CAMBRIDGE SYSTEMATICS

# California High-Speed Rail 2020 Business Plan

*Ridership and Revenue Forecasting*

technical

supporting document

*prepared for*

**DB Engineering & Consulting USA Inc. for the California High-Speed Rail Authority**
**for the California High-Speed Rail Authority**

*prepared by*

**Cambridge Systematics, Inc.**

January 2020

www.camsys.com

CALIFORNIA High-Speed Rail Authority

# SIGNATURE/APPROVAL SHEET

**TO:** **Joseph Hedges**

**FROM:** **Stefan Reul**

**SUBJECT: 2020 Business Plan Ridership & Revenue Forecasting Report**

**DESCRIPTION OF ENCLOSED DOCUMENT(S): 2020 Business Plan Ridership & Revenue Forecasting Report**

| REVIEWER | REVIEWER'S INTIALS/DATE: | COMMENTS |
|---|---|---|
| **Signer #1 Name (Print):**<br>**Joseph Hedges,**<br>**CHSRA Chief Operating Officer** | SIGNATURE ON FILE | |
| **Signer #2 Name (Print):**<br>**Brian Annis,**<br>**CHSRA Chief Financial Officer** | SIGNATURE ON FILE | |
| **Signer #3 Name (Print):**<br>**Jorge Rios,**<br>**DB E&C USA VP& ETO Project Director** | SIGNATURE ON FILE | |
| **Reviewer #1 Name (Print):**<br>**Frank Vacca**<br>**CHSRA Chief of Rail Operations** | SIGNATURE ON FILE | |
| **Reviewer #2 Name (Print):**<br>**Stefan Reul,**<br>**ETO, Lead Transportation Expert** | SIGNATURE ON FILE | |
| **Author #1 Name (Print):**<br>**Rachel Copperman,**<br>**Cambridge Systematics Project Manager** | SIGNATURE ON FILE | |

*technical supporting document*

# California High-Speed Rail 2020 Business Plan

## *Ridership and Revenue Forecasting*

*prepared for*

**DB Engineering & Consulting USA Inc. for the California High-Speed Rail Authority**

*prepared by*

**Cambridge Systematics, Inc.**
505 E. Huntland Drive, Suite 550
Austin, TX  78752

*date*

**January 2020**

# Table of Contents

# List of Tables

# List of Figures

# Executive Summary

Cambridge Systematics' (CS) approach to preparing forecasts for use in the California High-Speed Rail Authority's ("the Authority") 2020 Business Plan was predicated on the following concepts:

- The ridership and revenue (R&R) model should produce reasonable forecasts with reasonable sensitivities to changing conditions.
- Models are not perfect, and their imperfections need to be understood and reflected in the forecasts used for business planning purposes.
- Future conditions cannot be known with certainty. The forecasts used for business planning purposes need to recognize those uncertainties and present a reasonable range.

The resulting R&R forecasting process involved the following key steps. CS:

- Refined the high-speed rail service plans reflecting the updated strategy for implementation and sequencing of the Phase 1 system; this includes producing forecasts for a line that connects Silicon Valley to the Central Valley (defined as San Francisco to Bakersfield) for a 2029 opening year and forecasts for the Phase 1 system between San Francisco and Anaheim for 2033 (opening year) and 2040 (out year).
- Updated the conventional passenger rail networks to ensure consistency with current and planned routes and service.
- Incorporated revisions to socioeconomic growth assumptions (population, housing, and employment forecasts) using county-level forecasts developed by California Economic Forecast (CEF), Moody Analytics, and California Department of Finance (DOF), as well as developed a range of alternative forecasts for use in the risk analysis.
- Updated the previous risk analysis assumptions to include the factors that are believed to have the greatest influence on high-speed rail ridership and revenue. The ridership and revenue forecasts in this document are expressed in terms of probabilities that were developed using this approach.

## Summary of Ridership and Revenue Forecasts

Ridership and revenue forecast ranges with the probabilities of achieving certain values are shown in Tables ES.1 and ES.2, respectively. A 10-percent confidence level means that there is a 10-percent chance that the ridership/revenue will be lower than this value (or a 90-percent chance that it will be higher). The range in revenue for the Silicon Valley to Central Valley line year 2029 forecast between the $10^{th}$ and $90^{th}$ percentiles is $1,100 million, compared to $2,440 million for the Phase 1 year 2040 forecast.

The model analysis was performed at 2029, 2033 and 2040 model horizons to remain consistent with previous Business Planning efforts. The data presented in the 2020 Business Plan document was extrapolated from these raw model numbers to reflect the actual opening dates of the Silicon Valley to Central Valley and Phase 1 projects and includes ramp-up.

The data presented in this report shows raw model output only, without ramp-up deductions. All dollar figures presented in this document are base year as of June 2019.

## Table ES.1 Range of Annual Ridership by Implementation Step[a] (Millions)

| Confidence Level That Ridership Will Be Less Than Stated Value | Implementation Step—Silicon Valley to Central Valley line 2029 | Implementation Step—Phase 1 2033 | Implementation Step—Phase 1 2040 |
|---|---|---|---|
| Minimum | 3.7 | 8.6 | 8.9 |
| 1% | 6.5 | 14.6 | 15.2 |
| 10% | 9.6 | 21.1 | 22.2 |
| 25% | 12.4 | 27.0 | 28.3 |
| Median | 16.5 | 35.6 | 37.4 |
| 75% | 21.7 | 46.1 | 48.5 |
| 90% | 27.1 | 56.8 | 60.6 |
| 99% | 37.3 | 76.0 | 80.9 |
| Maximum | 53.4 | 109.8 | 118.1 |
| **Base Run** | **16.2** | **35.6** | **38.6** |

Source: Cambridge Systematics, Inc.

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

## Table ES.2 Range of Annual Revenue by Implementation Step[a] (Millions, 2019 Dollars)

| Confidence Level That Ridership Will Be Less Than Stated Value | Implementation Step—Silicon Valley to Central Valley line 2029 | Implementation Step—Phase 1 2033 | Implementation Step—Phase 1 2040 |
|---|---|---|---|
| Minimum | $279 | $668 | $707 |
| 1% | $436 | $1,027 | $1,077 |
| 10% | $626 | $1,477 | $1,547 |
| 25% | $803 | $1,866 | $1,961 |
| Median | $1,066 | $2,422 | $2,558 |
| 75% | $1,396 | $3,098 | $3,273 |
| 90% | $1,732 | $3,755 | $3,987 |
| 99% | $2,339 | $4,917 | $5,299 |
| Maximum | $3,298 | $6,694 | $7,643 |
| **Base Run** | **$982** | **$2,207** | **$2,410** |

Source: Cambridge Systematics, Inc.

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

# 1.0 Introduction

## 1.1 Overview

Since 2007, Cambridge Systematics (CS) has been supporting the California High-Speed Rail Authority ("the Authority") by producing ridership and revenue (R&R) forecasts for different high-speed rail service options using a state-of-the-art travel demand model. The "Version 1" model was originally estimated and calibrated using data from the 2000-2001 California Statewide Household Travel Survey and a 2005 Stated-Preference Survey to support alternatives analyses and project-level environmental work.

In 2010 and 2011, CS made numerous enhancements to the original Version 1 R&R model. The updated model was used to support the California High-Speed Rail Draft 2012 Business Plan.[1] After receipt of public comment, the Authority made changes to the high-speed rail scenarios being considered in the draft version of the 2012 Business Plan, and CS updated the model assumptions and prepared forecasts in support of the Final 2012 Business Plan.[2]

In 2012 and 2013, CS made additional enhancements to the R&R model to accommodate the evolving forecasting needs of the Authority. The enhanced model, known as Version 2, represented a major overhaul of all model components. It responded to the recommendations of the Authority's Ridership Technical Advisory Panel (RTAP), and considered comments from the Authority's Peer Review Group (PRG) and the Government Accountability Office (GAO). In addition to model enhancements, CS used a risk analysis approach to prepare and present ridership and revenue forecasts in support of the Final 2014 Business Plan.[3]

For the 2016 Business Plan, CS made additional changes to the Version 2 model.[4] The updated version:

- Fully incorporated findings of both stated-preference and revealed-preference surveys into the rider choice models.
- Refined the previous Version 2 model to reduce the number of trips that involve a relatively long trip to travel to or from the high-speed rail station, combined with a relatively short trip on the high-speed rail line itself by adding a variable to reflect the unattractiveness of those types of trips.

1 Cambridge Systematics, Inc., "California High-Speed Rail 2012 Business Plan, Ridership, and Revenue Forecasting, Draft Technical Memorandum," prepared for Parsons Brinckerhoff for the California High-Speed Rail Authority, October 19, 2011.

2 Cambridge Systematics, Inc., "California High-Speed Rail 2012 Business Plan, Ridership, and Revenue Forecasting, Final Technical Memorandum," prepared for Parsons Brinckerhoff for the California High-Speed Rail Authority, April 12, 2012.

3 Cambridge Systematics, Inc., "California High-Speed Rail 2014 Business Plan, Ridership, and Revenue Forecasting, Technical Memorandum," prepared for Parsons Brinckerhoff for the California High-Speed Rail Authority, April 18, 2014.

4 Cambridge Systematics, Inc., "California High-Speed Rail 2016 Business Plan, Ridership, and Revenue Forecasting, Technical Supporting Document," prepared for Parsons Brinckerhoff for the California High-Speed Rail Authority, April 8, 2016.

- Updated the conventional passenger rail and urban transit networks to ensure consistency with current routes and service.
- Made several other minor revisions to input variables and recalibrated the model.

The RTAP supported CS' efforts to estimate, calibrate, and validate this new model version, known as the BPM-V3. Documentation of the model and its calibration can be found in the *California High-Speed Rail Ridership and Revenue Model, Business Plan Model Version 3 (BPM-V3) Model Documentation*, which is posted on the Authority's website. The 2016 Business Plan ridership and revenue forecasts also went through a comprehensive review by Project Finance Advisory Limited (PFAL), who commented in their report that "We consider the CS forecasting model to be of good quality and can provide it with a clean bill of health in terms of its design and functionality."[5] The BPM-V3 was subsequently used to produce forecasts for the 2018 and 2020 Business Plans.

This technical memorandum documents the application of the BPM-V3 to produce ridership and revenue forecasts that support the 2020 Business Plan. Section 2.0 documents the implementation steps evaluated. Section 3.0 describes the assumptions related to other transportation systems. Section 4.0 summarizes the socioeconomic forecasts incorporated in the model. Section 5.0 documents the ridership and revenue forecasts, and Section 6.0 explains the risk analysis approach. Readers interested in learning more about the risk analysis process and the range of forecasts are directed to the *2020 California High-Speed Rail Business Plan—Ridership and Revenue Risk Analysis Technical Report.*

Based on the latest available research and economic forecasts, certain model input assumptions for each forecast year were updated for the 2020 Business Plan ridership and revenue forecasts.[6] These include:

- Population, household, and employment forecasts at the county-level and distribution of county-level totals to TAZs.
- Conventional passenger rail to ensure consistency with planned routes and service.
- Auto operating costs.
- Travel time, frequency of service, and stopping patterns for the different phases of the high-speed rail systems.

[5] https://www.hsr.ca.gov/docs/brdmeetings/2016/brdmtg_121316_item2_ATTACHMENT_Ind_Con_Draft_Memo_Ridership_Revenue_for_Valley_to_Valley_Line.pdf.

[6] The model is defined as the structure and parameters that have been estimated, calibrated, and validated to an observed base year (i.e., Year 2010). Updating model input assumptions does not change the model structure or parameters.

## 1.2 Scope of Forecasts

CS developed forecasts for two main phases of the project as specified by the Authority:

1. **Silicon Valley to Central Valley line:** San Francisco and Merced to Bakersfield opening in year 2029.
2. **Phase 1:** San Francisco and Merced to Los Angeles and Anaheim opening in 2033 and an out-year forecast of 2040.

Ridership and revenue forecasts were prepared for the opening year for each implementation step and a Phase 1 out year.[7] The 2040 forecast would reflect ridership and revenue on a mature system that would at the time have more than 10 years of operating history. The 2020 Business Plan lays out an implementation strategy that starts with the Silicon Valley to Central Valley line. The model results for each of these segments is reported.

### *1.2.1 Ridership and Revenue Adjustments to Account for "Ramp up"*

The ridership and revenue forecasts assume a mature high-speed rail system, where potential passengers are fully aware of the system. However, it usually takes some time for a new system to achieve this mature state. The 2020 Business Plan lays out the assumptions to reduce ridership and revenue in the early years of each phase to account for the "ramp up" of ridership and revenue over time.

## 1.3 Disclaimer

The information and results presented in this technical memorandum are estimates and projections that involve subjective judgments, and may differ materially from the actual future ridership and revenue. This technical memorandum is not intended, nor shall it be construed, to constitute a guarantee, promise, or representation of any particular outcome(s) or result(s). Further, the material presented in this technical memorandum is provided solely for purposes of the Authority's business planning and should not be used for any other purpose.

[7] CS also developed Year 2033 forecasts for Silicon Valley to Central Valley. These two sets of forecasts were interpolated to derive ridership and revenue forecasts for interim years.

# 2.0 Phased Implementation Scenarios and High-Speed Rail Assumptions

## 2.1 Scenario Overview

The ridership modeling for the 2020 Business Plan assumes that the California high-speed rail will open in phases and is modeled at 2029, 2033 and 2040 horizon years as described below. The actual implementation timeline differs for both scenarios and extrapolation was used in the Business Plan document to reflect the actual opening dates.

This process was chosen by the Authority to leave the existing modeling framework intact and to ensure comparability to previous business plan model output.

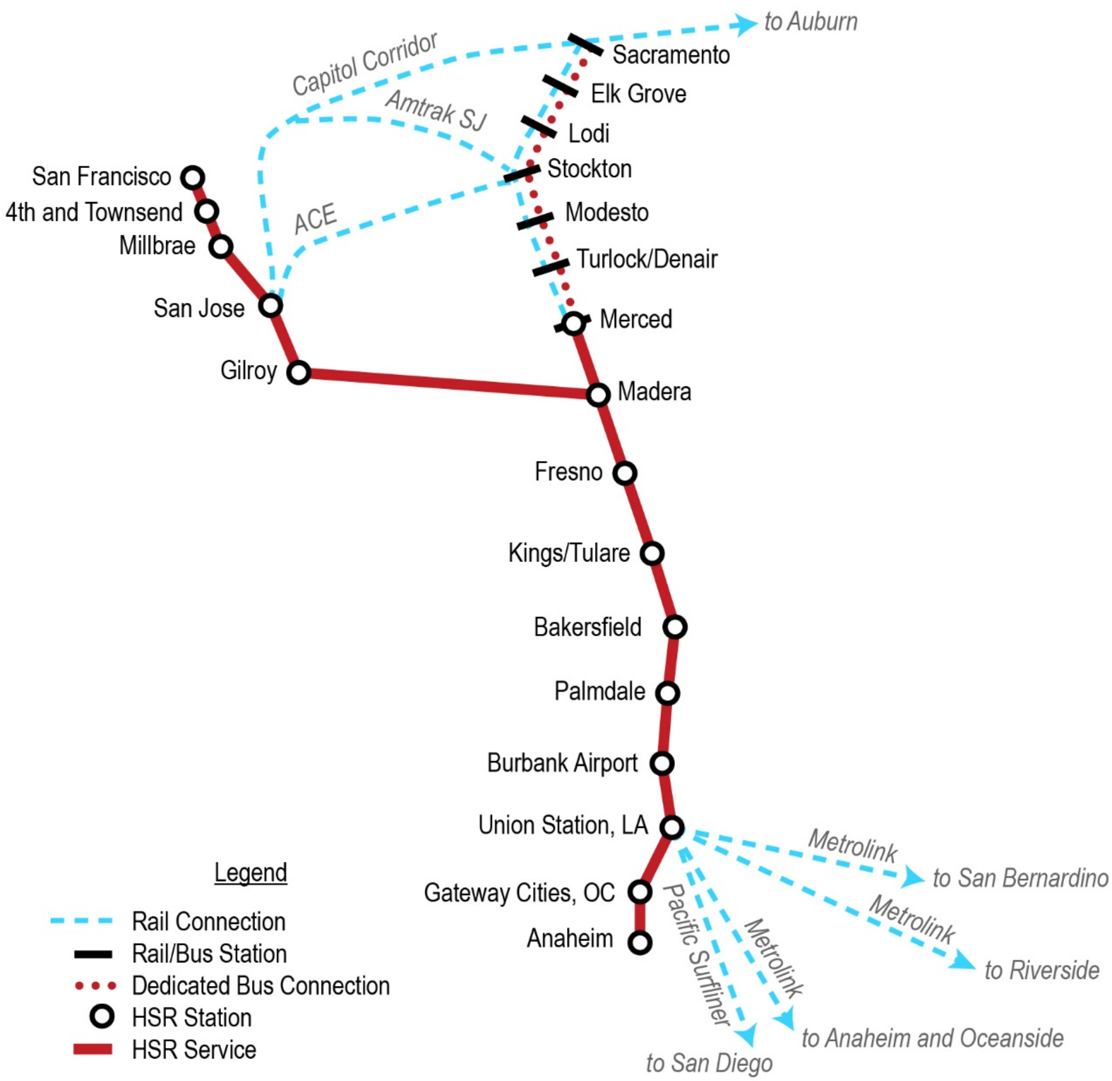
Further detail on the high-speed rail fares and station parking costs are provided in Section 2.2.

### *2.1.1 Silicon Valley to Central Valley Line—Open in 2029*

The Silicon Valley to Central Valley line is assumed to begin service in 2029. It is characterized by:

- A north terminal at San Francisco and Merced and a south terminal at Bakersfield (Figure 2.1).
- High-speed rail service will share the Caltrain corridor between San Francisco and San Jose, meaning that congestion on the corridor and service capacity are considered.
- Dedicated coach services that meet each train will be provided between the Merced station and the Sacramento region, as well as between the line's southern terminus and locations in the Los Angeles Basin (LA Basin).
- Connections with Amtrak at Merced to Sacramento would be coordinated.

## Figure 2.1 Silicon Valley to Central Valley Line




### 2.1.2 Phase 1

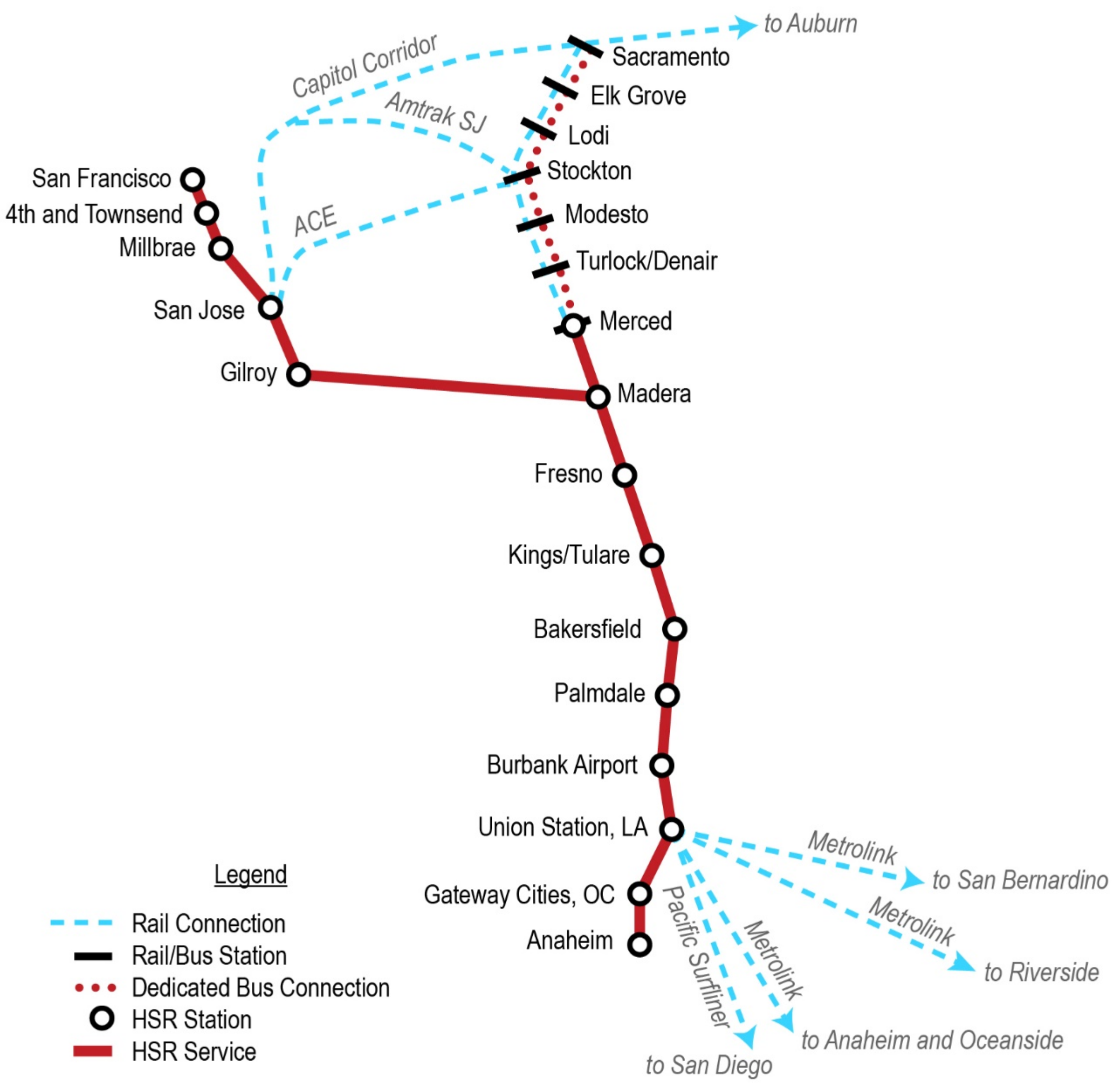
Phase 1 operations are modeled for 2033 and 2040 horizons where Phase 1 extends the high-speed rail system from a north terminal at San Francisco to the south terminal at Anaheim (Figure 2.2), with these characteristics:

- High-speed rail service will share the rail corridor between San Francisco and San Jose as well as Burbank and Anaheim, meaning that congestion on the corridor and service capacity are considered.
- Dedicated coach services would be provided from Merced to Sacramento.
- Cross-platform connections with Amtrak at Madera to the Bay Area and Sacramento would be coordinated.

- Connections with Metrolink feeder service at Los Angeles Union Station (LAUS) to LA Basin destinations would be coordinated.

Service assumptions vary by each implementation step. The details of the service frequencies are described in Table 2.1. The stopping patterns are provided in Appendix A.

## Figure 2.2 Phase 1



## Table 2.1 High-Speed Rail Service Plan Assumptions by Scenario

| Business Plan Scenario | North Terminus | South Terminus | High-Speed Rail Service Summary[a] | Dedicated Peak Bus Coach Connections[b]—North Terminus | Dedicated Peak Bus Coach Connections[b]—South Terminus | Conventional Rail Connections |
|---|---|---|---|---|---|---|
| Silicon Valley to Central Valley Line | San Francisco and Merced | Bakersfield | • 12 peak TPD between San Francisco and Bakersfield (10 in off-peak).<br>• 6 peak TPD between Merced and Bakersfield (11 in off-peak). | • 6 peak BPD between Merced and Sacramento (11 in off-peak). | • 13 peak BPD between Bakersfield and LAUS (20 in off-peak).<br>• 13 peak BPD between Bakersfield and West LA (20 in off-peak).<br>• 13 peak BPD between Bakersfield and Santa Anita (20 in off-peak). | • Connecting service with Amtrak at Merced. |
| Phase 1 | San Francisco and Merced | Los Angeles and Anaheim | • 12 peak TPD between San Francisco and Los Angeles (20 in off-peak).<br>• 12 peak TPD between San Francisco and Anaheim (10 in off-peak).<br>• 12 peak TPD between San Jose and Los Angeles (0 in off-peak).<br>• 12 peak TPD between Merced and Anaheim (10 in off-peak).<br>• 6 peak TPD between San Jose and Merced (10 in off-peak).<br>• 6 off-peak TPD between San Francisco and Merced (10 in off-peak). | • 12 peak BPD between Sacramento and Merced (10 in off-peak). | • None. | • Connecting service with Amtrak at Merced.<br>• Metrolink connections at LAUS. |

[a] TPD—Trains per Day.
[b] BPD—Buses per Day.

## 2.2 Additional High-Speed Rail Assumptions

### 2.2.1 High-Speed Rail Fares

High-speed rail fares for all 2020 Business Plan scenarios are identical to those in the 2018 Business Plan, but escalated from 2017 dollars to 2019 dollars. The fares are based on the formula below, with a $100 maximum fare in 2019 dollars (see Table 2.2):

- $36.30 + $0.224 per mile (in 2019 dollars) for interregional fares.
- $26.94 + $0.187 per mile (in 2019 dollars) for intraregional fares for the SCAG region.
- $17.45 + $0.150 per mile (in 2019 dollars) for intraregional fares for MTC regions.

High-speed rail bus fares in 2019 dollars, rounded to nearest dollar, are as follows:

- $1 from Stockton, Modesto, Denair to Merced.
- $11 from Sacramento, Elk Grove, and Lodi to Merced.
- $14 from Bakersfield to Southern California locations.

## Table 2.2 Assumed High-Speed Rail Fares (2019 Dollars, Rounded to Nearest Dollar)

| High-Speed Rail Stations | San Francisco (Transbay and 4th & King) | Millbrae | San Jose | Gilroy | Merced | Madera | Fresno | Kings/ Tulare | Bakersfield | Palmdale | Burbank Airport | Los Angeles Union Station | Gateway Cities/ Orange County | Anaheim |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| San Francisco (Transbay and 4th & King) | | $20 | $26 | $28 | $66 | $75 | $79 | $88 | $100 | $100 | $100 | $100 | $100 | $100 |
| Millbrae | | | $22 | $27 | $66 | $74 | $79 | $87 | $100 | $100 | $100 | $100 | $100 | $100 |
| San Jose | | | | $21 | $63 | $66 | $71 | $76 | $93 | $100 | $100 | $100 | $100 | $100 |
| Gilroy | | | | | $59 | $62 | $66 | $73 | $88 | $100 | $100 | $100 | $100 | $100 |
| Merced | | | | | | $46 | $50 | $59 | $75 | $95 | $96 | $100 | $100 | $100 |
| Madera | | | | | | | $43 | $51 | $69 | $88 | $89 | $93 | $99 | $99 |
| Fresno | | | | | | | | $45 | $63 | $84 | $85 | $88 | $91 | $94 |
| Kings/Tulare | | | | | | | | | $58 | $75 | $76 | $84 | $86 | $88 |
| Bakersfield | | | | | | | | | | $58 | $59 | $63 | $65 | $67 |
| Palmdale | | | | | | | | | | | $36 | $37 | $39 | $41 |
| Burbank Airport | | | | | | | | | | | | $30 | $34 | $36 |
| Los Angeles Union Station | | | | | | | | | | | | | $30 | $34 |
| Gateway Cities/ Orange County | | | | | | | | | | | | | | $30 |
| Anaheim | | | | | | | | | | | | | | |

Source: Cambridge Systematics, Inc.

### 2.2.2 High-Speed Rail Station Parking Costs

High-speed station parking costs for all 2020 Business Plan scenarios are identical to those in the 2018 Business Plan, but escalated from 2017 dollars to 2019 dollars. The parking costs, which are a combination of actual dollars spent and the value of time spent to walk from the parking garage to the station were developed as follows and shown in Table 2.3:

- San Francisco Transbay and 4th and King parking costs are equivalent to Year 2015 average daily parking cost at surrounding parking garages. These garages are assumed to be approximately a 10-minute walk from each station. An additional $3, or the equivalent to 10 minutes of out-of-vehicle time, have been added to the high-speed rail parking cost as a result.
- Millbrae, San Jose, Burbank Airport, Los Angeles, Gateway Cities/Orange County, and Anaheim parking costs are equivalent to the Year 2015 daily terminal parking costs at nearby airports. It is assumed these high-speed rail station parking garages would be a 10-minute walk from garage to station. An additional $3, or the equivalent to 10 minutes of out-of-vehicle time, have been added to the high-speed rail parking costs at these stations as well.
- Fresno and Bakersfield parking costs are equivalent to the Year 2015 daily terminal parking costs at nearby airports. However, it is assumed that these high-speed rail station parking garages would be just a 5-minute walk from garage to station. An additional $1, or the equivalent to 5 minutes of out-of-vehicle time, have been added to the high-speed rail parking costs at these stations.
- Gilroy, Madera, Kings/Tulare, and Palmdale parking costs were set identical to the parking cost at the Fresno high-speed rail station.

## Table 2.3 High-Speed Rail Station Parking Costs (2019 Dollars)

| Station | Daily Terminal Parking Cost at Nearest Airport[a] | Daily Parking Cost near Station | High-Speed Rail Station Parking Cost |
|---|---|---|---|
| San Francisco | | $38 | $44[b] |
| Millbrae | $34 | | $37[b] |
| San Jose | $34 | | $37[b] |
| Gilroy | | | $15[c, d] |
| Madera | | | $15[c, d] |
| Fresno | $14 | | $15[d] |
| Kings/Tulare | | | $15[c, d] |
| Bakersfield | $14 | | $15[d] |
| Palmdale | | | $15[c, d] |
| Burbank Airport | $29 | | $32[b] |
| Los Angeles | $34 | | $37[b] |
| Gateway Cities/ Orange County | $22 | | $26[b] |
| Anaheim | $22 | | $26[b] |

Source: Cambridge Systematics, Inc.

[a] Airport and downtown garage parking costs reflect Year 2015 prices adjusted to Year 2017 dollars.

[b] Additional 10 equivalent minutes of out-of-vehicle access time added to cost.

[c] Parking cost equivalent to Fresno parking cost.

[d] Additional 5 equivalent minutes of out-of-vehicle access time added to cost.

# 3.0 Service Assumptions for Other Modes and Background Networks

## 3.1 Air Service Assumptions

Average airfares and frequency of air service remain consistent with 2018 Business Plan assumptions. Table 3.1 provides base airfares and headways between select major airports.

### Table 3.1 Air Service Assumptions

| Origin Airport | Destination Airport | Assumed Airfare (2019 Dollars) | Assumed Headway (Minutes) |
|---|---|---|---|
| Burbank | San Francisco | $130 | 480.0 |
| Burbank | Sacramento | $126 | 150.0 |
| Los Angeles | San Diego | $267 | 32.0 |
| Los Angeles | San Francisco | $112 | 23.0 |
| Oakland | San Diego | $125 | 46.0 |
| Oakland | Los Angeles | $125 | 44.0 |
| Sacramento | Burbank | $126 | 150.0 |
| Sacramento | San Francisco | $336 | 141.0 |
| San Francisco | San Diego | $108 | 28.0 |
| San Francisco | Burbank | $130 | 480.0 |

Source: Aviation System Consulting.

## 3.2 Conventional Passenger Rail Service

Conventional rail (CVR) service, including travel times, frequency of service, and stations served, were updated to match with specifications received from DB.[8] The service plan was largely similar to the service assumed in the 2018 Business Plan which reflected the then latest 2018 California State Rail Plan (CSRP).[9] The largest service changes from current service include route changes and increased frequency of service on the San Joaquin line to reflect the introduction of high-speed rail, increased Caltrain service due to electrification, and increased service frequency on the Pacific Surfliner and Metrolink in the LA Basin. The operating routes and frequencies assumed in the 2020 Business Plan are summarized in Table 3.2. Fare assumptions for all CVR lines are consistent with online published fares in 2011, in real dollars.

[8] 2020 Business Plan Service Plan Assumptions V2.04 (Draft), 09-13-2019, DB Engineering & Consulting USA Inc.

[9] 2018 California State Rail Plan (Draft), October 2017, available at: http://www.dot.ca.gov/californiarail/.

### Table 3.2 CVR Operating Plan Service Frequencies

| Service and Route | Current Operation | Year 2029 (V2V) | Year 2033, 2040 (Ph1) |
|---|---|---|---|
| **San Joaquin Route (Daily Trains between city pairs)** | | | |
| Sacramento—Merced via San Joaquin Route | 0 | 9 | 9 |
| Sacramento—Bakersfield via San Joaquin Route | 2 | 0 | 0 |
| Oakland—Bakersfield via San Joaquin Route | 5 | 0 | 0 |
| Oakland—Merced via San Joaquin Route | 0 | 5 | 5 |
| **Ace Route (Daily Trains between city pairs)** | | | |
| San Jose—Stockton via ACE Route | 4 | 6 | 6 |
| San Jose—Merced via ACE Route | 0 | 1 | 1 |
| **Capitol Corridor (Daily trains between segments)** | | | |
| Auburn to Roseville | 1 | 1 | 1 |
| Roseville to Sacramento | 1 | 2 | 2 |
| Sacramento to Oakland | 15 | 20 | 20 |
| Oakland to San Jose | 7 | 7 | 7 |
| **Caltrain (Daily Trains between city pairs)** | | | |
| San Francisco to San Jose/Gilroy | 46 | 66 ($4^{th}$/King) | 66 (Transbay) |
| **Pacific Surfliner (Daily Trains between city pairs)** | | | |
| San Luis Obispo—Los Angeles | 2 | 3 | 3 |
| Goleta—Los Angeles | 3 | 5 | 5 |
| Los Angeles—San Diego | 12 | 14 | 14 |
| **Metrolink & COASTER (Daily Trains by line)** | | | |
| Antelope Valley | 15 | 17 | 17 |
| Perris Valley 91 | 7 | 10 | 10 |
| Orange County | 15 | 20 | 20 |
| Ventura County | 17 | 17 | 17 |
| Inland Empire | 8 | 14 | 14 |
| Riverside | 6 | 6 | 6 |
| San Bernardino | 19 | 20 | 20 |
| Oceanside—San Diego (Coaster) | 11 | 19 | 19 |

Source: Cambridge Systematics, Inc.

## 3.3 Urban Transit Network

Urban rail and bus service, including travel times, frequency of service, and stations served, are identical to those in the 2018 Business Plan.

## 3.4 Highway Network

CS used the same highway network assumptions as those used for the CSTDM 2.0 forecast years.[10] CS averaged AM and PM peak congested travel times derived from the CSTDM 2.0 for use when peak travel times were needed in the mode choice model. Similarly, CS averaged midday and off-peak congested speeds for when off peak travel times were needed.

Auto terminal times represent the average time to access one's vehicle at each end of the trip and are added to the congested travel time to get the total congested travel time skim. They are based on the area type of the trip ends and are assessed at both the origin and destination of the trip. The auto terminal time assumptions are consistent with the 2010 base year scenario, and discussed in the BPM-V3 model documentation.

Travel times for the modeled forecast years were obtained by interpolating between the closest forecast years, and are identical to those used in the 2018 Business Plan.

Auto costs (besides operating costs discussed in Section 3.5) are comprised of tolls and parking costs. Toll costs were imported from networks developed for the CSTDM 2.0. Tolls corresponding to single-occupancy vehicles were assumed in the auto skims. Peak and off-peak tolls were averaged where costs differed. The parking costs developed for the 2010 base year scenario, and discussed in the BPM-V3 model documentation, were used for all future year scenarios.

## 3.5 Automobile Operating Cost

The auto operating cost is multiplied by the distance traveled by auto to determine the cost incurred when a long-distance trip or an access or egress trip is made by auto. Future auto operating costs consider 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 14.2 percent higher than the national average (based on consistent patterns in past trends).
2. An estimate of the future market penetration rate of electric vehicles, along with accompanying costs for electricity, miles per gallon equivalent (MPGe) rating to determine energy costs for electric vehicles, and the cost of electricity. These estimates were developed from the 2019 Annual Energy Outlook produced by the EIA.
3. Additional fees due to Cap and Trade implementation.
4. The fuel economy of the entire "on the road" fleet, calculated from the 2019 Annual Energy Outlook (AEO).
5. Nonfuel costs, which were obtained from the Bureau of Transportation Statistics (BTS).

[10] Cambridge Systematics, Inc., "California Statewide Travel Demand Model, Version 2.0, Model Overview," prepared for the California Department of Transportation, June 2014.

The following formula was used to develop the auto operating cost:

*Auto Operating Cost = (1 - %EVs) * (CA Gas Price + C&T Impact) / Fuel Efficiency + %EVs * (CA Electricity Price * 33.7) / EV Fuel Efficiency + Nonfuel Operating Costs*

Table 3.3 gives the auto operating cost component values used in the above formula. More information on the development of each of these components can be found in the following sections.

## Table 3.3 Auto Operating Costs (2019 Dollars)

| Auto Operating Cost Component, by Year | Auto Operating Cost Components |
|---|---|
| **2029 Auto Operating Cost ($/mile)** | $0.26 |
| U.S. Gas Price ($/gal) | $3.25 |
| California Gas Price ($/gal) | $3.71 |
| California Electricity Price ($/kWH) | $0.16 |
| % Electric Vehicles | 5.81% |
| MPG | 27.8 |
| MPGe | 69.56 |
| Non-fuel cost ($/mi) | $0.12 |
| Cap & Trade ($/gal) | $0.27 |
| **2033 Auto Operating Cost ($/mile)** | $0.25 |
| U.S. Gas Price ($/gal) | $3.54 |
| California Gas Price ($/gal) | $4.04 |
| California Electricity Price ($/kWH) | $0.19 |
| % Electric Vehicles | 10.34% |
| MPG | 33.4 |
| MPGe | 83.62 |
| Non-fuel cost ($/mi) | $0.12 |
| Cap & Trade ($/gal) | $0.52 |
| **2040 Auto Operating Cost ($/mile)a** | $0.24 |
| U.S. Gas Price ($/gal) | $3.69 |
| California Gas Price ($/gal) | $4.21 |
| California Electricity Price ($/kWH) | $0.20 |
| % Electric Vehicles | 14.24% |
| MPG | 36.7 |
| MPGe | 91.77 |
| Non-fuel cost ($/mi) | $0.12 |
| Cap & Trade ($/gal) | $0.59 |

Source: Cambridge Systematics, Inc.

### 3.5.1 Fuel Component of Auto Operating Costs

Forecasts of future fuel costs are a function of the cost of fuel and vehicle fuel economy. Each of these is discussed below.

**Motor gasoline price forecasts.** The gasoline price forecast was based on the U.S. Energy Information Administration's (EIA) 2019 Annual Energy Outlook (AEO). CS updated the projected motor gasoline prices in California based on the 2019 AEO, which extends through 2050. The EIA provides average motor gasoline price forecasts for three different oil price scenarios: 1) reference, 2) low, and 3) high. Historically, California's retail gasoline prices have been higher than the U.S. average; the overall average for California prices over the U.S. average prices between 2001 and 2019 has been 14.2 percent. CS developed a forecast of California gasoline prices by taking the reference forecasts from EIA and increasing them by 14.2 percent.

On January 1, 2015, the Cap and Trade rules came into effect for the fuel sector in California. The California Legislative Analyst's Office (LAO) estimated in 2017 that Cap and Trade could add $0.15 to $0.63 per gallon to retail gasoline prices in 2021. Independent projections from the Ontario Energy Board (OEB) and the California Energy Commission (CEC) of the long-term price of carbon on the Cap and Trade auction market provided estimates of the most likely price per ton of carbon as a function of the maximum price through 2028. To determine the impact of Cap and Trade in the model forecast years, the maximum per gallon impact of Cap and Trade is interpolated from the 2017 LAO letter, and then multiplied by the OEB/CEC projections of price per ton of carbon.

**Fuel Economy Forecasts.** Fuel economy projections of all on-the-road light-duty vehicles was based on the 2019 AEO. These fuel economy forecasts were used to calculate the fuel economy of only the non-electric portion of the vehicle fleet.

The 2019 AEO also provides an estimate for the fuel efficiency of new vehicles sold in each year of the forecast, including electric and other alternative fuel vehicles. It does not, however, include an estimate for the fuel efficiency of the on-the-road fleet of alternative fuel vehicles. In order to capture the higher fuel efficiency of the electric fleet consistent with other assumptions, the equivalent miles per gallon efficiency of the electric fleet is set at 2.5 times the projected fuel efficiency of the Light Duty Stock Fleet, which reflects recent trends calculated from the 2019 AEO.

The market penetration of electric vehicles is estimated as the number of total electric fuel vehicles (including both cars and light trucks) divided by the total size of the stock vehicle fleet. Penetration rates were calculated under the 2019 AEO's reference case.

### 3.5.2 Non-Fuel Component of Auto Operating Costs

The 2010 model base year scenario uses 10-cents per mile non-fuel cost in 2019 dollars, as discussed in the BPM-V3 model documentation. The Bureau of Transportation Statistics (BTS) publishes historical average nonfuel auto operating costs which includes maintenance, tires, insurance, license, registration and taxes, depreciation, and finance. Non-fuel cost for all forecast years was assumed to be the average non-fuel cost from 1991 and 2015. Since this average is one cent higher than the Year 2010 BTS non-fuel cost, fuel cost was set at 11 cents per mile to reflect the change in non-fuel cost between Year 2010 and the forecasted value.

# 4.0 Socioeconomic Forecast

Updated socioeconomic projections were developed to support ridership and revenue forecasting for the 2018 Business Plan. Statewide forecasts of population, households, and employment were assembled from four independent forecasting sources. The sources include:

- The California Economic Forecast (CEF) for the Caltrans Transportation Economics Branch (2018).
- Moody Analytics (2019).
- MPO data (assembled by CS from plans available through April 2017).[11]
- California Department of Finance (DOF), Demographic Research Unit (Baseline 2016, updated in 2019).[12]

The population and employment forecasts showed high degrees of similarity despite a wide variation in household forecasts across the state. The differences in the household forecasts were caused by differing assumptions regarding future household sizes. The Moody Analytics forecasts assumed decreasing average household sizes over time, reflecting a continuing trend of lower birthrates in the U.S. and other developed countries. The CEF data assumed that average household sizes also decrease over time but at a slower rate than the Moody Analytics forecasts. For 2040, the Moody Analytics data, CEF data, and MPO data reflected average household sizes of 2.79, 2.98, and 2.88, respectively.

County-level forecasts of population, households, and employment were averaged to produce the county-level forecasts for the 2020 Business Plan. While data from the MPO plans may have been produced prior to 2016, they were included in the averages to incorporate local knowledge. By averaging households from the sources, the resulting statewide average household size was 2.90. Each of the county-level forecasts were disaggregated to TAZs based on the MPO TAZ-level forecasted distributions. Major market forecasts are shown in Table 4.1.

[11] Socioeconomic forecasts based on local MPO data was not updated for the 2020 Business Plan and is consistent with assumptions for the 2018 Business Plan. The MPO data forecasts were updated due to schedule considerations.

[12] Due to forecast availability, DOF data were used only to inform population projections, but not other economic forecasts.

## Table 4.1 Regional Socioeconomic Forecasts (Millions)

| Region | Population | | | Households | | | Employment | | |
|---|---|---|---|---|---|---|---|---|---|
| | 2029 | 2033 | 2040 | 2029 | 2033 | 2040 | 2029 | 2033 | 2040 |
| MTC | 8.44 | 8.68 | 9.12 | 3.05 | 3.13 | 3.27 | 4.61 | 4.72 | 4.90 |
| SCAG | 20.36 | 20.79 | 21.51 | 6.76 | 6.95 | 7.25 | 9.32 | 9.56 | 9.97 |
| San Joaquin Valley | 4.74 | 4.93 | 5.30 | 1.46 | 1.53 | 1.66 | 1.73 | 1.80 | 1.95 |
| SACOG | 2.79 | 2.88 | 3.09 | 1.01 | 1.06 | 1.13 | 1.24 | 1.28 | 1.38 |
| SANDAG | 3.58 | 3.66 | 3.78 | 1.27 | 1.31 | 1.36 | 1.74 | 1.78 | 1.85 |
| Other | 2.90 | 2.98 | 3.07 | 1.08 | 1.10 | 1.17 | 1.26 | 1.28 | 1.33 |
| **Total** | **42.81** | **43.92** | **45.87** | **14.63** | **15.08** | **15.84** | **19.90** | **20.42** | **21.38** |

Source: Cambridge Systematics, Inc.

# 5.0 Ridership and Revenue Forecast Results

## 5.1 Summary of Assumptions

Table 5.1 summarizes the input assumptions for each high-speed rail operating plan and forecast year.

## Table 5.1 Summary of High-Speed Rail Assumptions for Each Modeled Business Plan Phase

| Category | Year 2029 | Year 2033 | Year 2040 |
|---|---|---|---|
| High-speed rail Phase | Silicon Valley to Central Valley Line | Phase 1 | Phase 1 |
| Highway Network | Year 2029[a] | Year 2033[a] | Year 2040[a] |
| Auto Travel Time | Year 2029[b] | Year 2033[b] | Year 2040[b] |
| Auto Parking | Year 2010 | Year 2010 | Year 2010 |
| Air Travel Time | Year 2012[c] | Year 2012[c] | Year 2012[c] |
| Air Service Frequency | Year 2012[c] | Year 2012[c] | Year 2012[c] |
| Air Reliability | Year 2010[d] | Year 2010[d] | Year 2010[d] |
| Parking Cost at Airport | Year 2010 | Year 2010 | Year 2010 |
| CVR Service Plans | 2020 BP Assumptions[e] | 2020 BP Assumptions[e] | 2020 BP Assumptions[e] |
| CVR Fares | Year 2010 | Year 2010 | Year 2010 |
| CVR Reliability | Year 2010[f] | Year 2010[f] | Year 2010[f] |
| Parking Cost at CVR Station | Year 2010 | Year 2010 | Year 2010 |
| High-speed Rail Service Plan | 2020 BP for VtoV | 2020 BP for Phase 1 | 2020 BP for Phase 1 |
| High-speed Rail Fares | 2016 BP (83% of airfare) | 2016 BP (83% of airfare) | 2016 BP (83% of airfare) |
| High-speed Rail Reliability | 2016 BP (99%) | 2016 BP (99%) | 2016 BP (99%) |
| High-speed Rail Parking Cost | Year 2015 | Year 2015 | Year 2015 |
| Urban/Light Rail Service Plans | Consistent with 2017 MPO RTP forecasts | Consistent with 2017 MPO RTP forecasts | Consistent with 2017 MPO RTP forecasts |
| Other Transit Lines | Year 2017 | Year 2017 | Year 2017 |
| Socioeconomic Data | Year 2029 | Year 2033 | Year 2040 |
| Auto Operating Cost | 26 cents/mile | 25 cents/mile | 24 cents/mile |
| Air Fares | Year 2009[c] | Year 2009 | Year 2009 |

[a] The high-speed rail master highway network was developed based on the CSTDM 2.0 highway network for each respective forecast year. Thus, the highway “build” assumptions are consistent with those used for the CSTDM 2.0.

[b] The auto travel times for peak and off-peak were developed by transferring link speeds from the loaded CSDTM 2.0 AM peak and off-peak congested speeds for year 2020, 2035 and 2040 on to the corresponding year high-speed rail highway network, and then skimming the high-speed rail network to obtain peak and off-peak travel times. Travel times for the modeled forecast years were obtained by interpolating between the closest forecast years. The main mode auto times reflect an average of peak and off-peak travel times.

[c] Air service frequency, travel times, and fares remain consistent with the 2016 Business Plan, which were developed in 2011 by CS and ASC.

[d] Air reliability remains consistent with Bureau of Transportation Statistics published data for year 2010 (http://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp?pn=1).

[e] The CVR service plan, including travel times, frequency of service, and stations served, are based on the modified assumptions developed by DB for the 2020 Business Plan and based on the 2018 California State Rail Plan (SRP) and coordination with stakeholders.

[f] CVR reliability remains consistent with year 2010 reliability assumptions developed from information published by each CVR operator.

## 5.2 Summary of Ridership and Revenue Forecasts

The base case ridership and revenue forecasts are shown in Table 5.2. Ridership is presented in millions of annual passengers for each implementation step starting with the Silicon Valley to Central Valley line in year 2029 and Phase 1 in Years 2033 and 2040. Annual revenue is reported in millions of June 2019 dollars for the same implementation steps and forecast years.[13]

### Table 5.2 Annual Ridership and Revenue by Implementation Step (Millions)

| Ridership and Revenue | Silicon Valley to Central Valley 2029 | Phase 1 2033 | Phase 1 2040 |
|---|---|---|---|
| Ridership (in millions) | 16.2 | 35.6 | 38.6 |
| Revenue (in June 2019 dollars) | $982 | $2,207 | $2,410 |

Source: Cambridge Systematics, Inc.

## 5.3 Ridership and Revenue Forecast Comparisons by Implementation Step and Year

A comparison of forecasts for the Silicon Valley to Central Valley line in year 2029, Phase 1 year 2033, and Phase 1 year 2040 annual trips by major market is shown in Table 5.3. These values are shown for illustrative purposes to provide a sense of how ridership and revenue varies by project phase for particular region pairs and at particular stations. The values represent a mature system that do not account for the time it takes for customers to become fully familiar with a new service.

The Silicon Valley to Central Valley line is assumed to provide less frequent high-speed rail service compared to Phase 1. The Silicon Valley to Central Valley line provides two peak trains per hour (TPH) between San Francisco and Bakersfield. Dedicated coach services are assumed to be provided to Sacramento and the Los Angeles Basin. However, the coach service results in longer travel times to the Los Angeles Basin relative to Phase 1. The markets forecasted to have the highest high-speed rail mode shares in the Silicon Valley to Central Valley line include the longer-distance markets and those involving the MTC region.[14] For example, the MTC to SCAG market will have the highest mode share at 10.0 percent, followed by MTC to the San Joaquin Valley at 9.0 percent.

The lower high-speed rail mode share in the MTC to San Joaquin Valley market is partially explained by the market dynamics. MTC to San Joaquin Valley has about twice the number of total person trips as MTC to SCAG (50.1 versus 24.3 million for Year 2029) and is dominated by auto travel, which is forecasted to carry about 89 percent of the overall demand. The MTC to SCAG market, on the other hand, has a well-established air market compared to MTC to San Joaquin Valley. In longer-distance markets, high-speed rail diverts a smaller share from autos and a greater share from air travel. While the absolute number of high-speed rail riders in the MTC to San Joaquin Valley market is forecasted to be higher than MTC to SCAG, the

[13] State of California Department of Industrial Relations. *California Consumer Price Index Chart*. https://www.dir.ca.gov/OPRL/CPI/EntireCCPI.PDF.

[14] Mode share is defined as the percentage of the total travel market riding a particular mode. It is calculated by dividing the total person trips on high-speed rail by the sum of the person trips on all modes (auto person trips, conventional rail person trips, air person trips, and high-speed rail person trips).

mode share is lower because high-speed rail is not as competitive in shorter-distance markets where autos are the dominant mode.

Extending the high-speed rail system to Phase 1 (where it stretches from San Francisco and Merced to Los Angeles and Anaheim) provides more access to the most populous areas in the state. Compared to the Silicon Valley to Central Valley line, the high-speed rail mode share more than doubles to 25 percent in Year 2033 between MTC and SCAG. This increase is driven by the system's extension to SCAG and now provides a one-seat ride from northern California to southern California adding new opportunities for people to access stations closer to them within the state's largest metropolitan areas.

## Table 5.3 Comparison of Annual Ridership (Millions) and Revenue (Millions, 2019 Dollars) by Major Market[a]

| Market | | Year 2029 Silicon Valley to Central Valley | | | Year 2033 Phase 1 | | | Year 2040 Phase 1 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | High-Speed Rail Riders | High-Speed Revenue | High-Speed Rail Share | High-Speed Rail Riders | High-Speed Revenue | High-Speed Rail Share | High-Speed Rail Riders | High-Speed Revenue | High-Speed Rail Share |
| SACOG | SACOG | 0.0 | $0.0 | 0.0% | 0 | $0.00 | 0.00% | 0 | $0.00 | 0.00% |
| SACOG | SANDAG | 0.0 | $2.2 | 2.4% | 0.1 | $8.10 | 6.40% | 0.1 | $8.70 | 6.10% |
| SACOG | MTC | 0.6 | $12.6 | 0.9% | 0.9 | $19.70 | 1.30% | 1 | $22.00 | 1.30% |
| SACOG | SCAG | 0.3 | $26.9 | 4.6% | 0.8 | $80.80 | 10.20% | 0.9 | $89.70 | 10.10% |
| SACOG | San Joaquin Valley | 0.2 | $13.0 | 1.5% | 0.2 | $15.10 | 1.60% | 0.3 | $17.00 | 1.60% |
| SACOG | Other regions | 0.1 | $3.2 | 0.5% | 0.1 | $4.30 | 0.60% | 0.1 | $4.20 | 0.60% |
| SANDAG | SANDAG | 0.0 | $0.0 | 0.0% | 0 | $0.00 | 0.00% | 0 | $0.00 | 0.00% |
| SANDAG | MTC | 0.2 | $18.7 | 4.4% | 0.7 | $70.60 | 15.40% | 0.8 | $77.30 | 15.60% |
| SANDAG | SCAG | 0.0 | $0.0 | 0.0% | 2.8 | $86.80 | 2.10% | 3 | $92.30 | 2.10% |
| SANDAG | San Joaquin Valley | 0.1 | $8.3 | 3.5% | 0.4 | $33.70 | 10.30% | 0.4 | $37.00 | 10.20% |
| SANDAG | Other regions | 0.0 | $3.0 | 1.2% | 0.1 | $12.00 | 4.50% | 0.1 | $12.40 | 4.60% |
| MTC | MTC | 1.8 | $46.9 | 4.8% | 2.2 | $56.30 | 5.50% | 2.3 | $58.10 | 5.30% |
| MTC | SCAG | 2.4 | $237.3 | 10.0% | 6.4 | $642.30 | 25.20% | 7.2 | $716.50 | 25.90% |
| MTC | San Joaquin Valley | 4.3 | $314.4 | 8.6% | 4.6 | $336.40 | 8.70% | 5.2 | $374.00 | 8.80% |
| MTC | Other regions | 2.3 | $63.9 | 4.7% | 2.7 | $78.20 | 5.30% | 2.7 | $80.40 | 5.20% |
| SCAG | SCAG | 0.0 | $0.0 | 0.0% | 4.6 | $156.40 | 2.80% | 5.1 | $174.30 | 2.80% |
| SCAG | San Joaquin Valley | 0.9 | $59.1 | 2.6% | 4.6 | $338.70 | 12.50% | 5 | $367.80 | 12.20% |
| SCAG | Other regions | 0.4 | $33.3 | 1.3% | 1.2 | $102.40 | 3.80% | 1.2 | $106.80 | 3.90% |

| Market | | Year 2029 Silicon Valley to Central Valley | | | Year 2033 Phase 1 | | | Year 2040 Phase 1 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | High-Speed Rail Riders | High-Speed Revenue | High-Speed Rail Share | High-Speed Rail Riders | High-Speed Revenue | High-Speed Rail Share | High-Speed Rail Riders | High-Speed Revenue | High-Speed Rail Share |
| San Joaquin Valley | San Joaquin Valley | 1.6 | $93.1 | 7.2% | 1.8 | $103.80 | 7.50% | 1.9 | $109.80 | 7.10% |
| San Joaquin Valley | Other regions | 0.6 | $40.6 | 2.6% | 0.7 | $44.70 | 2.70% | 0.7 | $44.30 | 2.60% |
| Other regions | Other regions | 0.1 | $4.3 | 0.5% | 0.1 | $5.60 | 0.60% | 0.1 | $5.50 | 0.50% |
| Long-Distance Total | | 16.2 | $980.7 | 2.3% | 35.1 | $2,195.80 | 4.80% | 38.1 | $2,398.20 | 4.80% |
| MTC (< 50 miles) | MTC (< 50 miles) | 0.0 | $1.0 | 0.0% | 0.3 | $7.30 | 0.00% | 0.4 | $8.10 | 0.00% |
| SCAG (<50 miles) | SCAG (< 50 miles) | — | $0.0 | 0.0% | 0.1 | $3.8 | 0.0% | 0.1 | $3.8 | 0.0% |
| Short-Distance Total[b] | | 0.0 | $1.0 | 0.0% | 0.4 | $11.10 | 0.00% | 0.5 | $12.00 | 0.00% |
| **Total** | | **16.2** | **$981.71** | | **35.6** | **$2,207.00** | | **38.6** | **$2,410.10** | |

[a] With the exception of the SCAG and MTC regions, only long-distance trips (trips made to locations 50 or more miles from a traveler's home) are shown in the table. In the SCAG and MTC regions, separate summaries of intraregional trips made to locations less than 50 miles from the travelers' homes also are shown.

[b] Only short-distance auto, high-speed rail, and conventional rail modes are used to calculate mode share.

# 6.0 Risk Analysis

## 6.1 Approach

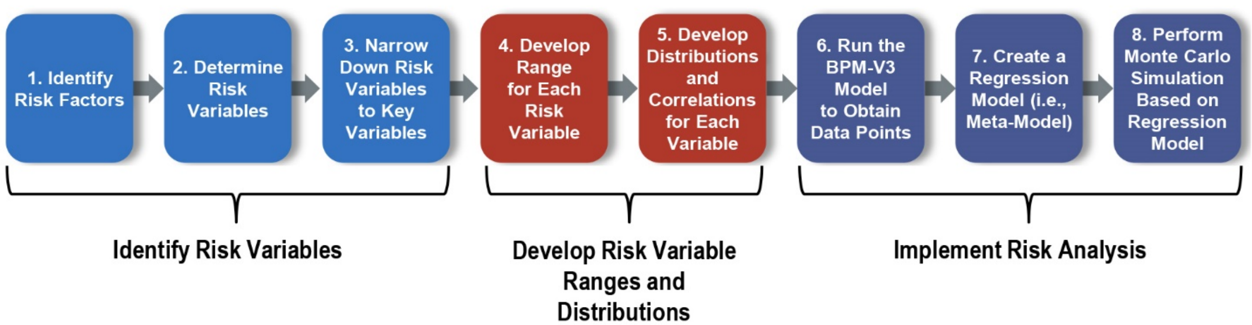
The purpose of the risk analysis is to incorporate the uncertainty associated with model inputs and assumed travel behavior into the 2020 Business Plan HSR ridership and revenue forecasting process. A risk analysis approach has been developed that expresses forecast results as probabilities of achieving different outcome levels. This approach is identical to the previous risk analyses performed for the 2018 Business Plan.

In order to develop full ranges of possible ridership and revenue forecasts, 150 full model runs were performed for each forecast year to estimate relationships between forecast revenue and ridership and selected input risk variables. These runs were used to create a "meta-model" to generate thousands of revenue and ridership forecasts over the entire ranges of identified risk variables without requiring computationally expensive and time-consuming full model runs. These thousands of revenue and ridership forecasts were used to develop probability distributions of total HSR revenue and ridership.

The initial step in the risk analysis was the identification of 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). Second, the impact of each risk factor was assigned to a model variable or 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. Third, the "meta-model" was coupled with researched distributions of the model inputs and used in a Monte Carlo simulation to develop 100,000 unique forecasts of revenue and ridership. Finally, probability distributions of total revenue and ridership were estimated from the results of the Monte Carlo simulation.

An eight-step risk analysis approach was employed to forecast a range of revenue and ridership forecasts for the 2020 Business Plan, as shown in Figure 6.1 and detailed below.

### Figure 6.1 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.
- The final risk factors for each forecast year were chosen based on their likelihoods of affecting ridership and revenue for the forecast year.

### *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 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.
- Based on the range and known sensitivity of the risk variables under consideration, final sets of risk variables were 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 2029 is predicted to range from $0.20 per mile to $0.34 per mile (stated in 2019 dollars), with a most likely value of $0.25 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 of possible values 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 $0.20 per mile or the maximum value of $0.34 per mile, but very likely it will be close to $0.25 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 defined sets of risk variable levels to obtain data points for estimation of two sets of regression models (i.e., meta-models) for each forecast year that estimates the values of the dependent variables, either HSR revenue or ridership, based on values of the selected input risk variables*

- The sets of BPM-V3 specified model runs were developed using a modified Latin hypercube sample design process to ensure that the data points represented the solution space effectively.[15]
- A Gaussian Process Regression (GPR) was used to develop the meta-model. GPR does not impose a restriction on the functional form of the output (e.g., it does not need to be linear or any particular defined nonlinear function). Instead, the functional form is developed on the reasonable assumption that, if two observations have inputs that are similar, then the output should also be similar.

### *Step 8. Perform a Monte Carlo simulation by running the GPR model 100,000 times with varying levels of the input variables based on the distributions assigned to the variables*

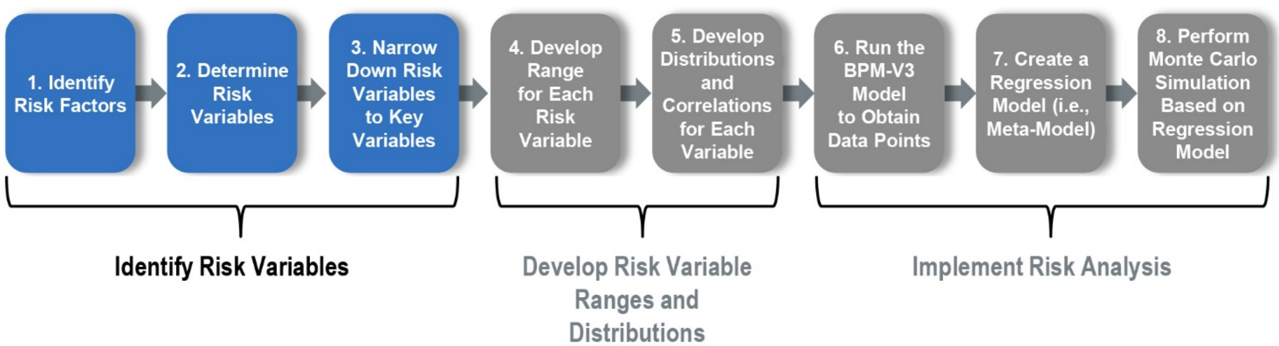
- The simulation results in probability distributions of HSR revenue and ridership.
- The results of the simulation were analyzed to determine the relative contribution of each risk factor on revenue and ridership.

The rest of this section is divided into three sections that provide insight into the steps taken to produce the simulation results: Identification of Risk Variables (*Steps 1 to 3*), Development of Risk Variable Ranges and Distributions (*Steps 4 to 5*), and Risk Analysis implementation (*Steps 6 to 8*).

## 6.2 Identification of the Risk Variables

This section details the steps taken to identify the risk variables included in the risk analysis, as shown in Figure 6.2 below.

**Figure 6.2 Eight-Step Risk Analysis Approach: Identifying Risk Variables (Steps 1 to 3)**




[15] Latin hypercube sampling is a statistical method for producing a close to random sample of values from a multidimensional distribution.

To develop a set of potential risk factors (*Step 1*), CS held a series of meetings with key stakeholders and staff to review the potential risks originally identified by a panel of experts for the 2018 Business Plan Risk Analysis and identify any changes to those potential risks or new 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 2029, 2033, and 2040? 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.

The list of potential risk factors generated through that discussion 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? How does one account for these risks in the model? Next, sensitivity runs of the BPM-V3 model were assessed for each risk variable that allowed for a quantitative comparison of the impacts of each risk variable on ridership and revenue. Based on the sensitivity analysis, risk variables that were determined to have the greatest effect on high-speed rail ridership and revenue and the highest potential uncertainty for each forecast year were selected for inclusion (*Step 3*). Sets of risk variables were included in the risk analysis for each forecast year, as shown in Table 6.1. This table also documents the risk factors that are represented by each risk variable. The risk variables are identical to those in the 2018 Business Plan.

## Table 6.1 Variables Included in Risk Analysis for Each Analysis Year

| Number | Risk Variable | Reasons for Considering Model Variable and Risk Factors Represented |
|---|---|---|
| **1 (All Years)** | 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 (All Years)** | Commute HSR Mode Choice Constant | |
| **3 (All Years)** | Recreation/Other HSR Mode Choice Constant | |
| **4 (All Years)** | 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 (All Years)** | Recreation/Other Trip Frequency Constant | |
| **6 (All Years)** | Auto Operating Costs | This variable reflects the inherent risks in forecasting future fuel costs, fuel efficiencies, the adoption of alternative fuels/electric vehicles, maintenance costs, changes in gas taxes, potential impacts of cap and trade on fuel costs, and for 2040, market penetration of autonomous connected vehicles, autonomous vehicle fuel economy, higher shares of "shared use" vehicles, and shared use vehicle operating costs. |
| **7 (All Years)** | 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 (All Years)** | HSR Frequency of Service | With final service plans expected to be developed by a private operator, 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 2029)** | 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. |
| **10 (Year 2033 and 2040)** | 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. |

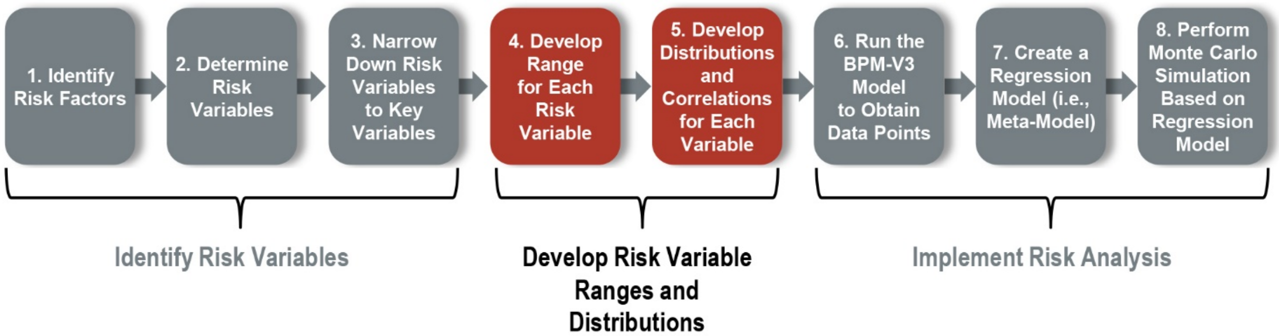
| Number | Risk Variable | Reasons for Considering Model Variable and Risk Factors Represented |
|---|---|---|
| **11 (All Years)** | Coefficient on Transit Access-Egress Time/Auto Distance Variable | Between some regions in California, especially in the Silicon Valley to Central Valley line scenario, 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 to estimate the uncertainty around the effect of these transfers on HSR ridership and revenue. |
| **12 (All Years)** | 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 over time. |
| **13 (Year 2040)** | Auto In-Vehicle Travel Time Coefficient | The introduction of autonomous vehicles may change the way travelers view auto travel due to the substitution of other activities such as sleeping, reading, Internet communications for the time spent driving. |
| **14 (All Years)** | HSR Reliability | Early implementation issues with equipment and operations could affect HSR reliability in the early stages of each phase. Overall HSR reliability may not match international experience on which the original 99 percent reliability assumption is based. |
| **15 (All Years)** | Long Access and Egress | Reliably estimating parameters for exceptionally long access and egress from currently available survey data is very difficult. There are very few observed trips with these attributes (e.g., there are no observations of access and egress by any mode over three hours). In addition, access and egress times to main modes are generally correlated: if your origin is very far from an airport, you are usually also very far from a train station, and vice versa. This will not necessarily be the case for HSR, since it is possible to be far from HSR but close to an airport or CVR stations. This risk variable is used to estimate the uncertainty around the effect of exceptionally long access and egress on HSR ridership and revenue.[a] |
| **16 (All Years)** | Additional Induced Travel | Induced travel forecasted by the model is low compared to what has been observed on international HSR systems. |
| **17 (All Years)** | Visitor Travel | The model-only forecasts travel by California residents. However, in 2016 there were 60 million annual visitors to California. These visitors, especially those that travel by air to arrive in California, may find HSR a desirable option for traveling between various locations in California. |

[a] Note that this risk variable is focused on exceptionally long access and egress by any mode in distance ranges where there were virtually no observed data. In contrast, Risk Variable 11 focused on transit access or egress in relation to the total trip distance; while observed data existed to estimate the coefficient, the applicability for ranges of access/egress to HSR is less certain.

## 6.3 Development of Risk 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 6.3, 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 6.3 Eight-Step Risk Analysis Approach: Develop Risk Variable Ranges and Distributions (Steps 4 to 5)**




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.

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 Monte Carlo simulation. The shape of the distribution can be triangular, PERT, uniform, or another form. The shape of the distribution around the minimum, most likely, and maximum values of each risk variable was determined based on the level of uncertainty surrounding each of the three data points.

Table 6.2, Table 6.3, and Table 6.4 identify the ranges of values and distribution for each risk variable for years 2029, 2033, and 2040, respectively. The "base run" values are presented for comparison purposes, but they are not directly used within the risk analysis.[16] More information on the research and methodology for developing the minimum, most likely, and maximum value can be found in the *2020 California High-Speed Rail Business Plan—Ridership and Revenue Risk Analysis Technical Supporting Document*.

[16] The "base run" is the revenue for the year and scenario forecast using the BPM-V3 model with the base input variable values.

## Table 6.2 Year 2029 Silicon Valley to Central Valley Risk Variable Ranges and Distributions

| Risk Variable | Base | Absolute Minimum | Most Likely | Absolute Maximum | Distribution |
|---|---|---|---|---|---|
| High-speed rail Constant | High-speed rail Calibrated Constant (Assumes Wait + Terminal Time = 25 min). | CVR bundled Constant + Assumed Wait + Terminal Time = 45 min. | High-speed rail Calibrated Constant (Assumes Wait + Terminal Time = 25 min). | High-speed rail Calibrated Constant + (HSR Constant—CVR Constant) + Assumed Wait + Terminal Time = 15 min. | Includes two components: Unexplained Variation and terminal and wait time.<br>Unexplained Variation: 50% Correlation between purposes; Distribution = Shape 4 PERT.<br>Terminal/Wait Time: 100% Correlation between purposes; Distribution = Triangle. |
| Business/Commute Trip Frequency Constant *(Annual business/commute round trips per person)* | 2.21 | 1.41 | 2.21 | 3.44 | Includes two components: Unexplained Variation and Economic Cycle.<br>Unexplained Variation: 50% Correlation between purposes; Distribution = Shape 4 PERT.<br>Economic Cycle: 100% Correlation between purpose; Distribution = Triangle. |
| Recreation/Other Trip Frequency Constant *(Annual recreation/other round trips per person)* | 5.86 | 4.83 | 5.86 | 7.12 | |
| Auto Operating Cost *($/mile in 2017 dollar)* | $0.25 | $0.20 | $0.25 | $0.34 | Distribution = Shape 5 PERT |
| High-speed Rail Fares *(Decimal Factor Difference from Base Fare)* | 1.0 | 0.74 | 1.0 | 1.42 | Distribution = Triangle |
| High-speed Rail Frequency of Service *(Roundtrips per day)* | 38 | 14 | 38 | 76 | Distribution = Triangle |

| Risk Variable | Base | Absolute Minimum | Most Likely | Absolute Maximum | Distribution |
|---|---|---|---|---|---|
| Availability and Frequency of Service of Conventional Rail and High-speed rail Buses that connect with High-speed rail | Scenario 3 = 2020 BP CVR Operating Plans; 66 Caltrains per day to 4th and King; w/ high-speed rail buses | Scenario 1 (5%) = 2017 CVR Operation level; 66 Caltrains per day to 4th and King; no high-speed rail buses. | Scenario 2 (40%) = ¾ CVR frequency between Base and 2017 CVR Operation level; SJQ local service provided in CV; 75% high-speed rail buses. | Scenario 3 (55%) [Same as Base] = 2020 BP CVR Operating Plans; w/ high-speed rail buses. | Distribution = multinomial. There are three 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. Note: The scenarios do not represent the minimum, most likely, and maximum values. |
| Coefficient on Transit Access-Egress Time/Auto Distance Variable | Calibrated Transit Penalty variable based on Air and CVR RP data. | Transit Penalty set to equal auto penalty based on International Experience. | Calibrated Transit Penalty variable based on Air and CVR RP data. | Calibrated Transit Penalty variable based on Air and CVR RP data. | Distribution = Shape 4 PERT. |
| Number and Distribution of Households throughout the State *(Total Households in Millions)* | Average of CEF, Moody's and MPO Forecasts. | Lowest available forecast for each county. | Blended. | Highest available forecast for each county. | Distribution = Triangular. |
| HSR Reliability (*Decimal Percent*) | 0.990 | 0.900 | 0.990 | 0.997 | Distribution = Shape 4 PERT. |
| Exceptionally Long Access and Egress (*Percent increase in disutility*) | 0% | 150% | N/A | 0% | Distribution = Uniform. Composed of a set of penalties that are added to the access/egress mode choice utilities after access/egress time reaches defined travel time thresholds for each access/egress mode. |
| Visitor Travel (*millions of HSR trips*) | 0 | 0.57 | N/A | 1.23 | Distribution = Uniform. 50% positive correlation with total California resident high-speed rail ridership. |
| Additional Induced Travel (*Percent of resident HSR ridership*) | 0% | 0% | 7.5% | 15% | Distribution = Triangular. 50% negative correlation with the trip frequency constant. |

## Table 6.3 Year 2033 Phase 1 Risk Variable Ranges and Distributions

| Risk Variable | Base | Absolute Minimum | Most Likely | Absolute Maximum | Distribution |
|---|---|---|---|---|---|
| High-speed rail Constant | High-speed rail Calibrated Constant (Assumes Wait + Terminal Time = 25 min). | CVR bundled Constant + Assumed Wait + Terminal Time = 45 min. | High-speed rail Calibrated Constant (Assumes Wait + Terminal Time = 25 min). | High-speed rail Calibrated Constant + (HSR Constant—CVR Constant) + Assumed Wait + Terminal Time = 15 min. | Includes two components: Unexplained Variation and terminal and wait time.<br>Unexplained Variation: 50% Correlation between purposes; Distribution = Shape 4 PERT.<br>Terminal/Wait Time: 100% Correlation between purpose; Distribution = Triangle. |
| Business/Commute Trip Frequency Constant *(Annual business/commute round trips per person)* | 2.28 | 1.46 | 2.28 | 3.54 | Includes two components: Unexplained Variation and Economic Cycle.<br>Unexplained Variation: 50% Correlation between purposes; Distribution = Shape 4 PERT.<br>Economic Cycle: 100% Correlation between purpose; Distribution = Triangle. |
| Recreation/Other Trip Frequency Constant *(Annual recreation/other round trips per person)* | 5.95 | 4.90 | 5.95 | 7.22 | |
| Auto Operating Cost *($/mile in 2017 dollars)* | $0.25 | $0.19 | $0.25 | $0.33 | Distribution = Shape 5 PERT. |
| High-speed Rail Fares *(Decimal Factor Difference from Base Fare)* | 1.0 | 0.74 | 1.0 | 1.42 | Distribution = Triangle. |
| High-Speed Rail Frequency of Service *(Roundtrips per day)* | 98 | 44 | 98 | 152 | Distribution = Triangle. |
| Airfares *(Decimal Factor Difference from Base Fare)* | 1.0 | 1.0 | 1.15 | 1.31 | Distribution = Triangle. |
| Coefficient on Transit Access-Egress Time/Auto Distance Variable | Calibrated Transit Penalty variable based on Air and CVR RP data. | Transit Penalty set to equal auto penalty based on International Experience. | Calibrated Transit Penalty variable based on Air and CVR RP data. | Calibrated Transit Penalty variable based on Air and CVR RP data. | Distribution = Shape 4 PERT. |

| Risk Variable | Base | Absolute Minimum | Most Likely | Absolute Maximum | Distribution |
|---|---|---|---|---|---|
| Number and Distribution of Households throughout the State *(Total Households in Millions)* | Average of CEF, Moody's and MPO Forecasts. | Lowest available forecast for each county. | Blended. | Highest available forecast for each county. | Distribution = Triangular. |
| HSR Reliability *(Decimal Percent)* | 0.990 | 0.900 | 0.990 | 0.997 | Distribution = Shape 4 PERT. |
| Exceptionally Long Access and Egress *(Percent increase in disutility)* | 0% | 150% | N/A | 0% | Distribution = Uniform. Composed of a set of penalties that are added to the access/egress mode choice utilities after access/egress time reaches defined travel time thresholds for each access/egress mode. |
| Visitor Travel *(millions of HSR trips)* | 0 | 1.67 | N/A | 3.60 | Distribution = Uniform. 50% positive correlation with total California resident high-speed rail ridership. |
| Additional Induced Travel *(Percent of resident HSR ridership)* | 0% | 0% | 7.5% | 15% | Distribution = Triangular. 50% negative correlation with the trip frequency constant. |

## Table 6.4 Year 2040 Phase 1 Risk Variable Ranges and Distributions

| Risk Variable | Base | Absolute Minimum | Most Likely | Absolute Maximum | Distribution |
|---|---|---|---|---|---|
| High-speed rail Constant | High-speed rail Calibrated Constant (Assumes Wait + Terminal Time = 25 min). | CVR bundled Constant + Assumed Wait + Terminal Time = 45 min. | High-speed rail Calibrated Constant (Assumes Wait + Terminal Time = 25 min). | High-speed rail Calibrated Constant + (HSR Constant—CVR Constant) + Assumed Wait + Terminal Time = 15 min. | Includes two components: Unexplained Variation and terminal and wait time.<br>Unexplained Variation: 50% Correlation between purposes; Distribution = Shape 4 PERT.<br>Terminal/Wait Time: 100% Correlation between purpose; Distribution = Triangle. |
| Business/Commute Trip Frequency Constant (Annual business/commute round trips per person) | 2.46 | 1.57 | 2.46 | 3.79 | Includes two components: Unexplained Variation and Economic Cycle.<br>Unexplained Variation: 50% |
| Recreation/Other Trip Frequency Constant (Annual recreation/other round trips per person) | 6.27 | 5.15 | 6.27 | 7.59 | Correlation between purposes; Distribution = Shape 4 PERT<br>Economic Cycle: 100% Correlation between purpose; Distribution = Triangle. |
| Auto Operating Cost ($/mile in 2015 dollars) | $0.24 | $0.19 | $0.24 | $0.32 | Composed of various components with PERT = Shape 5, uniform and triangular distributions. |
| High-speed Rail Fares (Decimal Factor Difference from Base Fare) | 1.0 | 0.74 | 1.0 | 1.42 | Distribution = Triangle. |
| High-Speed Rail Frequency of Service (Roundtrips per day) | 98 | 44 | 98 | 152 | Distribution = Triangle. |
| Airfares (Decimal Factor Difference from Base Fare) | 1.0 | 1.0 | 1.15 | 1.31 | Distribution = Triangle. |

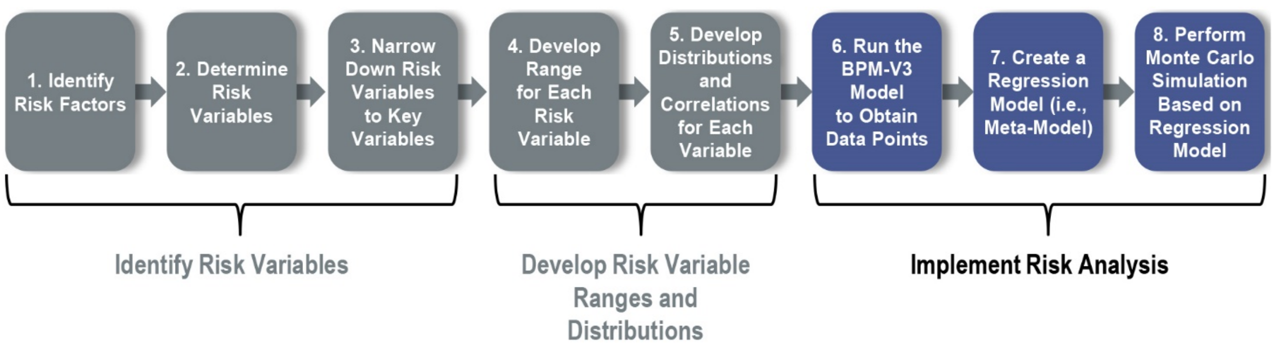
| Risk Variable | Base | Absolute Minimum | Most Likely | Absolute Maximum | Distribution |
|---|---|---|---|---|---|
| Coefficient on Transit Access-Egress Time/Auto Distance Variable | Calibrated Transit Penalty variable based on Air and CVR RP data. | Transit Penalty set to equal auto penalty based on International Experience. | Calibrated Transit Penalty variable based on Air and CVR RP data. | Calibrated Transit Penalty variable based on Air and CVR RP data. | Distribution = Shape 4 PERT. |
| Number and Distribution of Households throughout the State *(Total Households in Millions)* | Average of CEF, Moody's and MPO Forecasts. | Lowest available forecast for each county. | Blended. | Highest available forecast for each county. | Distribution = Triangular. |
| Auto In-Vehicle Time Coefficient (Decimal Factor from Base) | 1.0 | Alone = 0.5<br>Group = 0.8 | Alone = 0.75<br>Group = 0.90 | 1.0 | Distribution = Triangular. |
| HSR Reliability *(Decimal Percent)* | 0.990 | 0.900 | 0.990 | 0.997 | Distribution = Shape 4 PERT. |
| Exceptionally Long Access and Egress *(Percent increase in disutility)* | 0% | 150% | N/A | 0% | Distribution = Uniform.<br>Composed of a set of penalties that are added to the access/egress mode choice utilities after access/egress time reaches defined travel time thresholds for each access/egress mode. |
| Visitor Travel *(millions of HSR trips)* | 0 | 1.81 | N/A | 3.91 | Distribution = Uniform.<br>50% positive correlation with total California resident high-speed rail ridership. |
| Additional Induced Travel *(Percent of resident HSR ridership)* | 0% | 0% | 7.5% | 15% | Distribution = Triangular.<br>50% negative correlation with the trip frequency constant. |

## 6.4 Implementation of Risk Analysis

To fully understand the uncertainty in the high-speed rail forecasts of revenue and ridership, the full range of values for the risk variables was analyzed. To capture this full range, a Monte Carlo simulation of the BPM-V3 model is desired, but due to the BPM-V3's complexity, it is infeasible to run the model 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.[17] 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 developed meta-models replicate the results of the BPM-V3 model very well, indicating that the regression model forecasts of ridership and revenue match closely with the BPM-V3 forecasts of ridership and revenue given the same input values.[18]

As shown in Figure 6.4, 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 high-speed rail revenue or ridership. Each full BPM-V3 model run acts as one data point for use in estimating the regression models (*Step 7*). A Monte Carlo simulation, of 100,000 draws, is then run using the ridership and revenue regression meta-models and different combinations of values of the risk variables, with the values being drawn from the assigned risk variable distributions (*Step 8*). The revenue and ridership output from these runs is then used to develop the revenue and ridership range and probability of occurrence.

**Figure 6.4 Eight-Step Risk Analysis Approach: Implement Risk Analysis (Steps 6 to 8)**




### 6.4.1 BPM-V3 Model Runs

An experimental design for model runs lays out the number of model runs needed to support the risk analysis and the combination of risk variable values that compose each model run. For a complex model such as BPM-V3, it is important to design experiments to provide data to the risk analysis in an efficient manner, as the computational cost for each individual experiment is high.

[17] 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.

[18] The adjusted R-squared results from the linear regression models range from 0.976 to 0.983. The GPR cross-validation scores range from 0.747 to 0.988.

The BPM-V3 is a deterministic simulation model whose meta-models are best supported by a “space filling” design of experiments, such as Latin hypercube draws.[19] A Latin hypercube sample for one dimension is constructed by subdividing the distribution of each input factor into *N* equally probable ranges, and drawing one random sample within each range. For example, if an input factor is assumed to be uniformly distributed between 0 and 100, that distribution can be divided into four regions (0-25, 25-50, 50-75, and 75-100), and one random draw can be made in each region. This ensures better coverage of the entire input range than making 4 random draws from the full 0-100 space, which could result in a cluster of observations in one part of the space and a large void elsewhere. Generating a multi-dimensional Latin Hypercube sample for use with multiple input variables follows this same basic technique, although the various draws in each dimension are randomly reordered before being joined with draws in other dimensions, to avoid unintended correlation.

The Latin hypercube design does not demand any particular number of experiments. Adding a dimension without changing the number of experimental runs just marginally degrades the efficiency of the design. Conversely, increasing the number of experimental runs while holding the dimensionality of the problem constant can marginally improve the meta-model estimation. Practical experience across multiple domains has led to a “rule of thumb” that good results for prediction can be obtained from ten experimental data points per input variable dimension.[20] Using this “rule of thumb” at least ten BPM-V3 model runs were run for each risk variable for each model year and operating plan (150 model runs per forecast year).

Additional details of the development of the experimental design process are discussed in the *2020 California High-Speed Rail Business Plan—Ridership and Revenue Risk Analysis Technical Supporting Document*.

### 6.4.2 Final Revenue and Ridership Regression Models

The forecast revenues and ridership from the 150 BPM-V3 runs were used as data points for developing linear regression equations of the log of revenue as a function of the risk variables, used as the initial step in defining the meta-model for each forecast year. All models have r-squared values above 0.9, indicating that the linear regression model fits the BPM-V3 data points very well, and all of the signs and magnitudes of model coefficients are sensible.

The linear regression trend model provides an initial prediction for the (log of) revenue and ridership generated by the BPM-V3 model for each run. The difference between the linear regression prediction and the actual revenue observed for each run represents the residual, which is used as the dependent variable of a GPR model. The cross-validation results, ranging from 0.747 to 0.988, for the ridership and revenue GPR regression models indicates that the GPR provides a notable improvement in model fit above and beyond the linear regression model for each forecast. Additional details of the development of the ridership and revenue regression models are discussed in the *2020 California High-Speed Rail Business Plan—Ridership and Revenue Risk Analysis Technical Supporting Document*.

[19] Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical science*, 409-423.

[20] Loeppky, J., Sacks, J., & W.J., W. (2009). Choosing the sample size of a computer experiment: A practical guide. *Technometrics*, 366-376.

### 6.4.3 Revenue Results of the Monte Carlo Simulation

A Monte Carlo simulation using the regression meta-model outlined above was run 100,000 times using different combinations of values of the risk variables, with the values being drawn from the assigned risk variable distributions. It is important to note that some risk factors include multiple, sometimes correlated, 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 High-Speed Rail Mode Choice Constant risk variable. 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 100,000 Monte Carlo runs was used to develop the revenue range and probability of occurrence, as shown in Table 7.5. Revenue listed in the table does not include adjustments due to ramp-up. Short-distance trips of less than 50 miles in length within SCAG and MTC contribute approximately $3 million (2019 dollars) in revenue in year 2029, $11 million (2019 dollars) in revenue in year 2033 and $12 million (2019 dollars) in year 2040. This short-distance revenue was added to long-distance revenue for all probability levels to obtain total high-speed rail revenue.

The “base run” is the revenue for the year and scenario forecast using the BPM-V3 model with the base input variable values, as discussed in Section 5.

**Table 6.5 Year 2029–2040 High-Speed Rail Revenue Range and Probability of Occurrence[a] (2019 Dollars)**

| Probability | Revenue (Millions of 2019 Dollars)—2029 Silicon Valley to Central Valley | Revenue (Millions of 2019 Dollars)—2033 Phase 1 | Revenue (Millions of 2019 Dollars)—2040 Phase 1 |
|---|---|---|---|
| Minimum | $279 | $668 | $707 |
| 1% | $436 | $1,027 | $1,077 |
| 10% | $626 | $1,477 | $1,547 |
| 25% | $803 | $1,866 | $1,961 |
| Median | $1,066 | $2,422 | $2,558 |
| 75% | $1,396 | $3,098 | $3,273 |
| 90% | $1,732 | $3,755 | $3,987 |
| 99% | $2,339 | $4,917 | $5,299 |
| Maximum | $3,298 | $6,694 | $7,643 |
| **Base Run** | **$982** | **$2,207** | **$2,410** |

Source: Cambridge Systematics, Inc.

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

### 6.4.4 Ridership Results of the Monte Carlo Simulation

A Monte Carlo simulation using the ridership regression meta-model was applied to the same 100,000 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 6.6. Ridership listed in the table does not include adjustments due to ramp-up. Short-distance trips of less than 50 miles in length within SCAG and MTC contribute 0.16 million in ridership in year 2029, 0.44 million in ridership in year 2033 and 0.48 million in ridership in year 2040. This short-distance ridership was added to long-distance ridership for all probability levels to obtain total high-speed rail ridership. The "base run" is the ridership for the year and scenario forecast using the BPM-V3 model with the base input variable values, as discussed in Section 5. The percentages shown are where the original base ridership falls on the continuum of ridership forecasts produced by the various risk models.

**Table 6.6 Year 2029–2040 High-Speed Rail Ridership Range and Probability of Occurrence[a]**

| Probability | Ridership (Millions)—2029 Silicon Valley to Central Valley | Ridership (Millions)—2033 Phase 1 | Ridership (Millions)—2040 Phase 1 |
|---|---|---|---|
| Minimum | 3.7 | 8.6 | 8.9 |
| 1% | 6.5 | 14.6 | 15.2 |
| 10% | 9.6 | 21.1 | 22.2 |
| 25% | 12.4 | 27.0 | 28.3 |
| Median | 16.5 | 35.6 | 37.4 |
| 75% | 21.7 | 46.1 | 48.5 |
| 90% | 27.1 | 56.8 | 60.6 |
| 99% | 37.3 | 76.0 | 80.9 |
| Maximum | 53.4 | 109.8 | 118.1 |
| **Base Run** | **16.2** | **25.6** | **38.6** |

Source: Cambridge Systematics, Inc.

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

# Appendix A. High-Speed Rail Operating Plans

## A.1 Silicon Valley to Central Valley—2029

### *A.1.1 Dedicated Bus Connections—North*

Peak services (6 hours)

Bay Area Bus
(only planned to meet services to/from Bakersfield)

| Pattern # | BN-1 |
|---|---|
| Frequency | 60 |
| Run times from start in minutes | |
| Sacramento | 0 |
| Elk Grove | 20 |
| Lodi | 60 |
| Stockton | 85 |
| Modesto | 140 |
| Turlock/Denair | 165 |
| Merced | 200 |
| # of buses | 12 |

Off-peak services (12 hours)

Bay Area Bus
(only planned to meet services to/from Bakersfield)

| Pattern # | BN-1 |
|---|---|
| Frequency | 60 |
| Run times from start in minutes | |
| Sacramento | 0 |
| Elk Grove | 20 |
| Lodi | 60 |
| Stockton | 85 |
| Modesto | 140 |
| Turlock/Denair | 165 |
| Merced | 200 |
| # of buses | 22 |

### *A.1.2 High-Speed Rail Patterns*

Train

| Pattern # | V2V-02M | V2V-04M |
|---|---|---|
| Frequency | 60 | 30 |
| Run times from start in minutes | | |
| San Francisco | | 0 |
| Millbrae | | 17 |
| San Jose | | 39 |
| Gilroy | | 65 |
| Merced | 0 | \| |
| Madera | 12 | 99 |
| Fresno | 36 | 120 |
| Kings/Tulare | 58 | 142 |
| Bakersfield | 94 | 178 |
| # of trains | 12 | 24 |

Train

| Pattern # | V2V-02M | V2V-04M |
|---|---|---|
| Frequency | 60 | 60 |
| Run times from start in minutes | | |
| San Francisco | | 0 |
| Millbrae | | 17 |
| San Jose | | 39 |
| Gilroy | | 65 |
| Merced | 0 | \| |
| Madera | 12 | 99 |
| Fresno | 36 | 120 |
| Kings/Tulare | 58 | 142 |
| Bakersfield | 94 | 178 |
| # of trains | 22 | 19[a] |

Note:
[a] First northbound train starts at Fresno. Last southbound train terminates at Fresno.

### A.1.3 Dedicated Bus Connections—South

LA Basin Bus

| Pattern # | BS-1 | BS-2 | BS-3 |
|---|---|---|---|
| Frequency | 30 | 30 | 30 |
| Run times from start in minutes | | | |
| Bakersfield | 0 | 0 | 0 |
| Burbank Airport | 132 | \| | \| |
| LA Union Station | 160 | \| | \| |
| Van Nuys | | 140 | \| |
| West Los Angeles | | 160 | \| |
| Santa Anita | | | 160 |
| # of buses | 26 | 26 | 26 |

LA Basin Bus

| Pattern # | BS-1 | BS-2 | BS-3 |
|---|---|---|---|
| Frequency | 30 | 30 | 30 |
| Run times from start in minutes | | | |
| Bakersfield | 0 | 0 | 0 |
| Burbank Airport | 132 | \| | \| |
| LA Union Station | 160 | \| | \| |
| Van Nuys | | 140 | \| |
| West Los Angeles | | 160 | \| |
| Santa Anita | | | 160 |
| # of buses | 39 | 39 | 39 |

## A.2 Phase 1—2033

### A.2.1 Dedicated Bus Connections—North

Peak services (6 hours)

Bay Area Bus

| Pattern # | BN-1 |
|---|---|
| Frequency | 30 |
| Run times from start in minutes | |
| Sacramento | 0 |
| Elk Grove | 20 |
| Lodi | 60 |
| Stockton | 85 |
| Modesto | 140 |
| Turlock/Denair | 165 |
| Merced | 200 |
| # of buses | 24 |

Off-peak services (6 hours)

Bay Area Bus

| Pattern # | BN-1 |
|---|---|
| Frequency | 60 |
| Run times from start in minutes | |
| Sacramento | 0 |
| Elk Grove | 20 |
| Lodi | 60 |
| Stockton | 85 |
| Modesto | 140 |
| Turlock/Denair | 165 |
| Merced | 200 |
| # of buses | 18 |

## A.2.2 High-Speed Rail Patterns

| Pattern # | PH1-00 | PH1-02 | PH1-03 | PH1-04 | PH1-05 | PH1-06 | PH1-07 | PH1-08 | PH1-09 | PH1-10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Frequency | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60[a] | |
| Run time from start in minutes | | | | | | | | | | |
| San Francisco | 0 | 0 | 0 | | | 0 | | | | 0 |
| 4th and Townsend | \| | 4 | \| | | | 4 | | | | 4 |
| Millbrae | \| | 24 | \| | | | 24 | | | | 24 |
| San Jose | \| | 46 | 36 | 0 | 0 | 46 | | | 0 | 46 |
| Gilroy | \| | \| | \| | \| | 24 | 71 | | | 25 | 71 |
| Merced | \| | \| | \| | \| | \| | \| | 0 | 0 | 68 | 114 |
| Madera | \| | \| | \| | \| | \| | 106 | 15 | \| | | |
| Fresno | \| | \| | 93 | \| | \| | 127 | 36 | 24 | | |
| Kings/Tulare | \| | \| | \| | \| | \| | 150 | 59 | \| | | |
| Bakersfield | \| | \| | 138 | \| | \| | 185 | 94 | 69 | | |
| Palmdale | \| | \| | \| | \| | 125 | 225 | 134 | \| | | |
| Burbank Airport | \| | 175 | 191 | 128 | 152 | 252 | 161 | 122 | | |
| Union Station, LA | 174 | 193 | 209 | 146 | 170 | 270 | 179 | 140 | | |
| Gateway Cities, OC | | 239 | | | | 316 | 225 | 186 | | |
| Anaheim | | 251 | | | | 327 | 236 | 198 | | |
| # of trains | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 6[b] | 6[c] |

| Pattern # | PH1-01 | | PH1-03 | | | PH1-06 | PH1-07 | PH1-09 | PH1-10 |
|---|---|---|---|---|---|---|---|---|---|
| Frequency | 60 | | 60 | | | 60 | 60 | 60[a] | |
| Run time from start in minutes | | | | | | | | | |
| San Francisco | 0 | | 0 | | | 0 | | | 0 |
| 4th and Townsend | \| | | \| | | | 4 | | | 4 |
| Millbrae | \| | | \| | | | 24 | | | 24 |
| San Jose | 36 | | 36 | | | 46 | | 0 | 46 |
| Gilroy | \| | | \| | | | 71 | | 25 | 71 |
| Merced | \| | | \| | | | \| | 0 | 68 | 114 |
| Madera | \| | | \| | | | 106 | 15 | | |
| Fresno | \| | | 93 | | | 127 | 36 | | |
| Kings/Tulare | \| | | \| | | | 150 | 59 | | |
| Bakersfield | \| | | 138 | | | 185 | 94 | | |
| Palmdale | \| | | \| | | | 225 | 134 | | |
| Burbank Airport | 165 | | 191 | | | 252 | 161 | | |
| Union Station, LA | 183 | | 209 | | | 270 | 179 | | |
| Gateway Cities, OC | | | | | | 316 | 225 | | |
| Anaheim | | | | | | 327 | 236 | | |
| # of trains | 20 | | 20[d] | | | 19[e] | 18[f] | 9 | 14 |

Note:

[a] Only one of these two patterns runs in each hour in each direction.

[b] All run San Jose to Merced only.

[c] All run Merced to San Francisco only.

[d] First northbound train starts at Fresno; last southbound train terminates at Fresno.

[e] Last northbound train terminates at Fresno; fourth last southbound train terminates at LAUS, last 3 southbound trains terminate at Fresno.

[f] First northbound train starts at Fresno; last southbound train terminates at Los Angeles Union Station.

# A.3 Phase 1—2040

## A.3.1 Dedicated Bus Connections—North

Peak services (6 hours)

Bay Area Bus

| Pattern # | BN-1 |
|---:|:---:|
| Frequency | 30 |
| Run times from start in minutes | |
| Sacramento | 0 |
| Elk Grove | 20 |
| Lodi | 60 |
| Stockton | 85 |
| Modesto | 140 |
| Turlock/Denair | 165 |
| Merced | 200 |
| # of buses | 24 |

Off-peak services (6 hours)

Bay Area Bus

| Pattern # | BN-1 |
|---:|:---:|
| Frequency | 60 |
| Run times from start in minutes | |
| Sacramento | 0 |
| Elk Grove | 20 |
| Lodi | 60 |
| Stockton | 85 |
| Modesto | 140 |
| Turlock/Denair | 165 |
| Merced | 200 |
| # of buses | 18 |

## A.3.2 High-Speed Rail Patterns

| Pattern # | PH1-00 | PH1-02 | PH1-03 | PH1-04 | PH1-05 | PH1-06 | PH1-07 | PH1-08 | PH1-09 | PH1-10 |
|---:|:---:|:---:|:---:|:---:|:---:|:---:|:---:|:---:|:---:|:---:|
| Frequency | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60[a] | |
| Run time from start in minutes | | | | | | | | | | |
| San Francisco | 0 | 0 | 0 | | | 0 | | | | 0 |
| 4th and Townsend | \| | 4 | \| | | | 4 | | | | 4 |
| Millbrae | \| | 24 | \| | | | 24 | | | | 24 |
| San Jose | \| | 46 | 36 | 0 | 0 | 46 | | | 0 | 46 |
| Gilroy | \| | \| | \| | \| | 24 | 71 | | | 25 | 71 |
| Merced | \| | \| | \| | \| | \| | \| | 0 | 0 | 68 | 114 |
| Madera | \| | \| | \| | \| | \| | 106 | 15 | \| | | |
| Fresno | \| | \| | 93 | \| | \| | 127 | 36 | 24 | | |
| Kings/Tulare | \| | \| | \| | \| | \| | 150 | 59 | \| | | |
| Bakersfield | \| | \| | 138 | \| | \| | 185 | 94 | 69 | | |
| Palmdale | \| | \| | \| | \| | 125 | 225 | 134 | \| | | |
| Burbank Airport | \| | 175 | 191 | 128 | 152 | 252 | 161 | 122 | | |
| Union Station, LA | 174 | 193 | 209 | 146 | 170 | 270 | 179 | 140 | | |
| Gateway Cities, OC | | 239 | | | | 316 | 225 | 186 | | |
| Anaheim | | 251 | | | | 327 | 236 | 198 | | |
| # of trains | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 6[b] | 6[c] |

| Pattern # | PH1-01 | | PH1-03 | | | PH1-06 | PH1-07 | PH1-09 | PH1-10 |
|---|---|---|---|---|---|---|---|---|---|
| Frequency | 60 | | 60 | | | 60 | 60 | 60[a] | |
| Run time from start in minutes | | | | | | | | | |
| San Francisco | 0 | | 0 | | | 0 | | | 0 |
| 4th and Townsend | \| | | \| | | | 4 | | | 4 |
| Millbrae | \| | | \| | | | 24 | | | 24 |
| San Jose | 36 | | 36 | | | 46 | | 0 | 46 |
| Gilroy | \| | | \| | | | 71 | | 25 | 71 |
| Merced | \| | | \| | | | \| | 0 | 68 | 114 |
| Madera | \| | | \| | | | 106 | 15 | | |
| Fresno | \| | | 93 | | | 127 | 36 | | |
| Kings/Tulare | \| | | \| | | | 150 | 59 | | |
| Bakersfield | \| | | 138 | | | 185 | 94 | | |
| Palmdale | \| | | \| | | | 225 | 134 | | |
| Burbank Airport | 165 | | 191 | | | 252 | 161 | | |
| Union Station, LA | 183 | | 209 | | | 270 | 179 | | |
| Gateway Cities, OC | | | | | | 316 | 225 | | |
| Anaheim | | | | | | 327 | 236 | | |
| # of trains | 20 | | 20[d] | | | 19[e] | 18[f] | 9 | 14 |

Note:

[a] Only one of these two patterns runs in each hour in each direction.

[b] All run San Jose to Merced only.

[c] All run Merced to San Francisco only.

[d] First northbound train starts at Fresno; last southbound train terminates at Fresno.

[e] Last northbound train terminates at Fresno; fourth last southbound train terminates at LAUS, last 3 southbound trains terminate at Fresno.

[f] First northbound train starts at Fresno; last southbound train terminates at Los Angeles Union Station.