Discussion-piece #10

Additional Measures for Quality Control of New Starts Forecasts

Federal Transit Administration

June 6, 2006

1. Motivation. Forecasts are large sets of complex information. Forecasting models are complicated tools with many potential sources of error. Benefits – the deltas between two forecasts – are challenging to estimate accurately. Experience since 2002 has shown that forecasts have many previously unknown and unexpected properties, some of which are errors caused by: models, networks, service plans, and other sources. A full set of QC tests is important to ensure the reliability of forecasts for proposed New Starts projects and maintain a level playing field for the evaluation of all projects.

2. Existing QC measures. The standard application of the current version of Summit yields district-to-district tables of deltas, thematic mapping of user-benefits row-sums and column-sums, and summary of changes in transit-access opportunities. The tables and thematic mapping have provided substantial insights into unexpected properties of models, networks, and service plans. The transit-access summaries of ridership changes and user benefits have illuminated a substantial source of error in the evaluation of proposed projects: inconsistencies in the geographic coverage provided by the baseline and build alternatives. Together, these QC measures have had significant impacts on the transparency, reliability, and usefulness of travel forecasts used to evaluate New Starts projects.

3. Some gaps. Other experience over the past several years has highlighted two significant gaps that remain in the measures available for quality control of forecasts: (1) understanding of the service-level changes that produce ridership gains/losses and user benefits, and (2) understanding of the extent to which the proposed transit guideway is the specific cause of those ridership changes and user benefits. *Ad hoc* analyses recently completed for several projects have demonstrated the potentially large distortion of insights into project benefits that can arise in the absence of clear understandings in both of these dimensions. An analysis of the home-based-work forecasts for one project, for example, showed that 80 percent of the benefits were caused by changes in first-waits and transfer-waits and that 70 percent of the benefits occurred on zone-to-zone transit paths that did not involve the proposed project. Further analysis showed that a significant increase in area-wide bus service frequencies – included in the build alternative but largely unrelated to the project itself – were the source of most ridership gains and user benefits in the travel forecasts. A more complete set of quality-control tests is therefore needed to close these gaps in the review of ridership forecasts for New Starts projects.

4. A general strategy. The necessary tests could be implemented in any travel forecasting software package but, to simplify the discussion, are described in the context of the semi-independent forecasts using an incremental logit mode-choice model with two sets of "best" transit paths each for the baseline and build alternatives (see Discussion-piece #9). The first test identifies the individual contributions of individual service characteristics to overall benefits and ridership gains. The second test identifies the fraction of these benefits that are associated with transit paths that include the proposed project.

5. Test 1: sources of benefits. The incremental logit model computes the change in transit-access mode shares based on the relative changes in the utility of transit with walk-access and transit with drive-access:

$$(1) \quad P_{tw,build} = \frac{P_{tw,base} \times \exp(\Delta U_{tw})}{[\, P_{tw,base} \times \exp(\Delta U_{tw}) \,] + [\, P_{td,base} \times \exp(\Delta U_{td}) \,]}$$

where $P_{tw,build}$ is the probability of transit/walk in the build alternative;
$P_{tw,base}$ is the probability of transit/walk in the base alternative;
$P_{td,base}$ is the probability of transit/drive in the base alternative; and
ΔU_{tw} is the change in transit/walk utility, build minus base.

The change in the exponentiated utility of either transit mode is a function of changes in its individual service variables between the baseline and build alternatives:

$$(2) \quad \exp(\Delta U_{tw}) = \exp[\, b_{ivt}(\Delta ivt) + b_{ovt}(\Delta ovt) + b_{fare}(\Delta fare) + \ldots + b_n(\Delta var_n) \,]$$

or, equivalently:

$$(3) \quad \exp(\Delta U_{tw}) = \exp[b_{ivt}(\Delta ivt)] \times \exp[b_{ovt}(\Delta var_{ovt})] \times \exp[b_{fare}(\Delta fare)] \times \ldots \times \exp[b_n(\Delta var_n)]$$

where b_{ivt}, b_{ovt}, b_{fare}, and b_n are the coefficients on in-vehicle time, out-of-vehicle time, fare, and other variables in the utility expression. Equation (3) provides the basis for isolating the contributions of individual service characteristics. The overall exponentiated delta-utility can be computed as the product of the individual exponentiated delta-utility components. That contribution of any individual component can be determined by repeating the calculation but omitting that component. Consequently, the test on sources of benefits can be implemented through an efficient set of computations of ridership changes and user benefits where each computation in the set omits one of the transit delta-utility components. The ratio between each "partial" result and the full result will provide a direct indicator of the role of each service variable in generating new riders and user benefits. This test is expected to provide meaningful results regardless of the nature of the transit pathbuilding employed in the sponsor's model set. Changes in probability-weighted attributes of multiple paths should provide results that are as clear as those based on attributes from all-or-nothing single paths.

6. Test 2: association with the project. The second test may be straightforward or more complicated, depending on the nature of the transit path-builder employed in the sponsor's model set. With a single, best-path, all-or-nothing pathbuilder, the impedance values available for each transit path will include the in-vehicle time on each transit mode that can be tested for each zone-to-zone interchange to determine whether or not the project mode is involved in the path. For a starter line, a test of "zero, or not-zero" will be sufficient to identify paths using the proposed project. For a project that extends an existing line or adds a new line to an existing guideway system, one of several conditions will apply for each zone-to-zone path:

- Guideway time in the build alternative is zero, so the path does not use any guideway;

- Guideway time in the build alternative is equal to guideway time in the baseline alternative, so the guideway path is (presumably) the same in both alternatives; and

- Guideway time in the build alternative is less than guideway time in the baseline alternative (an unexpected result that might suggest some further checking), so the build path probably does not involve the proposed project;

- Guideway time in the build alternative is greater than the <u>zero</u> guideway time in the baseline alternative, so the build path probably does involve the proposed project and trips using that path would be new to the guideway system; and

- Guideway time in the build alternative is greater than the <u>non-zero</u> guideway time in the baseline alternative, so the build path probably does involve the proposed project and trips using that path would "existing" guideway riders using the guideway for a longer part of their trips.

The benefits and new transit ridership in zone-to-zone interchanges in the last two categories would be classified as project related, and comparison with total benefits and new riders would suggest the fractions of those impacts that are generated by the project itself. The differentiation between the last two categories – "existing" and "new" guideway riders – is useful because it potentially sheds light on the impacts of guideway-specific constants in the mode choice model and/or transit pathbuilding: existing guideway riders are likely to benefit from the guideway effect in both the baseline and build alternatives, while new guideway riders are likely to benefit from the effect only in the build alternative.

The probability-weighted impedances provided by multi-path pathbuilders pose challenges for this second QC test. The "zero, or not-zero" test of guideway in-vehicle times is much less useful because those times can be not-zero for set of paths that have a small – but non-zero – probability of guideway use. Consequently, the in-vehicle time test is much less able to distinguish actual guideway riders – and changes in guideway riders. Application of the second QC test with multi-path pathbuilders will require additional work by FTA and project sponsors to identify ways to provide the necessary insights.

7. <u>Implementation.</u> Summit version 1.5 will include the means to apply both of the new QC tests as part of the preparation of the semi-independent QC forecasts and in parallel with (or independent of) the computation of user benefits from the sponsors forecasts. FTA anticipates that testing of the new QC measures will be completed by late 2006 and that these QC measures will be a required element of requests for entry into preliminary engineering beginning in January 2007. The requirement may also apply to requests for advancement into final design for projects currently in preliminary engineering.