Memo to the Ridership Peer Review Panel **September 2011**

Prepared by Elizabeth Alexis, CARRD

We have a number of concerns about the existing ridership model. The peer review panel is addressing many of them like the discontinuities with urban, short and long trips. There are other items that we believe warrant further discussion.

The overall specific model behaviors that we are concerned about are:

1) The high rate of diversion from trips currently taken by car. The diversion from automobiles is almost as high as that from airplanes for SF – LA trips.
2) The high sensitivity to frequency at a normal level of headways.
3) The insensitivity of the model to access and egress issues.
4) The lack of sensitivity to significant socio-economic differences that exist between regions in California.
5) The treatment of longer distance commuters as high-end business travelers.
6) The lack of induced travel.
7) The presumption of high rates of population growth.

While some of these stem from the challenges of using a stated preference data set that oversampled those most likely to take a train, others come from the calibration phase and some are simply related to model inputs.

We believe that at a minimum the model should be re-estimated. Our preference given the importance of the model and the challenges in forecasting a new service in such a large and diverse state would be for the Authority to solicit proposals as to the best approach to the model and have the Peer Review panel help assess the relative merits.

We understand the time pressures involved with this project but we feel that the issues that have been identified with the model have a significant potential impact on forecasts for the Initial Operating Segment. We would also note that while the Authority has taken the stance that a higher forecast is more conservative for the purposes of environmental review, this is currently unsubstantiated.

Below we discuss certain specific issues in more detail.

Market definition

The CS model had an unusually broad definition of the market for high speed rail. It included all trips within the State of California[1]. This includes trips that are not served by HSR. For example, a trip from Eureka to Sacramento would be counted as part of the market.

An analysis comparing an adjusted number of trips[2] used in a previous ridership forecast by Charles River Associates which used a more traditional definition of market (from one county or region served by high speed rail to another also served by high speed rail) suggests that **50% or more** of the trips in the CS model would not be served by HSR.

This means that the baseline number of trips used to compute data like mode shares is inflated relative to other similar calculations.

It is very difficult to interpret the data in Figure 3 regarding the mode shares by trip distance from the 2011 CS memo given that HSR may not be a realistic option for the majority of trips.

For example, Table 4 gives the mode shares for business trips from "SCAG to MTC Interchange," better known as Los Angeles to the Bay Area. This trip is in the 375 mile range. While the mode shares vary somewhat by income, the auto share ranges from 10% to less than 5%. Figure 3, on the other hand, shows an auto share of about 40% for such trips.

The data in Tables 8, 9a and 9b from the same memo that gives mode shares by distance from station is somewhat better but still problematic. It limits the analysis to trips that start near a station but would presumably still include trips that do not end near another station.

For these types of analyses, it might be useful to limit the trips analyzed to those that had a viable HSR path. For short trips, CS limited HSR as an option to those with both access and egress shorter than 25 miles. For longer trips, this was limited to those with access less than 100 miles and the access and egress times each less than the train travel time.

This still includes many trips that seem unlikely to attract HSR passengers, but would be more helpful in comparing to the results to other systems.

[1] Except those to Lake Tahoe.
[2] The CRA analysis did not include commute trips and had data for 2000, 2015 and 2020.

Sensitivity to price changes

It is not a surprise that the model shows overall sensitivity to price changes. The calibration process that was undertaken endeavored not only to get reasonable replication of travel mode behavior but also achieve demand sensitivity.

Tables 5 and 6 in CS memo

The statewide averages for business and commute trips seem implausible given that only two small regions have trips that are lower than the average.

Differences between regions in California

The change in the frequency coefficient was motivated by a desire to dampen demand for air travel from certain airports, presumably those in the Central Valley as there is currently is a low level of service for intra-California flights from Central Valley cities.

The model used a more brute force method to further lower demand for air service.

The final model used large constants to represent the attributes of specific airport markets. For most of the Bay Area to Los Angeles area markets,these were large positive numbers,typically equivalent to $294 business / $143 non-business. They helped offset a large negative mode specific coefficient for air service (-$604 business/-$144 non-business).

There were no such constants for travel from any of the Central Valley airports, which clearly helped limit demand for air travel.

There are several problems with this approach. First, presumably the same types of factors that limit demand for air travel would apply in some manner to high speed rail, which is a close enough substitute for air that a nested model structure is used. There are however no similar dampening mechanisms for high speed rail demand from the Central Valley.

Second, a better approach would be to directly incorporate socio-economic data that differentiates Central Valley cities from others in California. Travel demand in general and high speed rail travel demand in particular are closely linked to income levels, educational attainment and certain types of employment. While the peer review group has focused on the importance of income, we would suggest using additional measures as the differences between regions are stark. Many of these metrics are analyzed in detail in this sobering report:
http://www.measureofamerica.org/california/

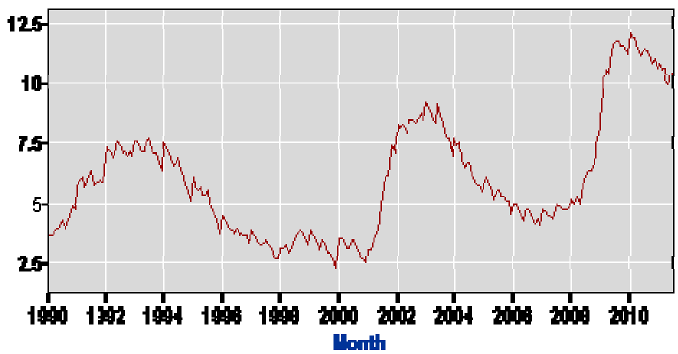
We have included a number of data points below.

1. Income and educational attainment levels[3]. Renfe (the Spanish HSR operator) indicated in a June 2011 presentation to CHSRA that 68% of their customers are college educated. The SP survey collected educational data.
2. Local economy data. We have both unemployment statistics and a breakdown of employment by the types of industries that have an affinity for high speed rail. These illustrate the large differences in California's regional economies. We do not have any data on hotel rooms but that would also be useful as a measure of tourism.

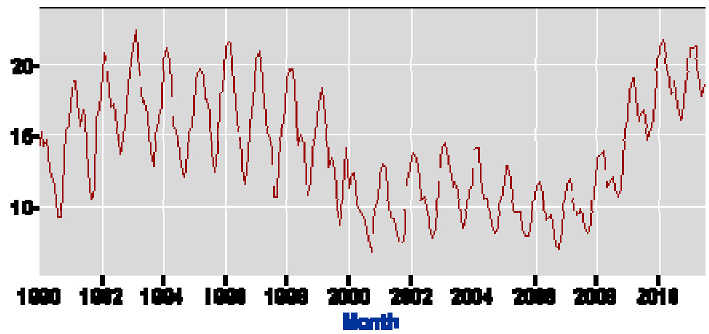
	Educational attainment	Household Income	Per capitaincome
San Francisco	**51%**	**$86,546**	**$44,573**
Palo Alto	**79%**	**$119,483**	**$68,944**
San Jose	**36%**	**$78,660**	**$33,165**
Gilroy	**23%**	**$67,317**	**$28,192**
Merced County	**13%**	**$43,848**	**$17,994**
Fresno County	**19%**	**$46,230**	**$20,375**
Kings County	**12%**	**$45,595**	**$17,416**
Bakersfield	**20%**	**$51,886**	**$22,601**
Palmdale	**15%**	**$54,840**	**$19,231**
Burbank	**34%**	**$62,255**	**$32,885**
Los Angeles	**29%**	**$48,570**	**$27,070**
Anaheim	**22%**	**$57,870**	**$22,522**
Irvine	**64%**	**$92,195**	**$42,255**

[3] Educational attainment is the percentage of adults over 25 with a college degree or higher.

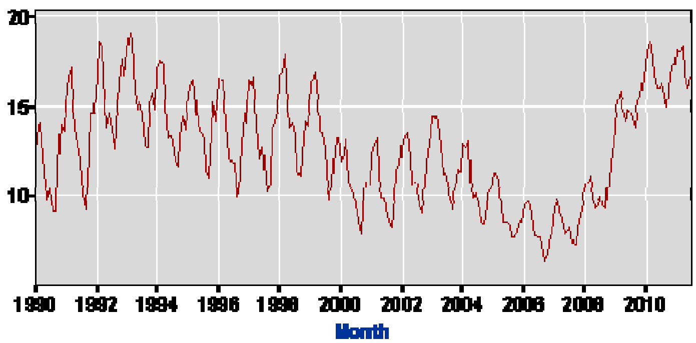
A. **San Jose** unemployment rate



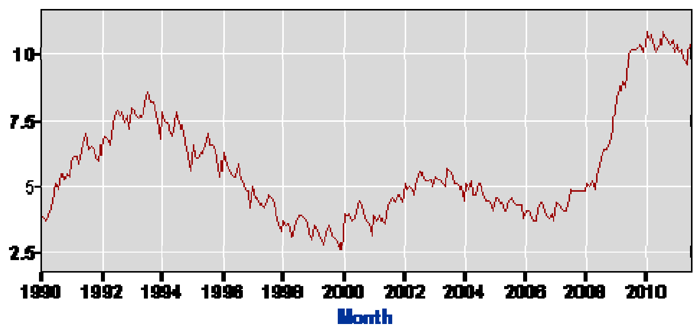
B. **Merced** unemployment rate



C. **Fresno** unemployment rate



D. **San Diego** unemployment rate



Source: http://data.bls.gov/cgi-bin/surveymost?la+06

Employment by sector

	Information	Financial Activities	Professional and Business Services	Total	Civilian Labor Force	% of labor force
San Francisco - Redwood City	**38.5**	**75.8**	**201.5**	**315.8**	**946.4**	**33%**
San Jose - Sunnyvale	**48.1**	**30.8**	**164.9**	**243.8**	**892.7**	**27%**
Merced	**1.3**	**1.6**	**3.7**	**6.6**	**106.6**	**6%**
Fresno	**3.7**	**13.2**	**25.8**	**42.7**	**434.9**	**10%**
Bakersfield	**2.8**	**7.9**	**23.5**	**34.2**	**364.2**	**9%**
LA-Long Beach - Santa Ana	**230.3**	**309.7**	**777.8**	**1317.8**	**6420.8**	**21%**

Source: May 2011 BLS http://www.bls.gov/eag/eag.ca.htm

Access and egress

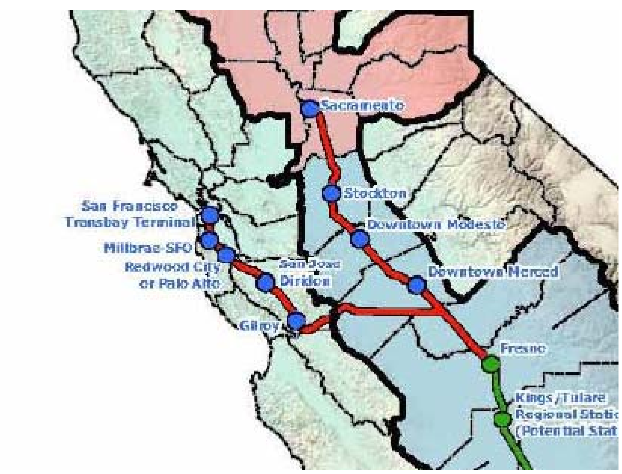
We have significant concerns that the model underestimates the important of train access. The previous model by Charles River Associates (available at http://cahighspeedrail.ca.gov/WorkArea/DownloadAsset.aspx?id=6588) only considered trips that were made by people who lived in the vicinity of a station and who were traveling to some place near another station.

The CS model includes all trips taken in the state of California (with the exception of those to Lake Tahoe) although the original SP sample specifically only included people who took a trip from one region served by high speed rail to another region served by high speed. The vast majority of the people participating in the survey lived quite close to a station. For example, residents of Monterey were excluded from the survey. Residents of half of California's counties were excluded from the survey.

It is a well-known attribute of high speed rail that ridership demand is closely tied to the accessibility of stations. It is even more sensitive than air service. "Beet root " stations often have disappointing ridership. Recent work by Mark Hansen and Reinhard Clever highlights the paramount importance of station location.[4]

In many ways, the model shows little of that sensitivity.

It can be seen in the extreme "stub" station effect wrt Merced and Anaheim.



In Phase 1, Merced will be the northern most station on the branch to Sacramento.The ridership forecast for Merced is **7,370** passengers per day. In phase 2, stations are scheduled to open in Modesto (40 miles away) and Stockton (70 miles away) and service will be extended to Sacramento and San Diego. Ridership for Merced is scheduled to drop to **1,558** passengers per day. The explanation given is that the passengers who will then use the Modesto (3,671 passengers) and Stockton (5,064 passengers) stations would have driven up to 100 miles on congested Highway 99 to Merced in Phase 1 to catch the train.

[4] "Interaction of Air and High-Speed Rail in Japan" Reinhard Clever and Mark M. Hansen*Transportation Research Record*, **2043**, 1-12 (2008)More at http://thinkmetric.com/diss/

The following assumptions were made in determining the "best path" high speed rail skims:

- All available local transit options in the system as well as auto and walk canbe used to access and egress high-speed rail stations;
- The maximum distance for auto access to high-speed is assumed to be 100miles (the same distance as air);
- High-speed rail will not be used if the auto access or auto egress time is greater than in-vehicle time;
- High-speed rail will not be used if the "best path" origin station anddestination station are the same;
- High-speed rail will not be used if the distance between the origin TAZ and destination TAZ is less than 100 miles and the access or egress distance is greater than 25 miles; and
- High-speed rail will not be used if the drive access plus drive egress time is greater than the time it would take to drive from the origin TAZ to the destination TAZ.

A similar phenomenon occurs with Anaheim.

The data that CS presented in the 2011 memo regarding mode share by distance from station clearly demonstrates the lack of sensitivity.

The model accounted for accessibility in two ways. First there was a link to calculated log sums from the access and egress (long trips only) mode choice models. These seem to have a limited role in the main mode choice.

Second, the high speed rail networks had explicit limits placed on them[5] [see side panel for excerpt]. These appear to be binding constraints in many cases .

For short trips, both auto access and egress were limited to 25 miles. In addition, the access and egress time each had to be less than the train trip time. This produces the more reasonable looking access data in Figure 7a and table 8. However, for trips that are less than 100 miles door-to-door, one would expect even more clustering of demand around stations.

For long trips, the constraints are less stringent yet still appear to have more influence on mode choice than the access and egress models.

CS is quite open about the limitations of the access/ egress components of the model. The problem is that the engineering consultants are using results from the model to help determine station location and even the number of stations. In addition, the center of information and financial firms is not always in the downtown, making the station decision more challenging.

The lack of sensitivity will be an issue with ridership forecasts for the Initial Operating Segment, where both ends will necessarily be "stub stations."

Frequency/ headway coefficient

"Frequency does not have a significant effect either for air or HSR."[6]

[5]Bay Area/California High-Speed Rail Ridership and Revenue Forecasting Study, Statewide Model Networks August 2007 page 5-14

As the Panel has noted, the current headway coefficient is inappropriate. But as the panel has also observed, the current model does not show extreme sensitivity to headways as measured by elasticities. This is not surprising, given the incredibly low absolute level of headways in the current operating schedule.

A 25% change in headways for long distance intercity travel with headways of 4 to 15 minutes means only 1 – 4 minutes change. This cannot be compared to the change in headways where hourly service is more the norm and a 25% change in headways would mean a 15 minute change. While at some ranges of headway the demand response may be logarithmic wrt changes in frequency, it is unlikely to be so at the extreme low levels of headways in the operating schedule.

The frequency elasticity would be more relevant if the model was re-run with the headways that were in the original SP survey of 30 minutes to 2 hours.

Frequency is a complicated variable. The sensitivity depends on the length of trip. For long trips with long headways, the demand response may be more logarithmic than linear. For short trips with moderate headways, the demand response may be more exponential than linear.

In the case of the California High Speed Rail model, the topic is especially fraught.

First, the use of the high frequency coefficient is being used to dampen demand for air service from the Central Valley and masks the real issues as to why there is low air (and likely HSR train) demand between certain markets.The low frequency of service between certain cities is not exogenous;it is in part a function of low demand that correlates with many excluded variables that would help explain differences in demand due to the differences in regional economies and other socio-economic factors.

Table 1 on page 4 from the CS 2011 memo gives validated air passenger rates.The total number of daily intra-California passengers between the Central Valley and the Bay Area/ Los Angeles is only 140 people. While some of the trips in this market are too short for air service, many are quite long. For instance, Bakersfield to San Francisco is 450 km and Anaheim to Fresno is 390 km. In other markets with these same distances, there is a much larger air market.

Second, the high number of trains in the operating schedule is an artifact of the high frequency coefficient. Over the last five years, service levels have been dramatically increased. This was done because at higher headways (headways more similar to those in Europe) the model showed very high sensitivity to headways. Parsons Brinckerhoff then methodically increased the number of trains until there were very limited additional gains to be wrung.

Next, we have concerns that this model may not be robust to the more normal operating schedules that for-profit operators would likely propose.

---

[6] CS Memo, Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Meeting Minutesfrom February 7, 2006

Finally, the high frequency coefficient completely skewed the ridership forecasts for the Altamont vs. Pacheco routing decision. The Altamont route had headways that were 150%- 300% higher than the Pacheco route. While we understand that the Panel has been told explicitly not to deal with this issue, our estimates (based on population and air demand) are that a route like Altamont that more directly serves the East Bay could have substantially higher ridership and we would recommend that the route be re-analyzed when a new model is developed.

Cost of driving

"Usually, auto travelers will consider their cost of travel to be only their out-of-pocket gas costs. Thus, in most intercity travel models, auto costs are generally in the range of $0.10 to $0.15 per mile. While higher per mile costs are more consistent with the true costs of driving (including operating, maintenance, and ownership costs), they are generally not considered by travelers for specific travel decisions."[7]

The model currently uses $0.24 per mile in 2005 dollars to calculate the cost of driving. In 2011 dollars, this cost is $0.27 per mile. This cost includes both fuel costs and other operating costs, assumed to be 60% of the fuel costs. The cost of gas is not explicitly included but can be calculated given an assumption of fuel mileage and a cost per mile of driving.

There are a couple of issues with this formulation.

First, it is not obvious that non-fuel operating costs should be included in a model of intercity travel. The ridership model is a behavioral model so it is important to have the perceived cost of travel, not the actual cost of driving. As CS points out in their review of a ridership model for another California high speed rail project, people do not typically consider costs like depreciation and repairs when taking occasional trips.

Second, if operating costs are included, there is no reason that they should be a function of gas prices. Our understanding is that this was a method used at one point by MTC but is no longer used by them. While that approximation may have been appropriate at one cost of fuel, it is subject to scale issues. While real costs of gas have increased substantially, there is no reason why maintenance costs should increase in sync with gas. What may have been appropriate when gas cost $1.10 per gallon may not be appropriate with gas closer to $4.00 per gallon. It is also confusing for sensitivity tests of gas price changes. The maintenance costs would also be increased, which may or may not be intended.

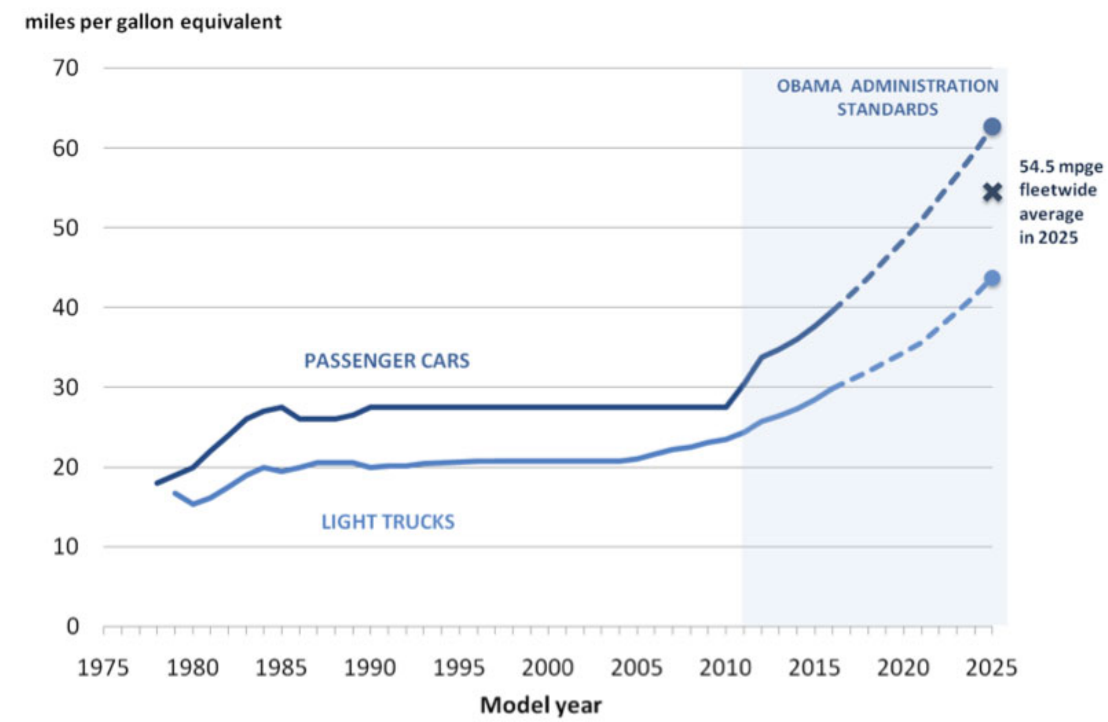
Finally, regardless of whether or not non-fuel costs are considered, the original SP survey was very specific that the costs were fuel only and the model was estimated accordingly. While the model was later calibrated using the more inclusive costs, this does not change the fact that all of the cost coefficients were estimated using just fuel costs. If people do mentally add in additional driving costs when making decisions, this would have been included in the regression coefficients.

---

[7] Cambridge Systematics Desert Xpress Ridership Forecast Review February 29, 2008

If it is determined that non-fuel costs should be included, then the model should be re-estimated adding in these costs to the cost of driving. Otherwise, the model should only use out of pocket costs like fuel.[8]

We would also point out that some significant increases in fuel efficiency have recently been mandated. Even if gas prices rise significantly, fuel efficiency gains may actually lower the cost of driving in the future, AKA the "Prius effect." This type of scenario should be considered along with higher driving cost runs.



MY1978-2011 figures are NHTSA Corporate Average Fuel Economy (CAFE) standards in miles per gallon. Standards for MY2012-2025 are EPA greenhouse gas emission standards in miles per gallon equivalent, incorporating air conditioning improvements. Dashed lines denote that standards for MY2017-2025 reflect percentage increases in Notice of Intent.

			MPG		
Cost of gas / gallon	20	25	30	40	50
$3.00	$0.15	$0.12	$0.10	$0.08	$0.06
$3.50	$0.18	$0.14	$0.12	$0.09	$0.07
$4.00	$0.20	$0.16	$0.13	$0.10	$0.08
$4.50	$0.23	$0.18	$0.15	$0.11	$0.09
$5.00	$0.25	$0.20	$0.17	$0.13	$0.10
$6.00	$0.30	$0.24	$0.20	$0.15	$0.12
$7.00	$0.35	$0.28	$0.23	$0.18	$0.14
$8.00	$0.40	$0.32	$0.27	$0.20	$0.16
$9.00	$0.45	$0.36	$0.30	$0.23	$0.18

[8] The case of business travelers is somewhat tricky. While federal reimbursement rates are very high, it is not clear that employees consider this as their cost.

Group travel

The CS model does not directly adjust the price of driving based on the number of passengers in the car. The standard treatment in other models is to calculate the cost by dividing driving costs by the number of passengers. This model does not do so.

For short trips, there does not appear to be any adjustment whatsoever.

For long trips, the CS model uses a group constant to represent both cost savings and a “road trip" effect in lieu of dividing the cost by the number of passengers.*The cost savings of traveling in a group has not been updated even as the assumed cost of driving has more than doubled.*

Given that most intercity car trips are with more than one traveler, this is a serious issue.In the SP survey, only 17% of all car trips were taken by single drivers.[9] Unlike local trips, the solo driver automobile is the exception rather than the rule and this effect is stronger the longer the trip.

Thus,the current model generally overestimates the cost of driving when there is more than one person in the car. This is a significant problem as a large majority of non-business intercity automobile trips are taken in groups.

In the original estimation, the consultant found that for long trips there was a group effect above and beyond the cost savings. There are several different plausible explanations for this finding. For example, multiple drivers can share driving and it is easier to deal with a whole family’s luggage with one’s own car.

For long distance non-business trips, the value of the group constantestimated using the SP survey data was $40 (2005 dollars). In the estimation, the average price of driving was about 11 cents per mile. Depending on the number of people in the car and how the group travel constant was attributed between cost savings and the “road trip” effect, this value gives very reasonable breakeven numbers.

Subsequent to the initial estimation and discussed in detail in the previous section, the price of driving has been significantly increased from the 11 cents per mile figure. It was 20 cents per mile in the initial calibration and is now 24 cents per mile. All these figures are in 2005 dollars.

However, the group constant has not been changed. It no longer gives a reasonable breakeven.

The below chart gives a typical per person cost of driving, assuming an operating cost of 24 cents per mile for a given number of passengers driving a given distance.

[99] CS Memo, Bay Area/California High Speed Rail Ridership and Revenue Forecasting Study Meeting Minutesfrom January 5, 2005

	100	200	300	400	500	600
2	$12	$24	$36	$48	$60	$72
3	$8	$16	$24	$32	$40	$48
4	$6	$12	$18	$24	$30	$36

The next chart shows the effective cost per person using the group coefficient method in the CS model and assuming that 75% of the value is attributed to cost savings. All numbers in red are higher than the per person driving cost a standard model would give.

	100	200	300	400	500	600
2	-$6	$18	$42	$66	$90	$114
3	-$6	$18	$42	$66	$90	$114
4	-$6	$18	$42	$66	$90	$114

For a 400 mile trip (Bay Area to Disneyland) with four people in the car, the difference is **substantial**. The model uses a cost of $66 per person vs. a traditional method cost of $24. For the household, this is **$264 vs $96** each way.

In addition, the issues with the price of driving identified previously compound the issues with the group travel coefficient.

Operating schedule

The current operating schedule has very frequent service and offers service to many different stations. Thus far, the operating schedule has been determined by the engineering program manager, Parsons Brinckerhoff, largely in response to a ridership model that showed excess sensitivity to normal frequencies. No actual train operator has been selected and there are several aspects about the schedule that a for-profit operator would likely change.

First, there would likely be many fewer trains. More service would be direct Bay Area to Los Angeles, in line with current demand for air service. And finally, there would be limited service available for regional commuter traffic.

Most non-subsidized long distance intercity rail systems in the US (Northeast Corridor) and in Europe try not to serve the regional commute market. They do this by setting fares very high, explicit prohibitions on short trips and scheduling.

This is done for very basic reasons. Because the demand for regional transit (Bakersfield to LA) outstrips the demand for long distance transit (San Francisco to Bakersfield), selling Bakersfield – LA tickets will often mean empty seats from San Francisco to Bakersfield. This dramatically lowers yield per mile operating results.

One solution to this issue is to run shorter distance trains to serve the regional markets. The commute market often has asymmetrical demand – with high demand in the morning in one direction and high demand in the evening in the other. This either requires additional trains or requires trains to run mostly empty in one direction. In addition, longer distance commuters tend to be sensitive to price so lower ticket prices need to be offered to attract significant numbers. In certain countries where there is a policy aim to spread out population away from the urban areas, there are very large discounts offered to long distance commuters.

Serving long distance commuters may serve policy goals but it is typically a money losing proposition. In general, this service is either offered through another subsidized service offering (AVE Avante[10]) or provided by a regional service (Metronorth, NJ Transit).

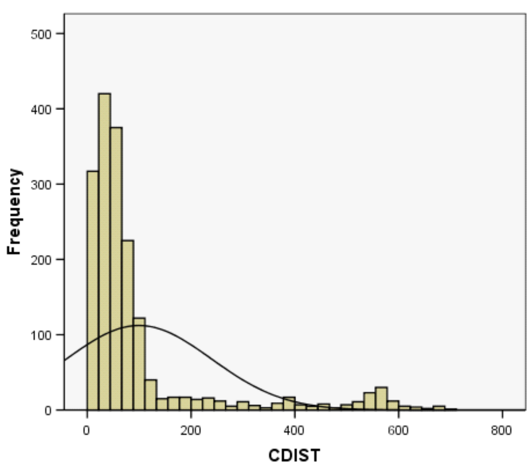
The CS model evenly spread out commuters throughout the morning and evening peak times, which is not likely to be realistic for most of the markets, particularly in phase 1 and the IOS.

In addition, there are no agreements that we are aware of with existing local train operators who may be competing for similar markets. For example, Caltrain's Baby Bullet service offers SJ to San Francisco service currently and this is one of the most profitable parts of their business. There are similar issues with Metrolink and the various Amtrak providers.

As the HSR operator has not been selected and no agreement has been reached with any local train operators as to who will serve the longer distance commuters, it makes sense to consider other operating schedules.

Commuters

The treatment of commuters in this model is problematic. Some of this stems from the initial data set. The vast majority of the "short" commuters surveyed were on a train to work. There were only 6 auto commuters. The original estimation revealed a value of time of only a couple of dollars per hour (this was later constrained to be higher), a clear sign that this was not a representative sample.



On the other hand, there were virtually no long distance commuters sampled. And of the 27 surveyed, a disproportionate number were flying from the Bay Area to LA. While there is a small group of such travelers and they would be very likely to consider HSR, a large majority of commuters over 100 miles are traveling just over that arbitrary limit.

[10] Cuidad Real in Spain has become known for its long distance commuters. The Avante commute service uses a lower end trainset and fares for commute times at more than a 75% discount to the service offered on the long distance trains.

While undocumented anywhere in the original reports, CS decided toaddress the small sample size by combining the long distance commuters with the business travelers, who were over-represented by those flying to and from the Bay Area.

This resulted in assigning a very high value of time to these commuters ($64), many of whom have a long commute because they decided that the lower cost of living outside of the Bay Area was worth the long time in the car.

While it is important to understand the potential of the high speed rail infrastructure to serve commuters, the actual level of service for commuters will likely depend critically on the available local subsidies. Realistically, a HSR operator will set fares high enough or alter schedules to limit demand during peak periods like Amtrak does with stations like New Haven and Princeton.

It may actually make more sense to separate out commute service from the long distance service, especially as it is unlikely that the long distance operator will be in a position to subsidize local service as significant project debt is envisioned and, by law, any extra money must help expand the system.

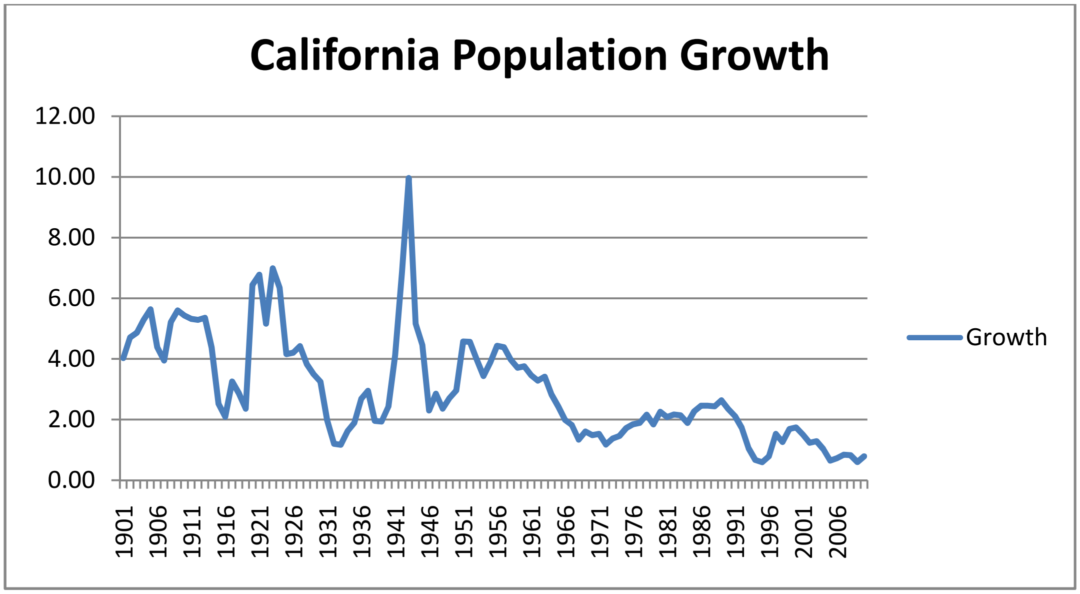
The regional transit authorities could then decide what type of regional transportation they want to support and then develop their own ridership numbers. The regional models have much more critical detail regarding access and egress and other connecting transit services.

This of course is only one possible approach but the current treatment of long distance commuters as high-end business travelers needs to be fixed in some way. This may also help solve some of the calibration issues with demand for air service from Central Valley airports.

Population

California experienced tremendous growth over the last century and current forecasts include quite high levels of growth. The growth rate has significantly slowed and growth over the last decade was half of the growth originally forecast. Previously California enjoyed high rates of migration. Migration accounted for about half of California's growth. Earlier this decade, preceding the economic slowdown, this trend dramatically reversed and more people are moving out of California than moving in. This is perhaps not surprising as California's schools are poor performing and there are few areas with both a good job market and affordable housing.

While population growth could resume at its previous trajectory, it is also possible that growth could occur at half the forecast rate. We would recommend that the model be run with high and low population forecasts to understand the sensitivity to this critical assumption.



Source: http://www.dof.ca.gov/research/demographic/reports/estimates/e-7/view.php
Additional detailed data for last decade:

http://www.dof.ca.gov/research/demographic/reports/estimates/e-6/2000-10/documents/E-6_Report_July_2000-2010_updated_with_2010_census.xls