

# Gamed Traffic Data Endangers High-Speed Rail Project

by Richard F. Tolmach

President, California Rail Foundation

February 8, 2010

Ridership claims of the California High-speed Rail Authority (HSRA) have strained the credulity of transportation industry observers for the past decade. However, until recently the public was unaware of any substantial problem with the data because supporting detail was largely hidden in technical supplements. Massive total figures were cited by the Authority, but breakdowns were not readily forthcoming. This changed recently because of legislative demands that the HSRA planning process become more transparent.

Mid-December publication of the 2009 High-Speed Rail Business Plan enlivened the ridership debate. For the first time in a number of years, HSRA disclosed specific projected inter-regional boardings by station. It provided some additional data not included in the 2008 Business Plan, although it did not go so far as to provide city-pair data, the ordinary metric of ridership analysis.

Changes from the prior Charles River Associates Independent Ridership Study, prepared for HSRA in January 2000 were extreme and very inconsistent.

The new figures from the Cambridge Systematics model prepared for the Bay Area's Metropolitan Transportation Commission (MTC) for the HSRA showed huge increases in San Francisco, Gilroy, Merced, Bakersfield, Palmdale, and Anaheim ridership combined with precipitous declines in Peninsula and Los Angeles boardings. (*figure at right*)

Gilroy and Bakersfield basically doubled while Merced rose by 895 percent from the Charles River Associates estimate. Palmdale and Anaheim, which were not even served by the 2000 network, suddenly became the 2 top Southern California stations for interregional (>100 mi.) ridership.

The details in the Business Plan led activists and politicians to become concerned, because the newly released boardings by station had immediately visible anomalies. In an attempt to understand the new numbers, CRF and the Peninsula Cities Consortium (PCC), representing Atherton, Burlingame, Belmont, Menlo Park and Palo Alto, constructed various comparisons of the data.

PCC produced a comparison with national Amtrak data showing that the HSRA projections for San Francisco, Anaheim, Los Angeles and Palmdale exceeded current New York City Amtrak ridership.

CRF produced visual comparisons of HSRA's projected annual per capita interregional rides (*next page*) showing unbelievable growth in projections for Gilroy, Merced, and Anaheim. These analyses helped make the model problem obvious, but left unanswered the question of how model projections went so wrong.

Average Daily Inter-Regional Boardings			
	2000 Charles River Associates Full Network for 2020	2009 Business Plan Starter Network for 2035	Data Change 2000 to 2009
Sacramento	9,866	—	
Stockton	1,488	—	
Modesto	1,785	—	
SF Transbay Term.	10,730	20,600	+92%
Millbrae	1,760	900	-49%
Redwood	3,140	2,400	-24%
San Jose	6,843	4,700	-31%
Gilroy	2,063	3,800	+84%
Los Banos	212	—	
Merced	563	5,600	+895%
Fresno	3,406	4,700	+38%
Bakersfield	2,545	5,300	+108%
Palmdale	—	5,500	
Santa Clarita	3,118	—	
Sylmar	—	3,300	
Burbank	4,062	800	-80%
Los Angeles	12,324	3,800	-69%
E San Gabriel Val.	6,636	—	
Norwalk	—	3,000	
Anaheim	—	19,100	
Ontario	1,404	—	
Riverside	2,626	—	
Temecula	1,594	—	
Escondido	2,510	—	
Mira Mesa	1,255	—	
San Diego	7,457	—	

## 2000 Charles River Associates 2020 Full Network Interregional Boardings Per Capita (Based on 2008 DOF Population)



## 2009 HSRA Business Plan 2035 Starter Network Interregional Boardings Per Capita (Based on 2008 DOF Population)



### New Plan and New Model

A diminished network, but increased ridership, including huge gains in per capita riders in unlikely cities.

The prior model had believable results in all markets except Gilroy. The new model fails the plausibility test because of its extreme results in Los Angeles, Burbank, Palmdale, Merced, Gilroy and Millbrae.

In January, Californians Advocating for Responsible Rail Design (CARRD), a Peninsula group with econometrics expertise, alerted CRF about problems in the sampling methods used in the surveys. These included using current California rail passengers for a disproportionate percentage of its sample in surveys to determine attitudes toward using high speed rail, which introduced obvious bias into the model.

However, sampling error could not explain other troubling data anomalies, such as the huge per capita ridership seen above in Merced, Gilroy, and Anaheim. To try to solve the riddle, CARRD asked for additional data including the final coefficients and constants used in the HSR Ridership and Revenue Model.

### Ridership Mystery Solved

CARRD received the final coefficients and constants in late January from Deputy Director Jeff Barker of HSRA, along with a surprise—a transmittal memo from George Mazur of Cambridge Systematics. This transmittal places direct blame on MTC for withholding these documents from the public for the past 33 months: “The client, MTC, elected not to update the Task 5a report nor to include the final coefficients and constants in the final project report.” This is a remarkable assertion for Cambridge Systematics, assigning responsibility for what may be fraud.

The final coefficients and constants are substantially changed from those peer reviewed and published, and the revised elements were never seen by the peer review panel. Jeff Barker of HSRA said in his transmittal e-mail that “... this material as presented did not previously exist and significant amounts of subconsultant staff time went into preparing it.” Until CARRD received the material, no member of the public nor any oversight agency had seen the technical underpinnings of the final ridership model. Without peer-reviewed documentation that a model accurately replicates current travel behavior, it is useless as a predictor of future travel.

Doug Kimsey, the MTC’s planning director, was quoted February 7 in the Palo Alto Daily News confirming that there were unpublished changes to the model. “My understanding is that the modifications were minor enough they didn’t need to be peer reviewed. ... Minor modifications to the evaluation tool don’t necessarily have to be peer-reviewed.”

## Is a 6000 percent change minor?

One notable covert change made commute and business travel sensitivity to departure frequency 6000 percent higher than the value reported to the peer group and the public. The peer-reviewed model figured an extra 30 minutes between departures would cut traffic by the same factor as 5 extra minutes of travel time. The revised model made that **5 extra hours**. In the resulting model, frequency clearly trumps travel time.

This change appears designed to tip the scales toward rail, which has speed challenges due to route detours added since 2000 but plenty of frequency. There are 99 total San Francisco-Los Angeles trains each direction, but only 6 achieve HSRA's claimed 2 hour 40 minute travel time. 74 trains stop five or more times and take an average of 3 hours 13 minutes. Only by tweaking frequency sensitivity do these trains become air-competitive.

The size of the frequency sensitivity change erases the obvious travel time advantages of air travel in distant markets such as Orange County and San Diego, and counteracts penalties for rail detours or transfers. The traffic increases at Anaheim, Merced, Gilroy and San Francisco seen on the prior page appear to be one result.

The disclosure also explains how Altamont's population and travel time advantages over the Pacheco Route were negated. Tweaked frequency sensitivity is a plausible explanation why the model showed ridiculously circuitous Sacramento-Chowchilla-San Francisco trains producing more revenues than direct Altamont trains.

The 60-times change is not a minor distortion, and would undoubtedly have caused the peer review panel to reject the model. Proper coefficients are calculations with substantiating data. The new frequency coefficient lacks any supporting evidence, documentation, or even informal explanation. This lack of documentation makes it hard to accept that the changes are any sort of honest corrections.

## **"Sledgehammer" Adjustment Signals Tampering with Model Results**

There are additional clear signs that the revised model is fatally flawed. Constants specific to air, high-speed rail and conventional rail are now very large relative to other parameters. This is a *de facto* admission that the model cannot accurately replicate passenger choice based on objective factors such as travel time and costs, so as one economist commented, "a sledgehammer is being used to bang them into place."

For travel demand specialists, use of large specific modal constants is an enormous red flag, because by their very nature they suggest bias and tampering. Models are supposed to work without big modal constants.

Because the peer review committee was kept in the dark and MTC published preliminary coefficients and constants which looked relatively reasonable, the gaming of the model workings was concealed from all. The true authorship of the numbers is an open question, but it appears that Cambridge Systematics wants to distance itself from all work that may have happened after April, 2007. Mazur's transmittal states that "There have been no changes to these model elements since April 2007." This raises more questions about authorship of the August 2007 revisions that lowered Altamont's traffic below Pacheco's, as well as subsequent changes.

Who is responsible for any revisions since April 2007 is still a matter of controversy. What is not controversial is that the final work is irrevocably flawed and needs to be redone from scratch.

## **Conclusions**

Transportation modeling consultant, Norm Marshall of Smart Mobility based in Norwich, Vermont, is preparing additional conclusions but has stated to us:

"I conclude that the final coefficients and constants introduce unacceptable biases into the model, and that the model as presented in the January 29, 2010 memo is invalid for forecasting future HSR ridership and revenue."

Now that HSRA has disclosed its use of this invalid model, CRF finds the following significant implications:

1. Applications for Federal ARRA funding were based on questionable ridership estimates.
2. Each Program and Project Level EIR which relied upon findings of the model calls for facilities which may be unjustified. Some of the recent design augmentations are quite significant, such as the requirements for 10-train hourly San Francisco Transbay Terminal and Los Angeles Union Station capacity, and quadruple tracking of San Francisco-San Jose and Los Angeles-Anaheim segments requiring use of eminent domain.
3. Claims that the High-Speed Rail Authority's project will not require public subsidy are no longer credible.
4. Without valid ridership and fiscal projections, the hundreds of millions of dollars spent on engineering and EIRs may have been wasted.