0.142	0.38	0.47	14.42%	20.15%	0.51
wkrs_sh	1-1dy	1,734	0.125	1.00	1.00		19.36%	1.00
wkrs_sh	2-2dy	3,138	0.097	0.55	0.60	20.57%	19.36%	0.60
wkrs_sh	3-All	4,521	0.088	0.38	0.50	18.16%	19.36%	0.50
wkrs_sr	1-1dy	1,734	0.131	1.00	1.00		26.72%	1.00
wkrs_sr	2-2dy	3,138	0.108	0.55	0.67	34.27%	26.72%	0.63
wkrs_sr	3-All	4,521	0.093	0.38	0.50	19.17%	26.72%	0.55

Table G-2-7 presents the variance ratios and the estimated *α* factors for the comparison of the two-day file estimates and the full-file estimates to the one-day estimates, ordered by variance ratio. For the two-day file, the estimated *α*'s range from negative values to a high of 59%. The median value is 20%. There is a very wide range of *α*'s. For the full-file the estimated *α*'s range from 2% to 61%, with a median value of about 15%.

Table G-2-7. Variance ratios and α factors for comparison of the two-day file to the one-day files, comparison of the full file to the one-day file.

VR Parameter Estimates Two-day to One-day	Correspon-ding two-day to one-day α factor	VR Parameter Estimates Full-file to One-day	Correspon-ding full-file to one-day α factor
0.453	-9.35%	0.397	2.06%
0.454	-9.20%	0.401	2.60%
0.538	7.69%	0.435	8.12%
0.544	8.76%	0.436	8.27%
0.544	8.80%	0.436	8.39%
0.546	9.13%	0.442	9.39%
0.558	11.61%	0.447	10.22%
0.574	14.72%	0.456	11.61%
0.589	17.90%	0.463	12.66%
0.593	18.69%	0.473	14.32%
0.597	19.48%	0.473	14.42%
0.603	20.57%	0.479	15.35%
0.603	20.60%	0.479	15.35%
0.609	21.77%	0.480	15.45%
0.612	22.43%	0.496	18.16%
0.629	25.88%	0.499	18.62%
0.631	26.24%	0.503	19.17%
0.645	28.95%	0.504	19.36%
0.650	29.96%	0.504	19.39%
0.650	30.07%	0.567	29.60%
0.671	34.27%	0.567	29.63%
0.698	39.55%	0.570	30.09%
0.712	42.32%	0.595	34.26%
0.742	48.40%	0.599	34.85%
0.774	54.75%	0.604	35.61%
0.797	59.33%	0.762	61.30%

Table G-2-8 presents the eigenvalues from the variance ratio matrix $\widehat{\mathrm{vr}}\left(\widehat{\boldsymbol{\beta}}^{(2)}, \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right) = \left\{v_J\left(\widehat{\boldsymbol{\beta}}^{(1)}\right)\right\}^{-\mathbf{1}} v_J\left(\widehat{\boldsymbol{\beta}}^{(2)}\right)$ and $\widehat{\mathrm{vr}}\left(\widehat{\boldsymbol{\beta}}^{(3)}, \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right) = \left\{v_J\left(\widehat{\boldsymbol{\beta}}^{(1)}\right)\right\}^{-\mathbf{1}} v_J\left(\widehat{\boldsymbol{\beta}}^{(3)}\right)$ (see Appendix F). These eigenvalues should generally track the simple univariate parameter variance ratios. In particular the product of the eigenvalues should be fairly close to the product of the univariate ratios as given in Table G-2-7. As one can see from Table G-2-8, this is in fact the case. The geometric means of eigenvalues and univariate parameter ratios are very close for both files. The eigenvalues have a larger range, but are centered around the same geometric mean.

Table G-2-8. Variance ratios and eigenvalues for comparison of the two-day file to the one-day files, comparison of the full file to the one-day file.

VR Parameter Estimates Two-day to One-day	Eigenvalues Two-day to One-day	VR Parameter Estimates Full-file to One-day	Eigenvalues Full-file to One-Day
0.453	0.287	0.397	0.228
0.454	0.372	0.401	0.275
0.538	0.398	0.435	0.300
0.544	0.415	0.436	0.310
0.544	0.450	0.436	0.329
0.546	0.480	0.442	0.350
0.558	0.508	0.447	0.366
0.574	0.533	0.456	0.398
0.589	0.541	0.463	0.417
0.593	0.577	0.473	0.421
0.597	0.602	0.473	0.440
0.603	0.620	0.479	0.466
0.603	0.633	0.479	0.489
0.609	0.663	0.480	0.518
0.612	0.694	0.496	0.534
0.629	0.706	0.499	0.550
0.631	0.733	0.503	0.591
0.645	0.761	0.504	0.659
0.650	0.792	0.504	0.676
0.650	0.799	0.567	0.694
0.671	0.811	0.567	0.722
0.698	0.857	0.570	0.754
0.712	0.879	0.595	0.777
0.742	0.963	0.599	0.806
0.774	0.995	0.604	0.880
0.797	1.089	0.762	0.973
Geometric Mean	Geometric Mean	Geometric Mean	Geometric Mean
0.610	0.628	0.497	0.499

Appendix G-3. Model Estimation Results for Worker Tour Generation

The tour generation model jointly predicts the number and purpose of tours made by individual. The models are specific to person type, and this section describes the models estimated for workers. Appendix I enumerates the possible alternatives, with workers restricted to the first 24 alternatives.

The models are estimated from each of the three files using weighted maximum likelihood estimation, and the jackknife weights were utilized to generate jackknife estimates for each of the three files. The weights for the base estimation are all 1, with the weights varying for each jackknife replicate, but still 1 for most observations. A weighted MLE parameter vector estimate and jackknife variance matrices are generated for each file, and comparisons made of the variance matrices. Tables G-3-1, G-3-2, and G-3-3 provide the parameter estimates from the full data set, the two-day data set, and the one-day data set respectively.

The model specification starts with a set of alternative specific constants, using stay at home as the reference alternative. Next, a set of demographic variables is related to the propensity to participate in certain types of tours. For example, part time workers are less likely to go to work on the travel day, and higher income workers are more likely to go to work. The presence and type of other household members affects the participation in certain types of tours, as does the highway logsum.

Table G-3-1. Parameter Estimates for Work Tour Generation Model: Full Data Set

Description	Alternative includes:	Coefficient estimate	Jackknife T-statistic
Alternative-Specific Constant	SH	2.1635	0.97
	SR	-5.3451	-3.55
	O	-2.2598	-9.80
	SH-SR	-2.7983	-1.01
	SH-O	0.3514	0.16
	SR-O	-6.7980	-4.43
	SH-SR-O	-4.2568	-1.53
	W	-1.8354	-5.50
	W-SH	-0.5804	-0.26
	W-SR	-7.4584	-4.83
	W-O	-4.6413	-10.76
	W-SH-SR	-6.7156	-2.39
	W-SH-O	-3.3763	-1.48
	W-SR-O	-10.2071	-6.45
	W-SH-SR-O	-9.7774	-1.06
	W-WB	-3.4729	-10.17
	W-WB-SH	-2.1588	-0.95
	W-WB-SR	-8.8868	-5.71
	W-WB-O	-6.2602	-13.20
	W-WB-SH-SR	-8.1454	-2.76
	W-WB-SH-O	-8.1454	-2.76
	W-WB-SR-O	-12.2336	-7.47
	W-WB-SH-SR-O	-9.5654	-1.02
Person is a part-time worker	Work tour	-0.7715	-7.80
	Work-based subtour	-0.8573	-2.80
Person is female	Work tour	0.2475	3.86
	Work-based subtour	-0.2933	-2.04
	Shopping tour	0.1384	1.81
	Social/recreation tour	-0.1177	-1.92
Person is under age 35	Shopping tour	-0.3014	-2.13
	Other tour	-0.3251	-3.02
Person is age 55 or older	Shopping tour	0.2673	2.89
Income $10-24k	Work tour	0.7695	2.20
	Other tour	0.8161	3.27
Income $25-49k	Work tour	1.3600	4.07
	Other tour	0.6129	2.52
Income $50-99k	Work tour	1.4352	4.32
	Other tour	0.6097	2.51
Income $100k+	Work tour	1.2761	3.88
	Other tour	0.6260	2.56
0 Vehicles	Work tour	-0.5450	-1.99
	Shopping tour	-0.8831	-2.80
	Social/recreation tour	-1.1039	-3.99
0 < Vehicles < Workers	Work tour	-0.3626	-3.02
	Shopping tour	-0.2420	-1.73
Other workers present in HH	Shopping tour	-0.2388	-2.57
	Social/recreation tour	-0.1257	-1.68
Non-working adults present in HH	Shopping tour	-0.2519	-2.45
	Social/recreation tour	-0.2338	-3.04
Children present in HH	Shopping tour	0.1779	1.64
	Social/recreation tour	0.3299	4.43
	Other tour	0.4215	5.61
Highway logsum	Shopping tour	-0.3059	-1.83
	Social/recreation tour	0.3164	2.80

Table G-3-2. Parameter Estimates for Work Tour Generation Model: Two-Day Data Set

Description	Alternative includes:	Coefficient estimate	Jackknife T-statistic
Alternative-Specific Constant	SH	1.8734	1.18
	SR	-6.8726	-3.71
	O	-2.0909	-8.23
	SH-SR	-4.7105	-1.95
	SH-O	0.1471	0.09
	SR-O	-8.2478	-4.38
	SH-SR-O	-6.1258	-2.52
	W	-1.3507	-3.89
	W-SH	-0.5383	-0.34
	W-SR	-8.6178	-4.56
	W-O	-4.0572	-9.16
	W-SH-SR	-8.2825	-3.33
	W-SH-O	-3.1920	-1.91
	W-SR-O	-11.3131	-5.86
	W-SH-SR-O	-11.0376	-1.18
	W-WB	-2.8980	-8.13
	W-WB-SH	-1.9049	-1.17
	W-WB-SR	-10.0037	-5.30
	W-WB-O	-5.6499	-11.19
	W-WB-SH-SR	-9.4502	-3.56
	W-WB-SH-O	-9.4502	-3.56
	W-WB-SR-O	-13.0724	-6.64
	W-WB-SH-SR-O	-13.0724	-6.64
Person is a part-time worker	Work tour	-0.8289	-7.86
	Work-based subtour	-0.8019	-2.45
Person is female	Work tour	0.2596	3.69
	Work-based subtour	-0.3496	-2.36
	Shopping tour	0.0875	1.00
	Social/recreation tour	-0.1637	-2.35
Person is under age 35	Shopping tour	-0.2864	-1.86
	Other tour	-0.2589	-2.08
Person is age 55 or older	Shopping tour	0.1941	1.95
Income $10-24k	Work tour	0.4994	1.38
	Other tour	0.8253	3.02
Income $25-49k	Work tour	1.0720	3.08
	Other tour	0.5295	2.01
Income $50-99k	Work tour	1.1444	3.33
	Other tour	0.4722	1.79
Income $100k+	Work tour	1.0293	3.01
	Other tour	0.5649	2.09
0 Vehicles	Work tour	-0.5519	-1.98
	Shopping tour	-0.6862	-1.96
	Social/recreation tour	-0.9919	-3.46
0 < Vehicles < Workers	Work tour	-0.3704	-2.73
	Shopping tour	-0.2181	-1.35
Other workers present in HH	Shopping tour	-0.1666	-1.60
	Social/recreation tour	-0.0609	-0.71
Non-working adults present in HH	Shopping tour	-0.1904	-1.70
	Social/recreation tour	-0.1623	-1.84
Children present in HH	Shopping tour	0.1054	0.90
	Social/recreation tour	0.3705	4.42
	Other tour	0.4460	5.36
Highway logsum	Shopping tour	-0.2704	-2.27
	Social/recreation tour	0.4397	3.14

Table G-3-3. Parameter Estimates for Work Tour Generation Model: One-Day Data Set

Description	Alternative includes:	Coefficient estimate	Jackknife T-statistic
Alternative-Specific Constant	SH	2.7641	1.54
	SR	-7.4541	-3.20
	O	-1.5677	-4.41
	SH-SR	-4.7468	-1.58
	SH-O	1.1185	0.62
	SR-O	-8.6814	-3.68
	SH-SR-O	-6.0528	-2.02
	W	-0.8532	-2.09
	W-SH	0.4247	0.24
	W-SR	-9.1037	-3.88
	W-O	-3.4299	-5.94
	W-SH-SR	-8.0848	-2.63
	W-SH-O	-2.1861	-1.17
	W-SR-O	-11.7595	-4.93
	W-SH-SR-O	-10.6070	-1.12
	W-WB	-2.5015	-5.98
	W-WB-SH	-1.2931	-0.70
	W-WB-SR	-10.7184	-4.61
	W-WB-O	-4.8270	-7.62
	W-WB-SH-SR	-9.1731	-2.89
	W-WB-SH-O	-9.1731	-2.89
	W-WB-SR-O	-13.1183	-5.44
	W-WB-SH-SR-O	-13.1183	-5.44
Person is a part-time worker	Work tour	-0.8081	-6.77
	Work-based subtour	-1.1069	-2.60
Person is female	Work tour	0.1936	2.23
	Work-based subtour	-0.1814	-1.08
	Shopping tour	0.0404	0.36
	Social/recreation tour	-0.1717	-2.14
Person is under age 35	Shopping tour	-0.3114	-1.62
	Other tour	-0.4385	-2.67
Person is age 55 or older	Shopping tour	0.0680	0.57
Income $10-24k	Work tour	0.6199	1.46
	Other tour	0.7022	1.81
Income $25-49k	Work tour	1.1774	2.87
	Other tour	0.3436	0.95
Income $50-99k	Work tour	1.2563	3.11
	Other tour	0.3603	1.00
Income $100k+	Work tour	1.0906	2.69
	Other tour	0.3558	0.97
0 Vehicles	Work tour	-0.6892	-1.97
	Shopping tour	-0.7432	-1.44
	Social/recreation tour	-0.7484	-2.24
0 < Vehicles < Workers	Work tour	-0.4136	-2.36
	Shopping tour	-0.1544	-0.74
Other workers present in HH	Shopping tour	-0.1941	-1.47
	Social/recreation tour	0.0799	0.75
Non-working adults present in HH	Shopping tour	-0.2163	-1.55
	Social/recreation tour	-0.0412	-0.36
Children present in HH	Shopping tour	0.1400	1.04
	Social/recreation tour	0.3263	3.21
	Other tour	0.5919	5.62
Highway logsum	Shopping tour	-0.2980	-2.19
	Social/recreation tour	0.5109	2.91

Table G-3-5 below present the parameter estimates as they are given in tables G-3-1, G-3-2, and G-3-3, but with the three day-file estimates together for each parameter. Also included are jackknife standard errors for the parameter estimates (Tables G-3-1 through G-3-3 present the t-statistics, which are the parameter estimates divided by the standard errors) Table G-3-5 presents degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). Note that in these tables we provided all estimates, even when the degrees of freedom were smaller than 30, to keep the parameter vector whole. But the variance calculations with low degrees of freedom should be treated with skepticism. Table G-3-4 presents a listing of short parameter names used in Tables G-3-5 and G-3-6 (to save space), linking back to the parameter descriptions in Tables G-3-1 through G-3-3.

Table G-3-4 (Part 1). Comparison of Full Parameter Names to Short Parameter Names for Worker Tour Generation Model Parameters

Full Parameter Name		Short Parameter Name
Description	Alternative includes:	Label
Alternative-Specific Constant	SH	asc2
	SR	asc3
	O	asc4
	SH-SR	asc5
	SH-O	asc6
	SR-O	asc7
	SH-SR-O	asc8
	W	asc9
	W-SH	asc10
	W-SR	asc11
	W-O	asc12
	W-SH-SR	asc13
	W-SH-O	asc14
	W-SR-O	asc15
	W-SH-SR-O	asc16
	W-WB	asc17
	W-WB-SH	asc18
	W-WB-SR	asc19
	W-WB-O	asc20
	W-WB-SH-SR	asc21_22
	W-WB-SH-O	asc21_22
	W-WB-SR-O/W-WB-SH-SR-O	asc23_24
Person is a part-time worker	Work tour	partime_w
	Work-based subtour	partime_wb
Person is female	Work tour	female_w
	Work-based subtour	female_wb
	Shopping tour	female_sh
	Social/recreation tour	female_sr
Person is under age 35	Shopping tour	age35u_sh
	Other tour	age35u_o
Person is age 55 or older	Shopping tour	age55p_sh

Table G-3-4 (Part 2). Comparison of Full Parameter Names to Short Parameter Names

Full Parameter Name		Short Parameter Name
Description	Alternative includes:	Label
Income $10-24k	Work tour	inc2_w
	Other tour	inc2_o
Income $25-49k	Work tour	inc3_w
	Other tour	inc3_o
Income $50-99k	Work tour	inc4_w
	Other tour	inc4_o
Income $100k+	Work tour	inc5_w
	Other tour	inc5_o
0 Vehicles	Work tour	veh0_w
	Shopping tour	veh0_sh
	Social/recreation tour	veh0_sr
0 < Vehicles < Workers	Work tour	vehlwk_w
	Shopping tour	vehlwk_sh
Other workers present in HH	Shopping tour	othwkrs_sh
	Social/recreation tour	othwkrs_sr
Non-working adults present in HH	Shopping tour	nwkrs_sh
	Social/recreation tour	nwkrs_sr
Children present in HH	Shopping tour	kids_sh
	Social/recreation tour	kids_sr
	Other tour	kids_o
Highway logsum	Shopping tour	hwylsum_sh
	Social/recreation tour	hwylsum_sr

Table G-3-5 (Part 1). Parameter Estimates for Worker Tour Generation Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
age35u_o	1-1dy	2,666	-0.438	0.164	27	0.130	0.224
age35u_o	2-2dy	4,794	-0.259	0.124	55	0.105	0.153
age35u_o	3-All	6,936	-0.325	0.108	71	0.093	0.129
age35u_sh	1-1dy	2,666	-0.311	0.192	56	0.162	0.236
age35u_sh	2-2dy	4,794	-0.286	0.154	50	0.129	0.191
age35u_sh	3-All	6,936	-0.301	0.142	42	0.117	0.180
age55p_sh	1-1dy	2,666	0.068	0.119	122	0.106	0.136
age55p_sh	2-2dy	4,794	0.194	0.100	135	0.089	0.113
age55p_sh	3-All	6,936	0.267	0.092	160	0.083	0.104
asc10	1-1dy	2,666	0.425	1.788	56	1.509	2.193
asc10	2-2dy	4,794	-0.538	1.600	30	1.279	2.139
asc10	3-All	6,936	-0.580	2.236	45	1.854	2.816
asc11	1-1dy	2,666	-9.104	2.346	52	1.969	2.902
asc11	2-2dy	4,794	-8.618	1.889	51	1.583	2.342
asc11	3-All	6,936	-7.458	1.545	45	1.281	1.946
asc12	1-1dy	2,666	-3.430	0.577	30	0.461	0.772
asc12	2-2dy	4,794	-4.057	0.443	21	0.341	0.633
asc12	3-All	6,936	-4.641	0.432	22	0.334	0.611
asc13	1-1dy	2,666	-8.085	3.076	49	2.570	3.834
asc13	2-2dy	4,794	-8.282	2.487	61	2.114	3.023
asc13	3-All	6,936	-6.716	2.815	56	2.377	3.454
asc14	1-1dy	2,666	-2.186	1.871	62	1.592	2.270
asc14	2-2dy	4,794	-3.192	1.669	31	1.338	2.219
asc14	3-All	6,936	-3.376	2.287	40	1.878	2.926
asc15	1-1dy	2,666	-11.760	2.384	53	2.004	2.943
asc15	2-2dy	4,794	-11.313	1.930	53	1.623	2.383
asc15	3-All	6,936	-10.207	1.584	48	1.321	1.978
asc16	1-1dy	2,666	-10.607	9.441	3	5.348	35.200
asc16	2-2dy	4,794	-11.038	9.382	3	5.315	34.982
asc16	3-All	6,936	-9.777	9.210	3	5.217	34.341
asc17	1-1dy	2,666	-2.501	0.418	19	0.318	0.611
asc17	2-2dy	4,794	-2.898	0.356	37	0.290	0.461
asc17	3-All	6,936	-3.473	0.341	34	0.276	0.447
asc18	1-1dy	2,666	-1.293	1.835	54	1.545	2.260
asc18	2-2dy	4,794	-1.905	1.630	31	1.307	2.167
asc18	3-All	6,936	-2.159	2.263	43	1.870	2.867

Table G-3-5 (Part 2). Parameter Estimates for Worker Tour Generation Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
asc19	1-1dy	2,666	-10.718	2.324	52	1.950	2.875
asc19	2-2dy	4,794	-10.004	1.889	53	1.588	2.332
asc19	3-All	6,936	-8.887	1.557	44	1.289	1.967
asc2	1-1dy	2,666	2.764	1.790	38	1.463	2.307
asc2	2-2dy	4,794	1.873	1.586	24	1.238	2.206
asc2	3-All	6,936	2.163	2.221	46	1.846	2.790
asc20	1-1dy	2,666	-4.827	0.633	39	0.519	0.813
asc20	2-2dy	4,794	-5.650	0.505	33	0.407	0.665
asc20	3-All	6,936	-6.260	0.474	31	0.380	0.631
asc21_22	1-1dy	2,666	-9.173	3.179	64	2.711	3.844
asc21_22	2-2dy	4,794	-9.450	2.654	72	2.282	3.171
asc21_22	3-All	6,936	-8.145	2.948	57	2.492	3.609
asc23_24	1-1dy	2,666	-13.118	2.413	53	2.029	2.979
asc23_24	2-2dy	4,794	-13.072	1.967	56	1.661	2.413
asc3	1-1dy	2,666	-7.454	2.328	59	1.973	2.839
asc3	2-2dy	4,794	-6.873	1.854	54	1.561	2.284
asc3	3-All	6,936	-5.345	1.505	45	1.248	1.896
asc4	1-1dy	2,666	-1.568	0.356	44	0.294	0.449
asc4	2-2dy	4,794	-2.091	0.254	32	0.204	0.336
asc4	3-All	6,936	-2.260	0.231	27	0.182	0.314
asc5	1-1dy	2,666	-4.747	3.001	60	2.547	3.653
asc5	2-2dy	4,794	-4.711	2.413	66	2.062	2.908
asc5	3-All	6,936	-2.798	2.766	54	2.328	3.407
asc6	1-1dy	2,666	1.118	1.817	44	1.504	2.295
asc6	2-2dy	4,794	0.147	1.604	27	1.268	2.183
asc6	3-All	6,936	0.351	2.240	44	1.855	2.830
asc7	1-1dy	2,666	-8.681	2.356	59	1.997	2.874
asc7	2-2dy	4,794	-8.248	1.882	54	1.584	2.318
asc7	3-All	6,936	-6.798	1.535	46	1.275	1.928
asc8	1-1dy	2,666	-6.053	3.002	61	2.551	3.648
asc8	2-2dy	4,794	-6.126	2.430	69	2.083	2.916
asc8	3-All	6,936	-4.257	2.784	54	2.344	3.430
asc9	1-1dy	2,666	-0.853	0.409	15	0.302	0.633
asc9	2-2dy	4,794	-1.351	0.347	29	0.276	0.467
asc9	3-All	6,936	-1.835	0.333	28	0.265	0.451

Table G-3-5 (Part 3). Parameter Estimates for Worker Tour Generation Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
female_sh	1-1dy	2,666	0.040	0.112	115	0.099	0.128
female_sh	2-2dy	4,794	0.088	0.088	126	0.078	0.100
female_sh	3-All	6,936	0.138	0.077	104	0.067	0.089
female_sr	1-1dy	2,666	-0.172	0.080	96	0.070	0.094
female_sr	2-2dy	4,794	-0.164	0.070	111	0.062	0.080
female_sr	3-All	6,936	-0.118	0.061	113	0.054	0.070
female_w	1-1dy	2,666	0.194	0.087	83	0.075	0.102
female_w	2-2dy	4,794	0.260	0.070	56	0.059	0.086
female_w	3-All	6,936	0.248	0.064	66	0.055	0.077
female_wb	1-1dy	2,666	-0.181	0.168	76	0.145	0.200
female_wb	2-2dy	4,794	-0.350	0.148	83	0.129	0.175
female_wb	3-All	6,936	-0.293	0.144	101	0.127	0.167
hwylsum_sh	1-1dy	2,666	-0.298	0.136	37	0.111	0.176
hwylsum_sh	2-2dy	4,794	-0.270	0.119	23	0.093	0.167
hwylsum_sh	3-All	6,936	-0.306	0.167	45	0.139	0.211
hwylsum_sr	1-1dy	2,666	0.511	0.175	56	0.148	0.215
hwylsum_sr	2-2dy	4,794	0.440	0.140	54	0.118	0.172
hwylsum_sr	3-All	6,936	0.316	0.113	46	0.094	0.142
inc2_o	1-1dy	2,666	0.702	0.388	65	0.331	0.468
inc2_o	2-2dy	4,794	0.825	0.273	46	0.227	0.343
inc2_o	3-All	6,936	0.816	0.250	38	0.204	0.322
inc2_w	1-1dy	2,666	0.620	0.423	24	0.331	0.589
inc2_w	2-2dy	4,794	0.499	0.361	39	0.296	0.464
inc2_w	3-All	6,936	0.770	0.350	41	0.288	0.446

Table G-3-5 (Part 4). Parameter Estimates for Worker Tour Generation Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
inc3_o	1-1dy	2,666	0.344	0.360	57	0.305	0.441
inc3_o	2-2dy	4,794	0.530	0.264	41	0.217	0.336
inc3_o	3-All	6,936	0.613	0.243	31	0.195	0.323
inc3_w	1-1dy	2,666	1.177	0.411	16	0.306	0.625
inc3_w	2-2dy	4,794	1.072	0.349	31	0.279	0.463
inc3_w	3-All	6,936	1.360	0.334	32	0.269	0.442
inc4_o	1-1dy	2,666	0.360	0.361	50	0.302	0.449
inc4_o	2-2dy	4,794	0.472	0.264	34	0.214	0.347
inc4_o	3-All	6,936	0.610	0.243	27	0.192	0.330
inc4_w	1-1dy	2,666	1.256	0.404	16	0.301	0.614
inc4_w	2-2dy	4,794	1.144	0.343	28	0.272	0.464
inc4_w	3-All	6,936	1.435	0.333	24	0.260	0.463
inc5_o	1-1dy	2,666	0.356	0.366	56	0.309	0.449
inc5_o	2-2dy	4,794	0.565	0.270	45	0.224	0.340
inc5_o	3-All	6,936	0.626	0.244	34	0.198	0.320
inc5_w	1-1dy	2,666	1.091	0.405	17	0.304	0.608
inc5_w	2-2dy	4,794	1.029	0.341	31	0.274	0.454
inc5_w	3-All	6,936	1.276	0.329	28	0.261	0.445
kids_o	1-1dy	2,666	0.592	0.105	96	0.092	0.123
kids_o	2-2dy	4,794	0.446	0.083	101	0.073	0.097
kids_o	3-All	6,936	0.421	0.075	107	0.066	0.087
kids_sh	1-1dy	2,666	0.140	0.135	135	0.121	0.153
kids_sh	2-2dy	4,794	0.105	0.117	116	0.104	0.134
kids_sh	3-All	6,936	0.178	0.109	105	0.096	0.125
kids_sr	1-1dy	2,666	0.326	0.102	115	0.090	0.117
kids_sr	2-2dy	4,794	0.371	0.084	88	0.073	0.098
kids_sr	3-All	6,936	0.330	0.074	108	0.066	0.086

Table G-3-5 (Part 5). Parameter Estimates for Worker Tour Generation Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for Std Err LB	CI for Std Err UB
nwkrs_sh	1-1dy	2,666	-0.216	0.140	92	0.122	0.163
nwkrs_sh	2-2dy	4,794	-0.190	0.112	100	0.098	0.130
nwkrs_sh	3-All	6,936	-0.252	0.103	136	0.092	0.117
nwkrs_sr	1-1dy	2,666	-0.041	0.114	61	0.097	0.138
nwkrs_sr	2-2dy	4,794	-0.162	0.088	75	0.076	0.105
nwkrs_sr	3-All	6,936	-0.234	0.077	64	0.066	0.093
othwkrs_sh	1-1dy	2,666	-0.194	0.132	80	0.115	0.157
othwkrs_sh	2-2dy	4,794	-0.167	0.104	95	0.091	0.121
othwkrs_sh	3-All	6,936	-0.239	0.093	83	0.081	0.110
othwkrs_sr	1-1dy	2,666	0.080	0.106	71	0.091	0.127
othwkrs_sr	2-2dy	4,794	-0.061	0.086	41	0.070	0.109
othwkrs_sr	3-All	6,936	-0.126	0.075	45	0.062	0.094
partime_w	1-1dy	2,666	-0.808	0.119	83	0.104	0.141
partime_w	2-2dy	4,794	-0.829	0.105	125	0.094	0.120
partime_w	3-All	6,936	-0.771	0.099	112	0.087	0.114
partime_wb	1-1dy	2,666	-1.107	0.426	19	0.324	0.623
partime_wb	2-2dy	4,794	-0.802	0.328	20	0.251	0.473
partime_wb	3-All	6,936	-0.857	0.306	27	0.242	0.416
veh0_sh	1-1dy	2,666	-0.743	0.517	7	0.342	1.052
veh0_sh	2-2dy	4,794	-0.686	0.349	12	0.250	0.577
veh0_sh	3-All	6,936	-0.883	0.315	11	0.223	0.536
veh0_sr	1-1dy	2,666	-0.748	0.334	13	0.242	0.538
veh0_sr	2-2dy	4,794	-0.992	0.287	14	0.210	0.453
veh0_sr	3-All	6,936	-1.104	0.277	8	0.187	0.530
veh0_w	1-1dy	2,666	-0.689	0.350	3	0.198	1.306
veh0_w	2-2dy	4,794	-0.552	0.278	4	0.167	0.800
veh0_w	3-All	6,936	-0.545	0.274	3	0.155	1.023
vehlwk_sh	1-1dy	2,666	-0.154	0.210	77	0.181	0.249
vehlwk_sh	2-2dy	4,794	-0.218	0.161	114	0.143	0.185
vehlwk_sh	3-All	6,936	-0.242	0.140	122	0.124	0.160
vehlwk_w	1-1dy	2,666	-0.414	0.176	52	0.147	0.217
vehlwk_w	2-2dy	4,794	-0.370	0.136	56	0.115	0.167
vehlwk_w	3-All	6,936	-0.363	0.120	56	0.101	0.147

Table G-3-6 below presents the variance ratios, *a* factors, and Pas design effects for the worker tour generation model, following the formulas as given preceding Table G-2-6. It should be noted that the degrees of freedom are not large for most of the parameter standard errors, so there is a lot of noise in these variance calculations that should lead to caution in interpreting the results.

Table G-3-6 (Part 1). Jackknife Standard Errors, *a* calculations, and design effects for Worker Tour Generation Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*a* Calculation	Estimated Within-Person *a*	Pas Design Effect
age35u_o	1-1dy	2,666	0.164	1.00	1.00		10.69%	1.00
age35u_o	2-2dy	4,794	0.124	0.56	0.57	14.17%	10.69%	0.55
age35u_o	3-All	6,936	0.108	0.38	0.43	7.21%	10.69%	0.45
age35u_sh	1-1dy	2,666	0.192	1.00	1.00		26.84%	1.00
age35u_sh	2-2dy	4,794	0.154	0.56	0.64	28.03%	26.84%	0.63
age35u_sh	3-All	6,936	0.142	0.38	0.54	25.64%	26.84%	0.55
age55p_sh	1-1dy	2,666	0.119	1.00	1.00		37.39%	1.00
age55p_sh	2-2dy	4,794	0.100	0.56	0.70	39.67%	37.39%	0.69
age55p_sh	3-All	6,936	0.092	0.38	0.60	35.12%	37.39%	0.61
asc10	1-1dy	2,666	1.788	1.00	1.00		80.11%	1.00
asc10	2-2dy	4,794	1.600	0.56	0.80	60.23%	80.11%	0.90
asc10	3-All	6,936	2.236	0.38	1.56	100.00%	80.11%	0.88
asc11	1-1dy	2,666	2.346	1.00	1.00		18.80%	1.00
asc11	2-2dy	4,794	1.889	0.56	0.65	29.63%	18.80%	0.59
asc11	3-All	6,936	1.545	0.38	0.43	7.97%	18.80%	0.50
asc12	1-1dy	2,666	0.577	1.00	1.00		23.00%	1.00
asc12	2-2dy	4,794	0.443	0.56	0.59	17.73%	23.00%	0.62
asc12	3-All	6,936	0.432	0.38	0.56	28.28%	23.00%	0.53
asc13	1-1dy	2,666	3.076	1.00	1.00		52.18%	1.00
asc13	2-2dy	4,794	2.487	0.56	0.65	30.76%	52.18%	0.76
asc13	3-All	6,936	2.815	0.38	0.84	73.61%	52.18%	0.71
asc14	1-1dy	2,666	1.871	1.00	1.00		79.58%	1.00
asc14	2-2dy	4,794	1.669	0.56	0.80	59.16%	79.58%	0.90
asc14	3-All	6,936	2.287	0.38	1.49	100.00%	79.58%	0.87
asc15	1-1dy	2,666	2.384	1.00	1.00		20.18%	1.00
asc15	2-2dy	4,794	1.930	0.56	0.66	31.13%	20.18%	0.60
asc15	3-All	6,936	1.584	0.38	0.44	9.23%	20.18%	0.51
asc16	1-1dy	2,666	9.441	1.00	1.00		94.85%	1.00
asc16	2-2dy	4,794	9.382	0.56	0.99	97.54%	94.85%	0.97
asc16	3-All	6,936	9.210	0.38	0.95	92.16%	94.85%	0.97
asc17	1-1dy	2,666	0.418	1.00	1.00		45.38%	1.00
asc17	2-2dy	4,794	0.356	0.56	0.73	45.01%	45.38%	0.73
asc17	3-All	6,936	0.341	0.38	0.67	45.75%	45.38%	0.66
asc18	1-1dy	2,666	1.835	1.00	1.00		78.95%	1.00
asc18	2-2dy	4,794	1.630	0.56	0.79	57.89%	78.95%	0.89
asc18	3-All	6,936	2.263	0.38	1.52	100.00%	78.95%	0.87

Table G-3-6 (Part 2). Jackknife Standard Errors, *a* calculations, and design effects for Worker Tour Generation Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*a* Calcu-lation	Estimated Within-Person *a*	Pas Design Effect
asc19	1-1dy	2,666	2.324	1.00	1.00		21.29%	1.00
asc19	2-2dy	4,794	1.889	0.56	0.66	32.10%	21.29%	0.61
asc19	3-All	6,936	1.557	0.38	0.45	10.49%	21.29%	0.52
asc2	1-1dy	2,666	1.790	1.00	1.00		78.48%	1.00
asc2	2-2dy	4,794	1.586	0.56	0.78	56.96%	78.48%	0.89
asc2	3-All	6,936	2.221	0.38	1.54	100.00%	78.48%	0.87
asc20	1-1dy	2,666	0.633	1.00	1.00		27.86%	1.00
asc20	2-2dy	4,794	0.505	0.56	0.64	27.09%	27.86%	0.64
asc20	3-All	6,936	0.474	0.38	0.56	28.62%	27.86%	0.56
asc21_22	1-1dy	2,666	3.179	1.00	1.00		58.32%	1.00
asc21_22	2-2dy	4,794	2.654	0.56	0.70	39.40%	58.32%	0.79
asc21_22	3-All	6,936	2.948	0.38	0.86	77.24%	58.32%	0.74
asc23_24	1-1dy	2,666	2.413	1.00	1.00		32.91%	1.00
asc23_24	2-2dy	4,794	1.967	0.56	0.66	32.91%	32.91%	0.66
asc3	1-1dy	2,666	2.328	1.00	1.00		16.15%	1.00
asc3	2-2dy	4,794	1.854	0.56	0.63	26.86%	16.15%	0.58
asc3	3-All	6,936	1.505	0.38	0.42	5.44%	16.15%	0.48
asc4	1-1dy	2,666	0.356	1.00	1.00		3.93%	1.00
asc4	2-2dy	4,794	0.254	0.56	0.51	1.97%	3.93%	0.52
asc4	3-All	6,936	0.231	0.38	0.42	5.89%	3.93%	0.41
asc5	1-1dy	2,666	3.001	1.00	1.00		52.41%	1.00
asc5	2-2dy	4,794	2.413	0.56	0.65	29.31%	52.41%	0.76
asc5	3-All	6,936	2.766	0.38	0.85	75.51%	52.41%	0.71
asc6	1-1dy	2,666	1.817	1.00	1.00		77.93%	1.00
asc6	2-2dy	4,794	1.604	0.56	0.78	55.87%	77.93%	0.89
asc6	3-All	6,936	2.240	0.38	1.52	100.00%	77.93%	0.86
asc7	1-1dy	2,666	2.356	1.00	1.00		17.02%	1.00
asc7	2-2dy	4,794	1.882	0.56	0.64	27.56%	17.02%	0.59
asc7	3-All	6,936	1.535	0.38	0.42	6.47%	17.02%	0.49
asc8	1-1dy	2,666	3.002	1.00	1.00		54.21%	1.00
asc8	2-2dy	4,794	2.430	0.56	0.66	31.08%	54.21%	0.77
asc8	3-All	6,936	2.784	0.38	0.86	77.33%	54.21%	0.72
asc9	1-1dy	2,666	0.409	1.00	1.00		44.79%	1.00
asc9	2-2dy	4,794	0.347	0.56	0.72	44.06%	44.79%	0.72
asc9	3-All	6,936	0.333	0.38	0.66	45.51%	44.79%	0.66

Table G-3-6 (Part 3). Jackknife Standard Errors, *a* calculations, and design effects for Worker Tour Generation Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*a* Calculation	Estimated Within-Person *a*	Pas Design Effect
female_sh	1-1dy	2,666	0.112	1.00	1.00		19.01%	1.00
female_sh	2-2dy	4,794	0.088	0.56	0.62	23.86%	19.01%	0.60
female_sh	3-All	6,936	0.077	0.38	0.47	14.17%	19.01%	0.50
female_sr	1-1dy	2,666	0.080	1.00	1.00		41.26%	1.00
female_sr	2-2dy	4,794	0.070	0.56	0.75	50.78%	41.26%	0.71
female_sr	3-All	6,936	0.061	0.38	0.58	31.73%	41.26%	0.64
female_w	1-1dy	2,666	0.087	1.00	1.00		29.20%	1.00
female_w	2-2dy	4,794	0.070	0.56	0.66	31.90%	29.20%	0.65
female_w	3-All	6,936	0.064	0.38	0.55	26.50%	29.20%	0.56
female_wb	1-1dy	2,666	0.168	1.00	1.00		56.19%	1.00
female_wb	2-2dy	4,794	0.148	0.56	0.78	55.50%	56.19%	0.78
female_wb	3-All	6,936	0.144	0.38	0.73	56.87%	56.19%	0.73
hwylsum_sh	1-1dy	2,666	0.136	1.00	1.00		77.06%	1.00
hwylsum_sh	2-2dy	4,794	0.119	0.56	0.77	54.12%	77.06%	0.89
hwylsum_sh	3-All	6,936	0.167	0.38	1.51	100.00%	77.06%	0.86
hwylsum_sr	1-1dy	2,666	0.175	1.00	1.00		16.39%	1.00
hwylsum_sr	2-2dy	4,794	0.140	0.56	0.64	27.52%	16.39%	0.58
hwylsum_sr	3-All	6,936	0.113	0.38	0.42	5.27%	16.39%	0.49
inc2_o	1-1dy	2,666	0.388	1.00	1.00		2.04%	1.00
inc2_o	2-2dy	4,794	0.273	0.56	0.50	-0.91%	2.04%	0.51
inc2_o	3-All	6,936	0.250	0.38	0.42	5.00%	2.04%	0.40
inc2_w	1-1dy	2,666	0.423	1.00	1.00		46.88%	1.00
inc2_w	2-2dy	4,794	0.361	0.56	0.73	45.48%	46.88%	0.73
inc2_w	3-All	6,936	0.350	0.38	0.68	48.27%	46.88%	0.67

Table G-3-6 (Part 4). Jackknife Standard Errors, *α* calculations, and design effects for Worker Tour Generation Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
inc3_o	1-1dy	2,666	0.360	1.00	1.00		9.32%	1.00
inc3_o	2-2dy	4,794	0.264	0.56	0.54	7.16%	9.32%	0.55
inc3_o	3-All	6,936	0.243	0.38	0.46	11.48%	9.32%	0.44
inc3_w	1-1dy	2,666	0.411	1.00	1.00		44.58%	1.00
inc3_w	2-2dy	4,794	0.349	0.56	0.72	44.12%	44.58%	0.72
inc3_w	3-All	6,936	0.334	0.38	0.66	45.04%	44.58%	0.66
inc4_o	1-1dy	2,666	0.361	1.00	1.00		9.00%	1.00
inc4_o	2-2dy	4,794	0.264	0.56	0.54	7.25%	9.00%	0.55
inc4_o	3-All	6,936	0.243	0.38	0.45	10.76%	9.00%	0.44
inc4_w	1-1dy	2,666	0.404	1.00	1.00		46.28%	1.00
inc4_w	2-2dy	4,794	0.343	0.56	0.72	44.72%	46.28%	0.73
inc4_w	3-All	6,936	0.333	0.38	0.68	47.84%	46.28%	0.67
inc5_o	1-1dy	2,666	0.366	1.00	1.00		9.20%	1.00
inc5_o	2-2dy	4,794	0.270	0.56	0.54	8.57%	9.20%	0.55
inc5_o	3-All	6,936	0.244	0.38	0.45	9.83%	9.20%	0.44
inc5_w	1-1dy	2,666	0.405	1.00	1.00		43.39%	1.00
inc5_w	2-2dy	4,794	0.341	0.56	0.71	42.00%	43.39%	0.72
inc5_w	3-All	6,936	0.329	0.38	0.66	44.77%	43.39%	0.65
kids_o	1-1dy	2,666	0.105	1.00	1.00		22.56%	1.00
kids_o	2-2dy	4,794	0.083	0.56	0.62	24.95%	22.56%	0.61
kids_o	3-All	6,936	0.075	0.38	0.51	20.18%	22.56%	0.52
kids_sh	1-1dy	2,666	0.135	1.00	1.00		46.11%	1.00
kids_sh	2-2dy	4,794	0.117	0.56	0.75	49.76%	46.11%	0.73
kids_sh	3-All	6,936	0.109	0.38	0.65	42.46%	46.11%	0.67
kids_sr	1-1dy	2,666	0.102	1.00	1.00		29.99%	1.00
kids_sr	2-2dy	4,794	0.084	0.56	0.68	35.50%	29.99%	0.65
kids_sr	3-All	6,936	0.074	0.38	0.54	24.49%	29.99%	0.57

Table G-3-6 (Part 5). Jackknife Standard Errors, *α* calculations, and design effects for Worker Tour Generation Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
nwkrs_sh	1-1dy	2,666	0.140	1.00	1.00		26.89%	1.00
nwkrs_sh	2-2dy	4,794	0.112	0.56	0.64	28.15%	26.89%	0.63
nwkrs_sh	3-All	6,936	0.103	0.38	0.54	25.63%	26.89%	0.55
nwkrs_sr	1-1dy	2,666	0.114	1.00	1.00		15.53%	1.00
nwkrs_sr	2-2dy	4,794	0.088	0.56	0.60	19.48%	15.53%	0.58
nwkrs_sr	3-All	6,936	0.077	0.38	0.46	11.59%	15.53%	0.48
othwkrs_sh	1-1dy	2,666	0.132	1.00	1.00		20.71%	1.00
othwkrs_sh	2-2dy	4,794	0.104	0.56	0.62	23.62%	20.71%	0.60
othwkrs_sh	3-All	6,936	0.093	0.38	0.49	17.80%	20.71%	0.51
othwkrs_sr	1-1dy	2,666	0.106	1.00	1.00		23.61%	1.00
othwkrs_sr	2-2dy	4,794	0.086	0.56	0.65	29.43%	23.61%	0.62
othwkrs_sr	3-All	6,936	0.075	0.38	0.49	17.80%	23.61%	0.53
partime_w	1-1dy	2,666	0.119	1.00	1.00		52.47%	1.00
partime_w	2-2dy	4,794	0.105	0.56	0.78	55.95%	52.47%	0.76
partime_w	3-All	6,936	0.099	0.38	0.69	48.98%	52.47%	0.71
partime_wb	1-1dy	2,666	0.426	1.00	1.00		19.50%	1.00
partime_wb	2-2dy	4,794	0.328	0.56	0.59	18.01%	19.50%	0.60
partime_wb	3-All	6,936	0.306	0.38	0.51	20.98%	19.50%	0.50
veh0_sh	1-1dy	2,666	0.517	1.00	1.00		-5.35%	1.00
veh0_sh	2-2dy	4,794	0.349	0.56	0.46	-8.69%	-5.35%	0.47
veh0_sh	3-All	6,936	0.315	0.38	0.37	-2.00%	-5.35%	0.35
veh0_sr	1-1dy	2,666	0.334	1.00	1.00		48.34%	1.00
veh0_sr	2-2dy	4,794	0.287	0.56	0.74	47.58%	48.34%	0.74
veh0_sr	3-All	6,936	0.277	0.38	0.69	49.11%	48.34%	0.68
veh0_w	1-1dy	2,666	0.350	1.00	1.00		31.66%	1.00
veh0_w	2-2dy	4,794	0.278	0.56	0.63	26.20%	31.66%	0.66
veh0_w	3-All	6,936	0.274	0.38	0.61	37.12%	31.66%	0.58
vehlwk_sh	1-1dy	2,666	0.210	1.00	1.00		13.90%	1.00
vehlwk_sh	2-2dy	4,794	0.161	0.56	0.59	18.25%	13.90%	0.57
vehlwk_sh	3-All	6,936	0.140	0.38	0.44	9.54%	13.90%	0.47
vehlwk_w	1-1dy	2,666	0.176	1.00	1.00		16.67%	1.00
vehlwk_w	2-2dy	4,794	0.136	0.56	0.60	19.73%	16.67%	0.58
vehlwk_w	3-All	6,936	0.120	0.38	0.47	13.61%	16.67%	0.49

Table G-3-7 presents the variance ratios and the estimated *α* factors for the comparison of the two-day file estimates and the full-file estimates to the one-day estimates, ordered by variance ratio. For the two-day file, the estimated *α*'s range from negative values to a high of 59%. The median value is 20%. There is a very wide range of *α*'s. For the full-file the estimated *α*'s range from 2% to 61%, with a median value of about 15%.

Table G-3-7. Variance ratios and α factors for comparison of the two-day file to the one-day files, comparison of the full file to the one-day file.

VR Parameter Estimates Two-day to One-day	Corresponding two-day to one-day α factor	VR Parameter Estimates Full file to One-day	Corresponding full-file to one-day α factor
0.457	-8.69%	0.372	-2.00%
0.495	-0.91%	0.415	5.00%
0.510	1.97%	0.417	5.27%
0.536	7.16%	0.418	5.44%
0.536	7.25%	0.421	5.89%
0.543	8.57%	0.424	6.47%
0.571	14.17%	0.429	7.21%
0.589	17.73%	0.434	7.97%
0.590	18.01%	0.441	9.23%
0.591	18.25%	0.443	9.54%
0.597	19.48%	0.445	9.83%
0.599	19.73%	0.449	10.49%
0.618	23.62%	0.451	10.76%
0.619	23.86%	0.455	11.48%
0.625	24.95%	0.456	11.59%
0.631	26.20%	0.468	13.61%
0.634	26.86%	0.472	14.17%
0.635	27.09%	0.494	17.80%
0.638	27.52%	0.494	17.80%
0.638	27.56%	0.509	20.18%
0.640	28.03%	0.514	20.98%
0.641	28.15%	0.535	24.49%
0.647	29.31%	0.542	25.63%
0.647	29.43%	0.542	25.64%
0.648	29.63%	0.548	26.50%
0.654	30.76%	0.559	28.28%
0.655	31.08%	0.561	28.62%
0.656	31.13%	0.580	31.73%
0.659	31.90%	0.601	35.12%
0.661	32.10%	0.613	37.12%
0.665	32.91%	0.646	42.46%
0.678	35.50%	0.660	44.77%
0.697	39.40%	0.662	45.04%
0.698	39.67%	0.665	45.51%
0.710	42.00%	0.666	45.75%
0.720	44.06%	0.679	47.84%
0.721	44.12%	0.682	48.27%
0.724	44.72%	0.686	48.98%
0.725	45.01%	0.687	49.11%
0.727	45.48%	0.735	56.87%
0.738	47.58%	0.838	73.61%
0.749	49.76%	0.849	75.51%
0.754	50.78%	0.860	77.24%
0.771	54.12%	0.861	77.33%
0.778	55.50%	0.952	92.16%
0.779	55.87%	1.494	100.00%
0.780	55.95%	1.515	100.00%
0.785	56.96%	1.520	100.00%
0.789	57.89%	1.521	100.00%
0.796	59.16%	1.540	100.00%
0.801	60.23%	1.564	100.00%
0.988	97.54%		

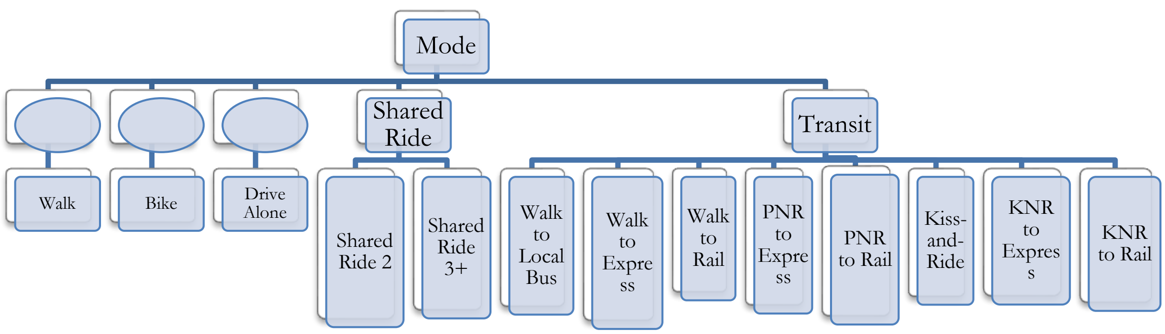
Table G-3-8 presents the eigenvalues from the variance ratio matrix $\widehat{\mathrm{vr}}\left(\widehat{\boldsymbol{\beta}}^{(2)}, \boldsymbol{V}_{\boldsymbol{\beta}}^{(\mathbf{1})}\right) = \left\{v_J\left(\widehat{\boldsymbol{\beta}}^{(1)}\right)\right\}^{-\mathbf{1}} v_J\left(\widehat{\boldsymbol{\beta}}^{(2)}\right)$ (see Appendix F). The corresponding matrix comparing the full file and the one-day file is not also presented, as there is a differing number of parameters (53 vs. 52). As with Table G-2-8, the geometric means of eigenvalues and univariate parameter ratios are very close for both files (0.661 for the univariate parameters and 0.669 for the eigenvalues). The eigenvalues have a larger range, but are centered around the same geometric mean.

Table G-3-8. Variance ratios and eigenvalues for comparison of the two-day file to the one-day files.

VR Parameter Estimates Two-day to One-day	Eigenvalues Two-day to One-day
0.457	0.294
0.495	0.300
0.510	0.329
0.536	0.343
0.536	0.351
0.543	0.375
0.571	0.388
0.589	0.407
0.590	0.429
0.591	0.448
0.597	0.465
0.599	0.480
0.618	0.482
0.619	0.502
0.625	0.519
0.631	0.534
0.634	0.540
0.635	0.565
0.638	0.573
0.638	0.597
0.640	0.601
0.641	0.618
0.647	0.634
0.647	0.643
0.648	0.648
0.654	0.664
0.655	0.672
0.656	0.684
0.659	0.710
0.661	0.734
0.665	0.777
0.678	0.786
0.697	0.801
0.698	0.829
0.710	0.853
0.720	0.868
0.721	0.879
0.724	0.891
0.725	0.916
0.727	0.943
0.738	0.949
0.749	0.993
0.754	1.002
0.771	1.038
0.778	1.073
0.779	1.083
0.780	1.108
0.785	1.171
0.789	1.242
0.796	1.272
0.801	1.333
0.988	1.374

Appendix G-4. Model Estimation Results for Mode Choice Models: Work Tours

The mode choice models predict the probability of a trip or tour using a specific travel mode. Appendix G-4 fits mode choices for work tours. The companion Appendix G-5 will fit social/recreation tours. The travel modes include: drive-alone, shared ride 2, shared ride 3+, walk, bike, walk to bus, walk to rail, drive to bus, and drive to rail. The mode choice model is a nested logit (NL) model to allow for greater competition among modes that share a common nest.



In the nested logit model, the probability of choosing an alternative *i* is the conditional probability of choosing *i* given that the nest $B(i)$ containing *i* is chosen, times the probability of choosing the nest $B(i)$.

$$\Pr(i) = \Pr\big(i \big| i \in B(i)\big) \cdot \Pr(i \in B(i))$$

The probability of choosing the alternative within the nest is:

$$\Pr\big(i \big| i \in B(i)\big) = \frac{\exp(U_i/\theta_{B(i)})}{\sum_{j \in B(i)} \exp(U_j/\theta_{B(i)})}$$

which is equivalent to the MNL probability with the addition of an estimated nesting parameter $\theta_{B(i)}$ bounded by 0 and 1. The probability of choosing the nest $B(i)$ is given by:

$$\Pr(i \in B(i)) = \frac{\exp(\theta_{B(i)}\Gamma_{B(i)})}{\sum_{m \in M} \exp(\theta_{B(m)}\Gamma_{B(m)})}$$

Where $\theta_{B(i)}$ is the same nesting parameter, *m* is a nest index from the set of all nests *M* at that level, and $\Gamma_{B(i)}$ is defined as:

$$\Gamma_{B(i)} = log\left(\sum_{j \in B(i)} \exp(U_j/\theta_{B(i)})\right)$$

- $\Gamma_{B(i)}$ is known as the logsum term, and represents the composite utility of all alternatives within the nest. As before, the utility can be expressed as $U_i = \beta X_i$ where β is the vector of estimated model coefficients and X_i is the vector of predictors, but the model also requires the estimation of θ.

As with the previous models, the mode choice models will be estimated using maximum likelihood estimation. The models will be estimated from each of the three files using weighted maximum likelihood estimation, and the jackknife weights will be utilized to generate jackknife estimates for each of the three files. The weights for the base estimation are all 1, with the weights varying for each jackknife replicate, but still 1 for most observations. A weighted MLE parameter vector estimate and jackknife variance matrices were generated for each file, and comparisons made of the variance matrices. Tables G-4-1, G-4-2, and G-4-3 provide the parameter estimates from the full data set, the two-day data set, and the one-day data set respectively.

The model parameters are described as follows:

- In-vehicle time is the total travel time spent in a car or transit vehicle. Walk mode and bike mode time is included with in-vehicle time.
- Out-of-vehicle time is the total time—walking, waiting, and drive access time—used in support of a transit trip, beyond what is in the transit vehicle.
- The cost coefficient is segmented by income. For the full sample estimation, the value of time is $2.66/hour for travelers in households earning $0-25,000, and $9.06/hour for travelers in households earning $25,000+. For comparison, the average hourly wage rate in the Cleveland region was $22.26/hour in May 2014 (Bureau of Labor Statistics 2015).
- There is a penalty applied if the path to a premium transit mode includes any in-vehicle time on a local bus.
- The mixed density measures make travelers less likely to drive or drive to transit.
- Travelers from larger households are more likely to carpool.
- Having a child in the household makes someone more likely to drive.
- There is a set of constants that is segmented by auto sufficiency (0 autos, 0<autos<workers, or autos>workers).
- There is an unsegmented set of constants that applies to specific transit modes.
- The nesting coefficient affects the relative cross-elasticities within the nests versus between nests.

Table G-4-1. Parameter Estimates for Work Tour Mode Choice Model: Full Data Set

Description	Applies to Alternatives	Label	Coeff	Jack-knife t-stat
In-Vehicle Time		ivt	-0.0328	-3.67
Out-of-Vehicle Time		ovt	-0.0196	-1.70
Cost, Income $0-$24,999		cost0_25	-0.0074	-2.31
Cost, Income $25,000+		cost25p	-0.0022	-2.12
Local bus used as access mode	Any express bus or rail	locpen	-0.4245	-1.59
Mixed density measure at home location	DA, SR2, SR3+, PNR, KNR	pmix_cardt	-0.1864	-1.20
Mixed density measure at work location	DA, SR2, SR3+	amix_car	-0.3648	-1.69
Household size	SR2	hhsize_s2	0.1682	1.40
Household size	SR3+	hhsize_s3	0.6698	2.18
Children in household	DA, SR2, SR3+	kids_car	0.5059	1.22
Constant: 0 auto or auto insufficient	Walk	autsuf1_wk	-0.3343	-0.58
Constant: auto sufficient	Walk	autsuf2_wk	-1.8682	-1.90
Constant: 0 auto or auto insufficient	Bike	autsuf1_bk	-3.5621	-2.26
Constant: auto sufficient	Bike	autsuf2_bk	-4.0136	-2.65
Constant: 0 auto or auto insufficient	SR2	autsuf1_s2	-2.5868	-2.68
Constant: auto sufficient	SR2	autsuf2_s2	-3.4970	-3.06
Constant: 0 auto or auto insufficient	SR3+	autsuf1_s3	-8.3483	-4.28
Constant: auto sufficient	SR3+	autsuf2_s3	-7.7559	-4.39
Constant: 0 auto	Any transit mode	autsuf0_t	2.4006	2.05
Constant: auto insufficient	Any transit mode	autsuf1_t	-1.8431	-2.09
Constant: auto sufficient	Any transit mode	autsuf2_t	-2.8577	-2.81
Constant: Park-and-Ride	Any PNR mode	asc_p	1.1450	1.76
Constant: Kiss-and-Ride	Any KNR mode	asc_k	-1.3928	-1.50
Constant: Express Bus	Any express bus mode	asc_e	-0.1884	-0.28
Constant: Rail	Any rail mode	asc_r	-2.2242	-3.68
Nesting Coefficient*	Transit, Shared Ride	theta	0.8666	3.07

Table G-4-2. Parameter Estimates for Work Tour Mode Choice Model: Two-Day Data Set

Description	**Applies to Alternatives**	**Label**	**Coeff**	**Jack-knife t-stat**
In-Vehicle Time		ivt	-0.0401	-3.29
Out-of-Vehicle Time		ovt	-0.0303	-1.83
Cost, Income $0-$24,999		cost0_25	-0.0087	-2.20
Cost, Income $25,000+		cost25p	-0.0029	-2.12
Local bus used as access mode	Any express bus or rail	locpen	-0.2004	-0.65
Mixed density measure at home location	DA, SR2, SR3+, PNR, KNR	pmix_cardt	-0.1670	-0.94
Mixed density measure at work location	DA, SR2, SR3+	amix_car	-0.4024	-1.46
Household size	SR2	hhsize_s2	0.1797	1.32
Household size	SR3+	hhsize_s3	0.7916	2.45
Children in household	DA, SR2, SR3+	kids_car	0.5532	1.07
Constant: 0 auto or auto insufficient	Walk	autsuf1_wk	-0.2388	-0.37
Constant: auto sufficient	Walk	autsuf2_wk	-1.9755	-1.61
Constant: 0 auto or auto insufficient	Bike	autsuf1_bk	-3.7007	-1.92
Constant: auto sufficient	Bike	autsuf2_bk	-4.5944	-2.26
Constant: 0 auto or auto insufficient	SR2	autsuf1_s2	-3.0933	-2.38
Constant: auto sufficient	SR2	autsuf2_s2	-4.1192	-2.58
Constant: 0 auto or auto insufficient	SR3+	autsuf1_s3	-9.8003	-0.66
Constant: auto sufficient	SR3+	autsuf2_s3	-8.8282	-4.01
Constant: 0 auto	Any transit mode	autsuf0_t	2.8608	1.72
Constant: auto insufficient	Any transit mode	autsuf1_t	-1.5058	-1.49
Constant: auto sufficient	Any transit mode	autsuf2_t	-2.5972	-2.13
Constant: Park-and-Ride	Any PNR mode	asc_p	0.8429	1.15
Constant: Kiss-and-Ride	Any KNR mode	asc_k	-1.9106	-1.57
Constnat: Express Bus	Any express bus mode	asc_e	-0.4197	-0.52
Constant: Rail	Any rail mode	asc_r	-2.1885	-3.40
Nesting Coefficient*	Transit, Shared Ride	theta	0.7375	2.64

Table G-4-3. Parameter Estimates for Work Tour Mode Choice Model: One-Day Data Set

Description	Applies to Alternatives	Label	Coeff	Jack-knife t-stat
In-Vehicle Time		ivt	-0.0385	-3.35
Out-of-Vehicle Time		ovt	-0.0219	-1.50
Cost, Income $0-$24,999		cost0_25	-0.0052	-1.53
Cost, Income $25,000+		cost25p	-0.0023	-1.84
Local bus used as access mode	Any express bus or rail	locpen	-0.5116	-1.37
Mixed density measure at home location	DA, SR2, SR3+, PNR, KNR	pmix_cardt	-0.0844	-0.54
Mixed density measure at work location	DA, SR2, SR3+	amix_car	-0.3990	-1.52
Household size	SR2	hhsize_s2	0.1463	1.17
Household size	SR3+	hhsize_s3	0.8309	2.77
Children in household	DA, SR2, SR3+	kids_car	0.7786	1.41
Constant: 0 auto or auto insufficient	Walk	autsuf1_wk	0.1245	0.22
Constant: auto sufficient	Walk	autsuf2_wk	-1.7609	-1.71
Constant: 0 auto or auto insufficient	Bike	autsuf1_bk	-2.7492	-1.74
Constant: auto sufficient	Bike	autsuf2_bk	-3.9646	-2.36
Constant: 0 auto or auto insufficient	SR2	autsuf1_s2	-2.9578	-2.33
Constant: auto sufficient	SR2	autsuf2_s2	-3.5774	-2.59
Constant: 0 auto or auto insufficient	SR3+	autsuf1_s3	-9.4101	-0.74
Constant: auto sufficient	SR3+	autsuf2_s3	-8.4351	-4.20
Constant: 0 auto	Any transit mode	autsuf0_t	2.4118	1.86
Constant: auto insufficient	Any transit mode	autsuf1_t	-1.6824	-1.69
Constant: auto sufficient	Any transit mode	autsuf2_t	-2.6774	-2.29
Constant: Park-and-Ride	Any PNR mode	asc_p	0.8077	0.82
Constant: Kiss-and-Ride	Any KNR mode	asc_k	-1.4310	-1.02
Constnat: Express Bus	Any express bus mode	asc_e	0.0396	0.04
Constant: Rail	Any rail mode	asc_r	-2.2584	-2.48
Nesting Coefficient*	Transit, Shared Ride	theta	0.8174	2.65

Table G-4-5 below present the parameter estimates as they are given in tables G-4-1, G-4-2, and G-4-3, but with the three day-file estimates together for each parameter. Also included are jackknife standard errors for the parameter estimates (Tables G-4-1 through G-4-3 present the t-statistics, which are the parameter estimates divided by the standard errors) Table G-4-5 presents degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). Note that in these tables we provided all estimates, even when the degrees of freedom were smaller than 30, to keep the parameter vector whole. But the variance calculations with low degrees of freedom should be treated with skepticism. Table G-4-4 presents a listing of short parameter names used in Tables G-4-5 and G-4-6 (to save space), linking back to the parameter descriptions in Tables G-4-1 through G-4-3.

Table G-4-4. Comparison of Full Parameter Names to Short Parameter Names for Work Tour Mode Choice Model Parameters

Full Parameter Name		**Short Parameter Name**
Description	**Applies to Alternatives**	**Label**
In-Vehicle Time		ivt
Out-of-Vehicle Time		ovt
Cost, Income $0-$24,999		cost0_25
Cost, Income $25,000+		cost25p
Local bus used as access mode	Any express bus or rail	locpen
Mixed density measure at home location	DA, SR2, SR3+, PNR, KNR	pmix_cardt
Mixed density measure at work location	DA, SR2, SR3+	amix_car
Household size	SR2	hhsize_s2
Household size	SR3+	hhsize_s3
Children in household	DA, SR2, SR3+	kids_car
Constant: 0 auto or auto insufficient	Walk	autsuf1_wk
Constant: auto sufficient	Walk	autsuf2_wk
Constant: 0 auto or auto insufficient	Bike	autsuf1_bk
Constant: auto sufficient	Bike	autsuf2_bk
Constant: 0 auto or auto insufficient	SR2	autsuf1_s2
Constant: auto sufficient	SR2	autsuf2_s2
Constant: 0 auto or auto insufficient	SR3+	autsuf1_s3
Constant: auto sufficient	SR3+	autsuf2_s3
Constant: 0 auto	Any transit mode	autsuf0_t
Constant: auto insufficient	Any transit mode	autsuf1_t
Constant: auto sufficient	Any transit mode	autsuf2_t
Constant: Park-and-Ride	Any PNR mode	asc_p
Constant: Kiss-and-Ride	Any KNR mode	asc_k
Constnat: Express Bus	Any express bus mode	asc_e
Constant: Rail	Any rail mode	asc_r
Nesting Coefficient*	Transit, Shared Ride	theta

Table G-4-5 (Part 1). Parameter Estimates for Work Tour Mode Choice Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for StdErr LB	CI for StdErr UB
amix_car	1-1dy	1,242	-0.399	0.262	55	0.221	0.322
amix_car	2-2dy	1,844	-0.402	0.276	57	0.233	0.338
amix_car	3-All	2,412	-0.365	0.216	58	0.183	0.264
asc_e	1-1dy	1,242	0.040	0.897	21	0.690	1.282
asc_e	2-2dy	1,844	-0.420	0.808	39	0.662	1.037
asc_e	3-All	2,412	-0.188	0.676	37	0.551	0.875
asc_k	1-1dy	1,242	-1.431	1.402	13	1.016	2.258
asc_k	2-2dy	1,844	-1.911	1.214	11	0.860	2.061
asc_k	3-All	2,412	-1.393	0.931	23	0.724	1.306
asc_p	1-1dy	1,242	0.808	0.982	21	0.755	1.403
asc_p	2-2dy	1,844	0.843	0.736	45	0.610	0.927
asc_p	3-All	2,412	1.145	0.652	43	0.539	0.826
asc_r	1-1dy	1,242	-2.258	0.910	13	0.660	1.466
asc_r	2-2dy	1,844	-2.188	0.643	35	0.522	0.839
asc_r	3-All	2,412	-2.224	0.605	35	0.491	0.789
autsuf0_t	1-1dy	1,242	2.412	1.297	56	1.095	1.591
autsuf0_t	2-2dy	1,844	2.861	1.667	18	1.260	2.465
autsuf0_t	3-All	2,412	2.401	1.170	29	0.932	1.573
autsuf1_bk	1-1dy	1,242	-2.749	1.577	8	1.065	3.021
autsuf1_bk	2-2dy	1,844	-3.701	1.925	10	1.345	3.379
autsuf1_bk	3-All	2,412	-3.562	1.574	11	1.115	2.673
autsuf1_s2	1-1dy	1,242	-2.958	1.268	62	1.079	1.538
autsuf1_s2	2-2dy	1,844	-3.093	1.300	58	1.101	1.589
autsuf1_s2	3-All	2,412	-2.587	0.967	49	0.808	1.205
autsuf1_s3	1-1dy	1,242	-9.410	12.743	3	7.219	47.514
autsuf1_s3	2-2dy	1,844	-9.800	14.782	3	8.374	55.117
autsuf1_s3	3-All	2,412	-8.348	1.952	26	1.537	2.675

Table G-4-5 (Part 2). Parameter Estimates for Work Tour Mode Choice Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for StdErr LB	CI for StdErr UB
autsuf1_t	1-1dy	1,242	-1.682	0.996	40	0.818	1.275
autsuf1_t	2-2dy	1,844	-1.506	1.011	52	0.849	1.251
autsuf1_t	3-All	2,412	-1.843	0.881	72	0.758	1.053
autsuf1_wk	1-1dy	1,242	0.124	0.561	40	0.461	0.718
autsuf1_wk	2-2dy	1,844	-0.239	0.646	41	0.532	0.824
autsuf1_wk	3-All	2,412	-0.334	0.580	33	0.468	0.763
autsuf2_bk	1-1dy	1,242	-3.965	1.680	28	1.333	2.272
autsuf2_bk	2-2dy	1,844	-4.594	2.030	40	1.666	2.597
autsuf2_bk	3-All	2,412	-4.014	1.517	35	1.231	1.979
autsuf2_s2	1-1dy	1,242	-3.577	1.382	55	1.165	1.699
autsuf2_s2	2-2dy	1,844	-4.119	1.598	46	1.328	2.008
autsuf2_s2	3-All	2,412	-3.497	1.141	47	0.950	1.429
autsuf2_s3	1-1dy	1,242	-8.435	2.009	25	1.575	2.773
autsuf2_s3	2-2dy	1,844	-8.828	2.200	37	1.794	2.846
autsuf2_s3	3-All	2,412	-7.756	1.767	23	1.374	2.479
autsuf2_t	1-1dy	1,242	-2.677	1.169	40	0.960	1.495
autsuf2_t	2-2dy	1,844	-2.597	1.221	51	1.023	1.514
autsuf2_t	3-All	2,412	-2.858	1.018	72	0.876	1.217
autsuf2_wk	1-1dy	1,242	-1.761	1.027	30	0.821	1.373
autsuf2_wk	2-2dy	1,844	-1.975	1.229	36	1.000	1.597
autsuf2_wk	3-All	2,412	-1.868	0.985	38	0.805	1.269
cost0_25	1-1dy	1,242	-0.005	0.003	22	0.003	0.005
cost0_25	2-2dy	1,844	-0.009	0.004	27	0.003	0.005
cost0_25	3-All	2,412	-0.007	0.003	19	0.002	0.005
cost25p	1-1dy	1,242	-0.002	0.001	33	0.001	0.002
cost25p	2-2dy	1,844	-0.003	0.001	45	0.001	0.002
cost25p	3-All	2,412	-0.002	0.001	46	0.001	0.001

Table G-4-5 (Part 3). Parameter Estimates for Work Tour Mode Choice Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for StdErr LB	CI for StdErr UB
hhsize_s2	1-1dy	1,242	0.146	0.125	67	0.107	0.150
hhsize_s2	2-2dy	1,844	0.180	0.136	67	0.117	0.164
hhsize_s2	3-All	2,412	0.168	0.120	66	0.102	0.145
hhsize_s3	1-1dy	1,242	0.831	0.300	14	0.220	0.473
hhsize_s3	2-2dy	1,844	0.792	0.323	20	0.247	0.467
hhsize_s3	3-All	2,412	0.670	0.307	14	0.225	0.484
ivt	1-1dy	1,242	-0.039	0.012	29	0.009	0.015
ivt	2-2dy	1,844	-0.040	0.012	38	0.010	0.016
ivt	3-All	2,412	-0.033	0.009	34	0.007	0.012
kids_car	1-1dy	1,242	0.779	0.552	36	0.449	0.717
kids_car	2-2dy	1,844	0.553	0.517	19	0.393	0.755
kids_car	3-All	2,412	0.506	0.415	32	0.334	0.549
locpen	1-1dy	1,242	-0.512	0.375	33	0.302	0.493
locpen	2-2dy	1,844	-0.200	0.309	47	0.258	0.388
locpen	3-All	2,412	-0.424	0.268	43	0.221	0.339
ovt	1-1dy	1,242	-0.022	0.015	35	0.012	0.019
ovt	2-2dy	1,844	-0.030	0.017	27	0.013	0.023
ovt	3-All	2,412	-0.020	0.012	24	0.009	0.016
pmix_cardt	1-1dy	1,242	-0.084	0.156	49	0.130	0.194
pmix_cardt	2-2dy	1,844	-0.167	0.178	36	0.145	0.231
pmix_cardt	3-All	2,412	-0.186	0.156	27	0.123	0.212
theta	1-1dy	1,242	0.817	0.308	40	0.253	0.394
theta	2-2dy	1,844	0.738	0.280	34	0.226	0.366
theta	3-All	2,412	0.867	0.282	29	0.224	0.379

Table G-4-6 below presents the variance ratios, *a* factors, and Pas design effects for the worker tour generation model, following the formulas as given preceding Table G-2-6. It should be noted that the degrees of freedom are not large for most of the parameter standard errors, so there is a lot of noise in these variance calculations that should lead to caution in interpreting the results.

Table G-4-6 (Part 1). Jackknife Standard Errors, *α* calculations, and design effects for Work Tour Mode Choice Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
amix_car	1-1dy	1,242	0.262	1.00	1.00		74.07%	1.00
amix_car	2-2dy	1,844	0.276	0.67	1.11	100.00%	74.07%	0.87
amix_car	3-All	2,412	0.216	0.51	0.68	48.13%	74.07%	0.84
asc_e	1-1dy	1,242	0.897	1.00	1.00		46.10%	1.00
asc_e	2-2dy	1,844	0.808	0.67	0.81	62.29%	46.10%	0.73
asc_e	3-All	2,412	0.676	0.51	0.57	29.91%	46.10%	0.67
asc_k	1-1dy	1,242	1.402	1.00	1.00		29.63%	1.00
asc_k	2-2dy	1,844	1.214	0.67	0.75	50.04%	29.63%	0.65
asc_k	3-All	2,412	0.931	0.51	0.44	9.22%	29.63%	0.57
asc_p	1-1dy	1,242	0.982	1.00	1.00		10.75%	1.00
asc_p	2-2dy	1,844	0.736	0.67	0.56	12.37%	10.75%	0.55
asc_p	3-All	2,412	0.652	0.51	0.44	9.12%	10.75%	0.45
asc_r	1-1dy	1,242	0.910	1.00	1.00		4.67%	1.00
asc_r	2-2dy	1,844	0.643	0.67	0.50	-0.06%	4.67%	0.52
asc_r	3-All	2,412	0.605	0.51	0.44	9.40%	4.67%	0.41
autsuf0_t	1-1dy	1,242	1.297	1.00	1.00		84.85%	1.00
autsuf0_t	2-2dy	1,844	1.667	0.67	1.65	100.00%	84.85%	0.92
autsuf0_t	3-All	2,412	1.170	0.51	0.81	69.69%	84.85%	0.91
autsuf1_bk	1-1dy	1,242	1.577	1.00	1.00		99.73%	1.00
autsuf1_bk	2-2dy	1,844	1.925	0.67	1.49	100.00%	99.73%	1.00
autsuf1_bk	3-All	2,412	1.574	0.51	1.00	99.45%	99.73%	1.00
autsuf1_s2	1-1dy	1,242	1.268	1.00	1.00		66.02%	1.00
autsuf1_s2	2-2dy	1,844	1.300	0.67	1.05	100.00%	66.02%	0.83
autsuf1_s2	3-All	2,412	0.967	0.51	0.58	32.03%	66.02%	0.79
autsuf1_s3	1-1dy	1,242	12.743	1.00	1.00		20.66%	1.00
autsuf1_s3	2-2dy	1,844	14.782	0.67	1.35	100.00%	20.66%	0.60
autsuf1_s3	3-All	2,412	1.952	0.51	0.02	-58.69%	20.66%	0.51

Table G-4-6 (Part 2). Jackknife Standard Errors, *α* calculations, and design effects for Work Tour Mode Choice Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
autsuf1_t	1-1dy	1,242	0.996	1.00	1.00		82.33%	1.00
autsuf1_t	2-2dy	1,844	1.011	0.67	1.03	100.00%	82.33%	0.91
autsuf1_t	3-All	2,412	0.881	0.51	0.78	64.66%	82.33%	0.89
autsuf1_wk	1-1dy	1,242	0.561	1.00	1.00		100.00%	1.00
autsuf1_wk	2-2dy	1,844	0.646	0.67	1.32	100.00%	100.00%	1.00
autsuf1_wk	3-All	2,412	0.580	0.51	1.07	100.00%	100.00%	1.00
autsuf2_bk	1-1dy	1,242	1.680	1.00	1.00		85.07%	1.00
autsuf2_bk	2-2dy	1,844	2.030	0.67	1.46	100.00%	85.07%	0.93
autsuf2_bk	3-All	2,412	1.517	0.51	0.82	70.13%	85.07%	0.91
autsuf2_s2	1-1dy	1,242	1.382	1.00	1.00		74.15%	1.00
autsuf2_s2	2-2dy	1,844	1.598	0.67	1.34	100.00%	74.15%	0.87
autsuf2_s2	3-All	2,412	1.141	0.51	0.68	48.29%	74.15%	0.84
autsuf2_s3	1-1dy	1,242	2.009	1.00	1.00		81.65%	1.00
autsuf2_s3	2-2dy	1,844	2.200	0.67	1.20	100.00%	81.65%	0.91
autsuf2_s3	3-All	2,412	1.767	0.51	0.77	63.30%	81.65%	0.89
autsuf2_t	1-1dy	1,242	1.169	1.00	1.00		80.44%	1.00
autsuf2_t	2-2dy	1,844	1.221	0.67	1.09	100.00%	80.44%	0.90
autsuf2_t	3-All	2,412	1.018	0.51	0.76	60.88%	80.44%	0.88
autsuf2_wk	1-1dy	1,242	1.027	1.00	1.00		93.41%	1.00
autsuf2_wk	2-2dy	1,844	1.229	0.67	1.43	100.00%	93.41%	0.97
autsuf2_wk	3-All	2,412	0.985	0.51	0.92	86.81%	93.41%	0.96
cost0_25	1-1dy	1,242	0.003	1.00	1.00		91.35%	1.00
cost0_25	2-2dy	1,844	0.004	0.67	1.36	100.00%	91.35%	0.96
cost0_25	3-All	2,412	0.003	0.51	0.89	82.70%	91.35%	0.95
cost25p	1-1dy	1,242	0.001	1.00	1.00		70.65%	1.00
cost25p	2-2dy	1,844	0.001	0.67	1.13	100.00%	70.65%	0.85
cost25p	3-All	2,412	0.001	0.51	0.64	41.30%	70.65%	0.82

Table G-4-6 (Part 3). Jackknife Standard Errors, *a* calculations, and design effects for Work Tour Mode Choice Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*a* Calcu-lation	Estimated Within-Person *a*	Pas Design Effect
hhsize_s2	1-1dy	1,242	0.125	1.00	1.00		93.54%	1.00
hhsize_s2	2-2dy	1,844	0.136	0.67	1.19	100.00%	93.54%	0.97
hhsize_s2	3-All	2,412	0.120	0.51	0.92	87.08%	93.54%	0.96
hhsize_s3	1-1dy	1,242	0.300	1.00	1.00		100.00%	1.00
hhsize_s3	2-2dy	1,844	0.323	0.67	1.16	100.00%	100.00%	1.00
hhsize_s3	3-All	2,412	0.307	0.51	1.04	100.00%	100.00%	1.00
ivt	1-1dy	1,242	0.012	1.00	1.00		67.75%	1.00
ivt	2-2dy	1,844	0.012	0.67	1.12	100.00%	67.75%	0.84
ivt	3-All	2,412	0.009	0.51	0.60	35.50%	67.75%	0.80
kids_car	1-1dy	1,242	0.552	1.00	1.00		52.34%	1.00
kids_car	2-2dy	1,844	0.517	0.67	0.88	75.33%	52.34%	0.76
kids_car	3-All	2,412	0.415	0.51	0.57	29.35%	52.34%	0.71
locpen	1-1dy	1,242	0.375	1.00	1.00		28.50%	1.00
locpen	2-2dy	1,844	0.309	0.67	0.68	36.59%	28.50%	0.64
locpen	3-All	2,412	0.268	0.51	0.51	20.41%	28.50%	0.56
ovt	1-1dy	1,242	0.015	1.00	1.00		69.58%	1.00
ovt	2-2dy	1,844	0.017	0.67	1.29	100.00%	69.58%	0.85
ovt	3-All	2,412	0.012	0.51	0.63	39.16%	69.58%	0.81
pmix_cardt	1-1dy	1,242	0.156	1.00	1.00		100.00%	1.00
pmix_cardt	2-2dy	1,844	0.178	0.67	1.31	100.00%	100.00%	1.00
pmix_cardt	3-All	2,412	0.156	0.51	1.00	100.00%	100.00%	1.00
theta	1-1dy	1,242	0.308	1.00	1.00		69.32%	1.00
theta	2-2dy	1,844	0.280	0.67	0.82	64.98%	69.32%	0.85
theta	3-All	2,412	0.282	0.51	0.84	73.66%	69.32%	0.81

Table G-4-7 presents the variance ratios and the estimated *a* factors for the comparison of the two-day file estimates and the full-file estimates to the one-day estimates, ordered by variance ratio. For the two-day file, the majority of the two-day variances actually exceed the one-day variance. This kind of inversion indicates a strong intra-person correlation: persons tend to have the same behavior for this particular model across the pair of days (in other words, mode choice for work tours tends not to change across a pair of adjacent days). We estimate the *a* value as 100% when the variance ratio exceeds 1 (assuming that the larger-than-1 value is a matter of noise in the variance estimates). The majority of the *a* values are therefore 100%. This can be interpreted as the true *a* values being large and close to 100%, without necessarily being 100% exactly.

For the full file, the variance ratios range from a low of 0.02 to a high of greater than 1 (four ratios being greater than 1), with a median value of 0.68. The corresponding *α* values range from a -59% (only one below 0) to a high value of 100% (four of these), with a median value of 48.3%. The *α* values range fairly well over the whole interval [0,1]. This differs from the two-day to one-day file case (previous paragraph). The *α* values indicate relatively sizeable within-person correlation of work tour mode choice across days, but not as extreme as the two-day to one-day case. The addition of a third day reduces this correlation. But all of these results should be taken with a grain of salt given the relatively small degrees of freedom for these jackknife variance estimates (see Table G-4-5). The eigenvalue analysis was not done for this model due to the unstable variance estimates.

Table G-4-7. Variance ratios and *α* factors for comparison of the two-day file to the one-day files, comparison of the full file to the one-day file (Work Tour Choice Model).

VR Parameter Estimates Two-day to One-day	Corresponding two-day to one-day *α* factor	VR Parameter Estimates Full file to One-day	Corresponding full-file to one-day *α* factor
0.50	-0.06%	0.02	-58.69%
0.56	12.37%	0.44	9.12%
0.68	36.59%	0.44	9.22%
0.75	50.04%	0.44	9.40%
0.81	62.29%	0.51	20.41%
0.82	64.98%	0.57	29.35%
0.88	75.33%	0.57	29.91%
1.03	100.00%	0.58	32.03%
1.05	100.00%	0.60	35.50%
1.09	100.00%	0.63	39.16%
1.11	100.00%	0.64	41.30%
1.12	100.00%	0.68	48.13%
1.13	100.00%	0.68	48.29%
1.16	100.00%	0.76	60.88%
1.19	100.00%	0.77	63.30%
1.20	100.00%	0.78	64.66%
1.29	100.00%	0.81	69.69%
1.31	100.00%	0.82	70.13%
1.32	100.00%	0.84	73.66%
1.34	100.00%	0.89	82.70%
1.35	100.00%	0.92	86.81%
1.36	100.00%	0.92	87.08%
1.43	100.00%	1.00	99.45%
1.46	100.00%	1.00	100.00%
1.49	100.00%	1.04	100.00%
1.65	100.00%	1.07	100.00%

Appendix G.5. Model Estimation Results for Mode Choice Models: Social/Recreation Tours

This Appendix G-5 discusses the models fit to social/recreation tours. The theory is very similar to that given in Appendix G-4 for work tours and will not be repeated here. Tables G-5-1, G-5-2, and G-5-3 provide the parameter estimates from the full data set, the two-day data set, and the one-day data set respectively.

Due to the low number of observations, drive to transit trips are not permitted for this purpose. The model parameters are described as follows:

- In-vehicle time is the total travel time spent in a car or transit vehicle. Walk mode and bike mode time is included with in-vehicle time.
- Out-of-vehicle time is the total time—walking, waiting, and drive access time—used in support of a transit trip, beyond what is in the transit vehicle.
- There is a single cost term. The value of time for the full sample model is $10.52. For comparison, the average hourly wage rate in the Cleveland region was $22.26/hour in May 2014 (Bureau of Labor Statistics 2015).
- Travelers from larger households are more likely to carpool.
- Having a child in the household makes someone more likely to drive.
- There is a set of constants that is segmented by auto sufficiency (0 autos, 0<autos<workers, or autos>workers).
- There is an unsegmented set of constants that applies to rail.

The nesting coefficients tested were not significant, so the model collapses to an MNL model.

Table G-5-1. Parameter Estimates for Social/Recreation Tour Mode Choice Model: Full Data Set

Description	Applies to Alternatives	Label	Coeff	Jack-knife t-stat
In-Vehicle Time		ivt	-0.0349	-8.59
Out-of-Vehicle Time		ovt	-0.0144	-1.10
Cost		cost	-0.0020	-2.26
Household size	SR2	hhsize_s2	0.1196	2.45
Household size	SR3+	hhsize_s3	0.8046	5.77
Children in household	DA, SR2, SR3+	kids_car	-0.3614	-2.01
Constant: 0 auto or auto insufficient	Walk	autsuf1_wk	-0.0647	-0.38
Constant: auto sufficient	Walk	autsuf2_wk	-1.0868	-5.34
Constant: 0 auto or auto insufficient	SR2	autsuf1_s2	-1.8926	-6.78
Constant: auto sufficient	SR2	autsuf2_s2	-2.2883	-14.66
Constant: 0 auto or auto insufficient	SR3+	autsuf1_s3	-7.0616	-9.10
Constant: auto sufficient	SR3+	autsuf2_s3	-7.5186	-10.79
Constant: 0 auto	Any transit mode	autsuf0_t	1.7218	1.98
Constant: auto insufficient	Any transit mode	autsuf1_t	-4.2588	-3.84
Constant: auto sufficient	Any transit mode	autsuf2_t	-5.6359	-4.27
Constant: Rail	Any rail mode	asc_r	-1.4801	-0.16

Table G-5-2. Parameter Estimates for Social/Recreation Tour Mode Choice Model: Two-day Data Set

Description	Applies to Alternatives	Label	Coeff	Jack-knife t-stat
In-Vehicle Time		ivt	-0.0315	-7.99
Out-of-Vehicle Time		ovt	-0.0133	-0.99
Cost		cost	-0.0017	-1.76
Household size	SR2	hhsize_s2	0.0493	0.88
Household size	SR3+	hhsize_s3	0.7364	4.96
Children in household	DA, SR2, SR3+	kids_car	-0.4220	-2.17
Constant: 0 auto or auto insufficient	Walk	autsuf1_wk	-0.0429	-0.24
Constant: auto sufficient	Walk	autsuf2_wk	-1.1334	-5.28
Constant: 0 auto or auto insufficient	SR2	autsuf1_s2	-1.5654	-4.92
Constant: auto sufficient	SR2	autsuf2_s2	-2.1403	-11.90
Constant: 0 auto or auto insufficient	SR3+	autsuf1_s3	-6.4777	-8.00
Constant: auto sufficient	SR3+	autsuf2_s3	-7.0989	-9.48
Constant: 0 auto	Any transit mode	autsuf0_t	1.6583	1.80
Constant: auto insufficient	Any transit mode	autsuf1_t	-4.0057	-3.62
Constant: auto sufficient	Any transit mode	autsuf2_t	-5.6644	-4.05
Constant: Rail	Any rail mode	asc_r	-1.3795	-0.14

Table G-5-3. Parameter Estimates for Social/Recreation Tour Mode Choice Model: One-day Data Set

Description	Applies to Alternatives	Label	Coeff	Jack-knife t-stat
In-Vehicle Time		ivt	-0.0276	-6.80
Out-of-Vehicle Time		ovt	-0.0075	-0.38
Cost		cost	-0.0017	-1.39
Household size	SR2	hhsize_s2	0.0139	0.23
Household size	SR3+	hhsize_s3	0.6979	4.97
Children in household	DA, SR2, SR3+	kids_car	-0.5909	-2.72
Constant: 0 auto or auto insufficient	Walk	autsuf1_wk	0.0648	0.33
Constant: auto sufficient	Walk	autsuf2_wk	-1.1093	-4.81
Constant: 0 auto or auto insufficient	SR2	autsuf1_s2	-1.4813	-4.16
Constant: auto sufficient	SR2	autsuf2_s2	-1.9777	-10.98
Constant: 0 auto or auto insufficient	SR3+	autsuf1_s3	-7.2676	-3.81
Constant: auto sufficient	SR3+	autsuf2_s3	-6.6638	-10.16
Constant: 0 auto	Any transit mode	autsuf0_t	1.8766	1.41
Constant: auto insufficient	Any transit mode	autsuf1_t	-4.8198	-3.17
Constant: auto sufficient	Any transit mode	autsuf2_t	-6.9331	-2.77
Constant: Rail	Any rail mode	asc_r	#N/A	#N/A

Table G-5-5 below present the parameter estimates as they are given in Tables G-5-1, G-5-2, and G-5-3, but with the three day-file estimates together for each parameter. Also included are jackknife standard errors for the parameter estimates (Tables G-5-1 through G-5-3 present the t-statistics, which are the parameter estimates divided by the standard errors) Table G-5-5 presents degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). Note that in these tables we provided all estimates, even when the degrees of freedom were smaller than 30, to keep the parameter vector whole. But the variance calculations with low degrees of freedom should be treated with skepticism. Table G-5-4 presents a listing of short parameter names used in Tables G-5-5 and G-5-6 (to save space), linking back to the parameter descriptions in Tables G-5-1 through G-5-3.

Table G-5-5. Comparison of Full Parameter Names to Short Parameter Names for Social/Recreational Tour Mode Choice Model Parameters

Full Parameter Name		**Short Parameter Name**
Description	**Applies to Alternatives**	**Label**
In-Vehicle Time		ivt
Out-of-Vehicle Time		ovt
Cost		cost
Household size	SR2	hhsize_s2
Household size	SR3+	hhsize_s3
Children in household	DA, SR2, SR3+	kids_car
Constant: 0 auto or auto insufficient	Walk	autsuf1_wk
Constant: auto sufficient	Walk	autsuf2_wk
Constant: 0 auto or auto insufficient	SR2	autsuf1_s2
Constant: auto sufficient	SR2	autsuf2_s2
Constant: 0 auto or auto insufficient	SR3+	autsuf1_s3
Constant: auto sufficient	SR3+	autsuf2_s3
Constant: 0 auto	Any transit mode	autsuf0_t
Constant: auto insufficient	Any transit mode	autsuf1_t
Constant: auto sufficient	Any transit mode	autsuf2_t
Constant: Rail	Any rail mode	asc_r

Table G-5-5 (Part 1). Parameter Estimates for Social/Recreational Tour Mode Choice Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	DF	CI for StdErr LB	CI for StdErr UB
asc_r	2-2dy	3,303	-1.380	9.752	3	5.524	36.360
asc_r	3-All	4,358	-1.480	9.484	3	5.372	35.360
autsuf0_t	1-1dy	2,185	1.878	1.330	8	0.898	2.548
autsuf0_t	2-2dy	3,303	1.658	0.923	12	0.662	1.523
autsuf0_t	3-All	4,358	1.722	0.872	14	0.638	1.375
autsuf1_s2	1-1dy	2,185	-1.485	0.356	75	0.307	0.424
autsuf1_s2	2-2dy	3,303	-1.565	0.318	80	0.276	0.376
autsuf1_s2	3-All	4,358	-1.893	0.279	72	0.240	0.334
autsuf1_s3	1-1dy	2,185	-7.333	1.907	4	1.142	5.479
autsuf1_s3	2-2dy	3,303	-6.478	0.810	30	0.647	1.083
autsuf1_s3	3-All	4,358	-7.062	0.776	34	0.628	1.017
autsuf1_t	1-1dy	2,185	-4.861	1.521	6	0.980	3.349
autsuf1_t	2-2dy	3,303	-4.006	1.106	15	0.817	1.712
autsuf1_t	3-All	4,358	-4.259	1.109	13	0.804	1.786
autsuf1_wk	1-1dy	2,185	0.073	0.195	25	0.153	0.270
autsuf1_wk	2-2dy	3,303	-0.043	0.179	24	0.140	0.249
autsuf1_wk	3-All	4,358	-0.065	0.169	27	0.133	0.229
autsuf2_s2	1-1dy	2,185	-1.973	0.180	54	0.152	0.222
autsuf2_s2	2-2dy	3,303	-2.140	0.180	59	0.152	0.219
autsuf2_s2	3-All	4,358	-2.288	0.156	85	0.136	0.184
autsuf2_s3	1-1dy	2,185	-6.607	0.656	11	0.464	1.113
autsuf2_s3	2-2dy	3,303	-7.099	0.749	12	0.537	1.236
autsuf2_s3	3-All	4,358	-7.519	0.697	13	0.505	1.123

Table G-5-5 (Part 2). Parameter Estimates for Social/Recreational Tour Mode Choice Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	DF	CI for StdErr LB	CI for StdErr UB
autsuf2_t	1-1dy	2,185	-6.979	2.504	6	1.613	5.513
autsuf2_t	2-2dy	3,303	-5.664	1.399	10	0.978	2.456
autsuf2_t	3-All	4,358	-5.636	1.318	10	0.921	2.314
autsuf2_wk	1-1dy	2,185	-1.099	0.231	41	0.190	0.294
autsuf2_wk	2-2dy	3,303	-1.133	0.215	26	0.169	0.294
autsuf2_wk	3-All	4,358	-1.087	0.203	28	0.161	0.275
cost	1-1dy	2,185	-0.001	0.001	9	0.001	0.002
cost	2-2dy	3,303	-0.002	0.001	12	0.001	0.002
cost	3-All	4,358	-0.002	0.001	23	0.001	0.001
hhsize_s2	1-1dy	2,185	0.017	0.060	68	0.051	0.072
hhsize_s2	2-2dy	3,303	0.049	0.056	51	0.047	0.070
hhsize_s2	3-All	4,358	0.120	0.049	79	0.042	0.058
hhsize_s3	1-1dy	2,185	0.713	0.141	27	0.111	0.191
hhsize_s3	2-2dy	3,303	0.736	0.148	11	0.105	0.252
hhsize_s3	3-All	4,358	0.805	0.139	13	0.101	0.225
ivt	1-1dy	2,185	-0.028	0.004			
ivt	2-2dy	3,303	-0.031	0.004	25	0.003	0.005
ivt	3-All	4,358	-0.035	0.004	24	0.003	0.006
kids_car	1-1dy	2,185	-0.600	0.217	75	0.187	0.258
kids_car	2-2dy	3,303	-0.422	0.194	29	0.155	0.261
kids_car	3-All	4,358	-0.361	0.180	36	0.147	0.234
ovt	1-1dy	2,185	-0.007	0.020	8	0.013	0.037
ovt	2-2dy	3,303	-0.013	0.013	9	0.009	0.024
ovt	3-All	4,358	-0.014	0.013	8	0.009	0.025

Table G-5-6 below presents the variance ratios, a factors, and Pas design effects for the worker tour generation model, following the formulas as given preceding Table G-2-6. It should be noted that the degrees of freedom are not large for most of the parameter standard errors (as for the work tour choice model as given in Appendix G-4), so there is a lot of noise in these variance calculations that should lead to caution in interpreting the results.

Table G-5-6 (Part 1). Jackknife Standard Errors, *α* calculations, and design effects for Social/Recreational Tour Mode Choice Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
autsuf0_t	1-1dy	2,185	1.330	1.00	1.00		1.79%	1.00
autsuf0_t	2-2dy	3,303	0.923	0.66	0.48	-3.71%	1.79%	0.51
autsuf0_t	3-All	4,358	0.872	0.50	0.43	7.29%	1.79%	0.40
autsuf1_s2	1-1dy	2,185	0.356	1.00	1.00		48.60%	1.00
autsuf1_s2	2-2dy	3,303	0.318	0.66	0.80	59.71%	48.60%	0.74
autsuf1_s2	3-All	4,358	0.279	0.50	0.62	37.49%	48.60%	0.68
autsuf1_s3	1-1dy	2,185	1.907	1.00	1.00		-49.73%	1.00
autsuf1_s3	2-2dy	3,303	0.810	0.66	0.18	-63.90%	-49.73%	0.25
autsuf1_s3	3-All	4,358	0.776	0.50	0.17	-35.56%	-49.73%	0.08
autsuf1_t	1-1dy	2,185	1.521	1.00	1.00		14.82%	1.00
autsuf1_t	2-2dy	3,303	1.106	0.66	0.53	5.77%	14.82%	0.57
autsuf1_t	3-All	4,358	1.109	0.50	0.53	23.86%	14.82%	0.48
autsuf1_wk	1-1dy	2,185	0.195	1.00	1.00		63.49%	1.00
autsuf1_wk	2-2dy	3,303	0.179	0.66	0.84	68.41%	63.49%	0.82
autsuf1_wk	3-All	4,358	0.169	0.50	0.75	58.58%	63.49%	0.78
autsuf2_s2	1-1dy	2,185	0.180	1.00	1.00		79.56%	1.00
autsuf2_s2	2-2dy	3,303	0.180	0.66	1.00	99.57%	79.56%	0.90
autsuf2_s2	3-All	4,358	0.156	0.50	0.75	59.54%	79.56%	0.87
autsuf2_s3	1-1dy	2,185	0.656	1.00	1.00		100.00%	1.00
autsuf2_s3	2-2dy	3,303	0.749	0.66	1.30	100.00%	100.00%	1.00
autsuf2_s3	3-All	4,358	0.697	0.50	1.13	100.00%	100.00%	1.00

Table G-5-6 (Part 2). Jackknife Standard Errors, *α* calculations, and design effects for Social/Recreational Tour Mode Choice Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
autsuf2_t	1-1dy	2,185	2.504	1.00	1.00		-27.48%	1.00
autsuf2_t	2-2dy	3,303	1.399	0.66	0.31	-37.51%	-27.48%	0.36
autsuf2_t	3-All	4,358	1.318	0.50	0.28	-17.44%	-27.48%	0.22
autsuf2_wk	1-1dy	2,185	0.231	1.00	1.00		68.52%	1.00
autsuf2_wk	2-2dy	3,303	0.215	0.66	0.87	73.21%	68.52%	0.84
autsuf2_wk	3-All	4,358	0.203	0.50	0.78	63.83%	68.52%	0.81
cost	1-1dy	2,185	0.001	1.00	1.00		30.08%	1.00
cost	2-2dy	3,303	0.001	0.66	0.67	34.55%	30.08%	0.65
cost	3-All	4,358	0.001	0.50	0.54	25.60%	30.08%	0.57
hhsize_s2	1-1dy	2,185	0.060	1.00	1.00		60.61%	1.00
hhsize_s2	2-2dy	3,303	0.056	0.66	0.88	76.43%	60.61%	0.80
hhsize_s2	3-All	4,358	0.049	0.50	0.66	44.78%	60.61%	0.76
hhsize_s3	1-1dy	2,185	0.141	1.00	1.00		98.79%	1.00
hhsize_s3	2-2dy	3,303	0.148	0.66	1.12	100.00%	98.79%	0.99
hhsize_s3	3-All	4,358	0.139	0.50	0.99	97.58%	98.79%	0.99
ivt	1-1dy	2,185	0.004	1.00	1.00		94.50%	1.00
ivt	2-2dy	3,303	0.004	0.66	0.95	89.00%	94.50%	0.97
ivt	3-All	4,358	0.004	0.50	1.00	100.00%	94.50%	0.97
kids_car	1-1dy	2,185	0.217	1.00	1.00		54.64%	1.00
kids_car	2-2dy	3,303	0.194	0.66	0.80	59.80%	54.64%	0.77
kids_car	3-All	4,358	0.180	0.50	0.69	49.48%	54.64%	0.72
ovt	1-1dy	2,185	0.020	1.00	1.00		2.00%	1.00
ovt	2-2dy	3,303	0.013	0.66	0.47	-6.59%	2.00%	0.51
ovt	3-All	4,358	0.013	0.50	0.45	10.58%	2.00%	0.40

Table G-5-7 presents the variance ratios and the estimated *α* factors for the comparison of the two-day file estimates and the full-file estimates to the one-day estimates, ordered by variance ratio. The variance ratios for the two-day file range from a low of 0.18 to a high of 1.31, with a median value of 0.80. The corresponding *α* values range from a -63.9% (four of these estimates are below 0) to a high value of 100% (two of these), with a median value of 59.8%. The *α* values range fairly well over a very wide interval. For social/recreational trips we might see a variety of mode choices (unlike work, where we would expect more consistency possibly across days), so 0% *α*'s, or even negative *α*'s is certainly possible. These results should be taken with a grain of salt given the relatively small degrees of freedom for these jackknife variance estimates (see Table G-5-5), and the wide range in the estimates probably indicate instability in the variance estimates. The eigenvalue analysis was not done for this model due to the unstable variance estimates.

For the full file, the variance ratios range from a low of 0.166 to a high of greater than 1 (two ratios being greater than 1), with a median value of 0.638. The corresponding *α* values range from a -35.6% (two below 0) to a high value of 100% (two of these), with a median value of 41.1%. Again there is a wide range in the *α* values. There is some evidence of smaller *α* values for the full file to one-day file comparison, as compared to the two-day to one-day file comparison, but the number of parameters are small, and the variance estimates unstable. The eigenvalue analysis was not done for this model due to the unstable variance estimates.

Table G-5-7. Variance ratios and *α* factors for comparison of the two-day file to the one-day files, comparison of the full file to the one-day file (Social/Recreational Tour Choice Model).

VR Parameter Estimates Two-day to One-day	Corresponding two-day to one-day *α* factor	VR Parameter Estimates Full file to One-day	Corresponding full-file to one-day *α* factor
0.180	-63.90%	0.166	-35.56%
0.312	-37.51%	0.277	-17.44%
0.467	-6.59%	0.429	7.29%
0.481	-3.71%	0.450	10.58%
0.529	5.77%	0.531	23.86%
0.673	34.55%	0.542	25.60%
0.799	59.71%	0.615	37.49%
0.799	59.80%	0.660	44.78%
0.842	68.41%	0.689	49.48%
0.866	73.21%	0.745	58.58%
0.882	76.43%	0.751	59.54%
0.945	89.00%	0.777	63.83%
0.998	99.57%	0.985	97.58%
1.117	100.00%	1.004	100.00%
1.305	100.00%	1.130	100.00%

Appendix G-6. Model Estimation Results for Destination Choice Models: Work Tours

The destination choice models predict the primary destination of tours. It is a multinomial logit model, with TAZs as alternatives. The utility of alternative *i*, takes the form: $U_i = \beta^1 X_i^1 + \ln(\beta^2 X_i^2)$. In this specification, β^1 is the standard vector of estimated model coefficients and X_i^1 is the standard vector of predictors. β^2 and X_i^2 are the estimated coefficients and predictors for the size term. The natural log transformations ensures that the probability of selecting an alternative changes linearly with the size term. Typically, the size is the employment by type in the TAZ, and doubling the employment will result in doubling the probability of selecting that TAZ, all else being equal.

Due to privacy restrictions of the employment data currently used in travel models in Ohio, the analysis will instead use employment data from the Longitudinal Employer-Household Dynamics (LEHD) program. LEHD employment data is published the US Census Bureau and provides estimates of employment by type at the geographic resolution of Census blocks. To maintain privacy, the data are made "fuzzy", but still provide a reasonable estimate of employment at the TAZ level.

This Appendix provides the destination choice model results for work tours.

As with the previous models, the destination models are estimated using maximum likelihood estimation. The models are estimated from each of the three files using weighted maximum likelihood estimation, and the jackknife weights will be utilized to generate jackknife estimates for each of the three files. A weighted MLE parameter vector estimate and jackknife variance matrices was generated for each file, and comparisons made of the variance matrices. Tables G-6-1, G-6-2, and G-6-3 provide the parameter estimates from the full data set, the two-day data set, and the one-day data set respectively.

The predictors included in the model are described as follows:

- The mode choice logsum is a generalized measure of impedance across all modes. The value lower than one implies a higher cross-elasticity across modes than destinations for work tours.
- Two additional terms are included, based on the log of distance. A linear distance term was tried, but the log value fit better. The distance term is segmented by income, with higher income travelers less sensitive to distance. This is probably because they tend to be more specialized, and thus have longer commutes.
- There is an additional log of distance term applied to part time workers. Part time workers are much more sensitive to distance than full time workers.
- The size term is based on total employment, with an additional factor applied if the person is a part time worker. Part time workers are more likely to be attracted to jobs in the retail or leisure industries. Leisure includes hotel and restaurant employment, as well as parks and recreation employment.

Table G-6-1. Parameter Estimates for Work Tour Destination Choice Model: Full Data Set

Description	Label	Coeff	Jack-knife
Mode choice logsum	lsum	0.7489	17.88
Log(distance), if income $0 to $75,000	logdst075	-0.6987	-3.44
Log(distance), if income $75,000+	logdst75p	-0.4973	-3.44
Log(distance), if part-time worker	logdist_pt	-1.1988	-3.62
Size term: total employment		1.0000	#N/A
Size term: retail and leisure employment, if part-time worker	rel_pt	1.1790	0.51

Table G-6-2. Parameter Estimates for Work Tour Destination Choice Model: Two-Day Data Set

Description	Label	Coeff	Jack-knife
Mode choice logsum	lsum	0.7490	18.45
Log(distance), if income $0 to $75,000	logdst075	-0.7436	-5.54
Log(distance), if income $75,000+	logdst75p	-0.5090	-3.57
Log(distance), if part-time worker	logdist_pt	-1.0670	-3.29
Size term: total employment		1.0000	#N/A
Size term: retail and leisure employment, if part-time worker	rel_pt	0.9443	1.36

Table G-6-3. Parameter Estimates for Work Tour Destination Choice Model: One-Day Data Set

Description	Label	Coeff	Jack-knife
Mode choice logsum	lsum	0.7034	16.76
Log(distance), if income $0 to $75,000	logdst075	-0.8334	-5.71
Log(distance), if income $75,000+	logdst75p	-0.5988	-3.77
Log(distance), if part-time worker	logdist_pt	-1.2207	-3.05
Size term: total employment		1.0000	#N/A
Size term: retail and leisure employment, if part-time worker	rel_pt	0.9505	1.31

Table G-6-5 below present the parameter estimates as they are given in tables G-6-1, G-6-2, and G-6-3, but with the three day-file estimates together for each parameter. Also included are jackknife standard errors for the parameter estimates (Tables G-6-1 through G-6-3 present the t-statistics, which are the parameter estimates divided by the standard errors) Table G-6-5 presents degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). Note that in these tables we provided all estimates, even when the degrees of freedom were smaller than 30, to keep the parameter vector whole. But the variance calculations with low degrees of freedom should be treated with skepticism. Table G-6-4 presents a listing of short parameter names used in Tables G-6-5 and G-6-6 (to save space), linking back to the parameter descriptions in Tables G-6-1 through G-6-3.

Table G-6-4. Comparison of Full Parameter Names to Short Parameter Names for Work Tour Destination Choice Model Parameters

Full Parameter Name	**Short Parameter Name**
Description	**Label**
Mode choice logsum	lsum
Log(distance), if income $0 to $75,000	logdst075
Log(distance), if income $75,000+	logdst75p
Log(distance), if part-time worker	logdist_pt
Size term*: total employment	L_S_M
Size term*: retail and leisure employment, if part-time worker	rel_pt

Table G-6-5. Parameter Estimates for Work Tour Destination Choice Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for StdErr LB	CI for StdErr UB
logdist_pt	1-1dy	1,240	-1.221	0.401	16	0.298	0.610
logdist_pt	2-2dy	1,841	-1.067	0.324	25	0.254	0.448
logdist_pt	3-All	2,408	-1.199	0.331	21	0.254	0.473
logdst075	1-1dy	1,240	-0.833	0.146	44	0.121	0.184
logdst075	2-2dy	1,841	-0.744	0.134	80	0.116	0.159
logdst075	3-All	2,408	-0.699	0.203	32	0.163	0.268
logdst75p	1-1dy	1,240	-0.599	0.159	18	0.120	0.235
logdst75p	2-2dy	1,841	-0.509	0.143	31	0.114	0.190
logdst75p	3-All	2,408	-0.497	0.145	62	0.123	0.175
lsum	1-1dy	1,240	0.703	0.042	80	0.036	0.050
lsum	2-2dy	1,841	0.749	0.041	78	0.035	0.048
lsum	3-All	2,408	0.749	0.042	92	0.037	0.049
rel_pt	1-1dy	1,240	0.951	0.727	31	0.583	0.967
rel_pt	2-2dy	1,841	0.944	0.692	55	0.583	0.851
rel_pt	3-All	2,408	1.179	2.323	18	1.755	3.435

Table G-6-6 below presents the variance ratios, *a* factors, and Pas design effects for the worker tour generation model, following the formulas as given preceding Table G-2-6. The estimates with low degrees of freedom (30 or below) should be treated with caution (though they are presented). The jackknife standard errors are generally (putting aside variability where the degrees of freedom are lower) similar for the 1-day, 2-day, and full files, leading to a estimates that are fairly high (40% to 100%). Work tour destination choice appears to be consistent across days in a way that reduces considerably the within-person correlations across days.

Table G-6-6. Jackknife Standard Errors, *a* calculations, and design effects for Work Tour Destination Choice Model (all files).

Parameter	Data File	Total Person -Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*a* Calcu-lation	Estimated Within-Person *a*	Pas Design Effect
logdist_pt	1-1dy	1,240	0.401	1.00	1.00		39.60%	1.00
logdist_pt	2-2dy	1,841	0.324	0.67	0.65	30.97%	39.60%	0.70
logdist_pt	3-All	2,408	0.331	0.51	0.68	48.23%	39.60%	0.63
logdst075	1-1dy	1,240	0.146	1.00	1.00		84.76%	1.00
logdst075	2-2dy	1,841	0.134	0.67	0.85	69.52%	84.76%	0.92
logdst075	3-All	2,408	0.203	0.51	1.93	100.00%	84.76%	0.91
logdst75p	1-1dy	1,240	0.159	1.00	1.00		66.28%	1.00
logdst75p	2-2dy	1,841	0.143	0.67	0.80	60.81%	66.28%	0.83
logdst75p	3-All	2,408	0.145	0.51	0.83	71.76%	66.28%	0.79
lsum	1-1dy	1,240	0.042	1.00	1.00		93.17%	1.00
lsum	2-2dy	1,841	0.041	0.67	0.93	86.98%	93.17%	0.97
lsum	3-All	2,408	0.042	0.51	1.00	99.36%	93.17%	0.96
rel_pt	1-1dy	1,240	0.727	1.00	1.00		90.53%	1.00
rel_pt	2-2dy	1,841	0.692	0.67	0.91	81.06%	90.53%	0.95
rel_pt	3-All	2,408	2.323	0.51	10.20	100.00%	90.53%	0.94

Appendix G-7. Model Estimation Results for Destination Choice Models: Social/Recreational Tours

This Appendix provides the destination choice model results for social/recreation tours. The theory behind these models is described in Appendix G-6 and will not be repeated here. Tables G-7-1, G-7-2, and G-7-3 provide the parameter estimates from the full data set, the two-day data set, and the one-day data set respectively.

The destination choice model for social/recreational tours is of a slightly different structure than for work tours. The initial trials could not estimate a model with an appropriate coefficient on the mode choice logsum term—it always estimated with a value greater than one. This violates the theory of choice models, and probably occurs because for non-work tours, travelers are more likely to trade-off destinations than modes. Therefore, the model for this purpose was specified to assume that the mode had already been chosen, and applied to only auto trips. This allowed mode specific level of service measures to be included in the model.

The predictors included in the model are described as follows:

- A negative and highly significant coefficient applies to travel time, as expected.
- The log of distance is included in this model as well, with a higher impedance for low income travelers.
- The size term considers two factors: households and leisure employment. Leisure employment includes hotels, restaurants, parks, recreation centers, and so forth, so is a key attractor of social and recreational tours. Travelers also visit friends and relatives within this purpose, so households is logical as well. The size term on leisure employment is segmented by income, with higher income travelers attracted at a higher rate to leisure employment, probably because they have more money to spend at such establishments.

Table G-7-1. Parameter Estimates for Social/Recreation Tour Destination Choice Model: Full Data Set

Description	Label	Coeff	Jack-knife
Highway time (min)	Time	-0.1270	-27.53
Log(distance), if income $0 to $25,000	logdst025	-0.9011	-4.61
Log(distance), if income $25,000+	logdst25p	-0.5443	-2.03
Size term*: households		1.0000	#N/A
Size term*: leisure employment, if income $0-$25,000	les025	1.2244	5.15
Size term*: leisure employment, if income $25,000+	les25p	1.7933	4.65

Table G-7-2. Parameter Estimates for Social/Recreation Tour Destination Choice Model: Two-Day Data Set

Description	Label	Coeff	Jack-knife
Highway time (min)	time	-0.1259	-23.01
Log(distance), if income $0 to $25,000	logdst025	-0.8420	-3.70
Log(distance), if income $25,000+	logdst25p	-0.6422	-2.00
Size term*: households		1.0000	#N/A
Size term*: leisure employment, if income $0-$25,000	les025	1.0787	3.74
Size term*: leisure employment, if income $25,000+	les25p	1.8741	4.35

Table G-7-3. Parameter Estimates for Social/Recreation Tour Destination Choice Model: One-Day Data Set

Description	Label	Coeff	Jack-knife
Highway time (min)	time	-0.1189	-18.69
Log(distance), if income $0 to $25,000	logdst025	-0.8019	-3.30
Log(distance), if income $25,000+	logdst25p	-0.8493	-1.80
Size term*: households		1.0000	#N/A
Size term*: leisure employment, if income $0-$25,000	les025	1.0135	2.89
Size term*: leisure employment, if income $25,000+	les25p	2.1893	3.73

Table G-7-5 below present the parameter estimates as they are given in Tables G-7-1, G-7-2, and G-7-3, but with the three day-file estimates together for each parameter. Also included are jackknife standard errors for the parameter estimates (Tables G-7-1 through G-7-3 present the t-statistics, which are the parameter estimates divided by the standard errors) Table G-7-5 presents degrees of freedom calculations for each jackknife standard error (see Section B-5 for formulas), and 95% confidence intervals for the standard errors (based on an assumed χ^2 distribution for variance estimates). Note that in these tables we provided all estimates, even when the degrees of freedom were smaller than 30, to keep the parameter vector whole. But the variance calculations with low degrees of freedom should be treated with skepticism. Table G-7-4 presents a listing of short parameter names used in Tables G-7-5 and G-7-6 (to save space), linking back to the parameter descriptions in Tables G-7-1 through G-7-3.

Table G-7-4. Comparison of Full Parameter Names to Short Parameter Names

Full Parameter Name	Short Parameter Name
Description	**Label**
Highway distance (mi)	time
Log(distance), if income $0 to $25,000	logdst025
Log(distance), if income $25,000+	logdst25p
Size term*: households	L_S_M
Size term*: leisure employment, if income $0-$25,000	les025
Size term*: leisure employment, if income $25,000+	les25p

Table G-7-5. Parameter Estimates for Social/Recreational Tour Destination Choice Model (all files), including Jackknife Standard Errors, Degrees of Freedom, Standard Error Confidence Intervals.

Parameter	Data File	Total Person-Days	Estimate	Jackknife Standard Error	Degrees of Freedom	CI for StdErr LB	CI for StdErr UB
les025	1-1dy	1,808	1.014	0.351	37	0.286	0.454
les025	2-2dy	2,862	1.079	0.288	61	0.245	0.350
les025	3-All	3,875	1.224	0.238	59	0.201	0.290
les25p	1-1dy	1,808	2.189	0.587	31	0.471	0.781
les25p	2-2dy	2,862	1.874	0.430	40	0.353	0.551
les25p	3-All	3,875	1.793	0.386	60	0.327	0.470
logdst025	1-1dy	1,808	-0.802	0.243	83	0.211	0.286
logdst025	2-2dy	2,862	-0.842	0.228	41	0.187	0.290
logdst025	3-All	3,875	-0.901	0.195	37	0.159	0.253
logdst25p	1-1dy	1,808	-0.849	0.472	19	0.359	0.690
logdst25p	2-2dy	2,862	-0.642	0.321	12	0.230	0.531
logdst25p	3-All	3,875	-0.544	0.268	15	0.198	0.415
time	1-1dy	1,808	-0.119	0.006	99	0.006	0.007
time	2-2dy	2,862	-0.126	0.005	34	0.004	0.007
time	3-All	3,875	-0.127	0.005	43	0.004	0.006

Table G-7-6 below presents the variance ratios, a factors, and Pas design effects for the worker tour generation model, following the formulas as given preceding Table G-2-6. The estimates with low degrees of freedom (30 or below) should be treated with caution (though they are presented). Unlike for the work tour destination choice model (see Table G-6-6), the a estimates are generally smaller. Social/recreation tour destination choice appears to be less consistent across days than work tours in a way that reduces considerably the within-person correlations across days.

Table G-7-6. Jackknife Standard Errors, *α* calculations, and design effects for Social/Recreational Tour Destination Choice Model (all files).

Parameter	Data File	Total Person-Days	Jackknife Standard Error	Inverse Person Ratio	Jackknife Variance Ratio	*α* Calcu-lation	Estimated Within-Person *α*	Pas Design Effect
les025	1-1dy	1,808	0.351	1.00	1.00		23.42%	1.00
les025	2-2dy	2,862	0.288	0.63	0.67	34.81%	23.42%	0.62
les025	3-All	3,875	0.238	0.47	0.46	12.02%	23.42%	0.53
les25p	1-1dy	1,808	0.587	1.00	1.00		7.52%	1.00
les25p	2-2dy	2,862	0.430	0.63	0.54	7.44%	7.52%	0.54
les25p	3-All	3,875	0.386	0.47	0.43	7.60%	7.52%	0.43
logdst025	1-1dy	1,808	0.243	1.00	1.00		58.95%	1.00
logdst025	2-2dy	2,862	0.228	0.63	0.88	75.42%	58.95%	0.79
logdst025	3-All	3,875	0.195	0.47	0.65	42.48%	58.95%	0.75
logdst25p	1-1dy	1,808	0.472	1.00	1.00		-8.73%	1.00
logdst25p	2-2dy	2,862	0.321	0.63	0.46	-7.29%	-8.73%	0.46
logdst25p	3-All	3,875	0.268	0.47	0.32	-10.17%	-8.73%	0.33
time	1-1dy	1,808	0.006	1.00	1.00		35.38%	1.00
time	2-2dy	2,862	0.005	0.63	0.74	47.86%	35.38%	0.68
time	3-All	3,875	0.005	0.47	0.53	22.90%	35.38%	0.60

References

Bureau of Labor Statistics (2015). Occupational Employment and Wages in Cleveland-Elyria-Mentor — May 2014, News Release 15-916-CHI, June 23, 2015, available at http://www.bls.gov/regions/midwest/news-release/occupationalemploymentandwages_cleveland.htm.

Appendix H. Cost-Benefit Analysis for Multi-Day Studies

The Pas (1986) paper (described in greater detail in Appendix A) develops an explicit cost model for comparing single-day and multiday studies. The cost of collecting T days of travel behavior from a single individual is assumed to be $C = p + qT$ where q is the cost of each collected day, and p is an 'overhead' cost for recruiting the individual. If N_M and N_S are the person-level sample sizes for the putative multiday and single-day studies, then the costs of these surveys using this simple cost model are

$$C_M = (p + qT)N_M, \quad C_S = (p + q)N_S \qquad \text{Eq(H} - 1)$$

Suppose C_S is the cost of a benchmark single stage study with sample size N_S that achieves set variance level V. Then $C_M = K_C C_S$ is the cost of a multiday study with T days that achieves the same variance level, with

$$K_C = \left(1 + \frac{q}{p}T\right)\left(1 + \frac{q}{p}\right)^{-1} \frac{1 + a(T - 1)}{T} \qquad \text{Eq(H} - 2)$$

Pas (1986) calls this a 'cost scale factor'. If K_C is greater than 1, then that means the single-day study that achieves the same precision is less expensive. If K_C is considerably smaller than 1, that means the multi-day study is less expensive. We can find the optimal T for given values of q, p, and a.

Table 6-1 presents four mean a factors as follows:

- Tabular a factors: weighted average of medians—25.5%
- Tabular a factors: weighted average of 75th percentiles—31.7%
- Model estimation a factors: weighted average of medians—51.9%
- Model estimation a factors: weighted average of 75th percentiles—63.9%.

This is a wide range of a factors. Tables H-1 through H-4 develop optimal designs then using a factors 25%, 37.5%, 50%, and 62.5% respectively. The first factor 25% is the most 'optimistic'; the last factor 62.5% the most 'pessimistic'.

The cost factor q/p will be determined by the particular study as well. We will use the range 0.05 through 0.20. A cost factor of 0.05 means that each extra day of travel collection costs $1/20^{th}$ the cost of recruiting the household and completing a baseline interview. A cost factor of 0.20 means that each extra day of travel collection costs $1/5^{th}$ the cost of recruiting the household and completing a baseline interview. The Appendix H tables below present the K_C factors for a range of T values for four different cost factors which may encompass future travel studies. The 'cost ratio' is $\left(1 + \frac{q}{p}T\right)\left(1 + \frac{q}{p}\right)^{-1}$ in Eq (H-2). The 'variance ratio' is $\frac{1+a(T-1)}{T}$. The K_C factor is the product of the cost ratio and the variance ratio, and the optimal value for T is the value that minimizes K_C. In some cases, two T values provide the same minimal K_C, and either (or both) are optimal.

Table H-1 (Part 1). K_C factors for an a value of 25% and a q/p value of 0.05, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.050	25.0%	1	1.000	1.000	100.00%
0.050	25.0%	3	1.095	0.500	54.76%
0.050	25.0%	5	1.190	0.400	47.62%
0.050	25.0%	6	1.238	0.375	46.43%
0.050	25.0%	7	1.286	0.357	45.92%
0.050	25.0%	**8**	1.333	0.344	**45.83%**
0.050	25.0%	9	1.381	0.333	46.03%
0.050	25.0%	10	1.429	0.325	46.43%

Table H-1 (Part 2). K_C factors for an a value of 25% and a q/p value of 0.075, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.075	25.0%	1	1.000	1.000	100.00%
0.075	25.0%	3	1.140	0.500	56.98%
0.075	25.0%	4	1.209	0.438	52.91%
0.075	25.0%	5	1.279	0.400	51.16%
0.075	25.0%	**6**	1.349	0.375	**50.58%**
0.075	25.0%	7	1.419	0.357	50.66%
0.075	25.0%	8	1.488	0.344	51.16%

Table H-1 (Part 3). K_C factors for an a value of 25% and a q/p value of 0.10, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.100	25.0%	1	1.000	1.000	100.00%
0.100	25.0%	3	1.182	0.500	59.09%
0.100	25.0%	4	1.273	0.438	55.68%
0.100	25.0%	**5**	1.364	0.400	**54.55%**
0.100	25.0%	**6**	1.455	0.375	**54.55%**
0.100	25.0%	7	1.545	0.357	55.19%
0.100	25.0%	8	1.636	0.344	56.25%

Table H-1 (Part 4). K_C factors for an a value of 25% and a q/p value of 0.15, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.150	25.0%	1	1.000	1.000	100.00%
0.150	25.0%	2	1.130	0.625	70.65%
0.150	25.0%	3	1.261	0.500	63.04%
0.150	25.0%	**4**	1.391	0.438	**60.87%**
0.150	25.0%	**5**	1.522	0.400	**60.87%**
0.150	25.0%	6	1.652	0.375	61.96%
0.150	25.0%	7	1.783	0.357	63.66%
0.150	25.0%	8	1.913	0.344	65.76%

Table H-1 (Part 5). K_C factors for an a value of 25% and a q/p value of 0.20, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.200	25.0%	1	1.000	1.000	100.00%
0.200	25.0%	2	1.167	0.625	72.92%
0.200	25.0%	3	1.333	0.500	66.67%
0.200	25.0%	**4**	1.500	0.438	**65.63%**
0.200	25.0%	5	1.667	0.400	66.67%
0.200	25.0%	6	1.833	0.375	68.75%

Table H-2 (Part 1). K_C factors for an a value of 37.5% and a q/p value of 0.05, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.050	37.5%	1	1.000	1.000	100.00%
0.050	37.5%	3	1.095	0.583	63.89%
0.050	37.5%	4	1.143	0.531	60.71%
0.050	37.5%	5	1.190	0.500	59.52%
0.050	37.5%	**6**	1.238	0.479	**59.33%**
0.050	37.5%	7	1.286	0.464	59.69%
0.050	37.5%	8	1.333	0.453	60.42%

Table H-2 (Part 2). K_C factors for an a value of 37.5% and a q/p value of 0.075, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.075	37.5%	1	1.000	1.000	100.00%
0.075	37.5%	3	1.140	0.583	66.47%
0.075	37.5%	4	1.209	0.531	64.24%
0.075	37.5%	**5**	1.279	0.500	**63.95%**
0.075	37.5%	6	1.349	0.479	64.63%
0.075	37.5%	7	1.419	0.464	65.86%

Table H-2 (Part 3). K_C factors for an a value of 37.5% and a q/p value of 0.10, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.100	37.5%	1	1.000	1.000	100.00%
0.100	37.5%	2	1.091	0.688	75.00%
0.100	37.5%	3	1.182	0.583	68.94%
0.100	37.5%	**4**	1.273	0.531	**67.61%**
0.100	37.5%	5	1.364	0.500	68.18%
0.100	37.5%	6	1.455	0.479	69.70%

Table H-2 (Part 4). K_C factors for an a value of 37.5% and a q/p value of 0.15, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.150	37.5%	1	1.000	1.000	100.00%
0.150	37.5%	2	1.130	0.688	77.72%
0.150	37.5%	**3**	1.261	0.583	**73.55%**
0.150	37.5%	4	1.391	0.531	73.91%
0.150	37.5%	5	1.522	0.500	76.09%

Table H-2 (Part 5). K_C factors for an a value of 37.5% and a q/p value of 0.20, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.200	37.5%	1	1.000	1.000	100.00%
0.200	37.5%	2	1.167	0.688	80.21%
0.200	37.5%	**3**	1.333	0.583	**77.78%**
0.200	37.5%	4	1.500	0.531	79.69%
0.200	37.5%	5	1.667	0.500	83.33%

Table H-3 (Part 1). K_C factors for an a value of 50% and a q/p value of 0.05, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.050	50.0%	1	1.000	1.000	100.00%
0.050	50.0%	2	1.048	0.750	78.57%
0.050	50.0%	3	1.095	0.667	73.02%
0.050	50.0%	**4**	1.143	0.625	**71.43%**
0.050	50.0%	**5**	1.190	0.600	**71.43%**
0.050	50.0%	6	1.238	0.583	72.22%
0.050	50.0%	8	1.333	0.563	75.00%

Table H-3 (Part 2). K_C factors for an a value of 50% and a q/p value of 0.075, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.075	50.0%	1	1.000	1.000	100.00%
0.075	50.0%	2	1.070	0.750	80.23%
0.075	50.0%	3	1.140	0.667	75.97%
0.075	50.0%	**4**	1.209	0.625	**75.58%**
0.075	50.0%	5	1.279	0.600	76.74%
0.075	50.0%	8	1.488	0.563	83.72%

Table H-3 (Part 3). K_C factors for an a value of 50% and a q/p value of 0.10, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.100	50.0%	1	1.000	1.000	100.00%
0.100	50.0%	2	1.091	0.750	81.82%
0.100	50.0%	**3**	1.182	0.667	**78.79%**
0.100	50.0%	4	1.273	0.625	79.55%
0.100	50.0%	5	1.364	0.600	81.82%

Table H-3 (Part 4). K_C factors for an a value of 50% and a q/p value of 0.15, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.150	50.0%	1	1.000	1.000	100.00%
0.150	50.0%	2	1.130	0.750	84.78%
0.150	50.0%	**3**	1.261	0.667	**84.06%**
0.150	50.0%	4	1.391	0.625	86.96%
0.150	50.0%	5	1.522	0.600	91.30%

Table H-3 (Part 5). K_C factors for an a value of 50% and a q/p value of 0.20, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.200	50.0%	1	1.000	1.000	100.00%
0.200	50.0%	**2**	1.167	0.750	**87.50%**
0.200	50.0%	3	1.333	0.667	88.89%
0.200	50.0%	4	1.500	0.625	93.75%

Table H-4 (Part 1). K_C factors for an a value of 62.5% and a q/p value of 0.05, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.050	62.5%	1	1.000	1.000	100.00%
0.050	62.5%	2	1.048	0.813	85.12%
0.050	62.5%	**3**	1.095	0.750	**82.14%**
0.050	62.5%	**4**	1.143	0.719	**82.14%**
0.050	62.5%	5	1.190	0.700	83.33%
0.050	62.5%	6	1.238	0.688	85.12%

Table H-4 (Part 2). K_C factors for an a value of 62.5% and a q/p value of 0.075, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.075	62.5%	1	1.000	1.000	100.00%
0.075	62.5%	2	1.070	0.813	86.92%
0.075	62.5%	**3**	1.140	0.750	**85.47%**
0.075	62.5%	4	1.209	0.719	86.92%
0.075	62.5%	5	1.279	0.700	89.53%

Table H-4 (Part 3). K_C factors for an a value of 62.5% and a q/p value of 0.10, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.100	62.5%	1	1.000	1.000	100.00%
0.100	62.5%	**2**	1.091	0.813	**88.64%**
0.100	62.5%	**3**	1.182	0.750	**88.64%**
0.100	62.5%	4	1.273	0.719	91.48%
0.100	62.5%	5	1.364	0.700	95.45%

Table H-4 (Part 4). K_C factors for an a value of 62.5% and a q/p value of 0.15, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.150	62.5%	1	1.000	1.000	100.00%
0.150	62.5%	**2**	1.130	0.813	**91.85%**
0.150	62.5%	3	1.261	0.750	94.57%
0.150	62.5%	4	1.391	0.719	100.00%

Table H-4 (Part 5). K_C factors for an a value of 62.5% and a q/p value of 0.20, with optimal T values in bold face.

q/p	Corre-lation a	Number of days	Cost ratio	Design effect	Kc factor
0.200	62.5%	1	1.000	1.000	100.00%
0.200	62.5%	**2**	1.167	0.813	**94.79%**
0.200	62.5%	3	1.333	0.750	100.00%
0.200	62.5%	4	1.500	0.719	107.81%

References

Pas, E. I. (1986). Multiday samples, parameter estimation precision, and data collection costs for least squares regression trip-generation models. *Environment and Planning A, 18*, 73-87.

Appendix I. Survey Data Processing

This appendix describes the processing to the 2012 Northeast Ohio Regional Travel Study, such that it could be used for estimating models of the key travel choices of interest. In all cases, the data include both the GPS-with-recall and GPS-only portions of the sample. The imputed mode, purpose, parking, fare and companion information are included in all cases where it is available, but the records are identified as imputed.

I-1. Converting Unlinked Trips to Linked Trips.

This section describes the processing necessary to the TRIP data table, which includes one record for each trip made.

I-1-1. Linking Trips

The starting data uses an "unlinked trip" format, where there is a new trip record each time a person changes mode or vehicles. This includes cases where that change is to walk or drive to a bus, or transfer between transit vehicles. Those transfer points are not of interest—what we care about is the initial starting point and final destination of the trip. To accommodate this, a second table was created that contains "linked trips", in which those transfer points are linked out. Thus, if a walking trip to a bus stop is observed immediately before a bus trip from that bus stop, those records would be merged into a single linked trip record from the origin of the first to the destination of the second.

Trips were linked in the following circumstances:

1. If the place type is a transit stop (PTYPE==5).
2. If the trip purpose is change travel mode/transfer (TPURP==4).
3. If two transit trips are immediately adjacent, with less than 20 minutes at that stop.
4. If walk, bike or auto trip is immediately before a transit trip, with less than 20 minutes at that stop.
5. If a walk, bike or auto trip is immediately after a transit trip, with less than 5 minutes at that stop.
6. If a walk trip is immediately before or after an auto trip, with less than 5 minutes at that stop.

In all cases, trips are only linked if they are made by the same traveler. More than two trips can be linked, and in many cases a single linked transit trip might consist of 3 or more unlinked trips. All modes used on unlinked trips are accumulated to the linked trip level, such that trip modes can be calculated.

I-1-2. Trip Modes

The survey modes are coded in a consistent manner with those used in the travel model. This is done at a linked trip level, based on the equivalencies shown in Table I-1-1.

Table I-1-1. Mode Equivalency

Model Mode	Survey Mode (MODE)	Access/Egress Modes (MODE)	Number of People on Trip (TOTTR)
DA – Drive Alone	5=Auto driver (car or small truck) 6=Auto passenger (car or small truck)	None 1=Walk	1
SR2 – Shared Ride 2	5=Auto driver (car or small truck) 6=Auto passenger (car or small truck) 7=Carpool/vanpool/other group ride	None 1=Walk	2
SR3 – Shared Ride 3+	5=Auto driver (car or small truck) 6=Auto passenger (car or small truck) 7=Carpool/vanpool/other group ride	None 1=Walk	>=3
Walk	1=Walk	None	Any
Bike	2=Bike	None	Any
WLKLOC – Walk to Local Bus (and Trolley and BRT)	9=Local bus (regular, standard, city) 13=Shuttle bus (public or employer-provided) 19=Airtrain or airport bus	None 1=Walk 2=Bike	Any
WLKEXP – Walk to Express Bus	10=Express bus (suburban, commuter, inter-city)	None 1=Walk 2=Bike	Any
WLKRAL – Walk to Rail	15=Subway (gcrta, etc)	None 1=Walk 2=Bike	Any
PNRLOC – Park-and-Ride to Local Bus (and Bus Rapid Transit)	9=Local bus (regular, standard, city) 13=Shuttle bus (public or employer-provided) 19=Airtrain or airport bus	5=Auto driver (car or small truck)	Any
PNREXP – Park-and-Ride to Express Bus	10=Express bus (suburban, commuter, inter-city)	5=Auto driver (car or small truck)	Any
PNRRAL – Park-and-Ride to Rail	15=Subway (gcrta, etc)	5=Auto driver (car or small truck)	Any
KNRLOC – Park-and-Ride to Local Bus (and Bus Rapid Transit)	9=Local bus (regular, standard, city) 13=Shuttle bus (public or employer-provided) 19=Airtrain or airport bus	6=Auto passenger (car or small truck) 7=Carpool/vanpool/other group ride	Any
KNREXP – Park-and-Ride to Express Bus	10=Express bus (suburban, commuter, inter-city)	6=Auto passenger (car or small truck) 7=Carpool/vanpool/other group ride	Any
KNRRAL – Park-and-Ride to Rail	15=Subway (gcrta, etc)	6=Auto passenger (car or small truck) 7=Carpool/vanpool/other group ride	Any

OTHER	3=Wheelchair/mobility scooter 4=Skates/skateboard/kick-scooter/Segway 8=Motorcycle/moped/motorized scooter 11=School bus 12=Charter bus (including employer-provided/ other contracted) 14=Paratransit service (access-a-ride, dial-a-ride, etc.) 16=Taxi 17=For-hire van/jitney/gypsy cab 18=Black car service/limo 97=Other (specify) 98=I don't know 99=Refused Any other modes	Any other combinations not counted above.	Any

In the table above, the survey mode is the main mode of the linked trip, and the access/egress modes are any modes used in support of that trip. On a transit trip, for example, it is allowed to walk or drive as a support mode to access the transit trip which is the main mode. Similarly, it is allowed to use a local bus to access an express bus or rail. The primacy of the mode is defined in a hierarchical fashion, with the modes in this list ordered from most to least important:

- 15=Subway (GCRTA, etc)
- 10=Express bus (suburban, commuter, inter-city)
- 9=Local bus (regular, standard, city)
- 13=Shuttle bus (public or employer-provided)
- 19=Airtrain or airport bus
- 7=Carpool/vanpool/other group ride
- 6=Auto passenger (car or small truck)
- 5=Auto driver (car or small truck)
- 2=Bike
- 1=Walk

The OTHER modes are left of this list, and are expected to be low-share options.

the NOACA model also includes alternatives for trolley and Bus Rapid Transit (BRT). For the purpose of this analysis, those modes are merged with local bus because 1) they are not readily identifiable from the mode information directly, and 2) it is expected that there will be a limited number of transit observations in the survey, making it difficult to support the more detailed break-out.

I-1-3. Productions and Attractions

The starting trip file was currently coded in origin destination (OD) format. While maintaining the origins and destinations of the trips, production attraction (PA) coding is added. The rules for coding are:

- If the place type of either end of the trip is home (PTYPE==1), then that home end become the production location, and the other end becomes the attraction location.
- If neither end of the trip is home, then there is no change.

Any trip-end attributes available in OD format are also coded in PA format. Specifically, this includes the location, place and purpose information. Any trips where the production end is at home are considered home-based trips, and any trips where the production end is not at home will be considered non-home-based.

I-1-4. Trip Purposes

Trip purposes are defined for trips coded in PA format. Table I-1-2 shows a summary of the trip purposes. The logic is such that any trip with the production end at home is a home-based trip, and subject to further disaggregation. Any trips with neither end at home are grouped into a non-home-based purpose. Non-home based trips with either end at work are considered NHBW and non-home based trips with neither end at work are classified as NHBO.

The trip purposes are listed in descending order of importance. If a trip has already been coded with a purpose higher in the list, it is not re-coded with a subsequent purpose. The person attributes represent a required filter. Only people who are employed can make HBW trips, and only people who are students at the appropriate level can make HBSC or HBU trips. For the first three purposes, the attraction place and the attraction trip purpose represent an OR condition, such that if someone goes to their work location, or says that they are working, they are considered to have made a HBW trip (assuming the other end is at home and the person is employed). For HBSH and HBSR trips, the attraction place type can be anything (that has not already been classified into another purpose), but the attraction trip purpose must be one of those listed.

Table I-1-2. Trip Purpose Equivalency

Model Trip Purpose	Production Place Type (PTYPE)	Person Attributes	Attraction Place Type (PTYPE)	Attraction Trip Purpose (TPURP)
HBW - Home-Based Work	1=HOME	Employed (EMPLY=1) OR Works (WORKS=1)	2=WORK 4=SECOND WORK	9=Work/doing my job
HBSC - Home-Based School	1=HOME	Student Status (STUDE=1 or STUDE=2) AND: Type of School (SCHOL>=1 and SCHOL<=4)	3=SCHOOL 6=DAYCARE	12=Attending class/studying
HBU - Home-Based University	1=HOME	Student Status (STUDE=1 or STUDE=2) AND: Type of School (SCHOL>=5 and SCHOL<=8)	3=SCHOOL	12=Attending class/studying
HBSH – Home-Based Shop	1=HOME	Any	Any	2=Shopping (on-line, catalog, or by phone) 16=Grocery/food shopping 17=Other routine shopping (clothing, convenience store, household maintenance) 18=Shopping for major purchases or specialty items (appliances, electronics, new vehicle, major hh repairs, etc)
HBSR – Home-Based Social/Recreational	1=HOME	Any	Any	21=Eat meal out at restaurant/diner 24=Outdoor recreation (jogging, biking, walking) 25=Indoor recreation (yoga, gym, etc) 26=Entertainment (movies, spectator sports, etc) 27=Social/visit friends/relatives
HBO – Home-Based Other	1=HOME	Any	Any other type	Any other purpose
NHBW – Non-Home-Based Work	Any Non-Home Location, with Either End at Work	Employed (EMPLY=1) OR	Any Non-Home Location, with Either End at Work (PTYPE=2	Any

	(PTYPE=2 or PTYPE=4)	Works (WORKS=1)	or PTYPE=4)	
NHBO – Non-Home-Based Other	Any Non-Home Location	Any	Any	Any

I-1-5. Time-of-Day

Each trip is assigned a time-of-day (TOD) based on the departure time from its origin. The TODs are:

- AM – 6:00-8:59 am
- MD – 9:00 am to 3:59 pm
- PM – 4:00-6:59 pm
- NT – 7:00 pm to 5:59 am

The AM and PM trips are further grouped into peak (PK) trips, and the MD and NT trips should be further grouped into off-peak (OP) trips.

I-1-6. TAZs

Coordinates at trip ends are replaced with TAZ IDs from 1597 TAZ system. This applies to both OD and PA, such that the file has PTAZ, ATAZ, OTAZ and DTAZ. The same is applied to the home, work and school locations in the person file.

Trip ends outside the model area are identified as external. If a trip has one end outside the model area it is assigned a flag of IX, and if it has both ends outside the model area it is assigned a flag of XX.

If the trip both starts and ends at home, it is identified as a LOOP.

I-1-7. Incomplete and Flagged Trips

Trips with missing information are to be excluded from certain analyses. Therefore, trips with the following conditions are flagged:

- Either end is external to the model area,
- Either end is otherwise not geocoded,
- The trip is a loop, or
- The mode is other.

I-2. Converting Linked Trips to Tours.

In addition to trip information, several analyses based on data compiled at a tour level. A tour is a chain of trips that starts and ends at home (i.e. a "round trip"). A work-based subtour is a chain of trips that starts and ends at work, without any stops at home. A typical work-based subtour is going out to lunch. Complete tours have at least two trips, but can have more trips as well. This data describes how the linked trip data are coded into tours.

I-2-1. Coding Tours

The tour coding starts from the linked trip file, sorted by person, by day, and by time.

The first step is to assign a main tour ID to each record in the linked trip file. For each person and each day, the main tour ID starts at 1. The main tour ID is then incremented each time the person departs from home (Origin PTYPE=1). Then, a second pass is made to identify any work-based subtours. Subtours are identified by tracking the last departure from work. If there is a return to work before a home location is encountered (i.e. it is on the same tour), then those trips between the departure from work and the return to work are flagged with a subtour ID. Finally, a combined tour ID is coded as a concatenation of the main tour ID and the subtour ID. For example, the trips on a main tour might have an ID of "1", those on the subtour might have an ID of "1-1", and those on a second main tour might have an ID of "2".

Any tours that do not start and end at home (other than subtours), are flagged as incomplete. Most often, these occur at the beginning or end of the travel day.

I-2-2. Tour Purposes

Tour purposes are coded as a function of the stop purposes included on the tour, based on the rules shown in Table I-2-1. The tour purposes are defined using the same rules as the trip purposes, with the exception that the place type and purpose can be for any stop encountered on the tour, and is not limited to the attraction location of an individual trip. This definition means that it is possible to have a work tour without any HBW trips, if there are intermediate stops both to and from work. The tour purposes are coded in a hierarchical fashion, with the most important purposes listed first in the table. Once a tour is classified as Work, it is not re-classified as any lower-importance purpose.

Table I-2-1. Tour Purpose Equivalency

Tour Purpose	Tour Type	Person Attributes	Place Type (PTYPE)	Purpose (TPURP)
W - Work	Home-Based	Employed (EMPLY=1) OR: Works (WORKS=1) AND: Age>=16	2=WORK 4=SECOND WORK	9=Work/doing my job
S - School	Home-Based	Student Status (STUDE=1 or STUDE=2) AND: Type of School (SCHOL>=1 and SCHOL<=4)	3=SCHOOL 6=DAYCARE	12=Attending class/studying
U - University	Home-Based	Student Status (STUDE=1 or STUDE=2) AND: Type of School (SCHOL>=5 and SCHOL<=8)	3=SCHOOL	12=Attending class/studying
SH - Shop	Home-Based	Any	Any	2=Shopping (on-line, catalog, or by phone) 16=Grocery/food shopping 17=Other routine shopping (clothing, convenience store, household maintenance) 18=Shopping for major purchases or specialty items (appliances, electronics, new vehicle, major hh repairs, etc)
SR - Social/Recreational	Home-Based	Any	Any	21=Eat meal out at restaurant/diner 24=Outdoor recreation (jogging, biking, walking) 25=Indoor recreation (yoga, gym, etc) 26=Entertainment (movies, spectator sports, etc) 27=Social/visit friends/relatives
O - Other	Home-Based	Any	Any other type	Any other purpose
SUB - Work-Based Subtour	Work-Based Subtour	Any	Any	Any

I-2-3. Tour-Level Aggregations

After the tour IDs are coded on the linked trip table, a separate tour table is created with one record for each tour ID. These records include a summation of the total trips on the tour, as well as the number of trips and stops of each purpose on the tour.

I-2.4 Incomplete and Flagged Tours

Tours with incomplete information are excluded from certain analyses. Therefore, tours with the following attributes should be clearly flagged:

- The tour does not both start and end at home (or at work for subtours),
- Any stop on the trip is outside the model area, and
- Any stop cannot be geocoded.

I-3. Person-Level Aggregations

The tour generation models operate at the person level. To accommodate this, a person type is defined, and tours are aggregated to the person level as described in this section.

I-3-1. Person Type

Table I-3-1 specifies the rules for defining person types. These are defined to be consistent with the person lifecycle categories found in Table 12-1 of the Survey Final Technical Compendium.

Table I-3-1. Person Type Equivalency

Person Type	**Employed (EMPLY)**	**Employment Status (PRIMA)**	**Number of Hrs Worked Per Week at Primary Job (HRS1)**	**Student Status (STUDE)**	**Age (AGE)**
1 - Full-Time Worker	1=Yes	Any	>=30	3=No 8=Don't know 9=Refused	>=18
2 - Part-Time Worker	1=Yes	Any	<30	3=No 8=Don't know 9=Refused	>=18
3 - University Student	Any	Any	Any	1=Yes, full-time 2=Yes, part-time	>=18

4 - Non-Worker	2= No
	8=Don't know
	9=Refused
		NOT:
		1=Retired
			Any	3=No
				8=Don't know
				9=Refused
					if PRIMA=1:
					>=18
					else:
					>=65

5 - Retiree	2= No 8=Don't know 9=Refused	1=Retired	Any	3=No 8=Don't know 9=Refused	>=18
6 - Driving-Age Child	Any	Any	Any	Any	>=16 AND <=17
7 - Pre-Driving-Age Child	Any	Any	Any	Any	>=6 AND <=15
8 - Preschool Child	Any	Any	Any	Any	<=5

I-3-2. Daily Trips and Tours

For each person day, the total number of tours by purpose and the total number of trips by purpose is aggregated. For each person travel day, a tour pattern is defined based on the combination of tour types that the traveler undertakes that day. Table I-3-1 shows the 40 tour patterns, and the list of person types allowed to make each tour pattern.

Table I-3-2. Tour Pattern Definitions

Person Types	Alt	Label	Work Tours W	Work-Based Subtrs. WB	School Tours SC	Univ. Tours U	Shop. Tours SH	Social / Rec. Tours SR	Other Tours O	Total Tours
Any	1	H	0	0	0	0	0	0	0	0
	2	SH	0	0	0	0	1+	0	0	1+
	3	SR	0	0	0	0	0	1+	0	1+
	4	O	0	0	0	0	0	0	1+	1+
	5	SH-SR	0	0	0	0	1+	1+	0	2+
	6	SH-O	0	0	0	0	1+	0	1+	2+
	7	SR-O	0	0	0	0	0	1+	1+	2+
	8	SH-SR-O	0	0	0	0	1+	1+	1+	3+
Full-Time Worker	9	W	1+	0	0	0	0	0	0	1+
	10	W-SH	1+	0	0	0	1+	0	0	2+
Part-Time Worker	11	W-SR	1+	0	0	0	0	1+	0	2+
	12	W-O	1+	0	0	0	0	0	1+	2+
	13	W-SH-SR	1+	0	0	0	1+	1+	0	3+
University Student (employed)	14	W-SH-O	1+	0	0	0	1+	0	1+	3+
	15	W-SR-O	1+	0	0	0	0	1+	1+	3+
	16	W-SH-SR-O	1+	0	0	0	1+	1+	1+	4+
Driving Age Child (employed)	17	W-WB	1+	1+	0	0	0	0	0	2+
	18	W-WB-SH	1+	1+	0	0	1+	0	0	3+
	19	W-WB-SR	1+	1+	0	0	0	1+	0	3+
	20	W-WB-O	1+	1+	0	0	0	0	1+	3+
	21	W-WB-SH-SR	1+	1+	0	0	1+	1+	0	4+
	22	W-WB-SH-O	1+	1+	0	0	1+	0	1+	4+
	23	W-WB-SR-O	1+	1+	0	0	0	1+	1+	4+
	24	W-WB-SH-SR-O	1+	1+	0	0	1+	1+	1+	5+
Driving Age Child	25	SC	0	0	1+	0	0	0	0	1+
	26	SC-SH	0	0	1+	0	1+	0	0	2+
Pre-Driving Age Child	27	SC-SR	0	0	1+	0	0	1+	0	2+
	28	SC-O	0	0	1+	0	0	0	1+	2+
	29	SC-SH-SR	0	0	1+	0	1+	1+	0	3+
Preschool Child	30	SC-SH-O	0	0	1+	0	1+	0	1+	3+
	31	SC-SR-O	0	0	1+	0	0	1+	1+	3+
	32	SC-SH-SR-O	0	0	1+	0	1+	1+	1+	4+
	33	U	0	0	0	1+	0	0	0	1+
University Student	34	U-SH	0	0	0	1+	1+	0	0	2+
	35	U-SR	0	0	0	1+	0	1+	0	2+
	36	U-O	0	0	0	1+	0	0	1+	2+
	37	U-SH-SR	0	0	0	1+	1+	1+	0	3+
	38	U-SH-O	0	0	0	1+	1+	0	1+	3+
	39	U-SR-O	0	0	0	1+	0	1+	1+	3+
	40	U-SH-SR-O	0	0	0	1+	1+	1+	1+	4+

3 Incomplete and Flagged Persons

Any flags identified at the trip or tour level are carried upwards to the person level.

I-4. Model Estimation Files

This section describes the additional survey processing needed to create estimation files for each of the models developed, as well as the structure of those files.

I-4-1. Auto Ownership Estimation File

The auto ownership estimation file is structured with one record for each household. In addition to the data included in the survey, TAZ data are merged based on the location of the household. These measures include county and area type codes, the density of households and employment within 1/2 mile of the TAZ centroid, a mixed density measure, and an approximate destination choice logsum value which serves as a measure of accessibility. The mixed density measure and logsums are similar to those used by Picado (2014) in a recent update of the NOACA auto ownership models. They are defined in Table I-4-1.

Table I-4-1. Mixed Density and Accessibility Measures (from Picado 2014)

Measure	Description & Formulas
Mixed Density Measures	*MDM = Ln {[Int*(Emp*a) * (HH*b)] / [Int+(Emp*a) + (HH*b)]}*
Mixed employment and household density	Where: *Emp*= Employment within 1/2 mile of centroid *HH*= Households within 1/2 mile of centroid *Int*=Intersections within ½ mile of centroid *a*= average Int / average Emp *b*= average Int / average HH
Accessibility Measures	*Accessibility variables are proportional to the number of opportunities (such as jobs or retail opportunities) that can be reached by auto, transit or walk means.*
Transit Accessibility Logsum	$TrLogsum_p = Ln\left(\sum_q \exp\left(-0.025 * TransitTime_{pq} + \ln(Emp_q)\right)\right)$ Where $TransitTime_{pq}$ is total transit time including a weight of 2 on all out-of-vehicle time components.
Auto Accessibility Logsum	$DrLogsum_p = Ln\left(\sum_q \exp\left(-0.025 * DriveTime_{pq} + \ln(Emp_q)\right)\right)$ Where $DriveTime_{pq}$ is total drive time during peak hour.
Composite Accessibility	Difference between Auto Accessibility and Transit Accessibility $CompLogsum_p = DrLogsum_p - TrLogsum_p$

I-4-2. Tour Generation Estimation File

The tour generation estimation file is structured with one record for each person-day. For the GPS-with-recall and log-only samples, there is one record for each person. For the GPS-only sample, there are 3 or 4 records for each person, depending on how many travel days they were assigned based on the starting day of week. Weekend days are excluded from the analysis, but processed nonetheless for completeness. In defining this structure, care is take to ensure that records are included when there is no travel on the day of interest. Density measures and accessibility measures are merged based on the home location in a manner similar to that used for the auto ownership estimation file.

I-4-3. Mode Choice Estimation File

The mode choice estimation file includes one record for each tour, with the primary tour mode defining the selected alternative. Level-of-service (LOS) measures are merged based on the round-trip impedance from the home TAZ (or work TAZ in the event of work-based subtours) to the primary destination TAZ, and back. LOS measures are included for all possible modes, not just the chosen mode. The analysis uses LOS skims derived from the NOACA model, based on the 1597 zone system. A mode is only considered to be available if there is a valid path in both directions. Accessibility and density measures are also merged.

I-4-4. Destination Choice Estimation File

The destination choice estimation file also includes one record for each tour. The chosen alternative in this case is the TAZ of the primary destination, selected from the set of all possible TAZs. For each record, rather than include all TAZs in the estimation file, a sample of 40 possible alternatives is included. This sample is selected based on the proximity to the home TAZ, and the employment and households in the TAZ being selected. A correction factor is applied in model estimation to ensure that the sampling does not bias the model estimation results.

For each sampled alternative, a set of size measures is added that includes the households and employment in the zone. Employment is segmented by industry, and aggregated to the TAZ level from the LEHD data. Impedance measures are joined for each sampled TAZ that included the round trip auto time and distance, and a mode choice logsum. The mode choice logsum is a composite measure of impedance across all available modes. It is added after the mode choice estimation is complete, running the destination choice file through ALOGIT using the preferred mode choice model.

References

Picado, R. (2014). Technical Memorandum #3: Upstream Model Updates. To: NOACA Model Update Project Team, From: Rosella Picado, Parsons Brinckerhoff, October 8, 2014.

Appendix J. Python Code for Jackknife Application

```
"""
gde 3 Feb 2016
NCHRP Project 08-36 Task 123

python script to produce jackknife variance estimates for choice models.
This is done by repeatedly running ALOGIT with different estimation weights,
then combining the results.

usage:

    python jackknife.py [conrolfilename.alo] [baseweight] [samplename]

    i.e.

    python jackknife.py ao21.py PERWGTU FullSample

"""

import sys
import os
import glob
import time
import subprocess
import shutil
import pandas as pd
import numpy as np

# constants
NUM_REPLICATES = 453


def readParameters(f12_filename, replicate_num):
    """
    Reads the parameter estimates from the specified file.  The file should
    be a .F12 file, as created by ALOGIT.  replicate_num is an integer
    used to specify the column name for the estimate.

    Returns a pandas dataframe with two columns--the name and the estimate.
    """

    # read the file
    df = pd.read_csv(f12_filename,
                     skiprows=3,
                     header=None,
                     delim_whitespace=True,
                     error_bad_lines=False,
                     warn_bad_lines=False)

    # drop the stuff at the end that doesn't fit our column structure
    df = df[df[0]==0]

    # keep only the columns of interest
    df = df.rename(columns={1 : 'NAME' , 3 : 'ESTIMATE' + str(replicate_num)})
    df = df[['NAME', 'ESTIMATE' + str(replicate_num)]]

    return df
```

```python
# main function call
if __name__ == "__main__":

    # get command line argument
    CONTROL_FILE = sys.argv[1]
    BASE_WEIGHT  = sys.argv[2]
    SAMPLE_NAME  = sys.argv[3]

    base_name = os.path.splitext(CONTROL_FILE)[0]

    # start by running the base model with standard weights
    f = open('weights.txt', 'w')
    f.write('weight = ' + BASE_WEIGHT)
    f.close()

    subprocess.check_call('alo4ec ' + CONTROL_FILE)

    # now run with each of the replicate weights
    for i in range(1, NUM_REPLICATES + 1):

        write_complete = False
        while not write_complete:
            try:
                f = open('weights.txt', 'w')
                f.write('weight = ' + BASE_WEIGHT + str(i))
                f.close()
                write_complete = True
            except (IOError):
                print IOError
                time.sleep(3)

        replicate_control_file = base_name + '_' + str(i) + '.alo'

        shutil.copy(CONTROL_FILE, replicate_control_file)
        subprocess.check_call('alo4ec ' + replicate_control_file)

    # get the full-sample parameter estimates
    params = readParameters(base_name + '.F12', 0)

    # get and merge each of the replicate estimates
    for i in range(1, NUM_REPLICATES + 1):

        replicate_params = readParameters(base_name + '_' + str(i) + '.F12', i)
        params = params.merge(replicate_params, how='left', on=['NAME'])

    # calculate the jackknife variance estimates
    params['JACKKNIFE_VARIANCE'] = 0.0
    jackknife_total = params['JACKKNIFE_VARIANCE']

    rf = pd.read_csv('repfactors.csv', index_col='index')['REPFACTORS']
    for i in range(1, NUM_REPLICATES + 1):

        jackknife_i = rf[i] * ((params['ESTIMATE' + str(i)] - params['ESTIMATE0']).a
        jackknife_total = jackknife_total + jackknife_i

    params['JACKKNIFE_VARIANCE'] = jackknife_total

    # write all params to a CSV file
    params.to_csv(base_name + '_' + SAMPLE_NAME + '_parameters.csv')

    # clean up unneeded files
    for f in glob.glob("*.F11"):
        os.remove(f)
```