AECOM Consult, An Affiliate of DMJM Harris
3101 Wilson Boulevard, Suite 400, Arlington, VA 22201
T 703.682.5100 F 703.682.5001 www.dmjmharris.com

Memorandum

Date:	July 2, 2008
To:	Sarah Sun, Fred Ducca (FHWA)
From:	Bill Woodford (ACI), Bill Davidson (PB Americas), Dave Schmitt (ACI)
Subject:	Webinar Summary Report

The "Shining a Light Inside the Black Box" webinar series, conducted as part of the Travel Model Improvement Program in the Spring of 2008, was an outgrowth of one of the recommendations contained in *Transportation Research Board Special Report 288: Metropolitan Travel Forecasting – Current Practice and Future Direction*. That report was commissioned to gather information on the state of the travel forecasting practice in the United States and make recommendations for improving these practices in the future.

In its evaluation of the state of practice, the committee noted that traditional travel demand models have inherent structural weaknesses that limit the ability of these tools to forecast future travel demand levels. More importantly, however, the committee noted shortcomings related to poor technical execution in the development and use of existing modeling techniques, which can have a negative impact on forecasting accuracy. Examples of poor execution related to existing modeling techniques include the failure to collect adequate data, optimism bias, and an insufficient emphasis on model validation. Furthermore, the committee said that deficiencies in the practice of travel forecasting will not be resolved solely by switching to more advanced models. Instead, the profession must also focus on improving the practice of travel demand forecasting.

The Federal Highway Administration (FHWA) has supported these findings, stressing that the basic practice of travel forecasting needs improvement regardless of the technical methods that are used. FHWA commissioned the "Shining a Light Inside the Black Box" webinar series to begin the process of raising awareness of best practices related to developing, maintaining, and applying travel forecasting tools.

Motivations
As an introduction to the webinar series and as an illustration of why good practice matters, we reviewed existing studies that compared past travel forecasts to the outcomes that actually occurred. The first of these studies, Flyvbjerg's 2005 study, examined 210 large transportation infrastructure projects completed between 1969 and 1998. The study found that half of all road traffic forecasts were wrong by more than 20%, and rail passenger forecasts were overestimated in 90% of the projects. Flyvbjerg investigated whether accuracy was better in more recent projects than older ones. He found that both road and rail forecasts have not improved over time. He concluded that the results "show that it is highly risky to rely on travel demand forecasts to plan and implement large transportation infrastructure investments."

The second study, *National Cooperative Highway Research Program's Synthesis 364*, analyzed predicted and actual toll road revenue for 26 U.S. toll road projects completed between 1986 and 2004. The authors found that actual revenues for 20 of 26 projects were less than 80% of projected revenues in their first years of operation. Only three projects had revenue greater than the predicted value. The authors also concluded that predictive techniques have not improved over time.

While not an academic study, the Federal Transit Administration began requiring detailed reporting of travel forecasts for proposed major transit projects and lower cost alternatives in 2002. Its reviews found significant problems for most projects in either the models or the comparability of the alternatives. It found strong indications that cursory reporting of travel forecasts hides major errors in many models and networks.

Finally, Rodier's 2003 study of the 1991 Sacramento travel forecasting model indicated that travel models were being used to distinguish among alternatives in cases where the differences were below the model's margin of error. Using actual land use for the year 2000, she found that the model overestimated Vehicle Miles Traveled (VMT) by 6%, Vehicle Hours Traveled (VHT) by 4% and Vehicle Hours of Delay (VHD) by 17%. If the original 2000 land use data set (as projected in 1991) was used, the model overestimated VMT by 12%, VHT by 13% and VHD by 38%.

These studies all show that travel forecasts often have large errors and that these errors typically overestimate travel demand. As a consequence, the results of these analyses are, at best, of limited usefulness to decision-makers. We believe that these poor results occurred (and continue to occur) because of at least one of the following:

- The project's characteristics, as constructed or implemented, did not reflect the planned and modeled characteristics.
- The future setting reflected in the model, most notably demographic and transportation infrastructure input assumptions, did not correctly reflect actual conditions.
- The models themselves did not fully or correctly understand the relationship between demographic conditions, transportation networks, and the resulting travel behavior.

The first two characteristics are beyond the direct control of the modeling community. Modelers have virtually no ability to impact how a project or program is actually constructed or implemented. Modelers have a limited ability to alter the future setting forecast, because that is usually developed by land use and transportation planners and finalized by the local Metropolitan Planning Organization (MPO). In both cases, modelers must respond by communicating the impact of uncertainties in project definition or future setting by describing how inaccuracies in the input assumptions affect potential outcomes.

Modelers do have direct control over the third characteristic of poor forecasting results—models that do not appropriately represent travel behavior. We can correct this in a number of ways: by improving models and modeling practice where we can, informing decision-makers about the insights gained from the forecasts, and recognizing and communicating the dimensions of uncertainty to decision-makers and the public. It is in these areas that we feel that modelers need to adjust current practice or techniques.

We identified four key issues related to good practice that are applicable to virtually all model sets, regardless of whether the models are very basic or extremely advanced. We believe that careful attention to each of these items would result in significantly better forecast accuracy:

- Accurate representation of transportation supply

- Travel distribution models that represent sub-regional variations in travel patterns
- Validation strategies that focus on matching observed, disaggregate travel patterns
- Communication strategies to inform decision-makers about the insights that can be obtained from model results

Accurate Representation of Transportation Supply
Realistic representation of the key characteristics of transportation supply (e.g., travel time, cost, capacity, and connectivity) is vital for developing models that properly reflect the relationship between these facilities and travel behavior. Too often, models rely on transportation network information that is coded incorrectly or is processed by the model in ways that result in highly inaccurate estimates of capacity and travel time. Improved practice requires a protocol to effectively check coded representations of transportation supply, such as the number of highway lanes, bus headways, and network connectivity.

Functional relationships must be logical and represent the real-world interrelationships between the physical attributes of the transportation system, the demand for using that infrastructure, and the resulting level of service. We encourage relating link speed and capacity to the physical characteristics of the roadway rather than the functional classification of the facility. Key physical attributes that affect speed and capacity include posted speed limit, design speed, signalization, horizontal and vertical alignment, lane widths, presence of a median, and presence of parking. On the transit side, bus travel times should be related to congested highway travel times plus an allowance for stop delay to reflect the conditions that cause buses to run more slowly than adjacent traffic. This approach results in more realistic bus speeds than the more typical factor on highway time, particularly in cases where highway congestion is expected to increase dramatically in the future. The focus of each of these functional relationships should be an accurate reflection of how and why the transportation system is influenced by congestion.

Finally, it is vital that the representation of the supply is fully validated. Supply validation techniques include comparing link-specific free-flow and congested speeds by mode against observed speeds, testing paths of key origin/destination paths by mode, and testing level-of-service (skim) matrices against observed travel times.

Travel Distribution Models That Represent Sub-Regional Variations in Travel Patterns
Distribution is the major input to mode choice and assignment models, but results from distribution models are rarely checked or investigated in great detail. Even when trip distribution models are carefully calibrated to match trip length frequency distributions, the resulting models often fail to replicate observed district-to-district flows. Common issues are related to the thoroughness of validation and model structures that understate the importance of non-transportation factors, such as residential and work place locations, in determining key trip anchors.

Typical distribution model development involves calibration using aggregate measures such as matching modeled and observed trip length frequency distributions stratified by trip purpose. Although a useful first step in the model development process, these distributions do not provide sufficient information to demonstrate that the distribution model fully represents regional trip-making patterns.

Good practice requires a detailed comparison of the modeled person trip flows to observed flows for each trip purpose and each socioeconomic class at a district-to-district level of detail. Unfortunately, this step is often bypassed due to insufficient data on observed travel flows. Regions depending on household surveys with very small sample sizes will not be able to make a meaningful verification of estimated travel patterns. Better and more data are needed to be able to make these comparisons.

The webinar discussed two structural issues with distribution models today. The first issue is an over-reliance on transportation as the driver of distribution. Distribution is a complex social behavior that relates transportation to other factors such as residential location choice, employment options, and non-work activities. Although transportation is an element of this decision process, other factors, such as housing cost, employment type, socioeconomic characteristics, and even past travel patterns, may be more important.

Many distribution models distribute all trips by purpose using a single measure of zone-to-zone impedance that ties back to travel time. This overreliance on transportation as a driving factor inflates the impact of the transportation system on travel distribution and may result in models where a small change in the roadway network leads to a disproportionate impact on distribution patterns.

The most practical way to introduce non-transportation factors into existing distribution models is by increasing the level of trip disaggregation. Most models treat all types of productions and attractions identically and attempt to find the generalized cost function that explains how travelers link origins and destinations. A classic problem occurs when a downtown area with high-income workers is surrounded by low-income residential areas. Unstratified models are likely to connect too many low-income residents with high-income jobs. Other examples of potential stratifications include shopping trips, which vary by retail type. Grocery stores have much shorter trip lengths than a regional mall that draws customers from a much larger area. We recommend that productions and attractions be disaggregated to the extent possible.

A second issue arises from the mechanism used to constrain the results of distribution models to the original production and attraction totals. In the webinar series, we showed an example where the balancing mechanism itself caused “decline” in travel between a stable inner-suburban area and downtown even when both areas experience modest growth. This occurs between rapidly growing exurban areas that take a greater share of travel to the Central Business District, leaving the inner-suburban workers to find employment elsewhere. We suggest that in some cases, models should not be constrained by input productions and attractions. In fact, the mismatch between original productions and attractions and the output of the distribution model may represent the economic attractiveness of those areas.

In all cases, validation of model results is vital. Distribution models should be validated using detailed comparisons of travel patterns by purpose and socioeconomic stratification at a district-to-district level of detail. A poor distribution model can be solved immediately replacing it with data entirely, such as the Census Transportation Planning Package (CTPP) 2000, or by using a large-sample travel survey, embarking on a more complete disaggregation of distribution models, or applying (with great care) K-factors as needed. In any event, modelers should be aware of what models are not capturing (or cannot capture) in travel distribution and include those aspects that can be captured.

Validation Strategies That Focus on Matching Observed, Disaggregate Travel Patterns

In our experience the traditional model development under-emphasizes validation, the step where the model is tested and refined to represent regional variations in travel patterns, which is the heart of developing a successful model. This often occurs because model estimation and calibration tend to consume a large amount of resources, leaving little time and few resources for extensive validation. Even when time and budget are not issues, adequate data may not be available to support extensive validation. Finally, validation efforts have become overly focused on traffic or transit line volumes, rather than validating all steps or phases of the model.

We strongly recommend that modelers avoid underestimating the resources and time needed to properly validate and test their models. This can be done by addressing schedule and resource issues before model development is initiated. In our experience, we find that model testing, validation, and documentation do not conform well to fixed timetables. This is partly because the issues that arise from this process are unknown beforehand. To this end, it is helpful that critical path modeling needs not be dependent on a timely model validation.

Collecting data to sufficiently test model results is best achieved when a plan is developed that relates the model structure and development to the data collection plan. For MPOs, we highly recommend that this be accomplished via their Unified Program Work Plans.

In the webinar series, we promoted an expansion of the model calibration/validation efforts beyond traffic counts and/or transit line volumes. Performing more meaningful model tests uncovers problems that would remain hidden if traditional calibration and validation practice is followed. We mentioned in the previous section that an inspection of person demand/travel flows is essential to test distribution models. Comparing point-to-point travel times is necessary to validate the transportation supply (namely, free-flow speeds) and model estimates (congested speeds). A comparison of estimated and observed trip tables is also a particularly useful test of the model. This test is extremely helpful using data from choice-based surveys, such as a transit on-board survey. The trip tables should be compared at as disaggregate a level as possible. Another test is to assign the observed trip tables to the network. This test can uncover many problems with the model's representation of transportation supply, problems that would be difficult to uncover otherwise.

It is important that the models themselves relate properly to traveler behavior. A good test is to interpret, in a qualitative explanation, the model's parameters, constants, coding conventions, and other decision rules as they relate to the transportation system and how travelers use it. This test helps to ensure that the various parameters, constants, and model algorithms tell a coherent story about travel behavior.

The travel demand model should also demonstrate reasonable predictions of change. Models should provide reasonable predictions of change between today and a future no-build condition, and between a future no-build condition and a realistic alternative (i.e., a change in the transportation system). These tests should be performed applying the model in full production mode (i.e., testing realistic alternatives using the full model set). Findings of these tests can highlight problems not prevalent in base-year conditions. High levels of congestion in future-year scenarios, for example, tend to amplify problems not visible in the validation year.

Uncovering problems within the model is the key undertaking of model testing. Once basic problems are identified and fixed, secondary problems with the representation of travel demand are revealed. Correcting these more subtle problems requires special attention. We believe that wherever possible model adjustments need to reflect the factors that truly affect travel behavior and not simply be an arithmetic alteration. We highlighted an example from Charlotte, North Carolina in the second webinar, where we altered the model's representation of auto-access to transit to represent observations suggesting that park-and-ride customers expect a direct express trip from the park-and-ride lot to their destination.

We suggest using a three-step process to improve the model's understanding of real-world traveler behaviors: (1) look at patterns in the travel data, (2) understand why the model does or does not properly value the trip, and (3) adjust the model.

In addition to reporting the calibration and validation results, the model documentation should describe the "readiness" of the model set for forecasting, including:

- The presentation of the significant travel markets and facilities that exist today
- The ability of the model set to describe the nature and magnitude of those markets
- The reasonableness of predicted changes in land use and transportation

- The identification of the model set limitations that restrict the correct representation of current travel markets, behaviors, and modes
- The identification of future travel markets and/or facilities and their representation for forecasting

The documentation should be devoid of technical jargon or equations so that inconsistencies between the model's parameters and traveler behavior can be easily identified.

Communication Strategies to Inform Decision-Makers
Improving the profession's ability to communicate model insights to decision-makers is becoming more and more important as stakeholder scrutiny increases. Frequently, we have found that modelers are unable to interpret model results in a way that leads to meaningful insights for decision-makers. Part of the issue is that we are overly focused on the numerical outcomes and not on telling a coherent story describing the meaning of the forecasts.

To begin to communicate with non-modelers, we must establish what the model really knows and what it does not know in each phase or step (i.e., networks, generation, distribution, path-building, mode choice, and assignment). It is also important to know the degree to which new programs or policies have a peer in the existing world, because models cannot reliably understand what does not appear today. The modeler then must identify sources of risk and uncertainty. The modeler must communicate in clear terms (i.e., without using technical jargon) to decision-makers.

Future Research Ideas
In addition to the near-term solutions presented earlier, we identified a number of areas where future research could be a great help in improving modeling practice. For instance, we know that disaggregate models make efficient use of small sample size travel surveys. However, we also know that larger sample sizes are needed in order to properly understand travel markets. In his landmark 1979 Transportation Research Record paper, Michael Smith suggested that extremely high household sampling rates would be needed to ensure statistically sound zone-to-zone travel patterns, assuming they are distributed normally. It would be helpful to verify this conclusion and research methods to calculate the minimum sample size needed to gather travel pattern information at a district-to-district level, where the districts are defined to be well below county level.

If large sampling rates are still found to be required, it is likely that alternative data collection strategies may be required. Research could be directed toward the effective collection and application of passively collected data. This type of data is very attractive because it is relatively inexpensive to acquire compared to traditional means and is not subject to common survey problems such as response biases and sampling errors. In some cases, such as freeway speed data and transit "smart cards," the data is very prevalent because it is collected continuously. We highlighted four examples from around the country where passively collected data is being used. We encourage other areas to examine passively collected data and use it in innovative ways.

Future research geared toward improving distribution models would be especially helpful, specifically in applying techniques that account for real estate and employer location decisions such as in land use models like UrbanSim. Exploring alternative strategies of capturing non-transportation elements in distribution models would be helpful to address the overreliance on transportation prevalent in today's distribution models.

Summary of Issues and Near-Team Solutions

Issue	Near-Term Solutions
Inattention to the Representation of Transportation Supply	Continuously check coded data, functional relationships, paths, and skim matrices
	Relate speed and capacity to the physical characteristics of the roadway
	Compare speeds to actual data
	Test paths
	Review zone sizes
Inattention to Travel Distribution Models	Validate distribution models using detailed comparisons
	Replace it with data entirely, such as CTPP 2000 or large-sample travel surveys
	Perform a more complete disaggregation of distribution models
	Apply K-factors as needed with caution
	Be aware of what models are not capturing (or cannot capture) in travel distribution and include those aspects that can be captured
	Explore alternative strategies of capturing non-transportation elements
Collective Problems with Traditional Model Development	Plan for sufficient time for validation
	Collect data to sufficiently test model estimates and results
	Expand model calibration and validation efforts
	Demonstrate reasonable predictions of change
	Make model adjustments that reflect real-world travel behavior
	Interpret model parameters, coefficients, etc. vis-à-vis traveler behavior
	Provide informative documentation of testing results and forecasting weaknesses
Inability to Effectively Communicate Insights from Model Forecasts to Decision-Makers	Establish what the model really knows and what it does not in each phase or step
	Know the degree to which new programs and policies have a peer in the existing world
	Identify sources of risk and uncertainty
	Communicate insights (not the results) in clear terms devoid of technical jargon

Summary of Future Research Ideas

Issue	Future Research Ideas
Inattention to Travel Distribution Models	Identify methods to determine minimum sample size for effective evaluation of travel patterns
	Investigate new data collection techniques to satisfy larger sample size requirements
	Improve distribution models by applying techniques that account for real estate and employer location decisions
	Explore strategies to capture non-transportation elements
Collective Problems with Traditional Model Development	Identify methods to determine minimum sample size for effective evaluation of travel patterns
	Investigate new data collection techniques to satisfy large sample size requirements
	Investigate methods to effectively collect and apply passively collected data

DISCLAIMER

The Web Series is part of a capacity building initiative of the Travel Model Improvement Program (TMIP). The Web Series presenters have extensive experience in all aspects of travel forecasting and the views presented in this Web Series are based on their experience. Those views do not represent official FHWA policy. The views expressed also do not represent the opinions of FHWA and do not constitute an endorsement, recommendation or specification by FHWA. Likewise they do not determine or advocate a policy decision/directive or represent specific recommendations regarding future research initiatives. FTA also supports the effort to provide a general picture of the state-of-the-practice on analytical work supporting the metropolitan and statewide transportation planning process. This presentation does not address the specific requirements used in rating and evaluating New Starts proposals.