Technical Memo

To:	SHRP2 - C10 Tri-Agency Project Implementation Files
From:	Elizabeth Sall, UrbanLabs LLC
	Suzanne Childress, PSRC
	Lisa Zorn, MTC
Date:	July 6th, 2016
Subject:	Developing an Application-Ready Route Choice Model

The main goal of Task 4 of this tri-agency SHRP2 C10 Implementation Assistance Program project is to deliver a transit route choice model that is sensitive to basic quality of service variables such as in-vehicle time and wait time as well as variables of interest to recent and forthcoming policy analyses such as reliability and crowding. Ideally, the transit route choice model would also account for various transit market segments and be straightforward to understand and implement. This memo documents a pivot in the Task 4 team's approach to accomplishing these goals.

The original workplan finalized in March 2015 presumed that the team could estimate route choice model parameters from revealed route choice data in the San Francisco Bay Area including the GPS records from the 2013 California Household Travel Survey (CHTS) and the routes from various on-board surveys (OBS) that the Metropolitan Transportation Commission (MTC) has been conducting throughout the region. However, after exploring several potential avenues for estimation, the Task 4 team now believes that the best path forward is to bifurcate the task into two paths: (1) articulating the issues with estimation methods that still require significant basic and applied research, and (2) implementing techniques that we can readily employ in the meantime without expending too much budget and effort on something that will likely be superseded by emerging research. The suggested applications-ready track, which is detailed in the second part of this memo, is an asserted, then calibrated route choice logit model. A scope for proposed future research will be developed in the coming months based on available budget and input from academic collaborators.

Background: How we got here

Earlier this year, the Task 4 Team became aware of some theoretical, computational, and behavioral drawbacks of using the trip-based hyperpath formulation with a route choice logit model. Jeff Hood, the contractor tasked with estimating the model, brought these issues to our attention and the Team collectively recommended pursuing a time-bounded implementation of the promising recursive logit model (see "Route Choice Estimation Steps Forward", dated April 4, 2016) formulation and reverting to a hyperpath-based route-choice logit model only if recursive logit proved infeasible.

Since then, the combined advice of route choice and transit modeling experts such as Mark Hickman, Jeff Hood and Michael Florian as well as other modeling experts such as Peter Vovsha and Yi Chang Chiu has been that the recursive logit model as currently formulated would take a significant amount of work and research to make it appropriate for transit, and that the fall-back approach of the hyperpath-based route-choice logit model would not be worth expending a great deal of effort on because the parameters that it would produce would be significantly biased. The Task 4 Team agrees with these experts and now suggests that model parameters supported by estimation be a parallel research track that is undertaken concurrently with an asserted, then calibrated logit route choice model.

The rest of this memo is organized in two parts. In the first part, we outline issues that require further research in order for a true estimation to be useful in an applications context and why we believe a partial approach with the current methods we have available is not worth implementing at this time. These issues will be pursued in a separate parallel research track and not part of the critical path on this project. In the second part of this memo, we discuss how we will develop a set of parameters that will achieve the near term performance goals of this project.

Route choice issues that require further research

The following issues are hurdles to the implementation of a statistically valid and behaviorally sound route choice logit model using currently available techniques and methodology:

- variables that are non-additive or interdependent across links such as fares or reliability;
- overlap in transit alternatives in a behaviorally consistent manner so as to not inflate utilities of overlapping paths;
- bias in the estimated coefficients that can occur because of the common line problem;
- bias introduced because of differences in the choice set for the observed route choices and the estimation choice set;
- stochasticity of vehicle arrivals.

In addition, the solutions to these problems must be able to be solved in a manner that does not overwhelm computational pragmatism. This is a particular hurdle for any method that uses route-enumeration.

The Team feels like they have sufficiently explored these issues to conclude that they are both consequential to a successful implementation and have not yet been sufficiently resolved in either research or practice. The academics that were consulted as well as the attendees at the peer review did not have a suggested approach for a near-term solution and agreed that this was one of the things that “made transit hard.” After a budget from remaining funds is identified, the Team proposes to define a reasonable scope of work to furthering this research. Because of the depth and breadth of this work, it is unlikely that any of these research challenges will be "solved" within the scope of this project. Rather, we expect to hold a gathering of researchers in order to further flesh out and prioritize the research needs and/or develop workarounds that minimize a given problem.

Application-Ready Approach

The Task 4 Team proposes an applications-ready approach that leverages existing research and data sources and minimizes additional risk. The general approach is to assert route choice parameters derived from existing estimated mode choice models, literature, and engineering judgement and calibrate these parameters to meet performance and validation targets. The steps to support this approach are outlined in more detail below and diagrammed in Figure 1 at the end of this memo.

Develop Performance Targets

Because both the assert-calibrate parameter development method as well as the hyperpath methodology in Fast-Trips itself requires a great deal of subjectivity, the Team will prioritize developing a set of objective measurements for how well the model is performing on a practical, behavioral, and theoretical basis. The Team will develop both metrics and performance targets for the following categories:

- Path-based quality of service: how well does Fast-Trips match reality for quality of service factors (e.g., in-vehicle time, wait time, crowding) ?
- Path set coverage: does Fast-Trips cover all the observed and reasonable paths that should be considered, but not more than are behaviorally reasonable or computationally efficient?
- Path probabilities: does Fast-Trips assign reasonably large probabilities to observed and high-likelihood paths and relatively low probabilities to less attractive paths?
- Assignment: does Fast-Trips assign the right number of people to various routes, stations, and screenlines?
- Computational efficiency: can we run Fast-Trips route choice in a reasonable time in a practical computing environment?

Develop Validation Data

The team will organize and format data that is needed in order to ascertain the performance of the Fast-Trips model across each Performance Target category. This will include route data from the CHTS and OBS, passenger volumes from Automatic Passenger Counters (APCs) as well as information from atypical sources such as Google Maps.

Develop Calibration Processes and Tools

The Team will develop a calibration dashboard that compares Fast-Trips output, available observed data, and the performance targets. An initial version of this exists in Tableau, but PSRC staff envision creating it using open source tools such as Python.

Identify Initial Parameters

The starting point for asserting parameters will be the Transportation Authority's estimated mode choice models which have parameters for various components of the access and egress walk (e.g. elevation gain, indirectness, and density) as well as more traditional transit quality of service variables. For this step in the process, we will focus on getting a model running quickly, as opposed to adding every possible variable.

Finalize Validation Run Inputs and Specifications

While a base year network the San Francisco Bay Area has existed for months, it may be refined based on the variables that are being used in the route choice model specification and based on its performance when compared with the observed path-based quality of service that emerges when it is put into Fast-Trips. This step also includes making sure that various types of demand are ready, documented, and formatted including: the base year regional demand, demand from the OBS, and demand from the CHTS. In addition, this step involves formatting the initial run parameters and deciding on specifications to use within Fast-Trips for the initial validation run.

First-Tier Calibration: Observed Routes

The first tier of validation will use demand and chosen route information from the OBS and CHTS and will evaluate the performance categories that require knowledge of routes including:

- Path-based quality of service,
- Path coverage, and
- Path probabilities.

In order to reasonably satisfy our targeted performance of Fast-Trips across these categories, we will utilize calibration techniques such as:

- Correction to base year networks or demand;
- Tweaking of run parameters such as the delta for the hyperpath;
- Addition or modification of the specification of various factors or their parameters.

If the team determines that new variables are needed, they will reference existing research in order to identify a good starting point. It should also be noted that some modifications to the observed route choice data may be necessary if they are deemed to have routes that were chosen based on illogical behavior or non-representative/outlier transit performance/route availability.

All levels of calibration will complete several sensitivity tests in order to make sure that changes to route choices based on quality of service changes fall within reasonable and believable tolerances. Where available, we will use observed data to ground our beliefs about what is reasonable.

Second-Tier Calibration: Regional-Scale

The second tier of calibration will utilize the full regional demand in order to assess the two remaining performance targets: the assignment as a whole as well as the computational efficiency.

SF-CHAMP Implementation

This step connects all the plumbing that allows SF-CHAMP to use Fast-Trips as its transit route choice model rather than the current Cube/Voyager processes. At this point, this process translates from being under the purview of Task 4 to Task 8.

Third-Tier Calibration: Regional-Scale with Full Model

The final tier of calibration will assess the performance of the Fast-Trips parameters within a full SF-CHAMP model run.

Finalization

The team expects that there will be quite a bit of give and take between the performance on a regional scale that can be ascertained only by blunt means such as screenlines in some cases, and the small amount of data that we have on actual transit routes. This step will balance these potentially competing objectives in order to achieve parameters that are theoretically and behaviorally palatable as well as useful to planners. Similarly, parameters and processes that may result in a more precise validation and desirable calibration my expend too much computing resources and need to be scaled back.

Figure 1. Route Choice Parameter Development Steps Sans Estimation