2016 California High-Speed Rail Business Plan Ridership and Revenue Risk Analysis

technical

report

prepared for

California High-Speed Rail Authority

prepared by

Cambridge Systematics, Inc.

April 8, 2016

technical report

2016 California High-Speed Rail Business Plan Ridership and Revenue Risk Analysis

prepared for

California High-Speed Rail Authority

prepared by

Cambridge Systematics, Inc.
555 12th Street, Suite 1600
Oakland, CA 94607

date

April 8, 2016

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Forecasts of California high-speed rail (HSR) ridership and revenue are estimated using a travel demand model. The Business Plan Model – Version 3 (BPM-V3) travel demand model predicts, for a specified forecast year, the number of annual trips made by households residing in California, where these trips are going within California, and the mode of transportation (i.e., auto, air, conventional rail, or HSR) used to make these trips. In order to predict this travel behavior, the model takes in as input predictions about what households and the transportation system will look like for each forecast year. However, there is uncertainty in what the future will actually hold for these input values (e.g. auto operating cost), as well as how travel patterns and people's choices will evolve over time. Thus, to fully understand the uncertainty in the HSR forecasts of revenue and ridership, the full range of probable values for these variables should be analyzed.

The purpose of the risk analysis is to incorporate the uncertainty associated with model inputs and assumed travel behavior into the 2016 Business Plan (BP) HSR ridership and revenue forecasting process. A risk analysis approach was developed that expresses forecast results as probabilities of achieving different outcome levels. This approach builds on and expands the previous risk analysis that was performed for the 2014 Business Plan (BP). In order to develop a full range of possible ridership and revenue, dozens of full model runs were used to estimate relationships between forecast revenue and ridership and select input variables (This created a "meta-model."). The approach began by identifying potential risk factors that could impact ridership and revenue forecasts (e.g., potential changes in auto operating costs or the impact of new technologies such as autonomous vehicles). These factors were estimated using model variables, and the variables were systematically narrowed to the set of inputs that would have the highest combination of uncertainty and impact on the forecasts. The "meta-model" was coupled with researched distributions of the model inputs and used in a Monte Carlo simulation to develop 50,000 unique forecasts of revenue and ridership. From this simulation, probability distributions of total revenue and ridership were estimated.

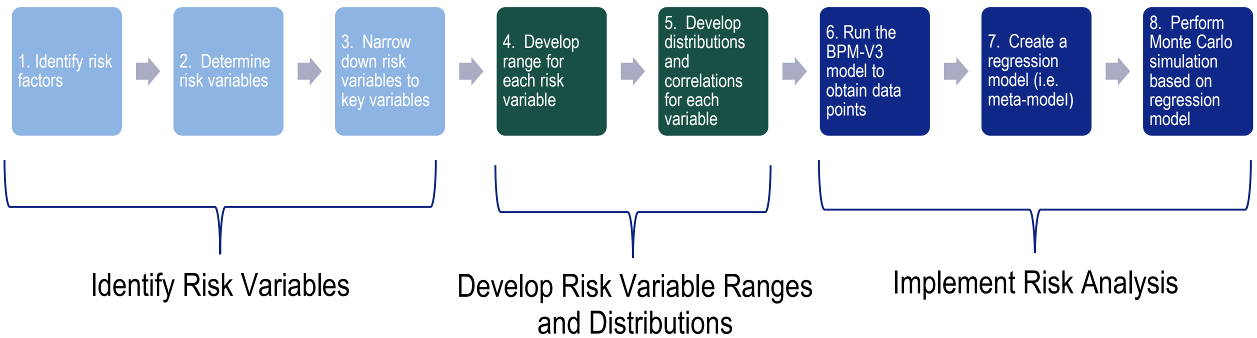
This methodology is similar to the methodology employed for the 2014 BP. It was refined and enhanced by:

- Refining the process for identifying the model inputs for inclusion in the analysis by developing a systematic and transparent methodology;
- Incorporating more model inputs and subcomponents within the model inputs, tailored to each forecast year and implementation step in the analysis, so that the exact uncertainty in each forecast can be examined; and
- Using a two-step analysis process to test interaction effects and to ensure that the full model runs replicated the BPM-V3 model for all levels of input values.

1.1 Overview of the Risk Analysis Approach

An eight-step risk analysis approach was employed to forecast revenue and ridership for the 2016 BP, as shown in Figure 1.1.

Figure 1.1. Eight-Step Risk Analysis Approach



The steps to identify the model assumptions are described below.

Step 1. Develop a list of possible risk factors to be considered for the revenue and ridership risk analysis

- Risk factors are defined as any circumstance, event, or influence that could result in the HSR revenue and ridership deviating from its forecasted value;
- A panel of experts was used to develop a set of potential risk factors that could impact future HSR ridership and revenue; and
- The identified risk factors differed between forecast years.

Step 2. Identify risk variables for each risk factor

- Risk variables are actual variables and constants that can be adjusted in the BPM-V3. As an example, auto operating cost (i.e., cost, in dollars, per vehicle mile driven) is a risk variable that can be adjusted in the model. To address the possibility that fuel cost and fuel efficiency may be higher or lower than predicted, auto operating cost may be increased or reduced in the risk analysis to test how these two risk variables affect ridership and revenue.
- The risk variables have been chosen to represent one or more risk factors identified in Step 1.

Step 3. Narrow risk variables to key variables for inclusion within each forecast year of analysis

- Sensitivity runs of the BPM-V3 were performed for each risk variable that allowed for a quantitative comparison of the impacts of each risk variable on ridership and revenue; and
- Based on the range and known sensitivity of the risk variables under consideration, a final set of 10 risk variables was selected for inclusion for each forecast year.

Steps 4 and 5. Develop a range and distribution for each risk variable under consideration

- The uncertainty associated with each risk variable was quantified by assigning a range and distribution for each variable. For example, based on the research on each risk factor affecting auto operating cost, such as fuel cost and fuel efficiency, auto operating cost in year 2025 is predicted to range from $0.15 per mile to $0.31 per mile, with a most likely value of $0.20 per mile.
- For each risk variable, the minimum, most likely, and maximum values for each forecast year were developed based on currently available research and analysis.
- The shape of the distribution for each variable determined the likelihood of the variable's value, within the set range, under random sampling. For example, it is very unlikely that auto operating cost will be the minimum value of $0.15 per mile or the maximum value of $0.31 per mile, but very likely it will be close to $0.20 per mile. The auto operating cost distribution is defined such that the most likely value will be chosen at a much higher rate than the extreme values, and thus the simulated model runs will be more representative of potential future outcomes.

Steps 6 and 7. Run the BPM-V3 using a defined set of risk variable levels to obtain data points for estimation of two sets of regression models (i.e., meta-models) that regresses the 10 risk variables on either HSR revenue or ridership

- The set of BPM-V3 specified model runs were developed to:
  - Test for the presence of two variable interaction effects,
  - Estimate nonlinearity of model variables,
  - Adequately capture the boundaries of the solution space, and
  - Ensure that data points do a good job of representing the interior of the solution space.

Step 8. Perform a Monte Carlo simulation by running the regression model 50,000 times with varying levels of the input variables based on the distributions assigned to the variables

- The simulation results in probability distributions of HSR revenue and ridership.
- The results of the simulation were analyzed to determine the relative contribution of each risk factor on revenue and ridership.

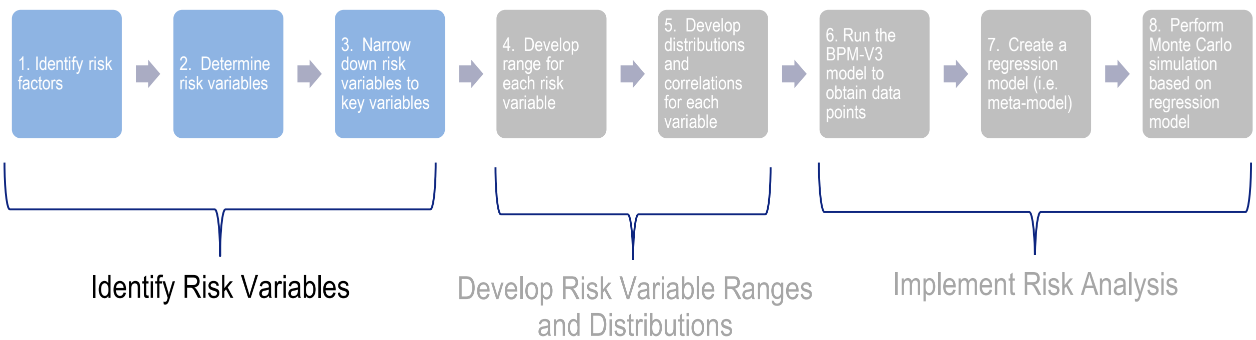
Each step in the risk analysis required thorough evaluation to ensure key risk factors were understood and addressed appropriately. The rest of this technical memorandum is divided into five sections that provide insight into the steps taken to produce the simulation results:

- **Section 2.0. Identification of Risk Variables** (Steps 1 to 3);
- **Section 3.0. Development of Risk Variable Ranges and Distributions** (Steps 4 to 5);
- **Section 4.0. Risk Analysis implementation** (Steps 6 to 8); and
- **Appendices A to J**.

2.0 Identification of Risk Variables

This section details the steps taken to identify the risk variables included in the risk analysis, as shown in Figure 2.1.

Figure 2.1 Eight-Step Risk Analysis Approach: Identify Risk Variables (Steps 1 to 3)



To develop a set of potential risk factors (*Step 1*), Cambridge Systematics, Inc. (CS) started by holding a series of meetings among Rail Delivery Partner (RDP) and Authority staff to brainstorm and identify potential risks that could impact ridership and revenue forecasts. The meetings sought to answer the following question: What real-world risks could impact ridership and revenue in years 2025, 2029, and 2040? These forecast years were chosen based on the HSR-phased implementation analyzed for the 2016 Business Plan, which is outlined in Table 2.1. As a result, the list of risk factors identified differed depending on the operating plan and forecast year under consideration. For example, the uncertainty and impact of HSR bus connections to actual HSR service are a concern for earlier years, while the likelihood of significant autonomous vehicle use affecting HSR ridership is not likely until 2040.

Table 2.1 Description of each Phase of the HSR System

Operating Phase	Year	High-Speed Rail Segment	Frequency of Service	HSR Bus and Conventional Rail Connections
Silicon Valley to Central Valley Line (VtoV)	2025	San Jose to North of Bakersfield	2 trains per hour during the peak period and 1 train per hour during the off-peak period	Includes Caltrain connections between San Jose and San Francisco, bus connections between North of Bakersfield and Los Angeles and rail and bus connections between Fresno and Sacramento
Phase 1 (PH1)	2029 & 2040	San Francisco and Merced to Los Angeles and Anaheim	Up to 8 trains per hour (from all destinations) during the peak period and 5 trains per hour during the off-peak period	Includes rail and bus connections from Merced to Sacramento and rail connections in Southern California

This list generated was used to identify risk variables (i.e., assumptions built into the BPM-V3 model) that could represent each risk factor (*Step 2*). The risk variables identified for each risk factor were determined by answering the following questions: What model inputs and variables drive these risks, and how do we account for these risks in the model? Sensitivity runs of the BPM-V3 model were run for each risk variable that allowed for a quantitative comparison of the impacts of each risk variable on ridership and revenue. Based on this sensitivity analysis, the risk variables that were determined to have the greatest effect on HSR ridership and revenue and the highest potential uncertainty for each forecast year were selected for inclusion (*Step 3*). A set of 10 risk variables was included in the risk analysis for each forecast year, as shown in Table 2.2. This table also documents the risk factors that are represented by each risk variable. Appendix A of this technical memorandum gives more detail, for each included risk variable, the reason for considering the model variable, the risk factors represented by the variable, and the sensitivity results and quantitative reasoning for including the variable. The appendix also highlights the list of risk analysis variables that were considered, but eventually excluded from the analysis.

Table 2.2 Variables Included in Risk Analysis for Each Analysis Year

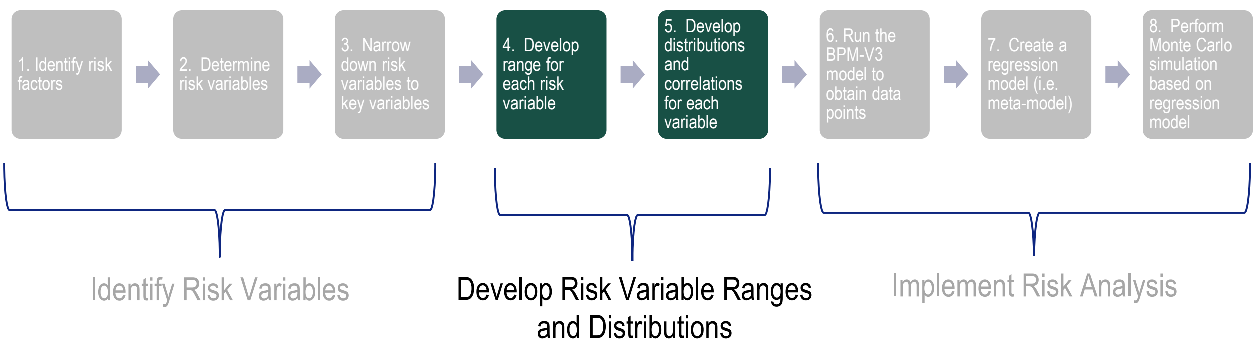
Number	Risk Variable	Reasons for Considering Model Variable and Risk Factors Represented
1	Business HSR Mode Choice Constant	The mode constants capture the unexplained variation in traveler mode choices after system variables and demographics are taken into account. Unexplained variation may include factors, such as comfort aboard trains, opinions regarding HSR, need for a car at the destination, level of familiarity with HSR, etc.
2	Commute HSR Mode Choice Constant	
3	Recreation/Other HSR Mode Choice Constant	
4	Business/Commute Trip Frequency Constant	The trip frequency constants capture the unexplained variation in the number of long-distance trips that travelers will take after accounting for household demographics and the accessibility of available destinations. Also, risks associated with the state of the economy are accounted for within the trip frequency constant risk variable.
5	Recreation/Other Trip Frequency Constant	
6	Auto Operating Costs	This variable reflects the inherent risks in forecasting future: fuel costs, fuel efficiencies, adoption of alternative fuels/electric vehicles, maintenance costs, changes in gas taxes, potential impacts of cap and trade on fuel costs, market penetration of autonomous connected vehicles, and higher shares of "shared use" vehicles.
7	HSR Fares	A number of issues could affect actual fares charged to travelers, especially as the system is being opened: institution of discount/premium fares (advance purchase, peak/off-peak, first/second class seating); adjustments needed to respond to changing auto operating costs or air fares; yield management strategies; etc.
8	HSR Frequency of Service	With final service plans expected to be developed by a private operator that has not been brought on board yet, there is uncertainty around the amount of service that will be provided based on the markets and strategies that the operator may employ.
9 (Year 2025)	Availability and Frequency of Service of Conventional Rail and HSR Buses that connect with HSR	Access to and egress from the system includes connections with both conventional rail services and HSR buses (as well as many other modes). Levels of conventional rail service are assumed based on the State Rail Plan, but there is some uncertainty around the availability of the exact amount of conventional rail service. Similarly, the amount of connecting bus service could be different than currently forecasted. These connections are most critical in the early years of the program when the high-speed rail system does not yet connect the whole State.
9 (Year 2029)	Airfares	Airfares change and fluctuate over time. Some possible reasons that airlines may change airfares from currently forecasted levels include changes in fuel or personnel costs or airport landing fees; changes in equipment or efficiency, such as NextGen technology; competitive response to HSR to maintain air market shares; acceptance of HSR as a replacement for inefficient; short-haul air service; etc.

Number	Risk Variable	Reasons for Considering Model Variable and Risk Factors Represented
10 (Year 2025 and Year 2029)	Coefficient on Transit Access-Egress Time/Auto Distance Variable	Between some regions in California, especially in the Silicon Valley to Central Valley line scenario, individuals who wish to travel primarily by transit to reach their destination must transfer from an HSR bus or conventional rail system before or after traveling on HSR. International experience has shown that there is uncertainty around how the need to make these transfers affects overall HSR ridership. The model includes a variable that makes HSR less attractive for trips that require a long access or egress trip in relation to the time spent on HSR (or another public mode such as conventional rail or air), and the variation in this variable was used as a way to estimate the uncertainty around the effect of these transfers on HSR ridership and revenue.
9 (Year 2040)	Number and Distribution of Households (HH) throughout the State	The forecasted number of statewide HHs can fluctuate for a variety of reasons, such as inherent uncertainty with population forecasts; national and statewide economic cycles; impacts of natural disasters, such as continuing draught; changes in U.S. immigration policy; etc. The uncertainty of population forecasts and the divergence between different forecasts increase the further out the forecasts make predictions. Based on a review of nine forecasts for 2020, the differences in predicted California population between the lowest and highest forecasts were only 840,000, while the differences for 2040 were 2.4 million.
10 (Year 2040)	Auto Travel Time	The introduction of autonomous vehicles is represented by decreases in auto travel times included within the model.

3.0 Development of Risk Variable Ranges and Distributions

To conduct the risk analysis, the uncertainty surrounding each risk variable must be quantified by assigning a range and distribution for each variable. As shown in Figure 3.1, determining the ranges of the risk variables corresponds to *Step 4*, and developing the distributions corresponds to *Step 5* of the risk analysis approach.

Figure 3.1 Eight-Step Risk Analysis Approach: Develop Risk Variable Ranges and Distributions (Steps 4 to 5)

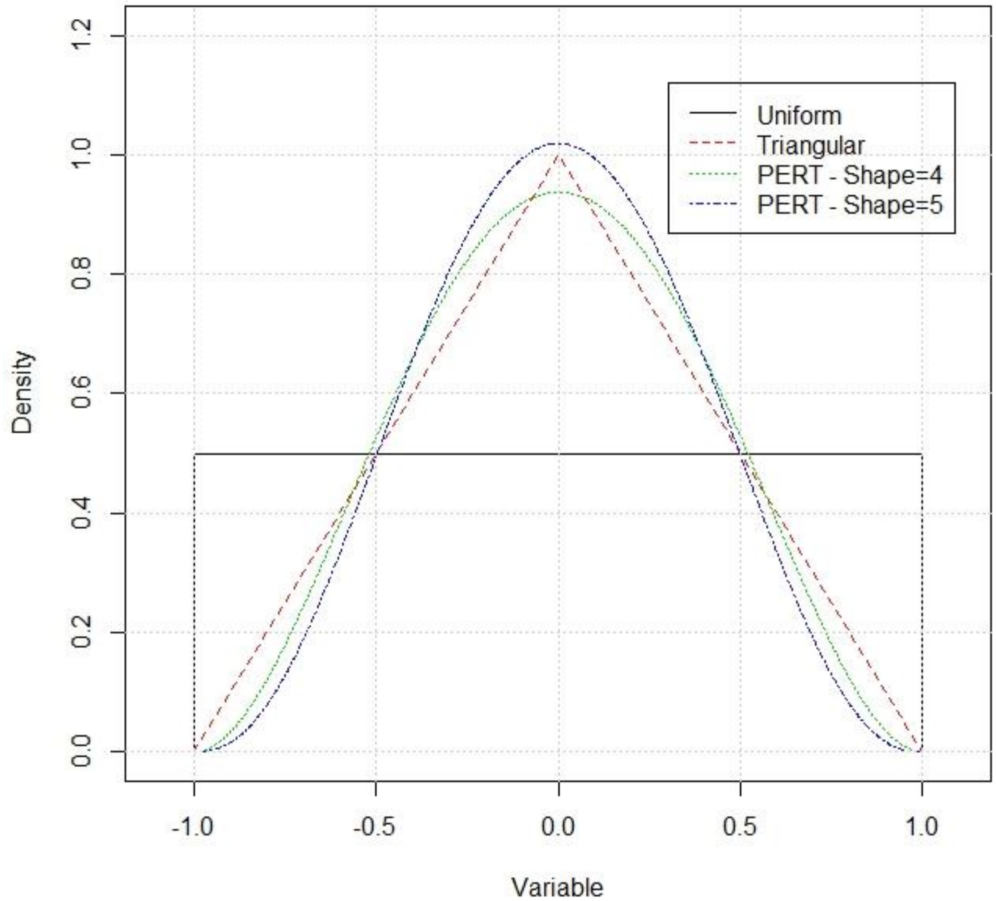


To perform the risk analysis, a range of possible values for each risk variable has to be established in order to quantify the uncertainty related to that variable. The absolute minimum and absolute maximum value of the variable sets the range of the variable's forecasted value, while the most likely represents the peak of the variable's distribution. For each risk variable, the absolute minimum, most likely, and absolute maximum values were driven by independent research and analysis.

A distribution around the minimum, most likely, and maximum values of each risk variable was determined based on the characteristics of these three points. The shape of the distribution determines the likelihood of the variable's value, within the set range, under random sampling. The most likely value has the greatest likelihood of occurring within the distribution. The shape of the distribution can be triangular, PERT, uniform, or another form. PERT distributions were used for variables where there are significant tails based on the values assumed for the minimum and maximum (i.e., the minimum and maximum are extreme values). A Shape = 4 PERT distribution was assumed to be standard with a higher Shape used for Auto Operating Costs, because the maximum and minimum involve several independent downside or upside events taking place at the same time, which results in a highly unlikely event (and longer tails). Triangular distributions were used where there is less information about

the exact shape, but values around the most likely are more likely to occur than the values closer to the minimum and maximum (though not to the same extreme as for the PERT distributions). Uniform distributions were used where there is high uncertainty regarding the forecast values for the risk variable. Figure 3.2 illustrates the shapes of the different distributions used with the risk analysis.

Figure 3.2 Shapes of the Distributions Used in the Risk Analysis



The following sections identify the ranges of values and distribution for each risk variable and summarize the research and methodology for developing the absolute minimum, most likely, and absolute maximum value.

3.1 HSR CONSTANT

The HSR constant for each of the four trip purposes (i.e., business, commute, recreation, and other) is composed of two components: 1) unexplained variation, and 2) terminal and wait time. The unexplained variation component represents the desirability to choose HSR that is not captured directly by the system variables (e.g., travel time, cost, etc.) included in the model. Terminal time is the out-of-

vehicle time spent traveling from the point of departure from the access mode to the train platform. Wait time is the out-of-vehicle time spent waiting on the platform for the train to arrive and the time spent waiting for the train to leave the platform once boarded. For all forecast years, the range for the HSR constant was defined as:

- **Minimum**. At the very worst case, HSR will be perceived as an equivalent mode to Conventional Rail (CVR), but terminal + wait time will be 45 minutes;
- **Most Likely.** Calibrated HSR constant with terminal + wait time of 25 minutes; and
- **Maximum**. Assumes unexplained variation is symmetrical, but terminal + wait time will be 15 minutes.

For BPM-V3 runs, terminal and wait time are included with the unexplained variation within the HSR constant. For Monte Carlo risk analysis, each component of the HSR constant is considered as a separate risk variable with completely independent distributions (i.e., the unexplained variation uses a PERT distribution while the terminal/wait time uses a triangular distribution). The former allows for estimation of a single regression model parameter, and does not require an additional risk variable in the experimental design framework. The latter allows for an understanding of the terminal/wait time's effect on ridership and revenue uncertainty independent from the HSR constant's effect on ridership and revenue uncertainty since the two variables do not necessarily move together. More information on the development of the range and distribution of the components of the HSR constant are detailed in Appendix B.

3.2 Trip Frequency Constant

The trip frequency constants include the unexplained variation in the propensity of households to make long-distance trips within California. Within the risk analysis, variation in the trip frequency constants (i.e., business/commute and recreation/other) also is developed to reflect the effect of the state of the economy on the proclivity of households to make trips. While "the economy" is an overarching risk that affects many different decisions regarding travel, one of the most direct and principal impacts on HSR ridership and revenue is whether a long-distance trip is even made. The state of the economy affects household income and employment levels; and the level of these variables directly affects trip frequency rates within the model (i.e., people who are out of work or have reduced income due to a recession make fewer long-distance trips). Instead of including these variables directly as a risk variable in the risk analysis model to account for changes in the state of the economy, the effects of these risk variables on trip frequency levels are accounted for within the trip frequency constant risk variable.

The unexplained variation range is based on the range seen in forecasted annual long-distance trip rates produced by the model. The most likely value for each forecast year is the calibrated constant. The minimum value of the trip frequency

constants is specified such that for year 2040, the trip frequency constants produce average trip rates equal to the 2010 rates by trip purpose long-term trends show people's propensity for making long-distance trips increasing over time). For the maximum value, the trip frequency constant is specified to mirror the deviations from the calibrated constants for the minimum values (i.e., symmetry of the constant offsets are assumed).

The range of constant offsets for the Economic Cycles provides proxies for the actual economic-cycle risk variables being considered. This approach provides a method for specifying a continuous range of outcomes rather than developing multiple input socioeconomic datasets. The economic-cycle range was developed by calculating the implied trip rates based on changes in the number of workers and income levels from the following scenarios:

- **Minimum**. Based on HSR implied trip rate decrease resulting from a three percent per year decrease in employment from the low-growth scenario for three years preceding the forecast year. The direct impact of the low economic cycle on trip frequency is determined by changing the distributions of households by number of workers and households by income group to reflect the three percent per year decrease in employment.
- **Most likely**. Resulting trip rates obtained using calibrated trip frequency constants.
- **Maximum**. Based on HSR implied trip rate increase resulting from a three-percent per year increase in employment from the high-growth scenario for five years preceding the forecast year. The direct impact of the high economic cycle on trip frequency is determined by changing the distributions of households by number of workers and households by income group to reflect the three percent per year increase in employment.

The offsets must be combined to represent the full range of possible outcomes for the development of the risk analysis regression equations. The constant offsets for the Unexplained Variation and Economic Cycle are added together, and the implied ranges of annual long-distance round trips per capita were estimated as shown in Table 3.1.

Table 3.1 Ranges of Implied Annual Round Trips per Capita for Full Model Runs Based on Trip Frequency Constant Offsets

Model Year	Purpose	Implied Annual Long-Distance Round Trips per Capita After Applying Offsets		
		Minimum	Most Likely	Maximum
2025	Business/Commute	1.30	2.16	3.35
	Recreation/Other	4.76	5.76	6.84
	Total	6.06	7.92	10.19
2029	Business/Commute	1.37	2.20	3.62
	Recreation/Other	4.84	5.85	7.13
	Total	6.21	8.05	10.75
2040	Business/Commute	1.45	2.44	3.97
	Recreation/Other	5.06	6.22	7.54
Total		**6.51**	**8.66**	**11.51**

For Monte Carlo risk analysis, each component of the trip frequency constant is considered as a separate risk variable with completely independent distributions (i.e., 0 percent correlation). The unexplained variation uses a PERT distribution, while the economic cycle uses a triangular distribution. A 50-percent correlation is assumed between the business/commute and recreation/other risk components for unexplained variation, since there is likely to be some relationship (though not perfect correlation) in changes to overall trip-making for different purposes. Perfect correlation is assumed between economic-cycle risk components for business/commute and recreation/other purposes. More information on the development of the range and distribution of the trip frequency constant components is detailed in Appendix C.

3.3 Auto Operating Cost

The auto operating cost for year 2025 and 2029 is assumed to be associated only with privately owned nonautonomous vehicles. Auto operating cost is calculated from the following components:

1. Retail fuel prices in California, which are projected using The U.S. Energy Information Administration (EIA) forecasts with an assumption that California prices are 11 percent higher than the National average (based on consistent patterns in past trends).

2. Additional fees and charges based on two scenarios:
    a. Cap and Trade implementation (i.e., 0 to 19 percent impact on retail fuel prices)[1]; and
    b. Potential increase in Federal excise tax.
3. Fuel economy of the entire "on the road" fleet, calculated from the EIA.
4. Nonfuel costs, which are obtained from the Bureau of Transportation Statistics.

More information on the development of each of these components can be found in Appendix D. The minimum, most likely, and maximum were set based on the combined impacts of these components. For example, the minimum combines the lowest EIA projection with the least impact from Cap and Trade, no increase in Federal taxes, high fuel efficiency, and low nonfuel costs. This approach is reflected in the following formulas, which were used to calculate the minimum, most likely, and maximum auto operating cost:

Minimum Auto Operating Cost = ((Low CA Gas Price + No C&T Impact + No Increase in Federal Gas Tax)/High Fuel Efficiency) + Low Nonfuel Operating Costs

Most Likely Auto Operating Cost = ((Most Likely CA Gas Price + Avg (C&T No Impact, C&T High Impact) + No Increase in Federal Gas Tax)/Most Likely Fuel Efficiency) + Most Likely Nonfuel Operating Costs

High Auto Operating Cost = ((High CA Gas Price + C&T High Impact) + Increase in Federal Gas Tax)/Low Fuel Efficiency) + High Nonfuel Operating Costs

Table 3.2 gives the auto operating cost component values and the resulting minimum, most likely, and maximum auto operating cost for each forecast year. Since auto operating cost comprises individual components that each has minimum and maximum values (as described above), auto operating costs utilize a Shape = 5 PERT distribution. This distribution has somewhat longer tails since the very low or high end of the range has to have each of the individual components end up on the low or high end, which is a very unlikely occurrence.

[1] The exact impact of Cap and Trade on fuel prices is unknown and could change over time based on the industry response to reduce emissions. The California Air Resources Board estimated in 2010 that gasoline price changes in 2020 could range between 4 percent and 19 percent due to Cap and Trade rules (http://www.arb.ca.gov/regact/2010/capandtrade10/capv4appn.pdf). The minimum assumption assumes that Cap and Trade would not result in an increase in gas prices.

Table 3.2 Range of Auto Operating Cost for each Forecast Year by Auto Operating Cost Component
2015 Dollars

	Minimum	Most Likely	Maximum
2025 Auto Operating Cost ($/mile)	$0.15	$0.20	$0.31
California Gas Price (CA Gas Price)	$2.78	$3.41	$5.28
Cap and Trade (C & T Impact)	$0.00	$0.32	$1.00
Increase in Federal Gas Tax	$0.00	$0.00	$0.12
Fuel Efficiency (mpg)	29.4	28.5	28.2
Total Fuel Operating Cost ($/mile)	$0.09	$0.13	$0.23
Nonfuel Operating Cost ($/mile)	$0.05	$0.06	$0.08
2029 Auto Operating Cost ($/mile)	$0.14	$0.19	$0.30
California Gas Price (CA Gas Price)	$2.83	$3.63	$5.73
Cap and Trade (C&T Impact)	$0.00	$0.35	$1.09
Increase in Federal Gas Tax	$0.00	$0.00	$0.12
Fuel Efficiency (mpg)	32.7	31.6	31.2
Total Fuel Operating Cost ($/mile)	$0.09	$0.13	$0.22
Nonfuel Operating Cost ($/mile)	$0.05	$0.06	$0.08
2040 Auto Operating Cost ($/mile)[a]	**$0.13**	**$0.19**	**$0.32**
California Gas Price (CA Gas Price)	$3.00	$4.51	$7.32
Cap and Trade (C & T Impact)	$0.00	$0.43	$1.39
Increase in Federal Gas Tax	$0.00	$0.00	$0.12
Fuel Efficiency (mpg)	38.6	37	36.1
Total Fuel Operating Cost ($/mile)	$0.08	$0.13	$0.24
Nonfuel Operating Cost ($/mile)	$0.05	$0.06	$0.08

[a] The 2040 auto operating costs presented in the table do not include adjustments for autonomous and shared-use vehicles. Once the adjustments for autonomous and shared-use vehicles are accounted for, the year 2040 auto operating cost ranges from $0.13 per mile to $0.37 per mile with a most likely of $0.21 per mile.

For year 2040, in addition to privately owned non-autonomous vehicles, it is possible that autonomous vehicles and shared-use vehicles will have high enough market penetration to affect the overall auto operating cost for long-distance trips. Appendix H provides background on auto operating costs for autonomous and shared use vehicles and their impacts on overall auto operating costs as used for the 2040 analysis. Based on the adjustments for autonomous and shared-use vehicles, the year 2040 auto operating cost ranges from $0.13 per mile to $0.37 per mile, with a most likely of $0.21 per mile.

3.4 HSR FARES

The original base average HSR fare was set at 83 percent of airfares for the Northern California to Southern California market, creating a cap on intermediate average fares of $89. This remains the most likely HSR fare scenario for each forecast year. Variability in HSR fares from 2025 to 2029 is assumed to be consistent with variability in airfares over the five-year period from 2009 to 2014. The maximum HSR fares for years 2025 and 2029 is based on the increase in weighted average of airfare from 2009 to a weighted average of airfare in 2014 for all Southern California to Northern California air travel. Specifically, air travel for the six main Southern California airports (i.e., Los Angeles International Airport (LAX), Burbank Bob Hope Airport (BUR), Long Beach Airport (LGB), John Wayne Airport (SNA), Ontario International Airport (ONT), and San Diego International Airport (SAN)); and four main Northern California airports (i.e., San Francisco International Airport (SFO), Sacramento International Airport (SMF), Oakland International Airport (OAK), and San Jose International Airport (SJC) was examined. It was assumed that HSR fares would increase by an equivalent amount to the highest airfare increase seen in the market in the five-year period.

The minimum for years 2025 and 2029 is based on the lower weighted average of airfare offered by a new market entrant (i.e., Virgin America), compared to a market incumbent (i.e., United Airlines) across two key, comparable segments (i.e., LAX/SFO and SAN/SFO). This calculation is based on the idea that to capture initial market share, HSR may offer lower fares than anticipated.

The minimum for year 2040 is based on the compound annual growth rate (CAGR) that is calculated assuming that the base HSR fare decreases to the minimum HSR fare within a five-year period. The minimum fare for year 2040 applies the derived CAGR from year 2027 to year 2040 assuming the base fare for year 2027. The maximum for year 2040 is based on the CAGR that is calculated assuming that the base HSR fare increases to the maximum HSR fare set for the year 2027. The maximum fare for 2040 applies the derived CAGR from year 2027 to year 2040 assuming the base fare for year 2027. Table 3.3 shows the range in HSR fares for each forecast year as percent change from the base fare. HSR fare has a triangular distribution.

Table 3.3 Range of HSR Fares
2015 Dollars

Forecast Year and Operating Plan	Minimum	Base/Most Likely	Maximum
2025 VtoV	-15.4% = $69 for San Jose to North of Bakersfield	$82 for San Jose to North of Bakersfield	+27.5% = $105 for San Jose to North of Bakersfield
2029 PH1	-15.4% = $75 for San Francisco to Los Angeles	$89 for San Francisco – Los Angeles	+27.5% = $113 for San Francisco to Los Angeles
2040 PH1	-35.3% = ~$58 for San Francisco to Los Angeles	$89 for San Francisco – Los Angeles	+88.1% = $167 for San Francisco to Los Angeles

3.5 HSR Frequency of Service

The number of roundtrip HSR trains that are actually scheduled may vary from the planned service levels. The most likely scenario matches the current planned levels of service in the base model runs. The minimum is based on the absolute least amount of service that could be expected to be run once the system is constructed. The maximum service frequency is based on the maximum amount of service that could be expected to run on a Silicon Valley to Central Valley line system for year 2025 and a level of service that approaches maximum track capacity, subject to a flexible service plan, for year 2029 and year 2040. Table 3.4 shows the range in trains per day for each forecast year. HSR frequency of service uses a triangular distribution.

Table 3.4 Range in HSR Frequency of Service

Forecast Year	Minimum[a] (Roundtrips/Day)	Base/Most Likely (Roundtrips/Day)	Maximum (Roundtrips/Day)
2025 (VtoV)	14	22	76
2029 & 2040 (Phase 1)	44	98	152

[a] For comparison, the Capital Corridor runs 15 roundtrips per day.

3.6 Availability and Frequency of Service of CVR and HSR Buses

The availability and frequency of service of CVR and HSR buses is a discrete variable that considers the presence, or lack of, specific improvements to connecting rail services and HSR bus connections. This risk variable is only considered for the year 2025 Silicon Valley to Central Valley line scenario because these connections have less of an impact once the Phase 1 system is completed.

The variable is composed of three potential future scenarios (1, 2, and 3) with a probability assigned to each scenario. Only one of the three scenarios is chosen for each draw of the Monte Carlo simulation. The scenarios consider the following CVR improvements:

- Construction of a new BUR Metrolink station that would be closer to the high-speed rail station, which is set to begin in 2015, and thus has a low risk of not being constructed.
- Electrification of Caltrain, which is funded and set to be completed in 2020 or 2021.
- Increasing San Joaquin service frequency to that which is projected in the 2013 California State Rail Plan. This proposed service frequency is above current capacity levels negotiated with freight railroads, thus, increasing capacity to this frequency level would require changes to those agreements and/or improvements to the line that are in the planning stages now.

The three distinct scenarios that were considered are as follows:

- **Scenario 1.** No service improvements are made to CVR above 2015 levels. No HSR buses are provided to meet HSR trains.
- **Scenario 2.** All future CVR improvements are completed, including the BUR Metrolink station and the Electrification of Caltrain, with the exception of the San Joaquin service improvements. About 75 percent of the originally planned HSR buses are in service to meet HSR trains.
- **Scenario 3.** All planned CVR improvements are completed, and all planned HSR buses are available to meet the HSR trains.

The discrete probabilities of each of these scenarios occurring were set at 10 percent for Scenario 1, 50 percent for Scenario 2, and 40 percent for Scenario 3.

3.7 Coefficient on Transit Access-Egress Time/ Auto Distance Variable

Between some regions in California, especially in the VtoV scenario, individuals who wish to travel primarily by transit to reach their destination must transfer from a HSR bus or CVR system before or after traveling on HSR. International experience has shown that there is uncertainty around how the need to make these transfers affects overall HSR ridership. The uncertainty in the impact of transfers can have a significant impact on ridership and revenue, especially when the CVR or HSR bus leg of the journey is relatively long in relation to the HSR travel length. Thus, this uncertainty was included as a potential risk variable.

The transit transfer uncertainty is addressed by varying the range for the parameters associated with transit access/egress travel times relative to origin-

destination (OD) distances variable. This variable appears in the access and egress modal utility functions as follows:

$$\boldsymbol{\beta} \times \max\left(\mathbf{0}, \frac{[\boldsymbol{Acc\ or\ Egr\ Time}]}{[\boldsymbol{OD\ Distance}]} - \boldsymbol{Threshold}\right)$$

In the base model, several threshold parameter options were tested in model estimation, and a value of 0.2 was ultimately identified. The values of beta (the variable coefficient) were estimated directly, and were found to be negative. Separate coefficients were estimated for auto access/egress modes versus non-auto access/egress modes (transit and walk/bike), with the magnitude of auto coefficients estimated to be much larger. This variable essentially provides a disincentive for selecting a main mode that requires a long access or egress time, relative to the entire trip length. The uncertainty associated with the variable is only applied for the HSR main mode (i.e., not air or CVR).

An example of the experience in France was researched. In the French experience, moving from a direct CVR connection between Paris and Grenoble to an HSR trip from Paris to Lyon and a connection to CVR from Lyon to Grenoble saved 90 minutes of total travel time, but did not result in increased ridership. The observed "90-minute penalty" in France served as a rough benchmark for determining a lower bound on the model parameters.

Appendix E details the process taken to develop the minimum parameter values for this variable. The minimum threshold value is set to 0.1, since a lower threshold would start to impact local transit access and other unrelated trips. The minimum coefficient value is set to -2.0 for business/commute purpose and -1.3 for recreation/other purpose. These are set to achieve penalty values of 51 and 66 minutes. These penalty value benchmarks come from the penalties the model suggests for the French scenario for drive access/egress modes. The lower bound on the transit penalty should not exceed the penalty suggested by the model for drive access/egress modes. A 51-minute and 66-minute penalty was used instead of the 90-minute penalty observed in the French experience because it offered more reasonable model behavior overall, and it was not desirable to change the long-distance models in unreasonable ways to match a single observed data point. The coefficient and threshold value vary in parallel (i.e., perfect correlation) for the full model runs and Monte Carlo simulation.

The maximum threshold and coefficient values are set to be identical to the calibrated base/most likely values since there is no evidence to suggest that the penalty to transfer from transit to HSR should be less than the penalty used for CVR and Air that was developed based on observed data. A PERT distribution was used for this variable.

3.8 Airfares

Airfares are only considered as a risk variable for year 2029. The airfare uncertainty is based on the variability in airfares from 2009 to 2014 from routes that serve the major airports of the Northern California-Southern California market. Mean, minimum, and maximum annual weighted airfares by route were calculated for each year between 2009 and 2014. Since, the base airfares (i.e., year 2009) represent the lowest point from the range analyzed, the base fares were set as the minimum value. The most likely value was set as the decimal factor difference from the base fare and the average of the calculated mean airfares across the analyzed routes (i.e., 20 percent higher airfares compared to the base fares). The maximum value was set as the decimal factor difference from the base fare and the average of the calculated maximum airfares across the analyzed routes (i.e., 33 percent higher airfares compared to the base fares). A triangular distribution was used for this variable.

3.9 Number and Distribution of Households throughout the State

Statewide population forecasts were assembled from various sources, as shown in Figure 3.3 and documented in detail in Appendix F. As shown in Figure 3.3, short-term forecasts (i.e., through year 2029) do not differ very much, while year 2040 has higher variation in forecasts. Thus, uncertainty associated with number and distribution of households throughout California is only considered for year 2040.

The maximum value was set based on sources with high projections that were adjusted further up based on possible, though unlikely events, such as increased lifespans, increased fertility rates, comprehensive immigration reform that allows more immigrants, and more balanced domestic migration. The minimum value was set based on sources with low projections that were adjusted further down based on possible, though unlikely events, such as substantial tightening of immigration policy and reduced lifespan. The Most Likely forecast uses a combination of mid-range forecasts. Table 3.5 describes the population forecast assumptions and CAGR for each forecast level.

Figure 3.3 Statewide Population Forecasts by Source of Forecast

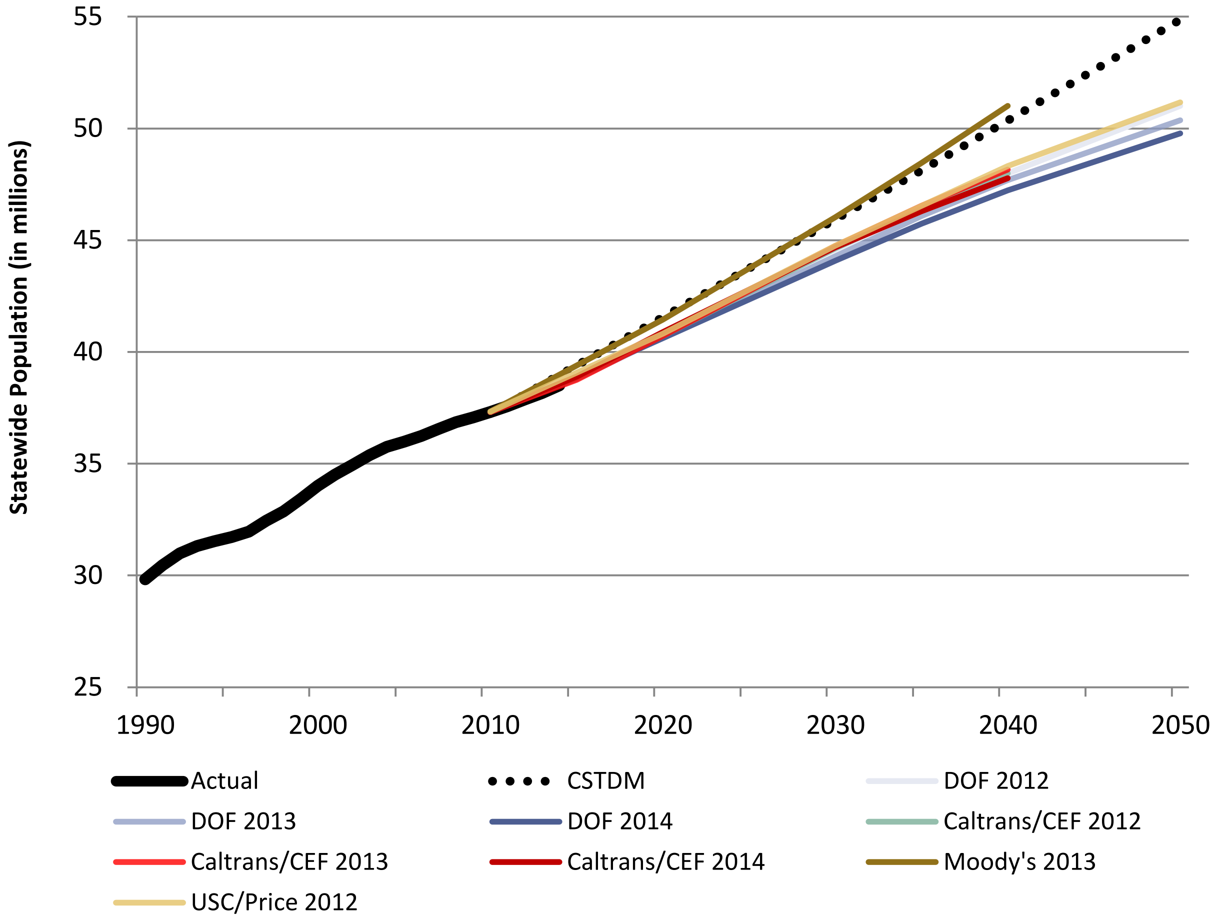


Table 3.5 Statewide Population Forecasts

Forecast	Source of Forecast	CAGR (2010 to 2040)	2040 Forecast California Population
Maximum	• Statewide population total matched to Department of Finance (DOF) "actuals" through 2014. • For 2015 to 2020, statewide population growth follows the California Statewide Travel Demand Model (CSTDM) 20-year moving average growth rate. • For 2021 to 2040, statewide population growth follows the CSTDM 20-year moving average growth rate plus additional 50,000 residents per year (2021), increasing to 150,000 residents per year in 2050.	1.16%	52 million
Mid Range	• Statewide population total matched to DOF "actuals" through 2014. • Between 2015 and 2023, statewide population growth follows the midpoint between the U.S. Census (National) and DOF 2014 (California) 20-year moving average growth rates. (The two growth rates converge at 0.82 percent in 2023.) • Beyond 2023, statewide population growth follows the DOF 2014 20-year moving average growth rates.	0.82%	47 million
Minimum	• Statewide population total matched to DOF "actuals" through 2014. • Between 2015 and 2020, statewide population growth follows the U.S. Census (National) 20-year moving average growth rate. • For 2021 to 2040, statewide population growth follows the 20-year moving average growth rate minus 25,000 fewer residents per year (2021), decreasing to 100,000 fewer residents per year in 2050.	0.58%	44 million

Once population forecasts were developed, the population forecasts were converted to household totals using the following assumptions:

- **Maximum and Most Likely**. Statewide household total calculated by applying the CSTDM household size assumptions to the statewide population totals; and
- **Minimum**. Statewide household total calculated by applying the average household size assumptions from three sources (i.e., CSTDM, CEF 2014, and Moody's 2013) to the minimum statewide population totals.

Household projections ranged from 14.98 million households to 17.84 million households with a most likely value of 16.13 million households for 2040. A triangular distribution was utilized for the distribution.

3.10 Auto Travel Time

By 2040, it is likely that autonomous vehicles (AV) will compose some share of all automobile travel. One of the promises of AV technology is to improve travel speeds by connecting vehicles, allowing them to travel much closer to one another

at high speeds, effectively increasing capacity and reducing congestion.[2] While AVs will improve travel times eventually, current forecasts show that it is unlikely that AVs will represent even a majority of auto travel by 2040, relative to non-AVs. Most of the travel time benefits of AVs rely on AVs representing a clear majority of autos, with the most benefits really being achieved once market penetration reaches about 75 percent (since a mix of AVs and non-AVs does not achieve the same ability to space vehicles closer together).[3] However, market penetration is expected to be between 10 percent and 75 percent based on available research.[3,4]

The current congested travel times forecast for 2040 are considered to be the maximum auto travel times that are likely to occur in year 2040. While it is possible that AVs, in the short term, increase congestion, there is minimal risk in that direction. On the other end of the spectrum, free-flow travel time is considered the absolute minimum travel times that could occur. In theory, it would actually be possible to achieve better speeds than free-flow speeds at very high levels of AV market penetration. However, that is unlikely by 2040 given limitations due to market penetration and highway design.

The AV effect on auto travel times was modeled using a weighted average of congested and free-flow travel times, using a travel time index varying between 0 and 1 based on the following:

- At 0, congested travel times are observed;
- At 1, free-flow travel times are observed; and
- At 0.5, the midpoint travel times between congested and free-flow are observed.

Appendix G details the methodology undertaken to develop the range in auto travel times based on two key sources of uncertainty that affect the travel time index: 1) market penetration, and 2) the impact of AV travel time at each market penetration level.

The effect of market penetration on travel time may be different depending on whether a vehicle is traveling on a freeway compared to an arterial. Speed improvements could be realized at lower market penetrations on freeways compared with arterials, where more advanced technology might be required to realize improved travel speeds. Thus, the relationship between market penetration and travel time is segmented across freeways and arterials. However,

---

[2] Bierstedt, J., A. Gooze, C. Gray, J. Peterman, L. Raykin, and J. Walters, 2014. Effects of Next Generation Vehicles on Travel Demand and Highway Capacity by FP Think Working Group Members. FP Think Working Group.

[3] Littman, T., 2015. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning. February 27, 2015. Victoria Transport Policy.

[4] Milakis, D., M. Snelder, B. van Arem, B. van Wee, and G. Correia. 2015. Development of Automated Vehicles in the Netherlands: Scenarios for 2030 and 2050. Delft, The Netherlands: Delft University of Technology.

since the experimental design can handle only a single auto travel time index variable, freeway and arterial indices are set to be perfectly correlated. Table 3.6 shows the AV market penetration range for year 2040 and the minimum and maximum travel time indices for freeways and arterials.

Table 3.6 Auto Travel Time Index Range for Freeways and Arterials

Risk Variable	Minimum	Most Likely	Maximum
AV Market Penetration	10%	35%	75%
TT Index – Freeway	0.00	0.06	0.78
TT Index – Arterial	0.00	0.03	0.40

4.0 Implementation of Risk Analysis

The approach to developing a risk analysis model for quantifying the uncertainty associated with HSR ridership and revenue can be formally described as developing a meta-model for uncertainty quantification. Uncertainty quantification refers to the process of propagating input uncertainty through a computational model in order to estimate and characterize the response uncertainty. For the 2016 BP risk analysis, input or model uncertainty is addressed by the risk variable ranges and distributions, and the response uncertainty is captured by the probability distribution of high-speed rail ridership and revenue.

This type of analysis is new to the field of travel demand forecasting, but relatively common in other fields. The literature was reviewed from these other fields to develop an experimental design appropriate for this type of analysis.[5] A meta-model is any relatively simple mathematical relationship between parameters and a response, often based on a subset of data. For this analysis, the relationship between parameters and response comes directly from the BPM-V3, which is a complex mathematical model. To capture risk, a Monte Carlo simulation of the model was needed, but due to the model's complexity, it was infeasible to run it thousands of times. Therefore, regression meta-models were developed to approximate the relationships between BPM-V3 revenue and ridership and model inputs and variables based on actual model runs. The regression model can be run very quickly (i.e., tenths of a second), while the BPM-V3 model takes hours to run.[6] Based on the model runs that were conducted, it is possible to test the regression meta-model's ability to replicate the results of the original model. The meta-models that were used were all able to replicate at least 90 percent of the variation in the base model, a very strong and sufficient relationship.

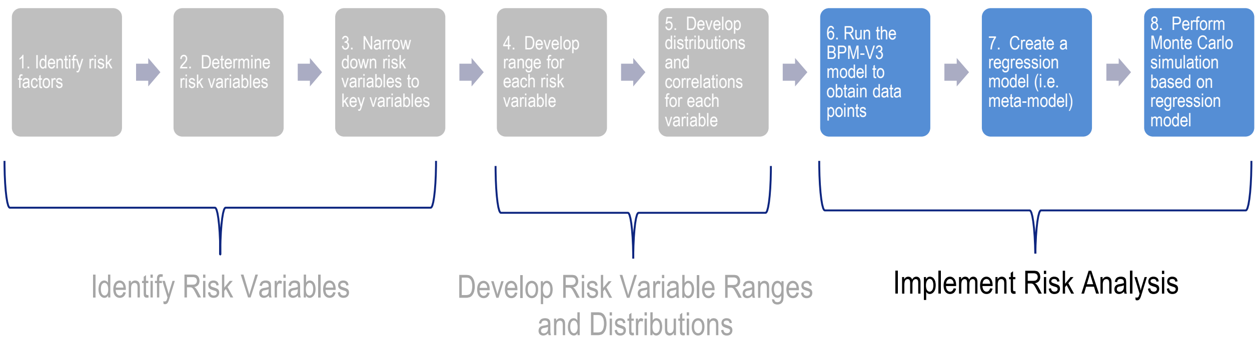
As shown in Figure 4.1, there are three steps that comprise the risk analysis implementation. The regression meta-model is developed from a set of full BPM-V3 runs (*Step 6*). The independent variables of the regression model are the risk analysis variables, and the dependent variable is either HSR revenue or ridership. Each full BPM-V3 model run acts as one data point for use in estimating the regression equations (*Step 7*). A Monte Carlo simulation, of 50,000 draws, is then run using the ridership and revenue regression meta-models and different

[5] One Hour Presentation: Surface (Meta-Model) Methods and Applications, B. M. Rutherford, L. P. Swiler, T. L. Paez, and A. Urbina, Sandia National Laboratories; presented at the 2006 IMAC-XXIV: Conference and Exposition on Structural Dynamics.

[6] It takes approximately 12 hours to run the BPM-V3 model using a one-thread set-up. It takes one hour to run the BPM-V3 model using a 12-thread set-up, which is the maximum possible threads that can be run on one standard computer.

combinations of values of the risk variables, with the values being drawn from the assigned risk variable distributions (*Step 8*). The revenue output from these runs is then used to develop the revenue range and probability of occurrence.

Figure 4.1 Eight-Step Risk Analysis Approach: Implement Risk Analysis (Steps 6 to 8)



4.1 BPM-V3 Model Runs

An experimental design was developed to determine the number of full BPM-V3 model runs needed and the combination of variable values that compose each BPM-V3 model run. The analysis used a combination of a Fractional Factorial design and a Sampling design to develop a two-step analysis process. Fractional Factorial designs are classical designs with a number of desirable properties:

- They do not require many runs per variable explored;
- They are powerful in their ability to distinguish those variables that are most important and those of lesser importance; and
- They can be designed to ensure that both main effects and interaction effects can be estimated from the results.

Sampling designs are particularly useful for computer experiments that exhibit systematic noise. The combination of the two for this analysis ensured interaction effects could be tested, guaranteed that the full model runs spanned the solution space, and kept the number of full BPM-V3 full model runs to a manageable number. Additional details of the development of the experimental design process are discussed in Appendix H.

As an initial step in the analysis, a fractional factorial three-level Resolution IV Design was used to test the existence of two-factor interaction effects. Three-level refers to the number of distinct values each risk variable can take. In this case, each risk variable was set to one of three levels (i.e., minimum, most likely, and

maximum) for each of the full model runs.[7] A three-level design allows for estimating a nonlinear functional form for each variable. Resolution IV indicates that two-way interaction effects can be detected and estimated. Two independent variables interact if the effect of one of the variables differs depending on the level of the other variable. A total of 81 full model runs is needed for this design. For all alternatives and forecast years, three regression models were developed from these designs. One regression was a main effects only regression that did not include interaction terms or nonlinear terms. Another was a main effects only regression that included both linear and nonlinear terms, but no interaction terms. The other regression tested interaction terms. We found that within the solution space tested, there were no two-way interaction effects that were strong enough to warrant including in the regression model. The results of these other models for the various operating plans and forecast years are discussed in Appendix I.

As a result of the findings based on the initial step, the final experimental design included 59 full model runs for each alternative and forecast year, as follows:

- 27 model runs using fixed minimum, most likely, and maximum values of risk variables specified using a three-level Resolution III fractional factorial design. Only 3 of these model run overlapped with the 81 model runs using the three-level Resolution IV Design
- 27 model runs sampled uniformly from low, mid, and high ranges of the risk variables using a random sampling design. These runs ensured that the interior of the solution space was well-represented and not biased toward the edges.
- 5 model runs representing extreme scenarios of full upside (3 runs) and full downside (2 runs); that is, all inputs in these runs were set to values that would either be toward the very favorable or very unfavorable end of the spectrum of HSR revenue and ridership. The runs correspond to the following percentiles for each risk variable: 10, 25, 75, 90, and 100. The 0^{th} percentile run was not added because the experimental design included this run already, where all inputs are set to the "min" value, and the Minimum value always corresponded to the absolute minimum, unfavorable value for HSR revenue or ridership.

Thus, the final experiment design includes both the Fractional Factorial design to help understand extreme values and tails of distributions, and the Sampling design which helps fill in the space in the middle of the distribution where most results fall.

---

[7] As stressed by the CAHSRA Ridership Advisory Technical Panel (RTAP) in their Report 11, it is important to set the bounds of the inputs used in the regression model to the same or wider than the range of inputs used in the Monte Carlo analysis.

4.2 Final Revenue Regression Models

The forecast revenues from the 59 BPM-V3 runs were used as data points for developing the meta-model linear regression equations of the log of revenue as a function of the 10 risk variables.[8] The final set of regression models for each model year and operating plan took the following functional form: $\ln(\text{Revenue}) = \text{Constant} + \beta_1 \times \text{Var}_1 + \beta_2 \times \text{Var}_2 \ldots + \beta_{10} \times \text{Var}_{10}$. This model is a main effects model with no interaction terms and matched the observed data well. For the 2040 model, a nonlinear transformation of the HSR fare variable was also found to be significant.[9] The estimated models are shown in Table 4.1. All models have r-squared values above 0.9, indicating that the regression model fits the BPM-V3 data points very well, and all of the signs and magnitudes of model coefficients are sensible. For example, a positive value on auto operating cost indicates that, as auto operating cost increases (i.e., it becomes more expensive to drive), HSR revenue also increases.

Table 4.1 Revenue Regression Model Results

Constant and Regression Model Variables	2025 – VtoV line		2029 – Phase 1		2040 – Phase 1	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Constant	20.595	89.7	21.171	68.9	21.071	114.0
HSR Mode Choice Constant – Business	0.230	15.5	0.180	14.5	0.194	11.4
HSR Mode Choice Constant – Commute	0.089	3.4	0.086	3.9	0.091	3.0
HSR Mode Choice Constant – Recreation/Other	0.428	16.3	0.390	17.7	0.391	12.9
Trip Frequency Constant – Business/Commute	0.506	6.3	0.587	8.9	0.596	6.8
Trip Frequency Constant – Recreation/Other	1.075	5.2	0.735	4.5	0.917	4.2
Auto Operating Cost	0.945	1.7	1.404	3.1	1.315	2.9
HSR Fare	-0.579	-3.3	0.067	0.5	0.847	1.9
HSR Headway	-0.244	-4.1	-0.156	-4.1	-0.160	-3.0

[8] The original 81 runs developed in the first step were not used for the development of the final regression equations in order to reduce the number of random sampling designs needed. To ensure the interior of the solution space was well-represented and not biased toward the edges, it is essential to perform the same number of three-level random sampling runs as fractional factorial runs.

[9] HSR fare was estimated as nonlinear for year 2040 because the range used for year 2040 fares was much wider than for the other model years and produced more of the parabolic relationship between fares and revenue.

Constant and Regression Model Variables	2025 – VtoV line		2029 – Phase 1		2040 – Phase 1	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
HSR Access/Egress Connecting Service – Scenario 1	-0.158	-2.4	n/a	n/a	n/a	n/a
HSR Access/Egress Connecting Service – Scenario 3	-0.073	-1.1	n/a	n/a	n/a	n/a
Airfare	n/a	n/a	-0.003	-0.0	n/a	n/a
HSR Access-Egress by Transit Variable	1.860	2.4	1.550	2.4	n/a	n/a
Number and Distribution of Statewide Households	n/a	n/a	n/a	n/a	0.075	0.9
Auto Travel Time Index	n/a	n/a	n/a	n/a	-0.097	-1.0
Exp (HSR Fare)	n/a	n/a	n/a	n/a	-0.212	-1.8
Model Statistics						
Sum of Squared Error	2.389		1.686		3.182	
R^2	0.942		0.945		0.909	

4.3 Final Ridership Regression Models

The forecast ridership from the 59 BPM-V3 runs were used as data points for developing the meta-model linear regression equations of the log of ridership as a function of the 10 risk variables. The final set of ridership regression models for each model year and operating plan took the following functional form: $\ln(\text{Ridership}) = \text{Constant} + \beta_1 \times \text{Var}_1 + \beta_2 \times \text{Var}_2 \ldots + \beta_{10} \times \text{Var}_{10}$. This model is a main effects model with no interaction terms. The estimated models are shown in Table 4.2. All models have r-squared values above 0.9, indicating that the regression model fits the BPM-V3 data points very well, and all of the signs and magnitudes of model coefficients are sensible. For example, a negative value on HSR fare indicates that, as HSR fare increases, HSR ridership decreases. Note that for revenue this is not always the case since for certain values of HSR fare; the increase in HSR ridership offsets the loss of revenue from a decrease in HSR fare.

Table 4.2 Ridership Regression Model Results

Constant and Regression Model Variables	2025 – VtoV line		2029 – Phase 1		2040 – Phase 1	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Constant	17.565	75.9	18.350	55.5	18.317	116.3
HSR Mode Choice Constant – Business	0.221	14.8	0.184	13.8	0.196	11.9
HSR Mode Choice Constant – Commute	0.097	3.7	0.103	4.3	0.111	3.8
HSR Mode Choice Constant – Recreation/Other	0.435	16.5	0.415	17.6	0.409	14.0
Trip Frequency Constant – Business/ Commute	0.502	6.2	0.595	8.6	0.603	7.1
Trip Frequency Constant – Recreation/ Other	1.046	5.0	0.710	4.1	0.862	4.1
Auto Operating Cost	0.885	1.5	1.117	2.3	1.026	2.3
HSR Fare	-1.416	-8.0	-0.770	-4.9	-0.679	-10.1
HSR Headway	-0.253	-4.2	-0.160	-3.9	-0.160	-3.1
HSR Access/Egress Connecting Service – Scenario 1	-0.143	-2.1	n/a	n/a	n/a	n/a
HSR Access/Egress Connecting Service – Scenario 3	-0.077	-1.2	n/a	n/a	n/a	n/a
Airfare	n/a	n/a	-0.056	-0.3	n/a	n/a
HSR Access-Egress by Transit Variable	2.136	2.8	2.273	3.3	n/a	n/a
Number and Distribution of Statewide Households	n/a	n/a	n/a	n/a	0.056	0.7
Auto Travel Time Index	n/a	n/a	n/a	n/a	-0.082	-0.8
Model Statistics						
Sum of Squared Error	2.425		1.871		2.977	
R^2	0.946		0.948		0.935	

4.4 Revenue Results of the Monte Carlo Simulation

A Monte Carlo simulation using the regression meta-model was run 50,000 times using different combinations of values of the risk variables, with the values being drawn from the assigned risk variable distributions. Note, some risk factors include multiple components that are sampled in the Monte Carlo analysis. For example, values are sampled from both the uncertainty component distribution and the terminal/wait time component distribution for the HSR Mode Choice

Constant risk variable. Appendix J details the components of each risk variable, the range of values and distributions for each component, and correlation between distributions of risk variables. Setting a positive correlation between two risk variable components results in the Monte Carlo simulation having a higher probability of sampling from the same point on the distribution (e.g., a 100-percent positive correlation would result in two risk variables always being chosen from the same percentile point on the distribution).

The revenue output from these 50,000 Monte Carlo runs was used to develop the revenue range and probability of occurrence, as shown in Table 4.3. Short-distance trips less than 50 miles within the Southern California Association of Governments (SCAG) and the Metropolitan Transportation Commission (MTC) contribute $12 million in revenue in year 2029 and 2040. This short-distance revenue was added to the year 2029 and year 2040 long-distance revenue for all probability levels to obtain total HSR revenue.

The "base run" is the revenue for the year and scenario forecast using the BPM-V3 model with the base input variable values. The percentages shown are where the original base revenue falls on the continuum of revenue forecasts produced by the various risk models.

Table 4.3 Year 2025 to 2040 HSR Revenue Range and Probability of Occurrence[10]

	Revenue (Millions of 2015 Dollars)		
Probability	**2025 VtoV line**	**2029 PH1**	**2040 PH1**
Minimum	$112	$634	$704
1%	$192	$950	$1,038
10%	$280	$1,303	$1,471
25%	$359	$1,619	$1,852
Median	$484	$2,082	$2,419
75%	$652	$2,691	$3,153
90%	$840	$3,359	$3,963
99%	$1,215	$4,610	$5,606
Maximum	$2,144	$6,628	$9,191
Base Run	$460 (46%)	$2,069 (49%)	$2,413 (50%)

Figures 4.2 plot the cumulative distribution of HSR revenue for years 2025, 2029, and 2040, respectively.

[10] The results are raw model output and do not account for ramp-up.

Figure 4.2 Year 2025 Cumulative Distribution of HSR Revenue
2015 Dollars

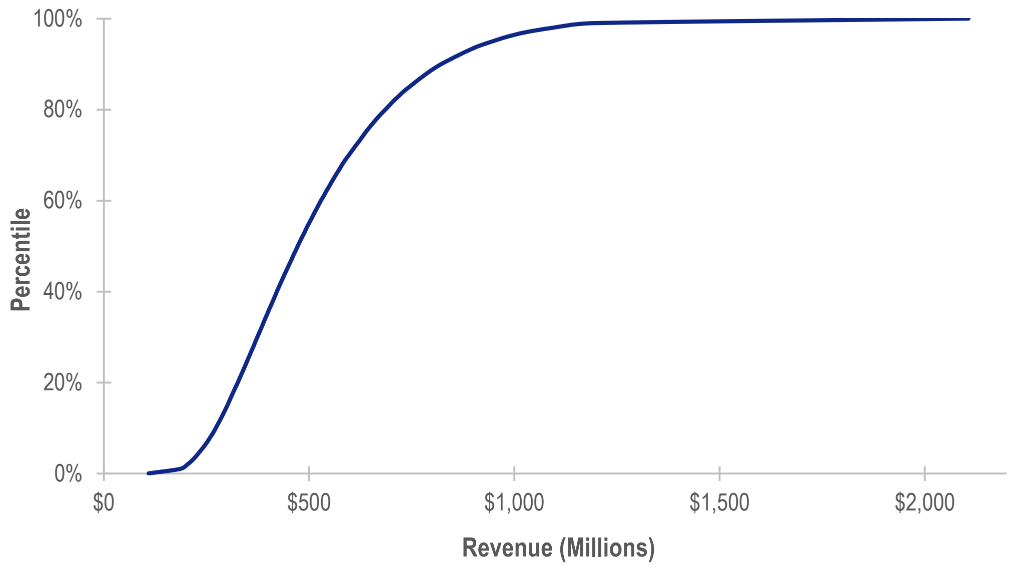


Figure 4.3 Year 2029 Cumulative Distribution of HSR Revenue
2015 Dollars

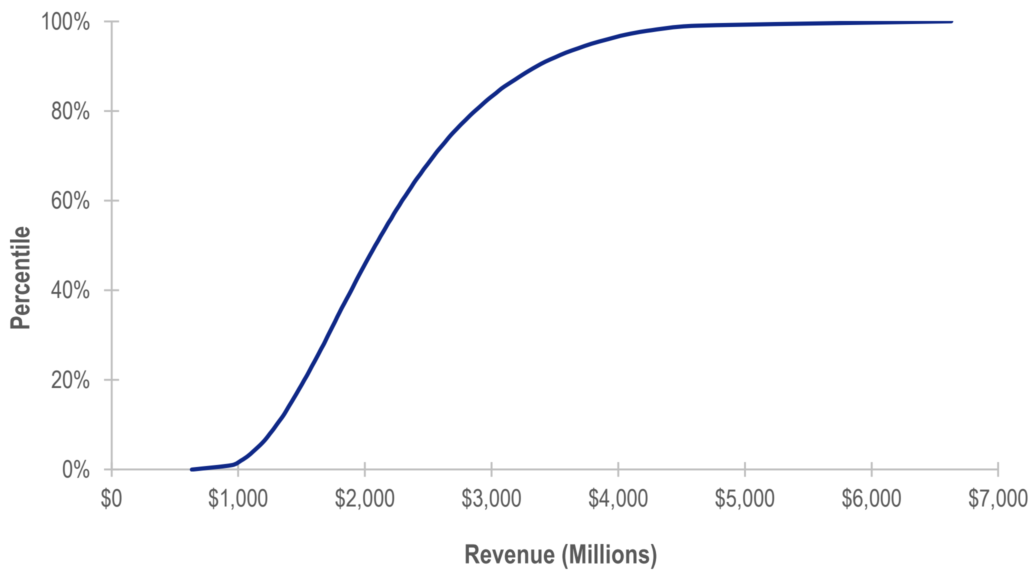
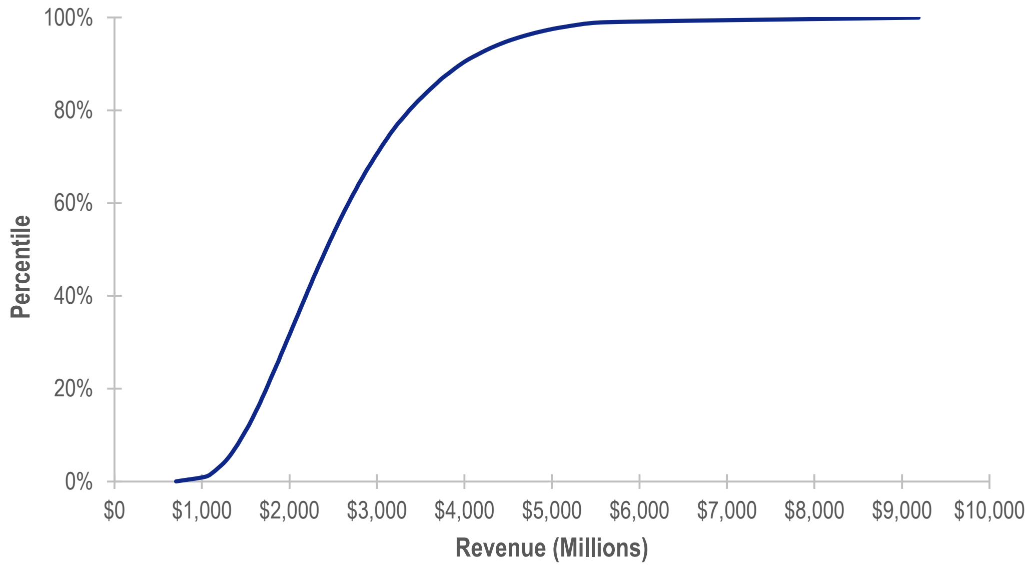


Figure 4.4 Year 2040 Cumulative Distribution of HSR Revenue
2015 Dollars



4.5 Ridership Results of the Monte Carlo Simulation

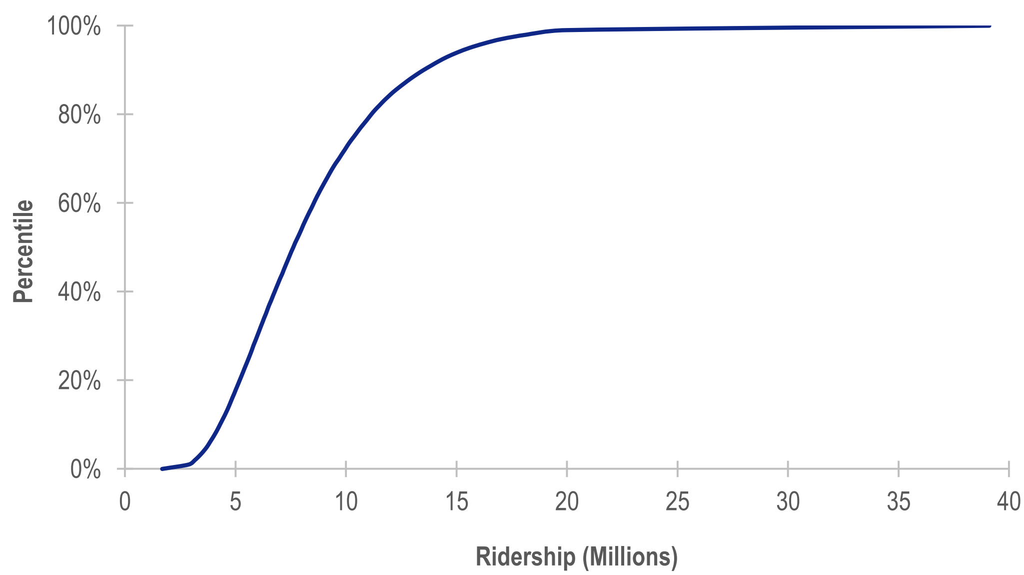
A Monte Carlo simulation using the ridership regression meta-model was applied to the same 50,000 runs developed for the revenue analysis. The ridership output from these runs was used to develop the ridership range and probability of occurrence, as shown in Table 4.4. Short-distance trips less than 50 miles within SCAG and MTC contribute 0.6 million in ridership in years 2029 and 2040. This short-distance revenue was added to the year 2029 and year 2040 long-distance revenue for all probability levels to obtain total HSR ridership. The "base run" is the ridership for the year and scenario forecast using the BPM-V3 model with the base input variable values. The percentages shown are where the original base ridership falls on the continuum of ridership forecasts produced by the various risk models.

Table 4.4 Years 2025 to 2040 HSR Ridership Range and Probability of Occurrence[11]

	Ridership (Millions)		
Probability	2025 VtoV line	2029 Ph1	2040 Ph1
Minimum	1.7	10.2	8.9
1%	3.0	16.3	15.8
10%	4.4	22.9	23.5
25%	5.7	28.7	30.3
Median	7.8	37.5	40.7
75%	10.6	49.1	54.7
90%	13.7	62.0	70.5
99%	20.2	86.6	104.1
Maximum	39.6	137.6	179.1
Base Run	7.5 (47%)	37.1 (49%)	42.8 (54%)

Figures 4.5 to 4.7 plot the cumulative distribution of HSR ridership for years 2025, 2029, and 2040, respectively.

Figure 4.5 Year 2025 Cumulative Distribution of HSR Ridership



[11] The results are raw model output and do not account for ramp-up.

Figure 4.6 Year 2029 Cumulative Distribution of HSR Ridership

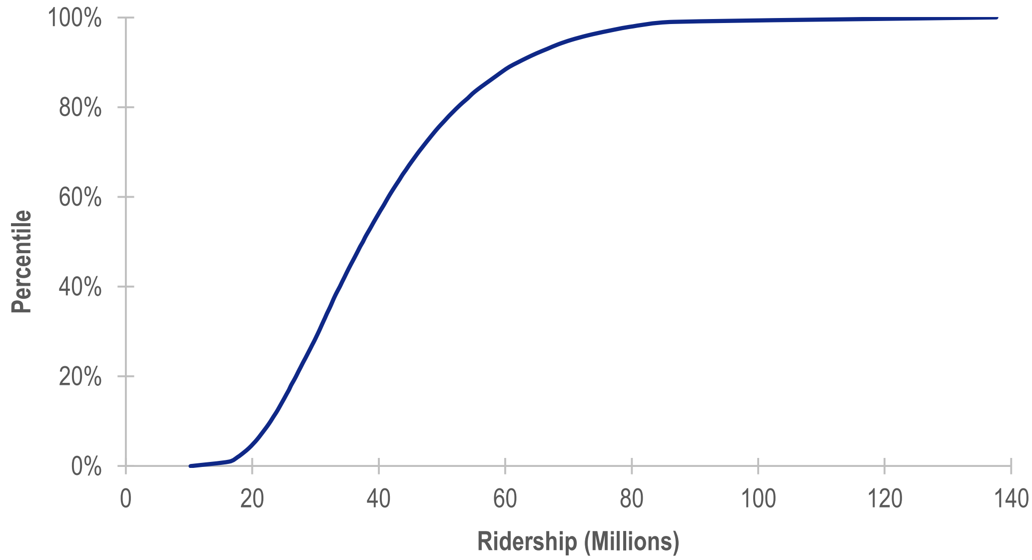
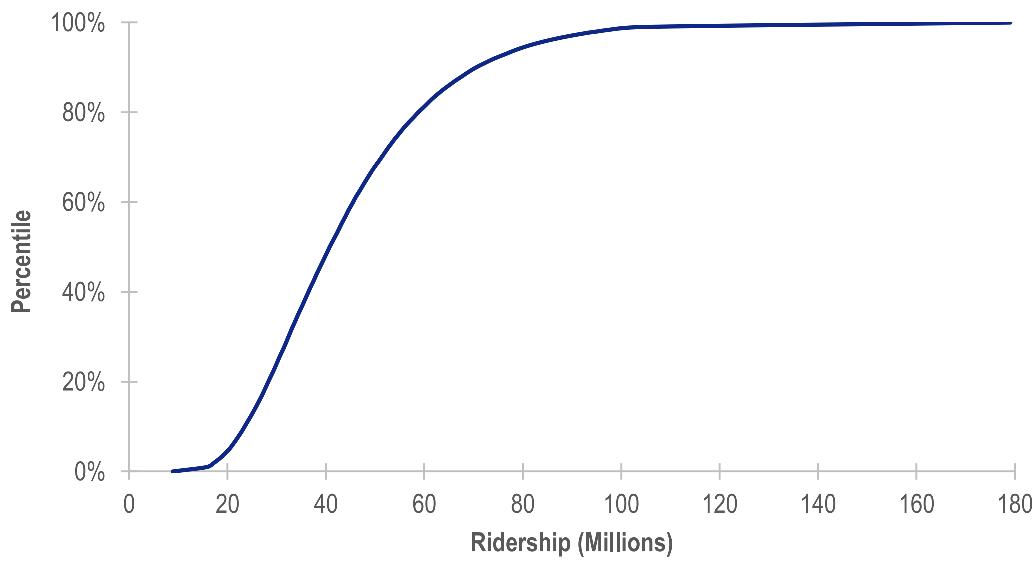


Figure 4.7 Year 2040 Cumulative Distribution of HSR Ridership



4.6 CONTRIBUTIONS TO RISK VARIANCE

One feature of the risk analysis approach taken here is that uncertainty forecasts of HSR ridership and revenue result from the underlying uncertainty in several variables that have direct impacts on HSR ridership and revenue. Each of those variables contributes to the uncertainty in different ways, which can be quantified by examining the variance in the forecasts.

The contribution of the variance of each risk variable component is shown in Table 4.5. The contribution to risk variance for each variable considers two features: 1) the risk variable distribution, and 2) the impact that a unit change in a risk variable has on revenue or ridership, which comes directly from the regression coefficients. The bigger the impact a variable has on revenue or ridership, the bigger its contribution to risk variance, all else being equal. Likewise, the wider a risk variable's distribution, the bigger its contribution to risk variance, all else being equal.

The calculation itself relies on the Monte Carlo simulation plus an additional error component. It adds an error term, which has a size consistent with the error observed from the actual regression model, to the simulated data points, and treats those as observed data points. The calculation comes from Lindeman, Merenda, and Gold (1980) and is called the LMG measure. The original calculation is based on obtaining importance measures for a regression model, which is why the error terms are added. The calculations are valid even when explanatory variables (e.g., risk variables) exhibit correlation among themselves, which is partly why this measure was chosen.

The HSR constants' unexplained variation contributes the most to the variance in the revenue distribution. This result reflects the large distribution on this risk variable component, as well as the large sensitivity of this variable to HSR revenue and ridership. The next set of variables that contributes the most to the variance in HSR revenue is the trip frequency constants' unexplained variation. These results make sense. There is simply a significant amount of uncertainty associated with how travelers will view HSR, because there is no way to observe and collect data related to it until HSR opens. On the other hand, the level of uncertainty associated with the HSR attributes and auto costs is much lower, because they are controllable in the case of the former, or there is a large amount of existing data to rely on in the case of the latter.

Overall, the range and distribution in revenue and ridership reflect the uncertainty associated with a number of the most important determinants across the forecast years. The variables were carefully examined and researched before assigning appropriate distributions to them. The demand model used for forecasting was constructed from and closely matches the results of a complex travel model system that has been vetted with industry experts over the course of several years.

Table 4.5 Contribution of HSR Revenue Variance of each Risk Variable Component

Risk Variable	Risk Variable Component	2025 VtoV	2029 PH1	2040 PH1
HSR Constant – Business	Unexplained Variation[a]	35.3%	32.0%	32.2%
HSR Constant – Commute	Unexplained Variation[a]	13.9%	14.1%	14.0%
HSR Constant – Recreation/ Other	Unexplained Variation[a]	38.4%	40.3%	37.8%
Terminal & Wait Time	Business[b], Commute[b] Recreation/Other[b]	1.4%	1.5%	1.5%
Trip Frequency Constant – Business/Commute	Unexplained Variation[c]	2.3%	3.4%	3.6%
	Economic Cycle[d]	1.3%	2.1%	2.6%
Trip Frequency Constant – Recreation/Other	Unexplained Variation[c]	2.2%	1.9%	2.4%
	Economic Cycle[d]	1.3%	2.1%	2.5%
Base Auto Operating Costs	n/a	0.2%	0.8%	0.6%
HSR Fares	n/a	1.5%	0.0%	0.6%
HSR Headway	n/a	1.9%	1.5%	1.3%
HSR Connecting Service	n/a	0.0%	n/a	n/a
HSR Access/Egress by Transit Variable	Index Variable	0.4%	0.3%	n/a
Airfares	n/a	n/a	0.0%	n/a
Number and Distribution of Statewide Households	n/a	n/a	n/a	0.2%
Automated Vehicle Market Penetration	Penetration	n/a	n/a	0.2%
Automated Vehicle Effect on Auto Travel Times	Alpha, Beta	n/a	n/a	0.0%
Automated Vehicle Fuel Economy		n/a	n/a	0.0%
Shared Use Vehicle Share		n/a	n/a	0.3%
Shared Use Vehicle Cost per Mile		n/a	n/a	0.4%

[a] 50-percent correlation for random draws from distributions in Monte Carlo simulation.
[b] 100-percent correlation for random draws from distributions in Monte Carlo simulation.
[c] 50-percent correlation for random draws from distributions in Monte Carlo simulation.
[d] 100-percent correlation for random draws from distributions in Monte Carlo simulation.

A. Risk Factors and Variables Considered

Table A.1 summarizes the risk factors and variables that were included in the risk analyses. Table A.2 summarizes the risk factors and variables that were considered for inclusion in the risk analyses, but eventually excluded as risk factors. The tables include a description of each variable, potential causes of uncertainty in the variable, and any qualitative or quantitative (such as sensitivity test results) information suggesting why the variable was or was not included in the risk analysis. Some variables that were originally considered for inclusion in the risk analysis were later removed based on additional analyses that concluded that their impact was smaller than originally anticipated (i.e., either the range of uncertainty was significantly narrower than what was tested in the sensitivity tests or the impact was too small for inclusion). Table A.3 lists risk factors that were excluded from consideration as potential risk variables.

Table A.1 Reasoning for Inclusion of Variables in Risk Analysis

Model Variable	Reasons for Considering Model Variable and Risk Factors Represented	Sensitivity Results/Qualitative Reasoning for Inclusion of Variable	Forecast Years
High-speed rail main mode choice constants (one for each trip purpose – business, commute, recreation/ other)	The mode constants capture the unexplained variation in traveler mode choices after system variables (travel cost, time, reliability, and frequency of service) and demographics (household income, autos owned, household size, and number of workers) are taken into account. Unexplained variation may include factors such as: • Comfort aboard trains, • Opinions regarding high-speed rail (HSR_, • Need for auto at destination, • Uncertainty regarding security screening procedures, and • Level of familiarity with HSR.	• The 2014 Business Plan (BP) Risk Analysis demonstrated high sensitivity of ridership/revenue to the HSR constants. • The HSR constants are asserted based on results of stated-preference surveys and cannot be calibrated; as a result, there is uncertainty with the constant itself. • There is uncertainty surrounding travelers' true perception of HSR given the mode currently is unavailable • Uncertainty regarding final access times through the terminal and wait times can be represented through the constant	All
Trip frequency constants (one for each trip purpose – business, commute, recreation, and other)	The trip frequency constants capture the unexplained variation in the numbers of long-distance trips travelers will take after accounting for household demographics and the accessibility of available destinations: • Impact of alternatives to travel (e.g., Skype, GoToMeeting, telecommuting, etc.); • Changing opinions regarding long-distance travel; and • Increased or decreased out-of-state travel replacing in-state travel.	• The 2014 BP Risk Analysis demonstrated high sensitivity of ridership/revenue to the trip frequency constants. • The trip frequency constants were calibrated to reproduce 2010 long-distance travel, which may not have been generally reflective of "typical" conditions since California was exiting a recession. • Even if 2010 represented the current norm, traveler opinions regarding long-distance travel may change over time. • Instead of including distributions of household and employment levels directly as a risk variable in the risk analysis model to account for changes in the state of the economy, risks associated with the state of the economy are accounted for within the trip frequency constant risk variable.	All
Auto operating costs	This variable reflects the inherent risks in forecasting future: • Fuel costs; • Fuel efficiencies; • Adoption of alternative fuels/electric vehicles;	• Most travelers divert to HSR from auto since auto is the dominant long-distance mode, carrying more than 95 percent of current long-distance trips. • HSR ridership and revenue are sensitive to changes in auto operating costs.	All

Model Variable	Reasons for Considering Model Variable and Risk Factors Represented	Sensitivity Results/Qualitative Reasoning for Inclusion of Variable	Forecast Years
	• Maintenance costs; • Possible phase in of vehicle mile traveled (VMT) charge program or increases in gas taxes; • Potential impacts of cap and trade on fuel costs; • Market penetration of autonomous connected vehicles; and • Higher share of "shared-use" vehicles	• For Sensitivity tests indicate that for VtoV a 100-percent increase in auto operating costs would produce a 28-percent increase in HSR ridership and a 30-percent increase in revenue, while a 50-percent decrease in auto operating costs would produce a 13-percent decrease in ridership and a 15-percent decrease in revenue. • For 2040 Phase 1 (PH1), increasing auto operating costs from 20 to 26 cents per mile produced a 6-percent increase in HSR ridership and an 8-percent increase in revenue, while decreasing auto operating costs to 11 cents produced a 13-percent decrease in ridership and a 15-percent decrease in revenue.	
HSR Fares	A number of issues could affect actual fares charged to travelers, especially as the system is being opened, such as: • Institution of discount/premium fares (advance purchase, peak/off-peak, first/second class seating…); • The HSR fare structure might change in response to changing auto operating costs or air fares; and • Yield management and marketing strategies might be used to by a concessionaire to increase efficiency of the service.	• HSR ridership and revenue are sensitive to fares. • Sensitivity tests indicate that for VtoV a 30-percent increase in HSR fares would produce a 21-percent drop in ridership and a 1-percent drop in revenue; a 15-percent decrease in HSR fares would produce a 9-percent increase in ridership and a 5-percent decrease in revenue. • For 2040 PH1, HSR ridership and revenue are sensitive to fares. In a test run, a 30-percent increase in HSR fares produced a 20-percent drop in ridership and a 1-percent increase in revenue; a 15-percent decrease in HSR fares produced a 9-percent increase in ridership and a 5-percent decrease in revenue.	All
HSR Frequency of Service	A number of issues could affect actual HSR frequency of service, especially as the system opens, such as: • Unexpected demand levels may compel the operator to increase or decrease service levels; • Yield management and marketing strategies might be used to by a concessionaire to increase efficiency of the service; and • Early implementation issues with equipment and operations or equipment delivery delays could preclude offering the planned frequency of service.	• HSR ridership and revenue are sensitive to HSR frequency of service. • Sensitivity tests indicate that for VtoV a 50-percent increase in HSR headways would produce a 11-percent drop in ridership and a 11-percent drop in revenues; a 50-percent decrease in HSR headways would produce a 15-percent increase in ridership and a 15-percent increase in revenue. • For 2040 PH1, a 50-percent decrease in HSR headways (time between trains) produced a 12-percent increase in ridership and a 11-percent revenue; a 50-percent increase in HSR	All

Model Variable	Reasons for Considering Model Variable and Risk Factors Represented	Sensitivity Results/Qualitative Reasoning for Inclusion of Variable	Forecast Years
		headways produced a 12-percent decrease in ridership and a 14-percent decrease in revenue.	
Availability and Frequency of Service of conventional rail (CVR) and HSR Buses that connect with HSR	The network assumes an increase (based on the State Rail Plan) in CVR service to HSR terminus stations from existing levels. In addition, the scenarios assume that HSR buses will be available to transport HSR users to/from various locations throughout the State.	• Sensitivity tests indicate that for VtoV using 2010 transit service and removing all HSR bus service would result in a 19-percent decrease in HSR ridership and revenue. • While there is some uncertainty in the service levels of CVR in 2029, PH1 service obviates the need to provide the HSR buses required to provide Sacramento to Los Angeles service under VtoV scenarios. As a result, this variable is only included in the VtoV risk analyses.	2025
Coefficient on Transit Access-Egress Time/ Auto Distance Variable	The model includes a variable that makes HSR less attractive for trips that require a long access or egress trip in relation to the time spent on HSR (the variable actually applies to air and CVR also). Given that the revealed-preference data did not include transferring from CVR (or other transit modes) to HSR, we do not have observed data to directly estimate a coefficient for HSR. Thus, the magnitude of this coefficient is inherently uncertain for HSR. In addition, the experience from HSR in France indicates that the impacts of transfers may carry additional uncertainty. This risk also may be addressed by adjusting the constant associated with transit access/egress for HSR. However, such a treatment would implicitly assume the transfer penalty is identical for all transit access/egress modes, including short local transit access or egress. The apparent transfer penalty effect (and the corresponding uncertainty) likely varies based on overall trip distance and access/egress distances, and thus this risk is better handled within the proposed variable.	• The French experience showed that replacing direct through service via CVR for trips between Paris and Grenoble with TGV between Paris and Lyon connecting to CVR between Lyon and Grenoble did not change total ridership even though there was a total time savings of 90 minutes. The full trip is about 300 miles (Paris to Grenoble) with 60 miles from Lyon to Grenoble. • The French experience was tested by modifying the transit access/egress constants to reflect the 90-minute penalty for a trip between San Francisco and Los Angeles on a VtoV system. The added penalty resulted in a 16-percent decrease in HSR ridership and revenue. Note, modifying the access/egress constant directly is not how we accounted for the "French Experience" risk; however, the results of the test indicate the risk should be analyzed. • With system expansion to the PH1 system, the risk of the penalty decreases as more of the State has a one-seat ride.	2025 & 2029
Airfares	Possible reasons airlines may change airfares from current levels include: • Changes in fuel or personnel costs or airport landing fees,	• Air carriers' response to HSR is unknown. A number of years ago, one airline vowed to "kill" HSR in another state if it was built by cutting fares. Meanwhile airlines in Europe have more recently embraced HSR to replace short-haul and feeder service to long-haul flights.	2029

Model Variable	Reasons for Considering Model Variable and Risk Factors Represented	Sensitivity Results/Qualitative Reasoning for Inclusion of Variable	Forecast Years
	• Changes in equipment or efficiency, such as NextGen technology, • Competitive response to HSR to maintain air market shares, and • Acceptance of HSR as a replacement for inefficient, short-haul air service.	• HSR ridership and revenue may be sensitive to changes in airfares. For year 2029 PH1, a general 30-percent increase in airfares produced a 6-percent increase in HSR ridership and an 8-percent increase in revenue, while a 30-percent decrease in airfares produced a 2-percent decrease in ridership and a 3-percent decrease in revenue. • Airfares are expected to have the largest variation in response to the introduction of HSR in direct competition to the in-state air market (mostly from Southern California to Northern California), so the risk is evaluated for the PH1 opening year forecast.	
Variation in the forecast number of statewide households	The forecasted number of statewide households can fluctuate for a variety of reasons, such as: • Inherent uncertainty with population forecasts; • National and statewide economic cycles; • Impacts of natural disasters such as continuing draught; and • Changes in U.S. immigration policy.	• There is a direct, almost one-to-one, impact of changes in population on changes in HSR ridership and revenue (i.e., a 1-percent change in population results in a 1-percent change in ridership and revenue). • Assumptions regarding the geographic distribution of the growth also affect the final impact of the growth. • The uncertainty of the forecast population and households increases with time and has been deemed to be most important for the 2040 forecast.	2040
Auto travel times	A host of risk factors might be represented through auto travel times: • Increased construction and use of managed lanes; • Significant road deterioration increasing congestion and slowing speeds; • Autonomous vehicles increasing capacity and, thus, speeds through shorter following distances; and • Changes in speed limits.	• While managed lanes are most likely to be constructed in urban areas, construction of managed lanes that have impact on interregional travel might occur by 2040. • The risk of serious road deterioration impacting the roadway system by 2025 or 2029 is unlikely; the risk of significant deterioration is a longer-term risk. • Sensitivity tests indicate VtoV HSR ridership is sensitive to auto travel time (a 20-percent increase in auto IVTT results in a 14-percent increase in HSR ridership, and a 20-percent decrease in auto IVTT results in a 13-percent decrease in HSR ridership). However, free-flow auto travel time between San Francisco and San Diego is approximately 6 percent lower than congested auto travel time (other regions showed lower percent differences). Thus, the expected range in auto travel times is not likely to be wide enough to result in a large	2040

Model Variable	Reasons for Considering Model Variable and Risk Factors Represented	Sensitivity Results/Qualitative Reasoning for Inclusion of Variable	Forecast Years
		change in HSR ridership or revenue until there are more significant changes to congested times or technology changes that impact traffic speed. • Research suggests that the introduction of autonomous connected vehicles could decrease congestion when the market penetration of these vehicles reaches about 75 percent, which could happen by year 2040. • Sensitivity tests indicate that HSR ridership in year 2040 PH1 is sensitive to auto travel time (a 20% increase in auto IVTT results in a 10-percent increase in HSR ridership and 12-percent increase in revenue; a 20-percent decrease in auto IVTT results in a 12-percent decrease in HSR ridership and revenue).	

Table A.2 Risk Analysis Variables Considered but Eventually Excluded from Risk Analysis Model

Model Variable	Risk Factors Represented	Sensitivity Results/Qualitative Reasoning for Exclusion of Variable
High-Speed Rail Reliability	A number of issues could affect actual HSR reliability, especially as the system is being opened: • Early implementation issues with equipment and operations; • Equipment delivery delays may reduce availability of backup equipment; and • The current assumption of 99 percent reliability may not be met.	• HSR ridership and revenue are sensitive to HSR reliability. In a test run, a 25-percent decrease in HSR reliability (from 99 percent on-time arrivals to 75 percent) produced a 21-percent decrease in ridership and a 21-percent decrease in revenue. • For Caltrain, which is a California CVR service that has close to dedicated right-of-way, 92.5 percent of trains have arrived within 5 minutes of scheduled time. • The capitol corridor, which is less than 550 miles is on-time, within 10 minutes of scheduled time, 93 percent of the time. 20 percent is due to freight interference (which would not be an issue for HSR). The San Joaquin, which is <550 miles is on-time, within 10 minutes of scheduled time, 74.2 percent of the time. About 50 percent of delays are due to issues with freight interference. • The above statistics use on-time performance of 5-10 minutes, rather than 15 minutes, suggesting that CVR percentage of on-time performance based on 15-minute arrival could be significantly higher. • International HSR has a strong history of being very reliable. Dedicated HSR track will inherently make HSR much more reliable than CVR. • Reliability was eliminated as a risk variable given that the range in on-time performance was very narrow. Based on sensitivity tests, potential realistic values for on-time performance for HSR (based on California CVR and international HSR experience) would not have a significant enough impact on ridership and revenue to be included.
Parking cost and location at HSR stations	Station planning currently is underway for both the stations and surrounding areas. The model assumes that parking will be provided with parking costs matching surrounding market conditions. These assumptions may change due to: • Insufficient land availability for parking around the HSR station; • Parking "yield" control; • Environmental concerns/neighborhood response; • Changes in planned land use development patterns; and	• A sensitivity test increased parking cost at all HSR stations by 50 percent, and also adjusted the park-and-ride access constant to reflect a 15-minute increase in terminal time (from outlying parking to the terminal). These adjustments resulted in a decrease in HSR ridership and revenue by 11 percent. • Two other sensitivity tests focused only on adjusting HSR parking cost. A 50-percent increase in HSR parking costs resulted in a 5-percent decrease in HSR ridership. A 50-percent decrease in HSR parking costs resulted in a 5-percent increase in HSR ridership. • Parking cost was eliminated as a risk variable given that in initial regression analysis testing, this variable had only a small contribution to uncertainty in ridership and revenue forecasts compared to other variables.

	• Higher amount of off-station parking requiring longer access time to reach station	
Types and numbers of jobs available within a region or statewide	• Development of new employment centers or industries in a region; and • Loss of employment due to recession or natural disaster.	• Sensitivity test indicates VtoV HSR ridership is not very sensitive to change in employment distribution. A ±20-percent change in Leisure/Hospitality employment in the San Joaquin Valley produced a ±1-percent change in HSR ridership and revenue. • The numbers and types of jobs have only an indirect effect on main mode choice through changes in destination choice. • Impact of this variable can be taken into account through widening of the range for the Trip Frequency constant.
Changes in spatial distribution of households and employment <u>within</u> metropolitan regions	• Transit-oriented development around HSR stations; and • Changes in major attractions in California (such as Disneyland).	• Changes in statewide numbers and locations of households are being considered in the risk analysis; this variable relates to very specific development patterns. • Deviations from projected land uses in 2025 and 2029 are unlikely to be sufficient to cause significant change in ridership in those scenarios in year 2040; sensitivity to types and numbers of jobs available within a region was minimal. • Temporary closure of Disneyland will not affect the VtoV scenario to the same extent as the PH1 scenario. Sensitivity tests on the PH1 scenario showed very low sensitivity of HSR revenue to opening and closing of a "Disney-like" special generator.
Airline service frequency	Possible reasons airlines may change service frequency or even markets served from current levels include: • Changes in fuel or personnel costs or airport landing fees at specific airports; • Changes in equipment or efficiency, such as introduction of NextGen allowing for more air traffic at an airport; • Competitive response to HSR to maintain air market shares; and • Acceptance of HSR as a replacement for inefficient, short-haul air service.	• HSR ridership and revenue are not very sensitive to changes in the frequency of air service. In sensitivity tests, a general 50-percent increase in air service headways (time between flights) produced a 2-percent increase in HSR ridership and a 3-percent increase in revenue, while a 50-percent decrease in headways between flights produced a 3-percent decrease in ridership and a 4-percent decrease in revenue.
Airline travel time	Airline travel time may change from current levels for the following reasons: • NextGen aviation system introduction decreasing travel time;	• NextGen implementation by 2029 is unlikely to change relative competitiveness in the short-haul air market within which HSR competes. • Sensitivity tests indicate that VtoV HSR ridership is not very sensitive to changes in air travel time. A 20-percent decrease in air travel time produced only a 1-percent decrease in HSR ridership and revenue.

	• Reliability of airlines could significantly improve decreasing the need for travel time padding; and • More or less nonstop service could be introduced between cities.	• Sensitivity tests show very low sensitivity of air travel time on HSR revenue. For year 2040 PH1, increasing air travel time by 20 percent increased HSR revenue by 0 percent, and decreasing air travel time by 30 percent decreased HSR revenue by 2 percent.
Auto mode characteristics (auto mode constant)	• Auto becomes more attractive compared to HSR due to autonomous vehicles	• The HSR constant measures the unexplained variation in the attractiveness of HSR in relation to the auto mode. Thus, the uncertainty surrounding changes in the perception of the auto mode already are captured via the HSR constant risk variable.
Auto ownership	This variable could reflect a number of risk factors: • Changes in the economy such as a prolonged recession; • Changes in travelers' perceptions of the auto; • Increase or decrease in the cost of auto ownership due to changes in auto manufacturing; and • Demographic changes (increased numbers of seniors in the population with less likelihood of auto ownership).	• While a number of the risk factors listed might take place, the overall change is likely to be gradual. • In a test run, 0 car households were increased by 20 percent with a commensurate decrease in 2+ car households. While the HSR ridership increased by 8 percent and revenue by 7 percent, the shift in car ownership represented is unlikely to occur by 2029. • The 7-year period from 2006 to 2013 only included a 5-percent increase in 0 car households' share of total households with minor changes in 1-car households. There does not appear to be enough uncertainty in this variable to warrant inclusion. • For year 2040, this risk factor is captured via the auto operating cost, as explained in Appendix J.
Household Income	This variable could reflect a number of risk factors: • Changes in the economy such as a prolonged recession; • Disproportionate in- or out-migration by income groups; and • Demographic changes (increased numbers of retirees on fixed incomes).	• In a test run, low-income households were increased by 10 percent with a commensurate decrease in high-income households (Note: This was coupled with a decrease in number of household workers.). The HSR ridership and revenue decreased by 2 percent. • Impact of this variable can be taken into account through widening of the range for the Trip Frequency constant.

Table A.3 Risk Factors Excluded from Risk Analysis for All Model Years and Operating Phases

Risk Factor	Reasoning for Exclusion of Variable
Earthquake recovery period and other natural disasters	• Typically, ridership (traffic) and revenue studies will assume no natural disasters. • Better place for modeling this risk would be in financial model.
Codeshare with airlines for international/longer trips	• The BPM-V3 model does not focus on this type of travel.
HSR used for freight transport	• Current policy framework for the system does not envision freight usage.
Labor relations/chance of strike/service disruptions risk	• Risk is impossible to model in BPM-V3 model without broad speculation. • Better place for modeling risk would be in financial model.
Hyperloop	• This mode currently is a speculative and unproven technology.

B. High-Speed Rail Constants

The high-speed rail (HSR) constant for each of the four trip purposes (i.e., business, commute, recreation, and other) is composed of two components: 1) unexplained variation, and 2) terminal and wait time. The unexplained variation component represents the desirability to choose HSR that is not captured directly by the system variables included in the model. Terminal time is the out-of-vehicle time spent traveling from the point of departure from the access mode to the train platform. Wait time is the out-of-vehicle time spent waiting on the platform for the train to arrive and the time spent waiting for the train to leave the platform once boarded. The risks associated with each of the components are different and should be specified separately for the Monte Carlo experiments, as discussed in the next sections.

For full model risk analysis runs, terminal and wait time are included with the unexplained variation within the HSR constant. For Monte Carlo risk analysis, each component of the HSR constant is considered as a separate risk variable with completely independent distributions. The former allows for estimation of a single regression model parameter, and does not require an additional risk variable in the experimental design framework. The latter allows for an understanding of the terminal/wait time's effect on ridership and revenue uncertainty independent from the HSR constant's effect on ridership and revenue uncertainty.

B.1 Unexplained Variation

An important part of any mode choice model is a modal constant that explains factors that are not quantifiable by the stated-preference (SP) and revealed-preference (RP) surveys. When dealing with existing modes, such as auto, conventional rail (CVR), and air, we can calibrate this constant by comparing the model outcomes to observed behavior. With a new mode like HSR in the California/U.S., this is impossible, and thus there is uncertainty in the asserted constant.

The asserted HSR constant value comes from averaging the values taken from two approaches. The first approach considered offsets from air and CVR constants derived from 2013 estimated SP constants. The second approach averaged the calibrated air and CVR constants used in model application. Details of the derivation of the HSR constant are documented in *California High-Speed Rail Ridership and Revenue Model Business Plan Model-Version 3 Model Documentation*. Both approaches were reasonable approaches to arrive at an HSR constant, but this analysis takes the average of these values. Since they are both considered reasonable approaches unto themselves, the values derived from each approach must fall within the uncertainty range considered in the risk analysis.

In order to better understand the uncertainty associated with the HSR constant, additional analyses of the 2013 RP/SP survey data was undertaken by performing additional mode choice model estimation using additional variables that were not included in the BPM-V3 model. There variables included demographic characteristics, trip characteristics, and attitudinal questions, as shown in Table B.1.

Table B.1 Additional Variables Considered in Analysis of HSR Constant

Demographic Characteristics	Trip Characteristics	Attitudinal Questions
Gender	Car not available for trip	Respondent's stated likelihood of ever using HSR service in the future.
Age	Car needed at destination	Respondent's perceived economic value of HSR to the State of California.
Worker Status	Duration of stay	Respondent's perceived environmental value of HSR to the State of California.
Highest education level achieved		Respondent's support/opposition level to HSR.
Schedule flexibility		Respondent's familiarity with conventional Amtrak, Acela services in the Northeast, and HSR in foreign countries.

Using the best model with these new variables, the HSR constant was recalibrated, assuming the same calibrated CVR and air constants.[12] The asserted HSR constants under this new model were nearly identical to those of the original model, suggesting that even after controlling for all these additional factors, the constants we would assert for the HSR mode would have been about the same in relation to the calibrated air and CVR constants. Several variables were found to be highly significant, and with expected signs and appropriate magnitudes. However, on their own, these coefficients do not say much about the size or magnitude of the HSR constant, or its relation with CVR or air constants.

Given that this additional model estimation did not provide additional insight into the uncertainty of the HSR constant, we had no basis to narrow the range in uncertainty from the range assumed in the 2014 Business Plan Risk Analysis. The 2014 Business Plan Risk Analysis uncertainty range was based on the assumption that the CVR constant represents an absolute worst case lower bound since there is no apparent reason that any of the unobserved characteristics for the HSR mode should be any worse than those for the CVR mode. In addition, there is no evidence that the uncertainty range is asymmetrical. Moreover, the asserted baseline constants come from averaging two reasonable approaches, as outlined above; both of which are equidistant from the asserted value and by extension,

[12] These constants would change under a different model specification, but this allowed for direct comparison of the resulting HSR constant to those of the original model.

should have the same likelihood of occurrence in the constant distribution. Thus, symmetry is assumed in the constant distribution. If symmetry is assumed, then the absolute minimum bound of the CVR constant also introduces an absolute maximum bound.

Each trip purpose (i.e., business, commute, recreation/other) is treated individually as separate risk factors since the difference between CVR and HSR base constants is different for each purpose, and thus, the CVR lower bound is different for each purpose. In addition, some parts of the uncertainty captured in the constants are considered to be likely correlated amongst trip purposes (i.e., 100 percent correlation), while others would be unrelated between purposes (i.e., 0 percent correlation). A 50-percent correlation between the HSR constant trip purposes was assumed to capture that a portion of the constants would be correlated, but not necessarily every aspect of them.

For the Monte Carlo simulation, a PERT is specified rather than a triangular distribution, because the CVR constant represents an absolute minimum possible value for the HSR constant, essentially a tail event. Since the triangular distribution does not have tails, it would overstate the likelihood of observing a very unlikely tail event.

B.2 TERMINAL AND WAIT TIME

Terminal Time

Terminal time is the out-of-vehicle time spent traveling from the point of departure from the access mode to the train platform. It currently is assumed that terminal times for CVR and air are 3 and 22 minutes, respectively; and for HSR, 10 minutes is assumed. A lower bound based on the CVR value is considered, but given that the size of HSR stations will be larger than many CVR stations, a lower bound for the risk analysis simulations of 5 minutes is more appropriate.

The upper bound on terminal time is based on the air terminal time. An upper bound of 22 minutes is used for HSR terminal time, which is identical to the terminal time assumed at airports. This conservative upper bound assumes that the time it takes to traverse an HSR station is similar to airports, and that HSR travelers will need to undergo security similar to current Transportation Security Administration (TSA) security at airports.

Wait Time

Wait time is the out-of-vehicle time spent waiting on the platform for the train to arrive and the time spent waiting for the train to leave the platform once boarded. The base wait and terminal times for HSR are set to 15 and 10 minutes, respectively. These were the terminal and wait times that were stated in both the 2005 and 2012/2013 RP/SP survey. Wait times are often related to service headways, except when headways grow too long. For long headways, people will

coordinate arrivals to coincide with train departure. For instance, Fan and Machemehl (2009) found that when headways exceeded 38 minutes, bus arrivals became fully coordinated (that is, no one arrived randomly). They also found that for headways under 11 minutes, arrivals became perfectly random. Assuming wholly random arrivals with headways no greater than 60 minutes for HSR, an absolute upper bound on average wait time would be 30 minutes (one-half of headway). However, based on the evidence, at 60 minutes, there would no longer be random arrivals. Therefore, an upper limit of 25 minutes is reasonable, as it represents the average maximum amount of time individuals would wait for an HSR train given nonrandom arrivals.

HSR headways also are at least 30 minutes in our current setup. With random arrivals, this would suggest an average wait time of 15 minutes, which is exactly what currently is used by the model. If it is assumed that with 30-minute headways, 25 percent of travelers have random arrivals with 15-minute average waits, 50 percent of travelers have coordinated arrivals with 10-minute average waits, and 25 percent of travelers have coordinated arrivals with 5-minute average waits; the overall average wait time is 10 minutes. Thus, 10 minutes is used as a lower bound on the distribution for risk analysis.

The wait time and terminal time risk variables for each trip purpose are 100 percent correlated with each other, since factors that contribute to shorter or longer terminal and wait times would not differ by trip purpose. The risk variable has a triangular distribution since the ranges do not reflect extreme or highly unlikely events.

C. Trip Frequency Constants

The trip frequency constants include the unexplained variation in the propensity of households to make long-distance trips within California. Within our risk analysis model, variation in the trip frequency constants is used to reflect the effect of the state of the economy on the proclivity of households to take high-speed rail (HSR). Instead of including distributions of household and employment levels directly as risk variables in the risk analysis model to account for changes in the state of the economy, risks associated with the state of the economy are accounted for within the trip frequency constant risk variable. The risks associated with each of the components are different and should be specified separately for the Monte Carlo experiments, as discussed in the next sections.

C.1 Unexplained Variation

The trip frequency model was calibrated to 2010 conditions and applied using forecast year socioeconomic data and networks. The changes in the demographic composition and the networks in the modeled forecast years result in an increase in annual long-distance trip rates compared to the year 2010 trip rates. This increase in annual long-distance trip rates is consistent with findings from the 1995 American Traveler Survey and the 2001 National Household Travel Survey, which found a 21-percent increase in annual round trips per household over the six-year period from 10.15 annual trips per household to 12.32 annual trips per household.[13] This occurred even though the economic conditions in 2001 were not as good as in 1995 due to the "dot-com" bust. In addition, since some surveys were collected after 9/11, the 2001 National Household Travel Survey (NHTS) trip rates may have been affected.

Annual long-distance trip rates over time are relatively stable and independent of disruptions caused by economic conditions, changes in technology, and changes in traveler perceptions and behavior. Information and communication technologies have been found to be a complement, and even be an incentive for, business trips.[14] During recessions and hard economic times, research has found that households choose to make more leisure trips closer to home for shorter

[13] Source: NCHRP Report 735.

[14] Aguilera, Anne (2008). Business Travel and Mobile Workers. Transportation Research Part A: Policy and Practice, Volume 42, Issue 8, October 2008, pages 1109 to 1116.

Mokhtarian, Patricia (2008). If Telecommunication is such a good substitute for travel, why does congestion continue to get worse? Transportation Letters, Volume 1, Issue 1, January 2009, pages 1 to 17.

periods of time, rather than taking longer trips that involve more days away from home.[15] As the baby boomers continue to move into retirement age, leisure travel also may increase due to fewer family obligations, higher incomes compared to their younger peers, and fewer necessary expenditures.[16] Research suggests that, if anything, long-distance travel may increase with changing technologies and demographics.

Since changes in economic conditions, technologies, and traveler perceptions and behaviors are not hypothesized as a significant risk to annual long-distance trip rates, the trip frequency constant risk factor range is based on the range seen in forecasted annual long-distance trip rates produced by the model. The most likely value for each forecast year is the calibrated constant. The minimum value of the trip frequency constants is specified, such that for year 2040, the trip frequency constants produce average trip rates equal to the 2010 rates by trip purpose. The maximum value of the trip frequency constant is specified to mirror the deviations from the calibrated constants for the minimum values (i.e., symmetry of the constant offsets is assumed).

For each trip purpose (i.e., business/commute, recreation/other), some parts of the uncertainty captured in the constants are considered likely to be correlated amongst trip purposes (i.e., 100 percent correlation), while others would be unrelated between purposes (i.e., 0 percent correlation). A 50-percent correlation between the trip frequency constant trip purposes was assumed to capture that a portion of the constants' uncertainty would be correlated, but not necessarily every aspect of it.

For the Monte Carlo simulation, a PERT is specified rather than a triangular distribution, because the minimum and maximum values represent unlikely events. Since the triangular distribution does not have tails, it would overstate the likelihood of observing a very unlikely tail event.

Table C.1 shows the approximate results in terms of annual long-distance round trips per capita resulting from the specification of the constant ranges to account for unexplained variation. Note that symmetry of the constant offsets does not produce symmetry of the implied trip rates. This is due to the trip frequency choice model being specified as a logit model with choices of no long-distance trip, one long-distance trip traveling alone, or one long-distance trip traveling in a group on a given day. Since the base shares for each of these choices are very low (e.g., ~0.2 percent), the model is more sensitive to the constants on the high end than the low end.

---

[15] Lamonda, Jeff, C. Bhat (2011). A study of visitors' leisure travel behavior in the northwest territories of Canada. Transportation Letters, Volume 3, Issue 1, January 2011, pages 1 to 19.

[16] Lamonda, Jeff, C. Bhat, and D. Hensher. An annual time use model for domestic vacation travel. Journal of Choice Modeling, Volume 1, Issue 1, 2008, pages 70 to 97.

Table C.1 Unexplained Variation of Trip Frequency Constants – Implied Annual Long-Distance Round Trips Per Capita

	Implied Annual Long-Distance Trips per Capita for 2025			Implied Annual Long-Distance Trips per Capita for 2029			Implied Annual Long-Distance Trips per Capita for 2040		
Purpose	Minimum	Most Likely	Maximum	Minimum	Most Likely	Maximum	Minimum	Most Likely	Maximum
Business/ Commute	1.64	2.16	2.85	1.67	2.20	2.90	1.86	2.45	3.23
Recreation/ Other	5.10	5.76	6.50	5.18	5.85	6.60	5.52	6.23	7.03
Total	**6.74**	**7.92**	**9.35**	**6.85**	**8.05**	**9.50**	**7.38**	**8.68**	**10.26**

C.2 Economic Cycle

Economic cycles potentially impact several different variables in the model, including the number of workers, household income levels, and overall trip making. However, incorporating each of these risk factors separately is infeasible. These impacts are interrelated and can be accounted for jointly. Sensitivity tests have shown that the economic-cycle variations can be reasonably accounted for by changes in trip frequencies. Thus, the effect of economic cycles on HSR ridership and revenue is accounted for as a separate risk component in the trip frequency constants.

In order to determine the appropriate range in the trip frequency constant, changes in employment and income need to be translated into changes in the trip frequency constants. The primary driver for long-distance trip-making in the BPM-V3 model is the number of households within the State. Households are stratified into 99 different groups based on 4 household size groups, 3 auto ownership groups, 3 number of workers groups, and 3 income groups. The 4 x 3 x 3 x 3 groups result in 108 strata; 9 of which are illogical (i.e., 2 or more worker, 1 person households for the 9 groups defined by auto ownership and income). Total trips are based on the modeled trip frequency and the numbers of households in the State.

Employment is the metric used to define the economic cycles for the State. Employment has a secondary impact on trip frequency and a more direct impact on destination choice. However, the employment levels also can be used to more directly impact the total numbers of trips through relationships with households by numbers of workers and households by income group. For a given forecast of households, the numbers of 0, 1, and 2+ worker households should vary so that total workers in the State track the total employment. Likewise, in a recession, it should be expected that the number of low-income households should increase at the expense of middle- and high-income households and, likewise, that the

number of middle-income households might increase at the expense of high-income households.

Suggested low and high employment levels representing the economic cycles are based on historic observations through 2014. The Great recession produced a -2.8-percent Compound Annual Growth Rate (CAGR) for employment in California between 2007 and 2010. Thus, for the low economic growth scenario, annual declines of 3.0 percent per year for the three years preceding the forecast year were assumed, with those decline being applied to the new "Low Scenario" statewide control total, as described in detail in Appendix G. The period from 1994 to 2000 was the high water period for job growth in California with a 3.0-percent CAGR for five years. Thus, for the high economic growth scenarios, annual increases of 3.0 percent per year for the five years preceding the forecast year were assumed, with the increase applied to the new "High Scenario" statewide control total.

It was assumed that the low employment forecast would result in a commensurate decrease in the number of household workers. This was accomplished through increasing in the number of 0 and 1 worker households, and decreasing 2+ worker households. The above changes could result from some households moving from 2+ worker households to 1 worker households, and 1 worker households moving to 0 worker households to reflect the increasing unemployment. It was assumed that the increase in 0 worker households would result in an increase in low-income households and a commensurate decrease in high-income households. The changes could result from some households moving from high-income households to middle-income households, and middle-income households moving to low-income households to reflect the increasing unemployment or underemployment.

It was assumed that the high employment forecasts would result in a commensurate increase in the number of household workers. This was accomplished through decreases in the number of 0 worker household and 1 worker households, and increases in 2+ worker households to maintain the statewide control total of households. Low-income households were assumed to decrease and middle-income households were assumed to increase. Table C.2 shows examples of the resulting joint distributions of households by number of workers and income group and the resulting factors for the base, low, and high employment scenarios.

In order to determine the compounding effects of income levels and number of workers on HSR long-distance trips per capita, the full model was run for each forecast year for the low and high economic growth socioeconomic datasets. The resulting trip rates for year 2025, 2029, and 2040 are shown in Table C.3.

Table C.2 Workers per Household by Income Group for Most Likely, Minimum, and Maximum Changes in Employment for 2040

Workers/ Household	Income Group			Total
	Low	Middle	High	
Base Scenario				
0	17%	9%	6%	**32%**
1	9%	12%	12%	**33%**
2+	3%	9%	23%	**35%**
Total	29%	30%	41%	**100%**
Minimum Economic Growth Scenario				
0	19%	11%	6%	**36%**
1	10%	14%	12%	**36%**
2+	2%	8%	18%	**28%**
Total	31%	33%	36%	**100%**
Maximum Economic Growth Scenario				
0	14%	8%	6%	**28%**
1	7%	11%	13%	**31%**
2+	2%	10%	29%	**41%**
Total	23%	29%	48%	**100%**

Table C.3 Annual Long-Distance HSR Trips per Capita for Most Likely, Minimum, and Maximum Employment Scenarios

	Base Scenario	Minimum Employment Scenario		Maximum Employment Scenario	
Purpose	Trip Rates	Trip Rates	% Change from Base	Trip Rates	% Change from Base
Year 2025					
Business/Commute	0.037	0.029	-21.6%	0.044	18.9%
Recreation/Other	0.047	0.043	-8.5%	0.049	4.3%
Total	0.084	0.073	-13.1%	0.093	10.7%
Year 2029					
Business/Commute	0.165	0.136	-17.6%	0.207	25.5%
Recreation/Other	0.243	0.227	-6.6%	0.263	8.2%
Total	0.408	0.363	-11.2%	0.470	14.9%
Year 2040					
Business/Commute	0.182	0.143	-21.4%	0.224	23.1%
Recreation/Other	0.259	0.238	-8.1%	0.278	7.3%
Total	0.440	0.381	-13.4%	0.502	14.1%

Since the risk analysis is focusing on HSR revenue and ridership, the per capita HSR trip rates shown in Table C.3 were used to factor the overall base trip rates (e.g., estimated overall trip rate x estimated low HSR trip rate/estimated base HSR trip rate). This effort produced alternate ranges for the overall trip rates that, in effect, included the compounding impacts of the input socioeconomic data and the transportation networks on HSR ridership. Table C.4 shows the ranges that result based on total trips and based on the adjustment for HSR shares. Due to the impacts of the various input data on destination choice and mode choice in the BPM-V3, the minimum and maximum values for the HSR adjusted rates could be different than the overall trip rates resulting from the application of the BPM-V3 (i.e., total forecast long-distance trips from the trip frequency model/total population). For example, the maximum values for the HSR adjusted trips for Business/Commute for all three forecast years are higher than the unadjusted maximum rates. The trip rate ranges are used to specify the ranges of trip frequency model constants for the full model runs that produce data for the calibration of regression models for the risk analysis. Thus, it is reasonable to use the minima and maxima for the ranges (as shown in bold italics in Table C.4). Table C.5 shows the range of annual long-distance round trips per capita resulting along with the constant offsets.

Table C.4 Range in Annual Total Round Trips per Capita Based on Total Trips and Based on Adjustment for HSR Shares

		Ranges Based on Total Trips			Ranges Based on Adjustment for HSR Trips		
Model Year	Purpose	Minimum	Most Likely	Maximum	Minimum	Most Likely	Maximum
2025	Business/Commute	1.76	2.16	2.50	***1.71***	2.16	2.54
	Recreation/Other	***5.31***	5.76	***6.15***	5.37	5.76	6.06
	Total	7.07	7.92	8.65	7.08	7.92	8.60
2029	Business/Commute	1.85	2.20	2.68	***1.80***	2.20	***2.75***
	Recreation/Other	***5.43***	5.85	6.36	5.47	5.85	6.32
	Total	7.28	8.05	9.04	7.27	8.05	9.07
2040	Business/Commute	1.98	2.44	2.94	***1.92***	2.44	***3.02***
	Recreation/Other	***5.72***	6.22	***6.73***	5.72	6.22	6.68
	Total	7.70	8.66	9.67	7.64	8.66	9.70

Table C.5 Minimum, Most Likely, and Maximum Economic-Cycle Trip Frequency Constant Offsets and Implied Trip Rates

		Constant Offsets (From Calibrated Constants)			Implied Trip Rates After Applying Offsets		
Model Year	Purpose	Minimum	Most Likely	Maximum	Minimum	Most Likely	Maximum
2025	Business/Commute	-0.233	0	0.165	1.71	2.16	2.54
	Recreation/Other	-0.070	0	0.052	5.37	5.76	6.06
	Total	–	–	–	***7.08***	***7.92***	***8.60***
2029	Business/Commute	-0.201	0	0.224	1.80	2.20	2.75
	Recreation/Other	-0.068	0	0.078	5.47	5.85	6.32
	Total	–	–	–	***7.27***	***8.05***	***9.07***
2040	Business/Commute	-0.246	0	0.209	1.92	2.44	3.02
	Recreation/Other	-0.087	0	0.071	5.72	6.22	6.68
	Total	–	–	–	***7.64***	***8.66***	***9.70***

C.3 Trip Frequency Constant Ranges

For full model risk analysis runs, economic-cycle effects are included with the unexplained variation in the range specified for the trip frequency constant. The range of constant offsets for the uncertainty analysis is directly related to the calibrated constants. The range of constant offsets for impacts of economic cycles provides proxies for the actual economic-cycle risk variable being considered. This approach provides a useful method for specifying a continuous range of outcomes rather than developing multiple input socioeconomic datasets. The offsets must be combined to represent the full range of possible outcomes for the development of the risk analysis regression equations. The constant offsets for the Unexplained Variation and Economic Cycle are added and the implied range of trip rates was estimated, as shown in Table C.6.

Table C.6 Range of Trip Frequency Constant Offsets and Implied Trip Rates for Full Model Runs

		Composite Trip Frequency Model Constant Offsets			Implied Trip Rates Based on Composite Constant Offsets		
Model Year	Purpose	Minimum	Most Likely	Maximum	Minimum	Most Likely	Maximum
2025	Business/Commute	-0.511	0	0.443	1.30	2.16	3.35
	Recreation/Other	-0.193	0	0.175	4.76	5.76	6.84
	Total	–	–	–	6.06	7.92	10.19
2029	Business/Commute	-0.479	0	0.502	1.37	2.20	3.62
	Recreation/Other	-0.191	0	0.201	4.84	5.85	7.13
	Total	–	–	–	6.21	8.05	10.75
2040	Business/Commute	-0.524	0	0.487	1.45	2.44	3.97
	Recreation/Other	-0.210	0	0.194	5.06	6.22	7.54
	Total	–	–	–	6.51	8.66	11.51

For Monte Carlo risk analysis, each component of the trip frequency constant is considered as a separate risk variable with completely independent distributions (i.e., 0 percent correlation). About 50-percent correlation is assumed between the business/commute and recreation/other risk components for unexplained variation. Perfect correlation is assumed between economic-cycle risk components for business/commute and recreation/other purposes.

D. Auto Operating Cost

The approach for forecasting auto operating costs for the 2016 Business Plan is consistent with the methodology used for the 2014 Business Plan, with updates to recognize the following:

- The most current motor gasoline price projections based on EIA's 2015 Annual Energy Outlook (AEO);
- Revised non-gasoline operating costs;
- The most current fuel efficiency projections of the on-the road vehicle fleet; and
- Effects of Cap and Trade rules in motor fuel prices and potential effects of an increase in the Federal excise tax rate.

The auto operating costs documented in this appendix are for privately owned vehicles. Appendix H provides background on auto operating costs for autonomous and shared use vehicles and their impacts on overall auto operating costs as used for the 2040 Phase 1 – Blended risk analysis.

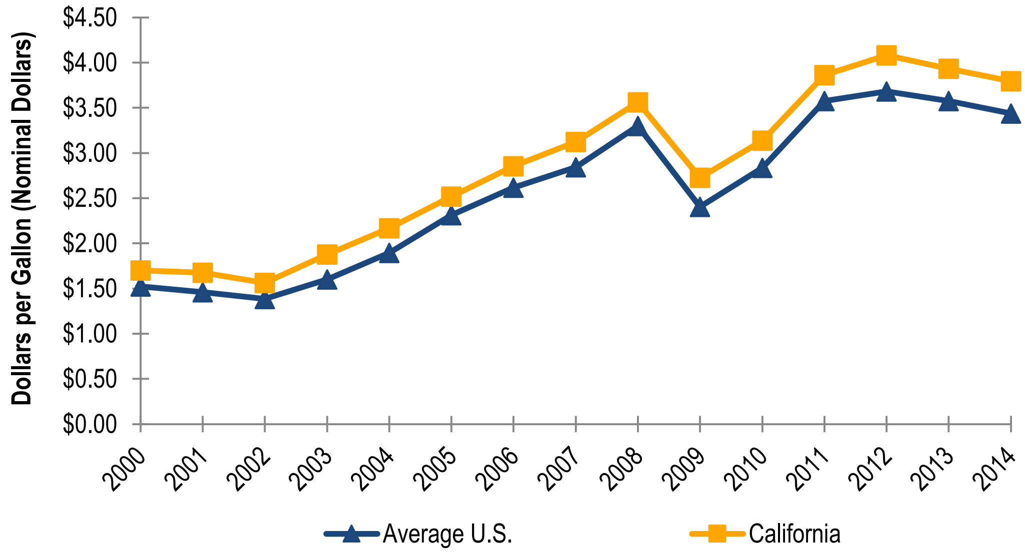
The following sequential steps were undertaken to calculate the auto operating cost:

1. Project retail fuel prices in California;
2. Adjust additional fees and charges based on two scenarios:
    a. Cap and Trade, and
    b. Potential increase in Federal excise tax;
3. Project fuel economy of the entire "on the road" fleet;
4. Estimate nonfuel costs; and
5. Combine fuel operating cost with nonfuel operating cost.

D.1 Fuel Prices

Historically, California retail gasoline prices have been higher than the U.S. average. As shown in Figure D.1, from year 2000 to 2014, the overall average for California prices was fairly consistently 11 percent higher than the U.S. average.

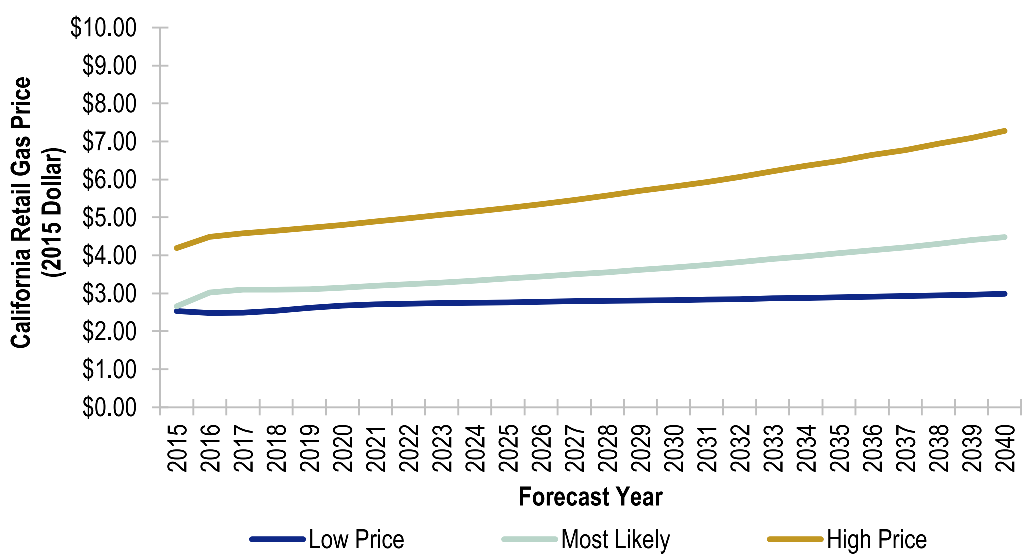
Figure D.1 Annual Retail Gasoline Prices



Source: U.S. Energy Information Administration: Annual All Grades All Formulations Retail Gasoline Prices http://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_nus_m.htm.

The U.S. Energy Information Administration (EIA) forecasts motor gasoline prices through 2040 for three different scenarios in its 2015 Annual Energy Outlook (AEO): reference, low, and high. The projections were increased by 11 percent to develop projections of retail gas prices in California, as shown in Figure D.2.

Figure D.2 Low, Reference, and High California Retail Gas Price



Source: EIA, AEO2015 National Energy Modeling System.

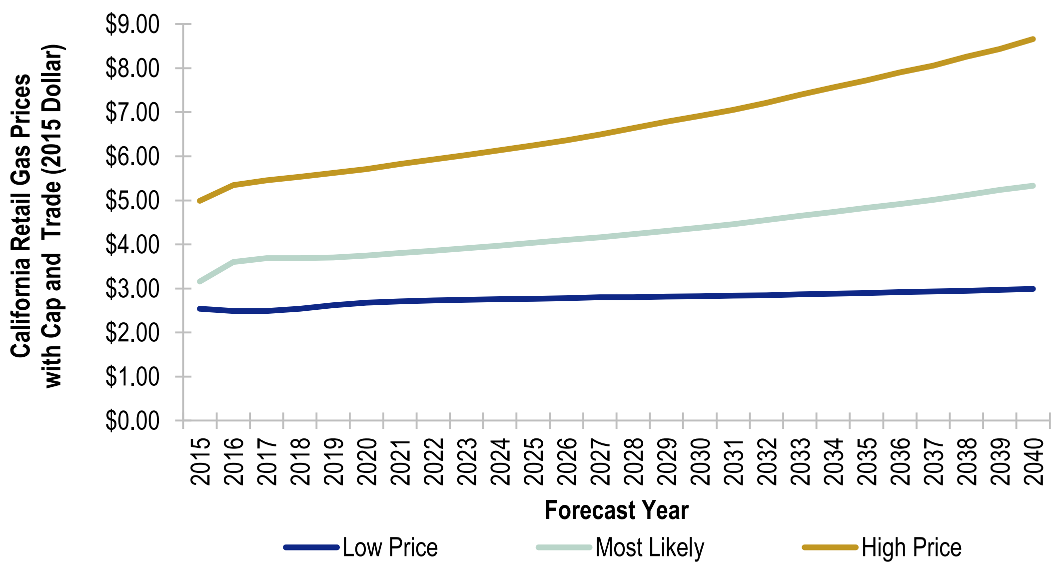
D.2 Cap and Trade Effects on Fuel Prices

On January 1, 2015, the cap-and-trade rules came into effect for the fuel sector in California. The California Air Resources Board estimated in 2010 that gasoline price changes in 2020 could range between 4 percent and 19 percent due to Cap and Trade rules.[17] The exact impact of Cap and Trade on fuel prices is unknown and could change over time based on the industry response to reduce emissions. Cap and Trade scenarios assumed the following impacts:

- Low scenario would assume a repeal of the Cap and Trade rules and, thus, no impact from Cap and Trade (i.e., 0 percent increase in retail gasoline price);
- High scenario would assume the maximum predicted effect of Cap and Trade (i.e., 19 percent increase in retail gasoline price); and
- Most likely scenario would assume the midpoint impact of Cap and Trade between the maximum and minimum (i.e., 9.5-percent increase in retail gasoline price).

Figure D.3 shows the low, most likely, and high total California fuel cost projections, including these offsets for Cap and Trade.

Figure D.3 Cap and Trade Scenario Total California Retail Fuel Price



[17] California Air Resource Board, http://www.arb.ca.gov/regact/2010/capandtrade10/capv4appn.pdf.

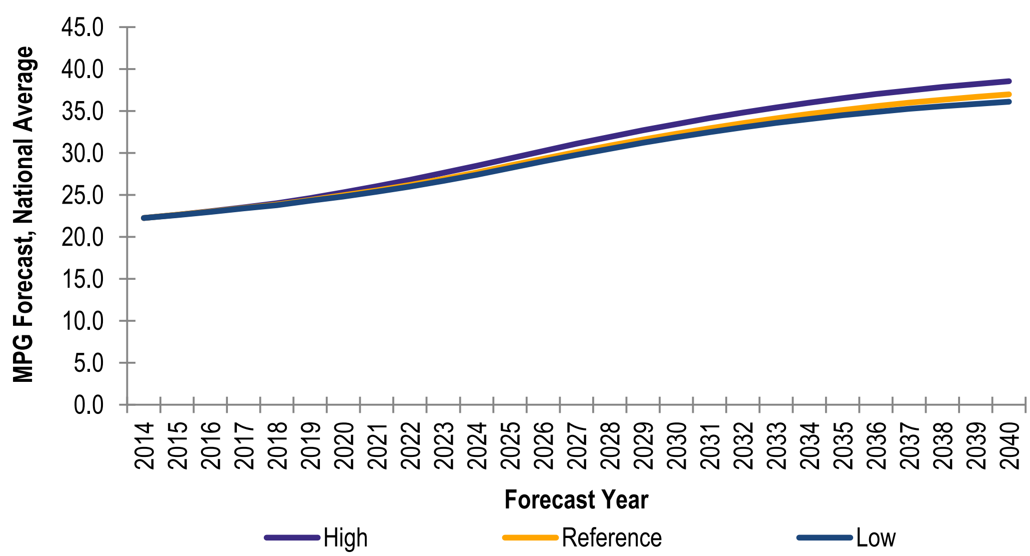
D.3 Federal Fuel Tax Increase Scenario

For the maximum auto operating cost scenario only, it is assumed that the Federal Government introduces a bill that links the Federal fuel tax to the Consumer Price Index. Today, the Federal Fuel Tax is $0.184 per gallon. If the Federal Fuel Tax is increased based on adjustment for Consumer Price Index (CPI) changes, which are assumed at 2.4 percent per year increase retroactive to year 1993 (i.e., last gas tax increase), then the Federal Fuel Tax would be $0.302 per gallon today. This results in the maximum scenario adding an additional $0.12 tax to the Fuel Cost projection (i.e., $0.30 – $0.18 = $0.12).

D.4 Projections of Fuel Economy of Light-Duty Vehicles

U.S. National Average shown in Figure D.4 is used for the assumptions of Fuel Economy projections in California.[18] For calculating the minimum auto operating cost, the high miles per gallon (MPG) forecast was coupled with the Low gasoline price forecast; and for the maximum auto operating cost, the low MPG forecast was coupled with the high gasoline price forecast.

Figure D.4 National Average Fuel Economy Forecasts

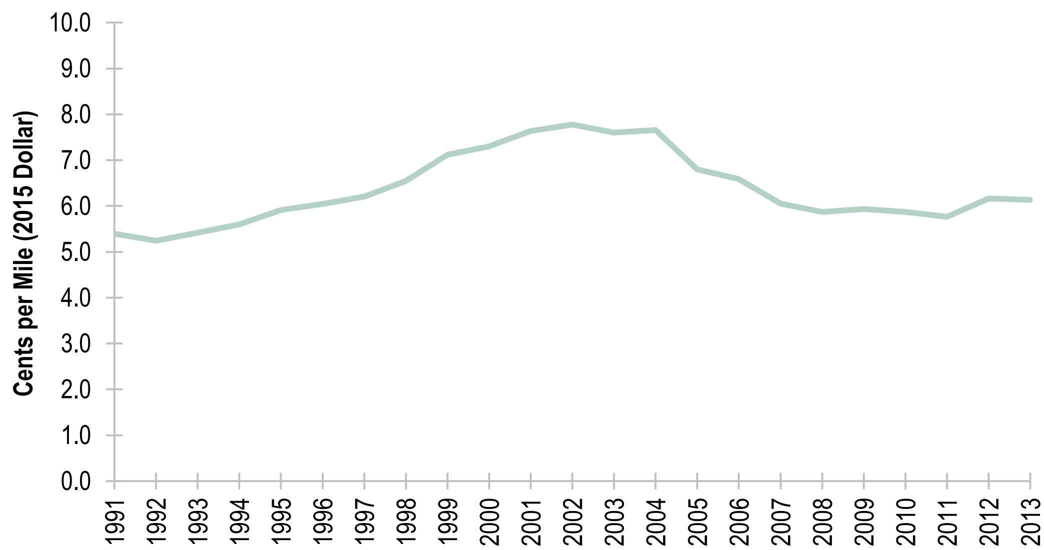


[18] U.S. Energy Information Association (2015). Annual Energy Outlook 2015 with projections to 2040. DOE/EIA-0383 (2015), April 2015.

D.5 NONFUEL OPERATING VEHICLE COST

The Bureau of Transportation Statistics (BTS) publishes historical average nonfuel auto operating costs. The total cost of owning and operating an automobile include fuel, Maintenance, Tires, insurance, license, registration and taxes, depreciation, and finance. Figure D.5 illustrates the nonfuel auto operating cost per mile between 1991 and 2014. The low nonfuel auto operating cost scenario is calculated as the minimum nonfuel cost between 1991 and 2014 (i.e., 5 cents per mile). The high nonfuel auto operating cost scenario is calculated as the maximum nonfuel cost between 1991 and 2014 (i.e., 8 cents per mile). The most likely value is the current nonfuel auto operating cost (i.e., 6 cents per mile).

Figure D.5 Historical Nonfuel Operating Vehicle Cost



Source: CPI, BLS, All Urban Consumer, National Average: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/html/table_03_17.html.

D.6 RANGE OF AUTO OPERATING COST

The following formulas were used to calculate the minimum, most likely, and maximum auto operating cost:

Minimum Auto Operating Cost = ((Low CA Gas Price + No C&T Impact + No Increase in Federal Gas Tax)/High Fuel Efficiency) + Low Nonfuel Operating Costs

Most Likely Auto Operating Cost = ((Most Likely CA Gas Price + Avg (C&T No Impact, C&T High Impact) + No Increase in Federal Gas Tax)/Most Likely Fuel Efficiency) + Most Likely Nonfuel Operating Costs

High Auto Operating Cost = ((High CA Gas Price + C&T High Impact) + Increase in Federal Gas Tax)/Low Fuel Efficiency) + High Nonfuel Operating Costs

Table D.1 gives the auto operating cost component values and the resulting minimum, most likely, and maximum auto operating cost for each forecast year before adjusting for the impact of autonomous and shared vehicles for 2040 forecasts.

Table D.1 Range of Auto Operating Cost for each Forecast Year by Auto Operating Cost Component

2015 Dollars

	Minimum	Most Likely	Maximum
2025 Auto Operating Cost ($/mile)	**$0.15**	**$0.20**	**$0.31**
California Gas Price (CA Gas Price)	$2.78	$3.41	$5.28
Cap & Trade (C & T Impact)	$0.00	$0.32	$1.00
Increase in Federal Gas Tax	$0.00	$0.00	$0.12
Fuel Efficiency (mpg)	29.4	28.5	28.2
Total Fuel Operating Cost ($/mile)	$0.09	$0.13	$0.23
Nonfuel Operating Cost ($/mile)	$0.05	$0.06	$0.08
2029 Auto Operating Cost ($/mile)	**$0.14**	**$0.19**	**$0.30**
California Gas Price (CA Gas Price)	$2.83	$3.63	$5.73
Cap & Trade (C & T Impact)	$0.00	$0.35	$1.09
Increase in Federal Gas Tax	$0.00	$0.00	$0.12
Fuel Efficiency (mpg)	32.7	31.6	31.2
Total Fuel Operating Cost ($/mile)	$0.09	$0.13	$0.22
Nonfuel Operating Cost ($/mile)	$0.05	$0.06	$0.08
2040 Auto Operating Cost ($/mile)	**$0.13**	**$0.19**	**$0.32**
California Gas Price (CA Gas Price)	$3.00	$4.54	$7.32
Cap & Trade (C & T Impact)	$0.00	$0.43	$1.39
Increase in Federal Gas Tax	$0.00	$0.00	$0.12
Fuel Efficiency (mpg)	38.6	37	36.1
Total Fuel Operating Cost ($/mile)	$0.08	$0.13	$0.24
Nonfuel Operating Cost ($/mile)	$0.05	$0.06	$0.08

E. Coefficient on Transit Access-Egress Time/Auto Distance Variable

Between some regions in California, especially in the VtoV scenario, individuals who wish to travel primarily by transit to reach their destination must transfer from a high-speed rail (HSR) bus or conventional rail (CVR) system before or after traveling on HSR. There is uncertainty around how the need to make these transfers affects the overall desirability of traveling by HSR. The uncertainty in the desirability of travel by HSR, when the CVR or HSR bus leg of the journey is relatively long in relation to the HSR travel length, has an impact on ridership and revenue. Thus, this uncertainty was included as a potential risk variable.

E.1 Options for Addressing Risk in Uncertainty Analysis

Two primary options were considered for addressing the transit transfer concern in the context of the risk analysis. The first option considers a range for the constant associated with the transit access/egress to the HSR main mode. The main advantage of this approach is its simplicity. The range used for the constant would come directly from conversion of a penalty value (in minutes) to utility. The main disadvantage is that the same range would need to be applied to all transfers between access/egress transit modes and HSR. This means that the penalty would apply equally to transfers between local transit (e.g., someone taking a city bus from their home to the station) and HSR and transfers between CVR or HSR bus and HSR with longer access trips. Transfer between local transit and CVR exist today and thus are accounted for within the model estimation of this variable, while transfers between CVR or HSR bus to HSR have not been observed in the estimation dataset. Moreover, it means the penalty would not vary on the basis of how long the trip was or how much of the trip was transit versus HSR.

The second option considers a range for the parameters associated with transit access/egress travel times relative to OD distances. This variable appears in the access and egress modal utility functions as follows:

$$\boldsymbol{\beta} \times \mathrm{MAX}\left(\mathbf{0}, \frac{[\boldsymbol{Acc\ or\ Egr\ Time}]}{[\boldsymbol{OD\ Distance}]} - \boldsymbol{Threshold}\right)$$

In the base model, several threshold parameter options were tested in model estimation, and a value of 0.2 was ultimately identified. The values of beta (the

variable coefficient) were estimated directly and were found to be negative. Separate coefficients were estimated for auto access/egress modes versus nonauto access/egress modes (transit and walk/bike), with the magnitude of auto coefficients estimated to be much larger. This variable essentially provides a disincentive for selecting a main mode that requires a long access or egress time, relative to the entire trip length.

The main advantage of the second option is that this differentiation would naturally occur between local transit and longer CVR or HSR bus connections. Since local transit connections would typically be very short distance and CVR or HSR bus may be short or long distance, the "penalty" associated with transit access/egress would reflect the access/egress mode's overall share of the total trip length. The second option is more appropriate for the risk analysis. The uncertainty associated with the variable is only applied for the HSR main mode (i.e., not air or CVR).

E.2 Development of the Range in the Risk Variable Parameters

Figures E.1 and E.2 show the variable's effect under the current model specification (in terms of equivalent minutes[19] of travel time) for the recreation/ other purpose (the results are very similar for business and commute trip purposes). The first plots penalty versus origin-destination (OD) distance for constant egress times, and the second plots penalty versus egress time for constant OD distance values. The same concepts apply to the access end of trips. The egress end is shown only as an example; the access time graph looks identical.

In both figures, certain regions of the graphs suggest very high penalties for certain types of trips. For instance, Figure E.1 shows very high penalties for the 100-minute egress line when OD distance is less than 100 miles. Likewise, Figure E.2 shows very high penalties for the 50-mile OD distance line when egress time is high[20]. Travelers typically do not make trips of this nature, since other main modes

[19] "Equivalent minutes of travel time" is estimated by dividing a constant or a variable by the coefficient associated with travel time. Equivalent minutes of travel time provides a convenient way to measure the magnitude of "unexplained variation" of a model constant using an understandable metric and to compare values among different models. Equivalent minutes of travel time is a derived measures that can be computed for any model variable. So, for example, a $72 HSR fare (2005 dollars) for an interchange in the recreation/other mode choice model would equate to 337 equivalent minutes of travel time while the implied equivalent minutes of travel time savings for group travel in an auto for the interchange would equate to a savings of 619 equivalent minutes of travel time. Note, however, these variables are important for their contributions to the mode choice utility function, not as direct measures of travel time.

[20] A chart of penalty versus OD distance for constant access time would look identical to the chart for egress time.

would be highly favored, so these penalty values are very unlikely to be actually applied.

Figure E.1 Penalty versus OD Distance for Constant Egress Times

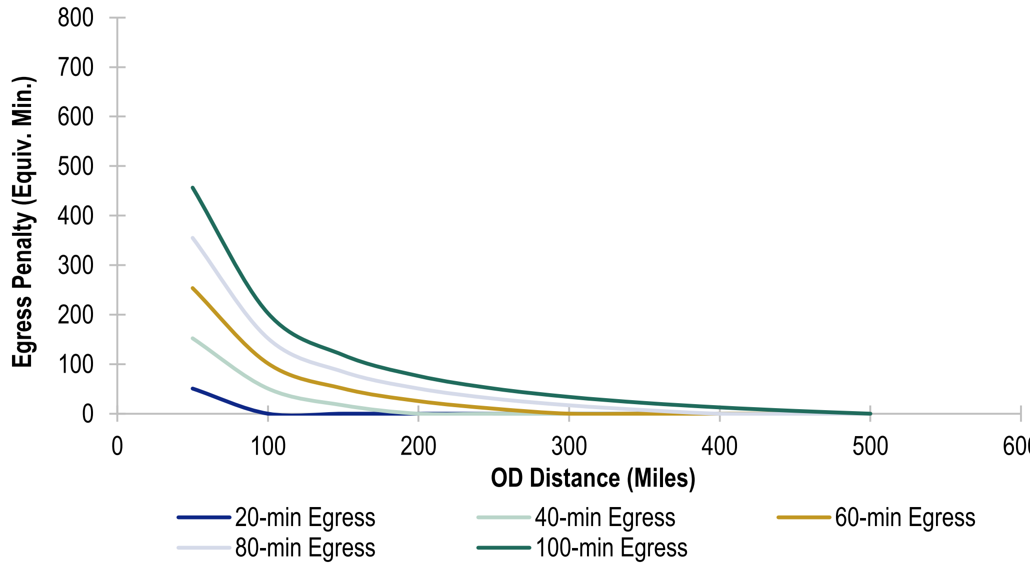
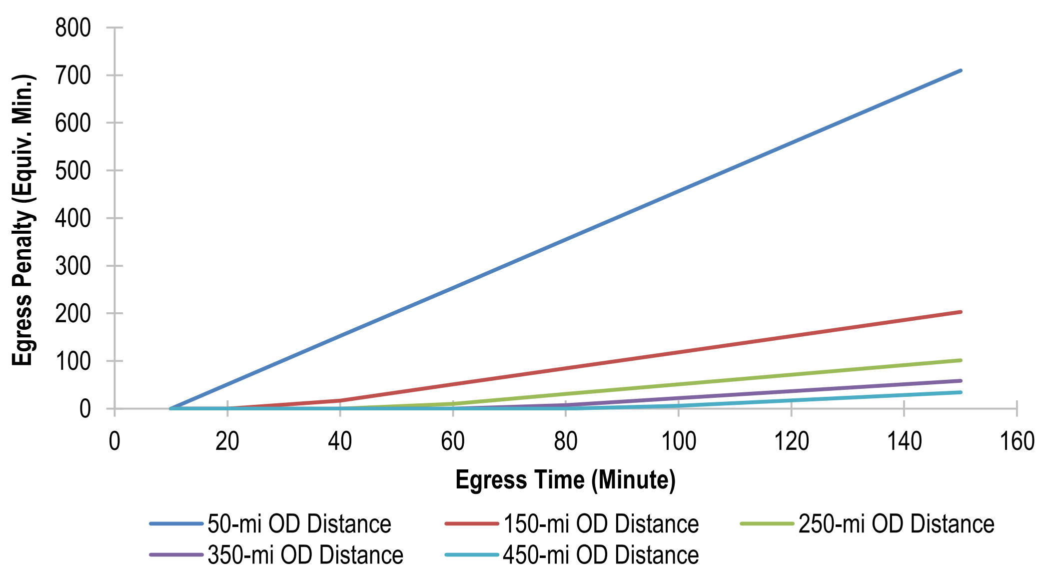


Figure E.2 Penalty versus Egress Time for Constant OD Distance Values



An example of the experience in France was researched to assist in developing a reasonable range for the variable parameters. In the French experience, moving from a direct CVR connection between Paris and Grenoble to an HSR trip from Paris to Lyon and a connection to CVR from Lyon to Grenoble saved 90 minutes

of total travel time, but did not result in increased ridership. The observed "90-minute penalty" in France served as a rough benchmark for determining a lower bound on the model parameters. In the French experience, the trip between Paris and Grenoble via Lyon is about 350 miles total, with about 90 minutes of conventional rail time (Lyon to Grenoble).[21] Using an OD distance of 350 miles, an egress time of 90 minutes, and the aforementioned 90-minute savings from the French experience, several approaches were tested to achieve an appropriate lower bound for the variable.

There are two ways uncertainty affects the variable. First is the coefficient associated with the variable. Second, the variable uses a threshold value, set such that the variable takes a value of zero when the ratio of access/egress time to distance is less than 0.2. The threshold value was set in model estimation by trial and error. The value of 0.2 was selected because it fit the data better than other potential values. However, like the coefficient, there is uncertainty associated with it.

Several options were considered for setting lower bounds for the threshold variable. Based on a review of potential options, threshold values of 0.05 and 0.10 were tested. In both cases, the coefficient on the variable was selected so that the value of the penalty was about 60 minutes for a case similar to the French example. A 60-minute penalty was used instead of the 90-minute penalty observed in the French experience, because it offered more reasonable model behavior overall, and it was not desirable to change the long-distance models in unreasonable ways to match a single observed data point. Figure E.3 and Figure E.4 plot the penalty versus egress time to OD distance ratios for baseline, drive access/egress variables, transit access/egress variables with threshold value of 0.05, and transit access/egress variables with threshold value of 0.10. Figure E.3 shows the results for the business/commute purpose, and Figure E.4 shows results for the recreation/other purpose. The drive access/egress variable is plotted for comparison purposes only, and has no bearing on the variable discussed in this section. It applies when the access/egress mode is an auto mode (rather than transit).

[21] The OD distance and egress times cited for the French experience are approximate, as it is based on Google maps and train timetables. While the network distance is about 350 miles between Paris and Grenoble, the straight line distance is only 300. And, some egress train options took longer than 90 minutes, up to and over 120 minutes.

Figure E.3 Business/Commute Penalty versus Egress Time to OD Distance Ratios for Baseline, Drive Access/Egress Variables, and Transit Access/Egress Variable Options

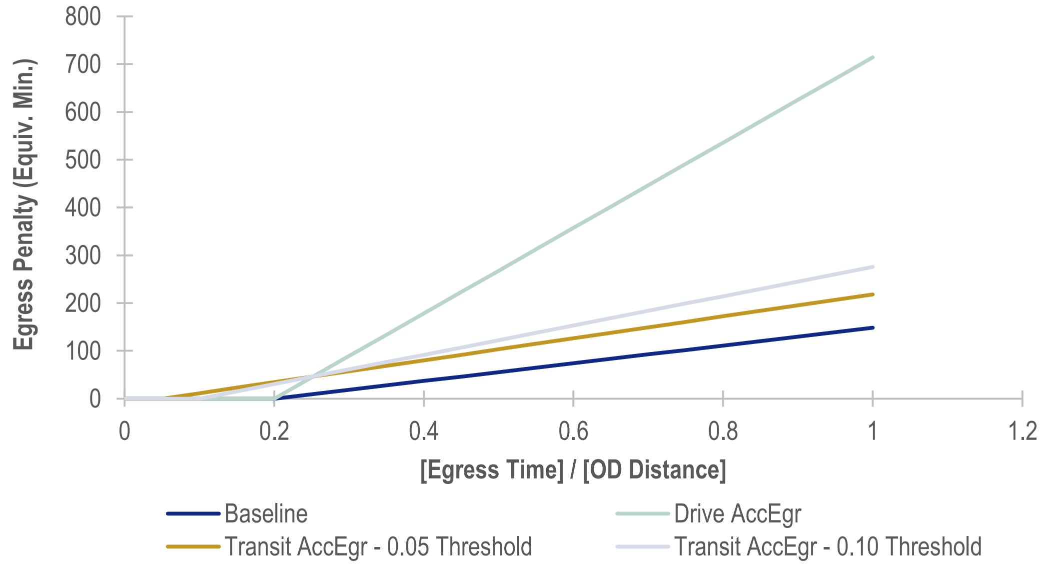
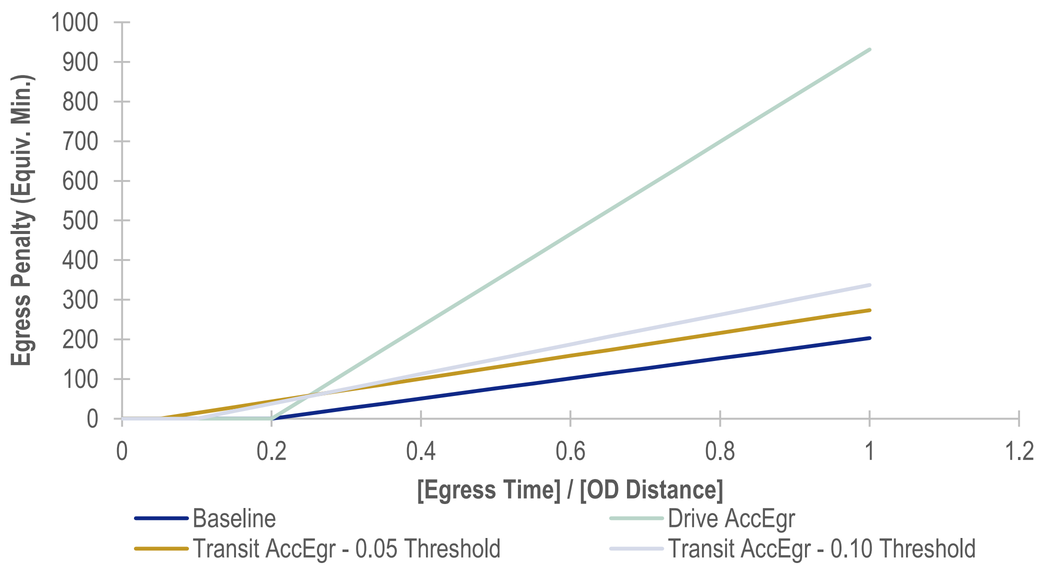


Figure E.4 Recreation/Other Penalty versus Egress Time to OD Distance Ratios for Baseline, Drive Access/Egress Variables, and Transit Access/Egress Variable Options



E.3 Range of Coefficient on Transit Access-Egress Time/Auto Distance Variable

The transit access/egress variable with threshold value of 0.10 was chosen as the low scenario. This threshold was chosen over 0.05, because it causes less disruption to the relationships between the drive and transit access/egress variables for trips with shorter access/egress (e.g., when the ratio of access or egress time to OD distance is around 0.1 to 0.2). The coefficient value is set to -2.0 for business/commute purpose and -1.3 for recreation/other purpose. As described above, these were set to achieve penalty values of about 60 minutes. The coefficient and threshold value are assumed to vary in parallel (i.e., perfect correlation) for the full model runs and Monte Carlo simulation.

The maximum threshold and coefficient values were set identical to the base/most likely values. There is no evidence to suggest that the penalty to transfer from transit to HSR should be less than that used for CVR and air.

F. Number and Distribution of Statewide Households and Employment

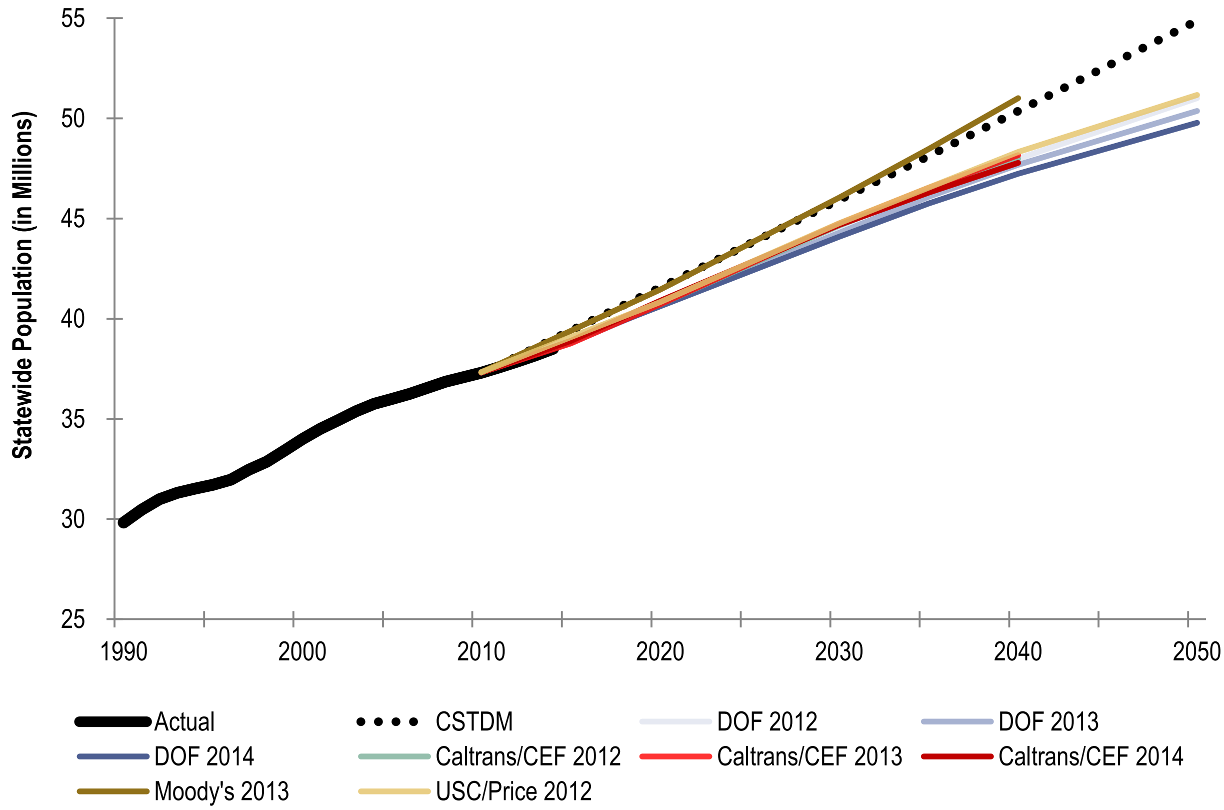
The risk analysis conducted to create the 2014 Business Plan (BP) forecasts of statewide population, household, and employment were used as a starting point for developing the 2016 BP risk analysis forecasts. The 2014 BP forecasts were improved by additional historical data to 2014 and obtaining new forecasts from a variety of sources:

- New national forecast was obtained from the U.S. Census Bureau, showing a reduction in national population projections (U.S. Census, 2014).
- New statewide forecasts were obtained from the Department of Finance (DOF, 2014) and the California Economic Forecast (CEF, 2013 and 2014). These forecasts were mostly unchanged through year 2025, but decreased beyond year 2025, resulting in a one-percent decrease in forecasted population in year 2040.
- New regional forecasts were obtained from the Sacramento Area Council of Governments (SACOG) and Southern California Association of Governments (SCAG).

Figure F.1 shows the range in statewide forecasts for each of the obtained data sources. County-level population forecasts were assembled from the following sources:

- California Statewide Travel Demand Model (CSTDM);
- U.S. Census Bureau;
- Moody's Analytics (Economy.com);
- California DOF;
- California Employment Development Department;
- CEF;
- University of Southern California (Price School); and
- Metropolitan Planning Organizations (MPO): Metropolitan Transportation Commission (MTC), SACOG, San Diego Association of Governments (SANDAG), SCAG, and the San Joaquin Valley MPOs.

Figure F.1 Statewide Population Forecasts by Source of Forecast

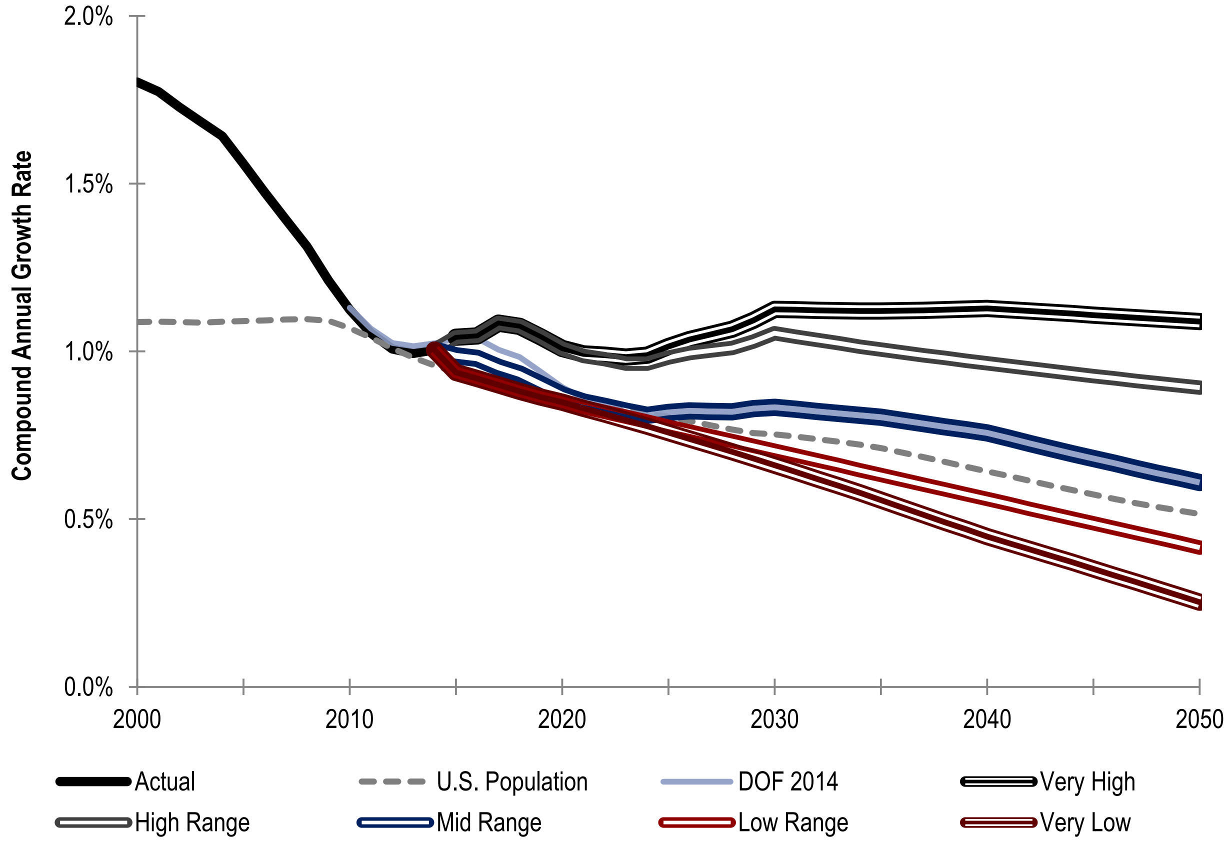


The High Range forecast follows the CSTDM compound annual growth rate (CAGR). The Maximum forecast was adjusted up from the High Range forecast to account for possible, but unlikely events, such as comprehensive immigration reform, increased lifespans, increased fertility rates, and balanced domestic migration. The Most Likely forecast is developed by combining 20-year moving average growth rates from the U.S. Census (National) forecasts and the DOF 2014 (California). . The Low Range forecast CAGR matches the U.S. Census (National) 20-year moving average growth rate in the short-term and decreases to a CAGR slightly below the U.S. Census rate in the long-term. The Minimum forecast was adjusted down from the Low Range forecast to account for possible, but unlikely events, such as substantial tightening of immigration policy and reduced lifespan. Table F.1 describes the population forecast assumptions, CAGR, and Year 2040 forecast for each forecast level. Figure F.2 shows the population growth rates for each of the range in forecasts.

Table F.1 Statewide Population Forecasts

	Source of Forecast	CAGR (2010 to 2040)	2040 Forecast California Population
Maximum	• Statewide population total matched to DOF "actuals" through 2014. • For 2015 to 2020, statewide population growth follows the CSTDM 20-year moving average growth rate. • For 2021 to 2040, statewide population growth follows the CSTDM 20-year moving average growth rate plus additional 50,000 residents per year (2021), increasing to 150,000 residents per year in 2050.	1.16%	52 million
High Range	• Statewide population total matched to DOF "actuals" through 2014. • For 2015 and beyond, statewide population growth follows the CSTDM 20-year moving average growth rate.	1.00%	50 million
Mid Range	• Statewide population total matched to DOF "actuals" through 2014. • Between 2015 and 2023, statewide population growth follows the midpoint between the U.S. Census (National) and DOF 2014 (California) 20-year moving average growth rates. (The two growth rates converge at 0.82 percent in 2023.) • Beyond 2023, statewide population growth follows the DOF 2014 20-year moving average growth rates, decreasing to 0.61 percent by 2050.	0.82%	47 million
Low Range	• Statewide population total matched to DOF "actuals" through 2014. • Between 2015 and 2020, statewide population growth follows the U.S. Census (National) 20-year moving average growth rate. • Beyond 2020, statewide population growth steadily decreases from the U.S. average national growth rate (0.85 percent) to 0.1 percent below the U.S. average national growth rate by 2050. The resulting 2050 growth rate is 0.42 percent (20-year moving average).	0.63%	45 million
Minimum	• Statewide population total matched to DOF "actuals" through 2014. • Between 2015 and 2020, statewide population growth follows the U.S. Census (National) 20-year moving average growth rate. (Same as low range.) • For 2021 to 2040, statewide population growth follows the Low Range 20-year moving average growth rate PLUS 25,000 fewer residents per year (2021) decreasing to 100,000 fewer residents per year in 2050.	0.58%	44 million

Figure F.2 Range of Population Growth Rates (20-Year Moving Average)



Once the range of population forecasts were developed, the population forecasts were converted to household totals using the following assumptions:

- **Maximum and Most Likely.** Statewide household total calculated by applying the CSTDM household size assumptions to the statewide population totals; and
- **Minimum.** Statewide household total calculated by applying the average household size assumptions from three sources (i.e., CSTDM, CEF 2014, and Moody's 2013) to the minimum statewide population totals.

Statewide employment total were calculated by applying the CSTDM jobs per capita assumptions to the statewide population totals (i.e., about 0.43 job per capita in 2012, increasing to 0.45 by 2020, and then tapering to 0.44 by 2035 and holding thereafter).

Table F.2 shows the maximum, most likely, and minimum population, household, and employment forecasts utilized for the risk analysis.

Table F.2 Maximum, Most Likely, and Minimum Population, Household and Employment Projections

	Maximum Projections			Most Likely Projections			Minimum Projections		
Year	Pop	HH	Emp	Pop	HH	Emp	Pop	HH	Emp
2040	52.013	17.840	22.928	47.022	16.128	20.728	44.022	14.977	19.406

G. Quantifying the Effects of Autonomous and Shared Use Vehicles on Year 2040 Risk Variables

By 2040, it is likely that autonomous vehicles (AV) and shared-use vehicles will compose some share of all automobile travel. AVs could have important features that change the auto mode's perception among travelers, while the increase in shared-use vehicles could directly affect the auto operating cost of travelers, which may impact HSR ridership and revenue. The risk analysis framework considers two key features of the auto mode that might change as a result of AVs and shared-use vehicles: 1) auto travel times, and 2) auto operating costs.

G.1 Autonomous Vehicle Background and Research

One of the promises of AV technology is to improve travel speeds by connecting vehicles, allowing them to travel much closer to one another at high speeds, effectively increasing capacity and reducing congestion. Most of the travel time benefits of AVs rely on AVs representing a clear majority of autos, with the most benefits really being achieved once market penetration reaches about 75 percent. It is possible that AVs could contribute to congestion in the near term, depending on the programs that control them and how well they are able to interact with non-AVs.[22]

Auto operating costs can be improved for AVs via better gas mileage, lower insurance premiums if crashes can be reduced, and reduced parking costs, as AVs could potentially drop a passenger off and find free or cheaper parking. The possibility of increased vehicle miles traveled (VMT) due to taxiing with no passenger (e.g., to park) could effectively increase operating costs, though this would require it being legal for AVs to travel without an operator, which could be further into the future than 2040.

There are other potential characteristics of AVs that could influence long-distance auto trips, but are not included in the risk analysis model. The driving experience may be less onerous, as one can engage in other activities during travel. While this

[22] Litman, Todd. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning. February 27, 2015. Victoria Transport Policy.

might be true, auto passengers already have this freedom, and typically do not consider them differently in travel models. Moreover, high-speed rail (HSR) riders also would have this freedom, but this distinction is not made for HSR travel times in the model either. Toll roads/lanes may be developed specifically for use by AV motorists. The benefit of such facilities to AV motorists would be higher speeds achieved with all vehicles in the lanes being connected. However, this is unlikely to occur by year 2040. It would overcoming many issues, including building or converting AV lanes at the detriment of general purpose lanes at a very significant scale to be useful for long-distance travel. Therefore, we excluded this uncertainty as well. The reliability of auto travel may increase due to a significant reduction in the number of traffic incidents. While we directly account for the reliability of the public modes in the model, auto reliability is embedded within the constants.

G.2 AV Market Penetration Assumptions

AV market penetration is a key risk variable that informs both the uncertainty in auto travel times and the uncertainty in auto operating costs. For instance, if market penetration of AVs is 0 percent, then we expect no change to travel times or operating costs. However, if market penetration is 50 percent, we expect improved travel times and some effect on operating costs. Bierstedt et al. (2014) estimate that, under the right circumstances, AVs could represent 50 to 75 percent of the auto market by 2035 to 2045.[23] Litman (2015) forecasts that 30 percent market penetration will occur in the 2040s (but 40 percent of all travel), 50 percent market penetration will occur in the 2050s, and 75 percent will not occur until 2060.[24] Milakis et al. (2015) estimates that market penetration will be between 1 percent and 11 percent by 2030 and 7 percent and 61 percent in 2050, depending on a number of factors.[25] Based on this research, it is assumed that the market penetration of autonomous vehicles among the owned vehicle market is a triangular distribution with minimum 10 percent, maximum 75 percent, and most likely 35 percent.

[23] Bierstedt, J., A. Gooze, C. Gray, J. Peterman, L. Raykin, and J. Walters, 2014. Effects of Next Generation Vehicles on Travel Demand and Highway Capacity by FP Think Working Group Members. FP Think Working Group.

[24] Littman, T., 2015. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning. February 27, 2015. Victoria Transport Policy.

[25] Milakis, D., M. Snelder, B. van Arem, B. van Wee, and G. Correia. 2015. Development of automated vehicles in the Netherlands: scenarios for 2030 and 2050. Delft, The Netherlands: Delft University of Technology.

G.3 Shared-Use Vehicle Market Penetration Assumptions

The shared-use market penetration was calculated using a series of assumptions.[26] It was asserted that the long-distance trip shared-use market would vary by area type of the household, with households in denser areas being more likely to use shared-use vehicles. For each area type, a low, most likely, and high value of shared-use vehicle usage was asserted based on professional judgment. From those assertions, a weighted low, most likely, and high value was computed based on long-distance trip shares, as shown in Table G.1. The market penetration share is assumed to have a triangular distribution.

Table G.1 Shared-Use Market Penetration by Area Type

Area Type	Long-Distance Trips	Long-Distance Trip Share	Long-Distance Auto Trips Using Shared-Use Vehicles in Year 2040		
			Low	Most Likely	High
CBD-Bay Area	74,684	3%	0.2	0.3	0.5
Urban-Bay Area	10,0619	5%	0.1	0.2	0.45
CBD-Other	*71,068*	*3%*	*0.1*	*0.2*	*0.45*
Urban-Other	226,623	11%	0.05	0.1	0.35
Small Urban	157,145	7%	0	0.1	0.2
Suburban	1,055,053	49%	0	0	0.1
Rural	461,141	21%	0	0	0.1
Weighted Total			**2%**	**5%**	**20%**

G.4 Development of Auto Operating Cost Uncertainty

As discussed in Appendix D, the auto operating cost for privately owned non-AVs is comprised of different components. These components are treated together as one auto operating cost, which is referred to as OC_{Base}. Additional uncertainty was added to pertinent subcomponents, representing the uncertainty in auto operating costs due to AV adoption and shared-use vehicles. Key variables we considered were the level of AV market penetration and shared-use market share.

[26] The market penetration rates used in this analysis are in addition to the year 2010 (i.e., model calibration year) market penetration for shared-use vehicles used for long-distance trips, such as rental cars.

Because the marginal cost of trips made by shared-use vehicle will include additional costs over and above typical operating costs, those additional costs were considered. For instance, shared-use trips will be charged a surcharge, similar to a toll either based on the amount of time the vehicle is used or distance traveled. To keep the surcharge in the same units as auto operating cost, it is assumed the surcharge is based on distance traveled and would incur a charge per mile traveled. Ranges for cost per mile are predicted by Litman to be between $0.60 and $1.00 per mile, though this seems high given that current shared-use costs are on the order of $0.15 to $0.60 per mile.[27] Given that AVs may dominate this market and might have higher purchase prices, it is conceivable that the costs will be higher by 2040, but probably not as high as forecast by Litman. It is, therefore, assumed that shared-use cost per mile is a uniform distribution with minimum $0.18 and maximum $0.85 (2014 dollars).

AVs are predicted to drive in a more energy efficient manner compared to non-AV drivers due to a decrease in stop-and-go tendencies. Fuel economy could increase by as much as 23 to 39 percent.[28] It is assumed that fuel economy improvements of AVs are uniform distribution with minimum 10 percent and maximum 50 percent. For the use in the risk analysis, the 50-percent improvement is a more conservative assumption than the Eno prediction. As discussed in Appendix D, fuel costs represent approximately 60 percent of the base auto operating costs. It is assumed that only fuel costs would be affected by fuel economy improvements resulting from AV use.

The overall average auto operating cost is computed as a blended average for each market as follows:

$$OC_{avg} = OC_{nonAVnonSV} + OC_{AVnonSV} + OC_{nonAVSV} + OC_{AVSV}$$

Here, $OC_{nonAVnonSV}$ is the portion of operating costs attributable to owned non-AVs and is computed as:

$$OC_{nonAVnonSV} = (1 - S_{SV}) * (1 - S_{AV}) * OC_{Base}$$

S_{SV} is the market share of long-distance trips that use a shared vehicle (i.e., $(1 - S_{SV})$ is the market share of long-distance trips that use a nonshared vehicle), S_{AV} is the market penetration of AVs among nonshared use vehicles, and OC_{Base} is the base value of operating cost that comes from the distribution described in Appendix D for other model years.

OC_{AV} is the portion of operating costs attributable to owned AVs and is computed as:

$$OC_{AVnonSV} = (1 - S_{SV}) * S_{AV} * OC_{Base} * \left(\left[1 - S_{Base,FC}\right] + \left[\frac{S_{Base,FC}}{1 + FE_{AV}}\right]\right)$$

[27] Littman, T., 2015. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning. February 27, 2015. Victoria Transport Policy.

[28] Eno Center for Transportation, 2013. Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations. Eno Center for Transportation, October 2013.

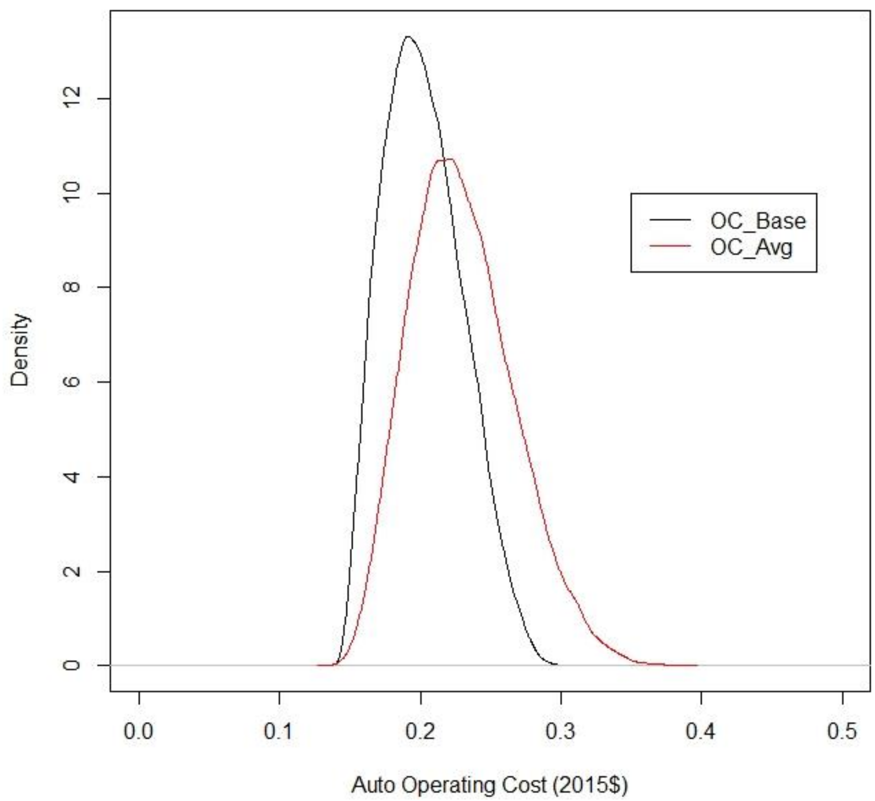
$S_{Base,FC}$ represents the share of base auto operating cost attributable to fuel. FE_{AV} is the fuel economy improvements achieved by AVs, on average. $OC_{sv}OC_{nonAVSV} + OC_{AVSV}$ is the contribution of shared-use vehicles to average auto costs and is computed as follows:

$$OC_{nonAVSV} + OC_{AVSV} = S_{SV} * (1 - S_{AV}) * CPM_{sv} + S_{SV} * S_{AV} * CPM_{sv}$$

CPM_{sv} is the cost per mile surcharge of shared vehicles.

Figure G.1 shows the distribution of auto operating costs in 2040. The black line corresponds to the distribution of the base auto operating costs, and the red line is the overall average auto operating cost distribution, based on the first, blended equation outlined above. The minimum auto operating cost is 13 cents per mile, the most likely is 21 cents per mile, and the maximum is 37 cents per mile (2015 dollars).

Figure G.1 Distribution of Auto Operating Costs in Year 2040



G.5 Development of Auto Travel Time Uncertainty

The current congested travel times forecast for 2040 are considered to be the maximum auto travel times that are likely to occur in year 2040. While it is

possible that AVs increase congestion in the short-term, there is minimal risk in that direction.

On the other end of the spectrum, free-flow travel time is considered as the absolute minimum travel times that could occur. In theory, it would actually be possible to achieve better speeds than free-flow speeds at very high levels of AV market penetration. However, that is unlikely by 2040 given limitations due to market penetration and highway design. Thus, free-flow speeds represent a reasonable set of minimum values.

The AV effect on auto travel times was modeled using a weighted average of congested and free-flow travel times, using a travel time index varying between 0 and 1 based on the following:

- At zero, congested travel times are observed;
- At one, free-flow travel times are observed; and
- At 0.5, the midpoint travel times between congested and free-flow are observed.

There are two key sources of uncertainty that affect the travel time index described: 1) market penetration; and 2) the impact of AV travel times at each market penetration level. The greater the market penetration, the closer travel times will be to free-flow travel times.[29] As discussed above, 75 percent market penetration is the absolute maximum that may be achieved by 2040.

While market penetration is important, there is no consensus of how much impact, for example, 50 percent AV market penetration will have on auto travel times. It is known that 0 percent market would result in no change to travel times, and 100 percent market would result in travel speeds that meet or possibly exceed free-flow speeds.[30] But there is uncertainty as to the effect of AVs for every other value in the middle.

Each source of uncertainty was included into the Monte Carlo analysis, using the following function:

$$index = \frac{\mathrm{B}(S_{AV};\alpha,\beta)}{\mathrm{B}(\alpha,\beta)}$$

---

[29] Effects of Next Generation Vehicles on Travel Demand and Highway Capacity by FP Think Working Group Members: Jane Bierstedt, Aaron Gooze, Chris Gray, Josh Peterman, Leon Raykin, and Jerry Walters. January 2014.

[30] Nevertheless, AV speeds will be constrained by design speeds of highways. While these are higher than posted speed limits, they will limit the maximum free-flow speeds AVs can achieve.

Here, B() is the beta function, S_{AV} is the market penetration of AVs, and α and β are random parameters.[31] This is the cumulative distribution function (cdf) for the beta distribution, but it is not being used strictly as a distribution here. As shown in the equation, the market penetration is one input. The other inputs are α and β, which are shape parameters to the beta function, controlling the shape of the travel time response to AV market penetration. They are treated as random parameters in the MC analysis. In other words, the actual impact that AVs have on travel times is taken to be uncertain. The distributions of these parameters are described in more detail below.

The effect of market penetration on travel time may be different depending on whether a vehicle is traveling on a freeway compared to an arterial. Speed improvements will likely be realized at much lower market penetrations on freeways compared with arterials, where more advanced technology might be required to realize improved travel speeds. Thus, the relationship between market penetration and travel time is segmented across freeways and arterials.

As described above, α and β are treated as random parameters to the MC analysis. They are responsible for the overall shape of the travel time response to AV market penetration. The distributions for each parameter were set using professional judgment by examining the resulting shapes of the travel time response to AV market penetration under different assumptions about the parameters. The mean values of α and β are set at 3 and 1.25 on freeways, and both are uniformly distributed with minimums of 2 and 0.75, respectively, and maximums of 4 and 1.75, respectively. For arterials, a fixed value of 0.5 is used to factor the values obtained for freeways. This is needed since the experimental design can handle only a single auto travel time index variable. By using a fixed factor, freeway and arterial indices are perfectly correlated, effectively making them a single variable.

Figure G.2 plots travel time index variables against market penetration for freeways and arterials. As evidenced by the figure, given a value of market penetration, there remains uncertainty in the indices, which results from randomness in α and β. The red lines in the figures show the mean index value at different levels of market penetration (using the mean values of α and β). These minimum, most likely, and maximum travel time index are shown in Table G.2.

[31] Note that the same AV market penetration variable was used here as in the auto operating costs. For auto operating costs, the market penetration referred specifically to AVs that are privately owned (non-shared-use). Here, market penetration is being used to refer to the overall population of long-distance trips. While, strictly speaking, these values would not typically be the same, here they are assumed to be identical, which is the same as assuming that AVs are represented equally in the shared-use market and privately owned market.

Figure G.2 Travel Time Index Variables versus Market Penetration for Freeways and Arterials

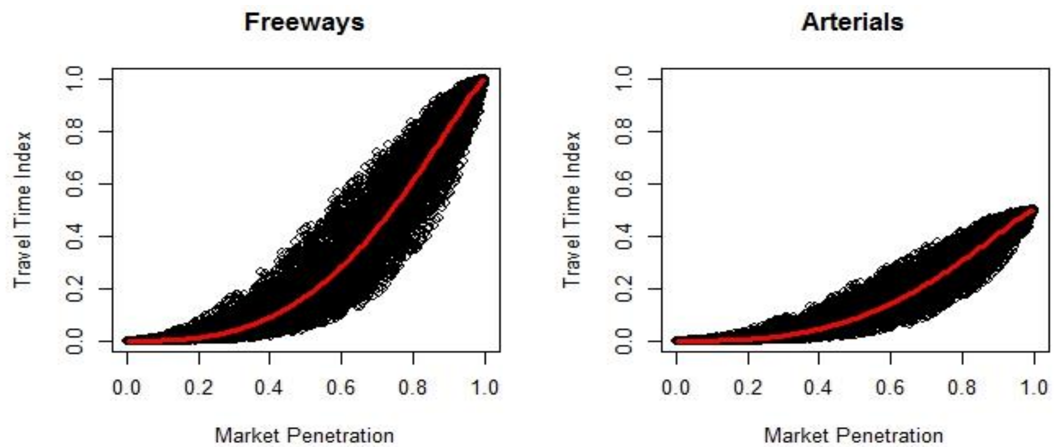


Table G.2 Auto Travel Time Index Range for Freeways and Arterials

Risk Variable	Minimum	Most Likely	Maximum
AV Market Penetration	0.100	0.350	0.750
TT Index – Freeway	0.000	0.060	0.796
TT Index – Arterial	0.000	0.030	0.398

H. Experimental Design

H.1 Fractional Factorial Design

Factorial designs are a classical design approach, originally developed for use with physical experiments. They have a number of desirable properties. First, they do not require a great number of runs per factor explored, at least if the number of factors is not too large. Second, they are powerful in their ability to distinguish which factors are most important and which are of lesser importance. And third, they can be setup to ensure that both main effects and interaction effects can be estimated.

Full factorial designs also typically require a significant volume of runs when the number of factors is large. In our case, we examined 10 factors, making a full factorial design infeasible. Fractional factorial designs greatly reduce the number of runs required, but confound some interaction effects with other interaction effects and/or main effects (meaning we cannot identify the distinct effects). Different resolutions can be chosen for fractional factorial designs, which set the level of confounding, each requiring different numbers of runs. In addition, factorial designs consider only two or three values, or levels, for each specific variable (the set of runs pairs different levels for each factor). In a two-level design it is only possible to estimate linear effects for continuous variables. For the 2016 BP analysis, a three-level design was used, where each variable takes on one of three values in each model run. This allows two different slopes to be estimated in the meta-model. The three-levels used to define the fractional factorial design correspond to the Minimum, Most Likely, and Maximum values for each variable discussed in the previous sections.

Table H.1 shows the number of runs required for a three-level fractional factorial design with 10 factors, at three different resolutions. About 10 risk variables were chosen for each model year, because including any additional variables in a Resolution IV design would require substantially more model runs (i.e., 243 runs).

Table H.1 Number of Runs for 10 Factors with Three Levels

Design Resolution	Runs[a]
Resolution III allows the estimation of main effects, but these may be confounded by two-factor interactions.	27
Resolution IV allows the estimation of main effects that are not confounded by two-factor interactions. It allows the estimation of the two-factor interaction effects, but these may be confounded by other two-factor interactions.	81
Resolution V allows the estimation of main effects that are not confounded by two-factor and three-factor interactions. It allows the estimation of two-factor interaction effects, which are not confounded by any other two-factor interactions.	243

[a] The required runs for 10 risk factors at different resolutions was obtained using SAS software.

H.2 Three-Level Random Sampling Design

One feature of factorial designs is that they are boundary designs. That is, all runs in the experimental design are at the boundary of the variable space. This is true even for most three-level designs, since every run will have at least one variable at a low or high boundary value. While this does tend to bound the dependent variable (i.e., high-speed rail revenue), it potentially is inefficient, particularly since the MC analysis will focus on draws of the independent factors from inside those bounds.

For computer experiments, so-called Sampling or "space-filling" designs also may be useful (Sacks et al., 1989). The sampling design developed for this analysis is derived from the fractional factorial design described above, but instead of selecting fixed low, middle, and high values for each run, the sampling design uses these levels to partition the variable space. This three-level random sampling design is necessary to better estimate nonlinearities and to ensure the entire solution space is represented. The three-level random design uses the same fractional factorial design, but with two important changes:

1. The input variables were reordered within the design. For instance, instead of setting the business purpose HSR constant as the experimental design variable, it was set to the fifth variable. This creates an entirely new set of experimental design runs with combinations of variables that did not appear in the original fractional factorial runs.

2. Different values for the variable were used instead of repeating the same value multiple times so that we can add useful information to the estimation of our regression model. Each time factor x uses level -1 (Low), a particular value for factor x was selected from the bottom third of its distribution. Likewise, for Level 0 (Middle), the middle part of its distribution was selected and for Level 1 (High) the upper third of its distribution was selected. This selection used a uniform distribution within each range level to help fill in more of the solution space.

To ensure the interior of the solution space was well-represented, and not biased toward the edges, it is essential to perform the same number of three-level random sampling runs as fractional factorial runs.

H.3 TWO-STEP RISK ANALYSIS PROCESS

The first step of the experimental design included running 81 runs using a fractional factorial three-level Resolution IV Design to estimate the existence of two-factor interaction effects. For all alternatives and forecast years, the model that included significant interaction terms demonstrated illogical results compared to the main effects only model (i.e., the sign on the effects of certain variables changed and became counterintuitive for certain combinations of variable values). In all cases, it was determined that a regression model without interaction terms produced more reasonable results than a model that also included interaction terms. The results of the interaction models for the various alternatives/forecast years are discussed in Appendix I.

As a result of the findings based on Step 1, the final experimental design for Step 2 was modified to include 59 full model runs for each alternative and forecast year, as follows:[32]

- 27 model runs used fixed minimum, most likely, and maximum values of risk variables specified using a three-level Resolution III fractional factorial design. Only three of these model runs overlapped with the 81 model runs using the three-level Resolution IV Design
- 27 model runs sampled from low, mid, and high ranges of the risk variables, using the three-level random sampling design.
- 5 model runs representing extreme scenarios of full upside (3 runs) and full downside (2 runs); that is, all inputs in these runs were set to values that would either be favorable or unfavorable to HSR revenue. The runs correspond to the following percentiles for each risk variable: 10, 25, 75, 90, and 100. The 0th percentile run was not added because the experimental design included this run already, where all inputs are set to the "min" value, and the Minimum value always corresponded to the absolute min, unfavorable value for HSR revenue.

The final experiment design includes both the Fractional Factorial design to help understand extreme values, tails of distributions, and the Sampling design, which helps fill in the space in the middle of the distribution where most results fall.

[32] The original 81 runs developed in the first step were not used for the development of the final regression equations in order to reduce the number of random sampling designs needed. To ensure the interior of the solution space was well-represented, and not biased toward the edges, it is essential to perform the same number of three-level random sampling runs as fractional factorial runs.

I. Regression Model with Interaction Effects

A regression model to test for the presence of two-factor interaction effects was estimated for each of the model years using the 81 runs using a fractional factorial three-level Resolution IV Design. This appendix lays out the consideration of regression meta-models with interaction effects, what their impact would have been, and the reasons why the non-interaction effect meta-models were more appropriate for use in the Monte Carlo simulation. The interaction effects regression model took the following functional form:

$$ln(Revenue) = Constant + \beta_1 \times Var_1 + \beta_2 \times Var_2 \ldots + \beta_{10} \times Var_{10} + \delta_1 \times Var_x \times Var_y$$

For each forecast year, the interaction model was compared against the main effects to understand how the variable coefficients changed with the introduction of interaction effects. The interaction model was developed using both a bottom-up and top-down approach. For the bottom-up approach each interaction term was added individually to the model. The original or new model was selected that had the highest R^2 value. These steps were repeated with the remaining interaction terms until no remaining interaction terms were statistically significant. The top-down approach, which introduced all interaction terms at one time and then systematically eliminated nonsignificant terms, produced the same results.

I.1 Interaction Effects Model Investigated for 2025 Silicon Valley to Central Valley Line

Table I.1 shows the year 2025 VtoV model with interaction effects. Compared to the main effects model the coefficients for variables not included in interaction terms remained stable while the coefficient for variables included in interaction terms changed substantially.

Table I.1 Year 2025 VtoV Interaction Effects Model

	Main Effects Model[a]		Interaction Model	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Constant	19.955	90.7	19.683	81.5
HSR Mode Choice Constant – Business	0.233	16.2	0.233	18.5
HSR Mode Choice Constant – Commute	0.071	2.8	0.071	3.2
HSR Mode Choice Constant – Recreation/Other	0.441	17.3	0.441	19.7
Trip Frequency Constant – Business/Commute	0.494	6.4	0.494	7.2
Trip Frequency Constant – Recreation/Other	0.556	2.8	0.556	3.1
Auto Operating Cost	1.094	2.0	2.314	3.9
HSR Fare	0.086	0.5	-0.123	-0.7
HSR Headway	-0.294	-5.1	-0.006	-0.1
HSR Access-Egress Connectivity = Low	-0.191	-2.6	-0.578	-1.5
HSR Access-Egress Connectivity = High	0.010	0.1	0.792	3.5
HSR Access-Egress by Transit Variable	1.151	1.6	-1.664	-1.2
Interaction Terms				
Auto Operating Cost & High Connectivity			-3.660	-3.5
HSR Fare & Low Connectivity			0.629	1.8
HSR Headway & Low Connectivity			-0.280	-2.3
HSR Headway & High Connectivity			-0.141	-1.1
HSR Headway & Acc/Egr Transit Variable			2.943	2.2
Model Statistics				
Sum of Squared Error	5.953		4.595	
R^2	0.892		0.917	

[a] The main effects and interaction models shown are based on 81 BPM-V3 runs using a factional factorial design. No range sampling has been performed. The main effects model used for the risk analysis is based on 59 BPM-V3 runs = 27 with fractional factorial design + 27 with range sampling + 5 extreme value runs

The interaction model produces results that are counterintuitive. While the main effects model results are less precise (in terms of fitting full model run results), model sensitivities are all appropriate. As one example, the Figure I.1 shows the

relative effect of the HSR access/egress transit variable on log revenue in the main effects, linear model. Figure I.2 shows the relative impact of HSR access/egress transit variable on log revenue in the interaction model, for different values of the HSR headway. The effect of the variable approaches zero as the slope of headway falls below 1.0. This (and other similar results) does not match the behavior of the full model, and thus, the interaction model was rejected as the best model to use for risk analysis.

Figure I.1 Relative Effect of the HSR Access/Egress Transit Variable on Log Revenue in Main Effects Model

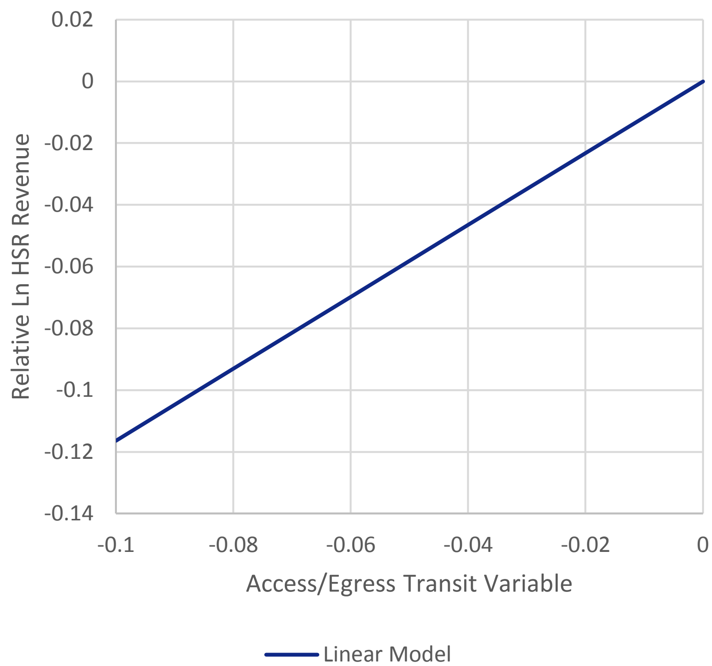
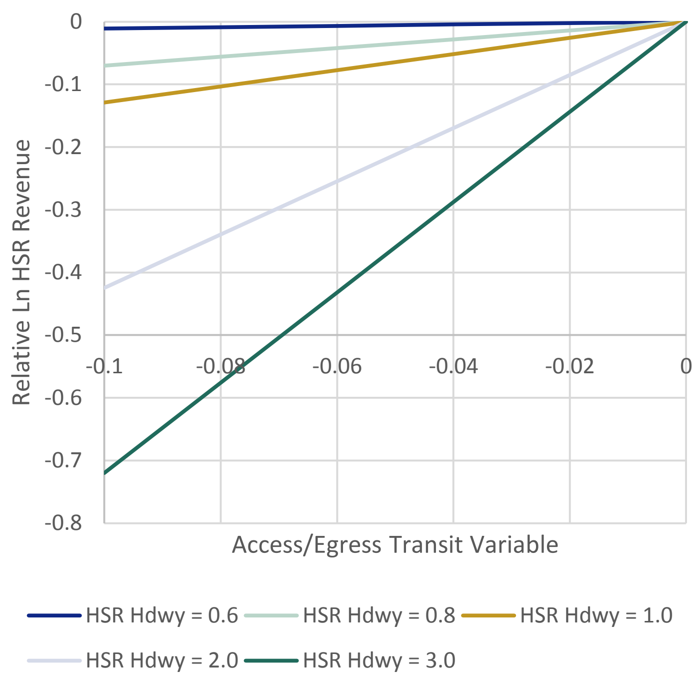


Figure I.2 Relative impact of HSR Access/Egress Transit Variable on Log Revenue in the Interaction Model for Different Values of the HSR Headway



I.2 Interaction Effects Model Investigated for 2029 Phase 1

Table I.2 shows the year 2029 Phase 1 model with interaction effects. Compared to the main effects model the coefficients for variables not included in interaction terms remained stable while the coefficient for variables included in interaction terms changed substantially and became difficult to interpret. Coefficients for auto operating cost, HSR fare, airfare, and HSR access/egress transit variable become inflated and are offset by interaction variables.

Table I.2 Year 2029 Phase 1 Main Effects and Interaction Effects Model

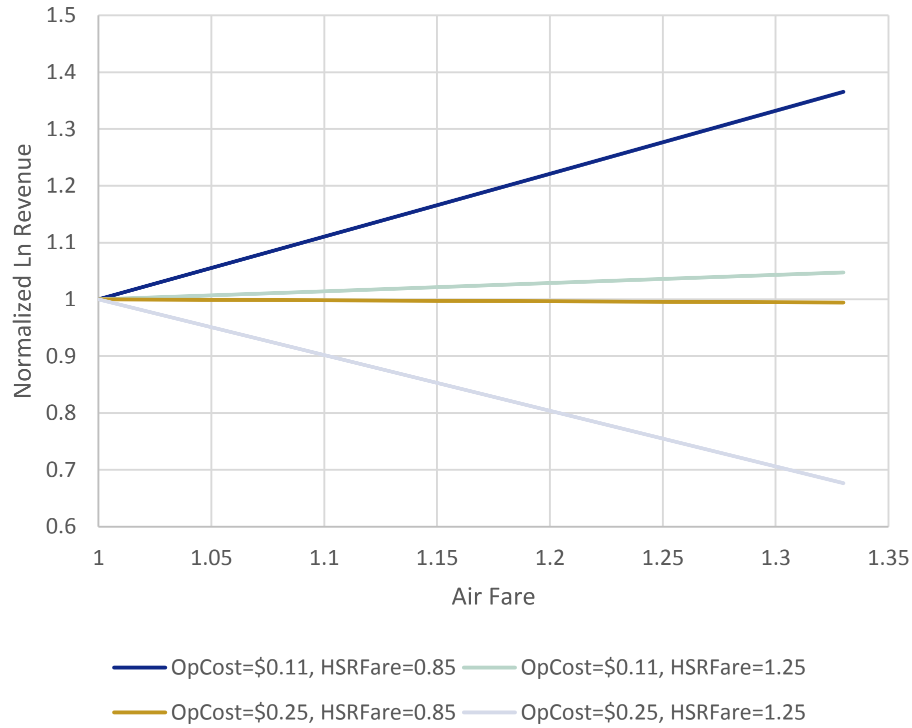
	Main Effects Model[a]		Interaction Model[a]	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Constant	20.902	84.4	16.122	12.6
HSR Mode Choice Constant – Business	0.167	15.9	0.167	18.2
HSR Mode Choice Constant – Commute	0.071	3.8	0.071	4.4
HSR Mode Choice Constant – Recreation/Other	0.402	21.6	0.402	24.8
Trip Frequency Constant – Business/Commute	0.467	8.4	0.467	9.7

	Main Effects Model[a]		Interaction Model[a]	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Trip Frequency Constant – Recreation/Other	0.592	4.3	0.592	4.9
Auto Operating Cost	1.339	3.5	10.788	3.7
HSR Fare	0.237	1.9	3.070	2.8
HSR Headway	-0.167	-5.2	-0.016	-0.3
Airfare	0.140	0.9	4.036	3.8
HSR A-E Transit Variable	1.199	2.2	-2.702	-2.4
Interaction Terms				
Auto Operating Cost & Airfare			-8.030	-3.3
HSR Fare & Airfare			-2.408	-2.6
HSR Headway & HSR A-E Transit Variable			3.001	3.8
Model Statistics				
Sum of Squared Error	3.189		2.415	
R^2	0.916		0.936	

[a] The main effects and interaction models shown are based on 81 BPM-V3 runs using a factional factorial design. No range sampling has been performed. The main effects model used for the risk analysis is based on 59 BPM-V3 runs = 27 with fractional factorial design + 27 with range sampling + 5 extreme value runs.

The interaction model produces results that are counterintuitive. While the main effects model results are less precise in terms of fitting full model run results, model sensitivities are all appropriate. Figure I.3 shows the relative impact of airfare on log revenue in the interaction model, for different values of the HSR fare and operating cost. The slope of airfare changes signs depending on the values for operating cost and HSR fare. These results are illogical. It does not make sense for an increase in airfare to result in a decrease in HSR revenue, as suggested when auto operating costs are $0.25 per mile and HSR fares are a factor of 1.25 from the base fare. The interaction term between HSR headway and HSR access/egress transit variable also results in illogical implications for the effect of HSR access/ egress transit variable on HSR revenue for certain values of HSR headway. Thus, the interaction model was rejected as the best model to use for risk analysis.

Figure I.3 Relative Impact of Airfare on Log Revenue in the Interaction Model for Different Values of the HSR Fare and Operating Cost



I think it is worth closing with one sentence that interprets the chart above. Or just moving some of the language in the paragraph above to below this figure.

I.3 INTERACTION EFFECTS MODEL INVESTIGATED FOR 2040 PHASE 1

Table I.3 shows the year 2040 Phase 1 model with interaction effects. Compared to the main effects model the coefficients for variables not included in interaction terms remained stable while the coefficient for variables included in interaction terms changed substantially and became difficult to interpret. Coefficients on auto operating cost and population/employment growth changed sign while the coefficient on the auto travel time index increased substantially.

Table I.3 Year 2040 Phase 1 Main Effects and Interaction Effects Model

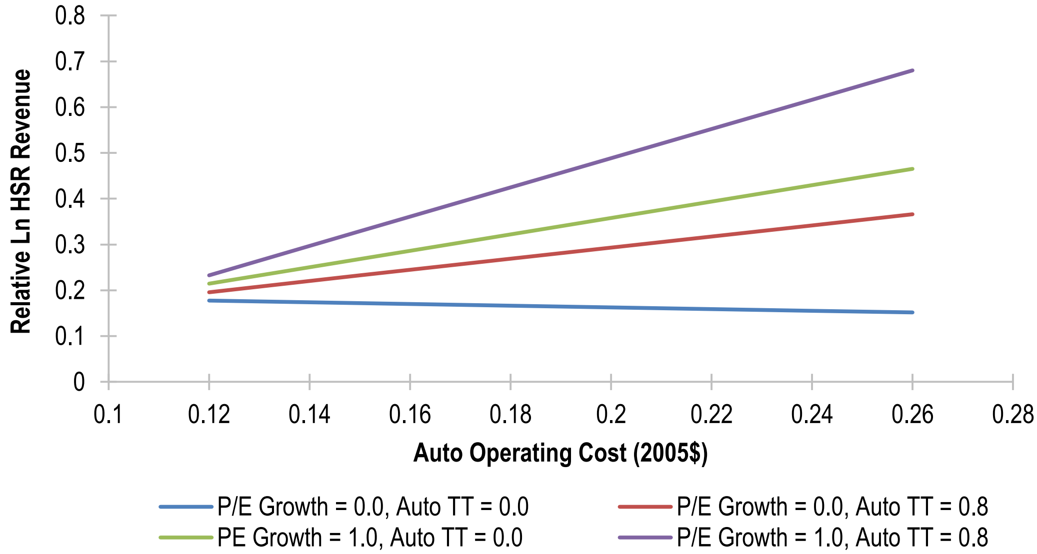
	Linear Model[a]		Interaction Model[a]	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Constant	21.086	191.9	21.274	156.8
HSR Mode Choice Constant – Business	0.179	14.4	0.179	15.8
HSR Mode Choice Constant – Commute	0.079	3.6	0.079	4.0

HSR Mode Choice Constant – Recreation/ Other	0.397	18.0	0.397	19.8
Trip Frequency Constant – Business/ Commute	0.481	7.6	0.481	8.3
Trip Frequency Constant – Recreation/Other	0.642	4.0	0.642	4.4
Auto Operating Cost	1.240	3.8	-0.186	-0.4
HSR Fare	0.137	2.7	0.212	3.6
HSR Headway	-0.173	-4.5	-0.173	-5.0
Population/Employment Growth	0.314	4.9	-0.068	-0.4
Auto Travel Time Index	-0.006	-0.1	-0.039	-0.2
Interaction Terms				
Auto Operating Cost & Population/ Employment Growth			1.976	2.7
Auto Operating Cost & Auto Travel Time			1.753	2.1
HSR Fare & Auto Travel Time			-0.260	-2.1
Model Statistics				
Sum of Squared Error	4.449		3.711	
R^2	0.896		0.913	

[a] The linear and interaction models shown are based on 81 BPM-V3 runs using a factional factorial design. No range sampling has been performed. The linear model used for the risk analysis is based on 59 BPM-V3 runs = 27 with fractional factorial design + 27 with range sampling + 5 extreme value runs.

The interaction model produces results that are counterintuitive. Figure I.4 shows the relative impact of auto operating cost on log revenue in the interaction model, for different values of the population/employment growth and auto travel time index. The slope of auto operating cost changes signs depending on the values for population/employment growth and auto travel time. These results are illogical. It does not make sense for an increase in auto operating cost to result in a decrease in HSR revenue, as suggested when the growth index is 0.0 and the auto travel time index is 0.0. In addition, the interaction terms between operating costs, HSR fares, and auto travel time also result in illogical implications for the effect of auto travel time on HSR revenue for certain values of operating costs and HSR fares. Thus, the interaction model was rejected as the best model to use for risk analysis.

Figure I.4 Relative Impact of Auto Operating Cost on Log Revenue in the Interaction Model, for Different Values of the Population/ Employment Growth and Auto Travel Time Index



J. Risk Variable Component Specification for Monte Carlo Simulation

Table J.1 details the components of each risk variable, the range of values and distributions for each component, and correlation between distributions of risk variables. Some risk factors include multiple components that are sampled in the Monte Carlo analysis. For example, values are sampled from both the error component distribution and the terminal/wait time component distribution for the HSR Mode Choice Constant risk variable. The sampled values are combined, as appropriate, prior to inputting the value into the regression model used for the Monte Carlo simulation. Setting a positive correlation between two risk variable components results in the Monte Carlo simulation having a higher probability of sampling from the same point on the distribution (e.g., a 100-percent positive correlation would result in two risk variables always being chosen from the same percentile point on the distribution).

Table J.1 Risk Variable Distributions Used in Monte Carlo Analyses

Risk Factor	Components	Years	Minimum	Most Likely	Maximum	Distribution	Notes
HSR Mode Choice Constant – Business	Error Component	All	-2.335	0.0	2.335	PERT – Standard (Shape = 4)	Unit = offset from calibrated coefficient. 50% Correlation with Commute & Recreation/Other HSR Error Components
	Terminal/Wait Time	All	-0.3264	0.0	0.1632	Triangular	Unit = offset from calibrated coefficient. 100% Correlation with Commute & Recreation/Other Term./Wait Times
HSR Mode Choice Constant – Commute	Error Component	All	-1.222	0.0	1.222	PERT – Standard	Unit = offset from calibrated coefficient. 50% Correlation with Business & Recreation/Other HSR Error Components
	Terminal/Wait Time	All	-0.3264	0.0	0.1632	Triangular	Unit = offset from calibrated coefficient. 100% Correlation with Business & Recreation/Other Term./Wait Times
HSR Mode Choice Constant – Recreation/Other	Error Component	All	-1.354	0.0	1.354	PERT – Standard	Unit = offset from calibrated coefficient. 50% Correlation with Business & Commute HSR Error Components
	Terminal/Wait Time	All	-0.1388	0.0	0.0694	Triangular	Unit = offset from calibrated coefficient. 100% Correlation with Business & Commute Term./Wait Times
Trip Frequency Constant – Business/ Commute	Error Component	All	-0.278	0.0	0.278	PERT – Standard	Unit = offset from calibrated coefficient. 50% Correlation with Recreation/Other Error Components
	Economic Component	2025 2029 2040	-0.233 -0.201 -0.246	0.0 0.0 0.0	0.165 0.224 0.209	Triangular	Unit = offset from calibrated coefficient. 100% Correlation with Recreation/Other Economic Component

Risk Factor	Components	Years	Minimum	Most Likely	Maximum	Distribution	Notes
Trip Frequency Constant – Recreation/Other	Error Component	All	-0.123	0.0	0.123	PERT – Standard	Unit = offset from calibrated coefficient. 50% Correlation with Business/Commute Error Components
	Economic Component[1]	2025	-0.070	0.0	0.052	Triangular	100% Correlation with Business/Commute Economic Component
		2029	-0.068	0.0	0.078		
		2040	-0.087	0.0	0.071		
Auto Operating Costs	Combined Components	2025	0.15	0.20	0.31	PERT – Shape=5 2040 n/a	Unit = 2015$/mile Full Model & Regression Model use 2005$, rather than 2015$. Conversion at following rate 202.6 / 250.404 based on CPI. 2040 values used in Full Model Runs but not in Monte Carlo
		2029	0.14	0.19	0.30		
		2040	0.13	0.21	0.37		
Auto Operating Costs Impacts of Autonomous and Shared-Use Vehicles	Owned Nonautonomous vehicle auto operating cost	2040	0.13	0.19	0.32	PERT – Shape=5	Unit = 2014$/mile Used in Monte Carlo but not used in Full Model Runs
	Owned Autonomous Vehicle Market Penetration	2040	0.10	0.35	0.75	Triangular	Unit = Decimal percent of owned AVs used for long-distance trips Used in Monte Carlo but not used in Full Model Runs
	AV Fuel Economy improvements	2040	0.10		0.50	Uniform	Unit = Decimal percent fuel economy improvements from base Used in Monte Carlo but not used in Full Model Runs
	Shared-use vehicle market share	2040	0.02	0.05	0.20	Triangular	Unit = Decimal percent of shared-used vehicles used for long-distance trips Used in Monte Carlo but not used in Full Model Runs
	Shared-use vehicle auto operating cost	2040	0.18		0.85	Uniform	Unit = 2014$/mile Used in Monte Carlo but not used in Full Model Runs

Risk Factor	Components	Years	Minimum	Most Likely	Maximum	Distribution	Notes
HSR Fares	n/a	2025	0.846	1.0	1.275	Triangular	Unit = Factor from Base/Most Likely Value
		2029	0.846	1.0	1.275		For 2040, more uncertainty regarding
		2040	0.647	1.0	1.881		
HSR Headway	n/a	2025	0.29	1.0	1.58	PERT – Standard	Unit = Factor from Base/Most Likely Value
		2029 & 2040	0.65	1.0	2.25		
HSR Connecting Service	n/a	2022 & 2024	Scenario 1 – 10%	Scenario 2 – 50%	Scenario 3 – 40%	Multinomial	Unit = 1 if Scenario is chosen, 0 otherwise
HSR Access/Egress Via Transit Variable	Business/Commute Coefficient	2025, and 2029	-2.0	-1.215	-1.215	PERT – Standard	Unit = coefficient Used in Full Model Run but not used in regression 100% Correlation with Recreation/Other coefficient & Threshold parameter
	Recreation/Other Coefficient		-1.3	-0.88	-0.88	PERT – Standard	Unit = coefficient Used in Full Model Run but not used in regression 100% Correlation with Business/Commute coefficient & Threshold parameter
	Threshold Parameter		0.1	0.2	0.2	PERT – Standard	Unit = threshold value Used in Full Model Run but not used in regression 100% Correlation with Business/Commute coefficient & Recreation/Other coefficient
	Index Variable		-0.1	0.0	0.0	PERT – Standard	Unit = index variable Not used in Full Model Runs but used in regression; Middle value set to 0.05 for Full Model Runs.
Airfares		2029	1.0	1.2	1.33	Triangular	Unit = Factor from Base
Number and Distribution of Households throughout the State	n/a	2040	0.0	0.402	1.0	Triangular	Unit = index variable

Risk Factor	Components	Years	Minimum	Most Likely	Maximum	Distribution	Notes
Auto Travel Time	Index Variable	2040	0.0	0.06	0.8	n/a	Unit = index variable Arterial travel time index =.5*Freeway index, with 100% correlation Used in Full Model Runs and Regression but not in Monte Carlo
	Autonomous Vehicle Market Penetration	2040	0.1	0.35	0.75	Triangular	Unit = Percent of AVs on roads Used in Monte Carlo but not used in Full Model Runs 100% correlated with Owned autonomous vehicle market penetration used in auto operating cost
	alpha	2040	2.0	3.0	4.0	Uniform	Unit = n/a Used in Monte Carlo but not used in Full Model Runs
	Beta	2040	0.75	1.25	1.75	Uniform	Unit = n/a Used in Monte Carlo but not used in Full Model Runs