CAMBRIDGE SYSTEMATICS

Response to the Review of *Bay Area/California High-Speed Rail Ridership and Revenue Forecasting Study*

David Brownstone, Mark Hansen, and Samer Madanat
Institute of Transportation Studies, University of California

final report

prepared for

California High Speed Rail Authority Board

prepared by

Dr. Lance Neumann, President and
Dr. Kimon Proussaloglou, Senior Principal
Cambridge Systematics, Inc.

July 8, 2010

www.camsys.con

BACKGROUND

Cambridge Systematics, Inc. (CS) was founded in 1972. Throughout the history of the firm, two basic principles have guided all of CS' work: 1) employing state-of-the-practice analytical skills; and 2) providing objective, unbiased results. CS has become an industry leader in the profession by developing a reputation for objective, state-of-the-practice travel demand forecasting. Today, CS employs the largest travel demand forecasting staff of any firm in the country. Principals associated with CS have pioneered many of the most significant advances in the profession. In the United States, CS has developed travel demand forecasting tools for over 16 states and over 30 metropolitan areas. CS has performed high-speed rail forecasting work both in the U.S. and abroad.

CS' national reputation in the field is such that the U.S. Department of Transportation has selected CS to develop much of the travel modeling guidance it has produced, including the *Travel Survey Manual*, the *Model Validation Guide*, and training courses on advanced travel forecasting techniques and the link between transportation and land use. This guidance material is actively being used by states, metropolitan planning organizations, and other firms across the country to inform travel demand modeling practices. Similarly, CS staff are actively engaged in research and theoretical considerations related to travel forecasting. They are active in every committee of the Transportation Research Board (TRB) that relates to travel forecasting theory and practice. A principal of Cambridge Systematics currently chairs the transportation demand forecasting committee of the TRB, which is part of the National Academies.

GENERAL APPROACH

Developing practical travel demand models requires an attention not only to theoretical and academic considerations, but also to real-world practical concerns such as schedule, budget, and policy-driven considerations. Therefore, our approach to developing models is first to understand the specific policy issues and decision-making context facing our clients. This understanding allows us to develop the most appropriate tool and compile the best possible information to help our client make those decisions. During the development of every travel model, modeling theory must be balanced with practical considerations. Both science and judgment are involved in every model development effort, and each model development effort can always use more data and additional model refinements. The challenge for the practical model developer is to design a workable, policy-driven tool within budget and schedule constraints that provides credible information upon which the client can base its decisions. The model should not be overcomplicated, nor should it have superfluous features just for the sake of modeling elegance.

EXPERT DEVELOPMENT TEAM

While under contract to the Metropolitan Transportation Commission (MTC), CS assembled an absolutely outstanding expert modeling team to support the work of the California High-Speed Rail Authority (Authority). This team included CS staff and consultants who are internationally and nationally recognized for their skill and experience in travel forecasting. It is a team with extensive high-speed rail forecasting experience in Europe, Australia, and North America. In addition, an independent peer review panel was established to provide guidance from additional national experts in travel forecasting and many of the most accomplished travel demand modeling practitioners in the State of California. The peer review panel offered another source of expertise to review key assumptions and to help the project team make sound methodological judgments. As a result of the work of this expert team, the ridership and revenue model does represent the state-of-the-art in travel forecasting.

The California High-Speed Rail Ridership and Revenue Forecasting model is a complex system of dozens of interrelated, state-of-the-art model components that span different geographies, different trip purposes, and different travel market segments. The model reflects an appropriate blend of theory and judgment, which is always required in real-world applications of travel forecasting models. The model produces realistic results that are sensitive to the key input variables. It is a model that CS stands behind, and it is an appropriate tool to support the Authority. CS welcomed the review from ITS and respects the reviewers. However, CS disagrees with the conclusions of the review.

RESPONSE TO ITS REPORT

Over the last two-and-a-half months, CS carefully reviewed a total of 30 questions that were posed by the ITS team, and provided clear and detailed responses to each question. In these responses, CS outlined the thinking involved in key decisions and explained how difficult problems were resolved during the estimation, calibration, validation, and application phases of model development. Of the 30 questions, 7 issues were discussed in the final ITS report. Although there is much common ground between ITS and CS on most of the original questions, significant disagreements remain, particularly when it comes to some of the ITS report's broad conclusions about data representativeness, estimation biases, and policy sensitivity. This document provides direct responses to these seven remaining issues.

1. Division into Short and Long Trips

The ITS team's first criticism is that "the sharp delineations between different trip categories seem arbitrary," an issue which they suggest could create discontinuity for trips around 100 miles in length. Segmenting the market into short and long trips is common modeling practice, since it is widely recognized that travel behavior differs by trip distance. Market segmentation recognizes the different levels of service offered by competing modes of travel, and allows a model to predict a traveler's choices differently based on the length of their desired trip. A trip distance of 100 miles was used as a convention to delineate long and short trips, since this distance generally reflects most long-trip versus short-trip market segments in California. Furthermore, the 100-mile breakpoint is also consistent with other data collection efforts such as the American Travel Survey and the National Household Travel Survey. Finally, even if the 100-mile benchmark were abandoned for some other reasonable benchmark, it would have a minimal impact on systemwide high-speed rail ridership for any analyzed alternative.

2. Assigning All Business Travel to Peak Periods

The ITS report states that assigning all business or commuting trips to peak-hour conditions is "potentially a serious problem." CS believes that the adopted approach is very reasonable for modeling interregional travel. Indeed, assigning business and commute trips to peak periods merely reflects that the vast majority of business travel occurs during the peak period; therefore, the overwhelming majority of intercity business and commute travelers face peak-period service levels.

The ITS reviewers acknowledge that assigning such trips to the peak period is standard practice in urban models, but state that this practice is incorrect for interregional travel since "quite a few business trips are made in the off-peak." As evidence of this claim, the ITS review posits that 25 percent of business and commute interregional trips in California take place outside of peak periods. However, the National Household Travel Survey indicates that this percentage is actually higher (40 percent) for urban area travelers. The actual similarity between the urban and interregional travel patterns conflicts with ITS' claims, and suggests that assigning business and commute travelers to peak-period service conditions is appropriate for both urban and interregional travelers.

3. Treatment of Panel Dataset

The ITS report postulates that "there is likely unobserved serial correlation" since "each respondent provided responses to several hypothetical choice situations." ITS claims that such correlation might have two potential implications: 1) that the statistical significance of the parameters is overstated; and/or 2) that the value of the parameters themselves might be affected. The report does not point out that only four choice experiments were asked of respondents – so although there is a slight possibility that statistical significance of a few coefficients in one model may have been overstated, the impact is very likely to be quite small. More importantly, there is no impact on the parameter values or the relative importance of factors that affect traveler choice behavior. As the ITS review states, under certain econometric assumptions, the "parameter estimates are still consistent." Therefore, the relative importance of policy-sensitive factors appears to be properly reflected in the estimated models.

4. Constraining the Headway Coefficient

The ITS report mentioned that the headway coefficients in the mode choice models were constrained, stating that "the modelers' expectation would be reasonable if this [change] was

[for] an urban travel demand model, but it is incorrect in the present context." In the model, coefficients for access times, wait times, terminal times, and egress times for each mode reflect observed current experiences and proposed service levels. However, the coefficient for headways – the frequency of service – was constrained, and as a result reflects the unique case of high-speed trains that offer far more frequent interregional service than is currently available on conventional rail services such as Amtrak.

The ITS report notes that the relative importance of the headway coefficient was constrained during model calibration, prior to performing any forecasts. When faced with different sources of conflicting data, modelers are compelled to use professional judgment. This is not done by imposing one's own beliefs, but rather by investigating the differences between the datasets. In an effort to better replicate existing travel patterns and create a policy sensitive model, the headway coefficient that reflects the convenience of a schedule was constrained. The adjustment made to the headway coefficient was within the range of reasonable values presented to peer review during the model development process.

5. Absence of Airport/Station Model

The ITS reviewers claimed that the assignment method used to allocate travelers from each geographic zone to rail stations and airports in the region was "behaviorally unrealistic [since], depending on their desired travel schedule, access/egress modes and other factors, travelers may choose different stations." Under the adopted modeling process, all travelers between the same origin and destination point and traveling for the same purpose are assigned to a single rail station and airport. Hence, travelers from the same area are allowed to choose different rail stations or airports depending upon their trip purpose and travel destination.

The ITS reviewers prefer a more elegant station choice model in which travelers between the same origin-destination point could be assigned to two or more rail stations or airports. However, for the purposes of high-speed rail forecasting, this difference would only be relevant to a small portion of the population served. Therefore, the influence of such a change on total ridership results would be minimal.

Depending on where a traveler lives, the station or airport that he/she is likely to choose depends on the level of service to access the airport or station, as well as the level of service that exists and is offered at each airport or station. The assignment method used by CS is behaviorally sound, since it considers these very same factors. To test the potential impact of the ITS panel's recommendation, CS examined the size of the travel markets affected by the CS method versus the approach proposed by the ITS team. CS' analysis suggests that ITS' proposed approach might influence ridership projections in a very small portion of the Bay Area between roughly Atherton and Sunnyvale. In essence, a more elaborate station choice model would simply take the riders in a zone and split them in a percentage manner to two or more stations rather than assigning all riders in a zone to the same station. CS found the potential impact of such a procedure to be less than one percent of the total estimated high-speed rail ridership.

As the ITS report states, using ITS' proposed method "would almost certainly reduce, although probably not eliminate, the ridership difference between the Pacheco and Altamont alignments found in the CS study." In reality, ITS' suggestion to create a more elaborate station choice model would be an interesting exercise, but one that would not ultimately change the model's outcome in a meaningful way.

6. Calibration of Constants

The ITS report made two interrelated claims regarding the calibration of constants in the CS model: 1) that key market segments were oversampled; and 2) that the panel disagreed with the method used to reflect the true travel market shares.

CS agrees with the ITS that key market segments were oversampled. However, this oversampling was in fact a requirement for reliable model estimation. Unfortunately, a purely random sample of the population is neither a realistic nor a cost-effective approach for a study of interregional travel and high-speed rail ridership. A random sample of all travelers would collect information mostly from automobile users, thereby underreporting data from other key segments of the population, such as current rail riders and air passengers. To alleviate this issue, model developers are often forced to issue supplemental surveys which necessarily overrepresent air and rail riders to gain statistically significant sample sizes. Then, proven methods are used to estimate and calibrate models that reflect the true market shares in the population.

At the outset of developing the ridership and revenue model, a 2000/2001 Caltrans travel survey of more than 17,000 randomly sampled California households was viewed as a key data source for model estimation. However, in these 17,000 households, a mere 25 interregional trips were made by air passengers and rail riders combined. Any modeler would agree that to estimate reliable mode choice models with so few observations from two critical markets segments would be impossible. To address this lack of data for key market segments, the sample was enriched by a new data collection effort. Approximately 3,000 new stated-preference surveys were collected reflecting travel by auto, rail, and air. These new observations were collected using a proven technique known as "choice-based sampling". Instead of randomly calling respondents, surveys were conducted on trains and at airports by randomly intercepting these travelers. These surveys were used to enrich the larger random sample by including more statistically significant response rates from segments of key interest to the project at hand.

However, since more observations were collected from rail riders and air passengers than their share of the interregional travel market, an adjustment had to be made once the models were estimated. The adjustment process is called a "calibration of mode constants". By calibrating mode constants, travel market shares are adjusted to reflect the true market shares in the population. CS' primary area of disagreement with ITS concerns the method it used for calibration. CS employed an academic method that has been proven a long time ago, has been used widely and consistently to calibrate models, and is well established in literature and in practice. The ITS reviewers maintain that this method has been superseded by new, unproven academic research that was published in 2008, after the ridership and revenue model was developed.

This is a technical, academic debate – and a classic example of the creative tension that often exists between theory and practice. On the one hand, a well established and proven econometric method was used by CS to calibrate the air and rail constants. On the other, ITS noted new emerging research which, though based on promising theory, has not been tested or used in practice.

In summary, a large randomly sampled survey data set was enriched using a supplemental survey to meet project objectives, and to reflect and quantify the decisions made by rail riders and air passengers. In addition, CS used the most tested and best available approach to calibrate the model to be more representative of the population. CS believes these methods were, and continue to be, both sound and appropriate.

7. Constraining of Coefficients

Finally, the ITS report suggests an "excessive" constraining of coefficients in the final models, and states that "restrictions were based on professional judgment *instead of* on observed data." CS respectfully disagrees on both counts.

During the calibration of any such model, the constants for modes, airports, and geographic locations are typically constrained to ensure that the resulting forecasts reflect observed travel patterns. This constraining of constants is a matter of accepted practice. In addition, CS constrained the parameters for a very small number of explanatory variables out of the total. For example, out of 24 total explanatory variables used as model parameters in the long-distance mode choice models, only three of them were constrained. One, the headway coefficient, has already been discussed. CS considers this level of constraining to be very modest given a model system of this size and complexity.

In the development and application of practical travel demand models, it is often the case that various sources of data need to be reconciled with different or conflicting empirical evidence from the model estimation. In these cases, it is absolutely necessary to use analysts' judgment to reconcile different data and arrive at the most practical model possible. The decision to constrain certain coefficients was made neither unilaterally nor arbitrarily, but was based on the best available data, published literature, and accepted practice. These judgments were further scrutinized by peer review during the model development process. Finally, all adjustments made to match existing travel patterns for auto, air, and rail travel were made before the model was ever applied to forecast HSR ridership.

CONCLUSION

In summary, as the ITS reviewers note, the CS approach to model development uses widely accepted methods and professional standards reflecting the theory and practice of model estimation, validation, and application. However, the core of the ITS and CS differences of opinion on these seven modeling issues appear to reflect a creative tension between academic theory and the practical tradeoffs that must be made in real-world model applications.

As in every debate, there remain notable areas of disagreement. Most importantly CS disagrees with the broad conclusions ITS has made regarding the presence of bias and the model's reliability. These conclusions have led to further misinterpretations of how the data were collected, how the model works, and how the model is applied. In practice, the model's validity is not compromised by the academic econometric issues that were raised in the ITS review. The model is policy-sensitive. It allows the Authority to address questions related to alignments and to levels of service. Its sensitivity to a range of different policies has been tested successfully. This sensitivity is the best proof of a carefully developed and calibrated model. It ensures that the California High-Speed Rail Ridership and Revenue Forecasting model has and will continue to provide the Authority with valuable information in the planning decision-making process.

Ridership and Revenue Forecasting
Response to Review

presented to

California High-Speed Rail Authority

presented by

Cambridge Systematics, Inc.
Lance Neumann, Ph.D.
President
Kimon Proussaloglou, Ph.D.
Principal

July 8, 2010

CAMBRIDGE SYSTEMATICS Transportation leadership you can trust.

Cambridge Systematics
A Legacy of Excellence and Trust

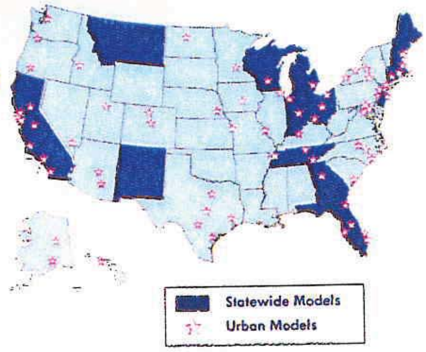
- Founded in 1972, independent and employee-owned
- Full-service transportation planning firm with real-world experience
- Hundreds of clients worldwide
  - » Depth of analytic skills
  - » Objectivity
- Research and practical applications

CAMBRIDGE SYSTEMATICS

2

Cambridge Systematics
Travel Demand Forecasting Leadership

- **Over 35 years of national and international experience**
- **Largest travel demand forecasting firm in the U.S. (50+ staff)**
- **Pioneered many of the most significant advances in the travel demand forecasting profession**
- **Practical worldwide experience**
  - » 16 statewide & 30+ urban models
  - » High-speed rail models in both the U.S. and abroad



CAMBRIDGE SYSTEMATICS

3

Cambridge Systematics
Model Development and Application Leadership

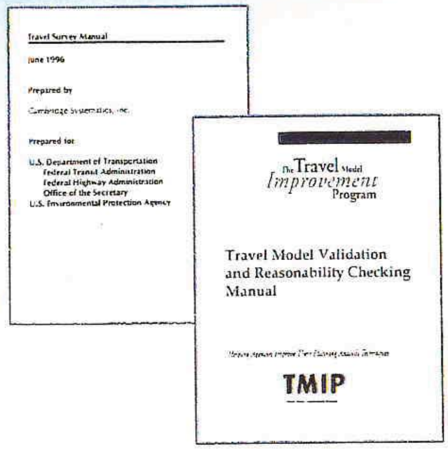
- **We *do***
  - » Consider the specific policy and decision-making context in determining the appropriate modeling approach
  - » Explain the necessary balance among model theory, practicality, complexity, and cost to our clients
  - » Ensure that the modeling approach is consistent with an agency's schedule and resource constraints while meeting appropriate professionals standards
- **We *do not***
  - » Assume a "one size fits all" theoretical approach is right for each ridership forecasting problem
  - » Include unneeded features that would adversely affect model performance and cost

CAMBRIDGE SYSTEMATICS

4

Cambridge Systematics
Training and Guidance Leadership

- **Travel Survey Manual (U.S. DOT)**
- **Model Validation Guide (U.S. DOT)**
- **Advanced Travel Demand Forecasting course (U.S. DOT)**
- **Transportation and Land Use course (U.S. DOT)**


Travel Survey Manual

June 1996

Prepared by

Cambridge Systematics, Inc.

Prepared for

U.S. Department of Transportation
Federal Transit Administration
Federal Highway Administration
Office of the Secretary
U.S. Environmental Protection Agency

The Travel Model Improvement Program

Travel Model Validation and Reasonability Checking Manual

TMIP

CAMBRIDGE SYSTEMATICS

5

Cambridge Systematics
Transportation Research Leadership

- **Transportation Research Board (National Academy of Sciences)**
  - » **Transportation Demand Forecasting**
  - » **Travel Survey Methods**
  - » **Travel Behavior and Values**
  - » **Travel Analysis Methods**
  - » **Statewide Transportation Data and Information Systems**
  - » **Intercity Passenger Rail**

CAMBRIDGE SYSTEMATICS

6

Cambridge Systematics
California High-Speed Rail Ridership and Revenue Forecasting

- **Expert Model Development Team**
  - Assembled internationally recognized team that has developed high-speed rail forecasts in Europe, Australia and the U.S.
  - Convened an independent peer review panel of academic and practitioner experts
  - Client project manager, Chuck Purvis, is a recognized national leader
- **Ridership and Revenue Model**
  - State-of-the-art
  - Appropriate blend of theory and judgment
  - Realistic, proven sensitivities to key inputs
- **Confident the model is the right tool to support the Authority**

CAMBRIDGE SYSTEMATICS

7

Response to ITS Review
Overview

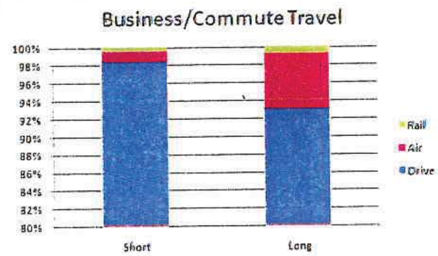
- **Initial review generated 30 questions**
- **Issues discussed in the final report**
  - Division into short and long trips
  - Assigning all business travel to peak period
  - Treatment of panel dataset
  - Constraining the headway coefficient
  - Absence of an airport/station choice model
  - Calibration of constants in mode choice models
  - Constraining of coefficients
- **A complex system of models**
- **Data, models, calibration, and sensitivity**

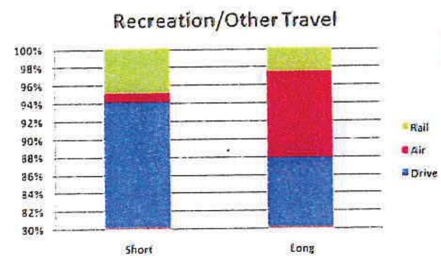
CAMBRIDGE SYSTEMATICS

8

Issue 1:
Division into Short and Long Trips

- Market segmentation
- Travel behavior by distance
- 100 miles as a cutoff point
- Consistent with nationwide FHWA surveys
- Reflection of market segments and traveler tradeoffs





CAMBRIDGE SYSTEMATICS

Issue 2:
Assigning All Business Travel to Peak Period

- Majority of business travel occurs during the peak
- Similar patterns in urban and interregional travel
- Model properly reflects
  - » Total market size
  - » Size of work and nonwork market segments
  - » Service and costs during peak and off-peak periods

CAMBRIDGE SYSTEMATICS

10

Issue 3: Treatment of Panel Dataset

- Two questions in the ITS review
  - The relative values of the policy sensitive parameters
  - The statistical significance of the estimated parameters
- Relative importance is key to policy-sensitive models
  - Parameters are consistent and free of bias
  - Relative importance of parameters is correct

CAMBRIDGE SYSTEMATICS

11

Issue 4: Constraining the Headway Coefficient

- Components of out-of-vehicle time
  - Access time, Wait time, Terminal time, and Egress time
  - Schedule convenience: Headway component
- High speed rail: a different paradigm of service frequency
  - Headways are shorter than best commuter rail operations
  - Headway coefficient within range discussed with peer panel
- Reasonable value leading to a policy-sensitive model

CAMBRIDGE SYSTEMATICS

Issue 5: Absence of Airport/Station Model

- CS method considers station and airport choice
  - » Access and level of service by station/airport
  - » Same station/airport is assigned to all travelers in the same zone
- A model assigns travelers to 2^{+} airports/stations
- Magnitude of impact is estimated at less than 1%

CAMBRIDGE SYSTEMATICS

Issue 6: Calibration of Constants

- Two inter-related questions
- Data: Represent all travel modes
  - » Oversampling key segments
  - » Requirement for reliable model estimation
- Method: Reflects true shares in population
  - » Calibration of mode constants
  - » Adjustment for oversampling by mode

CAMBRIDGE SYSTEMATICS

Issue 6: Calibration of Constants - Data

- Data options examined at outset of project
- A random sample for the study
  - » Caltrans household survey (N=17,000 households)
  - » A minimal sample size for air and rail riders (N=25)
- Enriched sampling
  - » New revealed and stated preference surveys
  - » 3,000 surveys with 1,500 auto users
  - » On-board and airport terminal surveys
  - » Data used to develop reliable choice models

CAMBRIDGE SYSTEMATICS

Issue 6: Calibration of Constants - Methodology

- Need to correct back to true population market shares
  - Method: Calibrating mode constants
  - Reflects true market shares in population
- What is the source of disagreement?
  - Proven method was used to calibrate models
    - ♦ Well established in literature and in practice
  - New academic research from 2008
    - ♦ Method not widely used in practice
- Data are enriched to meet project objectives
- Model is representative of the population

CAMBRIDGE SYSTEMATICS

Issue 7: Constraining of Coefficients

- Model calibration to match observed travel
  - Adjustments to mode and airport constants
  - Constraints only on few explanatory variables
- Empirical evidence was used extensively
  - Decisions made to reflect base-year results
  - Reconciling of different sets of data sources
  - Published literature and accepted practice
- Limited constraining of explanatory variables
- No impact on model validity

CAMBRIDGE SYSTEMATICS

Summary

- Creative tension
  - Academic approach vs. real-world application
  - We "followed generally accepted professional standards in carrying out the demand modeling and analysis"
- We disagree with other broad conclusions
  - Data reflect travel among California residents
  - Model validity is not compromised by econometric issues
  - A policy-sensitive model addresses planning questions
  - Model sensitivity has been proven in 3+ years of application
- We fully stand behind the CAHSRA travel demand model

CAMBRIDGE SYSTEMATICS