Journal of Choice Modelling, 3(1), pp. 58-83
www.jocm.org.uk

California Statewide Model for High-Speed Rail

Maren Outwater[1,*] Kevin Tierney[2,†] Mark Bradley[3,Ŧ]
Elizabeth Sall[4,§] Arun Kuppam[5,¶] Vamsee Modugula[6,β]

[1] Resource Systems Group, 917 E. Spooner Rd, Fox Point, WI 53217
[2] 206 Broad Meadow Rd, Needham, MA 02492
[3] Mark Bradley Research and Consulting, 524 Arroyo Avenue, Santa Barbara, CA 93109
[4] San Francisco County Transportation Auth, 100 Van Ness Ave, San Francisco, CA 94102
[5] Cambridge Systematics, 9015 Mountain Ridge, Suite 210, Austin, TX 78759
[6] Citilabs, A-2/14, Vanashree CHS, Plot No 1&2, Sector 58A, Nerul, Navi Mumbai, 400706

Received 19 March 2008, received version revised 22 December 2008, accepted 19 September 2009

Abstract

The California High Speed Rail Authority (CHSRA) and the Metropolitan Transportation Commission (MTC) have developed a new statewide model to support evaluation of high-speed rail alternatives in the State of California. The approach to this statewide model explicitly recognizes the unique characteristics of intraregional travel demand and interregional travel demand. As a result, interregional travel models capture behavior important to longer distance travel, such as induced trips, business and commute decisions, recreational travel, attributes of destinations, reliability of travel, party size, and access and egress modal options. Intraregional travel models rely on local highway and transit characteristics and behavior associated with shorter distance trips (such as commuting and shopping).

Keywords: discrete choice models, interregional, high-speed rail, induced demand, ridership forecasting, trip frequency, destination choice, mode choice

* Corresponding author, T: +1 425 269 9684, F: + 011 802 295 1006, moutwater@rsginc.com
† T: +1 617 839 5295, kevintierney@rocketmail.com
Ŧ T: +1 805 564 3908, F: + 011 805 564 3927, mark_bradley@cox.net
§ T: +1 919 469 3609, F: + 011 919 302 0265, elizabeth.sall@sfcta.net
¶ T: +1 512 691 8503, F: + 011 512 691 3289, akuppam@camsys.com
β T: +91 022 27525025, F: + 011 510 523 9706, vmodugula@citilabs.com

Licensed under a Creative Commons Attribution-Non-Commercial 2.0 UK: England & Wales License
http://creativecommons.org/licenses/by-nc/2.0/uk/

ISSN 1755-5345

1 Introduction

By 2030, California's population will grow to 50 million people, which will nearly double interregional travel to one billion trips per year. High-speed trains are being considered to alleviate the need to build – at a cost of nearly $100 billion – about 3,000 miles of new freeway, plus five airport runways, and 90 departure gates over the next two decades. Electric and fully separated from automobile traffic, California's high-speed train would provide a new transportation option available to more than 90 percent of the residents of the state. At full build-out, the system will run from San Diego north to Sacramento and San Francisco.

The project objectives were to develop a new ridership forecasting model that would serve a variety of statewide planning and operational purposes:

- To evaluate high-speed rail ridership and revenue on a statewide basis;
- To evaluate potential alternative alignments for high-speed rail into and out of the San Francisco Bay Area; and
- To provide a foundation for other statewide planning purposes and for regional agencies to better understand interregional travel.

The core model design feature is the recognition that interregional and urban area travel is distinct and should be modelled separately to capture these distinctions accurately. This led to our approach to develop separate, but integrated, interregional and intraregional models, as demonstrated in Figure 1. There are two primary reasons for developing separate models for interregional and urban area travel: first, the trip purposes are different and second, the interregional travel models need to explicitly estimate induced demand. These models are applied to both peak and off-peak conditions for an average weekday. Weekend travel demand and annual ridership estimates are developed using annualization factors developed from observed data on high-speed rail systems around the world.

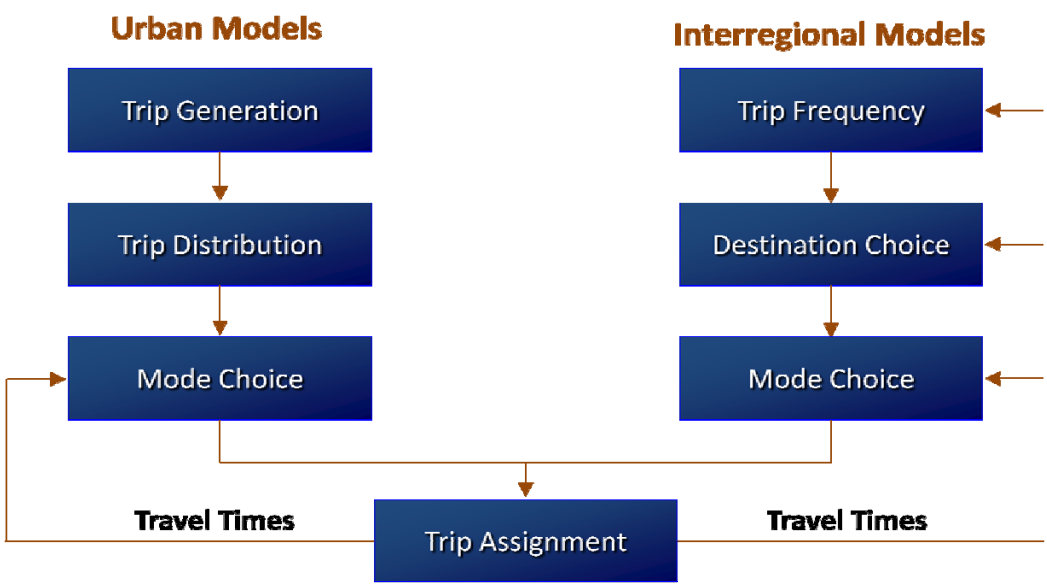

Figure 1. Integrated Modeling Process

There are fourteen regions established in the state that define interregional and intraregional travel:

- Interregional models estimate trip frequency, destination choice, and mode choice stratified by trip purpose (business, commute, recreation, and other) as well as by distance (trips greater than or less than 100 miles) and by trip type (trips made by residents of the four largest cities in California versus other trips). The interregional trip frequency models estimate induced travel based on improved accessibilities due to high speed rail options. The interregional models are similar in structure to models developed for the Australian Very Fast Train (VFT) project, except that there is more spatial detail (i.e. many more zones) and all models were estimated based on revealed preference with full and consistent nesting of the accessibility measures from access/egress models up through trip frequency models. There are also some similarities to the Norwegian and Swedish national models, which combine long-distance and short-distance travel and have similar modeling structures but have estimated simultaneous destination and mode choice with sampled destinations.
- Intraregional models are based on trip tables generated from the Metropolitan Planning Organization (MPO) models and estimate mode choice of urban area trips. These mode choice models reflect local urban area highway and transit systems as well as options for high speed rail within the region. Urban travel is stratified by trip purpose (work, school, college, other, and non-home-based).

The interregional and intraregional area models are based on travel survey data collected for these purposes.

2 Literature Review

Although few intercity high speed ground transportation (HSGT) systems have been implemented in the United States, planners have been performing ridership forecasting analyses and benefits assessments for proposed systems for more than thirty years. Almost all of the ridership forecasts have been in support of analyses of one of the 11 designated intercity corridors authorized by the Federal Railroad Administration in the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 and the Transportation Equity Act for the 21st Century (TEA-21) of 1998, including (Federal Railroad Administration 2006):

- Northern New England;
- Northeast Corridor;
- Empire Corridor;
- Keystone Corridor;
- Southeast Corridor;
- Gulf Coast Corridor;
- Florida Corridor;
- Chicago Hub Network;
- South Central Corridor;
- Pacific Northwest Corridor; and
- California Corridor.

The first U.S. HSGT ridership forecasts were developed to support investments to improve the Northeast Corridor (Koppelman *et al.* 1984). Since then, a variety of forecasting techniques have been used in corridor feasibility analyses, ridership and revenue evaluations, and environmental impact studies. The literature describing the analytical techniques employed in these various studies and the current states of forecasting practice over time have been reviewed by Koppelman *et al.* (1984), Miller (1992, pp. 378-389), Forinash (1992), and Horowitz (2006). The initial forecasting efforts relied on aggregate direct demand models, such as the Quandt and Baumol abstract mode model (1966), the Rand Corporation model developed by DeFerranti *et al.* (1973), and the Peers and Bevilacqua structural intercity model (1976).

As urban modelers adopted the four-step modeling process for regional forecasting, intercity models also became multi-step processes where the intercity trip tables are estimated and forecast, and separate mode choice models are used to determine modal shares under different service scenarios. Most of the more recent U.S. HSGT ridership forecasting efforts can be characterized by the specific analyses used to develop intercity trip tables, determine modal split, and the level to which the different model steps are integrated.

Modelers have employed many trip generation and trip distribution techniques to forecast future year intercity trip tables. For an early analysis of the Florida high speed rail corridor, PRC Voorhees used trip purpose specific growth factor models to forecast trip tables (1983). Other modelers, such as Thakuriah *et al.* (1999) and Cohen *et al.* (1978), have relied on gravity models for trip table prediction. One researcher identified by Horowitz used a time series model to forecast future intercity travel demand (Yu 1970). Many others have used a cross-sectional regression-based total travel demand (direct demand) formulation, including for instance, Booz-Allen and Hamilton (1989), TMS/Benesch (1991), and Brand *et al.* (1992). The previous analysis of the California high speed rail corridor relied on this modeling approach (Charles River Associates 2000).

Since the early 1980's, several different forms of disaggregate mode choice models have been developed for forecasting U.S. HSGT ridership. Early efforts, such as those by Grayson (1981), employed multinomial logit models, but subsequent models have expanded the technique in different directions. Cohen *et al.* (1978, pp. 21-25), Brand *et al.*(1992, pp. 12-18), and Charles River Associates (2000) formulated mode choice as separate binary diversion models where percentages of auto, air, and bus passengers are diverted to HSGT individually through binary models. This approach seeks to eliminate the troublesome outcome of the multinomial logit model's IIA property. Other researchers have relied on the use of nested logit models for intercity mode choice (TMS/Benesch 1991 and Forinash and Koppelman 1993). Researchers, most notably Chandra Bhat, have experimented with a wide variety of nested logit model formulations and variable combinations (1995) (1997) (1998). With advancements in effective ways for combining revealed preference and stated preference survey data (Ben-Akiva and Morikawa 1990) (Bradley and Daly 1997), intercity modelers have also begun to use stated preference survey methods to a greater extent for forecasting HSGT mode choice.

Most intercity transportation demand models have treated traveler's decisions regarding trip frequency, destination choice, mode choice, and route selection as separate sequential choices. The model system described in this paper seeks to connect the different model components by passing information from one choice component to the others during model development. When the multinomial logit and nested logit model components are applied, they are fully consistent with each other and are sensitive to each other's changes. Integrated modeling approaches have been proposed

by and implemented by Koppelman and Hirsh (1986), Koppelman (1989), Proussaloglou and Tierney (1999) and, outside the U.S., by Algers (1993) and Gunn *et al.* (1992). The latter approach was developed for an Australian VFT study and uses a very similar approach to the models developed for California, including the use of both stated and revealed preference data and linking mode, destination and trip frequency models through the use of accessibility measures.

3 Interregional Models

The interregional models are comprised of four sets of models: trip frequency, destination choice, main mode choice, and access/egress mode choice. The structure and contents of the interregional modelling system is presented in Figure 2.

The trip frequency model component predicts the number of interregional trips that individuals in a household will make based on the household's characteristics and location. The destination choice model component predicts the destinations of the trips generated in the trip frequency component based on zonal characteristics and travel impedances. The mode choice components predict the modes that the travelers would choose based on the mode service levels and characteristics of the travelers and trips. The mode choice models include a main mode choice, where the primary interregional mode is selected, and access/egress components, where the modes of access and egress for the air and rail trips are selected. A combined destination and mode choice model was initially considered, but the choice set for destination choice was all zones (4,667) and the combination of destinations and modes would have made this choice set too large.

3.1 Data for Estimation

There were three types of data compiled for the study: travel surveys, networks, and socioeconomic data. Some of the travel surveys were collected specifically for this study, three were available from MPOs around the state (Southern California Association of Governments (SCAG), San Francisco Metropolitan Transportation Commission (MTC), and Sacramento Association of Governments (SACOG)), and there was a Caltrans statewide survey available. The interregional models were based on revealed and stated preference surveys, collected specifically for this study, of air and rail travelers, as well as additional households in the state to capture auto travelers. These new data were collected in fourteen regions in California. These were combined with revealed preference surveys of households across the state collected by Caltrans and interregional travel extracted from the MPO regional travel surveys (San Francisco, Sacramento, and Los Angeles). Intraregional mode choice models were based on urban area travel surveys in combination with a stated preference survey for high speed rail conducted in Los Angeles. By combining the various available data sources, we were able to provide more robust data sets for model estimation than was otherwise possible. A summary of these sources shows the trip records derived from each:

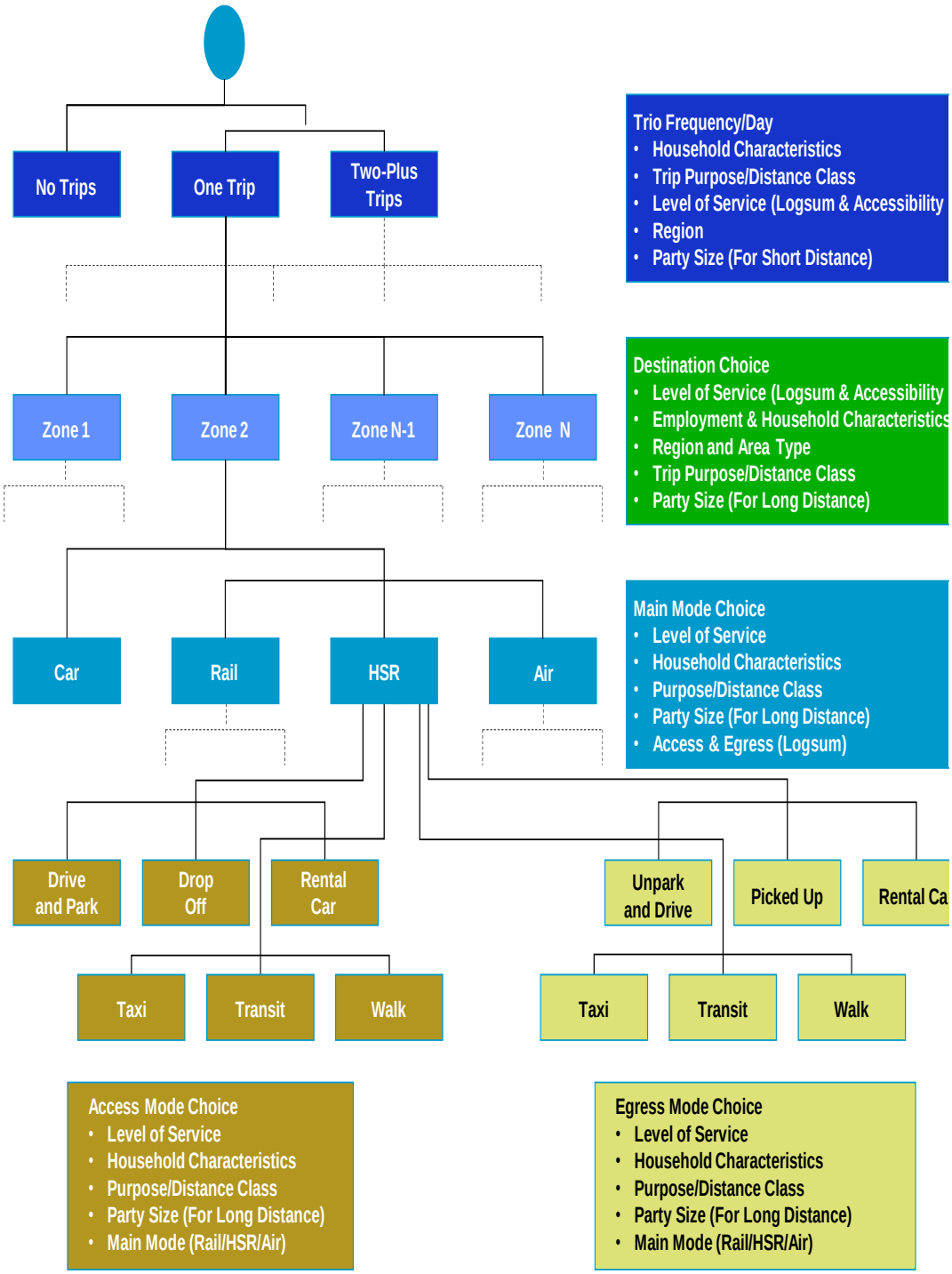

Figure 2. Interregional Model Structure

- Air, rail, auto passenger surveys 2678 trips
- Caltrans travel surveys 2820 trips
- SCAG travel surveys 343 trips
- MTC travel surveys 723 trips
- SACOG travel surveys 318 trips

After combining these surveys, 6,882 completed surveys were available to use for model estimation, as shown in Table 1. There was different estimation datasets used for each model component, depending on the requirements for the model. This is described in more detail in the Interregional Model System Development Report (Cambridge Systematics 2006).

There are highway, air, rail, and local transit networks to support both the urban area and interregional travel models. The socioeconomic data includes household data in four classifications (household size, income groups, number of workers, and vehicle ownership) and employment data by type.

3.2 Accessibility Measures

In the development of the trip frequency models, accessibility measures were estimated for all trips to approximate the destination choice logsum measure. In the final models, accessibility measures were retained for intraregional trips because the intraregional models maintained by the MPOs do not include destination choice models, which are necessary to produce logsum measures. Accessibility measures for interregional trips were replaced with logsum measures from the destination choice models in the final models, as described below. There were four accessibility measures calculated, as follows:

- Auto peak work trip accessibility

$$A_{peak_auto} = LN\left[1 + \sum_d TotalEmp_d \times \exp\left(\underline{-2 \times Time_{peak_auto}} / Time_{peak_mean}\right)\right]$$

Table 1. Total of Survey Interregional Trips by Mode, Distance, and Purpose

	Drive	Air	Rail	Bus	Other	Total
Long Trips						
Business	314	620	27	18	17	**996**
Commute	263	15	9	1	74	**362**
Recreation	1114	228	80	3	23	**1448**
Other	365	85	17	8	91	**566**
Short Trips						
Business	381	14	48	3	15	**461**
Commute	1136	0	168	9	108	**1421**
Recreation	873	2	29	3	52	**959**
Short Other	591	1	10	23	44	**669**
Total	**5,037**	**965**	**388**	**68**	**424**	**6,882**

Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Interregional Model Systems Development Report, Table 2.6, August 2006.

- Auto off-peak non-work trip accessibility

$$A_{offpeak_auto} = LN\left[1+\sum_d (Households_d + RetailEmp_d + ServiceEmp_d) \times \exp\left(\underline{-2 \times Time_{offpeak_auto}} / Time_{offpeak_mean}\right)\right]$$

- Non-Auto peak work trip accessibility

$$A_{peak_nonauto} = LN\left[1+\sum_d TotalEmp_d \times \exp\left(\underline{-2 \times Time_{peak_nonauto}} / Time_{peak_mean}\right)\right]$$

- Non-Auto off-peak non-work trip accessibility

$$A_{offpeak_nonauto} = LN\left[1+\sum_d (Households_d + RetailEmp_d + ServiceEmp_d) * \exp\left(\underline{-2 \times Time_{offpeak_nonauto}} / Time_{offpeak_mean}\right)\right]$$

where:

TotalEmpd = total employment at the destination zone;
Householdsd = total households at the destination zone;
RetailEmpd = retail employment at the destination zone;
ServiceEmpd = service employment at the destination zone;
Timepeak_auto = highway travel time during the peak (based on congested time) from the origin zone to the destination zone;
Timepeak_nonauto = transit travel time during the peak (based on congested time) from the origin zone to the destination zone;
Timeoffpeak_auto = highway travel time during the off-peak (based on free-flow travel time) from the origin zone to the destination zone;
Timeoffpeak_nonauto = transit travel time during the off-peak (based on free-flow travel time) from the origin zone to the destination zone;
Timepeak_mean = average travel time from the origin zone to all possible destination zones during the peak period, calculated from the average of survey respondents travel time based on peak network times; and
Timeoffeak_mean = average travel time from the origin zone to all possible destination zones during the off-peak period, calculated from the average of survey respondents travel time based on off-peak network times.

3.3 Logsum Measures

Logsum measures are a means to estimate a weighted average of travel time and cost that can be fed back from one component to another. A summary of the logsum measures for each model component is as follows:

- Trip frequency models use “logsum” measures from the destination choice models, which are intended to capture the fact that it is easier to make relevant interregional trips from some zones than from other zones. For initial model estimation, a synthesized network zone accessibility measure was used.
- Destination choice models use logsum measures from the main mode choice models that are intended to provide measures of the composite impedance across

all modes of travel between each of the zones. For initial model estimation, a mode choice logsum calculate from the Caltrans statewide model was used.

- Main mode choice models use a logsum from the access/egress mode choice models. This was estimated prior to the main mode choice models, so a substitute measure was not necessary.

This allows higher level model components to reflect accessibility measured accurately from lower level models.

3.4 Trip Frequency

We used a simple multinomial logit (MNL) model to predict interregional trip frequency. Eight trip frequency models predict interregional person-trips per day, segmented by trip purpose (business, commute, recreation, and other) and length (over or under 100 miles). The MNL formulation allows important explanatory variables such as accessibility measures to affect the propensity to make interregional trips. In this case, the composite log-sums from the destination choice model are fed back to the trip frequency model to account for "induced trip making" behavior. The trip frequency models are segmented by length to allow different model specifications and parameters for short and long trips. For each model, the choice set for each person is zero, one, or two or more interregional trips per day. The final model specification constrains the variable coefficients of one-trip and two-trip choices to be equal, while allowing the alternative-specific constants for one- and two- trip choices to be estimated individually. This overcomes some illogical individual variable coefficients for each market segment, but allows us to retain separate choices for interregional travel. The estimation dataset includes all trips (both interregional and other) from the survey for a total of 108,401 trips. Table 2 presents the estimation results for long interregional trips (short trips are not included for brevity).

Three types of variables were tested in the trip frequency models: socioeconomic, accessibility, and geographic region of residence. Even though the trip frequency models are estimated at the person-level, estimation variables were constrained to be at the household-level to be consistent with existing future year socioeconomic predictions. Socioeconomic variables that were tested in model specifications include: household size, household size greater than two dummy variable, number of household workers, zero-worker household dummy variable, number of household vehicles, number of household vehicles is less than the number of household workers dummy variable, zero-vehicle household dummy variable, high household income (greater than $75,000), medium household income (Between $35,000 and $75,000), low household income (less than or equal to $35,000), and a missing income dummy variable for survey records with no income collected. The missing income dummy variable is used during model estimation, but is not included in the final model specification for application. Some variables retained in the final models were not significant (i.e. *t*-stat were less than 1.9) but were reasonable in terms of size and sign, so were retained to provide sensitivity in the models to these variables. The mode/destination logsum parameters in the frequency models were in the expected 0 to 1 range. They were not very significant, but that is typical for frequency models. If the logsum parameters were too large and significant, there would be a huge generation effect, and not reasonable. The accessibility and logsum measures were constrained during model calibration to the coefficients from the commute trip because this produced conservative and stable estimates of induced travel.

Table 2. Trip Frequency Models for Long Trips

	Business		Commute		Recreation		Other	
Observations	108,401		108,401		108,401		108,401	
Final log-likelihood	-1,168.3		-1,823.7		-2,048.8		-1,865.3	
ρ^2(const)	0.08		0.10		0.04		0.09	
Variable	**Coeff.**	**(*t*-stat)**	**Coeff.**	**(*t*-stat)**	**Coeff.**	**(*t*-stat)**	**Coeff.**	**(*t*-stat)**
Level of Service								
Intraregion accessibility	-0.128[1]	(-1.5)	-0.217	(-4.0)	-0.4[1]	(-6.0)	-0.532[1]	(-7.4)
Mode/destination choice logsum	0.466[1]	(1.5)	0.123	(0.6)	0.656[1]	(2.8)	0.159[1]	(0.6)
Household Characteristics								
Medium income	0.527	(1.5)	0.188	(0.8)				
High income	1.139	(3.0)	0.291	(1.1)	-0.246	(-1.3)	0.393	(2.1)
Missing income[2]	0.955	(2.3)	0.34	(1.1)	0.282	(1.3)	0.158	(0.7)
Fewer cars than workers in HH	-0.412	(-1.0)	-0.457	(-1.6)	-0.922	(-2.4)	-0.915	(-2.2)
No cars in HH								
Fraction of HH who are workers	0.537	(1.9)	1.274	(5.8)				
No workers in HH	-2.098	(-3.4)	-2.668	(-3.7)			0.372	(2.4)
Household size								
1 person household							-0.424	(-2.0)
3+ person household					-0.482	(-3.9)	-0.378	(-2.8)
Location Variables								
SACOG resident[3]	0.976	(3.7)	0.918	(4.7)	1.084	(4.4)	2.527	(10.3)
SANDAG resident[3]	-0.704	(-1.1)	-0.419	(-1.0)	1.344	(3.5)	0.92	(1.8)
SCAG resident[3]	-1.176	(-3.6)	-1.644	(-6.3)	-0.031	(-0.1)	0.259	(0.8)
MTC resident[3]	-1.372	(-3.6)	-0.729	(-2.9)	1.011	(3.4)	1.134	(3.4)

[1] These coefficients were later constrained during model calibration to the value of the commute coefficient. [2] Missing income was not used in model application. [3] The location variables are similar to constants applied to a region and these were calibrated to match observed trips in the region. The final values range from -1.421 to 4.676. Alternative specific constants are not presented here for brevity. Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Interregional Model Systems Development Report, Table 3.3, August 2006.

The estimation results follow an intuitive pattern. More household workers increase one's propensity to make interregional business and commute trips, but decrease one's propensity to make interregional recreation and other trips. The income coefficients indicate that as income increases, more interregional trips are taken. Households with fewer cars than workers are less likely to have the resources to undertake interregional travel. Three-person households are less likely to undertake interregional recreation and other trips, perhaps substituting this type activity closer to home.

As discussed above, the trip frequency models include measures that capture the accessibility of all relevant travel opportunities from travelers' home zones. For each residence, we calculated three peak/work and three off-peak/non-work accessibility measures for destinations in (1) their home region, (2) outside their region, within 100 miles of home, and (3) over 100 miles from home. The final model specifications rely on synthesized accessibility measures for the within home region destinations and on logsums calculated from the destination choice models for the remaining accessibility measures. The synthesized accessibility measure is necessary within the home region since the urban area models are not destination choice models (they are gravity models) and are therefore not able to produce logsums for the destination choices within the region. Logsums are a means to produce a weighted average of all potential destinations.

A high calculated "regional accessibility" to jobs, goods, and services within one's region of residence indicates less need to travel outside of the region. Therefore, as expected, this variable has a negative effect on all interregional travel. Separate short (within 100 miles of residence and outside the residence region) and long (outside 100 miles of residence and outside the residence region) logsums were calculated to represent accessibility to goods and services outside of one's home region. A higher logsum outside a home region increases the likelihood that an interregional trip will be undertaken.

Regional dummy variables for the MTC, SANDAG, SACOG, and SCAG regions are included to account for the different interregional trip making patterns observed for residents of large, metropolitan areas compared to residents in the rest of California. These were calibrated to match observed trips in these regions. The final values range from -1.421 to 4.676.

3.5 Destination Choice

The destination choice models were estimated with a simple multinomial logit model structure using ALOGIT software. The dataset used for the trip frequency models was combined with the stated-preference (SP) survey (used in the mode choice models) to produce a combined estimation dataset for the destination choice estimation models. The addition of the SP dataset significantly increased the number of "long" (more than 100 miles) trips in the dataset (by nature, the household surveys are generally better at capturing the more typical "short" trips). Since the trip frequency models already differentiate between the two, we can use this information as a valuable input to the destination choice models. This not only constrains an individual's choice set based on destinations being greater or less than 100 miles, but it recognizes that an individual may value different trip characteristics for different distance-categories of travel.

The short trip destination choice models used all four trip purposes modeled in the trip frequency step: business, commute, recreation, and other. Due to sample size considerations, only two aggregate trip purposes were estimated for the long trip destination choice models: business/commute and recreation/other.

The models use multimodal composite impedance from the mode choice model (logsum). This variable measures the combined utility of all available modal choices and level of service characteristics. The coefficient turned out to be positive and significant at the 95 percent confidence level in the destination choice models, indicating that the destination zone is more attractive if it is better accessible.

In all of the destination choice models, we have used a distance power series including distance, distance-squared, and distance-cubed. While common sense would say that all distance coefficients should be negative, one cannot analyze the distance coefficients individually, but as their collective impact. In addition, a great deal of the impedance between origin-destination pairs is captured within the travel impedance term and coefficient. Therefore it is not wrong for the collective effect of distance to be either positive or negative. It should be noted that since we are estimating separate models for "short" and "long" trips, that the "short" trips are automatically capped at 100 miles from the origin. All short trip distance functions show a decreasing function up to 100 miles, which is consistent with our expectations. Both long trip distance functions show a decreasing function from 100 to 250 miles and then an increasing function for trips greater than 250 miles.

Each possible destination zone could have one of three basic area-types assigned to it: rural, suburban, or urban. In the destination choice models we chose "suburban" to be the base. Additionally, we created several interaction terms to capture whether travelers were starting and ending in the same area type: rural to rural, suburban to suburban, urban to urban. We expect that the sign on urban to urban to be positive, and the sign on rural to rural and suburban to suburban to be negative or close to zero.

Similar to the area type interaction variables, the location type interaction variables allow us to relate where you want to go, to where you currently are, based on the location of the origin and destination. We tested four origin-destination location type interaction variables for all the "long" destination choice models: Los Angeles to/from San Francisco, Sacramento to/from San Diego, San Francisco to/from San Diego, and Sacramento to/from Los Angeles. These were adjusted during model calibration to match observed travel. Since all of these locations are major population centers and destinations in the state we generally expect them to have a synergistic quality between them that these variables represent, and thus have positive coefficients (although it makes sense that this may not occur for all trip purposes).

Size functions are used to measure the amount of activity that occurs at each destination zone and incorporate this into the utility of alternative variables. This type of variable is frequently used in destination choice models to account for differences in zone sizes and employment levels. Four size variables are used in these models: retail employment, service employment, other employment, and households. Other employment is used as the base size variable for business and commute trips and is constrained to 1.0 while retail and service are further segmented by household income levels – low, medium, high, and missing. Households are used as the base size variable for recreation and other trips. Income is used as a per person variable as an interaction between employment and income to show that different income levels of the destination choices will affect the attractiveness of the zone for particular travelers. For commute trips, short and long, as income increases, retail employment has a bigger impact on destination choice than service employment.

Table 3 presents the model estimation results of the destination choice models for long trips by trip purpose: business/commute, and recreation/other (the short trip models are not shown for brevity). The distance power series of coefficients for these models are both decreasing functions as expected. All other variables have the sign and size we expect, except for the coefficient of rural-to-rural for recreation/other

trips, which is positive when we expect it to be negative, but it is not significantly different than zero. Mode choice logsum coefficients were constrained during model calibration to half their original value to produce more conservative and stable estimates of destination choice.

3.6 Mode Choice

There were two types of mode choice models developed for this study: access and egress models and main mode choice models. Models were estimated to predict the access and egress modes to and from airports and rail stations. The models were based on actual reported and hypothetical stated data. For people who were intercepted making actual air or rail journeys, the access and egress mode choices are the actual

Table 3. Destination Choice Models for Long Trips[1]

	Business/Commute		**Recreation/Other**	
Observations	1,342		1,922	
Initial log-likelihood	-12,102.6		-17,029.0	
Final log-likelihood	-11,475.4		-16,219.3	
ρ^2	0.052		0.048	
	Coeff.	**(*t*-stat)**	**Coeff.**	**(*t*-stat)**
Level of Service				
Mode choice logsum[2]	0.107[3]	(5.1)	0.103[3]	(6.7)
Distance (miles)	-0.024	(-8.5)	-0.031	(-11.7)
Distance squared/100	0.0070	(8.9)	0.0087	(10.8)
Distance cubed/10000	-0.0005	(-8.0)	-0.0007	(-9.5)
Area type				
Urban destination	0.724	(6.7)	0.810	(9.5)
Rural destination	0.222	(2.0)	0.607	(6.8)
Urban to urban	-0.010	(-0.1)	-0.096	(-0.8)
Suburban to suburban	-0.185	(-1.5)	-0.029	(-0.3)
Rural to rural	-0.112	(-0.7)	-0.036	(-0.3)
Size variables (exponentiated)				
Other employment	1.000	constrained		
Households			1.000	constrained
Retail employment-low income	2.889	(2.1)	0.960	(-0.1)
Service employment - low income	1.728	(1.5)	0.287	(-3.6)
Retail employment -med income	9.318	(4.9)	0.850	(-0.4)
Service employment - med income	2.292	(1.8)	0.373	(-3.3)
Retail employment -high income	7.338	(5.6)	1.385	(0.8)
Service employment - high income	2.525	(2.8)	0.393	(-2.4)
Retail employment -missing income[4]	100.000	(0.1)	0.001	(-0.1)
Service employment - missing income[4]	100.000	(0.1)	0.433	(-1.4)

[1]Does not include district constants and region-to-region constants. [2] Estimated without distance terms. [3]These were later constrained during model calibration by half the original value. [4]Not used in application. Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Interregional Model Systems Development Report, Table 3.9, August 2006.

reported ones. For people whose actual journey was by car, the air and conventional rail access/egress mode choices are hypothetical. Obviously, the HSR access and egress mode choices are hypothetical for all respondents.

For access, the majority of respondents reported either driving and parking at the station/airport or else getting dropped off. For egress, the reported mode shares varied more by purpose and distance, with transit more popular for short trips, and rental car and taxi more popular for long trips and business trips. In all, there were six modes considered for each. A nested structure was adopted, as shown in Figure 3. The auto modes—drive and (un)park, pick up/drop off, and rental car—are all in separate nests, while taxi, transit (bus or light rail), and walk are nested together. This nesting structure gave the most reasonable results for all purposes.

The access and egress mode choice results for the Long Distance segments are shown in Table 4. A reasonable value of time was asserted for each segment based upon a review of other research. As the survey was not designed primarily to estimate access and egress choice models, and the zone size is in a statewide model is quite large for this type of local choice, the fact that access and egress time and cost parameters had to be constrained is not surprising. Also note that the costs of options such as taxi and rental car and airport/station parking are not readily obtained from network data. Other results of note are:

- The out-of-vehicle time coefficients were estimated for most segments, and result in ratios of out-of-vehicle time to in-vehicle time that are in the range of 2.0 to 2.9.
- The pickup and drop off alternatives have an additional negative in-vehicle time effect, capturing the disutility of the driver that has to make the round trip to the airport.
- We did not include taxi cost explicitly, but did include an additional distance coefficient for taxi, which is significant and negative for most segments, typically with an equivalent value of over $1.00 per mile.
- For most segments, transit is less likely to be chosen if there is no reasonable walk access to transit, meaning that a drive to transit path was included instead.
- For most segments, transit, which can include rail and/or bus, is more likely to be chosen if rail is included in the best transit path.

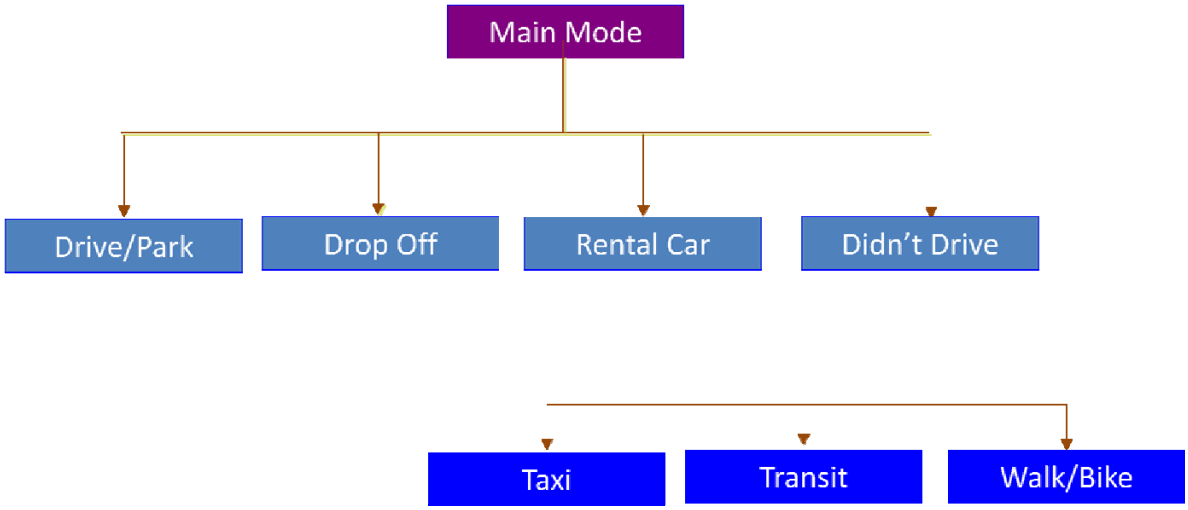

Figure 3. Access and Egress Mode Choice Model Structure

Table 4. Access and Egress Mode Choice Models for Long Trips[1]

	Access Models				Egress Models			
	Business/Commute		Recreation/Other		Business/Commute		Recreation/Other	
Observations	1,500		2,724		1,466		2,668	
Final log-likelihood	-1,662.3		-2,519.4		-2,121		-3,066.6	
$\rho^2(0)$	0.276		0.365		0.075		0.231	
ρ^2(cons)	0.003		0.068		-0.023		0.053	
	Coeff.	**(*t*-stat)**	**Coeff.**	**(*t*-stat)**	**Coeff.**	**(*t*-stat)**	**Coeff.**	**(*t*-stat)**
Level of Service								
Cost ($)	-0.075	constrained	-0.120	constrained	-0.075	constrained	-0.120	constrained
In-vehicle time (min)	-0.060	constrained	-0.030	constrained	-0.060	constrained	-0.030	constrained
Out of vehicle time (min)	-0.147	(-6.4)	-0.083	(-2.5)	-0.139	(-6.2)	-0.060	constrained
VOT IVT ($/hour)	$48.00		$15.00		$48.00		$15.00	
Ratio OVT/IVT	2.45		2.76		2.33		2.00	
Drive and (un)park								
Travel alone			-1.925	(-3.0)				
Fewer cars than persons	-1.547	(-2.2)	-1.903	(-2.8)				
Low income	-2.741	(-1.8)	-1.960	(-2.8)	-18.006	(-2.5)	-1.263	(-1.1)
High income	0.709	(1.6)	0.339	(1.4)				
To/from conventional rail					-9.490	(-2.5)		
To/from high-speed rail					-2.251	(-1.8)		
Airport is LAX	-3.128	(-3.8)	-1.275	(-1.7)				
Airport is SFO	-4.082	(-4.4)	-3.036	(-2.6)				
Airport is SJC			-1.479	(-2.1)				
Airport is SAN	-1.410	(-2.3)	-1.370	(-2.3)				
Rental car								
To/from conventional rail	-5.000[2]	constrained	-5.000	constrained	-3.522	(-2.4)	-1.176	(-3.1)
To/from high speed rail							-0.552	(-2.4)
No cars in HH	5.110	(3.2)						
High income	2.953	(2.4)						
Travel alone							-2.588	(-4.7)
Low income					-2.082	(-0.9)	-1.891	(-3.7)
Get dropped off/picked up								
In-vehicle time (min)	-0.014	(-2.5)	-0.031	(-3.1)			-0.015	(-3.9)
Household size	0.606	(2.9)	0.478	(2.8)	0.974	(2.8)		
Taxi								
Auto distance	-0.084	(-4.8)	-0.071	(-3.8)	-0.126	(-7.9)	-0.052	(-6.6)
To/from conventional rail	-2.827	(-2.6)	-2.265	(-2.4)				
To/from high-speed rail			-1.092	(-2.1)	2.507	(3.6)		
Travel alone			-0.877	(-1.8)			-2.768	(-4.6)
Low income	-3.010	(-1.9)			-3.002	(-2.3)	-1.038	(-2.3)
High income			0.849	(1.9)				
Transit								
No walk egress	-4.836	(-4.6)	-1.807	(-1.9)				
Rail used in path	3.689	(5.2)	1.727	(2.4)			2.960	(5.0)
To/from conventional rail					3.580	(5.2)	1.830	(2.8)
To/from high-speed rail					0.592	(0.7)	1.032	(1.9)
Travel alone			1.569	(2.3)				
No cars in HH			1.439	(1.7)				
Fewer cars than persons	1.480	(2.1)						
Low income			0.846	(1.0)			1.216	(1.9)
Walk								
To/from airport	-5.000[2]	constrained	-2.634	(-1.0)	-2.074	(-2.0)		
Nesting and scaling								
Nest- transit, walk, taxi	0.387	(5.9)	0.451	(3.3)	0.280	(6.9)	0.470	(5.3)
Scale on hypothetical choices	0.682	(15.9)	1.000	constrained	0.516	(9.8)	1.000	constrained

[1] Does not include alternative specific constants. [2] These were later reduced to -3.0 during model calibration. Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Interregional Model Systems Development Report, Table 3.12 and 3.13, August 2006.

- The long segments, taxi, parking, and rental cars are generally less desirable to rail stations than to airports, while transit is more desirable from rail stations. Walking is very rare to or from airports, capturing accessibility affects that are not captured well in the zone system.
- Drive and park access is less likely at the busiest airports – San Francisco (SFO), Los Angeles (LAX), and San Diego (SAN) – and somewhat at San Jose (SJC) as well. This may capture both cost and inconvenience effects at those airports.
- For most segments, those in larger households are more likely to be dropped off.
- In general, high income favors rental car, taxi, and drive and park, and low income slightly favors transit in some segments.
- There is a logsum coefficient less than 1.0 on the nest that includes transit, walk, and taxi. Each of the other three alternatives is in its own "nest," and scaled by the same logsum parameter to preserve equal scaling at the elemental level.
- The scale (the inverse of the residual error variance) for the hypothetical choices relative to the actual choices was significantly lower than 1.0 for most of the Egress model segments. This result indicates that many respondents have difficulty making an accurate assessment of mode choice options in less familiar surroundings at the non-home end of their trip, so that hypothetical choices should be weighted less in estimation than actual ones.

The main mode choice models produce probabilities that each trip will choose one of the main modes (auto, air, conventional rail, and high-speed rail). Several nesting structures were tested for the main mode choice models and the final nesting structure chosen is shown in Figure 4, with all the non-auto modes in a single nest. This structure provided the most logical and statistically sound nesting structure for the mode choice models.

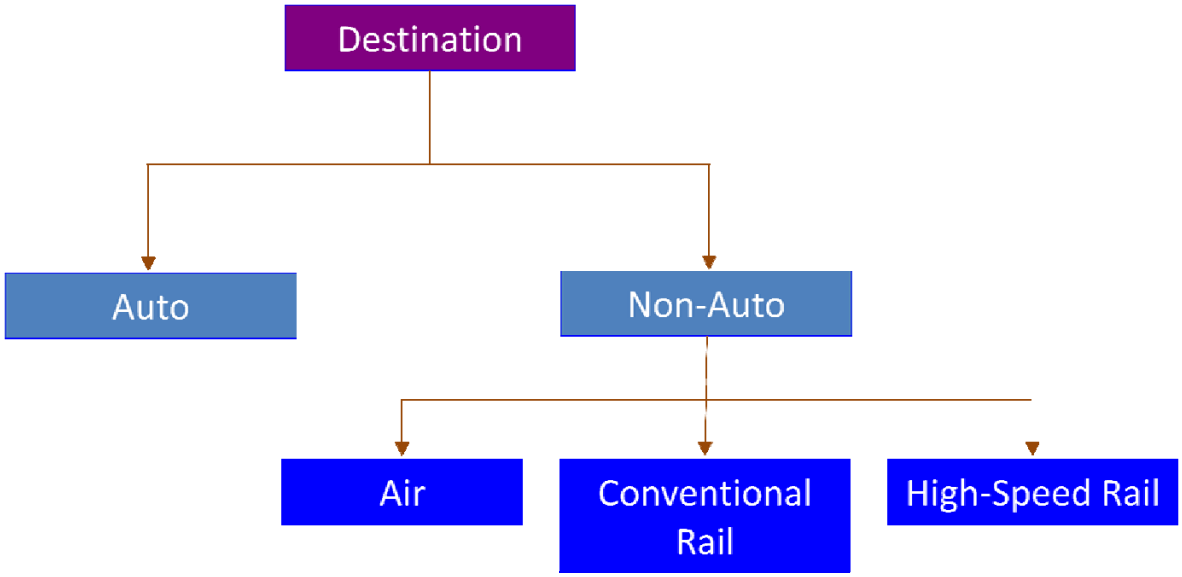

Figure 4. Main Mode Choice Model Structure

The main mode choice models were based on stated preference (SP) survey data. The overall choice shares in the SP data were around 50% for high speed rail, with most of the other choices for the respondents' actual chosen modes. The HSR choice share was highest for business trips and long trips, giving a first indication that HSR substitutes more closely with air than with car.

To prepare the data for estimation, the access and egress mode choice models were first applied to calculate access and egress mode logsums for each alternative. Then, a nested logit model was estimated across the four main modes for each of the segments (only three alternatives for the Short segments, as air was not available for those segments).

The estimation results for the Long Segments are shown in Table 5. Some results of note include the following:

- The cost and in-vehicle time parameters were not constrained during model estimation and produce reasonable values of time. In general, the value of time for the longer, more expensive trips is higher than for the shorter, more frequent trips. This is a typical result.
- The value of frequency (headway) is significant for all segments, but was only about 20 percent as large as the in-vehicle time coefficient. If wait time were half the headway and valued twice as highly as in-vehicle time, then we would expect the same coefficient on headway and in-vehicle time. For these modes, and particularly air, headway is less related to wait time than it is to scheduling convenience. Because none of the levels used in the SP had headways higher than a few hours, the implications for scheduling may not have been large enough to greatly influence mode choice. This coefficient was constrained to match in-vehicle time based on comments from the peer review panel.
- The value of reliability is fairly low for all segments, although with the correct sign. It is very difficult to measure the effect of reliability in a large-scale mailout SP survey, so we decided to use a somewhat higher effect of reliability in application, based on evidence from other models that this was reasonable.
- Those traveling with others are more likely to use car and less likely to use air. This effect was also tested on the cost coefficients and not found to be significant, so this relative mode preference appears to be related to more than just cost – such as the fact that people can share driving for long trips. Party size models were estimated to generate these data, but are not included here for brevity.
- People in larger households are more likely to use car. Even though we already have the group/alone segmentation, people in larger households are likely to be in larger groups.
- Higher income generally favors air and high-speed rail versus auto.
- Low auto availability within the household is related to a lower chance of choosing the auto.
- A nest with air, rail, and HSR, (with car in its own "nest") produced a logsum coefficient below 1.0 for all segments, indicating that this was a reasonable nesting structure for interregional trips.
- The access mode choice logsums were estimated with positive coefficients in the range of 0.14 to 0.46 for all segments.
- For the long trips, the egress mode accessibility seems to have somewhat more influence on mode choice than does the access mode. Travelers may be less

constrained at the home end, where they know the options and can use their own auto, than they are at the destination end.

Table 5. Main Mode Choice Models[1]

	Long Trip			
	Business/Commute		**Recreation/Other**	
Observations	2,918		5,075	
Final log-likelihood	-1,998		-3,936	
$\rho^2(0)$	0.380		0.309	
ρ^2(cons)	0.151		0.154	
	Coeff.	**(*t*-stat)**	**Coeff.**	**(*t*-stat)**
Main Mode Characteristics				
Level of Service				
Cost ($)	-0.017	(-12.8)	-0.035	(-18.5)
In-vehicle time (min)	-0.018	(-13.4)	-0.011	(-14.2)
Service headway (min)	-0.004[2]	(-3.9)	-0.003[2]	(-3.6)
Reliability (% on time)	0.023	constrained	0.005	(1.9)
Access Mode Choice Logsum	0.136	(3.4)	0.204	(3.7)
Egress Mode Choice Logsum	0.171	(3.9)	0.399	(7.1)
Implied Value of Time IVT ($/hour)	$63.64		$18.45	
Ratio Headway/IVT[3]	0.21		0.24	
Trip Characteristics				
Car – Travel in a Group 2+	1.086	(4.6)	1.43	(9.1)
Air– Travel in a Group 2+	-0.356	(-2.8)	-0.505	(-3.7)
Household Characteristics				
Car – Household Size	0.182	(1.2)	0.296	(4.4)
Air – High Income	1.18	(4.6)		
Conventional rail – High Income	0.613	(1.4)		
High-speed rail – High Income	1.147	(4.8)		
Car – Less than 2 Cars per 2+ Household			-0.308	(-2.3)
Nesting				
Nest – air, rail, high-speed rail	0.692	(10.4)	0.738	(13.0)

[1] Does not include alternative specific constants. [2] These were later constrained during model calibration to match in-vehicle time based on comments from the peer review and the modeling team. [3]After the headway coefficient was constrained, this ratio becomes 1. Source: The model was re-estimated when the headway measures were constrained and does not match previously published versions of this model - Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Interregional Model Systems Development Report, Table 3.15, August 2006.

4 Intraregional Models

Intraregional models will be used to forecast high speed rail trips with both ends within an urban area that has more than one proposed high-speed rail (HSR) station. These areas are the San Francisco Bay Area, Greater Los Angeles, and San Diego regions. Regional travel forecasting models in these areas will be modified to forecast urban high-speed rail trips for the San Francisco and Los Angeles areas. The market segments for intraregional travel include typical trip purposes such as home-based work, school, university, shopping, social-recreational, and other trips as well as work- and non-work-related non-home-based trips. San Diego is the only other region that contains the possibility of intraregional high-speed rail trips, but the estimate of these riders is very low relative to the other regions; and the level of effort to develop, calibrate, and apply the regional mode choice model is very high, so we decided to develop intraregional ridership for San Diego using a population-based estimate rather than a traditional mode choice model.

To model intraregional trips, we relied on the trip generation and distribution models in each of the urban areas and modified existing mode choice models. The urban mode choice models include a variety of transit modes, but not specifically a high-speed rail mode. The MTC urban mode choice models were modified to insert a high-speed rail mode based on coefficients and constants from the commuter rail mode, as a conservative estimate. The SCAG urban mode choice model was built from the MTC framework. Following is a brief description of the model implementation for each of the urban areas:

- San Francisco Bay Area - The San Francisco regional model was enhanced to include transit submodes (BART, commuter rail, light rail, ferry, local bus, and express bus) in the mode choice model. This allowed for easier inclusion of the high-speed rail mode in the model. The new mode choice model was validated at the regional level to match observed ridership numbers by operator.
- Southern California Association of Governments Region - The Southern California Association of Governments (SCAG) mode choice models were developed using the parameters and structure of the MTC model in combination with the SCAG networks and trip tables. This model was validated at the regional level to match observed ridership numbers by operator.

Urban trip tables from the MTC and SCAG metropolitan areas were added to the interregional trips for the assignment.

5 Model Application

5.1 Model Validation

The validation of the combined interregional and intraregional (urban) models was completed for the year 2000, because the available observed data for 2000 was more robust than for any other year. This statewide model was estimated from a combination of existing and new household and intercept traveler surveys collected in California and combined with intraregional trips generated from regional and statewide sources.

The validation work included the calibration process, development of data used for observed travel behavior, and documentation of the resulting calibration parameters for the interregional trips. In addition, this work included summaries and reasonableness checks on the intraregional trips derived from the MPO trip tables. These were not separately validated or calibrated, because each MPO has provided assurances that these trip tables were validated.

Trips by mode from the interregional models were combined with intraregional trips by mode to assign to the highway, air, and rail networks. Table 6 presents a summary of the 2000 interregional trips by mode and market.

Highway trips were converted from person trips to vehicle trips using vehicle occupancy factors derived from the Caltrans Statewide Travel Survey. In addition, highway trips were separated into peak and off-peak time periods so that peak and off-peak trip tables could be assigned separately to the highway network. This ensures that peak-period travel times would more accurately reflect congestion that occurs in the peak period.

Following the development of peak and off-peak auto vehicle interregional trips, these were combined with the auto vehicle intraregional trips. These intraregional trips come from four sources: MTC, SANDAG, SCAG, and Caltrans. The Caltrans Statewide Model is used to estimate intraregional trips for all the other regions (except MTC, SANDAG, and SCAG) so that the auto trip table will be representing all statewide travel. This ensures that congestion within each smaller urban area is adequately represented.

Validation of the base year assignments by mode involved detailed review of observed and modeled volumes. For air, these reviews focused on assignments for the major markets. For rail, these reviews focused on assignments by operator. For highway, these reviews focused on assignments by gateway and by region. A summary of the assignments by mode is provided in Table 7.

Table 6. 2000 Daily Interregional Trips by Mode

Market	Auto	Air	Rail	Total	Percent of Total
LA to Sacramento	7,479	4,935	-	12,414	1%
LA to San Diego	257,441	100	5,395	262,936	17%
LA to SF	28,031	26,867	-	54,898	4%
Sacramento to SF	137,739	25	1,816	139,580	9%
Sacramento to San Diego	175	2,858	-	3,033	0%
San Diego to SF	4,630	10,309	-	14,939	1%
LA/SF to SJV	205,205	3,393	926	209,524	14%
Other to SJV	281,750	243	344	282,337	19%
To/From Monterey/ Central Coast	275,794	3,532	1,105	280,431	19%
To/From Far North	184,506	3,005	16	187,527	12%
To/From W. Sierra Nevada	59,192	668	11	59,871	4%
Total	**1,441,942**	**55,935**	**9,613**	**1,507,490**	**100%**
Percent of Total	**95.7%**	**3.7%**	**0.6%**	**100%**	

Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Final Report, Table 5.1, July 2007.

Table 7. 2000 Daily Assignments by Mode

Mode	Units	Observed	Model[4]	Difference	Percent Difference
Air	Boardings	54,271[1]	54,876	605	1%
Rail	Boardings	16,710[2]	17,743	1,033	6%
Auto	Vehicle Counts	27,145,300[3]	25,206,373	(1,938,927)	-7%

[1]Source: U.S. Department of Transportation FAA O&D ten-percent sample database [2]Source: Interregional rail operators and MTC [3]Source: Caltrans, MTC and SCAG traffic count databases Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Final Report, Table 5.2, July 2007.

Even though the air and rail assignments were very small compared to auto, these were critical to the evaluation of high-speed rail, so a great attention to the validation of these modes was important. For the major markets and operators, these compared very well with observed numbers. Auto assignments were primarily validated based on gateways along the high-speed rail corridors. These compared very well to observed traffic counts. Additional validation effort to refine and improve the highway assignments is recommended if this model were to be used for highway planning purposes.

Comparison of the 2030 forecast to a No-Build scenario was completed for validation to ensure that the 2030 forecasts are reasonable for each model component. Overall, there is a 42 percent increase in households and a 51 percent increase in employment, and there is a 62 percent increase in interregional trips. The 2030 interregional trip table is presented in Table 8.

The higher percent of interregional trips compared to statewide household and employment growth is a reflection of the expansion of the regions beyond their regional borders, causing more travelers to make interregional travel instead of intraregional travel. The auto assignments (represented by total vehicle miles traveled) increase by 73 percent from 2000 to 2030, which is also caused by travelers having to go further to reach their destinations. These are presented in Table 9.

Table 8. 2030 Daily Interregional Trips by Mode

Market	Auto	Air	Rail	Total
LA to Sacramento	12,636	8,105	–	**20,741**
LA to San Diego	340,862	96	25,898	**366,856**
LA to SF	30,253	25,351	–	**55,604**
Sacramento to SF	174,844	26	11,798	**186,668**
Sacramento to San Diego	164	5,258	–	**5,422**
San Diego to SF	5,038	18,259	–	**23,297**
LA/SF to SJV	360,177	9,609	6,237	**376,023**
Other to SJV	553,466	1,944	4,792	**560,202**
To/From Monterey/Central Coast	426,056	5,886	2,077	**434,019**
To/From Far North	320,667	5,957	962	**327,586**
To/From W. Sierra Nevada	96,404	1,177	335	**97,916**
Total	**2,320,567**	**81,668**	**52,099**	**2,454,334**

Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Final Report, Table 5.3 July 2007.

Table 9. 2000 and 2030 Assignments by Mode

Mode	Units	2000 Model	2030 Model	Difference	Percent Difference
Air	Boardings	54,876	80,643	25,767	47%
Rail	Boardings	16,430	30,653	14,222	87%
Auto	Vehicle Miles Traveled	748,606,510	1,297,116,168	548,509,657	73%

Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Final Report, Table 5.4, July 2007.

Rail boardings increase at a higher rate than auto, indicating that as congestion increases; more travelers are taking rail, as expected. Air boardings do not increase as fast as rail or auto because the air fares increased and frequencies decreased between 2000 and 2005, making air a less attractive option. The 2005 observed air level of service was kept constant through 2030. The primary reason for significant changes in air service from 2000 to 2005 was the September 11 terrorist attacks in 2001, which affected air travel more than other modes.

5.2 Forecast Results

Table 10 presents a summary of the trips by mode and mode shares for the base year (2000) and the future year (2030) with and without the high-speed rail project. High-speed rail captures over 7 percent of the trips and draws from all other modes.

5.3 Sensitivity Tests

A series of sensitivity tests were conducted to test the impacts of changes in level of service on high-speed rail ridership and revenue. These tests were designed to assist in developing an improved operating plan, optimum fares, and to understand the impacts of potential changes in assumptions to the air and auto modes. The results of the sensitivity tests are provided in Table 11.

Table 10. Summary of Trips and Mode Shares for Base and Future Conditions

	2000 Base Year		2030 without HSR		2030 with HSR		2030 Difference	
	Trips	Mode Share	Trips	Mode Share	Trips	Mode Share	Trips	Pct of Total
Auto	1,441,942	95.7%	2,320,567	94.5%	2,193,248	89.2%	-127,319	-71%
Air	55,935	3.7%	81,668	3.3%	53,823	2.2%	-27,845	-16%
Rail	9,613	0.6%	52,099	2.1%	31,790	1.3%	-20,309	-11%
HSR					179,482		179,482	100%
Total	**1,507,490**	100.0%	**2,454,334**	100.0%	**2,458,343**	100.0%	**4,009**	

Note: The 4,009 difference in 2030 trips with and without HSR demonstrates how much induced travel is a result of HSR. Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Model Validation Report, Table 7.5, July 2007 and Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Ridership and Revenue Forecasts, Table 2.2, August 2007.

Table 11. Sensitivity Tests for High-Speed Rail

Sensitivity Test	Change in Level of Service	Percent Change from Base	
		Boardings	Revenues
High-Speed Rail Level of Service Tests			
Higher HSR Fares	25% increase	-13%	2%
Average Daily Headways	HSR headways[1]	-15%	-14%
Higher HSR Freq	100% increase	15%	16%
Express Service SF/LA	Double Freq SF/LA to SJV, SD/SF to SAC	22%	24%
Air and Auto Level of Service Tests			
Higher Air/Auto Times	6% increase[2]	6%	6%
Higher Air/Auto Costs	50% increase	46%	53%
Combined Level of Service Tests			
Higher HSR Fares and Higher Air/Auto Costs	25% increase in fares, 50% increase in costs	13%	19%
Higher HSR Fares and Higher Air/Auto Costs	50% increase in both	31%	40%
Higher HSR Fares and Higher Air/Auto Costs	100% increase in fares, 50% increase in costs	-6%	1%

[1] Average daily headways assume that the headway in the peak and off-peak periods are equal. This effectively increases peak headways and decreases off-peak headways. [2]The 6 percent increase in travel time was based on a 30-minute increase in travel time from San Francisco to Los Angeles by car. Source: Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Final Report, Table 7.1, July 2007.

The results show that improvements in high-speed rail frequencies can support much higher high-speed rail ridership; increased high-speed rail frequencies in the major corridors (San Francisco to Los Angeles, Los Angeles to San Joaquin Valley, San Diego to Sacramento, and San Francisco to Sacramento) were then retained for the alternatives analysis. These results also show that raising high-speed rail fares will not significantly increase revenues, unless this is combined with different assumptions of air and auto costs. Assumptions regarding air and auto cost increases remain a difficult issue, given the volatility in these costs in the past 5 years alone. The sensitivity tests do show that high-speed rail ridership is highly sensitive to the assumptions of air and auto costs and can increase as much as 46 percent with a 50 percent increase in air and auto costs, which seems quite reasonable compared to current trends in these costs.

6 Summary

The travel forecasting models developed for predicting high-speed rail alternatives for the state of California have several immediate benefits over previous ridership forecasting methods used in the state: they are network-based and provide more accurate assessments of time and cost tradeoffs with other modes, modal choices are sensitive to reliability, party size, and detailed access and egress options, induced travel is assessed based on changes in level of service for all modes, and intraregional travel is estimated based on detailed urban area models where interregional travel is estimated based on statewide models estimated from observed travel behavior. The intraregional and interregional models are integrated to assess impacts of congestion on other modes and to reflect differences in peak and off-peak conditions.

The primary advancement in this model is the additional level of detail (4,600 zones used for all modeling components without sampling), the inclusion of peak and off-peak assignments, and the consistent use of logsum accessibility measures at all levels of the models (from access and egress models up to trip frequency models). These models were estimated using revealed preference data and by combining multiple survey datasets, a more robust estimation dataset was possible.

There are some areas where these models may be improved for other statewide and regional planning activities. The trip frequency models could benefit from additional data on weekly or monthly long distance travel, because a one-day snapshot does not provide as strong a basis for travel decisions as longer-term data would provide. The destination choice models could also be improved by including data on special generators, such as Disneyland. Lastly, the mode choice models could benefit from a tour-based methodology, recognizing that decisions on mode are affected by both the outbound and return portions of the trip. In these cases, the models could benefit from additional data and resources that were beyond the original scope of the project.

These integrated statewide models offer a comprehensive tool to forecast long and short distance travel in California. The separation of travel into market segments based on distance (short and long), purpose (business, commute, recreation and other) and travel markets (inter- and intraregional) provide a robust and accurate assessment of multimodal travel at the statewide level.

7 References

Algers, S., 1993, Integrated Structure of Long-Distance Travel Behaviour Models in Sweden, *Transportation Research Record 1413*, Transportation Research Board, 141-149.

Ben-Akiva, M. and Morikawa, T., 1990, Estimation of Switching Models from Revealed Preferences and Stated Intentions, *Transportation Research Part A*, 24(6), 485-495.

Bhat, C.R., 1995, A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice, *Transportation Research Part B*, 29(6), 471-483.

Bhat, C.R., 1998, Accomodating Variations in Responsiveness to Level-of-Service Measures in Travel Choice Modeling', *Transportation Research Part A*, 32(7), 49-57.

Bhat, C.R., 1997, Covariance Heterogeneity in Nested Logit Models: Econometric Structure and Application to Intercity Travel, *Transportation Research Part B*, 31(1), 11-21.

Booz-Allen & Hamilton, 1989, Demand 'Model Estimation: Final Report', prepared for AMTRAK.

Bradley, M.A. and Daly, A.J., 1997, Estimation of logit choice models using mixed stated preference and Revealed Preference Information', in Stopher, P.R. and Lee-Gosselin, M. (Eds.) *Understanding Travel Behaviour in an Era of Change*, Oxford: Pergamon, 209-232.

Brand, D., Parody, T.E., Hsu, P.S. and Tierney, K., 1992, Forecasting High-Speed Rail Ridership', *Transportation Research Record 1342*, 12-18.

Cambridge Systematics with Mark Bradley Research and Consulting, 2006, Bay Area/California High-Speed Rail Ridership and Revenue Forecasting Study: Interregional Model System Development Report', prepared for the Metropolitan Transportation Commission and the California High-Speed Rail Authority.

http://www.cahighspeedrail.ca.gov/images/chsr/20090403111619_R5a_Model%20System%20Development_Final.pdf

Cambridge Systematics with Mark Bradley Research and Consulting, 2007, Bay Area/California High-Speed Rail Ridership and Revenue Forecasting Study: Ridership and Revenue Forecasts Report', prepared for the Metropolitan Transportation Commission and the California High-Speed Rail Authority. http://www.cahighspeedrail.ca.gov/images/chsr/20090403113249_R8b_FINAL_Ridership_Report_2007-26-10_AXK.pdf

Cambridge Systematics with Mark Bradley Research and Consulting, 2007, Bay Area/California High-Speed Rail Ridership and Revenue Forecasting Study: Final Report', prepared for the Metropolitan Transportation Commission and the California High-Speed Rail Authority. http://www.cahighspeedrail.ca.gov/images/chsr/20090403113328_R9a_Final_Report_FINAL.pdf

Charles River Associates, 2000, Independent Ridership and Passenger Revenue Projections for High Speed Rail Alternatives in California, prepared for the California High-Speed Rail Authority.

Cohen, G., Earlbaum, N.S. and Hartgen, D.T., 1978, Intercity Rail Travel Models', *Transportation Research Record 673*, 1978, 21-25.

DeFerranti, D., Goeller, B.F. and Chesler, L.G., 1973, The STAR Methodology for Short-Haul Transportation: Demand Prediction – Approach and Calibration', Rand Corporation Report R-1359/2-DOT, prepared for the Department of Transportation.

Federal Railroad Administration website 2006, http://www.fra.dot.gov/us/content/618

Forinash, C. and F.S. Koppelman 1993, Application and Interpretation of Nested Logit Models of Intercity Mode Choice, *Transportation Research Record 1413*, 98-106.

Forinash, C., 1992, A Comparison of Model Structures for Intercity Travel Mode Choice', M.S. Thesis, Northwestern University.

Grayson, A., 1981, Disaggregate Model of Mode Choice in Intercity Travel, *Transportation Research Record 835*, 36-42.

Gunn, H.F., Bradley, M. and Hensher, D.A., 1992, High-speed rail market projection: Survey design and analysis, *Transportation 19*, 117-139.

Horowitz, Alan 2006, NCHRP Synthesis 358: Statewide Travel Forecasting Models, *Transportation Research Board*, Appendix C.

Koppelman, F.S. and Hirsh, M., 1986, Intercity Passenger Decision Making: Conceptual Structure Data', *Transportation Research Record 1085*, 70-75.

Koppelman, F.S., 1989, Multidimensional Model System for Intercity Travel Choice Behavior', *Transportation Research Record 1241*, 1-8.

Koppelman, F.S., Kuah, G. and Hirsh, M., 1984, Review of Intercity Passenger Demand Modeling: Mid-60's to Mid-80's', The Transportation Center, Northwestern University.

Miller, E., 1992, Intercity Passenger Travel Demand Modelling: Present State and Future Possibilities, Canadian Transportation Research Forum 27th Annual Meeting Proceedings, University of Saskatchewan, 378-389.

Peers, J.B. and Bevilacqua, M., 1976, Structural Travel Demand Models: An Intercity Application, *Transportation Research Record 569*, 124-135.

PRC Voorhees, 1983, Florida High Speed Rail Technical Memorandum: Ridership, prepared for Bombardier, Inc.

Proussaloglou, K.E. and Tierney, K., 1999, An Integrated Disaggregate Model System for Intercity Travel, presented to the 7th TRB Conference on the Application of Transportation Planning Methods, Boston.

Quandt, R.E. and Baumol, W.J., 1966, The Abstract Mode Model: Theory and Measurement, *Journal of Regional Science*, 6(2), 13-26.

Thakuriah, P., Virmani, D., Yun, S. and Metaxatos, P., 1999, Estimates of the Demand for Inter-City Travel: Issues with Using the American Travel Survey, *Transportation Research Board Conference Proceedings- Personal Travel: The Long and Short of it*, June 28-July 1, 1999, Washington DC, 255-269.

TMS/Benesch High Speed Rail Consultants, 1991, Tri-State Study of High Speed Rail Service: Chicago-Milwaukee-Twin Cities Corridor, prepared for Illinois, Minnesota, and Wisconsin Departments of Transportation.

Yu, J.C., 1970, Demand Model for Intercity Multimode Travel, *The Transportation Engineering Journal of ASCE*, 96(TE2), 203-218.