Risk Analysis Approach to Forecasting California High-Speed Rail Ridership and Revenue for Business Planning

presented to

2018 TRB Innovations in Travel Modeling

presented by

Rachel Copperman

CAMBRIDGE SYSTEMATICS

Think Forward

June 24, 2018

Importance of Modeling Uncertainty

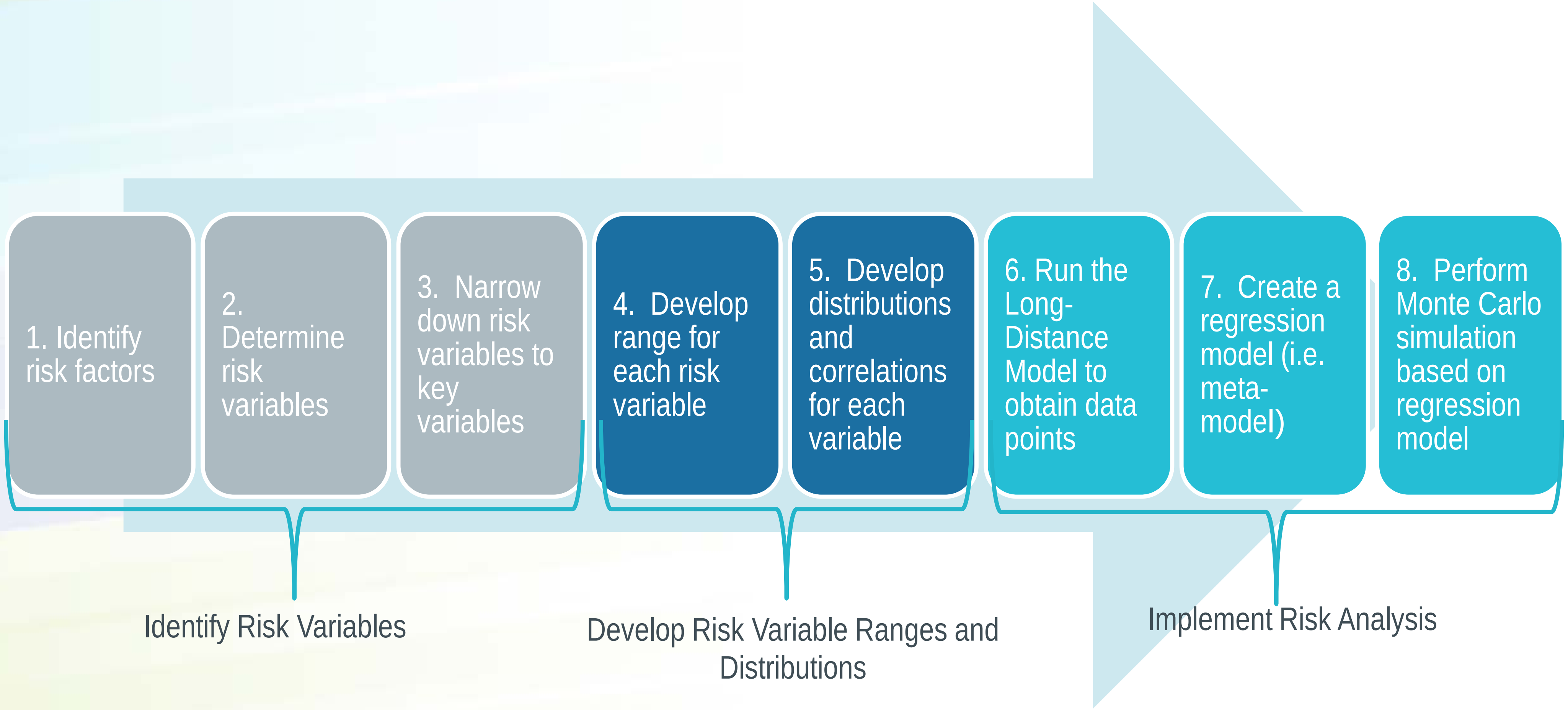
- Forecasting HSR ridership has inherent uncertainty:
  - » No HSR service currently exists
  - » Forecasts go out decades
- In that context, the range of possible outcomes (and their likelihood of occurring) becomes critical
- Understanding the uncertainty in the forecasts is important for:
  - » Analyzing the financial feasibility of different implementation strategies and other decisions
  - » Producing the required Breakeven Analysis for the Business Plan
  - » Analyzing maximum impacts for environmental clearance purposes
  - » Presenting a cohesive picture of potential ridership outcomes to decision-makers, stakeholders, and the public

High-Speed Rail System and Long-distance Model

- Long-distance Model:
  - Intra-California Trips > 50 miles
  - Trip-based model
- Modes:
  - Auto
  - Air
  - Conventional Rail
  - High-Speed Rail
- Two Phases considered in 2018 Business Plan:
  - San Francisco to Bakersfield (Year 2029)
  - San Francisco to Anaheim (Year 2033 and 2040)

Source: California High-Speed Rail Authority 2018 Business Plan

Eight-Step Risk Analysis Methodology



Step 1. Identify Possible Risk Factors
Step 2. Identify Variables for Each Risk Factor

- Step 1: Identify Possible Risk Factors
  - Circumstances, events, or influences affecting forecasts
  - Used panel of experts
  - Factors could vary by forecast year
- Step 2: Identify Variables for Each Risk Factor
  - Use actual variables and constants included in the Long-Distance Model
  - Variables represent one or more risk factors from Step 1

Risk Factors

- National fuel cost
- Fuel efficiency
- Vehicle maintenance costs
- Change in gas tax
- Impact of cap and trade on California fuel costs
- Market penetration of autonomous vehicles
- Share of shared-use vehicles

↓

Risk Variable

- Auto operating cost

Step 3. Narrow Key Variables by Forecast Year

- Long-Distance Model sensitivity runs performed for each risk variable
  - Quantitative comparison of impacts on ridership and revenue
- Final set of 17 risk variables selected (16 per forecast year)
  - Considered range and sensitivity of optional risk variables
  - Improvement in meta-model experimental design and enhanced automation of simulation has allowed for an increase from 2016 Business Plan of 10 risk variables

Sensitivity Tests (Sample)

- 100% increase in auto operating costs:
  - +28% HSR ridership
  - +30% HSR revenue
- 50% decrease in auto operating costs
  - -13% HSR ridership
  - -15% HSR revenue

Decision:

- Include Auto Operating Cost

Risk Variables Included in Risk Analysis

Year	Risk Variable
All Years (2029, 2033, 2040)	Business HSR Mode Choice Constant
	Commute HSR Mode Choice Constant
	Recreation/Other HSR Mode Choice Constant
	Business/Commute Trip Frequency Constant
	Recreation/Other Trip Frequency Constant
	Auto Operating Costs
	HSR Fares
	HSR Frequency of Service
	Coefficient on Transit Access-Egress Time/Auto Distance Variable
	Number and Distribution of Households throughout the State
	HSR Reliability
	Exceptionally Long Access and Egress
	Induced Travel
	Visitor Travel
Year 2029 Only	Availability and Frequency of Service of Conventional Rail and HSR Buses that Connect with HSR
Year 2033 and 2040	Airfares
Year 2040 Only	Auto In-Vehicle Travel Time Coefficient

Step 4. Develop a Range for Each Risk Variable

- For each risk variable / year:
  - minimum
  - most likely
  - maximum values
- Developed using available research and analysis

Research and Analysis (Auto Operating Cost)

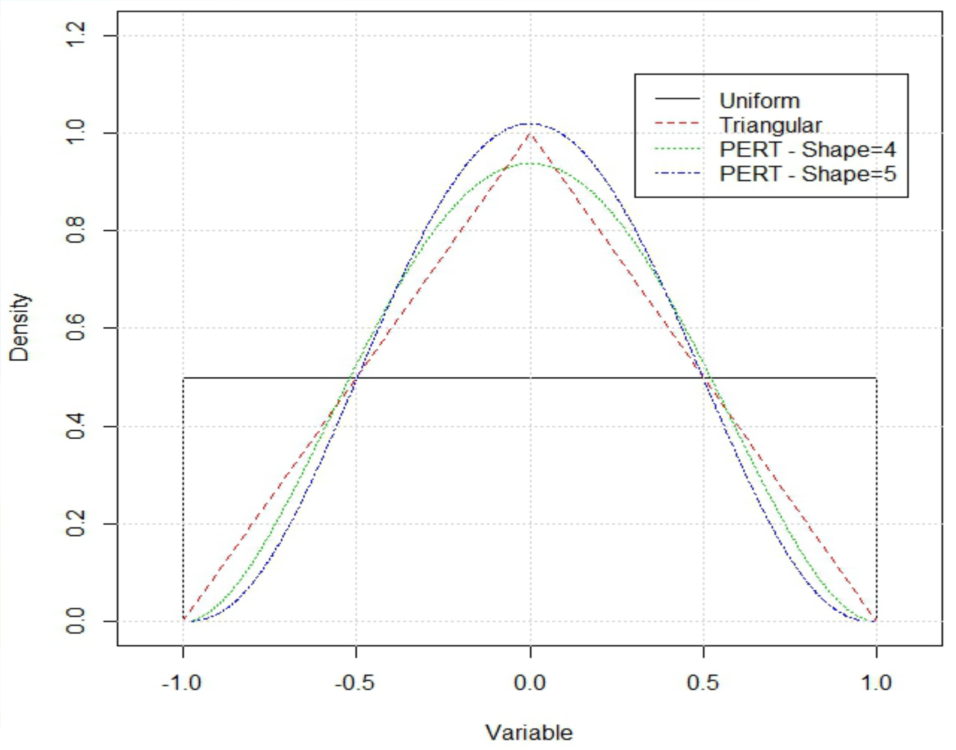
1. Projected retail fuel prices in California;
2. Adjust additional fees/charges based on two scenarios:
   - Cap and Trade, and
   - Potential increase in Federal excise tax;
3. Project fuel economy of "on the road" fleet;
4. Estimate nonfuel costs;
5. 2040 only, include effects of AV/CV and shared-ride vehicles

Range (Year 2033):

- Minimum: $0.17;
- Most Likely: $0.23;
- Maximum: $0.34

Step 5. Develop Distribution for Each Risk Variable

- Distribution shape for each variable determines the likelihood of the variable's value, within the set range, under random sampling



Likelihood of Range Values
- Minimum ($0.17): Unlikely
- Most Likely ($0.23): Very likely
- Maximum ($0.34): Unlikely

Distribution
- Most likely value chosen at a much higher rate than the extreme values
- PERT Distribution with Shape = 5

Step 6. Run the Long-Distance Model to Obtain Data Points

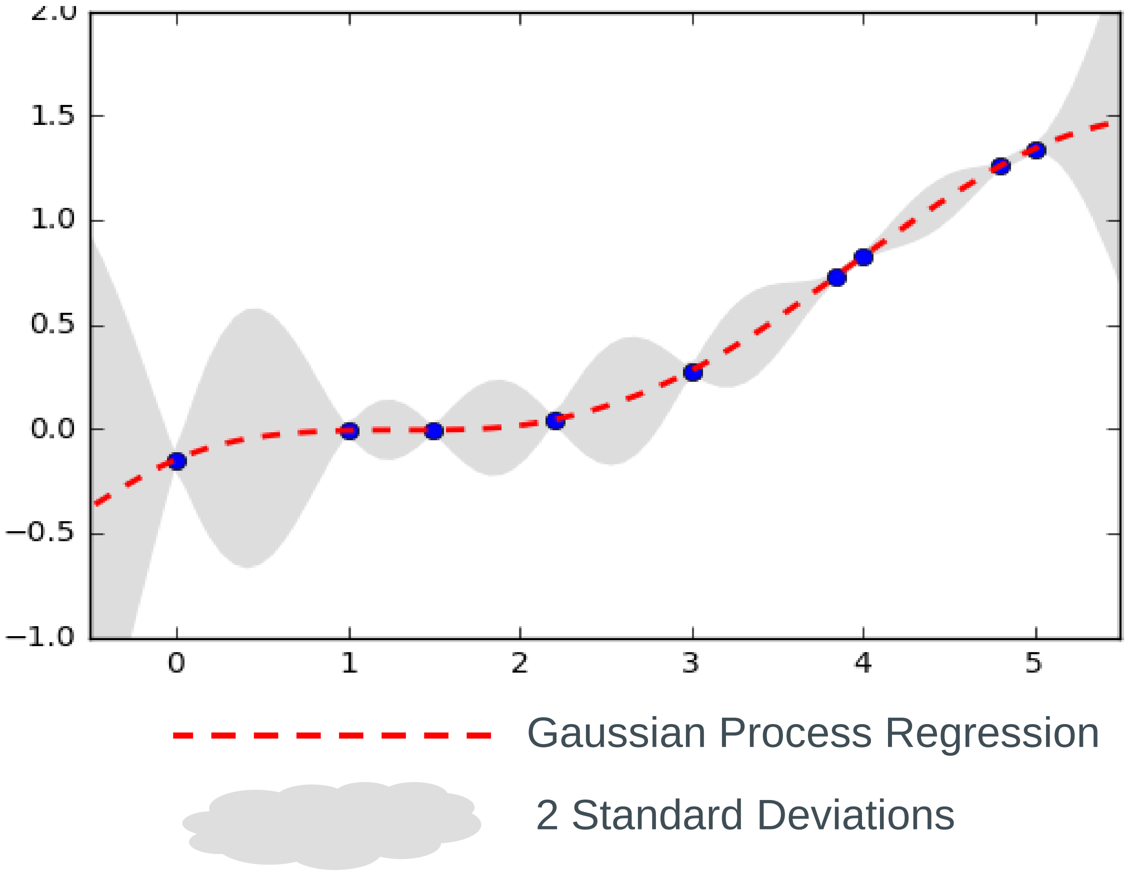
- Experimental design lays out the number of model runs, with varying levels of risk variable input values, needed to support building of the meta-model

- Latin-Hypercube Experimental Design
  - “Space-filling” to ensure maximum coverage and non-clustering around minimum and maximum
  - Allows for more risk variables to produce same number of data points (compared to fractional factorial design)
    - Rule of thumb: 10 data points per risk variable (~ 150 model runs per forecast year)

For more information attend Tuesday afternoon presentation:
Gaussian Process Regression for Risk Analysis of Travel Demand Forecasts

Step 7. Create Meta-Models

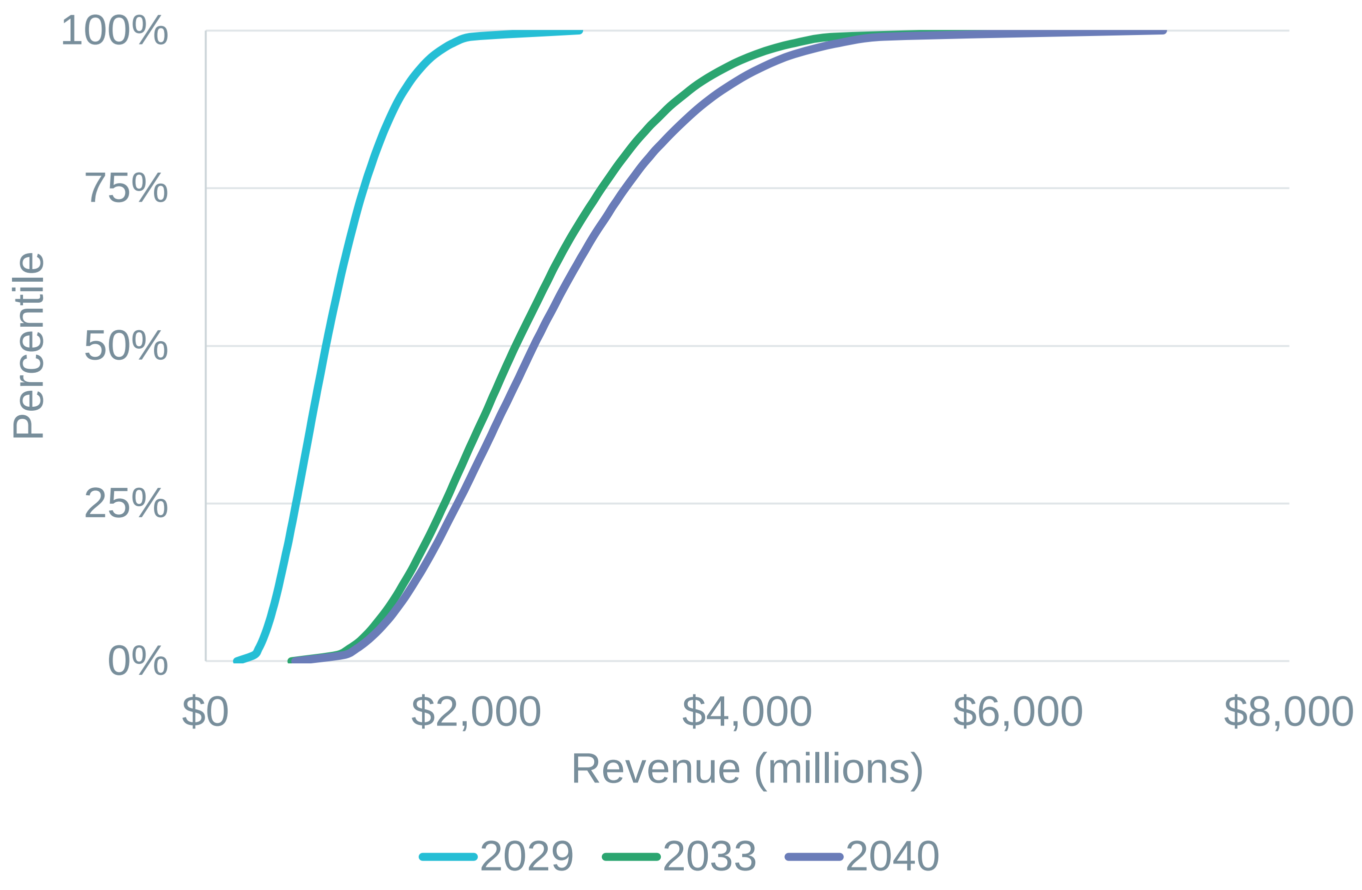
- Develop a main effects linear regression model from long-distance model runs
  - $\ln(\text{Revenue}) = \text{Constant} + \beta_1 \times \text{Var}_1 + \beta_2 \times \text{Var}_2 \ldots.$
- Apply a Gaussian Process regression
  - Non-parametric “machine learning” tool for regression analysis
  - Utilizes residuals from regression model as dependent variable in addition to full model runs



For more information attend Tuesday afternoon presentation:
Gaussian Process Regression for Risk Analysis of Travel Demand Forecasts

Step 8. Perform Monte Carlo Simulation Using Meta-Models

- 100,000 draws
  - » Risk variables drawn from respective distributions for each run
  - » Probability distributions for high-speed rail revenue is quantified

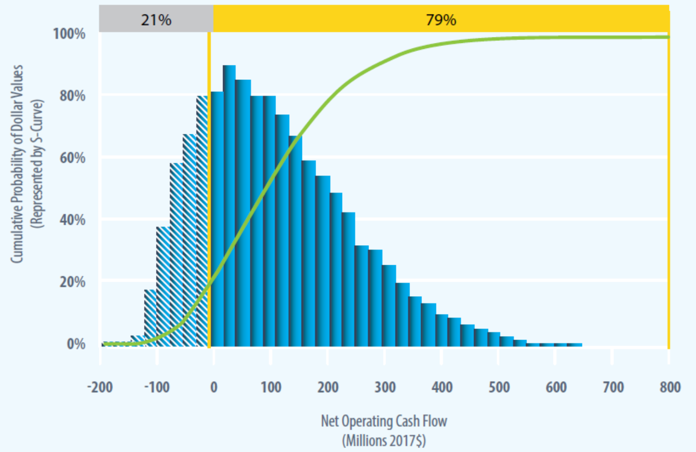


CAMBRIDGE SYSTEMATICS

Breakeven Analysis

EXHIBIT 7.12 BREAKEVEN ANALYSIS: OPENING YEAR SILICON VALLEY TO CENTRAL VALLEY (2029)
(IN MILLIONS OF $2017)

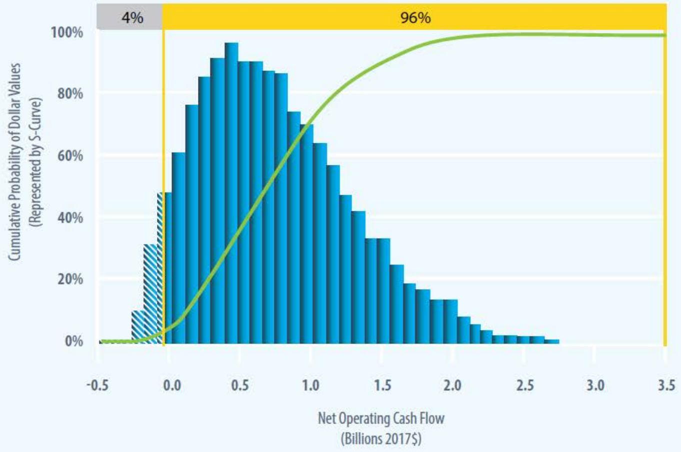
PROBABILITY DISTRIBUTION



DATA	
10%	($41)
25%	$15

EXHIBIT 7.13 BREAKEVEN ANALYSIS: OPENING YEAR PHASE 1 (2033)
(IN MILLIONS OF $2017)

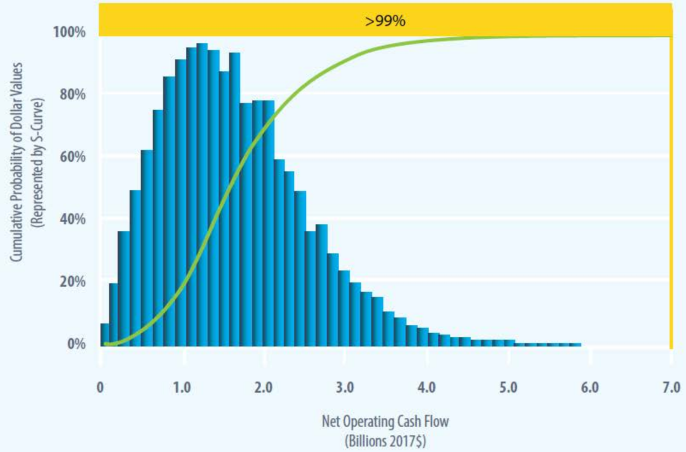
PROBABILITY DISTRIBUTION



DATA	
10%	$121
25%	$351

EXHIBIT 7.14 BREAKEVEN ANALYSIS: HORIZON YEAR PHASE 1 (2040)
(IN MILLIONS OF $2017)

PROBABILITY DISTRIBUTION



DATA	
10%	$662
25%	$1,065
Median	$1,636
75%	$2,321
90%	$2,998

Source: California High-Speed Rail Authority 2018 Business Plan

Resources and Acknowledgements

- E-mail: rcopperman@camsys.com

- California High-Speed Rail Authority 2018 Business Plan Ridership and Revenue Risk Analysis technical supporting document
  - » http://www.hsr.ca.gov/docs/about/business_plans/2018_CA_High_Speed_Rail_Business_Plan_Ridership_and_Revenue_Risk_Analysis.pdf

- California High-Speed Rail Authority 2018 Business Plan
  - » http://www.hsr.ca.gov/docs/about/business_plans/2018_BusinessPlan.pdf

- Team Members:
  - » David Kurth, Jeff Newman, Zeina Wafa, Moby Khan, Jason Lemp, *Cambridge Systematics*
  - » John Helsel, Matt Henley, *WSP*
  - » Boris Lipkin, *California High-Speed Rail Authority*