ARC Regional Transit Survey Project Work Plan

Conducted on behalf of:
Atlanta Regional Commission (ARC)

Prepared by:

ETC Institute
725 W. Frontier Circle
Olathe, Kansas 66061

March 25, 2019

1.0 PROJECT OVERVIEW

The Atlanta Regional Commission's (ARC) Regional On-Board Transit Survey serves several objectives, including:

- Compile statistically accurate information about transit customers' use of different transit services for planning purposes.
- Generate reliable linked Origin/Destination (OD) data to support the travel demand model (both trip & tour based) and transportation network simulation for long-range transportation planning activities.
- Understand differences in trip characteristics and ridership profiles from the region's 2010 survey.
- Assist ARC and regional transit operators in meeting Title VI Civil Rights Requirements.

The ARC Regional On-Board Transit Survey will include the following transit providers:

- MARTA
- SRTA
- Cobb County Transit
- Gwinnett County Transit
- Cherokee Area Transportation System
- Hall Area Transit

2.0 SCOPE OF WORK

2.1 Develop a Work Plan and Schedule

Upon the execution of a contract, the ETC project manager will make an on-site visit to meet with the ARC project manager, Advisory Team, and Transit Operators Subcommittee to conduct a project start-up meeting in anticipation of developing the project Work Plan. Following the initial on-site visit, ETC will develop a Work Plan to clearly define communication protocols and a detailed task level project schedule.

During the initial project start up meeting, ETC will work with ARC to develop specific data needs. We anticipate working with ARC and the transit agencies to capture the following:

- Ridership data – APC at the stop level and ridership accumulated at the route / direction / time of day cell level where APC data is available. This will be used in conjunction with the Farebox data to plan for the data collection and expand the data.

- GIS data – GTFS or Shape files / bus stop lists. These exist online but are preferred to come from the individual transit agencies as the online postings are slightly inaccurate.

- Schedule data – Trip level information synonymous with block or driver / runcut shifts to plan surveyor and interviewer shifts.

- Operational procedures – transit agency contacts, boarding procedures, command center potential locations, passes for interviewers, system ridership characteristics, dates not appropriate for surveying, difficult to reach populations, etc.

As part of the initial on-site visit, ETC and ARC will schedule a regular meeting time to help maintain good communication throughout the project. In addition, team members will work with the ARC to coordinate review of the project with FTA staff. ETC will coordinate directly with transit operations and garage staff well in advance of fielding the survey. Team members are well versed in this type of direct interaction and have built that communication time into the budget and schedule for this effort.

Task 1 Deliverables:
- Working Paper #1 – Work Plan

2.2　Survey Design and Sample Plan

2.2.1 On-to-Off Counts Sample Plan
On-to-Off (O2O) counts (also known as the Boarding-to-Alighting Counts) will be conducted prior to the OD intercept survey. The results of the O2O counts will allow ARC to understand ridership flows between stops, which would allow ETC to develop a more sophisticated sampling plan with sampling targets for Stop Groups and/or trip paths on each route. Not only would the O2O counts ensure that all types of trips, particularly short-trips, are well-represented in the OD survey, it would enhance ARC's ability to expand the data based on the trip-path rather than just the boarding or alighting location.

ETC Institute will develop a sampling plan for conducting On-to-Off counts on all routes with an average weekday ridership of 3,000 or more except the MARTA rail service because entry-to-exit data for the MARTA rail system is already collected.

The sampling plan will include the following:

- The routes to be surveyed
- The goal for completed surveys, which will be approximately 20% of the daily ridership
- The trips on each route that will be surveyed (approximately 25% of the trips on each route will be selected; the goal will be to get completed O2O counts from approximately 90% of the riders on each of the trip selected, which will yield a sample of 20% of the daily ridership for the route)

2.2.2 Origin-Destination Survey Sample Plan

As part of the process for developing the sampling plan, the ETC team will review the following:

- What are the proposed uses for the survey? (e.g., model calibration/validation, Title VI analysis, before/after study)
- What variables of interest do these uses require? (e.g., income, access/egress mode, trip purpose)
- What level of detail is needed for the variables of interest? (e.g., volume for each stop-group interchange, income-level distribution by route, or a combination of income-level by stop-group interchange)
- What is the desired level of statistical accuracy for large markets and small markets? How should large and small markets be defined?
- What are the best ways to control sampling rates (e.g., having different sampling rates for different markets, analyzing/collecting ancillary count data, and/or analyzing data from previous surveys to determine where cells could be combined)
- What biases exist that could impact data collection and how will they be addressed? Common biases include: Access Bias, Language Bias. Age Bias, Afternoon and Evening Fatigue Bias, Heavy Load Conditions Bias, Short Trip Bias, Non-Work Trip Bias

The sampling plan for the Origin-Destination survey will include the following:

- The routes to be included in the survey. [Note: ETC does not currently plan on conducting system-wide O-D survey for SRTA since ETC conducted a system-wide survey for SRTA in spring 2018, but the sampling plan for the 2019 survey will include goals for all SRTA routes to ensure that ARC modeling needs are met. If a sufficient number of surveys were collected in 2018, the 2018 surveys will be used to meet these goals. If not, more surveys will be conducted on SRTA routes as needed to ensure the goals for this project are met.]

- The sampling rate for each route, each transit agency, and the entire system

- The number of surveys that will be completed by time of day and direction on each the routes that are to be included in the survey.

- Methods for controlling sampling biases

- The definition of a "complete" survey that is useable for both modeling and Title VI analysis. Typically, this would mean that at least 90% of the Title VI related questions on the survey are answered and all the following components of the trip path will be geocoded to X, Y coordinates.
  - The trip origin location.
  - The location where the passenger initially entered the transit system from the origin.

- o All locations where transfers on transit took place for the one-way trip.
- o The location where the passenger finally exited the transit system to the destination.
- o The trip destination location.

- A list of the beneficiaries of protection under Title VI that are to be reached and methods for ensuring these populations are properly represented in the survey. At a minimum the sampling plan will address methods for reaching the following groups:
  - o LEP populations (it is anticipated at this time that Spanish, Korean, Chinese, and Vietnamese survey options will be needed)
  - o Non-Whites
  - o Ethnic minorities, particularly those of Hispanic/Latino ancestry
  - o Persons in low income households

2.2.3 Develop Data Analysis and Model Integration Plan

Since a key purpose of the on-board survey is to improve the transit forecasts generated by ARC's travel model. In this sub-task, the ETC Team will develop procedures to integrate the results from the on-board survey and ancillary data collection efforts into calibration/validation targets for the travel model, as well as propose data exploration tasks to help validate the assumptions about transit travel behavior built into the ARC activity-based model.

As noted above, the survey will collect data through the following three sampling frames: (a) phone interviews/web-based survey with Breeze customers; (b) intercept personal interviews with transit riders; and, (c) location traces from travel diary respondents. The table on the following page summarizes example metrics from each of these frames that can be used in model development efforts. In coordination with ARC, we will expand on these metrics as part of the project and then, for each of selected measures, we will prepare summaries from at least one other existing data source to create a set of expectations for the completed survey (the other data source may be the 2009 Atlanta on-board survey).

Frame	Measure	Notes
Breeze	Socio-demographics of riders	Can be used in validation
Breeze	Automobile ownership rates	
Intercept	Automobile ownership rates	
Intercept	Access/egress movement details	
Intercept	Transfer rates (bus to bus, bus to rail)	
Intercept	Path parameters	Calibrated by assigning the on-board survey records
Diary	Stops on transit tours	

Diary	Frequency of complex tour behavior	For example, how often do transit commuters take an Uber or Lyft home?
Diary	Frequency of automobile use for non-automobile owning households	
Diary	Refined estimates of walk access & egress distances	
Diary	Refined estimates of experienced wait times	

To save time on the back-end, we will build a cross-walk between the names of the transit routes used in the ARC travel models and the names used in the survey instrument. This will allow for easier comparisons to model outcomes when the data is processed.

One possible way of organizing the data analysis task is by formulating upfront hypotheses of transit rider behavior. The advantage of formalizing both what we think we know about transit rider behavior, and what we wished we knew, is that it can inform the specific information that is solicited from the survey respondents, and it focuses the data analysis from the beginning into generating useful insights, rather than general statistics. The end goal of this task is to construct a data-driven set of behaviors which can be used to demonstrate that the model gets it right, or to identify areas where the model needs improvement. It may also be useful to codify into hypotheses the behaviors that are already built into the model, so that the survey data analysis task produces statistics that are useful to prove or disprove these assumptions. As part of this task, we will collaborate with ARC and other stakeholders to formulate a set of hypotheses that will guide the data analysis undertaken in Task 7.

Task 2 Deliverables:
- Working Paper #2 – Survey Design and Sample Plan, which will include the following sections
    - Section 1: O2O Counts Sampling Plan
    - Section 2: Origin-Destination Survey Sampling Plan
    - Section 3: Data Analysis and Model Integration Plan

2.3   Develop Survey Instrument

2.3.1 On-to-Off Counts Instrument
ETC will conduct an O2O counts collection on fixed route weekday routes that are participating in this collection. The survey will be conducted from beginning to end of the eligible times on each route. The specific information collected on each survey includes the following:
- Route
- Direction of travel
- Boarding location
- Boarding time
- Alighting location

- Alighting time

The O2O counts will be collected using tablets with scan guns attached and bar-coded cards. This will allow for accurate time stamps and GPS locations to be captured. As riders board the vehicle, the counter at the front of the bus will scan a card, locking in time and location for the boarding, and distribute the card to the rider. As riders alight the vehicle, the rider will return the card to the second counter who will scan the card to capture the alighting information.

2.3.2 Origin-Destination Questionnaire

ETC will design a tablet-based personal interview survey as the primary survey medium and will use the questionnaire from ARC's previous survey effort as a starting point. The survey will focus on collecting 1) all elements of the "one-way" trip, including purpose, origin and destination addresses, boarding and alighting locations, access and egress modes, and transfers made, 2) information needed for the travel demand model and transit planning, including person and household demographics and activity-based model questions, and 3) Title VI information to be used for transit agency compliance. ETC understands this process will involve close and careful coordination with the ARC Project Team and the FTA.

For the full personal interview surveys, ETC will use tablet PCs integrated with GIS software to allow for accurate geocoding of most survey data as the survey is taken. At a minimum, the data to be geocoded will include:

- Boarding and alighting stops for current route
- All transfer locations for routes transferred to/from on the one-way trip
- Trip origin & destination locations
- Home address/location

In addition to the address information, the Traffic/Transportation Analysis Zone (TAZ) in which each location resides will be captured, which will minimize the need for extensive data entry/cleaning efforts after the survey is administered.

While the tablet-based interview will be the primary medium used for data collection, other mediums will also be utilized to make the best use of project resources. For local bus riders who don't have enough time to complete the interview, a phone option is available. For express routes, hard copy self-administered questionnaires may be employed. The rationale is two-fold. First, these riders are on the vehicle for a long period of time thus allowing more than enough time to complete the instrument. Second, these riders are generally more educated and thus have a better understanding, and appreciation, for the data collection purpose and elements needed. This was the methodology used for the 2018 SRTA Systemwide OD On-board Survey conducted by the ETC team.

For non-English speaking ridership, an attempt will be made to allocate bilingual interviewers to routes most likely to contain linguistically isolated Spanish speaking riders. Additionally, an option will be available if a non-bilingual interviewer is not on the vehicle

that will describe the purpose of the survey in their native language and allow a name and phone number to be provided so a callback attempt can be made in that language. It has been ETC's experience that hard copy foreign language questionnaires are rarely usable and programing the tablet in these languages is not very beneficial either as it still functions as a self-administered questionnaire. With that said, these options will be made available if desired by ARC.

Knowing that connectivity can be an issue on MARTA rail, testing using different service providers will be executed to determine which provider works best. If needed, an offline version of the questionnaire will also be designed for administration. This will be utilized for MARTA rail and any other services that have connectivity issues.

To verify that the sampling plan was effective at gathering a representative sample of beneficiaries of protection under Title VI, data items such as race, ethnicity, English proficiency, language spoken at home, household size, and household income will be included on the survey instrument. These data items will be used during Task 5 to evaluate how well Title VI beneficiaries are being reached and identify any potential changes to improve the response volumes during Task 6.

Task 3 Deliverables:
- Working Paper #3 which will include the following:
    - On-to-Off Tablet-based survey program that has been approved by ARC
    - Origin-Destination Survey Tablet-based survey program that has been approved by ARC
    - Copy of the survey questionnaires in English approved by ARC

2.4   Interview Training Manual

2.4.1 On-to-Off Counts Training
ETC will conduct training on how to administer the O2O counts. The training will involve both classroom training and field training for each employee. ETC will also be responsible for preparing all training materials, securing a facility for conducting the training, and providing field supervisors to conduct the training. The specific topics covered during the training will include:
- How to properly use the interviewing equipment
- Sampling procedures
- How to approach riders to ask them to participate
- How to react in various situations that may be encountered

Each counter will be required to demonstrate proficiency in conducting the O2O by the end of the training day. Counters who cannot demonstrate proficiency in all tasks related to the administration will be given additional training. If they still are not able to properly conduct the counts they will be replaced.

2.4.2 Origin-Destination (OD) Survey Training (Pilot and Full Surveys)
ETC will hold trainings on how to conduct the OD Survey. The training will involve class

room training, including mock scenarios, field training, and follow-up training per qualified interviewer. ETC will also be responsible for preparing all training materials, securing a facility for conducting the training, and providing field supervisors to assist in the training. Some of the topics that will be covered during the training will include, but will not be limited to:

- How to properly use the interviewing equipment
- Sampling procedures
- How to react in various situations that may be encountered
- Procedures for conducting debriefs with riders who complete printed versions of the survey
- How to conduct the survey if rider has at least 5 minutes for the survey
- How to conduct the survey if the rider has less than 5 minutes for the survey
- How to conduct the survey if the riders doesn't speak English
- What to do if a rider refuses to participate in the survey
- Safety protocols for boarding, riding, and alighting vehicles

Each interviewer will be required to demonstrate they can proficiently conduct the OD Survey. Interviewers who cannot demonstrate proficiency in all tasks related to the administration of the survey will be replaced if re-training is not successful.

Task 4 Deliverables:
- Working Paper #4 – Survey manual and training slides

2.5   Conduct and Evaluate Pre-Test

ETC will conduct a pre-test of the OD survey. The pilot test will involve the administration of at least 200 surveys and would replicate all tasks that would be conducted for the administration of the full survey. The pilot test will be conducted during the O2O collections using the survey staff best suited for the OD collection. ETC will document the results of the pilot test along with any changes to the survey instrument or survey methodology discovered as a result of the pilot test.

Using the pre-test data to inform the expected data specifications, the ETC Team will create the summary scripts necessary to create the model estimates identified in Task 2. As the full survey is conducted, these scripts will be executed on interim data deliverables to assist in monitoring data quality.

As part of the pilot test, ETC Institute will review response rates from LEP, minority, and low-income populations will be analyzed to ensure the sampling plan gathered a representative sample of Title VI beneficiaries. Based on the results of this analysis, the sampling plan will be adjusted to increase potential responses if necessary.

Task 5 Deliverable:
- Working Paper # 5 – Memo Summarizing the pilot test

2.6   Support Public Involvement Efforts

2.6.1 Support Public Involvement Efforts
Prior to the beginning of the collection, ETC will work with ARC and the transit agencies to ensure that both transit riders and internal staff are made aware of the upcoming effort. Specifically, examples of the following will be provided:

- Outreach tailored to transit police officers, operators and appropriate supervisors, including information on survey dates, times, procedures, interviewer identification, and who to contact in case of a problem or incident related to the survey.

- Outreach tailored to riders that increases their awareness and encourages their participation.

ETC will provide sample materials to support outreach efforts, but the production and distribution of the final materials will come through ARC and placement of any materials would be through the individual transit agencies.

2.6.2. Provide the Data Collection Team
ETC Institute will field the data collection teams to administer the On-to-Off and Origin and Destination surveys. The organizational structure of the team is described below.

- **Leadership Team -** ETC Institute's leadership team will consist of the project principal, project manager, assistant project manager, and data managers. The leadership team will be responsible for reviewing the performance of the data collection and ensuring that the sampling goals for each route are met.

- **Field Supervisors -** ETC Institute's field supervisors will oversee a group of approximately 20-30 counters/interviewers covering the entire service day. They will be deployed on buses / trains based on sample size requirements. The field supervisors will be responsible for ensuring interviewers are properly trained and equipped to conduct interviews, scheduling interviewers, inspecting work, and reviewing the data collected before submitting the data to the leadership team at the end of the day.

- **Data Collection Personnel -** Understanding that this will be a significant data collection, ETC will use its vast experience to cultivate quality staff over the life of the project. To accomplish this, ETC is continuing to partner with national staffing firm and DBE certified Stat Team along with ANIK International. These firms will be responsible for recruiting English, Spanish, and other foreign language speaking interviewers. ETC and Stat Team have successfully worked together for more than three dozen on-board transit studies conducted by ETC since 2014. This experience and understanding allows ETC to provide the ideal staff needed for this type of niche research.

2.6.3. Conduct the Surveys
The survey will be administered on non-Friday weekdays when schools are in session and not on holidays. Fridays are traditionally not surveyed due to the potential for atypical weekday travel patterns.

Interviewers will select riders at random to participate in the survey based on the sampling goals established for each route. Once an interviewer has randomly selected a rider for the survey, the interviewer will do the following:
- Approach the person who was selected and ask them to participate in the survey.
- If the person refuses, the interviewer will end the survey, but the refusal will be recorded on the tablet PC to help assess the overall response rate to the survey.
- If the rider agrees to participate, the interviewer will ask the respondent if he or she has at least five minutes to complete the survey.
- If a rider has at least five minutes, the interviewer will administer the full survey to the respondent as a face-to-face interview using a tablet PC.
- If the rider does NOT have at least five minutes, the interviewer will ask the rider to provide their name and phone number and collect as much of the remaining survey information as possible before alighting the vehicle. Shortly after this information is collected, a phone interviewer from ETC will contact the respondent and ask him or her to provide the remaining information from the questionnaire by phone. This methodology will ensure people who complete "short trips" on public transit are represented.
- If the rider does not speak English and the interviewer does not speak that foreign language, (most likely Spanish) the tablet will be programmed with instructions in that foreign language to provide their name and phone number and they will be contacted by a call center employee who speaks that language (similar to those who have less than five minutes).

ETC will utilize the survey tracking management developed for previous efforts that allows for a variety of tracking and monitoring features for data quantity and quality. The ETC staff will be able to view data in three primary ways: individual trip review, collection allocation, and collection distribution.
- Individual trip review – By visually reviewing the individual trip, with built in flags for inconsistencies, field supervisors can better track and instruct interviewers, thus continuing the training process.
- Collection allocation – As interviews are completed and verified, the website updates the progress to goals for both route/direction/time of day along with individual interviewers.
- Collection distribution – The collected data can be segmented by route, interviewer, trip attributes, and demographics in various combinations as a quality control measure.

This will allow for the field manager and supervisors to better understand and plan for:
- Training interviewers.
- Staffing to ensure an adequate number and coverage of interviewers.

- Tracking and weekly reporting interviewer performance in the field.
- Planning for remediation of poor interviewer performance.

ETC Institute will monitor income information and the response rates for income throughout the project and coach interviewers to maximize our ability to capture this information. If needed, ETC Institute will use imputation methods that developed for MTC to maximize the number of records that have income information.

Task 6 Deliverables:
- Working Paper #6 – ETC Institute will provide ARC with access to a data portal that will give ARC real-time access to the data collection progress in the field along with regular reports showing our progress on a weekly basis during data collection efforts

2.7   Data Processing and Analysis

2.7.1 Data Processing
ETC will develop a quality assurance/quality control (QA/QC) plan to guide all data collection and handling activities for the data collection. The QA/QC plan will address:

- How to ensure the collection of enough individual rider samples to assure statistical sampling goals in the sampling plan are met.
- The definition for a "complete and usable" survey.
- Procedures for conducting a random sampling of the completed survey records to assess the quality of the collected survey records and account for bias.
- Delineate how real-time QA/QC will be conducted and outline ETC's approach for ensuring that data collection ends with enough surveys to meet the guaranteed number of complete and usable surveys.

ETC will utilize proprietary mapping software to ensure reported origin, boarding, alighting, and destination locations are reasonable and consistent with reported transit route sequences to ensure all survey records are logical and geocodable. The software allows for the entire one-way trip to be mapped and reviewed for each portion of the trip. The following checks are displayed in map form:
- Visual review of data.
- Feasibility of transfers.
- Ratio of the access distance to transit from the rider's origin relative to the distance traveled on transit.
- Ratio of the egress distance from transit to the rider's destination relative to the distance traveled on transit.
- Total travel distance relative to the mode of access and egress.

In addition to systematic QA/QC checks, ETC implements a policy ensuring up to 10% of survey participants receive callbacks to evaluate their survey experience. This allows

ETC to validate that the interviewers are doing their job in a professional manner and completed records are not being artificially created.

- *In-Field Data Checks* - In-field checks will be addressed in two major ways. First, the tablets will be programmed so that many logic checks are automated. Second, and as described in the administration portion of the proposal, field supervisors play an extremely important role in the day-to-day monitoring of the production, in terms of quantity, quality, and distribution of samples collected.

- *Data Processing* - The ETC data review team will review all records through our ELVIS data review platform which allows the visual review off all interviews collected. If any of the required information is missing or incomplete, the data reviewers will forward the survey record, and corresponding name and phone number of the survey respondent, to ETC's call center. Interviewers working in ETC's call center will follow up with respondents who provided their name and phone number to retrieve the missing information by phone in the same manner as those who were only on the vehicle for a short period of time. When this is not possible, ETC has developed queries to analyze and visualize the data without additional respondent information. Once survey records have been classified as "complete," meaning all required information has been collected, the records will be forwarded to ETC data managers who conduct the final processing and geocoding.

2.7.2 Preliminary Data Expansion

Unlinked Trips. ETC will develop unlinked expansion factors for all bus and rail survey records. All records will be expanded by route, direction, and time of day. Most records will also be expanded by path (segment/station on-to-segment/station off).

While there are no specific guidelines for the expansion factor values, ETC uses a guideline of keeping expansion factors below 3 times the average expansion factor based on the sampling percentage. This is done to keep any one record from representing a markedly high number of riders in the system. The formula for determining this guideline is:

1 / (Sampling %) x 3 = Guideline Weight Factor

If the expansion factor for a boarding segment to alighting segment pair is greater than 3 times the average expansion factor, then it is aggregated into the adjacent boarding to alighting segment where it will have the least impact on the previously existing expansion factors. This guideline is standard for all the various expansion types.

Linked Trips. The linked trip expansion factor helps to account for the number of transfers that were made by each rider, so the linked expansion factors can better represent the overall system. Linked expansion factors are generated after the unlinked expansion factors are created.

The equation that is used to calculate the linked trip multiplying factor is shown below:

Linked Trip Multiplying Factor = [1 / (1 + # of transfers)]

If a rider did not make a transfer, the linked trip multiplying factor would be 1.0 because the person would have only boarded one vehicle. If a person made two transfers, the linked trip expansion factor would be 0.33 because the person would have boarded three transit vehicles during his/her one-way trip. Once the linked trip multiplier is created it is multiplied by the unlinked expansion factor to create the linked expansion factor.

Linked Trip Decomposition Analysis. Following the expansion, ETC will perform a linked trip decomposition analysis to understand how the linked-trip weights represent actual ridership. On a typical OD study, an unlinked-trip weight is calculated based on the average weekday ridership on the route in which the respondent was surveyed and does not consider whether they transferred to or from other routes during their trip. A second weight is calculated (the linked-trip weight) which does consider the number of transfers made. The decomposition analysis reviews all transit routes used by survey respondents and looks to see how many riders transferred to each route and from each route. This allows us to determine whether the total ridership estimated from the linked-trip weight using all the routes/lines adds up to the total boardings on a route as well as the total boardings for the entire system.

2.7.3 Secondary Data Expansion
After the preliminary expansion factors are developed, CTG will fine tune the expansion process to reduce biases that may occur in the survey due to challenges in collecting the survey data. Our significant experience with applying transit surveys for transit planning indicates that the secondary expansion is desirable to correctly represent important rider or trip types, such as students and park-ride trips, and to ensure the resulting trip patterns are consistent with how riders are using the system. In the description below, on-to-off flow counts are assumed, but the secondary expansion can be performed even without this data.

CTG utilizes three methods, described below, as part of the secondary expansion process. The preferred method can depend on:
    (1) The desired breakdown of rider and/or trip types that the survey records should match
    (2) The extent of response bias in the collected dataset, and/or
    (3) The corresponding data available to the rider/trip types.

A different method can be applied to each route, allowing the secondary expansion to maximize all data available from various ARC transit agencies.

Method #1: Standard Expansion Process (SP). This method follows the traditional expansion equation by dividing counts by the number of records collected.

$$SEF_a = \left\{ \left(\frac{NumTrips_a}{NumRecords_b} \right) x \left(\frac{NumTrips_a}{NumTrips_b} \right) \right\}, \forall \ a \in b$$

Where: SEF is the Secondary Expansion Factor for rider or trip type *a* (which is a subset of the larger group *b*), NumTrips is the number of trips for the rider/trip type, and NumRecords is the number of survey records for the rider/trip type.

A frequent example occurs where university students commonly use certain routes (see the following table). University students generally participate in surveys at a higher rate than the general population, so a traditional expansion of the survey records would over-estimate student ridership. In the table below, a route with 800 boardings is surveyed. This route also has counts by student/non-student via electronic fare card data. Using this data, Method #1 can correct the over-sampling of university students and report the right number of trips by rider type.

	Survey Records in Collected Dataset	Primary Expansion Factor	Estimated Boardings (from Preliminary Expansion)	Fare Card Counts	Expansion Factors (from Method #1)	Estimated Boardings (from Method #1)
University Students	50	10.0 (= 800 / 80)	500	250	5.0 (=250/50)	250
All Other Rider Types	30		300	550	18.3 (=550/30)	550
Total	80		800	800		800

The method is straightforward and easy to apply. However, it does not address any rider/trip type gaps in the collected dataset; that is, a rider type was not surveyed at all or was missed due to day-to-day variations in travel patterns. Also, the Method #1 computations can become unwieldy with more than two rider/trip types above those used in the primary expansion.

Method #2: Synthetic Survey Record Enhancement (SSRE). This method builds upon Method #1 by identifying gaps in the collected data – where rider/type trips are absent in the collected dataset – and resolving those gaps by using similar records from the collected dataset and modifying their data fields that are relevant to the data gap. For instance, assume that four types of rider counts are collected at a major transfer stop (see table below), but the fourth rider type is missing from the collected dataset (row 4, column 2). The primary expansion overlooks this data gap, while Method #1 is unable to compute a factor.

	Count	Survey Records in Collected Dataset	Primary Expansion Factor	Method #1 Expansion Factor	Method #2 Expansion Factor	Method #2 Total Survey Records
Rider Type 1	100	10		10.0	10.0	10
Rider Type 2	50	5		10.0	10.0	5
Rider Type 3	25	5	10.0 (= 200 / 20)	5.0	5.0	5
Rider Type 4	25	0		Cannot compute (= 25 / 0)	2.5 (=25/10 synthetic records from other rider types)	10
Total	200	20				30

Method #2 randomly samples the collected records that are as similar as possible to the data gap. In this case, it randomly selects 10 of the 20 records (at the same transfer stop) already collected and modifies the survey number and rider type data fields – keeping all other information identical – to reflect rider type 4's characteristics. This method assumes that riders who travel with type 4 riders are likely to have similar trip patterns and characteristics. The result is that Rider Type 4 riders are accounting for in the secondary expansion and the dataset now includes 30 records, 20 from the originally collected dataset and 10 synthetized records for rider type 4.

Method #3: Multivariate Error Minimization (MEM). This method is best applied when synthetic records are not desirable, but the collected dataset still needs – as best it can – reflect a wider array of rider and trip types. Method #3 uses an optimization approach to minimize error by considering multiple rider and trip types simultaneously, yielding a single error-minimized expansion factor for each survey record. Priority can be given to selected rider/trip types over others at ARC's discretion. The error optimization algorithm uses the initial expansion factors, the priority settings, and the counts for each rider/trip type. The procedure is flexible and can include multiple priority settings to help determine which setting minimizes the overall error.

A summary of all three methods is shown in the following table. Different methods can be applied to individual routes, depending on the available data, the diversity of rider/type types on the route, and the desired approach to handle data gaps. These methods and the ability to execute at the route-level provides ARC with maximum flexibility and produces a final dataset that can be used effectively.

	Standard Process (SP)	Synthetic Survey Record Enhancement (SSRE)	Multivariate Error Minimization (MEM)
Available counts/data for rider/trip types (in addition to types used in primary expansion)	1-2 types	3-5 types	5+ Types
Resulting error for all rider/trips types (weighted records)	0%	0%	>0%
Survey Records in Final Dataset	Collected records only	Collected + synthetic records	Collected records only
Addresses rider/trip types missing in primary expansion (data gaps)	X	√	X
Can be applied to specific routes	√	√	√

Before executing the secondary expansion process, CTG will present the expansion methods, a proposed application for all three methods, sources of observed data, and proposed priority settings to ARC. Agreement on these items with ARC will be required before expansion may begin.

After the expansion process is complete, CTG will perform the following tasks:

- Reviewing and updating transfer rates and access/egress modes at trip ends by performing analysis such as distance between origin and boarding location, alighting and destination location, origin to destination distance, trip purpose etc. This is important for determining the linked transit trips. For most analyses, (Before and After reports, mode calibration, etc.) linked transit trips is the critical variable and issues with it due to transfer rates or access/egress modes causes issues with utilizing survey for the FTA purposes. Often CTG has re-computed these for majority of the on-board surveys.

- Converting the linked transit trips to an observed survey trip table by path (rail only, rail plus bus, bus only etc.), time of day (as in the model), access/egress mode and assigning it to the network model to identify remaining issues with the survey such as geo-coding issues, time of day issues, mode of access/egress issues. These un-assignable survey records will be flagged and either corrected for or excluded for the final set of expansion factors.

- Developing district to district flows by critical variables of interest (income, trip purpose, access/egress mode, transfer rates, paths) to ensure sensibility with

cross-tabulations that will assist with utilizing expanded data for any B&A studies and model calibration/validation work. Special attention to the statistical accuracy of these cross-tabulations is needed, as our research work with FTA has strongly indicated that N-dimensional cross-tabulations, where N≥3, are likely to provide statistically insignificant information.

2.7.4 Analysis and Travel Model Network Assignments

Assigning the on-board survey trip records to the travel model networks accomplishes at least the following three key tasks:

- tests the integrity of the travel model network representation and coding;
- examines the integrity and logic of the survey record geo-coding; and,
- establishes and/or verifies the travel model path building parameters.

The ETC Team will create aggregate trip tables from the survey records per the ARC model transit assignment market segments. These trip tables will be assigned to the appropriate Cube model network. Analysis of the assigned and un-assigned trip statistics will reveal several potential actions, including the following:

- revisions to the procedures that create walk access/egress, walk transfer, and drive access connectors;
- refinements to the path building parameters; and,
- revisiting select survey record coding (to identify infeasible trips, for example).

Assigning the on-board survey can also be used to test hypotheses of transit user behavior, with the longer-term goal of modifying the ARC activity-based model's transit market segmentation. Such hypotheses may include examining alternatives to the current segmentation of transit modes into "all-transit" vs. "premium-transit", introducing user class segments (high value of time versus low value of time users, or slow walkers versus fast walkers, etc.), and tradeoffs between type of stop and proximity to stop.

ETC Institute will have prepared the model summary infrastructure using the pre-test data in task 5. As part of this task, ETC Institute will finalize the summaries.

Task 7 Deliverables:

- Final report, electronic version and five printed versions
- All data files and dictionaries
- PowerPoint presentation

2.8   Prepare Final Report and Datafiles

ETC will prepare a final report and submit all datafiles for the project.

2.8.1. Submit Written Report.

The final report will consist of the working papers that will be been submitted for each task along with a written summary of the major findings.

As part of this task, ETC Institute will prepare a technical memorandum detailing the methodology used to collect data necessary to ensure Title VI compliance, the sampling plan, data items collected, and the results of the pilot test conducted during Task 5 and any adjustments will be produced. Additionally, this technical memorandum will report the response volumes and confidence levels for surveys completed by LEP, minority, and low-income persons. This technical memorandum will be provided as part of the final documentation.

2.8.2. Prepare Data Portal
In addition to the written reports, ETC will create a data visualization portal to allow ARC to review the data. The elements of this portal are described below.

2.8.2.1 Data Dashboard
ETC will set up an exhaustive set of data access and visualization tools for use by ARC. These tools will be accessible from the Project Portal and will provide many additional pages of dynamic data views, summaries, charts, and maps, beyond the basic set up included in the analysis. The tool set will include these pages:

- Primary demographic page showing primary demographic data for collected surveys. Primary demographic data include respondent information such as age, race/ethnicity, gender, and income. These data can be filtered by date, route, as well as any of the primary demographic answer choices. Data are typically displayed in bar or pie charts.
- Two secondary demographic pages showing secondary demographic data for surveys collected. Secondary demographic data include respondent information such as employment status, student status, driver's license, fare type, disability status, number of autos in household, and number of workers in household. These data can be filtered by date, route, as well as any of the secondary demographic answer choices. Data are typically displayed in bar or pie charts.
- Trip data page - a page to include trip data such as address information, i.e. home address, origin address, destination address. Additionally, trip information such as access/egress modes, number of transfers before and after survey route, origin type place, destination type place. These data can be filtered by date, route, as well as any of the travel data answer choices. Data are typically displayed in bar or pie charts.

2.8.2.2 Data Filters
Exhaustive data filters and aggregators will be exposed to the user via the dashboard interface, essentially allowing the user to filter, summarize, and slice through the data using any logical combination of variables, such as: Household size, Income, Age, Gender, Time of day, Number of available household autos, Number of workers in household, Household life cycle, Age of respondent, Trip purpose, Number of transfers, Transit access mode, Distance to transit, Fare payment method, Ethnic background, etc.

2.8.2.3 GIS Maps

ETC will set up live maps and map views such as Origin-Destination locations heat maps and joined record-level (tabular) view of the captured trip stops and segments together with the corresponding trip map view.

2.8.2.4 Dynamic Flow Charts

Volumes of trips between Origins and Destinations and/or between the ON-segments and OFF-segments will be visualized via dynamic flow charts, showing trip totals split by transit type and time of day segments. These charts will allow the user to quickly and visually gain an understanding of rider flows between sub-regions of the study area and between segments of a single route.

Task 8 Deliverables:
- Electronic and hard copies of database and final report
- Access to the Data Portal

2.9 One-Day Tour Travel Diary Survey, Processing, and Visualization

ETC Institute will conduct a One-Day Tour Travel Diary Survey with a subsample of at least 1,000 riders who complete the full Origin-Destination (OD) intercept survey. The purpose of the One-Day Tour Travel Diary Survey will be to better understand tours completed by transit riders, which will be used to enhance ARC's ability to do Activity-Based Modeling (ABM) in the future. The data will be processed in a manner consistent with the model and the results will be analyzed and visualized using ARC's ABMVIZ travel and activity visualization tool - http://atlregional.github.io/ABMVIZ. The specific services that will be provided as part of this task will include the following:

- **Design the Survey.** The ETC team will work with ARC to design the survey. The survey will be designed to capture the types of tours and types of activities that are completed by transit riders. The ETC team will work with ARC to ensure that the survey questions are detailed enough to support the region's ABM needs without placing too much burden on the respondents to complete the survey.

- **Develop the Sampling Plan**. The full OD intercept will include a question that asks respondents if they would be willing to participate in follow-up research. ETC estimates that up to 5% of the full OD intercept survey participants will agree to participate in the follow-up research. ETC Institute will develop the sampling plan for the One-Day Tour Travel Diary Survey from the subsample who agree to participate in the follow-up research. The sampling plan will be designed to ensure that a minimum of 1,000 surveys are completed and that the distribution of these survey is representative of the region's overall population of transit riders. As previously described, the ETC team used fare card data to recruit respondents to participate in trip and tour-based sample. This resulted in a highly biased sample due to low minority and low-income participation. By changing the sample frame to those who have already participated in the trip-based interview this bias will be mitigated, but special attention will be paid to Title VI riders' participation.

- **Conduct a Pilot Test.** The bulk of the surveys will be administered using an app that can be downloaded on the respondent's smart phone. To ensure the survey procedures and smart phone application work properly, ETC will conduct a pilot test with at least 50 participants before the survey begins. Based on the results of the pilot test, ETC will adjust the survey administration procedures and smart phone application as needed. ETC will also monitor the distribution of the tour diaries collected by Title VI riders and overall geography.

- **Conduct the Survey.** ETC will administer the survey to at least 1,000 transit riders as follows:

    o Riders who completed the full OD intercept survey and agreed to participate in the follow-up research will be contacted by phone, email, and/or mail to see if they are willing to complete the One-Day Tour Travel Diary.

    o Those who agree to complete the diary will be told that they will receive a $25 gift card upon successful completion of the survey. The $25 gift card will be emailed to the participant once ETC has verified their completion of the survey.

    o Most participants will participate in the survey using an application that can be downloaded on their smart phone. Riders who do not have smart phones will be given the option of providing their travel information by phone the day after their travel day was scheduled.

- **Conduct Analysis.** The ETC team will code tours from trips, analyze results by travel markets, such as household and person types, auto sufficiency, tour purpose, OD district, mode, and times-of-day. Data will be processed in a manner consistent with ARC's travel model. The data set will be a good supplement to the primary survey and will be useful for expansion, understanding traveler tour patterns, and other analytics.

- **Visualization of Results**. The ETC team will integrate select analytics from the one-day tour-based survey with ARC's ABMVIZ travel data visualization tool. The tool will supplement ETC's transit survey visualization dashboard by focusing on analytics other than transit trip records. The final set of visuals to be implemented will be agreed upon by the project team, but the initial set includes:

    o Transit tours by purpose and person-type from the one-day tour survey
    o O-D flows using chord ring diagrams from the one-day tour survey
    o Transit riders in time and space throughout the day from the one-day tour survey
    o Dynamic select transit line or stop rider analysis from the one-day tour survey

The ETC team will contribute the updated ABMVIZ functionality to ARC's existing GitHub project and publish the results using the existing GitHub pages platform. There is no additional cost to host this web application since it is already online.

2.10 Optional Tasks

ETC Institute understands that the following optional tasks will <u>not</u> be included at the time ARC gives ETC Institute a notice to proceed. However, ARC will reserve the right to add and/or modify the optional tasks at any time during the project.

2.10.1 Optional Task A: Additional Integration with ARC's ABMVIZ Visualization Tool (RSG Lead)

In addition to the visualization work completed under Task 9, the ETC team will integrate additional analytics from the full on-board survey with ARC's ABMVIZ travel data visualization tool. This optional task will draw inspiration from MBTA's efforts – http://ctps.org/apps/mbtasurvey2018 and http://mbtaviz.github.io – to interactively visualize transit ridership using the latest open source web-based data science and GIS toolkits. The ETC team will process the expanded survey data to implement an agreed upon number of additional interactive visualizations, implement the visuals, and publish the tool using ARC's existing GitHub presence.

The cost for this work will range based on the number of additional visualizations developed, with each visual priced at approximately $25,000. The ETC team suggests developing between four and ten additional visuals, for a task cost of between $100,000 and $250,000. The final set of visuals to be implemented will be agreed upon by the project team, but examples include:

- Schematic transit network-based (i.e. link-node) select transit line or stop rider analysis from the full on-board survey
- Station/stop analytics by time-of-day, line, person-type, etc. from the full on-board survey
- Commute analytics by location, time-of-day, etc. from the full on-board survey
- Interactive visualization of analytics by service type – MARTA rail, streetcar, bus
- Interactive visualization of Breeze card analytics by OD, time-of-day, etc.
- User defined spatial queries (i.e. select polygons) of custom analytics from the full on-board survey
- Rider market analytics by person-type, income, etc. from the full on-board survey

The ETC team will contribute the updated ABMVIZ functionality to ARC's existing GitHub project and publish the results using the existing GitHub pages platform. There is no additional cost to host this web application since it is already online.

2.10.2 Optional Task B: Compare Diary Outcomes to Third-party Big Data (WSP Lead)

The use of the smartphone application provides a window into the usefulness of so-called Big Data to inform or improve on-board transit surveys. The location traces made available by location-based services aggregators such as SafeGraph and Cubiq (the latter provides the data that partially informs StreetLight data) reveal the same information as the data collected by the smartphone application. The important difference is that the application:

- targets the population of interest (i.e., transit riders in Atlanta) through a known sampling frame;
- provides the opportunity to ask additional questions; and,
- provides a pathway to continuing the relationship with the customer.

As an optional task, WSP recommends comparing the inferences made using the smart phone app data to inform the ARC travel model to inferences made from data obtained from SafeGraph or Cubiq. Doing so can provide insight into the potential usefulness of "big data" to augment on-board surveys. It may be that the third-party data reveals estimates in line with those created by the application data, which suggests that the smart phone application is only needed in subsequent survey efforts if it is cost competitive with purchasing data. If the estimates derived from third-party data are very different than the smart phone application estimates, it suggests that big data sources may not be representative of transit riders in the greater Atlanta Region.

WSP estimates the following cost for this task, with ARC able to execute any subset at lower cost:

- Assist ARC in specifying data product and negotiating data purchase with SafeGraph or Cubiq: **$2,000**
- Small Data purchase (direct expense pass through to SafeGraph or Cubiq): **$25,000** (estimate, needs to be negotiated)
- Larger Data purchase (direct expense pass through to SafeGraph or Cubiq): **$50,000** (estimate, needs to be negotiated)
- Initial Data Exploration Scripts: **$15,000**
- One Set of Inferences (e.g., walk access distance distribution): **$5,000**
- Full Set of Inferences for Calibration: **$50,000**
- Set-up Pipeline for on-going Monitoring of Inferences: **$50,000**

3.0   COMMUNICATION/REVIEW MEETINGS/DOCUMENTATION

3.1   Internal Communication

Open and frequent communication between ARC and the Project team will be essential to the success of the project. Internal communication must also be well coordinated to minimize confusion and duplication of efforts.

Project Team Meetings

The Project team will meet as needed throughout the project to discuss the project status and upcoming events. Each Project team member is responsible for disseminating information during the meeting to others in their respective organizations.

Project Team Communication

Members of the Project team may contact one another freely; however, any substantive communication should be copied to the ETC and the ARC project managers. If the communication is not easily copied, an email should be sent to alert the project manager of the communication. It is critical that the ETC and the ARC project managers be copied on e-mail correspondence and made aware of telephone or personal communication, as necessary, to be kept informed of project activities and to initiate necessary follow-up actions.

3.2   External Communications

Most project communication with external audiences will happen through organized outreach, and thus will be captured through meeting documentation. However, members of the public, media, and public officials may contact the Project Team if they have questions, concerns, or need project information. It is critical that these contacts be managed to ensure that questions are properly answered including appropriate documentation. All external contacts will be coordinated with both the ARC and ETC Team project managers.

3.3   Project Documents

Filing Naming

Keeping track of multiple drafts of various documents can be difficult without the use of a consistent naming protocol. The format for naming a file document will be the file name followed by the date (in month, day, and year). Please attempt to keep the name as short as possible while ensuring the content of the file is easily identifiable. Draft or Final or other file name definition will be used as appropriate.

For example, a Work Plan submitted on March 25th, 2019 would be named: **ARC On-Board Survey Working Paper #1 - Work Plan 20190325.**

Revised documents will be named and saved as a new document with a new date using the above format. The earlier version of the document will be saved in a separate file folder.

3.4   Meetings/Reviews

Project Kick-off Meeting

A project Kick-off meeting was held on March 13th, 2019. The purpose of the Project Kick-off meeting was to review project goals and objections, confirm specific work elements, review the project schedule and define lines of communication.

Periodic Project Team Meetings

During the project, progress meetings will be held in person or via conference call as needed based on the discretion of the ARC's project manager. These meetings will discuss key project activities to ensure the project is progressing as desired within budget and on schedule. Identified issues will be addressed quickly and efficiently. Technical meetings will be held more frequently if necessary, particularly during the survey development phase of the project.

3.5   Document Retention

The ETC project manager has initiated a filing system to identify, collect, maintain, and safe-keep official active project files. The project manager is also responsible for identifying documents needed to be retained for the duration of the project and those documents requiring long-term storage.

The project manager is also responsible for maintaining quality control for the project and performing audits to validate that project documentation control is being performed in accordance with filing system requirements.

4.0   PROJECT SCHEDULE

Given the March 2019 start date, ETC Institute anticipates finishing the project during the spring of 2020 as shown in the tables below and on the following page. The first table shows the schedule for 2019. The table on the following page shows the schedule for 2020.

Project Schedule (Mar-Dec 2019)

Task #	Task Description	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	Develop a Work Plan and Schedule	█									
2.1	On-to-Off Counts Sample Plan	█									
2.2	Origin-Destination Survey Sample Plan	█	█								
2.3	Develop Data Analysis and Model Integration Plan	█	█								
3.1	On-to-Off Counts Instrument	█									
3.2	Origin-Destination Questionnaire	█	█								
4.1	On-to-Off Counts Training	█									
4.2	Origin-Destination (OD) Survey Training (Pilot and Full Surveys)	█					█				
5	Conduct OD Survey Pre-Test	█									
6.1	6.1 Support Public Involvement Efforts	█	█	█	█	█	█	█	█	█	█
6.2	6.2. Provide the Data Collection Team	█	█	█	█	█	█	█	█	█	█
6.3a	6.3. Conduct the Surveys (On-to-Off Counts)	█	█								
6.3b	6.3. Conduct the Surveys (OD Survey)		█	█	█	█	█	j	█	█	█
7.1	7.1 Data Processing		█	█	█	█	█	█	█	█	█
7.2	7.2 Preliminary Data Expansion										█
7.3	7.3 Secondary Data Expansion										
7.4	7.4 Analysis and Travel Model Network Assignments										
8.1	8.1. Submit Written Report.										
8.2	8.2. Prepare Data Portal										
9	Task 9: One-Day Tour Travel Diary Survey, Processing, and Visualization										
Option A	Additional Integration with ARC's ABMVIZ Visualization Tool (RSG Lead)										
Option B	Compare Diary Outcomes to Third-party Big Data (WSP Lead)										

Project Schedule (Jan-May 2020)

Task #	Task Description	Jan	Feb	Mar	Apr	May
1	Develop a Work Plan and Schedule					
2.1	On-to-Off Counts Sample Plan					
2.2	Origin-Destination Survey Sample Plan					
2.3	Develop Data Analysis and Model Integration Plan					
3.1	On-to-Off Counts Instrument					
3.2	Origin-Destination Questionnaire					
4.1	On-to-Off Counts Training					
4.2	Origin-Destination (OD) Survey Training (Pilot and Full Surveys)					
5	Conduct OD Survey Pre-Test					
6.1	6.1 Support Public Involvement Efforts					
6.2	6.2. Provide the Data Collection Team					
6.3a	6.3. Conduct the Surveys (On-to-Off Counts)					
6.3b	6.3. Conduct the Surveys (OD Survey)	▓				
7.1	7.1 Data Processing	▓	▓	▓		
7.2	7.2 Preliminary Data Expansion	▓	▓	▓		
7.3	7.3 Secondary Data Expansion	▓	▓	▓	▓	
7.4	7.4 Analysis and Travel Model Network Assignments			▓	▓	▓
8.1	8.1. Submit Written Report.			▓	▓	▓
8.2	8.2. Prepare Data Portal			▓	▓	▓
9	Task 9: One-Day Tour Travel Diary Survey, Processing, and Visualization			▓	▓	▓
Option A	Additional Integration with ARC's ABMVIZ Visualization Tool (RSG Lead)				TBA	
Option B	Compare Diary Outcomes to Third-party Big Data (WSP Lead)				TBA	