

FINAL REPORT

Independent Peer Review of the California High-Speed Rail Ridership and Revenue Forecasting Process

Findings and Recommendations from August-December 2011 Review Period

February 15, 2012

The Peer Review Panel held its third formal meeting on August 9-10, 2011 at the offices of the San Francisco County Transportation Authority. The Panel subsequently conducted discussions via electronic mail, teleconference, and videoconferencing. This report covers their activities and deliberations from August through December, 2011.

The panelists include:

- Frank S. Koppelman, PhD, Professor Emeritus of Civil Engineering, Northwestern University (chair)
- Kay W. Axhausen, Dr.Ing., Professor, Institute for Transport Planning and Systems, ETH Zurich (Swiss Federal Institute of Technology Zurich)
- Billy Charlton, San Francisco County Transportation Authority
- Eric Miller, PhD, Professor, Department of Civil Engineering and Director, Cities Centre, University of Toronto
- Kenneth A. Small, PhD, Professor Emeritus, Department of Economics, University of California-Irvine

All panelists were present for the August meeting except for Dr. Axhausen, who attended via video-conferencing. Rick Donnelly, PhD, AICP of Parsons Brinckerhoff served as facilitator and recorder for the Panel. In this capacity he serves at the convenience of the chair rather than as a representative of the project management team. The Panel invited several others to attend some portions of the August meeting. These included:

- Roelof van Ark from the California High Speed Rail Authority (present for opening session with panel and wrap-up on second day)
- Hans Van Winkle, Gregg Albright, and Nick Brand from Parsons Brinckerhoff (representing the Program Management Consultant; all attended the wrap-up session, and Gregg attended all of the open portions of the meeting)
- Jeff Buxbaum, David Kurth, and Kimon Prousaloglou, from Cambridge Systematics. Dr. Prousaloglou was present during first open portion of the meeting only, while Mr. Buxbaum and Mr. Kurth attended all open portions of the meeting.

The subsequent dialogue of the Panel was closed except to members

1 Review of the Panel's Scope and Mission

Mr. van Ark was invited to comment on the work of the Panel to date, and to reaffirm the Authority's mission for the Panel. This occurred during the August meeting in San Francisco. The role of the Panel in guiding the future evolution and application of the model was endorsed, with appreciation expressed for the influence to date the Panel has exerted on the process. A continuing urgent deadline was identified with respect to completion of the 2012 business plan. This must be received by the California legislature by January, 2012, which in turn must be preceded by a 60-day public review. Thus, the Authority and its consultants were intensely focused upon an October deadline for completion of the draft business plan. Updated forecasts generated by Cambridge Systematics (CS) over the past two months are informing the business planning efforts.

It was noted that the possibility of private investment is likely to be at least five years into the future. Thus, adequate time appears to be available to implement the Panel's desire to steer the modeling process towards an explicit representation of risk and uncertainty in forecasting. While not a hallmark of current practice in North America, such measures are known to be of large concern to investors and their advisers. It was noted that the Spanish AVE system was operational and generating revenue before outside investors were brought in. The implication from a ridership and revenue standpoint is that the Authority will need an investment-grade model capable of demonstrating that the investment in high-speed rail (HSR) is sound for the State of California. Such a model must also inform the Authority in their negotiations with investors. While the latter are expected to undertake their own modeling as part of their due diligence, the Authority's own analyses will enhance its negotiating position. In particular, the need to support investment-grade decisions will likely emerge as part of the 2014 update of the business plan. While the current model is adequate for supporting the analyses conducted to date, the Panel reinforced the need for the next model version to address the long-term issues identified in its first (January-March, 2011) report.

The Panel and Mr. van Ark affirmed their shared desire to make the language and structure of the peer review reports more accessible to the public. The panel discussed this issue at length, both during this discussion and in subsequent closed parts of the meeting. It was agreed that the written reports are the only enduring record of the breadth and depth of issues investigated by the Panel. As such a delicate compromise must be struck between accessibility of the terminology and findings to a lay audience and the Authority's need for convincing evidence that the panelists fully examined detailed issues not necessarily of interest to lay readers. In light of the controversies surrounding the model the panel elected to err on the side of being more comprehensive than less, and more technical than expository, when required. However, it was readily agreed that the Panel would do everything possible to accommodate the needs of both audiences.

This portion of the meeting closed with questions from the Panel about anticipated future modeling milestones. It was noted that the next generation of the model, discussed further in Section 3 below, will need to be built, tested, and operational in time to support the 2014 business plan update. Between now and then the next big decision the model must inform will be the choice of whether to connect the initial operating system (IOS) in the Central Valley to the Bay Area to the north or Los Angeles Basin to the south. It is anticipated that an enhanced model (as defined below) will most likely be used to inform this decision unless the next generation model is completed sooner than expected.

2 Model Nomenclature and Versioning

The majority of the work of the Panel to date has focused upon a detailed examination of the line haul mode choice model. CS has explored several alternative specifications in response to guidance from the panel about the model's further development. As in the second meeting, the Panel discussed in closed session how these possible revisions might be incorporated into the modeling process. It was noted that every model has its strengths and weaknesses. The Panel endeavors to help CS find the best approach, taking into account the anticipated uses and audiences for future model development.

This discussion quickly brought into focus continued confusion about the various model designations in use by the Panel, CS, the Authority, and other consultants. At least four versions of the model have been discussed at various points:

- The *original model* developed by CS, initially under contract to the Metropolitan Transportation Commission (MTC), to include coefficients revised during the model calibration and validation processes. As noted in the Panel's second report (April-July, 2011), this model has now been extensively reviewed and found to perform satisfactorily for its intended uses of supporting planning and environmental analyses at the system and corridor level. This is the version of the model documented on the Authority's website.
- The original model was updated to include revised socioeconomic inputs and minor revisions to the trip generation model parameters, as described in the Panel's second report. It was used to produce the updated forecasts included in the 2012 business plan. This has been described as the *interim model*.
- An *enhanced model* will build upon the interim model by incorporating changes made to address the Panel's short-term concerns expressed in Section 4 of its first report. These changes will be incremental, and incorporated into the modeling process as each becomes available. This model will be recalibrated (if required) or re-validated (otherwise) only if required for interim forecasting. We recommend that such use be avoided, if possible, until all model changes have been completed for the second generation model.
- A *second generation model* addressing all of the short and long-term issues previously identified by the Panel is the eventual goal of the Authority. This has sometimes been referred to as an enhanced model.

The Panel recommends adopting a version numbering strategy to more clearly define which instance of the model is under discussion. A suggested numbering system is shown in the Appendix.

3 Review of the Interim Model

As noted, the most pressing deadline facing the Authority is completion of the 2012 business plan, which will include revenue projections, operating cost, and other performance measures derived from the ridership forecasts. The interim model was used to complete these forecasts, which CS delivered to the Authority as the Panel was convening for its third meeting. Most of the Panel's work to date has focused on the model itself and the data and assumptions that informed its development. The Panel had not fully reviewed the underlying assumptions and data used in the preparation of the forecasts using the model. The Panel began examining such during its third meeting.

3.1 Level of Service and Operating Characteristics

The Panel was particularly interested in assumptions about anticipated fares, both for HSR and competing modes. Geoffrey Gosling, a researcher at the University of California and private consultant, recently developed assumptions on air and high-speed rail (HSR) fare inputs to the forecasting process. This work will help address the Panel's concern, expressed in their first report, about inadequate consideration of the competition between HSR and air carriers. CS

briefly described the results of Gosling's work during the meeting. His final report to CS is under review by the Panel at this writing.

There was considerable discussion about fares and the impact of assumptions about them on the model. The panelists noted that the earlier runs of the CS model used fares in which HSR was assumed to cost one-half to three-quarters of competing air service. This range of fares was tested using the original model. The forecasted ridership dropped by a third when fares were increased from 50 to 83 percent of assumed competing airline fares. The Panel felt this was a reasonable response and within expected ranges. Moreover, it was acknowledged that fares will be dynamic and driven by market conditions, both of which are outside of the realm of model forecasting but amenable to investigation by sensitivity analysis. Nick Brand also reported that earlier tests of the effects of varying peak and off-peak fares had not resulted in large changes in forecasted fare revenues, presumably because changes in ridership and in revenue per rider tend to offset each other. After considerable discussion the Panel concluded that that the best approach for examining the likely range of future air and HSR fares and fuel prices is the use of expert judgment. The work by Gosling was considered a definitive source of such information.

The Panel explored several other assumptions used in existing runs of the model. The panel was reassured to learn that Phase 1 (HSR service from downtown Los Angeles to downtown San Francisco) was studied using numerous model runs looking at different network configurations. The panelists also discussed whether the model appropriately reflected the effects of increased roadway congestion, worsened roadway conditions, or the effects of dramatically higher fuel prices that might occur. Fuel prices, will be addressed in our next report.

3.2 Revised Socioeconomic Inputs

The socioeconomic data used to calibrate the original model were assembled from the metropolitan planning organizations (MPOs) within the state. Allocations of county-level population and employment estimates from the California Department of Finance were employed for areas outside of the MPOs. All of these data were assembled as part of the original model development, which predated the economic downturn on 2008. As such, it was natural to question whether the socioeconomic forecasts developed before the downturn represented a rosier outlook than is likely to emerge given current conditions.

CS obtained more recent data from Woods and Poole and Moody's Analytics for use in the interim model. These were characterized as short-term improvements to the socioeconomic forecasts rather than as being definitive, but still better than outdated data. Of particular interest was the difference in growth in population and employment totals in these data. Both sources show smaller 2030 forecasts for population, households, and employment than the earlier sources based on data from around 2000. As pointed out to us by Cambridge Systematics, the difference in forecasted employment is considerably larger than that in forecasted population (in the case of Moody's Analytics) or forecasted households (in the case of Woods and Poole). The Panel discussed possible implications of these forecasts at length. Total long-distance trip-making in the original and interim models is driven by changes in population rather than employment. Therefore, as noted by CS in presenting this analysis, the model will not be very sensitive to these changes in employment forecasts because it does not directly incorporate workers as trip generators. But in reality trip-making is sensitive to employment, and so it is possible that the model will over-estimate trip-making if future employment rates are significantly lower than the

base conditions for which the model has been calibrated. The extent of any such over-prediction cannot be easily assessed, given the complex relationships underlying long-distance travel generation, but it is likely to be small in the current application. Nevertheless, the issue of model sensitivity to employment should be addressed in the development of the second-generation model.

3.3 New data collection on long-distance travel

CS contracted with Harris Interactive to conduct an Internet poll of their panel, based on selected criteria, to collect data on long-distance travel within California in 2011. The results have been used to adjust trip generation rates to reflect more conditions more recent than those from which the model was originally estimated and calibrated. The Panel noted that the amount of time CS had to receive, process, and interpret the data amounted to only a few weeks. As with some other updated data sets recently incorporated, the results from Harris Interactive represent an improvement over the original data used to build the model, but lack the coverage and depth of the original surveys and so cannot substitute for them entirely.

The Panel discussed at length the process CS used to “expand” the data, i.e. to adjust the totals by various categories to reflect information in these new data sources. The Panel felt that the expansion methodology first proposed by CS was too simple, and especially needed to account for income. CS subsequently revised the expansion methodology, and has now provided an expanded explanation of it, including an explanation of why the particular two-month survey period chosen was a good proxy for annual trip patterns.

The expanded data on long-distance travel show a much lower share of commuting trips than did the original data, with a corresponding increase in the share of “other” trips. CS has correctly pointed out a number of factors making comparisons between years difficult, and other factors that make long-distance commuting less frequent in 2011 than in 2000, especially the end of the “dot-com” boom in the San Jose area. Furthermore, CS has now provided information making a good case that it was probably the original data that contained an unrealistically high proportion of long-distance commuting trips, rather than that the new data are too low.

3.4 Use of new socioeconomic and travel data

The panel believes that both the new expansion methodology and the use of the Harris Interactive travel data to adjust long-distance trip making are appropriate interim measures for current uses. The panel and CS agree that going forward, it will be better to recalibrate the model (either further developed interim models or the enhanced model that is under development) using the 2008 and 2011 data rather than the adjustment procedure now used. We also note that the draft Business Plan addresses the current uncertainty in baseline long-distance trip-making by using the original calibration for the “high” forecast and the adjustments based on newer data in the “low” forecast.

4 Suitability of the Model for Intended Uses

It is worth remembering that the original model was designed several years ago in order to enable the Metropolitan Transportation Commission to initially assess the feasibility of high-speed rail in the Bay Area and its assumed connection to the rest of the state. The rapid pace with which the Authority was subsequently formed, coupled with closely spaced and quickly

encountered state and federal legislative deadlines for various studies and reports, precluded the design and development of a newer model with enhanced capabilities to meet the expected long-term analytical needs of the Authority. Thus, the Authority's initial work to date has had to rely on incremental improvements to the original model and data. Thus, the distinction between an ideal model and one that is adequate for the tasks at hand must be made.

Our previous reports have described the short and long-term improvements required for an ideal model. They also describe how many of our concerns have been met through clarifications about the methodologies used, sensitivity testing, and comparisons to estimation results of alternative formulations suggested by the Panel. The recently completed validation work using updated inputs and the results of the Harris Interactive Poll, described earlier, have also increased our confidence in the interim model for its applications to date. The Panel has more recently also carefully considered the criticisms published by others, and researched those aspects of the model more closely. Our review of these findings lead us to conclude that the model and forecasts to date appear adequate for the tasks of assessing the feasibility of the system and building an initial business plan to guide its further development.

While adequate for their needs to date, the interim model and possible enhanced versions cannot meet anticipated future needs of the Authority. A second-generation model as described earlier is an essential evolution of the model that will further increase our confidence in the ridership projections and enable it to address planning and operational scenarios at higher levels of behavioral, temporal, and spatial resolution. It will address long-term issues identified in Section 5 of our first report and will require the collection of additional revealed and stated preference survey data. It will resolve controversies about the current model, incorporate the latest thinking in the design of discrete choice models and their estimation, and will be informed by expanded and more recent survey data. Such a model will be more closely compatible with the statewide travel model being separately developed by Caltrans, as it will share data used in the statewide model's development. Finally, it will facilitate more accurate quantification of the risk and uncertainty surrounding forecasts, an essential capability for its use in investment-grade forecasting.

5 Conclusion

The Panel is satisfied with the modeling work completed to date, the responses by CS to earlier published findings and recommendations, and the general direction the Authority is headed in updating the model for future anticipated uses. With the bulk of the urgent work of preparing forecasts for the 2012 Business Plan behind them, the Authority should now turn to a careful design of the next steps in the evolution of their ridership and revenue forecasting process.

The Panel continues to believe that future enhancements of the current interim model must address the long-term issues previously identified:

- Model validation and reasonability checks: These would include (1) comparisons to other HSR forecasts, (2) checks against newer survey data, and (3) extensive sensitivity testing of assumed HSR levels of service, congestion in competing modes of transport, differing socioeconomic forecasts, and assumptions about parking availability at planned HSR stations.

- Potential stated preference (SP) bias, which can be examined through comparisons with other models and studies using similar techniques, greater use of internal validation, comparisons with revealed preference (RP) studies, and collection of new SP and RP data from a sample of travelers in corridors currently well served by conventional rail.
- Testing whether newer econometric methods for estimation with choice-based sampling, discussed in Section 6 of the first report, would likely make any improvement in model accuracy. At this point, the panel feels they probably would not, but this should be confirmed in future model re-estimations

To meet the short and long term objectives of the Authority, it is essential that all parties agree on a time line for completion of models with different levels of enhanced capabilities. The Authority should take primary responsibility for this. Work tasks and assignments should be modified appropriately if the Authority requires or desires assistance from the Panel or CS in this regard.

The Panel endorses plans to continue to improve the current model by addressing the short-term issues previously identified, on which CS has made commendable progress. Further, the Panel also stresses the need to move quickly to address the long-term issues in a second-generation model, which will impart greater confidence in future modeling work and forecasts by both internal and external constituencies.

The panel believes that the most important future work in developing the proposed enhanced model is addressing the long-term issues previously identified. Addressing some of these issues may require some changes in the current structure of the model. It is possible, for example, that the mode and destination choice models would benefit from being jointly estimated. The goal should be to have the second-generation model operational in time for use in the 2014 revision of the business plan.

Appendix: Recommended version numbering

Recommended version numbering

Version	Corresponds to	Major features	Anticipated uses
1.0	Original model	Original model structure and formulation, updated through 2010 with calibrated/validated parameters	All applications through June 2011
1.1	Interim model	Original model updated with 2008 SE data from economy.com and adjusted trip tables based on 2011 Harris Interactive Poll	2012 Business Plan
1.2 through 1.9	Enhanced model (minor version number increases each time a major change is made to the model structure and re-calibration and re-validation has been completed)	Interim model with successively more improvements that address short-term issues identified by the Panel, and using newer model inputs (e.g., statewide travel survey, update Caltrans networks) as they become available	All applications after Business Plan submittal until the Version 2.0 model becomes operational
2.0	Second (next) generation modeling platform	New model specification that incorporates short-term improvements in Version 1.2 through 1.9, addresses the long-term issues identified by the Panel, and is designed to support investment-grade forecasting	2014 Business Plan and subsequent analytical requirements