

## HSIPR Best Practices: Ridership and Revenue Forecasting

April 2011

**Prepared for:**  
Office of Inspector General  
US Department of Transportation

**Prepared by:**  
Steer Davies Gleave  
883 Boylston Street, 3rd Floor  
Boston MA 02116  
(617) 391-2300 - [www.steerdaviesgleave.com](http://www.steerdaviesgleave.com)  
Under subcontract to Charles River Associates



## ACKNOWLEDGEMENTS

This report was prepared for the USDOT Office of Inspector General by Steer Davies Gleave Inc., working under a subcontract to Charles River Associates Inc. for order number DTOS59-10-F-10085.

The Steer Davies Gleave personnel who participated in this work include Jon Bottom, Leo Eyles, Masroor Hasan, Felicity Hulme, Lucile Kellis, Scott Prentice, Lars Rognlien and Tessa Wordsworth. Mark Kiefer, an independent consultant, also contributed to the writing. Masroor Hasan was the project manager and Jon Bottom was the project director.

Martin Baynam-Knight of Steer Davies Gleave and Dan Brand, a senior advisor at Charles River Associates, reviewed the report and contributed useful suggestions.



# CONTENTS

<b>1</b>	<b>OVERVIEW OF RIDERSHIP AND REVENUE FORECASTING</b>	<b>1</b>
	INTRODUCTION	1
	HIGH-LEVEL DESCRIPTION OF MODELING STEPS	1
	<i>Definition of study time frame</i>	2
	<i>Definition of study area</i>	2
	<i>Definition of level of geographic detail</i>	2
	<i>Trip table preparation</i>	3
	Trip segmentation	3
	Base year trip table development	3
	Forecast year trip table development	4
	<i>Base and forecast year modal network preparation</i>	4
	Choice of network representation	4
	Preparation of skim tables	5
	<i>Mode choice modeling</i>	5
	Choice modeling	5
	Diversion choice modeling	6
	<i>Induced travel modeling</i>	6
	<i>Assignment</i>	6
	<i>Feedback</i>	6
	<i>Revenue calculations</i>	7
	<i>Sensitivity analysis</i>	7
	COMPUTATIONAL SUPPORT	8
	OUTLINE OF THE REMAINDER OF THE REPORT	8
<b>2</b>	<b>TRIP TABLE PREPARATION</b>	<b>9</b>
	TRIP TABLE SEGMENTATION	9
	<i>Geographic detail</i>	9
	<i>Universe of trips</i>	9
	<i>Time period</i>	10
	<i>Market segmentation</i>	10
	DATA ISSUES	10
	APPROACHES TO TRIP TABLE PREPARATION	10
	<i>Trip table development from base data</i>	11
	<i>Trip table synthesis from trip generation/distribution models</i>	12
	<i>Discussion</i>	13
	POTENTIAL IMPACTS OF TRIP TABLES ON HSIPR FORECASTS	14
	METHODS FOR COLLECTING MODAL TRIP TABLE DATA	14
	<i>Auto</i>	15
	Household travel surveys	15
	Intercept surveys	16
	Trip table estimation from traffic counts	17
	License plate surveys	17
	Cell phone based data	18

Other sources.....	18
<i>Rail</i> .....	18
Amtrak data .....	18
Commuter rail data .....	19
<i>Air</i> .....	19
Data reporting issues .....	20
International air traffic data .....	21
<i>Air access</i> .....	22
<i>Bus</i> .....	22
<i>Other</i> .....	23
Tourism surveys and studies.....	23
Business surveys and studies.....	23
SEGMENTING TRIP TABLES BY PURPOSE .....	23
FORECASTING FUTURE YEAR TRIP TABLES FROM BASE YEAR TABLES.....	24
<i>Direct demand models</i> .....	24
<i>Growth rates from pre-existing forecasts</i> .....	25
<i>Assumed growth rates</i> .....	25
<b>3 MODAL LEVEL OF SERVICE DATA.....</b>	<b>27</b>
AIR.....	27
<i>Line haul travel times</i> .....	27
<i>Access/egress times</i> .....	28
<i>Air fares</i> .....	28
<i>Access/egress costs</i> .....	29
<i>Frequency</i> .....	30
<i>Reliability</i> .....	31
RAIL.....	32
AUTO.....	33
BUS.....	34
FUTURE LOS CHARACTERISTICS .....	34
POTENTIAL IMPACTS OF LOS CHARACTERISTICS ON HSIPR FORECASTS.....	35
<b>4 MODE CHOICE MODELING.....</b>	<b>37</b>
MODE CHOICE MODELS.....	37
<i>Choice modeling</i> .....	37
<i>Diversion choice modeling</i> .....	39
<i>Other approaches</i> .....	40
METHODS FOR OBTAINING MODE CHOICE DATA .....	40
<i>Revealed Preference (RP) data</i> .....	40
<i>Stated Preference (SP) data</i> .....	41
<i>RP and SP surveys</i> .....	41
Sample design.....	41
Questionnaire design .....	43
Estimation issues .....	44
<i>Combined RP/SP data</i> .....	44

MODEL VARIABLES AND SPECIFICATIONS.....	44
<i>Tripmaker-specific variables</i> .....	44
<i>Trip (LOS)-specific variables</i> .....	45
<i>Mode-specific constants</i> .....	45
SPECIFIC MODELING ISSUES.....	45
<i>Generalized cost</i> .....	45
<i>Reliability</i> .....	46
<i>HSR service classes</i> .....	46
<i>Service frequency</i> .....	46
<i>Network effects</i> .....	47
MODE CHOICE MODEL RESULTS AND STATISTICS.....	47
POTENTIAL IMPACTS OF MODE CHOICE MODELING ON HSIPR FORECASTS .....	48
<b>5 OTHER MODELING STEPS IN RIDERSHIP AND REVENUE FORECASTING.....</b>	<b>49</b>
INDUCED TRAVEL .....	49
<i>Typical range of induced demand magnitudes</i> .....	49
<i>Modeling approaches</i> .....	50
<i>Issues</i> .....	51
TOTAL HSIPR RIDERSHIP.....	51
ASSIGNMENT AND FEEDBACK .....	51
REVENUE CALCULATION .....	52
SENSITIVITY ANALYSIS .....	52
<b>6 SPECIAL TOPICS.....</b>	<b>55</b>
INTERFACE WITH URBAN (MPO) MODELS.....	55
<i>Issues</i> .....	55
AIR/RAIL COMPLEMENTARITY .....	55
<i>Possible competitive responses of the airlines to the introduction of HSI PR service</i> .....	55
<i>Use of HSR as an airport access mode</i> .....	56
<i>Connections between intercity high speed rail and air services</i> .....	57
Actual experience in Europe.....	57
<i>Assessment of HSR proposals serving airports</i> .....	59
<i>Combined choice of airport and access mode in multi-airport regions</i> .....	59
TOURISM TRAVEL.....	60
<b>7 BEST PRACTICE FORECASTING FOR DIFFERENT STUDY STAGES .....</b>	<b>61</b>
PRELIMINARY STAGE.....	62
<i>Best practice</i> .....	62
<i>Common errors and pitfalls</i> .....	64
INTERMEDIATE STAGE .....	64
<i>Best practice</i> .....	65
<i>Common errors and pitfalls</i> .....	67
FINAL STAGE .....	67
<i>Best practice</i> .....	68
<i>Common errors and pitfalls</i> .....	71

## FIGURES

Figure 1-1. HSIPR ridership and revenue forecasting flowchart.....	1
Figure 3-1. Example of auto travel times calculated from GIS-based highway network.....	34
Figure 4-1. Multinomial choice model example .....	38
Figure 4-2. Nested choice model example.....	39
Figure 4-3. Diversion choice model example .....	40

## TABLES

Table 2-1. An example trip production table .....	12
Table 2-2. An example trip attraction table .....	12
Table 2-3. An example zone to zone trip table .....	12
Table 2-4. Advantages and disadvantages of different data sources for air trip tables .....	22
Table 2-5. Advantages and disadvantages of travel growth rate development methods.....	26
Table 3-1. Advantages and disadvantages of data sources for air level of service .....	32
Table 5-1. Induced demand percentages in US HSIPR studies .....	50
Table 5-2. Induced demand percentages in international HSR systems and studies .....	50
Table 6-1. Summary of important factors in assessing HSR-air connections .....	58
Table 6-2. Summary of existing HSR-air connections in Europe.....	58
Table 6-3. Specification of HSR-air connection by study level.....	59
Table 7-1. Preliminary stage ridership modeling approaches .....	63
Table 7-2. Intermediate stage ridership modeling approaches .....	66
Table 7-3. Final stage ridership modeling approaches.....	70
Table 8-1. Study time frame checklist .....	73
Table 8-2. Study area network checklist .....	74
Table 8-3. Definition of travel markets checklist .....	74
Table 8-4. Definition of market segments checklist .....	74
Table 8-5. Trip tables checklist.....	75
Table 8-6. Demographic and socio-economic characteristics checklist.....	75
Table 8-7. HSIPR alternatives definition checklist .....	76
Table 8-8. LOS characteristics of the competing modes checklist .....	77
Table 8-9. LOS characteristics of the access modes checklist.....	78
Table 8-10. Mode choice models and underlying data checklist .....	79
Table 8-11. Induced demand models checklist.....	80
Table 8-12. Revenue calculation checklist .....	80

# 1 Overview of ridership and revenue forecasting

## Introduction

This report provides a high-level description of the steps typically involved in forecasting the ridership and revenue of a high-speed/intercity passenger rail (HSIPR) service. While particular studies may, for a variety of reasons, use approaches and methods that differ somewhat from those described here, the description here is believed to be an accurate representation of current standard practice in HSIPR forecasting.

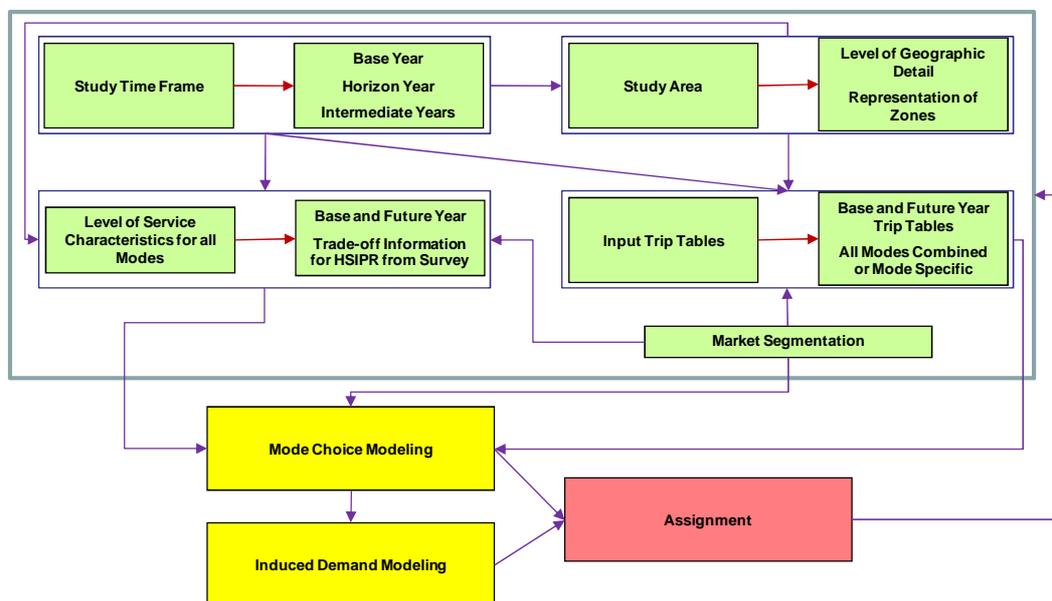
The report is intended for non-specialists who may be called upon to review HSIPR ridership forecasting studies prepared by others. It provides information on the range of data and methods used in HSIPR forecasting at different stages of study, and flags particular areas or subjects that will generally require in-depth examination by subject area experts. The intent is to provide information and guidance that will assist generalist reviewers to understand and evaluate forecasting studies. Similar reports have been prepared in the areas of HSIPR public benefits assessment and operating cost estimation.

It should be noted that, in a HSIPR study, forecasts will typically be prepared for both a “build” situation that includes the proposed project, as well as a “no-build” situation without the project. The specific definition of the no-build situation needs to be agreed with study reviewers. Project benefits and costs are generally defined in terms of differences between the two situations.

## High-level description of modeling steps

The flowchart in Figure 1-1 shows the high-level forecasting steps and the relationships between them.

FIGURE 1-1. HSIPR RIDERSHIP AND REVENUE FORECASTING FLOWCHART



### ***Definition of study time frame***

Defining the time frame of a forecasting study establishes the base data year, the first year of HSR operation, the final (horizon) year and perhaps one or a few intermediate years that will be explicitly modeled or analyzed in detail. The latter are particularly needed if the proposed system is to be constructed and opened in stages. Forecasts for intermediate years with no new added services are typically derived by interpolation (between the base and horizon years) and extrapolation (beyond the horizon year) of the results obtained for the explicitly-modeled years.

In almost all cases, the specific choice of the base data year and forecasting years is strongly influenced by available data and forecasts of required input or other needed data; many metropolitan planning organizations (MPOs), for example, prepare socio-economic and other forecasts at five-year intervals to a twenty to thirty-year horizon, so the years used in these forecasts would be natural candidates for use in HSIPR forecasts as well.

The first forecast year would typically be as close as possible to the expected opening year of the project being studied. The study horizon for a major transportation project would typically be 20 to 25 years after the first forecast year, although specific features of a project may modify this. For example, if bonding of future revenues is anticipated, a 30 to 35 year horizon forecast year after opening may be desired. Again, there may be a mismatch between years for which detailed modeling is possible because of data availability, and the project opening year or study end date; in such cases interpolation and extrapolation must be used.

### ***Definition of study area***

The study area is the geographic extent from which the rail service is likely to draw riders. This will be roughly defined by the rail station locations; the population and activity centers that the stations serve; and connection options to other transportation modes. It follows that the study area definition will be based on the characteristics of the rail line, access/egress modes, connection possibilities to other modes (e.g. airport or other rail lines) and the geographic distribution of socio-economic activities around the stations.

There is no definitive rule establishing the extent of the catchment area around a station: this will depend in part on the nature of the rail service, the location of the station relative to other stations on the same line and on competing lines, and the ease of access/egress. In general it is preferable to define a larger area so as not to limit artificially the extent of the market to be analyzed; if some portions of the study area are unlikely to generate rail demand, this will be found in subsequent forecasting steps.

As a practical matter, the limits of the study area will typically be chosen to coincide with already-defined and relevant geographic boundaries such as counties, Census geographic units or traffic analysis zones used in other modeling efforts that cover all or portions of the study area.

Where the rail service connects directly or indirectly to another longer-distance transportation mode (particularly air via a rail station at an airport or via a transfer connection), the locations served by the other mode (e.g. other airports and their service areas) are typically not considered part of the rail study area; rather, as will be seen below, the demand generated by this “extended” service is forecast separately.

### ***Definition of level of geographic detail***

The study area is subdivided into a number of non-overlapping geographic units, called traffic analysis zones or TAZs, that are considered to be the origins and destinations of trips. Put differently, travel demand is defined as being from one TAZ to another.

Travel that takes place entirely within the boundaries of a single TAZ is generally not represented explicitly in a forecasting model. Similarly, multiple transportation facilities within a TAZ tend to be represented approximately, if at all; for example, if a TAZ includes two rail stations, most models would not analyze demand at the two stations to the same level of detail as they would if the stations were in different TAZs.

Basic travel demand forecasting data is required at the TAZ level: both socio-economic data that describe the demand-generating activities within zones, and data on TAZ-to-TAZ flows. While it may seem preferable to define and use small zones because of the geographic precision that this would allow, the limited availability of data at a detailed geographic level, and the difficulty of accurately forecasting such data to future years, tend to favor larger zone sizes.

TAZs are usually defined in terms of pre-existing geographic units, typically ranging from entire counties down to the detailed traffic zones used in other modeling efforts in the study area. The appropriate zone size depends on the intended level of detail and accuracy of the forecasting effort, with larger sizes being more suitable for preliminary studies and conversely.

### *Trip table preparation*

In order to calculate the number of trips that divert to the new or improved HSIPR mode, the level of trip making without the new or improved HSIPR mode must first be established. This is the total amount of travel demand, in all the existing modes, that exists in the study area in the no-build situation. The changes to this level of travel demand produced by the HSIPR project will then be calculated through application of the mode choice model (described later).

As noted above, for forecasting purposes travel demand is defined as being from one TAZ to another. A rectangular trip table (or matrix) is frequently used to summarize travel demand, where the rows represent the different origins, the columns represent the different destinations, and the individual cells show the number of trips from the corresponding origin to the corresponding destination. The trip tables will vary depending on the study time frame, the study area and the level of detail in which the study area is represented. Preparation of trip tables is a central task of travel demand forecasting. Defining and characterizing the types of trips to be studied is a necessary first step in this process.

### *Trip segmentation*

Travel demand forecasting usually considers separately trips of different types, with different characteristics and behavioral responses to new HSIPR or other transportation system changes. Total demand is then the aggregation of the forecasting results for the various individual trip types.

The definition of the different trip types - the trip segmentation - depends closely on the specifics of the situation being analyzed. Trips are normally distinguished by mode of travel and trip purpose, because for example business travelers on rail can be expected to react differently than tourists in cars to new travel options. A distinction by trip length (with a threshold at 75-100 miles for example) is also common. Additional segmentation may be based on traveler characteristics such as income. The objective is to identify travel groups that are meaningfully different in the way they respond to travel options. As before, however, excessive segmentation (a large number of segments, many of which represent relatively small numbers of trips) leads to problems in collecting and forecasting the corresponding trip tables.

### *Base year trip table development*

#### Trip table development from direct data

Base year trip tables are generally prepared by using a variety of sources of data on actual trip making patterns and levels to derive the required zone-to-zone trip volumes. Travel surveys can be designed to sample and estimate the volume of all zone-to-zone movements in a study area with the appropriate trip segmentation; for example, questions formerly in the Census long form questionnaire and now in the American Community Survey attempt to develop this kind of data for journey-to-work trips. More generally, however, it is rare to have available a single data source that provides complete information at the required level of geographic detail for all trip segments; some processing and conflation of data from different sources is typically required. Sources and methods for developing trip tables are discussed further below.

### Trip table synthesis

A second approach to base year trip table preparation applies a sequence of modeling steps that ultimately result in a trip table. These steps begin by estimating how many total trips depart from and travel to each individual TAZ, a process called trip generation. The total trips departing from a zone (its trip production) can be viewed as the sum of the elements of the corresponding row of the trip table; similarly, the trip trips traveling to a zone (its trip attraction) can be viewed as the corresponding column sum. Trip generation typically takes account of a zone's socio-economic and accessibility characteristics when calculating the number of its departing and arriving trips.

Trip distribution computes the volume of zone-to-zone trips, given the trip productions and attractions computed during the trip generation step. Put differently, trip distribution fills in the elements of the trip table in a way that the given row and column sums are respected. Some methods of trip distribution take account of the relative time or cost of travel between zones, while others start from an approximate table (perhaps one based on historical trip patterns) and factor it to obtain the required row and column sums. Furthermore, some methods adjust the productions and/or attractions in order, for example, to ensure a balanced (symmetric) trip table, or to accord greater confidence to the predictions of productions compared to the attractions.

Trip table synthesis methods are very common in metropolitan and regional travel demand forecasting, and there are many variations both on the basic steps and in the ways the steps are combined and integrated with other components of the forecasting process. However, these models sometimes perform less well for intercity travel than they typically do for urban travel.

#### *Forecast year trip table development*

Future year trip tables cannot, of course, be directly developed from observational data. A common approach to preparing forecast year tables is to estimate the growth in trip productions and attractions between the base and future years (using socio-economic forecasts, for example), and then use a method that applies these growth factors to the base year trip table. In contrast, one of the advantages of trip table synthesis is that after trip generation and distribution methods are developed using base year data, they can usually be applied without change to forecast year trip tables if the required inputs for the future year (e.g. zonal socio-economic data) are available.

#### ***Base and forecast year modal network preparation***

Travel demand forecasts require information on the characteristics of the various travel modes that compete to serve travel demand between origins and destinations. This is used to provide a quantitative determination of the service characteristics offered by the various available modes, and to represent the way in which trips will follow specific modal routes, links, stations and other facilities they go from origin to destination. These modal service characteristics and representation of the network, in turn, also depend on the study time frame and the study geography.

#### *Choice of network representation*

Urban transportation planning models typically employ detailed representations of street and transit networks that include the location, alignment, connections and service characteristics of the transportation facilities in the study area. Specialized transportation modeling software packages facilitate the input, editing and checking of this representation.

Intercity passenger forecasting models use a similar network representation approach. The complexity of intercity networks is typically less than that of urban networks, so the advantages of this approach, in terms of having specialized software perform many of the required computations, may be less compelling.

Although the initial input of network data can be very time consuming, this burden is reduced if data from already-developed network models (from states or local MPOs for example) can be imported and used.

Alternatively, the network representation may be more *ad hoc*, using for example spreadsheets to organize and process information on network facilities. Modeling efforts that are not concerned with routing issues (perhaps because the structure of the network is very simple) might choose to not develop an explicit representation of the network at all, and instead focus directly on zone-to-zone level of service data, as described below.

#### *Preparation of skim tables*

Travel demand forecasting requires information on the time, cost and other service characteristics of the various modes that are available for a trip from origin to destination. The process of determining the end-to-end service characteristics for all modes and origin-destination pairs is known as skimming, and the results of this process are conveniently organized in skim tables, similar to trip tables but containing the values of service variables rather than trips.

Transportation planning software packages have functions that create skim tables from a network representation. Studies that do not use an explicit network representation may prepare skim tables from available sources of data on modal service characteristics. Some studies can re-use skim tables that are available for the study area from prior modeling efforts, without referring to the network representation that was used to develop these skim tables.

#### ***Mode choice modeling***

Mode choice modeling refers to a process that predicts the outcome of the decision process that travelers apply to choose the mode(s) that they will take to go from origin to destination. Different models are typically developed for the different trip segments identified earlier (and in fact the results of the model development may influence the definition of these segments). The trip tables and modal service characteristics discussed above are the inputs to the mode choice modeling process; in forecasting these would relate to the different study years and alternatives under consideration.

The mode choice model itself is typically developed via a statistical analysis of the behavior of travelers in different situations that allows the preferences and tradeoffs of travelers to be understood and modeled. The statistical analysis can be based on traveler behavior observed in actual travel situations (revealed preference data), or on behavior observed in hypothetical situations presented to travelers in a survey (stated preference data), or both. The models take into account the service characteristics of the modes being considered, as well as characteristics of the traveler and the trip.

Forecasts of intercity passenger rail demand typically apply either of two approaches for mode choice modeling, described in the following sections.

#### *Choice modeling*

Choice modeling represents the decision outcomes of travelers considering choices from among the full set of available mode alternatives; the mode used by travelers in the base situation is not considered. The output of a choice model is a set of mode share fractions (probabilities), one associated to each of the modes available to a particular trip segment for travel from a particular origin to a particular destination. By applying these fractions to the total number of travelers who are in that segment and travel between the origin-destination pair, the corresponding modal volumes are obtained. This calculation is repeated for each trip segment and each origin-destination pair to obtain trip tables for each segment and mode. The modal trip tables can also be assigned to a network to obtain the volumes that use different links, stations and other network facilities.

### *Diversion choice modeling*

Diversion choice modeling represents the decision outcomes of travelers who use a particular mode in the no-build situation and are considering an alternative (new or improved) mode in the build situation - for example, current automobile travelers considering HSR. An individual diversion choice model considers only two modes: the one in use in the base situation, and the alternative being considered for diversion. Distinct models are typically developed for different base modes as well as for different trip segments (e.g. purposes). Each such distinct model is applied separately for each origin-destination pair. In each case, the output of the model is the fraction of travelers who will divert from the base mode to the alternative mode. By applying this fraction to the number of travelers using the corresponding base mode, the volume of diversions to the alternative mode is obtained.

### *Induced travel modeling*

Induced travel refers to trips that only occur as a result of a transportation improvement, and that were not made prior to the improvement. Conceptually, the improved travel conditions result in movement along the demand curve, and the extent of this movement defines the volume of induced travel. Improved travel conditions can also bring about socio-economic development that shifts the demand curve outward, and this will also typically result in additional trips. This situation will be treated separately in this document, although in practice the distinction between the two is sometimes difficult to make.

Demand forecasting typically predicts induced travel using a model developed specifically for that purpose. Elasticity-based methods are a common approach, relating a percentage change in demand to a corresponding change in generalized cost, accessibility or other measure of travel conditions. The induced demand model is directly tied to the mode choice process as the generalized cost or accessibility measures are usually calculated from the mode choice models. Induced demand can also be calculated as a percentage of the HSIPR demand.

The combined results of the mode choice and induced travel modeling constitute the total origin-destination HSIPR demand forecasts. These forecasts may be final, or may be modified through a feedback process.

### *Assignment*

Assignment refers to the process by which normal and induced origin-destination demand is associated with a specific path or paths over the network. In rail studies, assignment is frequently a matter of tracking boarding to alighting station demand and doing the simple bookkeeping needed to compute the total ridership on track sections between adjacent stations, and the total boarding and alighting passenger volumes at stations. This can be done with a spreadsheet or by specialized software.

Where multiple comparable paths are available to travel from an origin to a destination, assignment may also split the demand among these paths in a way that attempts to represent traveler choices. Transportation planning software packages usually have capabilities to do this in transit networks. In a rail context, however, services in such situations have significantly different characteristics (a local and an express service, say, with very different travel times and fares), and predicting travelers' choices between them is a task that is usually handled by the mode choice rather than the assignment model.

Where congestion is important (e.g. on highways), assignment also takes account of the interplay by which travelers' path choices affect congestion levels and conversely.

### *Feedback*

Feedback is an attempt to equilibrate a multi-step travel forecasting model through an iterative process in which the outputs of the modeling process are "fed back" as inputs to earlier stages, either directly or with modification, and the process is repeated until changes in travel volumes or conditions between iterations are deemed insignificant.

As an example, the ridership attracted to a high-speed rail line may depend on the level of highway congestion. To develop forecasts, the model assumes a particular level of congestion, but the number of road users who choose rail may be sufficiently large to lower highway congestion below the initially assumed value. Feedback would then take the level of congestion predicted in one iteration after accounting for diversion to rail, and use it as input to the mode choice model of the next iteration. The result will presumably be lower rail ridership and higher highway congestion than in the prior iteration. The process continues until results of successive iterations are deemed sufficiently close.

Unless a rail project is anticipated to have significant within- or cross-modal congestion impacts, feedback is usually not incorporated in rail forecasting processes.

### ***Revenue calculations***

The diverted HSIPR trips calculated from the mode choice models and the induced HSIPR trips are added to produce the total HSIPR ridership for any OD pair or for the whole system. The HSIPR fare revenue produced by a particular alternative is calculated for each OD pair from this ridership forecast and the corresponding HSIPR fare. Detailed forecasting studies may consider multiple fare classes (e.g. premium vs. regular service) for which the calculation should be repeated. The sum of the fares generated by each OD pair and fare class (if applicable) constitutes the total HSIPR fare revenue in the modeling analysis period (e.g. an average day), which is generally converted to an equivalent annual value through application of a suitable annualization factor.

Intermediate or final stage studies may attempt to determine the level and structure of fares that maximize HSIPR fare revenues. This can be done, at least approximately, by systematically varying fares input to the ridership forecasting model and examining the corresponding revenue levels. Higher fares produce more revenue per passenger but fewer total passengers than lower fares, and conversely. Accordingly, the curve of total fare revenue vs. fare level (appropriately defined) typically has an inverted “U” shape, from which the maximum revenue and corresponding fare level can be determined. It is not uncommon for this curve to be relatively flat near its maximum, so that modest variations in fare levels around the maximizing fare do not have much impact on overall fare revenues. Hence, it is often suggested to set fare levels a little lower than the revenue maximizing levels. This will produce more ridership and hence more public benefits than the revenue maximizing fare levels without significantly reducing total revenue.

### ***Sensitivity analysis***

All ridership and revenue forecasting studies should incorporate an analysis of the sensitivity of forecast results to key inputs and modeling assumptions. Sensitivity analysis typically involves using the demand model to perform forecasts with modified inputs and parameters, noting the outputs corresponding to each model run. The modified values may be developed mechanically (e.g. central case value +/- 10%) or through more elaborate procedures (e.g. Delphi methods that solicit expert opinions). Probabilities of occurrence are associated in some way with each modified value in order to characterize the distribution of the output values. The effort devoted to this analysis generally increases with the study stage. Final stage studies often spend a significant effort investigating forecast sensitivity.

Sensitivity analyses serve a number of useful purposes. First and perhaps foremost, they indicate the reliability associated with the model output forecasts. Results of this analysis will identify the factors that most directly affect project ridership and revenue, and project proponents will want to focus their attention on the identified factors over which they have some control. Information from the sensitivity analysis regarding the distribution of potential project financial outcomes can also be of considerable use in preparing a suitable financial plan.

## Computational support

Travel demand forecasting can entail a significant amount of computation, and only very simple situations and rough forecasts can realistically be expected to be handled manually. In practice, some form of computational support for the forecasting calculations is almost always used. Commonly-applied computational tools include:

- Spreadsheets;
- Commercial off-the-shelf transportation network modeling software; and
- Custom software applications.

The choice between these options will depend on the amount of data to be manipulated, which is related in turn to the size of the study area and the level of geographic detail.

The data processing requirements of a large and complex forecasting study are significant. Managing the many input, output and intermediate files produced by a forecasting process requires careful organization. Commercial software packages typically automate the computational and file handling tasks to some extent.

## Outline of the remainder of the report

The remainder of this document consists of:

- More detailed descriptions of the various steps of HSIPR ridership and revenue forecasting process in Chapters 2, 3, 4, 5 and 6;
- HSIPR ridership and revenue forecasting best practices for different study stages in Chapter 7; and
- Ridership and revenue forecasting checklists for use by reviewers of HSIPR studies in Chapter 8.

The discussions of ridership and revenue forecasting steps in the following chapters are intended to be high-level summaries for generalist reviewers that cover the main issues, as well as the advantages and disadvantages of the principal options available at each step. The focus of this report is completely on ridership and revenue forecasting of long-distance HSIPR services; consideration of intra-urban rail travel is beyond the scope of this report.

## 2 Trip table preparation

The end products of a HSIPR demand forecasting effort are predictions of the number of trips that will use the new or improved HSIPR service. As part of this process, the level of total trip making in the absence of the HSIPR alternative is established first at the origin-destination level. Then the demand for the HSIPR alternative is calculated from this total demand through the application of mode choice models (described later). The following sections describe in detail various processes and issues involved in quantifying the level of total trip making (i.e. the trip tables) in the absence of the proposed new HSIPR service or improvements in the HSIPR alternative.

Perhaps the simplest form of travel forecasting involves the extrapolation of observed usage levels on a specific transportation facility (say on a portion of a rail route, or a section of highway). This forecasting method is clearly infeasible for new services, and has limited ability to account for the effects on demand of exogenous or policy changes even for existing services. For these reasons, the preferred forecasting method is to identify first the patterns and levels of demand from origin to destination, and then determine how this demand will choose to route itself over the changed transportation network. The pattern and level of demand are most often represented in the form of a rectangular trip table or trip matrix, in which the element in row  $i$  and column  $j$  of the table represents the demand from origin  $i$  to destination  $j$ , the sum of the elements in row  $i$  is the total travel produced by zone  $i$ , and the sum of the elements in column  $j$  is the total travel attracted to zone  $j$ . Preparing the trip table(s) is a key aspect of most forecasting studies.

### Trip table segmentation

The above basic definition needs to be supplemented with additional details about trip table specification and preparation. For example, depending on the forecasting methodology, study requirements and data availability, trip tables may be segmented in various ways: for example, combined for all modes or by the individual modes, for various time periods (e.g. yearly, daily, hourly), and for different trip purpose or market segments.

These are discussed in the following paragraphs.

#### *Geographic detail*

As mentioned, the different rows and columns of the trip table correspond to the different traffic analysis zones defined for a forecasting study. Accordingly, the steps discussed above of defining the study area and the level of geographic detail of its representation essentially determine the identity and interpretation of trip table rows and columns.

#### *Universe of trips*

Depending on the intended mode choice modeling approach, a trip table might represent either total trips by all modes, or trips by a particular mode; in either case the table may correspond to some subset of the universe of trips since it will likely be segmented by trip purpose or other segmentation variables. One mode choice forecasting approach predicts the probability that the user of a particular base mode will divert to (say) a new high-speed rail mode; in this case, an individual trip table will represent trips by a given base mode, and there will be at least as many different trip tables as there are relevant modes in the no-build situation. Another mode choice forecasting approach considers travelers to choose from among all available modes without taking account of their prior mode choice. In this case, the trip table should notionally represent trips by all modes, although in practice the details of model calibration may require nonetheless the preparation of trip tables by mode.

### ***Time period***

Trip tables may represent total travel volumes over a full year, on representative days (e.g. a “typical” workday or weekend day), or during a portion of a day (e.g. a peak period), with less detailed studies typically preferring to represent the longer time periods without additional specificity. Preparation of trip tables that correspond to representative days or to periods within a day might be undertaken when a study wishes to take account of their differences in demand patterns and levels (e.g. to represent seasonal or weekend demand), different fare policies (e.g. peak vs. off-peak fares), or differences in the service levels provided by competing modes (e.g. periods of congestion on the highway network). In some cases, the availability of travel data for one of these time frames (e.g. from prior studies or modeling efforts) may favor the preparation of trip tables based on that time frame.

### ***Market segmentation***

Market segments are subgroups of travelers having distinct behavioral or other characteristics that set them apart from other subgroups in terms of their response to a change in the transportation system such as, for example, the introduction of a high-speed rail service. The response of a market segment to such a change is typically forecast using a model developed (“calibrated” or “estimated”) for that segment, and different from the models used for other segments. Criteria frequently used in defining market segments include trip purpose, trip length, traveler income, travel party size, and others. The number and identity of market segments defined in a study will depend on the study level of detail and data availability; some studies might only distinguish travelers by trip purpose (business vs. non-business), for example. Separate trip tables need to be developed for each market segment (whether the tables are for total trips or trips by mode), so the level of effort required to prepare the trip tables will be roughly proportional to the number of segments.

As a specific issue in market segmentation for high-speed rail forecasting, it has been found important to distinguish travelers who need to use an automobile at intermediate stops along a trip (e.g. business travelers or tourists making intermediate stops) or at the destination; the former are termed en route auto captive, while the latter are called destination auto captive. The response of each of these segments to high-speed rail is likely to be significantly different from the response of other market segments.

### **Data issues**

Ideally, the data sources used to prepare trip tables include details that allow the corresponding market segment to be directly determined; this is typically the case, for example, for trip tables developed from traveler surveys, in which questions about trip purpose, traveler characteristics, etc. can be included. On the other hand, collecting reliable travel data for an excessively large number of market segments may be difficult or impossible because individual origin-destination flows will tend to be small and difficult to sample accurately. Moreover, in some cases the available data on travel flows (traveler or vehicle counts, ticket sales, etc.) simply do not provide details about individual market segments. In these cases, it may be necessary to factor tables of total trips into trips by market segment using exogenous information. This is less accurate than obtaining per segment information would be, but may be the only option in some cases.

### **Approaches to trip table preparation**

In general terms, there are two main approaches to trip table preparation:

- Direct development from base data; and
- Synthesis from trip generation and distribution models.

Each of these will be discussed in turn.

### ***Trip table development from base data***

“Direct” trip table development methods estimate base year zone to zone flows explicitly, using a variety of data sources related to zone-to-zone trip making. Such sources may include publicly available data collected through regular travel survey or business reporting processes, as well as special-purpose surveys developed and undertaken for the purpose of a particular transportation study. Methods for collecting and processing data from these various sources are discussed in detail in the following section.

Some of the standard sources of publicly-available general-purpose travel survey data that have been used for intercity passenger travel demand forecasting in recent studies are noted below. Modal trip data from ticket sales on existing modes (air, rail etc.) which are extremely important in intercity trip table development, are described in a later section.

- **American Travel Survey:** The American Travel Survey (ATS) was completed in 1995 and contains information on long distance trip volumes and patterns as well as other trip characteristics: for example, trip purpose, vehicle occupancy, trip durations, and the number of stops en route. The ATS data was collected over a period of one year from 80,000 households across the U.S. who were interviewed during four survey-waves.  
Issues: However, the survey is now dated and needs to be updated with more recent data; and it often lacks sufficient detail to be an adequate representation of trip making behavior in a relatively limited area, such as an HSIPR corridor.
- **National Household Travel Survey:** The National Household Travel Survey (NHTS) was first completed in 2001 and has been updated periodically, with the last round of data collection in 2009. The NHTS focuses on trips taken in a 24 hour period, and has a relatively low sample rate, which limits its applicability for corridor planning.  
Issues: The 2009 NHTS does not have a good degree of geographic disaggregation, and must be used in conjunction with other data sources to gain meaningful insights on the specificities of trip travel patterns in a region. It can be a useful data source for adjusting trip tables and for better understanding trends in personal daily travel.
- **Census Journey to Work data:** The Census Transportation Planning Package (CTPP) is a set of special tabulations of Census data tailored to meet the data needs primarily of urban transportation planners. Tabulations are provided for various levels of geography including counties. Three sets of tabulations provided, with those in the third set of particular interest because of the detail that it provides about trips (e.g. origin, destination, mode, departure time) and network conditions (mean and median travel time) associated with the journey to work. The geographic detail allows information on longer-distance (including interstate) JTW trips to be obtained.  
Issues: As JTW trips are primarily urban and suburban, the CTPP information is of marginal relevance to intercity trip making. The CTPP was compiled from the long form of the decennial Census survey through 2000. The long form has now been replaced by the American Community Survey (ACS), as described next.
- **New Journey to Work data from ACS:** The American Community Survey (ACS) is a continuous statistical survey carried out by the U.S. Census Bureau that is sent to approximately 3 million households annually. The ACS provides less statistically significant data than the CTPP due to its smaller sample size. Nevertheless, 3- and 5-year summaries of the ACS are now being used to provide more up-to-date data and associated data products. In July 2010, the newest update of the CTPP was released based on 3 year ACS summary data.  
Issues: This update of the CTPP will include journey to work data aggregated at the county level but is restricted to counties with at least 20,000 residents. The US Census Bureau is in the process of using the 2006-2010 ACS data to develop the first CTPP based on a 5-year summary of ACS data; this will be the first CTPP using ACS that provides small area tabulations.

- █ FHWA intercity trip table development study: In recognition of the difficulties of obtaining high-quality data on intercity automobile travel in the US, the FHWA has recently sponsored a study of the issues and approaches to address them. This study is still in its early stages and it is not yet clear what specific results will become available from it.

***Trip table synthesis from trip generation/distribution models***

Conventional urban travel forecasting model systems synthesize trip tables using a sequence of computational steps that predict (i) the number of trips produced by and attracted to each individual zone; then (ii) the number of trips that will travel from one specific zone to another i.e., the zone to zone trip table, given the production and attraction totals determined in the first step. The total trips departing from a zone (its trip production) can be viewed as the sum of the elements of the corresponding row of the trip table; similarly, the trip trips traveling to a zone (its trip attraction) can be viewed as the corresponding column sum. The step that determines zonal production and attraction totals is called trip generation, while the step that predicts zone-to-zone trips given the trip generation results is called trip distribution.

Examples of trip production and trip attraction tables are shown in Table 2-1 and Table 2-2, respectively. In Table 2-1, there are 500, 200 and 300 trips traveling out of zones 1, 2 and 3, respectively. Similarly, Table 2-2 shows that zones 1, 2 and 3 attract 300, 400 and 300 trips, respectively. Based on these trip attraction and production numbers, the resulting zone to zone trip table is calculated as shown in Table 2-3. As mentioned above, the row sums and the column sums of the zone to zone trip tables are the trip productions and attractions of the corresponding zones, respectively.

**TABLE 2-1. AN EXAMPLE TRIP PRODUCTION TABLE**

Zone 1	Zone 2	Zone 3	Total
500	300	200	1,000

**TABLE 2-2. AN EXAMPLE TRIP ATTRACTION TABLE**

Zone 1	Zone 2	Zone 3	Total
300	400	300	1,000

**TABLE 2-3. AN EXAMPLE ZONE TO ZONE TRIP TABLE**

Origin/Destination	Zone 1	Zone 2	Zone 3	Total
Zone 1	-	300	200	500
Zone 2	200	-	100	300
Zone 3	100	100	-	200
Total	300	400	300	1,000

Standard methods have been developed for each of these steps. For trip generation, linear regression models or look-up tables are typically used to relate a zone’s total trip production and attraction to its socio-economic characteristics; in some cases, a zone’s accessibility (appropriately defined) to other zones may also influence its production and attraction. Some studies prefer to develop and apply disaggregate models of trip making level by individuals or households, which predict the probabilities of making particular numbers (including 0) of trips of a given kind (e.g. purpose) as a function of tripmaker characteristics,

accessibility and other factors; these results are then aggregated to the zonal level, taking into account the distribution of tripmaker characteristics, to obtain the total zone production of that kind of trip.

For trip distribution, standard methods include:

- Growth factor models, which modify a prior-year trip table to account for growth in zonal productions and attractions between the prior and the forecast years;
- Gravity models, which relate zone-to-zone flows to the corresponding zonal productions and attractions and to some measure of the difficulty of travel between them. For example, similar to Newton's law of gravitation, a gravity model assumes that the number of trips between a zone pair is directly proportional to the number of trips produced by the origin zone and the number of trips attracted by the destination zone and inversely proportional to some measures of travel impedance (i.e., travel costs, times etc.) between the origin and the destination zones; and
- Destination choice models, which explicitly model the probability that a tripmaker will choose a particular zone as the destination of her trip from a particular origin zone, given the tripmaker's characteristics, the respective zonal characteristics and the difficulty of travel between them; zone-to-zone flows are then obtained by aggregating the predictions concerning different types of tripmaker.

Validation of the base year trip tables prepared in this way is typically done via indirect checks (since a complete empirical trip table is usually not available to compare the predicted table against). For example, predicted zonal productions and attractions are compared against results obtained from travel surveys, and the distribution of trip lengths or costs implied by the predicted table is compared against sampled values.

With this method of trip table preparation, predictions of forecast year trip tables are prepared in a way that is essentially identical to that used to develop base year tables, using the same trip generation and distribution relationships. Of course, the future-year values of socio-economic, accessibility and other variables involved in these relationships must be available in order to perform these predictions.

### *Discussion*

Trip tables for the common carrier modes (air, rail, bus) will generally represent station to station (or airport to airport) flows. Depending on how they are estimated, auto trip tables may include flows at levels of geography ranging from zip codes to entire metropolitan areas. HSIPR studies will typically use some disaggregate level of geography for analysis purposes, such as zone-zone flows, and trip tables will therefore need to be converted to this level of geography for forecasting purposes.

The overall geographic region defined as the study area will have already been determined in the early phases of the study. However, the process of disaggregation will still need to identify the subset of zones to which the trips will need to be distributed. If airports/stations are assigned to zones simply using engineering judgment (based on some notion of the likely catchment area of the airport/station), then the trips should be distributed to those zones to which the airport/station is assigned. If the assignment is done using a station/airport choice model, estimated with actual trip end distribution data, then the trips should be distributed to all zones according to the model predictions.

Trips should be distributed to finer levels of geography in a way that accurately takes into account likely differences in the propensity for trip making across these finer geographic units. The propensity for trip making will generally follow the socio-economic characteristics of a zone, such as population and employment, but the number of trips that originate or terminate in a given area may also depend on factors such as the number of hotel rooms in that area. Importantly, previous research has shown that trip rates tend to increase markedly with income, and thus the distribution of trips should account for this effect.

No single approach is currently in wide use to address the issue of distributing station or airport pair volume data to individual zone pairs. For preliminary studies with limited resources, it may be sufficient to distribute trips based on zone level population and/or employment. Final stage studies should employ a formal trip distribution model that incorporates variables affecting the level of trip making that have been estimated through regression analysis. The reverse gravity model method (i.e. use of a gravity model to estimate origin-to-station and station-to-destination flow distributions) is one such method that is more sophisticated.

In cases where auto trip tables are compiled at the zip code level, there may be instances in which the boundaries of zip codes do not correspond exactly with zone boundaries (that is, a portion of the zip code may be within the zone or county and a portion may be outside of it). In these instances the conversion of zip code level trips to zone level trips should employ a method that is systematic, reasonable, and defensible, and that could be replicated if needed by an outside party.

Two implications of direct trip table development should be noted. First, because all the relevant information about zone-to-zone trip making may have been used in preparing the table, indirect verification, using additional data sources such as summary statistics or samples, may be used in a reasonableness checking process. The purpose is to demonstrate that the prepared trip table is consistent with this exogenous information about trip making.

The second implication is that a separate trip table forecasting procedure will be needed when the base year trip table is directly prepared from available data. Unlike trip table synthesis methods, which develop both base and forecast year trip tables from the same relationships, only base year trip tables can be directly prepared. Growth factor methods are frequently used to expand directly prepared base year trip tables to future years. The zonal (production and attraction) growth rates used by these methods might be derived from independent estimates of travel growth over time, or may be developed from socio-economic growth drivers. In the latter case, both trip generation-type relationships and elasticity-based approaches are commonly used.

*Advantages and disadvantages:* Direct preparation of trip tables from base year travel data has the great advantage of being based on observed travel data. On the other hand, there is then a need to develop a separate procedure to forecast the base year trip table to forecast years. Synthesizing trip tables through trip generation and distribution relationships is a more indirect procedure, but may be all that is feasible if available data sources do not provide a sufficiently complete picture of trip making patterns and levels to allow direct estimation. One advantage of trip table synthesis methods is that the relationships developed to prepare a base year trip table can be applied, typically without change, to forecast future year tables as well, without requiring additional methodological developments.

### **Potential impacts of trip tables on HSIPR forecasts**

The base and forecast year input trip tables very directly impact a study's ridership (and hence revenue) forecasts as the forecasts are calculated directly as diversions from these trip tables. It follows that any over- or under-estimate of the trip tables will translate to high or low forecasts of HSIPR ridership.

### **Methods for collecting modal trip table data**

Reliable information on modal travel volumes is a prerequisite for valid ridership and revenue forecasts. The competing modes from which HSIPR draws its shares are mainly private automobile, rail, bus and air. There are various methods for collecting trip table data for these modes. Brief descriptions of these methods are provided below.

## **Auto**

One of the biggest data challenges in HSIPR ridership and revenue forecasting is in the preparation of intercity auto trip tables. There is unfortunately no standard source of information about intercity auto trip making in US that is sufficiently detailed to be used in project-level forecasting. Given the large volumes of intercity auto trips, preparation errors can result in significant discrepancies in the predicted HSIPR ridership and revenue numbers, even though the percentage rail diversion from the auto mode is likely to be small in general. It follows that the development of intercity auto trip tables using the best available information and methods (consistent with the project scope and budget) is very important. There are various methods of collecting data on intercity auto trip tables.

### *Household travel surveys*

Household travel surveys ask household members about their recent/current travel behavior. For a HSIPR study, survey respondents might be asked about any recent intercity trips that they have made in the geographic area under consideration, including the origin, destination, number and purpose of all such trips. Responses are then weighted appropriately to obtain estimates from the population as a whole from the survey's sample statistics. The weighting factors depend on the relationships between the household samples and the population under consideration.

Household surveys are typically administered through mail-back questionnaires (where questionnaires are given to respondents to be filled out later and mailed back), or via the internet (where respondents are informed how to access a web-based questionnaire). There are various ways of recruiting the potential survey respondents, such as through examination of license plates, or through intercepts (on the roadside or in other places like employment centers, shopping malls etc. in the corridor). Pre-recruited market research panels can also be used for household travel surveys and are described below.

*Advantages and disadvantages:* Household travel surveys can potentially be good sources of data on the level of intercity auto trip making. However, the biggest issue about these surveys is the difficulty of finding enough respondents who have made "candidate" intercity trips in the corridor under consideration and ensuring the representativity of the recruited respondents. As such, the sample size requirements (and hence costs) of these surveys can be very high. Moreover, the prohibitive cost of household or intercept surveys could be another deterrent to their use.

### *Market research panels*

Another option to undertake household travel surveys is through the use of members/households belonging a pre-recruited consumer panel. A market research firm enrolls a group ("panel") of people who have expressed their willingness to participate in surveys on a wide variety of topics when requested. The market research firm may compensate panel members for their participation and loyalty (for example, with small gifts and points programs) but this is done in a way that does not influence their responses to survey questions. When the firm is hired to conduct market research on a particular topic, it surveys a random selection from qualified participants in its pre-recruited panel rather than drawing a new sample from scratch. Alternatively, travel questions may be added to panel participants located in the study area during one or more of the panels' regularly scheduled surveys.

Panel members can take the survey through internet, mail-back questionnaire or even telephone. Until about ten years ago, there were a number of large national pre-recruited panels that could be used for mail and telephone surveys, whereas internet surveys are now the norm. For local or regional transportation planning purposes such use has been limited by the number of panel members within the area of interest. However, for intercity travel between large metropolitan areas this is less of a problem.

*Advantages and disadvantages:* These panels have been a feature of the US market research industry for a long time, and they have several pros and cons. In summary:

- The panels are often so large that it is feasible to achieve samples that mirror quite closely the demographic/socio-economic characteristics of the population of interest;
- Response rates in the 60% to 70% range can be routinely achieved among panel members, whereas most surveys using a random selection from the general population typically have difficulty reaching a 25% response rate;
- Depending of course on the nature of the survey, the cost per respondent of surveying panel members may be much lower than other types of survey.

However,

- Almost all panels are of an “opt-in” nature, comprising people who have volunteered to be on a panel. No matter how closely the sample demographic characteristics may represent those of the parent population, there is always a concern that the volunteers may be atypical in some way that is significant to the nature of the survey;
- The number of households that must be surveyed to achieve any specific target number of corridor travelers depends critically on the percentage incidence of corridor travel over a reasonable recall period, and this can only be approximately estimated. Whether any one panel will be sufficiently large to meet the needs, or whether it would be necessary to use the panels from more than one vendor, is always a critical question;
- Over the last ten to fifteen years, in response to the growth of online access and declining cooperation rates for telephone-based surveys, the US market research industry has increasingly migrated to online surveys, and panel vendors are no exception. However, between a quarter and a third of the US population lack regular Internet access, and there is no online equivalent to the random sampling of street addresses or telephone numbers. As a consequence, online survey samples exclude a significant portion of the population with no regular Internet access. However, given that intercity travel rates and Internet use increase rapidly with income and education levels, this may be less of a problem than otherwise.

#### *Intercept surveys*

One way of obtaining more detailed information on the origin-destination patterns of vehicles on study area roadways is by conducting an intercept survey. Survey respondents may be intercepted at the road-side (stopping them during their travel), or in other places (e.g. workplace, shopping malls, restaurants etc.) within the study corridor where they have already stopped.

Road side intercept surveys are one of the most effective ways of collecting data to prepare the intercity auto trip tables. By using short personal interviews, mail-back self-completion questionnaire or, if available, by handing out the internet address of the survey questionnaire to be filled out later, respondents are asked to describe the characteristics of their trip while it is quite fresh in their mind. To maximize the incidence of intercity auto travelers in the survey sample size, these surveys should be administered on highways at the entrances and exits of toll facilities, or by stopping traffic on the principal interstate highways and other primary routes or on on-ramps to the highways. This typically requires the involvement of relevant state or local authorities and law enforcement personnel to ensure that the interception is done in a thoroughly safe manner, does not create congestion delays, and yet maintains the statistical integrity of the sample.

Roadside intercept surveys are sometimes conducted at rest areas by interviewing the drivers of a sample of vehicles parked there. This has the potential to cause a significant over sampling of non-business travelers, travelers with significantly lower values of time, and travelers making longer trips. These biases will affect the resulting trip tables and any mode choice models developed using these data.

*Advantages and disadvantages:* Roadside intercept surveys (if conducted appropriately and at appropriate locations) can be a very effective method of collecting data to prepare intercity auto trip tables. As

respondents are intercepted during their trip, this method maximizes the incidence of intercity auto travelers in the survey sample size. However, the logistics and costs involved for such surveys can be a significant deterrent in undertaking such surveys. Surveys of travelers stopped at rest areas produce biased samples and should be avoided. Interception of respondents at other non en route locations has the risk of not capturing enough intercity travelers (in the corridor under consideration) in the survey sample.

#### *Trip table estimation from traffic counts*

A variety of methods have been proposed that process roadway link-level traffic count data to estimate auto trip tables. These methods all attempt to find a trip table such that, when it is assigned to the network, the predicted link flows closely match the link count data. These methods must address the fact that the estimation problem is “underdetermined” - a given set of traffic counts does not generally imply a unique matrix. Since, in general, this problem does not have a unique solution, most such methods allow a “seed” matrix to be specified; this matrix is used as an initial trial solution, and then is iteratively modified until a solution consistent with the link counts is found.

The seed or input matrix can take any values - starting from a prior trip table in the region to arbitrary small positive values for every cell, if there is no prior information on the flows. The accuracy of trip tables estimated this way depends on the accuracy and quality of seed matrix and traffic count data as well as the coverage of the area by traffic counts.

In choosing traffic count locations, great care should be taken to select locations that capture mostly intercity, long distance flows (and not local, short trips) in order to obtain counts that relate directly to the demand of interest for HSIPR forecasting.

*Advantages and disadvantages:* Traffic counts are regularly maintained and updated by state DOTs and are made available to public so this method of estimating trip tables is fairly inexpensive. However, the accuracy of trip tables thus developed can be questionable because 1) it is often difficult to find a reliable “seed” matrix to start with; 2) the matrix estimation problem does not have a unique solution; and 3) it can be difficult to find a sufficient number of locations to obtain a good coverage of the study area while only counting intercity auto trips. Moreover, the count data and estimated matrix only refer to total auto volumes, with no further breakdown into trip or market segment. Nevertheless, this can be an attractive option when other sources to develop intercity auto trip tables are not available.

#### *License plate surveys*

The use of license plate data to obtain information about trips and trip making patterns is becoming increasingly widespread. Technology has advanced to the point where, under good observation conditions, vehicle registration numbers can be captured from license plates using video cameras, and the numbers can be automatically extracted from the video images using optical character recognition software. The extracted registration numbers can then be used in a variety of ways.

For example, the address of the registered owner can be obtained and a travel survey can be mailed. However, this requires the cooperation of the vehicle registration authorities and raises privacy concerns; moreover, the response rate to surveys of this type is typically not good. On the other hand, the addresses of the vehicle owners can provide a strong indication of the proportions of through vs. local traffic on rural interstate highways, and even of intercity OD volumes with which to populate seed matrices for estimating trip tables from traffic counts.

In addition, if registration numbers are obtained at multiple locations, it is possible to develop information about vehicles’ paths through the network. The problem here is that the accuracy of the method depends on not losing track of vehicles, so developing such information in a large and/or complex network requires recording license plate data for all vehicles at many locations over a given time period, and maintaining very high registration number extraction rates throughout.

*Advantages and disadvantages:* With the improvement in video technologies and optical character recognition, reliable vehicular movement data can be obtained which can then be used in trip table developments. However, the accuracy of this method is completely dependent on the capture rate of the vehicles. The higher the capture rate, the better the accuracy. Moreover, in order to obtain enough information for trip table development the coverage area for video capture has to be broad. Also, the coverage time has to be sufficient to capture, for example, weekday and weekend variations. Above all, the cost factor is a very important issue for license plate surveys. The cost for such surveys can quickly escalate with the increase in the coverage area and coverage period. The cost of these surveys is a major impediment to the use of this method.

#### *Cell phone based data*

An increasingly popular option for obtaining automobile origins and destinations is the use of cell phone based data. Currently in the US a few companies have permission to collect and process cell phone data. These companies have patented, developed and deployed technologies that mine anonymous signaling data from wireless networks to detect the location and movement of mobile devices. This source of data is currently being used to develop auto trip tables in rail studies in the U.S.

*Advantages and disadvantages:* There are some outstanding issues as yet about this method:

- Guidelines on the time and frequency of detecting any particular cell phone to be included in the dataset are still being developed and have not matured yet;
- Lack of continuous cell phone coverage in a study corridor could result in identification problems;
- A limited subset of cell phone data is currently available as some major cell phone carriers have not yet joined the program;
- Above all, the technology is still evolving.

This is a relatively inexpensive way of getting useful information to develop the intercity auto trip tables. If the conversions of data from the sample to the population is done properly and the above mentioned issues are resolved, this method has the potential to be a very reliable source of data on intercity auto movements.

#### *Other sources*

Other commercial vendors (e.g. TomTom, Google etc.) have also started to provide information on auto trip tables. The cost and quality of the data from these sources is still being assessed.

### **Rail**

Rail travel of interest in HSIPR studies include trips on Amtrak and (to a lesser extent) commuter services operating in the study corridor. Good quality trip table data are maintained both by Amtrak and commuter rail agencies.

#### *Amtrak data*

Amtrak maintains detailed historic information on the number of trips between each station pair for all its services in the country. Amtrak maintains this data for individual train runs, but ridership information at this level of detail is typically not publicly available.<sup>1</sup> Amtrak, upon request, has made detailed station pair level data available for some HSIPR studies. Since the desired trip tables are between traffic analysis zones (TAZs), these station pair rail volumes generally need to be disaggregated to TAZ pair volumes. This involves the calculation of the relative proportions of the total number of trips at the station that are originating

---

<sup>1</sup> Information on total ridership and revenue by month and by route are available on Amtrak's website. In addition, boarding and alighting data are also available at station level.

from and going to each TAZ within the catchment area of each station. This is usually done using information on the socio-economic and demographic characteristics of the TAZs as well as the relative proximity of the TAZs to the station under consideration.

#### *Commuter rail data*

Commuter rail agencies in the country typically maintain some form of ridership data that may be obtained through specific requests. Commuter rail trips are generally less relevant for HSIPR studies and should only need to be accounted for when 1) the improvements in track conditions may improve commuter rail services that share the same track with the HSIPR service; and 2) the HSIPR service may have the potential to attract riders from commuter rail services that run between the same station pairs as the HSIPR service (either on shared or separate tracks).

*Advantages and disadvantages:* Rail travel volume information from both Amtrak and commuter rail agencies should be used to develop base year rail trip tables whenever available. Otherwise, more aggregate data (route level data, station boarding and alighting etc.) can be used depending on the level of accuracy required for any given HSIPR study. Care should be taken to disaggregate station pair data to the TAZ level to represent TAZ level rail volumes as accurately as possible.

#### *Air*

Excellent data is available for building air trip tables due to federal regulations requiring the airlines to report the details of their operations to the USDOT. Specifically, carriers are required to file a document known as Form41. In addition to data about costs, revenues, and financial performance, the Form41 filing includes information on the total number of passengers carried between each airport pair served. The USDOT maintains two types of databases containing the traffic data reported on Form41:

- Airline Origin and Destination Survey (DB1B), which contains the reported 10% sample of actual flight coupons; and
- Air Carrier Statistics (Form 41 Traffic), which contains the reported total passengers flown between airport pairs.

The DB1B database, because it is based on a 10% sample of actual flight coupons and the sampled coupons filed with USDOT, contains the complete itinerary of a given airline ticket. Therefore, it can be used to differentiate between truly “local” traffic and connecting traffic traveling between a given airport pair. It should therefore be used for constructing the air trip table for an HSR ridership and revenue forecasting study. The DB1B data is available in three distinct data tables, as follows:

- **DB1BCoupon:** This table provides coupon-specific information for each domestic itinerary of the Origin and Destination Survey, such as the operating carrier, origin and destination airports, number of passengers, fare class, coupon type, trip break indicator, and distance.
- **DB1BMarket:** This table contains directional market characteristics of each domestic itinerary of the Origin and Destination Survey, such as the reporting carrier, origin and destination airport, prorated market fare, number of market coupons, market miles flown, and carrier change indicators.
- **DB1BTicket:** This table contains summary characteristics of each domestic itinerary on the Origin and Destination Survey, including the reporting carrier, itinerary fare, number of passengers, originating airport, roundtrip indicator, and miles flown.

The DB1B data is available on a quarterly basis and historical data is available on the USDOT Bureau of Transportation Statistics’ Transtats website back to 1993. Earlier data may be requested from the Office of Airline Information. The raw data may be downloaded free of charge from the Transtats website but is also available for sale from data vendors that provide the means to make targeted queries of the database (the

raw data files are generally quite large, and significant processing may be required to generate trip tables from the raw data).

The Form41 traffic data is compiled by USDOT into the T-100 database, which contains air traffic by airport pair and carrier. The T-100 database is available in two separate types of data tables, as follows:

- **T-100 Market:** This table represents an accounting of travel between pairs of airports, regardless of the routing. That is, it represents both nonstop flights as well as connecting flights and flights that have an intermediate stop. It includes data on carrier, origin airport, destination airport, and service class for enplaned passengers, freight and mail.
- **T-100 Segment:** This table represents an accounting of only nonstop flights between two airports. It includes data on carrier, origin airport, destination airport, aircraft type, and service class for transported passengers, freight and mail, available capacity, scheduled departures, departures performed, aircraft hours, and load factor.

The T-100 data is available on a monthly basis and historical data is available on the Transtats website back to 1990 (earlier years may be requested from the Office of Airline Information). Like DB1B data, the basic T-100 data may be downloaded free of charge from the Transtats website, and may also be purchased from data vendors that post-process the data and allow targeted queries of the database.

The T-100 database is an excellent source of information on capacity and can be used to calculate average load factors for an airport pair route. But because these data represent total onboard passengers regardless of their ultimate origin and destination, T-100 does not provide an accurate accounting of the number of passengers that are truly “local” to a given airport pair. That is, it cannot be used to identify only those passengers that are originating and terminating their trips within a given corridor. This can be a very important distinction in a HSIPR study, particularly in a corridor containing a large hub airport that serves a large amount of connecting air traffic.

#### *Data reporting issues*

The air trip tables used in HSR proposals for corridors where a significant fraction of the air service is provided by small commuter carriers should be given special scrutiny due to certain data reporting limitations in the USDOT databases.

The USDOT databases contain data only from carriers required to report Form41. For T-100 data, this includes all carriers classified by USDOT as either large certificated air carriers, small certificated air carriers, or commuter air carriers, which means any airline or air taxi operator carrying passengers on service with published schedules on at least five round trips per week on at least one route between two or more points.

However, carriers required to report 10% of their flight coupons for the DB1B database include only the large certificated air carriers, defined as those US airlines operating aircraft with a maximum passenger capacity of more than 60 seats or with a maximum payload capacity of more than 18,000 pounds, or that operate international service. Traffic carried on carriers too small to meet these criteria will not be counted in DB1B but will appear in the T-100 database. However, while the traffic on these carriers will be represented in T-100, this database does not allow the identification of the portion of onboard passengers that are truly “local” (originating and terminating their travel within the corridor).

In these cases, the HSR proposal will need to estimate what portion of the T-100 traffic on these carriers is connecting vs. local, and the reasonableness of the estimate should be examined carefully. To the extent that these smaller commuter carriers are providing code-sharing services under the brand name of a major network carrier and serving a major hub of that network carrier, it may be reasonable to assume that most

of their passengers are connecting traffic. A list of the carriers reporting to the DB1B database is available on the Transtats website.

Prior to 2003 the traffic for these smaller commuter carriers was compiled into a separate database known as 298-C. From 2003 onward these data have been included in the Form41 traffic statistics as described above.

Potential high volume HSR corridors are in almost all cases served by large air carriers that report to the DB1B database, and hence good and reliable estimates of true OD volumes for air trips are available. The limitations of the T-100 data should not be a problem for such corridors. Similarly, to the extent that a HSR project is intended to serve longer distance (“intermodal”) access to major hub airports, and thus replace small air carriers serving connect traffic, the T-100 data can be very useful.

#### *International air traffic data*

HSR proposals that involve international service should be given special scrutiny because of other limitations in the USDOT air data. While rare, these circumstances would include situations where the HSR system serves connecting air passengers at a major airport, including those passengers connecting from international flights.<sup>2</sup>

In these situations, HSR proposals need to recognize that the DB1B database includes only data from US carriers, and therefore may not provide a complete accounting of international air trips in a given corridor. In addition, the USDOT restricts access to the international portion of the DB1B database, and permission must be obtained from the Office of Airline Information before any of these data can be obtained.

The T-100 database does not have such restrictions, and it also includes traffic reported by foreign flag carriers. However, because it represents only total onboard passengers as described above, it likewise cannot be used by itself to provide an accurate accounting of truly “local” (that is, “true origin/destination”) international passengers in a given corridor.

*Advantages and disadvantages:* Table 2-4 provides a summary of the data sources available for the construction of air trip tables and their respective advantages and disadvantages. The table shows that both data sources have features that make them attractive for building air trip tables but also that each has issues that must be addressed to ensure an accurate result. The primary advantage of the T-100 database is that it is an unambiguous count of 100% of the passengers traveling on scheduled flights between two airports, regardless of the size or other characteristics of the airline.

However, the T-100 database’s biggest weakness is that it does not distinguish passengers traveling “locally” between the airports from those that may be connecting on one end. As such, it may significantly overstate the size of the “local” travel market. The DB1B database, by contrast, is specifically designed to distinguish between “local” and connecting traffic. At the same time, because all airlines are not required to report data to DB1B, this database has the potential to understate the size of the local travel market. The successful construction of an air trip table therefore requires a strategy for leveraging the best features of each of these data sources while mitigating the effects of their limitations.

---

<sup>2</sup> A practical example of the latter situation is the French TGV service at Roissy Charles de Gaulle Airport in Paris, which provides connecting “intermodal” service to points within France and the continent for passengers arriving from outside the country.

**TABLE 2-4. ADVANTAGES AND DISADVANTAGES OF DIFFERENT DATA SOURCES FOR AIR TRIP TABLES**

Source	Database	Periodicity/Availability	Advantages	Disadvantages
Air carrier statistics (Form41 traffic)	T-100	Monthly 1990-present	<ul style="list-style-type: none"> <li>• Detail by airport pair and carrier</li> <li>• 100% of onboard passengers transported</li> <li>• All airlines</li> </ul>	<ul style="list-style-type: none"> <li>• No ability to distinguish “local” vs. connecting traffic</li> <li>• May overstate size of “local” market</li> </ul>
Airline Origin-Destination Survey	DB1B	Quarterly 1993-present	<ul style="list-style-type: none"> <li>• Detail by airport pair and carrier</li> <li>• “True origin/ destination” traffic</li> <li>• Ability to measure amount of local vs. connecting traffic</li> </ul>	<ul style="list-style-type: none"> <li>• Potential sampling issues</li> <li>• May miss traffic on small commuter carriers</li> <li>• No foreign carriers</li> <li>• Restrictions on international data</li> </ul>

**Air access**

Studies that examine the potential of HSIPR to serve as an airport ground access mode will require data on the geographic and modal distribution of airport access trips. Airports generally compile statistics on total passenger volumes, and these figures are also reported by the FAA in its Terminal Area Forecast database. Airport access trips will equal the total number of originating passengers (that is, not through or non-transfer passengers) using the airport. Most airports will have an estimate of the number or fraction of connecting passengers, and in any case this fraction is likely to be quite small for airports that do not serve as a hub for a major carrier.

Detail on airport access trips by mode may be available directly from the airports. Ground access surveys of departing air passengers will generally collect these data, but only very large airports conduct these surveys with any regularity. Other departing passenger surveys, such as those conducted by the airport to measure customer satisfaction may also provide this information. Local MPOs that have modeled airport access trips as part of their planning process may also be able to provide estimates on the mode shares of airport access trips. In general, at smaller airports where these data are least likely to be available, most airport access trips are likely to be by auto.

If information on access trips by mode is not available from any of these existing sources, a new departing passenger survey will be required to collect it. However, given the scale required to produce a representative sample, these surveys tend to be expensive. It is therefore most economical for airport access mode information to be collected as part of departing passenger surveys conducted for the purpose of collecting data for intercity mode choice model estimation. Such new surveys would only be expected to be conducted for a Final stage HSIPR study.

**Bus**

Intercity bus service in the US is provided by commercial operators, who tend to treat ridership information for individual routes or services as commercially sensitive. Accordingly, such information is typically difficult or impossible to obtain.

Because intercity bus ridership is typically a small fraction of the corresponding automobile, air or rail volumes, and because it is sometimes felt that bus riders have low values of time and are unlikely to divert to high-speed rail, many studies consider rail-bus competition to be negligible. Against this must be set the recent successful development and expansion on some routes (for example in the Northeast Corridor) of intercity bus services providing high quality (comfortable seats, WiFi access, en route refreshments) connections at relatively low cost. Of course, these services still operate over the highway network and are exposed to the risks of congestion. On routes where such services attract significant ridership, their competition with high-speed rail will need to be explicitly investigated as part of a rail forecasting study.

When it is desired to estimate bus ridership, use of published schedules combined with observations at terminals and stops of load factors may be the most suitable approach to develop useable estimates.

### ***Other***

The above methods and sources do not exhaust the range of possibilities for collecting base year trip table data to be used in intercity passenger rail forecasting studies. In some cases, rail travel may be bundled with other activities for tourists or for business events. In other cases, HSIPR may itself be viewed as the attraction. These should be investigated in greater detail using sources focused on those activities.

#### *Tourism surveys and studies*

A number of HSIPR projects in the US have explicitly focused on serving tourists and recreational travelers to specific major tourist destinations; this is perhaps understandable given, for example, the novelty of high-speed rail in this country. For such projects, it is important to understand the tourist market and the likelihood of its use of rail. General purpose surveys conducted by area business groups or tourism associations can be useful for sizing the potential market, but additional surveys, specifically designed to assess the likely reaction of tour packagers and tourists to a rail travel option, are typically required for more detailed levels of study. Results from such surveys may be used to develop separate ridership forecasting models that focus exclusively on the tourism market segment.

#### *Business surveys and studies*

Similarly, a rail project may be intended in large part to serve business travelers in a particular corridor. Again, it would be important in such a case to develop data and models that accurately represent the characteristics of this market, and predict the response of business travelers to a new or improved rail option. As before, specially-designed surveys would normally be used to obtain this information at more detailed levels of study.

### **Segmenting trip tables by purpose**

Because values of travel time and demand elasticities are known to vary between trip purposes (e.g. business and leisure travelers), HSIPR forecasts should, at the very least, segment trip tables by trip purpose. Auto trip tables developed from household travel surveys, intercept surveys and license plate surveys can be segmented based on the trip purpose information that is also collected as part of the surveys. However, auto trip tables calculated by trip matrix estimation generally do not include information sufficient to accomplish the segmentation. Similarly, rail and bus trip tables calculated from Amtrak and other operating agencies and the databases described above used for air trip tables also lack segmentation information.

The most reliable source of trip purpose information is a carefully designed survey - household travel surveys, intercept surveys, license plate surveys for auto trips; on-board and intercept surveys for rail and bus trips and departing passenger surveys conducted at the major airports for air trips - conducted in the corridor of interest to the HSR ridership and revenue forecasting study. These surveys may be carried out for the primary purpose of collecting stated preference information for use in mode choice modeling, but it is very important that they also collect some revealed preference information (such as trip purpose) as it can be very helpful in creating the trip tables.

If, due to budgetary or other constraints, the HSR forecasting study does not include new primary data collection, there are other sources that may provide trip purpose distribution or other relevant revealed preference information. These include the following:

- **Existing surveys:** Data from existing travel surveys may already be available in the corridor under consideration. These can be very useful source of input for market segmentation by trip purposes.

- American Travel Survey:** As mentioned before, this database contains national data on the nature and characteristics of long-distance personal travel, from a household survey conducted by BTS. The data is also now quite dated, however, as it was last conducted in 1995. The survey also includes detailed data for a select subset of metropolitan areas, though its applicability in estimating corridor specific trip purpose distributions will depend on whether the metro areas of interest were included in the detailed sample, and if so, whether trips between these areas were sufficiently sampled to produce reliable results.
- Existing studies:** Existing travel demand forecasting studies in the corridor under consideration or similar corridor elsewhere may also have information on trip purpose distribution for trip tables of various modes. In the absence of any other source, these may be valuable for information on market segmentation.
- Airport ground access surveys:** Many major airports conduct periodic, large-scale departing passenger surveys for the purposes of ground access planning. These surveys typically collect a wealth of revealed preference information, including trip purpose, group size, access mode, and trip origin location, as well as data on carrier and flight number, and destination airport. At larger airports these surveys tend to employ large samples, but care should still be taken when applying the results as the number of responses for a given airport pair may not always be large enough to produce reliable results.

### **Forecasting future year trip tables from base year tables**

Forecast year trip tables are generally created from base year trip tables by the application of growth rates derived from one of the following three sources:

- Direct demand models;**
- Pre-existing forecasts; and**
- Assumed growth rates.**

Each of these potential sources is discussed in more detail below.

#### ***Direct demand models***

Direct demand models, as the name implies, predict the number of trips “directly” based on various factors known to influence travel behavior. The trips could be total trips for all modes combined or, alternative, total trips for each specific mode. They are estimated econometrically using a regression analysis that typically relates total observed trip levels to socio-economic factors such as population and/or income and level of service factors such as costs, scheduled travel times or generalized costs. Direct demand models are often estimated from time-series/cross sectional data incorporating observations for several years as well as for all or a sample of zone pairs in the corridor of interest. When using time series data, recession years and years with other unusual events (e.g. gas price spikes, etc.) need to be recognized and accounted for appropriately in the model development.

Historical data on auto trips are almost never available. As a result, direct demand models estimated for auto trips often use only cross sectional data on base year auto trip making. Historical information on rail, bus and air trips can be obtained from Amtrak, the bus operating agencies and the same USDOT sources used for creating the base year trip tables (described above), respectively. Historical socio-economic data such as population, employment, income and other variables can be obtained for specific metropolitan areas from the US Census Bureau, local Metropolitan Planning Organizations (MPO) and other private vendors (e.g. Woods and Poole) etc. Historical information on modal service characteristics can be obtained from, for example, the following sources:

- Auto LOS characteristics from MPO and statewide models, past studies and traffic counts;
- Rail LOS characteristics from Amtrak; and
- Air fares from the DB1B database described above, and historical data on other airline level of service features such as travel time from the Official Airline Guide (OAG).

While attractive for their specificity and close connection to the HSR corridor of interest, the use of custom-developed direct demand models will not necessarily always be the preferred method for forecasting future modal trips. The results must still be tested for reasonableness, and if models that conform to prior expectations about the factors that influence modal travel cannot be developed (or differences cannot be satisfactorily explained), then an alternative approach may be warranted.

#### ***Growth rates from pre-existing forecasts***

If time or budgetary resources do not allow for the development of a direct demand model (such as in a preliminary study), one alternative for growing the base year trip table into a future year trip table is the application of growth rates from existing forecasts. Care should be taken to recognize the difference in growth rates between urban and intercity travel, especially by auto. In most cases, these forecasts are at higher level of geographic aggregation than the zone pair level calculated by direct demand models. Sources of these growth rates include the following:

- **MPO and statewide model forecasts and forecasts from existing studies:** Local MPOs and existing statewide studies produce forecasts that can be used to calculate growth rates. These growth rates may be at a zone pair or more aggregate level.
- **Airport forecasts, FAA Terminal Area Forecasts (TAF) and Aircraft industry forecasts.** Larger airports will often commission their own forecasts on a periodic basis, typically to support the development of an Airport Master Plan or for other general planning purposes. The Federal Aviation Administration (FAA) regularly publishes forecasts for all airports in the US with commercial air service. These forecasts are generally not OD market-specific, but rather contain only growth rates for the single airport. The major OEM aircraft producers (Boeing and Airbus) both publish annual forecasts of world air travel demand by region, as do some makers of regional aircraft (e.g. Bombardier) and aircraft engines (e.g. Honeywell). The forecasts are not airport specific.

Both economic conditions and the nature and extent of modal competition can vary widely among metropolitan areas or regions. As a result, growth rates can be expected to vary, and sources that apply to a region or the entire country will not capture these potentially important differences. Growth rates that are as specific as possible to the markets of interest are therefore preferable.

#### ***Assumed growth rates***

Another alternative if resources are not available for the development of a new direct demand model and zone or city pair-specific forecasts are not available (or are deemed inappropriate) is the application of assumed growth rates based on some corridor-specific indicators of the likely future trend in modal trips.

The direct demand models described above are typically based on socio-economic data, as are the models used by airports, the FAA, and the aircraft industry. Socio-economic projections for individual metropolitan areas can be used along with assumptions about the likely corridor specific relationship between these measures and modal travel to estimate mode specific travel growth rates for the zone pairs within a given corridor. More specifically, for example, based on the particular economic and demographic circumstances of a corridor, possible available direct demand models of intercity travel in similar corridors in that part of the country, and the nature and extent of existing service for different modes, one can assume that travel for any specific mode will grow faster, slower, or at about the same rate as some particular socio-economic variable (population, employment, regional product, etc.) Forecast growth rates for mode specific travel

can then be based on projections of the socio-economic measures. In this way, trip tables are sometimes grown based on specific demographic growth rates (e.g. non-business trips may be grown using population growth rates; business trips may be grown using employment growth rates; tourist trips may be grown using growth in hotel rooms, etc.)

Projections of population for individual Metropolitan Statistical Areas (MSAs) are available from the US Census Bureau, and income forecasts by MSA are available from private data vendors and in some cases from states and local Metropolitan Planning Organizations (MPOs).

*Advantages and disadvantages:* The choice of method for forecasting future modal trips will depend to a significant extent on the study level and the project’s available resources. The development of custom econometric direct demand models, while potentially the most precise method, may also be the most expensive and therefore may be most appropriate for final stage studies. These models are also not guaranteed to produce reasonable results. Because the assumed growth rates have a significant impact on the level of forecast HSR ridership and revenue, the credibility and defensibility of the growth rates used, apart from the methodology from which they were derived, should be an important determinant in the choice of the growth rates used and a key focus of proposal evaluations.

Table 2-5 below provides a summary of sources from which growth rates may be obtained for the purposes of creating a future year mode specific trip table from the base year trip table, a summary of their respective advantages and disadvantages, and an indication of which methods are most appropriate for each study level. The table shows that each method has both positive and negative features, and that some methods are more appropriate for different types of study levels. The credibility and specificity provided by developing new econometric direct demand models must be balanced against the time and expense required and the risk that this effort may ultimately not be successful. The use of pre-existing forecasts or assumed growth rates can be much more economical, but at the same time more limiting if these methods lack the credibility required to help a proposal move to the next stage.

**TABLE 2-5. ADVANTAGES AND DISADVANTAGES OF TRAVEL GROWTH RATE DEVELOPMENT METHODS**

Category	Advantages	Disadvantages	Appropriate for Study Level
Direct demand model	<ul style="list-style-type: none"> <li>Developed specifically for the HSR study</li> <li>Corridor and market-specific</li> <li>Most rigorous method</li> </ul>	<ul style="list-style-type: none"> <li>Can be expensive and time consuming</li> <li>Not guaranteed to produce reasonable results</li> </ul>	Final
Pre-existing forecasts	<ul style="list-style-type: none"> <li>Often available at city pair level and based on local data</li> <li>Official or industry standard sources provide credibility</li> <li>Limited time and expense required</li> </ul>	<ul style="list-style-type: none"> <li>May lack geographic specificity</li> <li>May not represent latest market conditions</li> </ul>	Preliminary/ Intermediate
Assumed growth	<ul style="list-style-type: none"> <li>Most economical</li> <li>Can be made corridor/market specific</li> </ul>	<ul style="list-style-type: none"> <li>Based in part on judgment rather than objective data</li> <li>Lacks credibility that comes from industry standard sources or econometric analysis</li> </ul>	Preliminary

### 3 Modal level of service data

Travel demand forecasts require information on the level of service (LOS) characteristics of the various travel modes that compete to serve travel demand between origins and destinations. This is used to provide a quantitative determination of the service characteristics offered by the various available modes and the proposed HSIPR mode, and to represent the way in which trips will follow specific modal routes, links, stations and other facilities as they go from origin to destination.

Modal LOS characteristics are the inputs to the mode choice modeling process where these are taken into account directly in determining the relative attractiveness of the HSIPR mode and the other available modes. In addition, these data are also used in direct demand model development as well as in validation of the network representation (if used) of the transportation system.

Various LOS characteristics data are required for all the modes in any HSIPR studies. The levels of detail of these data items usually vary depending on the study stages - preliminary, intermediate, final. LOS data are required for all modes and for base and forecast years. The following sections discuss the required data items for HSIPR ridership and revenue forecasting studies by mode in more detail.

#### Air

Data on air Level of Service (LOS) characteristics are readily available from a variety of sources. Here we discuss six types of LOS data, as follows:

- Line haul travel times;
- Access/egress times;
- Fares;
- Access/egress costs;
- Frequency; and
- On-time performance.

Each of these areas is discussed in more detail below.

#### ***Line haul travel times***

Airport-to-airport line haul travel times should be derived directly from airline schedules. The airlines' published schedules show the time the flight is scheduled to leave the gate at the departure airport and the time the flight is expected to reach the gate at the arrival airport, and are therefore directly analogous to the published timetables of rail services. Line haul travel times can therefore be derived by subtracting the scheduled arrival and departure times (and accounting for any time zone differences, if applicable).

Airline schedules can be accessed free of charge on the airlines' respective websites, and many airlines also provide a complete timetable for their entire system that can be downloaded in electronic format. These sources, while free, can also be somewhat cumbersome to access, as they may require multiple queries to multiple websites (to obtain schedules for multiple airlines for airport pairs, for example), and significant post processing, as neither the websites nor the electronic timetables typically allow the direct export of the data to spreadsheet or database formats.

Another source of airline schedule data that is generally accepted as the industry standard is the *Official Airline Guide* (OAG). OAG regularly publishes the schedules of all carriers providing commercial service, receiving the data directly from the airlines and maintaining information for scheduled flights several months in advance. The OAG is available in both a printed version as well as in a database in electronic

format. The entire database may be purchased for a flat rate subscription fee, and the company also provides ad hoc queries that can be purchased on a one-off basis.

### ***Access/egress times***

Travel times to and from the airport from various points within a given metropolitan area should be derived from the networks typically maintained by local Metropolitan Planning Organizations (MPOs) and/or state Departments of Transportation (DOTs). These times should not only include the access options by auto but also by any transit service that may be available and by walking. These same networks may be used to derive access/egress time to HSIPR stations (or existing rail stations), and for consistency it is important that the same source be used. Because of the importance of airports to regional transportation planning, the local MPOs may already have derived the travel times from each zone to each of the airports and therefore may be able to supply these times directly.

The accuracy and suitability of the times derived from these networks will depend on several factors, the most important being the level of geographic detail at which the local network is represented and the ability of the network to represent performance of the roadway system under the range of actual traffic conditions. The level of geographic detail is a function of the number of zones into which the region is divided -the finer this geographic detail the more closely the network will resemble the actual roadway system, and therefore the more accurately it will represent point-to-point distances and travel times.

MPOs will often have network files that represent peak and off-peak conditions in base and forecast years. The reasonableness of the data in these files should be tested by examining travel times for a sample of zone pairs and the average speeds implied by these times and the corresponding distances. In addition, base year times should be spot checked against the times given by trip planning websites such as Google Maps, Mapquest, and other sources that use empirical data.

Trip planning websites can also be very useful sources for access/egress time data in cases where detailed networks are not available (often the case for small metropolitan areas) or where resources do not allow the collection and processing of network data (such as in preliminary studies). Care should be taken when using these data, however, as they do not provide information on how the times are calculated, and short of actually driving the routes themselves it is not possible to independently verify the reported travel times. The reasonableness of the travel times should be tested by computing the implied average travel speeds for a sample of the zone pairs.

Notwithstanding these issues, however, these sites do provide this information free of charge, are simple to use, and are well known both within and outside the transportation planning profession, so the use of these sources would be understood even by a nontechnical audience. They also have the advantage in some cases that the travel times may conform well to the perceived travel times of the local population which is the basis for their travel decisions.

### ***Air fares***

The DB1B database (discussed above) is the preferred source of airline fare information for US carriers. Indeed, it is the source of the official air fare reports published regularly by the US Department of Transportation (a summary of average air fares by airport and the national level Air Travel Price index). There are nevertheless several issues that must be noted when using air fare information from this database, as follows:

- ***Small carriers.*** The 10% coupon sample includes only large certificated air carriers, and therefore will not include fares from airlines operating aircraft with no more than 60 seats. This could be an issue in corridors that include air service to smaller cities, which are more likely to be served by regional carriers, some of which may not be large enough to report data to DB1B.

- *Foreign flag carriers.* The DB1B database likewise includes only US carriers. This should be less of an issue in HSR studies, as under federal law foreign flag carriers are not allowed to transport passengers locally between two points within the US.
- *Airline fees.* The fares in DB1B include all taxes and government imposed fees, but they do not include any fees assessed by the airlines or revenues collected by the airlines for onboard services. This has become increasingly important given the rising prevalence of fees for checked baggage, inflight entertainment systems, and food for purchase options.
- *Service class.* The DB1B data includes information on service class, but not trip purpose, so the data cannot be used directly to estimate the average fares paid by business vs. non-business travelers.

Overall average fares (“yields”) are available from the DB1B database. Because this database is compiled from a 10% sample of actual flight coupons, the average fares derived from it will indicate the average of the prices actually paid on a route, and the average can be taken across all fare classes and for a given period of time (such as one a one-year period) to eliminate any seasonality effects. Moreover, the database identifies tickets purchased using frequent flyer miles or other “non-revenue passengers” (such as airline employees), and these tickets can be removed from the average if desired. This database also represents the true origin and destination of each traveler, and can therefore be used to isolate the average fares on a route of only those passengers that are “local” to that airport pair (that is, one can exclude the fares of passengers connecting to or from other flights on one end or the other).

In those cases where an HSR corridor is served by small carriers not reporting to DB1B, fares for a given airport-pair market are readily available on airline or travel websites, and can also be obtained by contacting airline reservation centers or travel agents. The fare information from these sources is available free of charge, and is an accurate representation of the current, actual prices being charged in a given market. But care must be taken to query these sources in a manner that is most likely to provide a good proxy for an average fare. Multiple queries should be made to represent different degrees of advance purchase, and to obtain a representative mix of, for example, first class and coach fares.

There is no uniform source of information on average additional costs due to baggage fees and inflight services in a given market (the airlines may report this information in their public filings but only in summary form for the company as a whole). The rules governing checked baggage fees are available on the airline websites, and some airport ground access surveys will collect information on the number of bags checked by travelers. In corridors where such surveys have been conducted, it may be possible to estimate an average of these fees, which in theory could then be added to the average fares obtained as described above.

#### ***Access/egress costs***

Access/egress costs will include both the cost of traveling to and from the airport as well as parking charges for those travelers that leave a car at the airport during their trip. Trips to and from the airport may generally be made by one of several modes:

- Private car (either parked at the airport or drop-off/pickup);
- Taxi or limo;
- Public transportation (bus, urban rail);
- Express bus (public or private); and
- Shared ride services.

Airport websites will typically have information about parking rates (including short term, long term, economy, valet, etc.). If ground access surveys have been conducted by the airport, these will often have information about parking lot use and duration, travel party size and other characteristics of interest; only large airports do these surveys with any regularity, however. Local transit agency websites will have

information about transit fares, and fares for private express bus companies and shared ride or limo services will also be available on these companies' websites. Local MPO networks generally have the access/egress auto cost information (i.e., tolls, distances, auto operating cost per mile etc.) and transit fare information. Hence, such data can also be available from the MPOs but should always be cross checked with transit agency data directly available from their websites and other relevant sources. City government websites will also have information about taxi fares (some cities have flat rates for airport trips, and/or extra fees).

Preliminary or intermediate HSR studies may simply use a cost per mile for access/egress costs that captures only the approximate cost of driving to and from the airport. For simplicity these studies may also assume parity between parking costs at the airport and at HSR stations. Such assumptions may fail to capture differences in parking rates, and parking costs at large airports can be significant. This approach may also ignore the possibility of airport access/egress by taxi or public transport, the costs of which may be quite different than driving and parking or being dropped off.

As a practical matter, however, the vast majority of air travelers at the vast majority of airports will be traveling to and from the airport by car. In larger cities, many will come by taxi, but even in these cities the share of trips using public transport is typically quite low (less than 10%). As a result even a carefully constructed weighted average access/egress cost that takes into account all of the above modes will be heavily weighted toward the cost of auto. It follows that in many cases the use of auto access/egress costs for simplicity may in fact be a reasonable approximation of actual conditions. Furthermore, if HSR stations are to be located in dense, downtown areas where parking is likely to be relatively costly, it may well be that parking prices would be similar to those at the airport - ultimately this is a policy decision for the HSR operator, and thus this should also be taken into account during ridership and revenue forecasting.

Given these considerations, there can be important *differences* between airport and HSR station access in the access/egress cost data. Thus to the extent that parking rates are materially different, or that good quality transit access is available to the HSR station but not the airport (or vice versa), attention should be focused on capturing these differentials. To the extent that these differences are small, the collection of precise access/egress cost data is of second order importance to HSR proposals. Conversely, if HSR has downtown Central Business District (CBD) stations, their advantage in access/egress time to business travelers over outlying airports will be a key difference, and careful attention should be paid to capturing this differential.

### **Frequency**

Air service frequencies (the number of scheduled flights per day) may be derived from the same sources used to determine line haul travel times described above. The timetables published by the airlines and the schedules published by OAG both contain all of the flights scheduled between a given airport pair on a given date. In addition, the USDOTs T-100 database described above also contains information about scheduled capacity (flights and seats), and can therefore be used to derive service frequencies for an airport pair market.

While it is generally straightforward to derive service frequencies from these sources, there are a few caveats that should be noted, as follows:

- *Code-sharing of flights.* Because of code-sharing or joint marketing agreements, some flights may be listed multiple times, each under a different airline;
- *Days of operation.* Some flights are operated only on certain days of the week, or may operate at different times on certain days; and
- *Seasonality.* In the US, there are generally more flights in summer months than in the winter.

Given these issues, care needs to be taken to ensure that in counting daily frequencies any duplicate flight records are eliminated, and that the computed average presents a reasonable approximation of the average frequency that can be expected over the course of a year (or at least represents a time of year that is not atypical of the amount of service on a given route). In this regard, the T-100 database may be a better source for calculating average daily flight frequencies for an entire year, as the format of the OAG schedules does not lend itself easily to such a calculation.

### **Reliability**

Importantly, the sources of information on line haul travel times described above include only *scheduled* travel times and the discussion above on service frequency concerns only the number of *scheduled* flights. The actual travel times experienced by air travelers may of course be quite different due to flight delays, and flight cancellations will reduce the frequency of service actually provided. On-time performance and the extent of cancellations likewise can vary significantly between airport pairs, carriers, and even individual flights at different times on a given route.

Mode choice models generally do not capture on-time performance directly as an explicit parameter. However, the consistently high schedule reliability of HSR services in Europe and Japan is well documented, and is certainly a characteristic that differentiates HSR from air service. To some extent this difference may be captured in the modal constants, at least to the extent that respondents to the stated preference surveys from which mode choice models are derived are aware of the difference and take it into account in making their choices. In some HSR forecasting studies, *ad hoc* adjustments have been made to the modal constants in an attempt to more explicitly account for HSR's greater on-time performance.

In any case, data on the schedule reliability of the airlines is readily available from the USDOT. These data come in two basic forms, including:

- **On-time performance statistics.** These data represent the percentage of flights arriving within 15 minutes of their scheduled arrival time and the percent of flights cancelled. These are compiled by carrier and by airport, but not by airport pair market; and
- **Delay statistics.** These data represent the average amount of delay experienced by the airlines, and are compiled from data from individual flights. This information is consolidated into the Airline System Performance Metrics (ASPM) database, access to which must be requested from the FAA.

Each of these types of data has its advantages and disadvantages. On-time performance statistics provide one measure of reliability, indicating how often flights are late or cancelled on average. On the other hand, they convey only the likelihood that a flight may be delayed (or cancelled) without giving data on the *distribution* of those delays (the extent of the delays associated with those flights that are late). One therefore cannot derive from these data an expected amount of delay. Conversely, the average delay statistics, while conveying an expected average amount of delay, do not indicate how often the flights are actually late (the same average delay would result from a flight that experiences a five-hour delay one time in ten, or from a flight that is delayed by half an hour ten times in a row, for example).

Given these considerations, the most appropriate measure for on-time performance will likely depend on how this aspect of the level of service is to be incorporated into the mode choice models. In general, because on-time performance is rarely incorporated directly into HSR ridership and revenue forecasting, the choice of these data is of second order importance in the evaluation of HSR proposals.

In summary, different possible sources for air LOS characteristics and the advantages and disadvantages of each are listed below in Table 3-1.

**TABLE 3-1. ADVANTAGES AND DISADVANTAGES OF DATA SOURCES FOR AIR LEVEL OF SERVICE**

Category	Source	Advantages	Disadvantages
Line haul travel time	Official Airline Guide (OAG)	<ul style="list-style-type: none"> <li>Definitive source</li> </ul>	<ul style="list-style-type: none"> <li>Relatively expensive</li> </ul>
	Airline websites	<ul style="list-style-type: none"> <li>Free data</li> </ul>	<ul style="list-style-type: none"> <li>Time consuming</li> </ul>
Access/ egress times	MPO networks	<ul style="list-style-type: none"> <li>Definitive source</li> </ul>	<ul style="list-style-type: none"> <li>Need to validate for reasonableness</li> </ul>
	State DOT networks	<ul style="list-style-type: none"> <li>Definitive source</li> <li>May have data for multiple locations</li> </ul>	<ul style="list-style-type: none"> <li>May not have all locations</li> </ul>
	Mapping websites	<ul style="list-style-type: none"> <li>Free data</li> <li>Easily accessible</li> </ul>	<ul style="list-style-type: none"> <li>No transparency</li> </ul>
Fares	USDOT DB1B database	<ul style="list-style-type: none"> <li>True o/d fares</li> <li>Ability to average over long period</li> <li>Detail by fare class</li> </ul>	<ul style="list-style-type: none"> <li>May not include all airlines</li> <li>Sampling issues</li> </ul>
	Travel websites	<ul style="list-style-type: none"> <li>Multiple airlines</li> <li>Free data</li> </ul>	<ul style="list-style-type: none"> <li>Can't observe actual average fare paid</li> </ul>
	Airline websites	<ul style="list-style-type: none"> <li>Free data</li> <li>Primary source</li> </ul>	<ul style="list-style-type: none"> <li>Can't observe actual average fare paid</li> <li>Need to query multiple sites</li> </ul>
Access/ egress costs	City government websites (taxi fares)	<ul style="list-style-type: none"> <li>May show "official" rates</li> </ul>	<ul style="list-style-type: none"> <li>Not available in all cities</li> </ul>
	Public transport websites	<ul style="list-style-type: none"> <li>Primary source</li> </ul>	<ul style="list-style-type: none"> <li>Time consuming</li> </ul>
	Websites of private bus companies (express bus fares)	<ul style="list-style-type: none"> <li>Primary source</li> </ul>	<ul style="list-style-type: none"> <li>Can't get average prices actually paid</li> </ul>
	Airport ground access surveys	<ul style="list-style-type: none"> <li>Very detailed data</li> <li>Covers all modes</li> </ul>	<ul style="list-style-type: none"> <li>Only available for larger airports</li> <li>Data may not be recent</li> </ul>
	Airport Authorities (parking costs)	<ul style="list-style-type: none"> <li>Primary source</li> </ul>	<ul style="list-style-type: none"> <li>None</li> </ul>
Frequency	Official Airline Guide (OAG)	<ul style="list-style-type: none"> <li>Definitive source</li> <li>All airlines</li> </ul>	<ul style="list-style-type: none"> <li>Relatively expensive</li> </ul>
	Airline websites	<ul style="list-style-type: none"> <li>Free data</li> </ul>	<ul style="list-style-type: none"> <li>Must query multiple sites</li> </ul>
	Travel websites	<ul style="list-style-type: none"> <li>Free data</li> <li>Primary source</li> </ul>	<ul style="list-style-type: none"> <li>Time consuming</li> </ul>
On-time performance	USDOT databases	<ul style="list-style-type: none"> <li>Definitive source</li> <li>Detailed and recent data</li> </ul>	<ul style="list-style-type: none"> <li>Multiple metrics</li> <li>May not be market-specific</li> </ul>

**Rail**

Rail level of service data will need to take into account both the characteristics of any existing intercity rail services in the HSIPR corridor of interest, as well as the characteristics of the proposed HSIPR service. Similar to the air mode, there are six types of rail LOS data that are usually considered - line haul travel times, access/egress times, fares, access/egress costs, frequency and on-time performance.

Travel times and service frequencies for existing rail services should be derived from current schedules for these services, and likewise fare data should indicate the current pricing structure in sufficient detail to distinguish between business and non-business passengers. These data are generally available directly from Amtrak. Access/egress times and costs should be calculated in the same way as was discussed for the air mode in the previous section.

Information on schedule reliability (delays) should likewise be obtained Amtrak, as they are the only source of these data. Care should be taken to interpret and apply the schedule reliability data in a manner that is consistent with the specification of the mode choice models.

Level of service data for proposed HSIPR services is generally produced by project proponents and/or the engineering consultants retained to design the system. For preliminary studies, it may be sufficient to use only data produced by project proponents, understanding that these data are only estimates expected to be significantly refined in later stage studies. Intermediate and final stage studies, however, should use LOS data developed by engineers from analyses of the HSIPR technology, the system alignment, operating plan, and station locations.

Final stage studies should use travel time data derived from simulations of the chosen HSIPR technology and a complete engineering specification of the alignment, including elevation, curvature, grade separation, etc. Service frequencies in final stage studies should be based on a detailed operating plan that is consistent with the performance specifications for the chosen HSIPR technology (acceleration/ deceleration, speed, signaling and train separation, reliability, maintenance requirements and turnaround times, etc.)

Access/egress times should be calculated from data derived from local MPO and/or statewide networks, consistent with that used for airport access times described above. Access/egress costs should likewise represent network-derived distances, HSIPR station parking costs, and where applicable, the likely share of access by taxi, public transport, or other modes other than private auto (though this is a second order consideration since most access will likely be by private car).

Fares for the proposed HSIPR service will necessarily be determined by assumption, but should be based on a reasoned and well-documented pricing policy decision by the project proponents. It is common for multi-modal forecasting models to simplify fare policies and use distance-based averages for each mode; however, such simplifications can have a significant impact on modeled choices when a broad range of fares exist. In final stage proposals, fares for the proposed service should explicitly take into account the tradeoff between fares that maximize revenues, and lower fares that increase ridership and thus the public benefits that justify any necessary public subsidies. This means that the fare policy determination should consider the impact of fares on both revenue generation and the public benefits of the proposed project.

### **Auto**

For preliminary studies, auto travel times can be estimated through actual experience, trip planning software such as Mapquest and Google Maps, and other sources such as real time travel time monitoring websites. For intermediate or final studies, GIS-based networks such as the National Highway Planning Network should be used in combination with the networks maintained by states and local MPOs if available. Figure 3-1 presents an example screen shot of origin to destination travel times calculated from a GIS based representation of the highway network. Travel times should be validated through checks of selected zone or county pairs using other sources, and the reasonableness of implied average speeds should be checked using the travel times and network distances.

FIGURE 3-1. EXAMPLE OF AUTO TRAVEL TIMES CALCULATED FROM GIS-BASED HIGHWAY NETWORK

	19267	30118	33449	44167	50507	57645	63658	73749	75553
19267	-	106.95	167.59	181.07	289.51	291.25	203.32	175.24	262.55
30118	106.86	--	65.31	86.83	187.23	188.97	101.04	72.96	168.65
33449	168.38	66.29	--	58.31	126.14	127.07	71.04	42.95	115.79
44167	181.32	86.37	57.92	--	181.25	153.44	123.53	95.44	95.81
50507	290.92	188.83	128.43	182.25	--	129.41	155.74	145.45	221.56
57645	290.13	188.04	126.48	151.71	125.57	--	175.82	155.13	140.02
63658	202.30	100.21	70.88	122.89	153.86	177.42	--	42.41	180.95
73749	171.10	69.01	39.68	91.69	144.40	158.39	39.38	--	146.30
75553	262.95	168.33	118.14	95.88	218.92	139.17	180.25	145.85	--

For the purpose of demand estimation, per-mile auto costs should be the costs perceived by respondents and used to develop the mode choice models. Travel surveys should also be investigated as possible sources, and where possible any new auto traveler survey data collection conducted as part of the study should obtain data on perceived auto travel costs. Similarly, existing travel surveys (either local, region-wide or state-wide) may also contain relevant information. These perceived out-of-pocket auto costs are much less than the per mile costs estimated by organizations such as the AAA, which include vehicle depreciation and other life cycle costs. Importantly, any estimates of auto costs must also include any applicable tolls and parking charges, which may be a function of the routing and exact destination locations assumed in deriving the travel time estimates.

Finally, mode choice models estimate the choice behavior of individuals, while auto cost estimates generally represent total costs *per vehicle*. These costs must therefore be converted to a per-person basis, which can be accomplished through the application of average group size estimates derived from travel surveys. Wherever possible, these group size estimates should be taken from surveys administered specifically for the HSIPR study or, at the very least, from available travel surveys in the study area.

### Bus

Intercity bus fares, travel times, and service frequencies are generally available from the websites of the private carriers that operate these services. However, care must be taken to ensure that frequencies and travel times are accurately derived from the published schedules, as some routes may involve transfers or operate only on certain days of the week.

Bus fares do not typically have the range of rules and restrictions associated with yield management that is common in the airline industry, and to a lesser extent with Amtrak rail service (discounts are offered for fares purchased well in advance and/or for itineraries involving Saturday night stays, for example). Likewise, buses generally offer only one class of service, and therefore do not have separate business or first class fares. Published fares and fare policies should therefore be examined to determine if a useful distinction can be made between business and non-business fares for the purposes of mode choice modeling.

### Future LOS characteristics

For both clarity and simplicity, HSIPR forecasting studies often assume that forecast year LOS characteristics for the existing modes will be the same as those prevailing in the current or base year. This may sometimes be a reasonable assumption, as it is difficult to predict many years in advance what the nature and extent of competition will be among the now-existing modes, or to say with any certainty what prices will be, or what amount of service will be offered.

At the same time, however, the transportation planning process in major metropolitan areas does provide some sources of information about projected future levels of roadway and airport congestion, and these sources can be used to assess potential differences in certain forecast year LOS characteristics.

Specifically, as has been described, local MPOs and state DOTs often maintain networks for both a base year and forecast years, with the future year network files incorporating both expected future changes to the infrastructure (the addition of new roads or lanes and the like) as well as the impact of expected higher future traffic levels. Each link in the network is coded with its free flow speed and capacity, and the forecast speeds in the future year networks will be calculated based on the traffic and capacity expected then. These future year networks can be used to derive forecast year intercity auto travel times and/or access/egress times, subject again to the types of reasonableness checks prescribed above in the discussion of base year times.

Likewise, major airports often commission detailed forecasts of activity (passengers and operations) as part of their master planning process. These forecasts sometimes include a detailed future year schedule that indicates the number of flights expected to be operated at the airport, by route and carrier. To the extent that these forecasts do provide this level of detail, they can be used to derive air frequencies for the forecast year. However, these forecasts will represent a number of assumptions that may or may not be consistent with the forecasts underlying the HSR study.

Indeed, the likely competitive response of common carrier service providers to the introduction of HSR service is impossible to predict with any certainty in advance. Absent information to the contrary, it is generally assumed that future common carrier LOS characteristics will mirror base year conditions, so the precision of future year level of service characteristics is of second order importance. Accordingly, the impacts of potential changes in future year service frequencies, fare levels or other LOS variable may be most appropriately examined in the context of a sensitivity analysis.

### **Potential impacts of LOS characteristics on HSIPR forecasts**

The relative LOS characteristics (both line haul and access and for both base and forecast years) of HSIPR and the competing modes directly impact the HSIPR's attractiveness and ridership diversion from the competing modes. These LOS characteristics are the main components of the modal attractiveness measures (known as utility functions<sup>3</sup>) that are used in mode choice models (discussed in the next section). Hence, accurate and realistic representation of the base and forecast year LOS characteristics is of paramount importance for accurate HSIPR ridership forecasting.

---

<sup>3</sup> Measure of attractiveness of a given mode, as used in mode choice models, calculated using the LOS characteristics of the modes and other relevant measures (e.g., income of traveler) and model parameters.



## 4 Mode choice modeling

Mode choice modeling predicts the outcome of the decision process by which travelers choose the mode(s) to take from origin to destination. This is a key step in HSIPR ridership and revenue forecasting, as it predicts the fraction of trips that will divert to the HSIPR project from the no-build modes. Different mode choice models are typically developed for the different market segments selected for the study (and in fact the results of the model development may influence the definition of these segments). The fractions output by the mode choice model, for each market segment and OD pair, are combined with the corresponding trip table values to predict the corresponding travel volumes by HSIPR and the other modes.

### Mode choice models

Mode choice models are typically developed via a statistical analysis of the behavior of travelers in different situations that allows the preferences and tradeoffs of travelers to be represented. The statistical analysis may be based on behavior observed in actual travel situations (revealed preference data), or in hypothetical situations presented to travelers in a survey (stated preference data), or both. Mode choice models take into account the service characteristics of the modes being considered, as well as characteristics of the traveler and the trip.

Selection of a proper mode choice model *structure* is important to obtain reliable and credible ridership and revenue estimates. Two types of mode choice model are commonly used in HSIPR demand forecasting - choice models and diversion choice models. Although both these model types are based on similar underlying theory, they are different in the way that they are estimated and applied in practice.

### Choice modeling

In choice models, travelers are assumed to choose from among all feasible mode options (including HSIPR in the build situation), without reference to the travelers' no-build situation mode choice. A choice model simultaneously predicts the share of users traveling on each of the considered modes. Choice models can be applied directly to predict mode shares, or incrementally<sup>4</sup> to predict changes in base situation mode shares; the latter can sometimes reduce the scope for forecasting errors by referencing model outputs to observed mode shares.

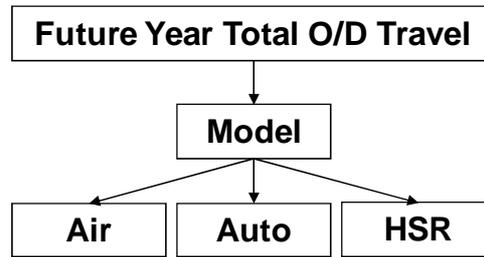
HSIPR studies have used various forms of choice model. Standard practice includes multinomial logit models and nested logit models; more complex forms are mostly used by academics. The advantages of this approach are common to all of these forms, but each form can have specific disadvantages.

*Advantages:* Choice models have been used extensively by the profession. They are constantly being adapted and applied through extensive experience. Model validation is straightforward, by comparing base mode shares with the results of applying the choice model to the base situation. Choice models imply a theoretically rigorous measure of user benefit and system accessibility.

---

<sup>4</sup> This means that the choice model is applied twice: first to a base situation for which actual mode shares have been observed, then to the situation being considered. The model-predicted change in mode share between the two situations is then algebraically added to the observed share, and the result is used to determine the corresponding modal volumes. This approach (also called pivot point) has the advantage of tying the forecasts to actual mode shares.

FIGURE 4-1. MULTINOMIAL CHOICE MODEL EXAMPLE



Multinomial logit models. In a multinomial choice model (MNL), all modes are considered comparably by the traveler (as shown in Figure 4-1). MNL models were used for mode choice calculations in the Northern New England Corridor<sup>5</sup> and Keystone Corridor<sup>6</sup> HSR studies.

*Disadvantages:* MNL models have a property known as independence from irrelevant alternatives (IIA). This property implies that individuals using a new or improved mode will be drawn from other modes in direct proportion to the share of trips made on those other modes. For example, if 80% of the trips between an OD pair are made by auto then, because of the IIA property, the MNL will predict that 80% of the trips on a new HSR mode will be diverted from auto, which is generally unrealistic.

Nested logit models. Probably the most popular form of choice model used in HSIPR studies is the nested or hierarchical logit model. A traveler's mode choice decision is modeled as if it followed a hierarchical decision process involving individual or nests (groups) of alternatives. In Figure 4-2, for example, a traveler decides between auto and a common carrier nest, which in turn includes air and HSR modes. Intercity bus and conventional rail modes, if present, might also be included in this latter choice set or perhaps in another (e.g. low speed common carrier). For a given set of modes, a number of different nesting structures might reasonably be specified, with the selection of a preferred structure from among them based on each specification's intrinsic reasonableness and statistical performance: there is no *a priori* "true" nesting structure. Nested logit models reduce somewhat the issues associated with MNL models: IIA still holds for choices within a given nest, but not for choices in different nests. Nested logit mode choice models have been used in various HSIPR studies including the most recent California Corridor HSR Study<sup>7</sup>, the Southeast Corridor<sup>8</sup>, the Midwest Corridor<sup>9</sup>, the Pacific Northwest Corridor<sup>10</sup> and others.

---

<sup>5</sup> "Ridership Forecasts for the Boston-Montreal High Speed Rail Corridor Feasibility Study", Brief Working Paper, by Cambridge Systematics. [CS Northern New England]

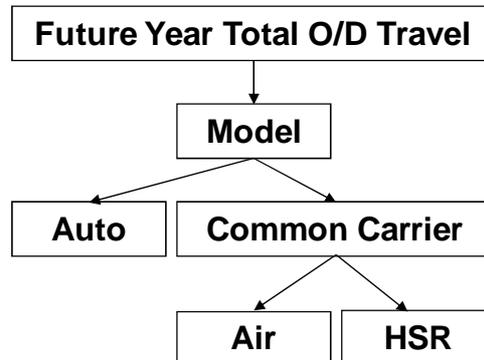
<sup>6</sup> "Market Demand: Pennsylvania High Speed Rail Feasibility Study," prepared for the Pennsylvania High Speed Intercity Rail Passenger Commission, July 1986, by Gannett Fleming [GF Keystone, 86].

<sup>7</sup> "Bay Area/California High-Speed Rail Ridership and Revenue Forecasting Study", Final Draft Report for Metropolitan Transportation Commission and the California High-Speed Rail Authority, July 2007, by Cambridge Systematics, Inc. with Corey Canapary & Galanis, Mark Bradley Research & Consulting, HLB Decision Economics, Inc., SYSTRA Consulting, Inc., and Citilabs. [CS CA Corridor 07]

<sup>8</sup> "Southeast High Speed Rail Market and Demand Study," Final Report, August 1997, by KPMG (now AECOM). [AECOM SE Corridor 97].

<sup>9</sup> "Midwest Regional Rail Initiative, Strategic Assessment and Business Plan", Final Report, August, 1998, by TEMS. [TEMS MWRRRI 98]

FIGURE 4-2. NESTED CHOICE MODEL EXAMPLE



*Disadvantages:* Because of their somewhat complex form, it can be difficult to check the reasonableness of nested logit models (i.e. they lack transparency), particularly those with several levels of nesting. For the same reason, nested logit models can be cumbersome to apply in practice unless appropriate software is available. There is no *a priori* correct nesting structure corresponding to a given set of modes, so specification and choice of a nesting structure is to some extent subjective, yet this choice can significantly affect ridership forecasts.

More complex model forms. More advanced choice models are also available (e.g. mixed logit models); these are significantly more complex in nature and are mostly used in academic research. These have not yet entered standard HSIPR forecasting practice because of their complexity and relative newness.

#### ***Diversion choice modeling***

This approach involves a set of models, each of which predicts the share of trips that will divert from a specific individual base mode to HSIPR. In this sense, the market segmentation used in the models is based in part on base situation mode choices. It considers that travelers' base mode choices reveal information about their valuation of travel time and other attributes, and implicitly incorporates this information in the models that predict their response to a HSIPR alternative. By their nature, diversion choice models have a binary form, as shown for example in Figure 4-3. When designing stated preference (SP) surveys to develop diversion choice models, questionnaires are prepared separately for each group of no-build mode users to obtain information about how they trade off the characteristics of their existing and the project mode. Like choice models, diversion choice models can also be applied directly or incrementally. Diversion choice modeling is also widely used in HSIPR studies, including the California corridor<sup>11</sup>, the Orlando-Tampa corridor<sup>12</sup>, the South Central corridor<sup>13</sup> and others.

---

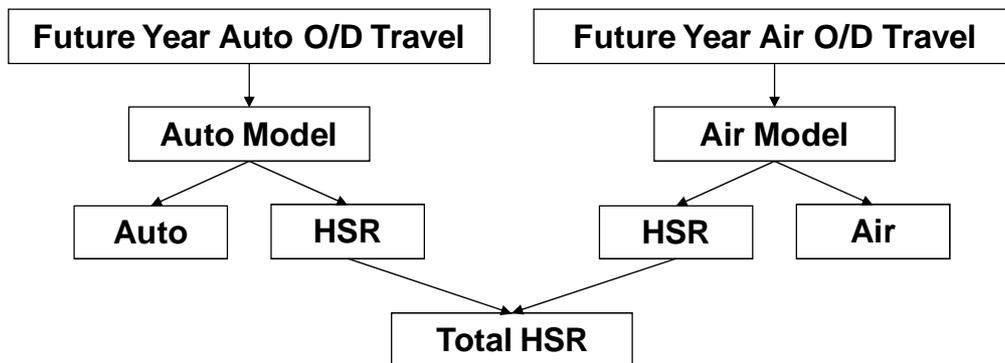
<sup>10</sup> "Pacific Northwest Rail Corridor, Ridership and Revenue Forecasts in support of the Amtrak Cascades Plan for Washington State 2003-2023 Update", Report, July 2003, by AECOM. [AECOM Pacific Northwest 03]

<sup>11</sup> "Independent Ridership and Passenger Revenue Projections for High Speed Rail Alternatives in California", January 2000 and July 1996, by Charles River Associates. [CRA CA Corridor 00 and 96]

<sup>12</sup> "Investment Grade Ridership Study Summary Report," November 20, 2002, by AECOM and Wilbur Smith Associates. [WS Orl-Tampa 02].

<sup>13</sup> "Independent Ridership and Passenger Revenue Projections for the Texas TGV Corporation High Speed Rail System in Texas", September 1993. [CRA South Central 93]

FIGURE 4-3. DIVERSION CHOICE MODEL EXAMPLE



*Advantages:* Diversion choice models are relatively simple to develop, requiring only the statistical estimation of binary logit models; nesting structure selection issues are entirely avoided. Their simple and transparent form makes it easy to check the reasonableness a set of diversion choice models and to efficiently adjust any that are found to produce unrealistic forecasts. Diversion choice models are equally simple to apply in practice, and do not require specialized or custom software.

*Disadvantages:* The binary diversion choice approach is most applicable to situations where a single mode is introduced or improved, with other modes remaining unchanged; more complex situations can be difficult to analyze. Since these models predict diversions, validation requires data on mode shares before and after a change, which may not be available. Estimation of models intended to predict diversions to a new mode must rely heavily on stated preference survey results, and so is subject to the biases of SP data. There is no broadly-accepted user benefit or accessibility measure corresponding to a set of diversion choice models.

**Other approaches**

HSIPR studies in the US have traditionally involved one of the approaches described above, but this is not necessarily the case for studies of existing conventional intercity rail services, where an elasticity based approach may suffice. This approach is used in the UK PDFH, for example<sup>14</sup>. However, useable data on rail elasticities may not exist in the US. Elasticity-based approaches are generally not adequate for a complex forecasting exercise, but they may provide useful cross-checks of forecasts developed by other means.

**Methods for obtaining mode choice data**

The input data required for mode choice modeling include historic and future values of socio-economic variables, information about existing and future travel modes including their level-of-service characteristics and levels of trip making, and data about user behavior obtained from surveys that may include hypothetical tradeoff scenarios involving the proposed rail mode.

**Revealed Preference (RP) data**

Travel volumes by modes and traveler valuations of the attributes that are typically considered in mode choice modeling - fare, line-haul travel time, access and egress times, service frequency, etc. - can be inferred from revealed preference data: the travel choices that people are actually observed to make in the marketplace. The collection of revealed preference data (including mode volumes, mode levels of service and traveler socio-economic and demographic characteristics) is discussed earlier in this report.

---

<sup>14</sup> Passenger Demand Forecasting Handbook.

*Advantages and disadvantages:* RP data is a valuable source of information in any HSIPR study. However, when collected from travel surveys (e.g. mail, telephone or internet), biases may also exist in the respondents' responses (similar to SP surveys), due for example to a desire of respondents to justify their chosen mode. RP data may also not be available in some instances.

### ***Stated Preference (SP) data***

Since true HSR does not exist in the US, it is not possible to use revealed preference data alone to determine how travelers actually behave in response to HSR. To solve this problem, stated preference data have been used in many HSR studies to assess likely traveler responses to a new service. SP data are collected through surveys of study area travelers in which they are asked to choose between hypothetical situations. A number of HSR studies (California Corridor HSR Study<sup>15,16</sup>, the Midwest Corridor<sup>17</sup>, the Florida Corridor<sup>18</sup> among others) have used SP surveys to address the non-existence of HSR in the base situation.

*Advantages and disadvantages:* The major advantage of SP surveys is that they provide information about likely traveler responses to modes or services that do not currently exist, and this information can be used to develop forecasting models for these modes and services.

On the other hand, SP surveys can easily produce unreliable data unless they are designed and administered with great care. For example, people do not necessarily do in reality what they say they will do during a survey. (This is particularly a problem if the new mode is described in a less than dispassionate way, encouraging favorable responses to questions about it.) Similarly, it is easy for a survey respondent to respond favorably to a hypothetical new mode, when in reality it may be more difficult to change habitual behavior. During a survey, respondents may overlook situational constraints such as convenience, bad weather or accessibility issues that affect mode choice decisions in reality, making their responses an opinion statement more than an indication of likely behavior. In some cases respondents may give answers to indicate that they would prefer a new mode because they think that the mode would be good for others, even though they do not intend to use it themselves.

These disadvantages and others can be mitigated to some extent through careful survey design and execution, and also by combining SP and RP data, as discussed further below.

### ***RP and SP surveys***

Both RP and SP data are collected through survey research. The value of such data depends critically on three broad considerations: the design of the sample, the design of the questionnaire, and the detailed procedures and quality control employed in carrying out the survey.

#### *Sample design*

Methods of selecting respondents from the population often depend on the mode of survey administration: that is, the possible means of identifying and soliciting cooperation from respondents often constrains the mode of survey administration, and *vice versa*. SP surveys ask respondents to express their preferences among two or more hypothetical alternatives, each described in terms of its salient attributes (travel times, cost, reliability, etc.). The nature of these questions is such that, for ease of respondent understanding, it is valuable to provide visual cues (in hard copy or on a computer display) summarizing the values of the

---

<sup>15</sup> Ibid., [CRA CA Corridor 00 and 96], page 39, footnote 11.

<sup>16</sup> Ibid., [CS CA Corridor 07], page 38, footnote 7.

<sup>17</sup> Ibid., [TEMS MWRRRI 98], page 38, footnote 9.

<sup>18</sup> Ibid., [WS Orl-Tampa 02], page 39, footnote 12.

attributes under consideration. Only certain modes of administration are conducive to this. If the sampling and interviewing are done by telephone, for example, the visual aids will need to be mailed or faxed to respondents, or made available on-line.

#### Intercept samples

Often the most cost-effective way of identifying and sampling travelers with a common characteristic (traveling in the same corridor or by the same mode, for example) is to intercept them in the course of their travel. In some studies, highway intercept surveys are undertaken and roadside interviews are carried out. There are also proxy intercept methods at natural bottlenecks (toll booths, for example) on highways: distributing self-completion mail-back questionnaires, providing information about a website containing the survey questionnaire, or recording license plate numbers for later contacts. These were discussed earlier. The earlier discussion also emphasized the importance of not conducting highway intercept surveys at roadside rest areas. Roadside rest area survey locations will, in all probability, cause a significant over sampling of non-business travelers, travelers with significantly lower values of time, and travelers making longer trips. These biases will likely show up in the trip table, as well as in lower values of time derived from SP survey questionnaires.

For air travelers, the most common intercept method is to interview representative samples of departing passengers in gate lounge areas.

#### Other approaches to sampling HSR corridor travelers

By contrast with intercept methods, interviewing random samples of the general public provides a much less efficient - and hence significantly more expensive - sampling approach. For screening purposes, respondents would be asked whether they had made a qualifying trip within the corridor within a credible recall period: say, the preceding three or six months. Persons passing the screening questions can be further surveyed, but the incidence of the qualifying behavior within a random sample of households resident in the study corridor is generally not known in advance. It is likely that finding a few hundred qualified respondents would require issuing questionnaires to several thousand households and possibly many more. Contacting and screening this number of households - by telephone or mail, for example - can be expensive.

Sometimes, “short cuts” are used in these situations. First, if private vehicle flows in the corridor are significantly influenced by a relatively few large trip generators and/or attractors (very large employment centers, for example), that information can be used to target the sample (through workplace-based surveys, for example).

Some studies tend to ignore these problems and opt to interview a “convenience sample” of qualified travelers. For instance, they will set up computer terminals in the lobbies of certain buildings with heavy pedestrian flows - typically, public buildings, large office buildings or shopping malls - and seek volunteers to “take a survey.” The key problem here is that the principle of random sampling has been abandoned completely, and one has no statistical understanding of how representative the achieved sample is of all travelers in the population of interest.

Another approach is to undertake a general population survey using members of a pre-recruited consumer panel as described earlier. For local or regional transportation planning purposes, such use has been limited by the geographical density of panel members within the area of interest, so this type of resource has not been used frequently in transportation studies except at the national or state level.

#### Sampling strategy issues

Survey efforts apply a sampling strategy to partition the population of potential respondents into distinct groups based on their characteristics and/or their choices. The strategy determines a target number of surveys for each group, and respondents are selected randomly from within each group. Two contrasting

sampling strategies can be distinguished: exogenous sampling, which is based on respondent attributes such as income or other socio-economic variable; and endogenous (or choice-based) sampling, which is based on respondent behavior such as mode choice.<sup>19</sup> In practice, survey data are often obtained from a mixture of both exogenous and endogenous sampling, for example by oversampling low-income transit users.

The issues associated with model estimation from data collected using various sampling strategies can be quite complex for all but the simplest model forms. The review of this component of an HSIPR study is best left to specialists, who themselves need to ensure that they remain current with evolving academic and applied research in this area.

### Sample size

It is important to determine the sample sizes required to achieve a desired accuracy level for the sample values of the variables of interest. In practice, sample size quotas or targets are often set for market segments such as trip purpose (e.g. commute, non-commute), time of day (e.g. AM peak, PM peak, off peak), actual mode used (e.g. auto, transit) and vehicle occupancy (e.g. drive alone, drive with passengers). The sample size to be used in any specific context needs to be determined from the level of accuracy desired for each market of interest.

SP survey sample sizes tend to be smaller than the sample sizes that are typical for RP household travel surveys. Purely random sampling is rarely adequate, and it is often necessary to set quotas for specific sub-samples to ensure that the obtained data are sufficient data to estimate separate models and/or parameters for key market segments. It is common to aim for a minimum of 150 respondents per SP exercise (a tradeoff question asked as part of the SP survey) per market segment. While it may often be possible to estimate credible, robust mode choice models based on fewer respondents, this rule of thumb provides a margin of safety that has proved valuable in some applications of the method.

### *Questionnaire design*

An SP exercise is defined by the set of product attributes being considered by respondents within a set of tradeoff questions. In any one question, the respondent should not be asked to be holding in his or her mind so many pieces of information that the question becomes overly complex. Usually, the product of the number of choices and the attributes under consideration should not exceed 10 to 15 (e.g. 3 choices and 4 to 5 attributes). If the number of attributes under consideration is large, this can be resolved by dividing the attributes into two or more separate sets or “exercises” (each including at least one common, most important elements - typically travel time or out-of-pocket cost in the travel choice context), administering any one respondent only one set, and multiplying the target sample size by the number of sets to ensure that each attribute receives adequate respondent attention. The resulting exercise-specific datasets are then pooled for the estimation of the mode choice models.

To elicit reliable data, an SP survey must present to respondents a reasonably realistic simulation of an actual choice situation. Choices presented in the survey must involve tradeoffs, so that improvement in one aspect of a modal alternative can only be obtained at the expense of some other aspect. The attribute levels that characterize each alternative should be specified so that (i) no one alternative constitutes an obviously better (dominant) choice over another; and (ii) attribute values are relatively uncorrelated (to avoid statistical difficulties disentangling the respective effects of each). Designed in this way, the results

---

<sup>19</sup> The term *endogenous* is used because the criterion used to define a group represents the same behavior that the survey is intended to collect information about. For example, a sampling strategy may deliberately sample transit users at a higher rate than the population as a whole.

obtained from an SP survey can provide considerable information into the preferences and relative valuation of different modal attributes, while mitigating some of the biases associated with SP data.

#### *Estimation issues*

SP survey respondents are presented multiple exercises in which they are asked to choose between alternatives in hypothetical choice situations. The responses obtained from each respondent may be affected by unobserved respondent characteristics, so should not be treated as independent. Statistical model estimation software usually includes methods to address this issue.

#### **Combined RP/SP data**

Stated Preference (SP) data can effectively complement RP data. The advantage of SP data is that hypothetical choice experiments can accommodate travel choice options that currently do not exist and these experiments can be designed to contain as much variation in each attribute as is thought appropriate. The advantage of RP data is that they provide information on travel choices actually made. By combining RP and SP data, the advantages of each can be obtained while mitigating their respective limitations.

Key advantages of combining RP and SP data are:

- Efficiency: joint estimation of preference (or attribute importance) parameters from all available data;
- Bias correction: helping to compensate for known SP biases; and
- Identification: estimation of preferences for new products or services and for new attributes or attribute levels that are not identifiable from RP data.

Due to the statistical complexities involved in model estimation from combined RP and SP data, this technique is relatively infrequently used in practical applications of travel demand forecasting. Indeed, none of the HSIPR studies reviewed as part of this study except the Southeast Corridor<sup>20</sup> used the joint RP/SP technique. However, this technique is very popular among academic practitioners and is widely used for academic research.

#### **Model variables and specifications**

An individual's choice among alternatives depends on the characteristics of each alternative and also on the socio-economic characteristics of the individual. In a mode choice context, the relative attractiveness (utility) of each mode is therefore usually expressed as a function of a mode-specific preference term (also known as a modal constant), socio-economic (SE) characteristics of the tripmaker and level-of-service (LOS) attributes of each mode.

#### ***Tripmaker-specific variables***

Tripmaker socio-economic (SE) characteristics such as income are included in utility specifications to account for the effect of these characteristics on modal preferences. In some cases, SE variables are added to mode utilities as mode-specific attributes; for example, income may be included in the transit mode utility function in a way that decreases transit utility as traveler income increases. There are other possibilities; for example, travel cost variables may be combined with traveler income (e.g. cost/income) to represent decreasing sensitivity to cost with increasing income.

---

<sup>20</sup> Ibid., [AECOM SE Corridor 97], page 38, footnote 8.

### ***Trip (LOS)-specific variables***

Typical LOS attributes include travel time and cost. The ratios of travel time and cost parameters show the tradeoffs that decision-makers are willing to make, in the form of measures such as the implied value of time (VOT).

It has long been observed that travelers value time savings on some parts of their trip more than time savings on others. The time taken by a typical trip can be divided into several parts: access/egress times, wait time, and in-vehicle time. (Access/egress and wait times do not generally apply to car or bicycle trips, or to walk-only trips, so there is no division into parts of trips by these modes.) A typical HSR trip with transit access and transit egress can be divided into as many as eight parts:

- Walk time from the origin;
- Wait time for the access mode (transit);
- Access mode in-vehicle time;
- Transfer/wait time for the main mode (HSR);
- Main mode (HSR) in-vehicle time;
- Transfer/wait time for the egress mode (transit);
- Egress mode in-vehicle time; and
- Walk time to the destination.

This breakdown applies to public transport trips as well. Each of these trip components may potentially entail a different value of time, with access/egress and transfer/wait times typically having a higher value than in-vehicle time.

### ***Mode-specific constants***

Mode choice model utility functions often include a constant term, the mode-specific constant, that represents the inherent attractiveness of a mode, everything else (e.g. time, cost, and other LOS attributes) equal. Mode-specific constants are considered to represent the average effect of all factors that influence its attractiveness but are not explicitly included in the utility specification. For example, factors such as modal comfort, safety, privacy and reliability may be accounted for in this way. Because they represent a fixed adjustment to a mode's utility that is independent of the magnitude of the other utility components, mode constants can sometimes distort model predictions in applications over a wide range of utility values. In some cases this may be desired: large auto constants can have the effect of discouraging short trips from diverting to rail for example. In others, they may be used to account for "lumpy" effects (e.g. dislike of airport security), using other methods (e.g. differential values of time) to account for otherwise unrepresented modal attributes, particularly ones whose effects are likely to vary with trip duration (e.g. comfort). Mode-specific constants can be translated to equivalent time or monetary values to check their reasonableness. If the mode-specific constant(s) are very large compared to the contribution to utility of time and cost and other LOS attributes, they will strongly influence the predicted mode shares.

## **Specific modeling issues**

### ***Generalized cost***

Generalized travel cost is the sum of the monetary and non-monetary costs of a trip. Monetary (or "out-of-pocket") costs may include fare on a public transport trip, or the perceived costs of vehicle operation (fuel, wear and tear) plus the actual cost of tolls or charges on a car trip. Non-monetary costs refer to the time spent undertaking the trip. Time is converted to monetary values using values of time, which may vary according to traveler income, trip purpose, mode and trip component. Generalized travel time, which is sometimes used, is the equivalent of generalized cost converted to a time value.

## ***Reliability***

In Europe and Asia, HSR services on dedicated track have shown themselves to be highly reliable, and traveler response to this reliability can be an important factor affecting HSR ridership. Reliability can be quantified in a variety of ways, including average delay, average delay due to cancellations, percentage of trips later than a threshold number of minutes, and the standard deviation of delay. Each measure relates to a different aspect of people's perception of reliability. In a mode choice model, the parameter associated with travel time unreliability should reflect its onerousness relative to other travel time components. It is not advisable to incorporate more than measure of unreliability in the utility function, as it can be difficult to disentangle the effect of different measures.

An important issue with the use of HSR reliability measures is the unavailability of such data in the US. Moreover, SP surveys conducted for a new HSR service do not generally ask respondents about reliability, other than by including it as an inherent positive characteristic of the service. Hence, these studies have not been able to estimate parameters for reliability measures other than via the HSR modal constant. However, Amtrak maintains very detailed reliability data for all its trains including the Acela service; in principle, revealed preference data on Acela use can be used to estimate parameters associated with reliability. This may be very useful in modeling the reliability of HSR and passenger rail modes.

## ***HSR service classes***

Passengers' HSR service class choice can be modeled as a two-level choice structure: intercity transportation mode choice first, and HSR service choice second. HSR service choice may consist of decisions about train schedule, train class, and vehicle class. If the LOS characteristics are not widely different, a specific service choice model may not be warranted. Alternatively, HSR service classes can be modeled via attributes within a single HSR utility function, for example by modeling time or cost parameters crossed with HSR service classes, or simply by introducing HSR class-specific constants.

The decision to model HSR service classes is made in light of both the goal of the study and the scale of the expected variation in utility of the different HSR classes. Existing HSIPR studies in the US have not explicitly accounted for HSR service class differences in their models. In countries where HSIPR systems are already in service, ridership by HSR service class can be an important desired output of the demand modeling. For example, the UK PDFH explicitly accounts for different service classes in ridership and revenue forecasting for intercity passenger rail services.

## ***Service frequency***

Many mode choice models express a mode's attractiveness (utility) in terms of waiting time rather than service frequency. The waiting time may be calculated by a simple transformation of service frequency: average headway (the time between services) is computed by dividing the length of the operating day by the service frequency, and average waiting time is estimated as some fraction of the headway. Because the operating day length directly affects the computed waiting time, it is important that this value is consistent between the common carrier modes (air, bus and HSIPR).

It is common in urban transportation modeling to assume that the waiting time is half the headway, since public transportation modes (e.g. buses) operate at relatively high but irregular frequencies, and passengers tend to arrive at stops randomly between services. In the intercity context, however, particularly in the case of regular but less frequently scheduled services, passengers are more likely to time their arrival at a stop or station according to the schedule (with a cushion to allow for possible delays), and so wait for less than half the headway. Mode choice models for such services sometimes incorporate frequency (or some transformation of it) rather than waiting time.

There may also be instances where intercity services are offered at very high frequencies, such as during peak periods in very large markets. In these cases, services may be operated within only a few minutes of

each other, or even at the same time (e.g. when competing airlines serve the same route). In these cases a transformation of frequency to waiting time based on half the headway may underestimate the amount of waiting that would occur in reality. To account for this, some models incorporate damped frequency measures, for which the incremental reduction in waiting time diminishes as more services are added. For example, damped frequency formulations were used in the mode choice model in the Southeast Corridor HSR Study<sup>21</sup> and the Orlando-Tampa HSR Study<sup>22</sup>.

In any of these cases, the most important factors to consider in assessing a mode choice model are (i) that the method chosen to incorporate frequency data is reasonable and defensible for the conditions in the corridor of interest; (ii) that it is applied consistently across the common carrier modes (air, rail, bus, HSIPR); and (iii) that it is applied in a manner that is consistent with the way in which the model's frequency or waiting time parameter was originally estimated.

### **Network effects**

Forecasting the demand for common carrier modes, including HSIPR proposals, requires consideration of the entire door-to-door trip when estimating or compiling LOS data. This means explicitly accounting for any transfers that may be required between the HSIPR service and access/egress modes or other intercity rail (or air or bus) services. It also means that LOS data used in HSIPR proposals should explicitly account for any planned improvements to public transport modes and/or the roadway network that are expected to be made as part of the project.

For routes where transfers are required within the intercity portion of the trip (e.g. trips involving more than one HSIPR corridor, say a trip from St. Louis to Detroit via St. Louis and Chicago, and Chicago and Detroit), travel times should represent the total origin-destination time, including the time between legs, as well some estimate of a *transfer penalty*. This transfer penalty (typically expressed in terms of time) is over and above any actual travel or connecting time, as transfers are often the most onerous aspect of a trip, something that travelers would like to avoid if possible. Transfer penalties between intercity rail lines can range from at least a half hour to over an hour. Service frequencies for these routes should also take into account realistic connection possibilities at the intermediate points for through travel. Fares should be based on the prices offered by the operators for the entire intercity trip.

Likewise, LOS characteristics of common carrier modes should explicitly incorporate the waiting time required for access/egress modes, as well as appropriate penalties for transfers to and from these modes. The waiting times for the access/egress modes will in turn depend on their frequencies and how their schedules are coordinated with that of the intercity mode.

### **Mode choice model results and statistics**

An HSIPR forecasting study report should present the mode choice model structures, the estimated utility function coefficient values, and any additional assumptions or procedures needed to estimate or apply the models. Reports should also present a number of other statistical measures, including the standard errors and *t* statistics<sup>23</sup> of the coefficient value estimates. These can be useful for quick evaluation of the statistical validity of the mode choice models. However, detailed review of the model development process and results should be done by specialists in the field.

---

<sup>21</sup> Ibid., [AECOM SE Corridor 97], page 40, footnote 8.

<sup>22</sup> Ibid., [WS OrL-Tampa 02], page 42, footnote 12.

<sup>23</sup> *t* statistics are calculated by dividing the model parameters by their corresponding standard errors.

### **Potential impacts of mode choice modeling on HSIPR forecasts**

Mode choice models are the tools that calculate ridership on HSIPR vs. competing modes. Hence, any shortcomings of these models and their various components will directly impact the quality and accuracy of the HSIPR forecasts. Mode choice models are characterized by their structure and utility functions.

Selection of a proper mode choice model structure is important to obtain reliable and credible ridership and revenue estimates. For example, the nesting structure of nested logit models can significantly affect rail ridership and revenue forecasts. HSIPR is in direct competition with other common carrier modes (air, conventional rail and bus) and hence would be expected to draw the majority of its ridership from them. Placing the HSIPR mode in the same nest as the auto mode could result in unrealistically high auto diversions to HSIPR.

Utility function specifications and estimated coefficient values must also be reasonable and defensible. Mode-specific constants are among the estimated values, representing a fixed adjustment to a mode's utility that is independent of the magnitude of the other utility components. Similarly, model parameters for LOS attributes such as travel time and cost dictate the sensitivity of the mode choice models (and hence HSIPR ridership forecasts) to these LOS characteristics.

Revealed preference (RP) and stated preference (SP) travel surveys are used to gather behavioral data needed for mode choice model estimation. Great care is needed to prepare and conduct these surveys, as biased responses by respondents may lead to erroneous mode choice models which will ultimately lead to erroneous ridership and revenue forecasts.

## 5 Other modeling steps in ridership and revenue forecasting

Once total diversions to a HSIPR project are calculated using mode choice models, other modeling steps may be undertaken to complete the ridership and revenue forecasts. These include estimation of induced HSIPR travel, the assignment and feedback process, calculation of fare revenue, and analyses of the sensitivity of ridership and revenue forecasts to input values and model assumptions.

### Induced travel

An HSIPR project will improve the overall level of service for intercity travel within a given corridor. This improvement will make conditions more favorable for travel and decrease its disutility. Trips will therefore be taken on HSIPR that might not otherwise have been made using any of the current modes. These new trips are commonly referred to as induced trips.

Induced HSIPR trips can be calculated as:

$$\text{Induced Travel} = \text{Total Travel}_{\text{Build}} - \text{Total Travel}_{\text{No Build}}$$

It is necessary to distinguish between two very different sources of induced travel:

- People decide to not make a trip when the disutility of travel (resulting from its time, cost, inconvenience, etc.) is greater than the benefit that they derive from the trip. A transportation system improvement decreases the disutility of travel, so when people re-assess their former decision to not make a trip, some may find that the trip has now become worthwhile and decide to make it.
- Over time, the mobility and accessibility changes brought about by a transportation system improvement will produce changes in the type, intensity and location of land uses and economic activities in the improvement's impact area. The transportation improvement will affect the socio-economic system. Increased population and economic activity will lead to increased travel.

Generally, HSIPR studies consider the former form of induced travel. It is usually beyond the scope of a ridership forecasting study to predict the land use and economic changes that might result from the presence of HSIPR service, and attempts to do so may risk double counting the project benefits.

### *Typical range of induced demand magnitudes*

Table 5-1 shows the range of induced demand percentages estimated in various US HSIPR studies. An upper limit on induced travel of approximately 10% of total HSR trips is widely accepted for proposed HSR systems in the US.

Induced demand percentage estimates also vary in HSR studies outside US. Moreover, induced demand information is also available for some HSR services that are currently in operation. Induced demand magnitudes for actual HSR systems and HSR studies outside US are shown in Table 5-2. The relatively high European numbers may be due to generally higher levels of automobile congestion there, as well as specific factors such as the inconvenience of crossing the English Channel by car and ferry.

**TABLE 5-1. INDUCED DEMAND PERCENTAGES IN US HSIPR STUDIES**

Study	Type	Year	Induced Demand % (approximate)
Orlando-Tampa Investment Grade Ridership Study <sup>24</sup>	Forecast	2002	7% (AECOM) 4% (WSA)
California HSR Study by Cambridge Systematics <sup>25</sup>	Forecast	2007	2%
California HSR Study by Charles River Associates <sup>26</sup>	Forecast	1996, 2000	6%
Midwest Regional Rail Initiative <sup>27</sup>	Forecast	1998	7%
Boston to Montreal HSR <sup>28</sup>	Forecast	2003	0.4% - 3.9%

**TABLE 5-2. INDUCED DEMAND PERCENTAGES IN INTERNATIONAL HSR SYSTEMS AND STUDIES**

Project	Flow	Type	Year	Induced Demand %
New Lines	London-Birmingham	Forecast	2030	18%
New Lines	London-Manchester	Forecast	2030	23%
New Lines	London-Edinburgh	Forecast	2030	67%
LGV	Paris-Lyon	Observed	1985	15%
Eurostar	London-Paris	Observed	1995	27%
Eurostar HS1	London-Paris	Observed	2008	6% <sup>29</sup>

**Modeling approaches**

The attributes of a new mode that affect its attractiveness to existing travelers will also likely influence the amount of new travel that it induces. A new mode that is able to capture a large portion of an existing market is very likely to induce additional trips that are not currently being made. Alternatively, if it attracts only a small portion of existing trips, it is unlikely that much induced demand will materialize. Because of this relationship between induced travel and modal attractiveness, the method used to forecast induced travel should be generally consistent with the method of predicting mode choice. To ensure this, methods used to calculate induced demand frequently refer to the utility functions from the mode choice models.

In this approach, LOS variables from the appropriate mode choice models may be used to calculate the change in total travel utility (attractiveness of the HSIPR mode or lack thereof) resulting from the HSIPR project. For each OD pair, this travel utility change is used to estimate the corresponding induced demand, using for example an elasticity of trip making with respect to travel utility. Since it was the HSIPR project that produced the change in travel utility, all of the additional trips are assumed to be made by HSIPR.

<sup>24</sup> Ibid., [WS Orl-Tampa 02], page 39, footnote 12.

<sup>25</sup> Ibid., [CS CA Corridor 07], page 38, footnote 7.

<sup>26</sup> Ibid., [CRA CA Corridor 00 and 96], page 39, footnote 11.

<sup>27</sup> Ibid., [TEMS MWRRRI 98], page 38, footnote 9.

<sup>28</sup> Ibid., [CS Northern New England], page 38, footnote 5.

<sup>29</sup> Incremental improvement in service.

The elasticity is generally estimated from data obtained from responses to questions in SP surveys about induced demand or from direct demand models. In the absence of an elasticity estimate derived from local data, studies may assume a value. Moreover, some studies simply assume an induced demand percentage without reference to LOS or utility changes and corresponding elasticities. As before, the induced demand percentage should be modest, and generally not exceed 10% of the HSR trips calculated from the mode choice models.

### **Issues**

Induced demand is an important component of the ridership and revenue forecasts and should be included in HSIPR studies. The travel induced by a new mode can be expected to be tied to its attractiveness relative to existing modes. This suggests that the methodology for forecasting induced travel should be consistent with the models for forecasting mode choice. Induced demand should be calculated at the OD pair level to properly account for the impact of the HSIPR service on individual OD pairs. Aggregate induced demand percentages (calculated as a percentage of the total HSIPR demand and not of HSIPR demand at the individual OD level) should generally be avoided. To the extent possible, values (e.g. elasticities) used for the induced demand calculation should be statistically estimated and not simply asserted. The induced demand component should not generally constitute a major portion of the total HSIPR ridership and revenue forecasts.

### **Total HSIPR ridership**

For each OD pair, the sum of HSIPR ridership forecast by the mode choice and the induced travel models is the total ridership. This can be presented in the form of an HSIPR trip table, summarized in a variety of ways (total system ridership or passenger-miles, etc.) or post-processed to compute additional variables of interest.

### **Assignment and feedback**

In general, *assignment* refers to the component of the travel forecasting process in which OD trips are routed over specific paths through the network based on an assumed behavioral principle (e.g. travelers choose the path having minimum generalized cost). When congestion is present, assignment generally requires an iterative approach because the paths on which OD trips are first routed (“loaded”) may no longer have the minimum cost due to the congestion on them produced by the loaded trips. Assignment tends to spread out trips over available paths.

Because of the importance of congestion effects, and the complexity of network structure, travel forecasting on *highway* networks usually requires an assignment model of the type described above. Rail systems, on the other hand, generally have a much simpler structure and congestion effects on them may be much less pronounced. Accordingly, assignment models for HSR, if used at all, tend to limit themselves to a simple bookkeeping of the trips that use the available lines through the rail network. The one exception to this - which is rarely used in HSR studies, although it is not uncommon in more general transit studies - concerns a situation where multiple lines with similar service characteristics are available to serve a particular origin-destination movement; a particular trip, for example, might have available an infrequent express service or a frequent local service, with comparable total times and costs. In this case, transit assignment splits the OD trips among the available comparable services. Situations such as this may become more common as more complex service plans are investigated. When the available alternatives have very different service characteristics (e.g. a “premium” and “conventional” service with different times, fares, and service qualities), it is usually more appropriate to use a mode choice model to predict the ridership on each.

When forecasts predict substantial diversions of automobile travel to rail, it may be appropriate to consider the resulting highway decongestion impacts. A highway assignment model can be used for this purpose, although there may be simpler yet acceptably approximate methods to achieve the same end. If, however, a highway network model has already been developed (for example to prepare highway level of service data), then application of the same network model to predict decongestion impacts is a natural extension. It should be noted that instances where automobile diversions to HSIPR result in significant highway congestion relief are expected to be relatively rare. This is because highway congestion is normally concentrated in urban areas; the proportion of intercity trips on congested highways is small, and of auto trips diverted to HSIPR is smaller still.

Model *feedback* refers to the process of iterating between distinct model steps in an attempt to achieve consistent equilibrium conditions across all modeling steps. For example, rail forecasts may be based on an assumed highway level of service. If it appears that rail will divert automobile travel in amounts sufficient to improve highway congestion levels, then a feedback process would use the improved highway times and costs as next iteration inputs to the rail ridership model. This may result in fewer diversions to rail and somewhat higher highway congestion levels, which could again be in a next iteration. The feedback process would typically continue until some convergence criterion (e.g. small changes in ridership or congestion levels from one iteration to the next) is met.

Feedback is rarely performed in HSIPR forecasting because of the typically small decongestion impacts of highway diversions to rail.

### **Revenue calculation**

As noted in Section 1, the forecasting of HSIPR fare revenues is essentially a simple calculation that uses, for each OD pair and fare class, the corresponding total HSIPR ridership forecast and fare level. The accumulation of each of these revenue components across all OD pairs and fare classes gives the total fare revenue generated by the HSIPR project. This may need to be converted from its value for the model time period (e.g. a representative day or part of day) to an equivalent annual revenue using appropriate annualization factors, which may be obtained from analysis of data on travel volume variations over the year, or more simply by adopting values from comparable projects.

Because HSIPR ridership varies with fare in a way that depends on the modal competition situation and traveler characteristics of each OD pair, prediction of the total revenue impacts of a fare change is best done using the full HSIPR ridership forecasting model. By applying the model to a variety of alternative fare structures and levels, it is possible to identify, at least approximately, revenue-maximizing fares; this is frequently done as part of intermediate and final stage studies.

Because the curve of total fare revenue vs. fare level is often relatively flat near its maximum, it is sometimes argued that fares can be set somewhat lower than their revenue maximizing values in order to increase the public benefits of HSIPR through higher ridership, without adversely impacting the generated fare revenues. Investigation of this question for a particular HSIPR project requires the combined efforts of the ridership forecasting and public benefits assessment teams.

### **Sensitivity analysis**

All HSIPR ridership and revenue forecasting studies should include an analysis of the sensitivity of forecast results to key input values and model assumptions including fare, running time, service frequency, station locations and assumptions about socio-economic and travel growth in forecast years. The effort devoted to this analysis should increase with the level of detail of the study. A significant fraction of the total effort in a final stage study will often be spent on sensitivity analysis.

Sensitivity analysis is conducted by varying, more or less systematically, selected forecasting model inputs, parameters or assumptions (e.g. inflation rate or fuel cost) around their “standard” value, running the model, and examining the variation in outputs. Sometimes the term is used to imply that only one value is changed at a time, although more generally multiple values might be changed in each model run. Because of interactions between model components, the effect of a change in one variable can in general be expected to depend on the level of other variables.

Specific values used in sensitivity analyses may be chosen through systematic perturbations of the standard values (e.g. fixed percentage changes +20%, +10%, -10%, -20% to input values) or alternatively selected via a more deliberative process (by soliciting opinions from experts and other sources, also known as the Delphi method). Sensitivity analysis also requires that the probability of each alternative input or parameter value be assessed, in order to compile the corresponding probability distribution of output values. When multiple variables are changed simultaneously, the assignment of a probability to the set of values should take account of possible correlations between the different inputs and/or parameters, and not automatically assume that they are independent.

The uses of sensitivity analysis are many, ranging from overall assessment of the reliability of forecasting results, through identification of critical factors that should be monitored and influenced, to preparation of a project financing plan.



## 6 Special topics

This chapter briefly discusses some special topics that may need to be taken into consideration in HSIPR ridership and revenue forecasting.

### Interface with urban (MPO) models

Existing and proposed HSIPR services generally have major urban or activity centers as termini as well as intermediate stopping points. Hence, even though most of HSIPR trips are inter-urban, these trips frequently involve urban access/egress components to/from the HSIPR stations. Proper ridership and revenue forecasts must adequately represent the urban components of HSIPR trips. Moreover, when a HSIPR project includes multiple stations in a metropolitan area, the service may be available for urban (e.g. commute) trip making. Similarly, HSIPR service with an airport station may be viable for urban airport access. Note that serving urban travelers is not the best use of high-speed intercity trains, and that urban service should not reduce the attractiveness of HSIPR to its higher fare paying intercity passengers, for example by stopping at many urban stations.

In order to accurately model the urban component of HSIPR ridership, studies often use urban travel demand models maintained by the local MPOs. As intercity travel demand models normally do not include urban trips, the separate use of the MPO models for this purpose is more accurate and so is particularly appropriate for final stage studies. Such applications typically require some modification to the MPO model to incorporate the HSIPR mode.

Intermediate and preliminary stage studies usually avoid working with urban models because the effort involved might exhaust the available budget and time. In these less-detailed studies, urban access/egress characteristics can be calculated from readily available sources (e.g. trip planning software) and more aggregate and high level estimates of the local travel market can be overlaid on the inter-urban forecasts. These estimates can be based on (i) the potential population catchments around the stations; (ii) the level of commuting in these catchments; and (iii) recent relevant forecasts of the potential share of this market that the HSIPR service or alignment might capture.

### *Issues*

It is acceptable to use both intercity and MPO models in the same study as long as they are used in a consistent manner. For example, if access/egress and urban area auto travel characteristics are both obtained from the MPO models, they must be treated consistently when used as inputs to the intercity models. Also, care must be taken not to double count any market segment in the models. For example, if the local travel market is modeled using the MPO model, it should not be included as part of the intercity model.

*Advantages and disadvantages:* MPO models represent urban areas in great detail and hence are suitable for situations where it is important to have such detail (e.g. for analyzing detailed station location alternatives). Due to the detailed representation of the urban modes, LOS attribute values calculated from MPO models are also more precise and accurate. However, use of MPO models is resource- and time-consuming, and the models must be modified to incorporate the HSIPR mode.

### Air/rail complementarity

#### *Possible competitive responses of the airlines to the introduction of HSIPR service*

An HSR project may include a station at a major airport that serves as a hub for one or more major airlines. These hub airports may have hundreds of flights per day, serving a very large number of markets ranging

from smaller cities relatively close to the hub airport to international destinations. The location of a HSR station at a hub airport therefore raises the possibility that the HSR system could provide a direct connection to this large array of air services, greatly expanding its intercity reach. This is especially true for HSR connections between a major hub airport and a smaller airport, where the HSR service may have the potential to attract air passengers who fly from the smaller airport to the hub to connect to their final destinations. A new HSR service may also connect major airports with significant true origin-destination air travel between them. The advent of the HSR service may potentially divert a major portion of this air market. As such, the HSR service may have the potential to pose a significant competitive threat to the airlines operating in the markets described above.

As a practical matter, even at final stage studies, the airlines' response to the HSIPR project will likely be unknown. Moreover, the range of possible airline responses is quite broad: conceivably, the airlines could react to the advent of an HSR service in one or more of the following ways:

- Outright opposition, including legal action;
- Strategic competitive response, including increased frequency and reduced fares;
- Limited competitive response, but without cooperation;
- Friendly competition with cooperation through codesharing/joint marketing;
- Friendly competition with cooperation through codesharing/joint marketing and integrated baggage handling; and
- Cooperation through code sharing/joint marketing, integrated baggage handling, and the elimination of competing air services on routes served by HSR, with the HSR itself operated by the airline (this was proposed for a Texas TGV project in the late 1980s).

The responses of airlines to an HSR project are rarely analyzed in detail because of the uncertainties involved in predicting these responses. It is often assumed that the air service status quo in the markets under consideration will be maintained, and the ridership and revenue forecasting task is performed accordingly. In later stage studies, a scenario or what if analysis may be used to determine the importance to HSR ridership and revenue of airline cooperation. Scenarios should be constructed to cover the range of possible competitive responses described above, and each of the scenarios modeled to determine their ridership and revenue impact.

#### ***Use of HSR as an airport access mode***

Airports are typically located at the periphery of major metropolitan areas, in contrast to many HSR proposals to locate a station in the Central Business District. Many HSR proposals also include additional stations serving heavily populated suburban areas. The inclusion of an airport HSR station therefore creates the possibility of using the HSR system as an airport access mode, with travelers using the HSR service to reach the airport from downtown or suburban stations within the same metropolitan area. It is important that a HSR proposal specify in detail how the system would specifically serve airport access trips.

Airport access is a complex and relatively well-studied issue in transportation planning. As was the case for intra-urban trips, use of HSR for airport access trips should be modeled separately from intercity trips, because the factors and behavior that affect mode choice decisions in the two situations are quite different.

There are important tradeoffs inherent in the use of an intercity HSR system for airport access within a metropolitan area, and it is likewise important that the modeling of HSR ridership and revenue be adequately sensitive to these tradeoffs. Airport access passengers will pay lower fares than intercity passengers. Unless there is spare capacity, these lower fare airport access passengers may displace higher fare paying intercity passengers. Additional capacity could be added to handle the additional passengers, but the revenue gains may not outweigh the additional costs.

In addition, a large volume of airport access passengers boarding at downtown or suburban stations may require much longer dwell times at these stations, increasing the total travel time for intercity HSR trips passing through these stations. These longer travel times will in turn reduce ridership from higher fare paying intercity passengers.

Finally, the location of the HSR station at the airport may have a significant impact on its use for both intercity trips as well as an airport access mode. For intercity trips, airports are generally located in low density suburban areas rather than in the middle of major suburban activity centers and residential areas which have the potential to generate high volumes of intercity travel on HSR. This must be compared to the potential of the HSR line to divert air trips with an airport station. If the airport station is located at a significant distance from the airport's passenger terminals and thus requires an additional transfer, this may make the HSR service less competitive for airport access than alternative modes. At the same time, however, a more convenient station location may have very significant cost implications, and these costs may not be offset by the additional revenue generated by the more favorable location.

#### *Connections between intercity high speed rail and air services*

As with the use of HSR as an airport access mode within a metropolitan area, a number of important issues must be addressed when assessing the ability of a proposed HSR system to serve connecting air passengers (when the HSR service is connecting two airports as mentioned above). These issues include the following:

- Location of the HSR station;
- Baggage handling;
- Scheduling;
- Marketing and distribution; and
- Pricing and revenue sharing.

A summary of the implications of these issues for ridership and revenue forecasts modeling is provided in Table 6-1. The estimation of HSR ridership and revenue derived from the capture of connecting air passengers must be modeled separately from the other potential sources of patronage because the mode choice decision for *intercity* trips that require a connection is generally made jointly rather than incrementally. Sophisticated models are required to choose between HSR and air connecting alternatives including connections through other hubs.

#### *Actual experience in Europe*

Another potential aid to the evaluation of HSR proposals that include a station at a major hub airport is the experience of HSR-air connections in Europe. While many European airports have some kind of rail access, only two airports, Frankfurt and Paris-Charles De Gaulle, have direct connections to their respective country's HSR networks.

Table 6-2 summarizes the main characteristics of these two examples of HSR-air connections. The table shows that both connections feature schedule coordination and airline codesharing, though only Frankfurt has integrated baggage handling facilities. The table also shows that HSR ridership is substantially higher at Paris-Charles de Gaulle, likely the result of the much larger number of cities served and higher service frequency compared to the service at Frankfurt Airport.

**TABLE 6-1. SUMMARY OF IMPORTANT FACTORS IN ASSESSING HSR-AIR CONNECTIONS**

Category	Factors to be detailed in HSR proposal	Implications for modeling
Location of HSR station	<ul style="list-style-type: none"> <li>In the airport terminal vs. outside the terminal</li> <li>Characteristics of the air-rail transfer</li> </ul>	<ul style="list-style-type: none"> <li>Specification of transfer penalty</li> <li>Specification of waiting time and total trip time</li> </ul>
Baggage handling	<ul style="list-style-type: none"> <li>Integrated through baggage handling</li> <li>Security procedures</li> </ul>	<ul style="list-style-type: none"> <li>Specification of transfer penalty</li> <li>Specification of waiting time and total trip time</li> </ul>
Schedule coordination	<ul style="list-style-type: none"> <li>HSR frequency</li> <li>Coordination with hub airport's connecting banks</li> <li>Possible replacement of existing short-haul air services</li> </ul>	<ul style="list-style-type: none"> <li>Specification of HSR frequency</li> <li>Specification of waiting time and total trip time</li> </ul>
Marketing and distribution	<ul style="list-style-type: none"> <li>Sale of single through tickets for combined air-rail trip</li> <li>Codesharing/joint marketing agreements</li> <li>Distribution channels</li> </ul>	<ul style="list-style-type: none"> <li>Simulation of CRS and other website screen presence</li> </ul>
Pricing and revenue sharing	<ul style="list-style-type: none"> <li>Responsibility for pricing</li> <li>Portion of through ticket fare revenue received by HSR operator</li> </ul>	<ul style="list-style-type: none"> <li>Calculation of HSR revenue</li> </ul>

**TABLE 6-2. SUMMARY OF EXISTING HSR-AIR CONNECTIONS IN EUROPE**

	Frankfurt	Paris
Airport	Frankfurt Airport (FRA)	Paris-Charles de Gaulle (CDG)
Distance from city center	8 miles	17 miles
Rail station location	Beneath Terminal 1 Concourse B	Beneath Terminal 2
Cities served by HSR	<ul style="list-style-type: none"> <li>Cologne</li> <li>Siegburg/Bonn</li> <li>Stuttgart</li> </ul>	<p>NORTH:</p> <ul style="list-style-type: none"> <li>Lille</li> <li>St. Pierre des Corps/ Tours</li> </ul> <p>EAST:</p> <ul style="list-style-type: none"> <li>Champagne</li> <li>Ardenne</li> <li>Lorraine</li> <li>Strasbourg</li> </ul> <p>SOUTH:</p> <ul style="list-style-type: none"> <li>Lyon</li> <li>Valence</li> <li>Avignon</li> <li>Aix-en-Provence</li> <li>Marseille</li> <li>Toulon</li> <li>Montpellier</li> <li>Nîmes</li> </ul> <p>WEST:</p> <ul style="list-style-type: none"> <li>Nantes</li> <li>Rennes</li> <li>Angers St-Laud</li> <li>Le Mans</li> <li>Poitiers</li> <li>Bordeaux St-Jean</li> </ul>
HSR frequency	<ul style="list-style-type: none"> <li>Cologne &amp; Siegburg/Bonn: 12 trains/day</li> <li>Stuttgart: 5 trains/day</li> </ul>	<ul style="list-style-type: none"> <li>North: 31 trains /day</li> <li>South: 16 trains/day</li> <li>East: 3 trains/day</li> <li>West: 12 trains/day</li> </ul>
Integrated baggage handling	Yes	No
Schedule coordination	With Lufthansa schedules	With key international arrival/departure times
Through tickets available?	Yes	Yes
Codesharing with airlines?	Lufthansa	Air France
Annual HSR ridership	170,000	3,000,000

Source: Frankfurt Airport and Aéroports de Paris.

### ***Assessment of HSR proposals serving airports***

A guide for assessing HSR proposals that include an airport station to be used for airport access purposes and/or for capturing traffic from connecting air passengers can be found in Table 6-3. The table shows the types of information that should be developed for each HSR study stage. For preliminary studies it may be sufficient to simply indicate the types of services that are envisioned, while final stage studies will need to develop the detailed data describing the characteristics of the HSR-air connection, and model the ridership and revenue for a set of detailed scenarios describing possible service features and levels of airline cooperation.

**TABLE 6-3. SPECIFICATION OF HSR-AIR CONNECTION BY STUDY LEVEL**

Issue	Preliminary	Intermediate	Final
Location of HSR station	<ul style="list-style-type: none"> <li>Statement as to whether airport station is desired</li> <li>Notional feasibility of on-airport location</li> </ul>	<ul style="list-style-type: none"> <li>Specific station location alternatives</li> <li>Implications for alignments and travel times</li> </ul>	<ul style="list-style-type: none"> <li>Exact station location</li> <li>Nature and extent of air-rail transfer</li> <li>Detailed level of service data including transfer times and penalties</li> </ul>
Baggage handling	<ul style="list-style-type: none"> <li>None needed</li> </ul>	<ul style="list-style-type: none"> <li>Description of pros/cons of integrated service</li> <li>Notional estimate of cost and revenue implications</li> </ul>	<ul style="list-style-type: none"> <li>Detailed description of baggage handling features, if any</li> <li>Plan for addressing security issues</li> <li>Ridership and revenue estimates with and without service</li> </ul>
Scheduling	<ul style="list-style-type: none"> <li>Notional estimate of HSR frequency/ headway</li> </ul>	<ul style="list-style-type: none"> <li>Description airport access service</li> <li>Description of HSR-airline schedule coordination</li> </ul>	<ul style="list-style-type: none"> <li>Detailed operating plan</li> <li>Full train schedule</li> <li>Trip times with and without airport access service</li> </ul>
Marketing and distribution	<ul style="list-style-type: none"> <li>None needed</li> </ul>	<ul style="list-style-type: none"> <li>Likelihood of airline cooperation</li> <li>Goals for coordinated operation</li> </ul>	<ul style="list-style-type: none"> <li>Simulation of CRS and other website screen presence</li> <li>Trip times with and without codesharing</li> <li>Ridership estimates with and without codesharing</li> </ul>
Pricing and revenue sharing	<ul style="list-style-type: none"> <li>None needed</li> </ul>	<ul style="list-style-type: none"> <li>Notional airport access fares and revenue</li> <li>Notional revenue sharing arrangement for connecting air service</li> </ul>	<ul style="list-style-type: none"> <li>Pricing of airport access and rail-air connection trips</li> <li>Capacity analysis</li> <li>Revenue forecast from capture of airport access trips and connecting air passengers</li> <li>Cost/benefit analysis for inclusion of airport access service</li> </ul>

### ***Combined choice of airport and access mode in multi-airport regions***

HSIPR proposals that provide a station at a major airport have the potential to influence both the airport access modes as well as the choice of airport in multi-airport regions. The HSIPR system may capture some airport access passengers, and it may also improve the access characteristics of the airport sufficiently to make that airport a more attractive option for some area travelers.

There is a growing literature studying the factors determining airport choice in multi-airport regions, and the consensus of this literature is that *access time* and *flight frequency* are the two most important factors considered by travelers.<sup>30</sup> More recently this literature has incorporated more complex formulations that allow for the modeling of the choice of access mode along with the choice of airport. This is an important advancement, since these choices are generally understood to be closely related in practice, along with a number of other decisions such as choice of airline, desired arrival time, fare class, etc.

Results from mostly academic literature in this area suggest that HSIPR studies that incorporate a station at a major airport in a multi-airport region should attempt to account for the potential influence of the HSIPR system on airport choice. The literature suggests that sophisticated modeling tools are required to represent this joint/sequential decisionmaking process, and estimates of airport access ridership should therefore be scrutinized to ensure that they incorporate the latest methodological developments.

### **Tourism travel**

A special category of HSIPR forecasting concerns tourism travel. There may be instances where the HSIPR system is patronized not for transportation purposes (whether for intercity travel, airport access, or intra-urban commuting), but rather for its novelty and entertainment value. That is, some patrons may use the system simply for the enjoyment of experiencing a cutting-edge, novel technology that has not previously been available in the US.

This type of ridership is relatively rare in US HSIPR proposals, and is most likely to occur with projects that involve the most advanced technologies (maglev or 220+ mph high speed rail) and in corridors that attract significant numbers of tourists and have a station configuration that allows for short distance rides. HSIPR proposals involving Las Vegas or Orlando might attract some amount of entertainment ridership -tourists that visit these places may wish to add a ride on the HSIPR system to the other recreational activities available to them during their stay.

Like airport access ridership, commuter ridership, or ridership derived from the capture of connecting air passengers, this type of ridership must be estimated separately as it represents distinct behaviors not otherwise modeled in traditional forecasting models. Like these other potential sources of ancillary ridership, tourism ridership should not generally be expected to constitute a significant fraction of total HSIPR patronage.

Tourism ridership can perhaps be estimated by analogy to existing fixed guideway transportation systems that are known to attract this type of patronage. Examples include the Seattle Monorail, Las Vegas Monorail, and the Detroit People Mover. The fixed guideway system used as the point of comparison should most closely resemble the characteristics of the proposed HSIPR system for the purposes of forecasting (similar size metropolitan area, similar amount of tourist travel to the region, similar operating characteristics of the service, etc.).

---

<sup>30</sup> Basar, G. and Bhat, C.R. (2002), "A Parameterized Consideration Set model for Airport Choice: An Application to the San Francisco Bay Area," Technical Paper, Department of Civil Engineering, University of Texas at Austin, Austin, TX.

## 7 Best practice forecasting for different study stages

Earlier in this report, three notional stages of HSIPR project study were distinguished:

- **Preliminary stage** - a screening or feasibility-type study that considers a potentially wide range of project alternatives defined as concept-level designs. The study is typically based more on assumptions than on primary data collected in the study area. The output is the identification of a smaller set of alternatives to carry to intermediate stage analyses, together with supporting evidence to make the case for a more detailed consideration of the proposal;
- **Intermediate stage** - a more in-depth examination of a smaller set of alternatives, intended to reliably select one or very small number of alternatives to advance to final stage analysis. Some primary data collection - including notably travel surveys in the study area - will usually be undertaken. Forecasting work at this stage tends to involve considerable interaction with other members of the study team (engineers, operations planners, forecasters, economic evaluators) to develop a coherent project definition and a cogent presentation of project benefits and impacts. The project analysis may still involve large numbers of assumptions, but these will now be specific rather than generic; and
- **Final stage** - as accurate a definition and analysis of the planned project as possible, documenting its characteristics and impacts for final regulatory or environmental review by government authorities, and/or to support decision making by potential project stakeholders.

It should be recognized that a particular study may incorporate elements that are characteristic of different study stages. For example, a preliminary stage study may need somewhat more refined ridership estimates typically associated with an intermediate stage study to eliminate some alternatives in view of HSR system development objectives. Nevertheless, distinguishing the three study stages is a useful way to highlight differences in data sources, modeling approach and outputs associated with different degrees of project development and the corresponding study resource implications.

In this chapter, the discussion HSIPR ridership and revenue forecasting data sources and methods of the preceding chapters is focused into recommendations about best practices for each of the study stages. Recommendations take into account the time and resource constraints typically associated with each stage, as well as the expectations regarding study output accuracy and the management of forecasting risk. The chapter is organized into sections corresponding to the different study stages. For each study stage, a discussion of modeling issues is followed by a table that summarizes typical approaches used in studies of that stage to address different elements of the modeling process. Final comments identify common errors and pitfalls that study reviewers should be alert to.

The modeling elements that are presented in the summary tables, and referred to in the discussion of modeling issues at each stage, are the same as those identified earlier:

- Definition of the study time frame;
- Definition of the study area;
- Definition of the level of geographic detail;
- Base and forecast year trip table preparation;
- Base and forecast year modal network preparation;
- Mode choice modeling;
- Induced travel modeling;
- Assignment;

- Feedback; and
- Computational tools.

### **Preliminary stage**

Recall that HSIPR ridership and revenue forecasting for preliminary stage studies is generally conducted for purposes such as project candidate screening and feasibility analysis. In such studies, it is typically important to develop rough ridership and revenue estimates for a number of project alternatives in a relatively short time and with relatively limited resources. The objective is typically to assess the relative performance of the candidate alternatives relative to each other, and to establish the basic feasibility and approximate performance characteristics of the most promising alternatives, with sufficient reliability to support a decision about progressing some or none to the next stage of study. Study proponents accept some risk associated with the inherently less accurate forecasts: it is preferred to analyze a larger number of candidates in less detail, in the hopes of identifying promising superior alternatives that can be studied with greater accuracy in subsequent stages. Because of the considerable variability possible in the scopes, number of alternatives, geographic extent, amount and nature of readily-available information and other factors, it is difficult to be specific about the level of effort required by a preliminary stage ridership and revenue forecasting study. As a very rough indication, however, such a study might take three to six months of effort by an expert team.

#### ***Best practice***

In general terms, best ridership and revenue forecasting practice in preliminary stage studies consists of methods that:

- Allow the rapid estimation of HSIPR project ridership and revenue;
- Do not require a very detailed description of project characteristics or extensive data collection; and
- Provide results that correctly indicate the relative performance of project alternatives, and that are sufficiently accurate to identify the project candidates that will be useful to study further.

The level of study - in terms of the representation of travel supply and demand - tends to be aggregate and less detailed rather than the opposite. For example, TAZs tend to be relatively large geographic units, and the representation of access/egress modes is typically done in summary rather than detailed fashion. Similarly, the number of individual market segments (or trip purposes) is usually limited, and secondary travel markets (for example, intra-urban or local airport access travel) are frequently analyzed using approximation methods.

The avoidance of extensive primary data collection - usually imposed by the project time frame and/or budget limitations - means that key components of the modeling process, such as the mode choice model, tend to be based on methods and values transferred from other studies, with coefficient values asserted/transferred (i.e. directly chosen) rather than estimated statistically from locally-collected data. Nonetheless, it is important to make some effort to adapt and adjust these to actual conditions in the study area using readily-available data; adjusting the values of time used in the model based on local wage rates is a common practice, for example. The models are usually less elaborate than those developed from extensive primary data collected locally, and typically do not handle (or handle in a very simplified way) effects such as decongestion impacts on competing modes.

Table 7-1 summarizes some of the main features of preliminary stage ridership modeling approaches.

**TABLE 7-1. PRELIMINARY STAGE RIDERSHIP MODELING APPROACHES**

<b>Modeling Element</b>	<b>Description</b>
Definition of study time frame	Typically 20 or more years following the service opening year. Modeling will typically either (i) analyze two years explicitly, with interpolation and extrapolation for other years; or (ii) analyze one year, with other years derived from growth factor methods.
Definition of study area	Defined as a geographic band centered on the project alignment. Catchment areas around stations are defined in terms of crow-fly distance or access time or cost thresholds. Connections with other common carrier modes identified but not defined in detail.
Definition of level of geographic detail	Typically relatively aggregate zones; could be whole metropolitan areas. Little consideration of intra-metropolitan travel markets. Summary consideration of external zones and their demand generation.
Base and forecast year trip table preparation	<p>Trip segmentation</p> <ul style="list-style-type: none"> <li>• Few trip purposes (e.g. business / non-business) at most</li> <li>• Limited additional segmentation</li> </ul> <p>Trip table development</p> <ul style="list-style-type: none"> <li>• Limited or no primary travel data collection: for base year trip tables reliance on available public sources for air and rail trips; traffic counts for highways at lowest volume intermediate points with suitable assumptions; and previous studies; can ignore bus trips</li> <li>• Simple growth factor methods used to develop future year trip tables</li> </ul>
Base and forecast year modal network preparation	<p>At the preliminary stage it is acceptable for the base and future year service characteristics of rail, competing and access/egress modes to be developed without preparation and processing of an explicit network representation. Base year skim tables of zone-to-zone service characteristics are often developed directly:</p> <ul style="list-style-type: none"> <li>• For the rail project, from a description of proposed project characteristics (rail simulation models are typically not needed)</li> <li>• For auto, from actual travel time runs combined with data from travel planners and other sources; access/egress times are average auto times for the large zones; cost attributes often factored from distance</li> <li>• For air and conventional rail, from published fares and schedules</li> </ul> <p>Future year service characteristics are estimated based on planned or anticipated changes in modal service provision, with congestion increases accounted for via growth factors (and not explicitly determined).</p>
Mode choice modeling	<p>Limited or no primary behavioral data collection</p> <p>Mode choice models typically developed by transferring models used elsewhere and asserting (rather than statistically estimating from local data) coefficient values</p> <p>At most limited adaptation of models to study area conditions via incorporation of area data (e.g. local wage rate to determine value to time)</p>
Induced travel modeling	Typically a fixed percentage applied to non-induced trips, or estimated via simple elasticity methods
Assignment	Typically not done
Feedback	Typically not done
Computational tools	Frequently spreadsheet-based; specific components of commercial transportation modeling software packages, or custom software developed for this level of analysis may sometimes be used; use of commercial software packages to implement a complete forecasting model is unusual at this stage.

### ***Common errors and pitfalls***

Preliminary stage HSIPR ridership and revenue forecasting involves tradeoffs inherent in developing an approach that on the one hand can be implemented relatively quickly and applied to a potentially large number of alternatives, and on the other hand provides sufficient accuracy in the forecasting outputs to guide decisions about subsequent stages of project development. This may mean that forecasting work at this stage requires more experience, judgment and intuition than is typically called for in later stages.

Base year trip table development is challenging at this stage because of the general inadequacy of data, particularly for intercity automobile travel (air and conventional rail trip tables present less of a problem in this regard, as discussed above). Reviewers of a preliminary stage study should closely examine the approach used to develop the auto trip table and be prepared to challenge its reasonableness.

For example, some studies attempt to infer the auto trip table from the more easily developed tables for other modes such as air by factoring the latter based on anticipated auto/air mode splits. While this is not necessarily incorrect, the mode splits used for a particular origin-destination pair in this approach should realistically represent the modal competitive situation specific to that pair (based for example on distance and generalized costs); applying a single factor to all origin-destination pairs would certainly be incorrect.

As noted, the mode choice models and coefficients used in a preliminary stage study are typically transferred from other studies rather than statistically estimated from local data. As such, the reasonableness of the model in the intended study area should be carefully scrutinized. A preliminary stage study should adapt a transferred model to local conditions through recognition of local socio-economic factors (e.g. the local wage rate affecting model values of time) or transportation conditions (e.g. limited adjustments to mode-specific constants to represent local mode service qualities); absence of such adaptation efforts could be a warning to reviewers.

The HSIPR project that is being studied may connect to future rail projects in other corridors. These other projects may have already been studied, or will be studied in a separate effort. The question is how the ridership from the other project should be incorporated in the forecasts for the project under study; including ridership between city pairs served by the separate HSR projects. Clearly, results from other studies should not simply be overlaid on project ridership forecasts without a critical review of the other studies' assumptions, methods and reasonableness of results. If the project ridership from external sources is a significant fraction of the total, reviewers should be alert to potential inconsistencies, and should consider requiring a more integrated study (that treats the different connecting corridors within a single study framework) at subsequent stages.

### **Intermediate stage**

HSIPR ridership and revenue forecasts for intermediate stage studies are generally conducted to refine and improve the accuracy of preliminary stage results in order to better assess candidate project alternatives, with the purpose of eliminating those that perform less well and providing a more reliable basis for the selection of the preferred alternative that will be advanced to the final stage of study. In such studies, it is important to be able to efficiently analyze the set of alternatives retained after the preliminary stage, as well as to develop forecasts that are sufficiently accurate to discriminate among potentially similar alternatives, having some confidence that the overall forecasts will not be significantly changed by the more detailed final stage studies. Intermediate stage studies attempt to reconcile the two contrasting objectives of efficiency and accuracy, so there is still more than minimum risk associated with the forecasts. As a very rough indication of the required level of effort, an intermediate stage ridership and revenue forecasting study might take four to eight months of effort by an expert team.

Of course, it is also possible that some studies will skip the intermediate stage of forecasting, and proceed directly to final stage studies, while other studies may view the intermediate stage study as an attempt to

reduce the uncertainty in only one part of the overall forecasts. Such omissions are permissible as long as the final stage study is well done.

### ***Best practice***

As noted above, intermediate stage forecasting studies must find an appropriate balance between the need to analyze a number of alternatives and the requirement for better than preliminary stage accuracy in that analysis. Usually this means that best practice for this stage requires:

- Some level of effort spent collecting and analyzing primary data for the service area to make possible improved forecasting accuracy;
- Increased levels of information regarding service characteristics of the project and the access/egress times. This often involves more interaction with project designers and operations planners than is common in preliminary stage studies, with intermediate forecasting results influencing the project design and service plan, and *vice versa*; and
- Identification of the key factors that influence demand, and some degree of increased attention to these specific factors through specific data collection and analysis activities and development of focused models to be applied within the overall forecasting process.

The level of study - in terms of the representation of travel supply and demand - tends to be more detailed rather than in preliminary stage work, with greater disaggregation around stations and/or major demand generators. In these areas, TAZs tend to be smaller geographic units, and the representation of HSIPR access/egress times (including parking time and cost) might be relatively explicit. A network representation may be prepared for some modes (e.g. auto), although this is not always the case. The number of market segments (or trip purposes) may be expanded to include key segments and their unique characteristics, beyond (say) simply business/non-business purposes, and secondary travel markets (for example, intra-urban or local airport access travel) will often be separately analyzed in some detail where appropriate.

Although intermediate stage studies typically involve some primary data collection (including rigorous traveler surveys), study time and budget constraints may limit the extent of such effort compared to a final stage study; on the other hand it may be known in advance that a single data collection effort will need to serve for both levels of study, and in this case the intermediate level data collection and analysis effort would be much closer to that of a final level study. The level of effort and sophistication in the study's mode choice model development will typically determine the corresponding features of the data collection effort. The analysis might take account of effects such as decongestion impacts on competing modes and feedback to account for changed levels of service on demand for the various modes, particularly if an explicit network representation is available; note however that these effects are frequently not sufficiently important to merit separate consideration in this stage of the modeling process.

Table 7-2 summarizes some of the main features of intermediate stage ridership modeling approaches.

**TABLE 7-2. INTERMEDIATE STAGE RIDERSHIP MODELING APPROACHES**

<b>Modeling Element</b>	<b>Description</b>
Definition of study time frame	Typically 20 or more years following the service opening year. Modeling will typically analyze two years explicitly, with interpolation and extrapolation for intermediate or later years.
Definition of study area	Catchment areas around stations are defined in terms of access time or cost thresholds. Connections with other common carrier modes may be explicitly represented.
Definition of level of geographic detail	Somewhat detailed zoning around stations and major demand generators. Important sub-markets (e.g. intra-metropolitan travel, local airport access, tourism or convention travel) may be separately modeled. External zones and their demand generation may be explicitly considered.
Base and forecast year trip table preparation	<p>Trip segmentation</p> <ul style="list-style-type: none"> <li>• Major trip purposes are broken out (e.g. tourism)</li> <li>• Some additional segmentation (e.g. income, travel party size)</li> </ul> <p>Trip table development</p> <ul style="list-style-type: none"> <li>• Combination of some primary travel data collection with use of public sources, previous studies for base year trip tables</li> <li>• Growth factor or synthetic methods used to develop future year trip tables</li> </ul>
Base and forecast year modal network preparation	<p>Either direct reference to data sources (published schedules and fares; travel planners) or an explicit network representation (especially for the automobile mode) might be used to develop base and future year service characteristics of rail, competing and access/egress modes. Rail project levels of service are typically provided by the study's engineering and operations planning teams, although detailed rail simulation modeling is typically not needed. Rail level of service information may be approximately derived from the service plan but may not represent it in complete detail.</p> <p>Future year service characteristics are estimated based on planned or anticipated changes in modal service provision, with congestion increases accounted for via growth factors or through use of an explicit network representation where available.</p>
Mode choice modeling	<p>Some primary behavioral data collection (e.g. stated preference surveys) would typically be done for intermediate stage studies. The scale of effort (e.g. sample size) would typically be limited unless intended to also serve for final stage studies.</p> <p>Some econometric model development based on the stated preference survey results, including separate models for the major trip purposes and/or market segments</p>
Induced travel modeling	Typically a fixed percentage applied to non-induced trips, or estimated via simple elasticity methods
Assignment	Typically not done
Feedback	Often not done, unless an explicit network representation is available and the effects are felt to be important
Computational tools	Varies: may be some combination of commercial transportation planning software packages, custom developed software, and multi-purpose application software such as spreadsheets and databases.

### ***Common errors and pitfalls***

Compared to preliminary stage studies, intermediate stage HSIPR ridership and revenue forecasting should increase the incorporation of study area context and project-specific factors, while at the same time preserving a relatively efficient forecasting work stream that can be applied to investigate a number of project alternatives. Reviewers should judge how effectively a particular study addresses and trades off these contrasting objectives, and assess the degree of confidence in the ridership and revenue forecasts that support the selection of a preferred alternative (or very small group of similar alternatives) that will be advanced to final stage study.

To begin with, reviewers should particularly focus on the design, execution and analysis of the primary data collection effort (especially including travel surveys) that is normally involved in intermediate stage studies. While detailed review of such efforts is a task for subject matter experts, non-specialist reviewers should judge whether the major market segments and trip purposes have been adequately sampled, if the survey instruments stringently avoid the use of language or graphics that might bias responses, and if survey questions relate reasonably to the ordinary travel experience of respondents and avoid far-fetched and ill-defined hypotheticals.

Some effort should be made in an intermediate stage study to prepare base year trip tables (particularly for the automobile mode) from primary data sources such as counts, surveys or anonymous vehicle or mobile phone tracking data; such sources will typically be combined with other publicly-available data on trips by non-automobile modes, as well as any other relevant and reliable datasets (e.g. trip table estimates from prior studies). Methods for trip table preparation are still under development and have not been standardized. Reviewers will need to assess the reasonableness of the resulting trip table itself obtained as much as the specifics of the data collection and processing methods. Examination of trip table values at the level of large geographic aggregations (e.g. city-to-city rather than TAZ-to-TAZ) can be useful for this purpose.

The mode choice models used in an intermediate stage study may result from a combination of econometric analysis of travel survey results, and results transferred from other studies. In all cases, reviewers should judge the reasonableness of the models by, among other things, examining the implied values of time (and comparing this to local information on incomes and wage rates), the relationships of in- and out-of-vehicle times, the cost or time equivalents of any mode-specific constants, and the demand elasticities computed with the models for the levels of service and mode shares characterizing the study situation.

An intermediate stage HSIPR study should consider access/egress issues in some detail, so as to develop accurate values for access/egress levels of service (e.g. travel time and cost, transfer time and penalties, parking impedances, etc.) Reviewers should examine the assumptions made about future year access/egress characteristics, particularly if substantial improvements are assumed. Reviewers should also check that the levels of service of competing modes (e.g. automobile and air) are sensible and represent prevailing values, as obtained for example from online trip planners or similar tools.

### **Final stage**

By the final stage of HSIPR project study, a single preferred project definition will have been retained. (Occasionally a very few closely related project alternatives will be carried to this level of study.) Accuracy is paramount in ridership and revenue forecasting studies at this stage. All aspects of the study effort should focus on attaining the highest possible level of forecast accuracy in order to minimize the uncertainties and risks associated with the project implementation. Final stage studies examine projects in greater detail than in earlier stage studies: the representation of the service plan and fare policy, of access/egress modes and their service levels, of user market segments and TAZ geography are all typically more detailed than in preliminary or intermediate stage studies. This greater level of detail is needed to support more detailed

project design and operations planning, rolling stock purchasing decisions, fare policy decisions and other aspects of the project definition, as well as to provide commensurately detailed inputs to final stage project financial planning, environmental impact analyses and public benefits assessment. Primary data collection, in the form of rigorously designed and executed surveys of potential project users, is usually necessary to attain the required level of forecasting accuracy and detail. Final stage forecasting studies are substantial undertakings that, to fix ideas, might take a year or more (including data collection) for an expert team to complete.

### ***Best practice***

In general terms, best practice in final stage studies consists of methods that allow the preparation of accurate, detailed ridership and revenue forecasts. These methods frequently require more time and resources to develop and/or apply than those typically used in earlier stage studies, but this tradeoff is necessary. Key characteristics of best practice final stage studies include:

- Careful identification of all key market segments likely to contribute to project ridership;
- A high-quality primary data collection effort focused on understanding the likely behavior of the key market segments in response to the HSIPR project;
- A serious effort to develop base and forecast year trip tables for the study area;
- Detailed and accurate representation of key aspects of the project service characteristics, as well as those of important access/egress modes;
- Mode choice models that are appropriately sensitive to the influence of these characteristics on the travel behavior of the key market segments, as determined from the primary data collection effort;
- Investigation of the project's induced travel potential incorporating detailed characteristics of the study area; and
- Extensive reasonableness checking and sensitivity analyses to identify the factors with the largest impact on forecast results, and to quantify the likely range of forecast uncertainty. These analyses may lead to additional study efforts to reduce the uncertainty associated with identified key factors.

In summary, final stage studies are characterized by intensive efforts to represent with as much detail and accuracy as possible the specific features of the proposed project and its service area, to understand and capture the likely response of the different types of traveler in the service area to the project, and to assess as realistically as possible the range of uncertainty associated with the forecasts. All important aspects of the project and its potential ridership should be identified and analyzed in terms of the local context. Reliance on modeling results from studies conducted elsewhere should be kept to a minimum.

As noted, the representation of travel supply and demand needs to be detailed. TAZs in key portions of the study area, such as at activity centers or around stations, should be relatively small geographic units, and there is a detailed representation of the service characteristics of the rail project and access/egress modes. This frequently involves interaction with the engineering and operations planning teams within the overall HSIPR project study. Similarly, the number of individual market segments (or trip purposes) is usually expanded to include all significant distinct groups or purposes, and secondary travel markets (for example, intra-urban or local airport access travel) are analyzed explicitly using separate models.

The reliance on extensive primary data means that the quality of the data collection effort - including its design, execution and analysis - is an important determinant of the quality of the forecasts. Stated preference surveys of potential HSIPR users, in particular, should be rigorously conducted in terms of respondent recruitment, survey instrument design, careful removal data outliers, and state-of-practice statistical analysis of survey responses. A considerable effort would normally be devoted to the preparation of base year trip tables - particularly automobile trip tables - including through methods such as counts, travel surveys and/or anonymous vehicle or cell phone tracking. Because of the emphasis on locally-

collected data, there should be relatively little need to invoke coefficient values or modeling results from other studies (although reference to the actual features and characteristics of comparable operational HSIPR systems may be quite useful). Any forecasting issue felt to be potentially important should be investigated in sufficient detail to quantify its contribution to the forecasting results, and to define an appropriate treatment of it within the forecasting process, rather than handled through assumptions.

In a final stage forecasting effort, effects that might have been treated summarily in earlier stages - for example, the project's decongestion impacts on highways and airports - should be treated in detail if they are likely to be important (though this does not mean that analysis of such effects should necessarily be included in every final stage study - the need must be assessed based on the specifics of the project and the competing modes). Analysis of decongestion impacts may require development of an explicit representation of the other modes' networks and facilities, and detailed predictions of the competitive impacts on these of the HSIPR project.

Table 7-3 summarizes some of the main features of final stage ridership modeling approaches.

**TABLE 7-3. FINAL STAGE RIDERSHIP MODELING APPROACHES**

<b>Modeling Element</b>	<b>Description</b>
Definition of study time frame	Typically 20 or more years following the service opening year. Modeling should analyze at least two years explicitly; additional years may be appropriate to model time staging of significant events (e.g. completion of additional project segments, connections to other projects, development of major demand generators).
Definition of study area	Defined as a geographic band centered on the project alignment. Catchment areas around stations are defined in terms of detailed access time and cost level of service, based on an accurate representation of connections to other modes.
Definition of level of geographic detail	Relatively detailed zones, particularly around stations and travel demand generators. Potentially detailed consideration of intra-metropolitan travel markets including local airport access where applicable. Careful consideration of external zones, connections to other projects, and the ridership increments that they might generate.
Base and forecast year trip table preparation	<p>Trip segmentation</p> <ul style="list-style-type: none"> <li>• Trip purposes sufficient to represent all potentially important types (e.g. commuting / business / non-business / airport access / tourism / other)</li> <li>• Potential additional segmentation (e.g. automobile availability and trip dependency; income; travel party size)</li> </ul> <p>Trip table development</p> <ul style="list-style-type: none"> <li>• Potentially extensive primary travel data collection, particularly to prepare automobile trip tables (counts, surveys, anonymous tracking)</li> <li>• Detailed forecasts of future year trip tables based on anticipated socio-economic growth</li> </ul>
Base and forecast year modal network preparation	<p>Base and future year service characteristics of rail, competing and access/egress modes might be developed through preparation and processing of an explicit network representation validated against base year conditions. Base year skim tables of zone-to-zone service characteristics may be developed directly (from timetable and fare data) or through use of the explicit network representation (particularly for the automobile mode). HSIPR project speeds and times would typically be developed by study operations planners.</p> <p>Future year service characteristics are estimated based on planned or anticipated changes in modal service provision. Highway congestion level increases are generally forecast from a network representation; air congestion changes from airport volume/delay functions.</p>
Mode choice modeling	<p>Extensive primary behavioral data collection, particularly stated preference surveys of potential users</p> <p>Mode choice models estimated from collected survey data using state-of-practice econometric methods</p> <p>Typically elaborate models that represent local context and demand influencing factors</p> <p>Distinct sub-models corresponding to different market segments and trip types</p>
Induced travel modeling	Elasticity or more elaborate methods
Assignment	May be done depending on the detail and complexity of the rail networks and operations plan
Feedback	May be done depending on anticipated magnitude of effects
Computational tools	Forecasting process may be implemented using commercial travel forecasting software packages, potentially interfacing with other software tools. Custom software developed for this level of analysis may sometimes be used. Spreadsheet-based methods may be cumbersome and error-prone considering the amount of detail and data typically manipulated in a final stage study.

### ***Common errors and pitfalls***

Reviewers of final stage HSIPR ridership and revenue forecasting studies should expect to be furnished sufficient documentation of the model development and application effort to allow a full understanding of the process, methodology and results. In general, reviewers should be alert to anything that suggests that the accuracy and level of detail of the study have been compromised through insufficient representation of travel demand and supply conditions, inadequate data collection, facile assumptions, and/or improper methodology. A final stage study should provide insight into the robustness of the study results, including sensitivity analyses of forecasts under changes in input values or model assumptions.

Base year trip table preparation - particularly for automobile trips - should be based on primary data collection of some sort, unless readily-available data provide sufficient detail and reliability. Even when such data collection is carried out, reviewers of a final stage study should closely examine the approach used to process the collected auto data and be prepared to challenge its reasonableness - methodology for this purpose is still evolving and has not been standardized.

Forecast year trip tables should be developed in part based on anticipated socio-economic growth in the project service area. Elasticities or similar factors used for this purpose should be documented and justified using area-specific data - for example, a time series analysis of historical travel growth.

Mode choice models used in a final stage study should normally be estimated from original data, notably in the form of stated preference surveys of potential HSIPR project users. While the detailed review of an SP survey design normally requires in-depth technical knowledge, non-specialist reviewers should be alert to common errors including inadequate control of the respondent recruitment process (e.g. recruitment of an unrepresentative "convenience" sample), and survey questions that, by their wording or graphical presentation, tend to encourage a particular response. Typically, respondents should be required to choose between modal alternatives having service characteristics that require a tradeoff (e.g. between travel time and cost), without one alternative dominating all others. The choice situations represented in an SP survey should represent actual trips made by the respondent, and not represent a completely fictitious situation to which the respondent cannot readily relate.

Detailed review of the econometric analysis used to develop mode choice models from the collected data requires technical knowledge. However, non-specialist reviewers can assess the care with which the collected data are free of outliers and errors, and can form an opinion about the reasonableness of some aspects of the estimated models and their utility functions - for example, the implied values of the components of time for different market segments and trip purposes, the equivalent monetary (or time) value of mode-specific constants, computed demand elasticities and the like. Reviewers should expect to find the details of this information in the forecasting methodology documentation.

Lack of a network based representation of the modal alternatives and their service levels in future years is not necessarily an error but, if not used, reviewers should closely examine the study methodology and assumptions made regarding forecast year levels of service by mode. This is particularly important if the study considers feedback and excepts substantial decongestion effects (benefits) from the HSIPR project.

Review of the induced demand results is similar in all stages of study. Methodology in this regard has not been standardized, so a review should focus on the reasonableness and soundness of the approach, and verify that the predicted amount of induced travel is not excessive: a forecast of induced demand that is greater than (say) 10% of total HSIPR travel should be examined closely.



## 8 HSIPR ridership and revenue checklists

This chapter presents a ridership and revenue checklist in the form of a series of twelve tables intended to be used by reviewers of HSIPR studies. Each table corresponds to one component of a HSIPR ridership and revenue forecasting study, and lists various items related to the component that reviewers should be alert to. The twelve components are:

1. Study time frame
2. Study area network
3. Definition of travel markets
4. Definition of market segments
5. Trip tables
6. Demographic and socio-economic characteristics
7. HSIPR alternatives definition
8. LOS characteristics of the competing modes
9. LOS characteristics of the access modes
10. Mode choice models
11. Induced demand models
12. Revenue calculation

The checklist tables provide a reasonably detailed and comprehensive listing of items under each main component. Not all the items are expected to be included in all study stages. Preliminary and intermediate stages in particular might not include some of these items. The discussions in Chapters 2 and 3 will serve as guides in determining whether or not these items should be included in a study under review.

**TABLE 8-1. STUDY TIME FRAME CHECKLIST**

<b>Required Item</b>	<b>Methods</b>
Initial (base) model year	Based on assumptions used
Final (horizon) model year	Based on assumptions used
Any intermediate years and years beyond the horizon year (if used)	Based on assumptions used

**TABLE 8-2. STUDY AREA NETWORK CHECKLIST**

Required Item	Methods
Study area boundary	Based on assumptions used
Geographic units or Traffic Analysis Zones (TAZ)	<ul style="list-style-type: none"> <li>Based on assumptions used</li> <li>Finer level of zones (e.g., TAZs used for MPO level modeling) for final stages</li> <li>Aggregate level of zones (e.g., Metropolitan Statistical Areas, counties etc.) for preliminary stages</li> <li>Combination of finer and aggregate level of zones or one or the other depending on the study context for intermediate stages</li> </ul>
Road network within the study area boundary (both existing and planned forecast year networks)	<ul style="list-style-type: none"> <li>Based on assumptions used</li> <li>Mainly for final stages and in some cases for intermediate stages</li> </ul>
Transit network within the study area boundary (both existing and planned forecast year networks)	<ul style="list-style-type: none"> <li>Based on assumptions used</li> <li>Mainly for final stages and in some cases for intermediate stages</li> </ul>

**TABLE 8-3. DEFINITION OF TRAVEL MARKETS CHECKLIST**

Required Item	Methods
Intercity	Always included
Urban	As applicable depending on the HSIPR alignment and station locations
Airport access	As applicable depending on the HSIPR alignment and station locations
Airport choice (in a multi-airport region)	As applicable depending on the HSIPR alignment and station locations
Within corridor and corridor to corridor	If multiple HSIPR corridors are present in the study area

**TABLE 8-4. DEFINITION OF MARKET SEGMENTS CHECKLIST**

Required Item	Methods
Business and non-business	<ul style="list-style-type: none"> <li>Based on assumptions</li> <li>At a minimum for all stages</li> </ul>
Existing mode users	<ul style="list-style-type: none"> <li>Based on assumptions</li> <li>As applicable if diversion choice model is used</li> </ul>
Resident and non-resident	<ul style="list-style-type: none"> <li>Based on assumptions</li> <li>For final or intermediate stages</li> <li>As applicable if the travel characteristics in the corridor require</li> </ul>
Home-based work and home-based other	<ul style="list-style-type: none"> <li>Based on assumptions</li> <li>For final or intermediate stages</li> <li>As applicable if the travel characteristics in the corridor require</li> </ul>
Non-home based	<ul style="list-style-type: none"> <li>Based on assumptions</li> <li>For final or intermediate stages</li> <li>If required by travel characteristics in the corridor</li> </ul>
Tourism and others	<ul style="list-style-type: none"> <li>Based on assumptions</li> <li>If required by travel characteristics in the corridor</li> </ul>

**TABLE 8-5. TRIP TABLES CHECKLIST**

<b>Required Item</b>	<b>Methods</b>
Combined for all modes or mode specific trip tables at TAZ to TAZ level by <ul style="list-style-type: none"> <li>• Travel markets (as applicable)</li> <li>• Market segments (as applicable)</li> </ul>	<ul style="list-style-type: none"> <li>• Based on assumptions used</li> <li>• Depending on the mode choice models used (choice models vs. diversion choice models)</li> </ul>
Person or vehicle trip tables (Person trips = vehicle trips * vehicle occupancy)	Based on survey data or assumptions
Both directional or one way trips	Based on calculation
Seasonality demand profiles (traffic levels by month of year and time of day – peak and off-peak)	Based on existing data or assumptions
Annualization factor to calculate yearly trip tables	Based on assumptions or traffic count data
Direct demand models <ul style="list-style-type: none"> <li>• Model specifications</li> <li>• Model parameter values</li> <li>• Statistical properties of the model</li> <li>• Implied growth rates</li> </ul>	<ul style="list-style-type: none"> <li>• Based on models used</li> <li>• Usually used in final stages and in some cases intermediate stages</li> </ul>

**Note:** Checks to be performed for base and forecast years.

**TABLE 8-6. DEMOGRAPHIC AND SOCIO-ECONOMIC CHARACTERISTICS CHECKLIST**

<b>Required Item</b>	<b>Methods</b>
Population at TAZ level	For all stages (based on assumptions)
Employment at TAZ level	For all stages (based on assumptions)
Household income at TAZ level	If applicable, for final or intermediate stages (based on assumptions)
Number of hotel rooms at TAZ level	If applicable, for final or intermediate stages (based on assumptions)
Number of visitors at TAZ level	If applicable for the corridor travel market (based on assumptions)
GDP at available geographic unit level in the study area	If applicable for aggregate growth measures mainly in preliminary stages (based on assumptions)
Any assumed growth rates	Based on assumptions

**Note:** Checks to be performed for base and forecast years.

**TABLE 8-7. HSIPR ALTERNATIVES DEFINITION CHECKLIST**

<b>Required Item</b>	<b>Methods</b>
Various alignments	Based on study scope and engineering assumptions
Various technology options (e.g. maglev, steel wheel on steel rail etc.)	Based on study scope and engineering assumptions
Operating hours	Based on study scope and service planning assumptions
Station locations	Based on study scope and engineering assumptions
Stopping patterns (potential to have multiple stopping patterns e.g. express and local service)	Based on study scope and service planning assumptions
Train frequency (variable frequencies based on stopping pattern, week of day etc. if applicable)	Based on study scope and service planning assumptions
Wait times	Half the headway or other estimates
Station to station run time	<ul style="list-style-type: none"> <li>• Based on study scope and engineering assumptions</li> <li>• From detailed train simulation or detailed calculations based on operating characteristics) in final and some cases intermediate stages</li> <li>• From high level estimates generally in preliminary stages</li> </ul>
Dwell time at stations	Based on study scope and service planning assumptions
Fares	<ul style="list-style-type: none"> <li>• Based on study scope and service planning assumptions</li> <li>• Average fares based on yield</li> <li>• By market segments <ul style="list-style-type: none"> <li>○ At least business and non-business for all stages</li> <li>○ Other segments if applicable (e.g. intercity fare, commuter fare based on monthly passes etc.) for final and intermediate stages</li> <li>○ Fares by classes of HSR service (only in final stages) – very rarely used</li> <li>○ Revenue maximizing fares in final or in some cases intermediate stages</li> </ul> </li> </ul>
Reliability of service (amount of delay)	<ul style="list-style-type: none"> <li>• Based on assumptions</li> <li>• In final and some cases intermediate stages</li> </ul>
Classes of service (first and second class) if applicable	<ul style="list-style-type: none"> <li>• Based on assumptions</li> <li>• Only in final stages (but rarely used)</li> </ul>
Amenities on-board the HSIPR service	<ul style="list-style-type: none"> <li>• Based on assumptions</li> <li>• Used in SP surveys while describing the HSIPR service (final or intermediate stages)</li> </ul>
HSIPR station facilities	<ul style="list-style-type: none"> <li>• Based on assumptions</li> <li>• Used in SP surveys while describing the HSIPR service (final or intermediate stages)</li> </ul>
Other notable amenities of the HSIPR service	<ul style="list-style-type: none"> <li>• Based on assumptions</li> <li>• Used in SP surveys while describing the HSIPR service (final or intermediate stages)</li> </ul>

**Note:** Checks to be performed for each phase if the project is implemented in phases.

**TABLE 8-8. LOS CHARACTERISTICS OF THE COMPETING MODES CHECKLIST**

Required Item	Methods
Auto travel times	<ul style="list-style-type: none"> <li>• Network based detailed calculations followed by checks from internet based sources (e.g., Mapquest, Google maps etc.) in final and some cases intermediate stages</li> <li>• Based on high-level estimates (based on free flow speed and distance) or internet based sources (e.g., Mapquest, Google maps etc.) in preliminary stages</li> </ul>
Auto travel costs (person costs or vehicle costs using vehicle occupancy values)	Based on calculation using auto per mile cost and auto distance and tolls if any
Travel times for common carrier modes (rail, air and bus)	Based on published schedules <ul style="list-style-type: none"> <li>• Detailed calculation based on weighted average taking into account daily schedule and seasonal variation in final or some cases intermediate stages</li> <li>• High-level estimates based on generally simple averages in preliminary stages</li> </ul>
Fares for common carrier modes	At least for business and non-business <ul style="list-style-type: none"> <li>• Airfares are based on USDOT DB1B data in final or some cases intermediate stages</li> <li>• Rail fares are based on weighted average of different Amtrak fare levels in final or some cases intermediate stages</li> <li>• Based on high-level calculations of air and Amtrak fares from various search engines for air travel planning and Amtrak website, respectively</li> </ul>
Frequencies for common carrier modes	Based on published schedules <ul style="list-style-type: none"> <li>• Detailed calculation based on weighted average taking into account daily schedule and seasonal variation in final or some cases intermediate stages</li> <li>• High-level estimates based on generally simple averages in preliminary stages</li> </ul>
On-time performance/delay information (as applicable)	<ul style="list-style-type: none"> <li>• Based on detailed calculations of air delay from USDOT data sources in final or some cases intermediate stages</li> <li>• Based on train level or more aggregate Amtrak data on on-time performance in final or some cases intermediate stages</li> <li>• High level aggregate level estimates of on-time performance of the rail and air modes from public sources (internet, Amtrak or airline websites) in preliminary stages</li> </ul>
Wait times for common carrier modes	Half the headway or other estimates based on published schedules (i.e., information on frequency) <ul style="list-style-type: none"> <li>• Detailed calculation based on weighted average taking into account daily schedule and seasonal variation in final or some cases intermediate stages</li> <li>• High-level estimates based on generally simple averages in preliminary stages</li> </ul>
Security delays at airports	Based on assumptions and available data from survey mainly in final and some intermediate stages

**Note:** Checks to be performed for base and forecast years.

**TABLE 8-9. LOS CHARACTERISTICS OF THE ACCESS MODES CHECKLIST**

Required Item	Methods
Auto access times	<ul style="list-style-type: none"> <li>• Network based detailed calculations followed by checks from internet based sources (e.g., Mapquest, Google maps etc.) in final and some cases intermediate stages</li> <li>• Based on high-level estimates (based on free flow speed and distance) or internet based sources (e.g., Mapquest, Google maps etc.) in preliminary stages</li> </ul>
Auto access costs (person costs or vehicle costs using vehicle occupancy values)	Based on calculation using auto per mile cost and auto distance, tolls and parking costs if any
Walk time to HSIPR stations after parking	<ul style="list-style-type: none"> <li>• Based on detailed engineering calculations or high level estimates in final or intermediate stages</li> <li>• Based on high level estimates in preliminary stages</li> </ul>
Taxi, limo, shared ride travel time to HSIPR stations	Based on auto access times as mentioned above
Taxi, limo, shared ride travel costs to HSIPR stations	Based on estimates using per mile cost and highway distance (calculated either from network based representation sometimes in case of final or intermediate stages or internet mapping sites for preliminary stage(if used at all))
Walk time to HSIPR stations after getting off taxi, limo and shared ride service	<ul style="list-style-type: none"> <li>• Based on detailed engineering calculations or high level estimates in final or intermediate stages</li> <li>• Based on high level estimates in preliminary stages (if used at all)</li> </ul>
Wait time for boarding transit access modes to go to the HSIPR stations	Half the headway or other value as appropriate
Travel time on the transit access modes to reach HSIPR stations	Based available schedule data in final and intermediate stages and high level assumptions in preliminary stages (if used at all)
Fares on transit access modes to reach HSIPR stations	Based on available fare data from transit agencies
Transfer penalty (in the form of time or cost) to transfer from transit access modes to HSIPR mode or between multiple transit access modes	Based available schedule data in final and intermediate stages and high level assumptions in preliminary stages

**Note:** Checks to be performed for base and forecast years.

**TABLE 8-10. MODE CHOICE MODELS AND UNDERLYING DATA CHECKLIST**

Required Item	Methods
Revealed Preference (RP) data summary	<ul style="list-style-type: none"> <li>Based on new survey data collected in final and some cases intermediate stages</li> <li>Based on existing off-the shelf survey data from earlier surveys in preliminary (if used at all) and some cases intermediate stages</li> </ul>
Stated Preference (SP) questionnaire	<ul style="list-style-type: none"> <li>Based on new SP survey conducted in final and some cases intermediate stages</li> <li>Based on existing off-the shelf SP questionnaire from earlier surveys (if available) in some intermediate stage studies</li> </ul>
Stated Preference (SP) data summary	<ul style="list-style-type: none"> <li>Based on new SP survey data collected in final and some cases intermediate stages</li> <li>Based on existing off-the shelf SP survey data from earlier surveys (if available) in some intermediate stage studies</li> </ul>
Survey sample designs	<ul style="list-style-type: none"> <li>Based on new survey data collected in final and some cases intermediate stages</li> <li>Based on existing off-the shelf survey data from earlier surveys in preliminary (if used at all) and some cases intermediate stages</li> </ul>
Survey sample characteristics	<ul style="list-style-type: none"> <li>Based on new survey data collected in final and some cases intermediate stages</li> <li>Based on existing off-the shelf survey data from earlier surveys in preliminary (if used at all) and some cases intermediate stages</li> </ul>
Surveying methods and survey locations	<ul style="list-style-type: none"> <li>Generally in final and some cases intermediate stages when new survey is conducted</li> <li>Information from existing off-the shelf survey in preliminary (if used at all) and some cases intermediate stages</li> </ul>
Model estimation techniques and procedure	In final stages and some cases intermediate stages only when new mode choice models are estimated
Model assumptions	Based on assumptions for all stages (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)
Model structures	<ul style="list-style-type: none"> <li>Choice models vs. diversion choice models; MNL vs. nested logit models; nesting structures etc.</li> <li>For all stages (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)</li> </ul>
Utility function specifications	As used (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)
Model parameter values	As used (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)
Values of times for different travel time components (e.g., line haul time, waiting time, access time etc.)	Based on calculations using the ratios of travel time (different travel time components) and travel cost parameters in the mode choice model (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)
Values of the mode specific constants	Absolute values or mode specific constants expressed in equivalent time or cost values by dividing the constants by corresponding time and cost model parameters (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)

Required Item	Methods
All the model variables and the specifications used for these variables	As used (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)
Statistical properties of the models (e.g., standard errors, <i>t</i> statistics)	Based on calculations (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)
Generalized cost and/or time as implied by the model parameters and variables	Based on calculations (from newly developed models in final and some cases intermediate stages and existing models transferred from other locations/studies in preliminary/intermediate stages)
Mode shares after the application of mode choice models	Based on calculations (for all stages)
Source of HSIPR ridership (diversion values and percentages from each existing mode)	Based on calculations (for all stages)

Note: Checks to be performed for each market segment and travel market

**TABLE 8-11. INDUCED DEMAND MODELS CHECKLIST**

Required Item	Methods
Calculation procedure	<ul style="list-style-type: none"> <li>Model based in final or some cases intermediate stages</li> <li>Simply based on percentage of HSIPR diversions from the mode choice models in preliminary and some cases intermediate stages</li> </ul>
Induced demand models <ul style="list-style-type: none"> <li>Model specifications</li> <li>Model parameter values</li> <li>Statistical properties of the model</li> </ul>	In final or some cases intermediate stages
Model estimation techniques and procedure	In final or some cases intermediate stages
Model assumptions	Based on assumptions (in final or some cases intermediate stages)
Portion of the Stated Preference (SP) questionnaire (if used) used for induced demand calculation	Based on questionnaire used (in final or some cases intermediate stages)
Induced demand estimates	Based on calculations for all stages
Induced demand as a % of HSIPR ridership	Based on calculations for all stages

Note: Checks to be performed for each market segment and travel market

**TABLE 8-12. REVENUE CALCULATION CHECKLIST**

Required Item	Methods
Ramp up period to steady state operation of HSIPR service introduction or improvements	Based on assumptions for all stages
Results of sensitivity analyses	Based on calculations (definitely in final and intermediate stages but may also be used for preliminary stage)
Revenue maximizing analyses	Based on calculations (generally only in final stage)
Price base year (all revenues, fares, costs expressed in this year's dollars)	Based on assumptions and calculations for all stages
Revenue and cost escalation percentage for forecast year values	Based on assumptions for all stages

