UNDERSTANDING AND ACCOMMODATING RISK AND UNCERTAINTY IN TOLL ROAD PROJECTS: A REVIEW OF THE LITERATURE

Jason D. Lemp
Graduate Student Researcher
The University of Texas at Austin - 6.508. Cockrell Jr. Hall
Austin, TX 78712-1076 jdl880@mail.utexas.edu

Kara M. Kockelman
(Corresponding author)
Associate Professor and William J. Murray Jr. Fellow Department
of Civil, Architectural and Environmental Engineering The
University of Texas at Austin - 6.9 E. Cockrell Jr. Hall
Austin, TX 78712-1076
kkockelm@mail.utexas.edu
Phone: 512-471-0210 & FAX: 512-475-8744

The following paper is a pre-print and the final publication can be found in
Transportation Research Record No. 2132: 106-112, 2009.
Presented at the 88th Annual Meeting of the Transportation Research Board, January 2009

Key Words: travel demand forecasting, project risk, highway tolling, forecast uncertainty

ABSTRACT

Forecasting traffic and toll revenues for new highway projects involves great uncertainty due to the inherent uncertainty in the models used to make forecasts. As private investment becomes more common in project financing, quantifying the levels of risk and uncertainty associated with such projects becomes critical. This paper represents a review of many key studies and reports dealing with uncertainty in traffic and revenue forecasts for highway projects. These studies found that tolled projects tend to suffer from substantial optimism bias in forecasts, with predicted traffic volumes exceeding actual volumes by 30% or more about half of the time. Moreover, projects with greater uncertainty tend to overestimate year-one traffic volumes more and stabilize at lower final traffic volumes. But after controlling for added optimism bias in traffic forecasts (compared to non-tolled projects), there is little difference in uncertainty levels between tolled and non-tolled forecasts. A typical way to address uncertainty in traffic forecasts is through sensitivity testing, via variations in key inputs and parameters. A more extensive and less arbitrary version of this, Monte Carlo simulation, can provide probability distributions of future traffic and revenue, though it tends to require many simulations, which demand greater computational effort and time, unless networks are streamlined. Nonetheless, if reasonable assumptions for model input and parameter distributions can be made, Monte Carlo simulation generates a variety of useful information, and establishes the actual likelihood of loss (rather than more basic win/lose indicators from a limited set of "stress tests").

INTRODUCTION

Considerable uncertainty exists in traffic forecasts for new highway projects. While such uncertainty is not unexpected, in many projects it is largely ignored by designers and transportation planners. Compounding the issue, rising congestion and scarcity of open space in urban areas means that forecasting errors can be quite costly. Recently, Flyvbjerg et al. (2005 and 2006) sought to quantify and document the level and nature of uncertainty in traffic forecasts using data from highway and transit projects across the globe. The absolute difference between forecast and actual traffic is more than 20% for about half of the highway projects examined, and about 40% for roughly one-quarter of projects. However, they did not consider uncertainty in traffic predictions for tolled highways specifically.

Uncertainty in demand for *tolled* roadways is compounded by the introduction of more unknown variables. Yet such new understanding can be critical, since private investment generally depends on cost recovery through toll collection. In order to begin to address this clear gap in the literature, Standard & Poor's (Bain and Wilkins 2002, Bain and Plantagie 2003 and 2004, Bain and Polakovic 2005) and Fitch Ratings (George et al. 2003 and George et al. 2007) produced a series of studies that examine the risk and uncertainty of tolled highway projects. This paper summarizes key elements of those results and investigates methods for accommodating (or at least recognizing) uncertainty in the forecasting process. The resulting synthesis is intended to offer guidance in planning and decision-making processes of tolled roadway projects. The first section of this paper describes the observed frequency and magnitude of traffic volume mispredictions (forecast versus actual), while the second explains the various sources of risk and uncertainty in traffic forecasts and how these relate to project financing. The third section describes methods for recognizing and incorporating uncertainty in models of travel demand.

FREQUENCY AND MAGNITUDE OF MISPREDICTIONS

Standard & Poor's (S&P's) study of traffic forecasts began in 2002 with data on 32 toll road projects from around the world. The sample was then increased to 68 and 87 projects in 2003 and 2004, respectively. However, in both updates the conclusions remained largely the same.

In the first study, Bain and Wilkins (2002) found that traffic forecasts for new toll roads suffer from substantial optimism bias, a finding that is supported in the subsequent studies. The average ratio of actual-to-forecast traffic volumes in the first year of operation was about 0.73 (versus 0.74, 0.76, and 0.77 in the 2003, 2004, and 2005 studies). Figure 1 shows the distribution of forecasting errors in the 2005 update. (Comparisons to non-tolled projects are drawn later in this section.) Of course, due to the nature of averaging ratios such as these, traffic forecasts for toll roads may be over-predicting actual volumes by even more than 33% (implied by an actual-to-forecast ratio of 0.75).[1] Moreover, the 2002 study found that 78% of actual-to-

[1] A volume-weighted average of ratios (essentially the sum of predicted values over the sum of actual values) yields a much more robust indicator of the average percentage error, reflecting whether an investor will win (average >1) or lose (<1) − on average, across projects. Essentially, the issue is that the ratios are non-negative and bounded by zero, leaving a right-side skew that can tends to bias averages high. For instance, if predicted-to-actual ratios for two projects are 0.5 and 2.0, the average is 1.25, suggesting predictions are biased high. If we invert the ratios first and

forecast traffic volume ratios were less than 0.9 while only 12% were over 1.05. In the 2003 study, 63% of such ratios were less than 0.85, and 12% were over 1.05. Essentially three quarters of first-year traffic forecasts for tolled facilities are overestimated by 10% or more, suggesting that planners, bankers, and communities should be wary, and modelers need to improve their methods.

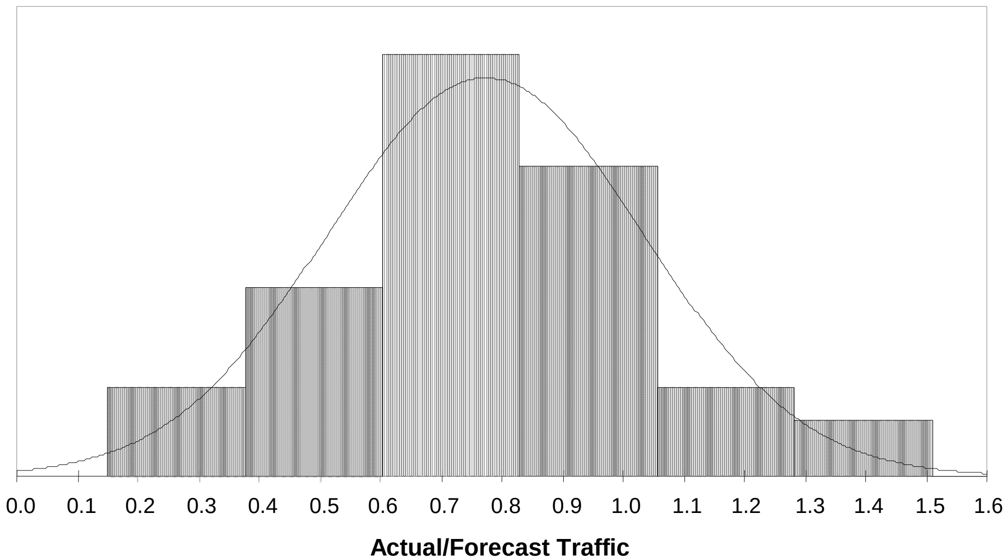

Figure 1: Distribution of actual-to-forecast traffic volumes (Source: Bain and Polakovic 2005, Chart 1)

One of the main diagnostics to come out of the 2002 study was S&P's Traffic Risk Index (TRI). While the exact details for its estimation are proprietary in nature (and thus not provided), the index attempts to predict the amount of project risk based on many project attributes, as discussed later in this paper. Based on the TRI, Bain and Wilkins (2002) determined a risk level (low, average, or high) for each project, and divided its discussion by forecast source: those commissioned by banks versus those commissioned by others.

The findings suggest that actual-to-forecast traffic volume ratios in the first year of operation average about 0.9 for "low-risk" bank-commissioned projects, and 0.8 for "low-risk" projects commissioned by others. Both types of low-risk projects had average ramp-up durations[2] of about 2 years (after which actual volumes track forecasts, on average). For "average-risk" projects, year one volume ratios were found to be 0.8 and 0.65 for bank- and non-bank-commissioned projects, respectively. Ramp-up duration was about 5 years in both cases. However, those commissioned by banks ramped-up to about 95% of forecast volumes over those first five years, while others ramped-up to only 90%. For "high-risk" projects, the volume ratios were just 0.7 and 0.45, respectively, and ramp-up durations were about 8 years. After ramp-up, bank-commissioned projects reached about 90% of forecast volumes while other projects

---

then average, the result is again 1.25, but the interpretation is that predictions are biased low. Thus, one must use caution when dealing with averages of ratios.

[2] The ramp-up period is the period in which traffic volumes rise to a relatively stable or equilibrium level. This period may require several years.

reached approximately 80% of forecast. What this suggests is that projects with greater uncertainty (and thus risk) overestimate initial traffic volumes by a greater amount, on average, experience a longer ramp-up duration (to reach stable volumes), and stabilize at lower final traffic volumes (versus predictions). Moreover, the magnitude of risk is greater for projects not commissioned by banks, which is not so surprising given that banks are much more directly accountable for investors' monies than are public agencies. Moreover, other project commissioners (public agencies, interest groups, and bidders) may have interests that are best served when predicted traffic volumes are high (Bain and Wilkins 2002).

With the 2003 study's increased sample size, Bain and Plantagie (2003) were able to conduct several less aggregate analyses. Multiple factors were investigated, but only one with significance was found, in distinguishing countries with and without a tolling history. The findings suggest that actual-to-forecast volume ratios in the first year of operations averaged 0.81 in countries with a history of tolling, but just 0.58 in other countries. Thus, forecast risks appear much higher in countries without a history of tolling. This is intuitive, given that user adoption will be much faster (thanks to existing toll tag and manual payment experiences) and that contractor and operator familiarity will be higher. In several U.S. regions (e.g., Florida, Southern California, New York, and Houston), flat-rate tolling is already well-established; so, in these regions it may be reasonable to expect first-year ratios in the neighborhood of 0.8. However, most other U.S. regions may dramatically under-perform if more appropriate modeling assumptions are not used (particularly for the ramp-up period).

In the 2004 update (Bain and Plantagie 2004), traffic forecasts along new tolled highways were compared to those of new non-tolled facilities. The sample size was increased to 87 highway projects, with all data for non-tolled facilities coming from Flyvbjerg et al.'s (2005 and 2006) work. The comparisons suggest that new non-tolled roadways exhibit little optimism bias, though the same amount of uncertainty or spread in the distribution (of volume ratios) remains. Figure 2 shows how the two distributions appear similar, but with the distribution of tolled road (forecast-to-actual) volume ratios shifted to the left by about 0.2 units (essentially indicating a 20-percent optimism bias). This suggests that, after controlling for the added optimism bias of tolled projects, there may be little difference in the accuracy of traffic forecasts for tolled and non-tolled projects.

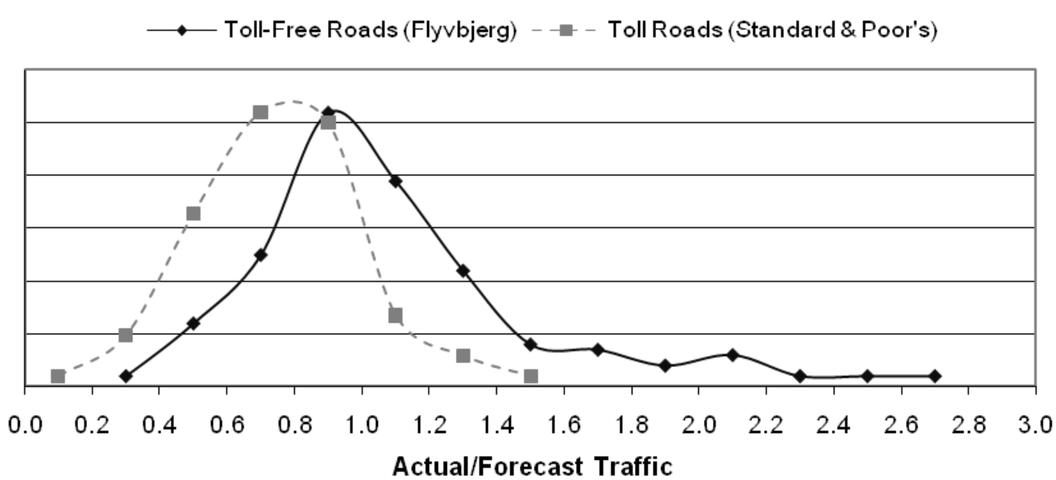

Figure 2: Distribution of actual-to-forecast traffic volumes for tolled and non-tolled projects (Source: Bain and Plantagie 2004, Chart 3)

Of course, S&P's uncertainty estimates for tolled roads were not exactly those published by Flyvbjerg et al. (2005 and 2006). Flyvbjerg et al. (2006) looked at 183 road projects (tolled and non-tolled), about half of which had actual-to-forecast volumes less than 0.8 or greater than 1.2 in their first year of operation (whereas S&P's studies [Bain and Wilkins 2002, Bain and Plantagie 2003 and 2004] found that roughly 65% of toll road projects fall into these two tails), and 25% had actual-to-forecast ratios less than 0.6 or greater than 1.4 (about 30% for tolled projects, according to S&P [Bain and Wilkins 2002, Bain and Plantagie 2003 and 2004]). Flyvbjerg et al. (2005 and 2006) found that the average actual-to-forecast ratio for non-tolled roads is 1.09 (with a 95% confidence interval on this value lying between 1.03 and 1.16). Of course, this average ratio is higher than if a weighted average were taken (as discussed previously). A weighted average ratio would likely be very close to zero since there appears to be approximately the same number of projects falling above and below the break even ratio of 1.0 (this corresponds to the 0% difference shown in Figure 3). Figure 3 shows Flyvbjerg et al.'s distribution of actual-to-predicted volumes.

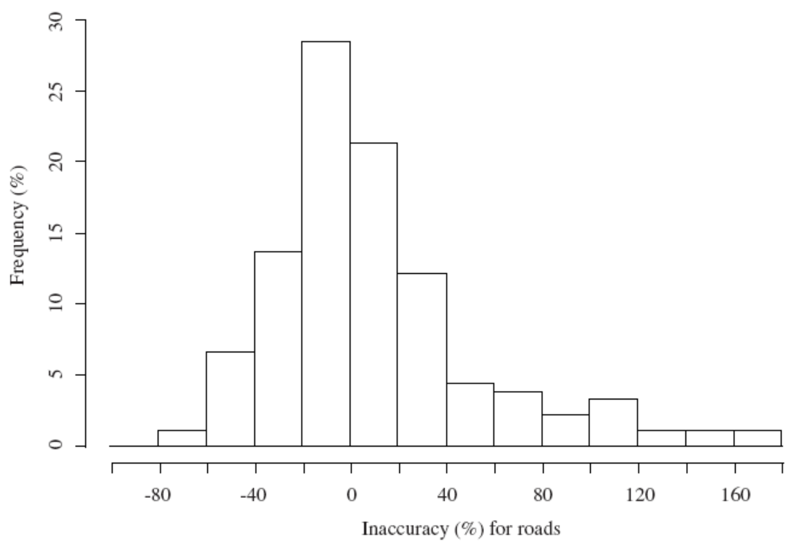

Figure 3: Distribution of actual-to-forecast traffic volume ratios for non-tolled road projects (Source: Flyvbjerg et al. 2006, Figure 1)

In Standard & Poor's 2005 update (Bain and Polakovic 2005), the uncertainty in project ramp-up years[3] was investigated in greater depth. The expectation is that uncertainty falls slightly from opening year forecasts, since traffic demand would have an opportunity to stabilize, as drivers learn of route alternatives and obtain toll accounts, for example. The sample size was just 25 projects for years 1 through 5, and the hypothesis was not supported (Bain and Polakovic 2005). The mean ratio (of actual-to-forecast traffic volumes) was 0.77 in year 1, and 0.79 (negligibly higher) in year 5. These results suggest that traffic performance generally remains much less than forecast, even into year 5 of operation. While Vassallo and Baeza's (2007) much smaller sample (of Spanish toll roads) identified similar optimism biases, forecast ratios generally

[3] Ramp-up years are those immediately following opening year.

improved following year one. So there is room for differences in average results, due to regional economic conditions, marketing campaigns or other factors.

SOURCES OF RISK AND UNCERTAINTY IN TRAFFIC FORECASTS

While significant uncertainty in traffic forecasts clearly exists, the causes of such uncertainty vary. Numerous studies have identified and examined several sources of forecast error (see e.g., Flyvbjerg et al. 2006 and 2006, Bain and Wilkins 2002, George et al. 2003, George et al. 2007), and for the most part, these are similar for tolled and non-tolled highways, but differences do exist.

Flyvbjerg et al. (2005 and 2006) interviewed project managers who identified a variety of sources, including several travel demand modeling components. The two top-stated sources of error for toll-free road projects are estimates of trip generation and land development, though trip distribution and the forecasting model are close runners-up. Flyvbjerg et al. (2005 and 2006) attribute much of the modeling uncertainty to dated data used in model calibration. Land Transport New Zealand (2006) also notes the importance of quality and relevance of data used in the forecasting model.

Zhao and Kockelman (2002) tracked the propagation of uncertainty through a four-step travel demand model. They controlled the uncertainty of model inputs and parameters, and performed 100 simulations of the model. Assuming coefficients of variation (CoVs) of 0.3 in all model inputs and parameters, Figure 4 illustrates the range of CoVs in intermediate and final model outputs (across the 100 simulations), including 5% and 95% bounds on these. Figure 6 suggests that modeling error in effect “grows” through the application of trip generation, trip distribution, and mode choice models (as one’s scale of resolution gets finer, essentially − to the number of trips by mode between each origin-destination pair). However, the final step of traffic assignment enjoys a drop in uncertainty (at the link-flow level), thanks to overlap in different trips’ routings and mode and trip distribution choices across all travelers, along with congestion feedbacks (which moderate the presence of high link-demand values). Overall, Zhao and Kockelman’s (2002) work suggests that link-flow estimates enjoy the same level of uncertainty as inputs and parameters, and simple regressions of outputs on inputs (and aggregations of inputs) offer very high predictive power, suggesting that prime sources of forecast uncertainties can be rather quickly deduced − and exploited, for better prediction.

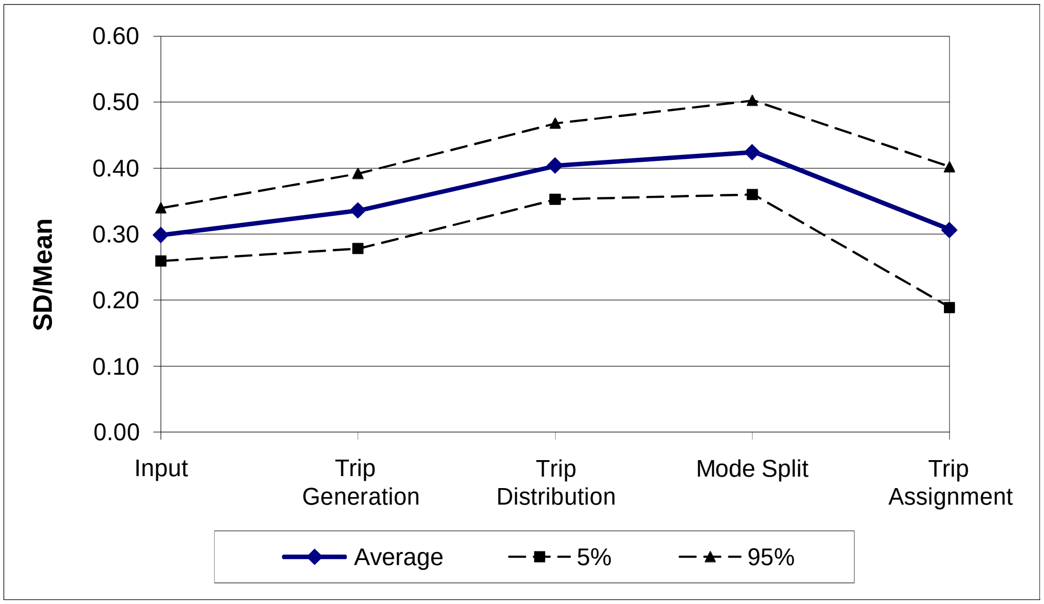

Figure 4: Uncertainty propagation through a four-step travel demand model (Source: Zhao and Kockelman 2002, Figure 5)

Zhao and Kockelman (2002) also point out that models are abstractions of reality and the entire modeling paradigm is a source of error in traffic forecasts. While their study did not consider tolled roads, one can imagine that output variability may rise, as toll-technology adoption rates and heterogeneity in value of travel time savings introduce more uncertainty. In fact, for tolled roads, Bain and Wilkins (2002) noted the importance of data used to calibrate travel demand models, both in terms of currency (more recent is better) and the ease with which data were collected (affecting data quality and quantity).

Network attributes can also play a key role in forecast reliability. Analysts do not know the actual future network, and coded networks are significant simplifications of actual networks (generally ignoring local streets, signal timing plans, turning lane presence and lengths, etc.). Forecasts that depend on future network changes (such as nearby highway extensions) tend to be less reliable (Bain and Wilkins 2002). Traffic congestion is also key. As noted by Bain and Wilkins (2002) and Zhao and Kockelman (2002), uncongested networks often are more difficult to anticipate flows on, since congestion feedbacks distribute traffic more evenly over space and time while establishing something like an upper bound (due to inherent capacity limitations) on all links. Thus, low-volume corridors tend to have greater uncertainty in their forecasts (Bain and Wilkins 2002).

Another key source of error in traffic forecasts comes from uncertainty in land development patterns (Rodier 2003, Flyvbjerg et al. 2005 and 2006, Land Transport New Zealand 2006). Rodier's (2003) application of the Sacramento, California travel demand model for year 2000 conditions found that about half of the 11-percent overestimation of VMT was due to demographic and employment projections, which serve as inputs to the demand models. The other half was due to the model itself. With forecasts anticipating demand 10-plus years out, Flyvbjerg et al. (2005 and 2006) suggest that more of the error may stem from uncertainty in coming land development patterns. For tolled roads, Bain and Wilkins (2002) argue convincingly that land development forecasts are regularly critical, and the more stable a region's economy, the better its land-use (and, thus, its travel demand) forecasts. Such forecasts are generally based on land use plans and expert judgment, which are simply educated guesses and tend to evolve over time. Another option is land use modeling, which, of course, is also fraught with a variety of uncertainties (see e.g., Pradhan and Kockelman 2002, Rodier and Johnston 2002, Krishnamurthy and Kockelman 2003, Rodier 2005, Clay and Johnston 2006, Sevcikova et al. 2007, and Kockelman et al. 2008).

While the sources of error described above apply for projects of any type, there are many others that are rather specific to tolled roads. One such source identified by Bain and Wilkins (2002) and George et al. (2007) is tolling design − i.e., whether shadow tolls or user-paid tolls[4] are used. With shadow tolls, the government pays the concessionaire an amount based on toll road use. So from the user perspective, it is very similar to a toll-free road. With user-paid tolls, the toll charge is quite transparent to the user. Since driver willingness to pay is more complex and difficult to understand, projects with user-paid tolls carry more forecasting risk. Moreover, George et al. (2007) suggest that user fees make a tolled road more susceptible to changes in demand caused by economic downturns/recessions, toll rate increases, and escalating fuel costs. Other special or relatively rare events (e.g., natural disasters or acts of terrorism among other events) are often key sources of uncertainty as well (George et al. 2007). Of course, such events are difficult to predict, though HLB Decision Economics (2004) suggests that the number and duration of recessions in the forecast period should be considered in investment grade studies.

Another important consideration in understanding project risk is the "tolling culture" of a region (Bain and Wilkins 2002). This is essentially the degree to which tolls have been used in the past. In nations and regions where tolling has not previously been used, there is greater uncertainty surrounding traffic forecasts. If travelers are accustomed to paying tolls for other road facilities, forecasts tend to be much more reliable. As noted earlier, this appears to result in 20% greater average optimism bias (Bain and Plantagie 2003).

Of course, travel demand model imperfections are a key source of error in traffic forecasts. For instance, the robustness and heterogeneity (across travelers and trip types) of value of travel time (VOTT) estimates are generally ignored, but may be crucial in producing accurate forecasts. The use of imported parameters (calibrated for other regions or even other countries) can also cause much error (Bain and Wilkins 2002). Another important issue in modeling deals with how the actual tolls are modeled. If a complex tolling regime is to be used (e.g., variable tolls or HOT lanes that are free at certain hours), models fully recognizing such complexity can be quite difficult to specify and calibrate (Bain and Wilkins 2002), introducing further uncertainty.

Facilities enjoying a competitive advantage of some sort also tend to offer more reliable forecasts (Bain and Wilkins 2002, George et al. 2007). For instance, forecasts for projects in dense, urban networks (with many alternative routes) generally will be less certain than those for projects with a clear competitive advantage over alternatives (e.g., a corridor with the only river crossing in a region). Moreover, many privately financed projects rely on protection against competition in

[4] Only 4 of the 32 projects investigated in the 2002, Bain and Wilkins study had shadow tolls.

the future. If protection is provided (via non-compete clauses, for example), long-run traffic forecasts tend to be more reliable (Bain and Wilkins 2002). Of course, such clauses may be contentious, as discussed in Perez and Sciara (2003), Poole (2007), and Ortiz et al. (2008). However, non-compete clauses generally do not ban planned improvements (Ortiz et al. 2008) and typically do not prohibit new free roads. But they may allow for compensation when toll revenues fall due to improvements on nearby non-tolled facilities (Poole 2007).

Meaningful distinctions can also arise in the context of user attributes. Bain and Wilkins (2002) assert that toll facilities serving mostly a small market segment of travelers allow for more reliable traffic forecasts. This is because smaller markets are easier to model than more heterogeneous populations (Bain and Wilkins 2002). For example, beltways (orbital style facilities) are likely to carry more forecasting risk than radial facilities (which typically carry a high share of commuters into and out of the city center, for work purposes). In addition, if there is a single origin-destination (O-D) pair that constitutes the majority of trips made on the facility, forecasts errors fall, as a result of the relatively homogeneous makeup of such travelers. However, George et al. (2007) warn that, when only a small market segment constitutes the majority of toll road users, the road's traffic and revenues will be more susceptible to any forms of downturn affecting that small segment.

Of course, road location and configuration also affect levels of forecast error. When the preferred alignment of a new tollway is constrained by external factors (e.g., land use patterns, nature and location of existing development, land/right-of-way availability, topography, geological sensitivities, engineering limitations, and politics), traffic forecasts become more uncertain (Bain and Wilkins 2002). Bain and Wilkins (2002) also assert that facilities with proper connectors to the rest of the network have more reliable estimates. If the toll road terminates in the downtown area and long queues await travelers joining the local network and/or if travelers must take circuitous routes to enter the tollway, the competitive advantage of the toll road can be compromised, and greater forecast errors can emerge. Demand variations over times of day and days of the year also affect forecast reliability. If a road serves a stable demand profile, forecasts tend to be more reliable (Bain and Wilkins 2002). Commercial users of the tolled facility also can play an important role. In particular, if most commercial vehicles are independent truckers, there is added risk in traffic forecasts since their behavior is less well understood. However, if most commercial truckers work for fleet owners, the opposite is true. (Bain and Wilkins 2002) Moreover, dependence on commercial travel carries more risk since commercial travel is more susceptible to economic downturns (George et al. 2007)

Overall, Bain and Wilkins (2002) indicate seven top drivers of forecast failure: poorly estimated VOTTs, economic downturns, mis-prediction of future land use conditions, lower-than-predicted time savings, added competition (e.g., improvements to competing roads or the addition of new roads), lower than anticipated truck usage, and high variability in traffic volumes (by time-of-day or day of the year). Bain and Plantagie (2003) added several other top drivers: complexity of the tolling regime, underestimation of the duration and severity of the ramp-up period, and reliance on a single VOTT (as opposed to segmenting user groups). However, it did not offer any information regarding the magnitude of added uncertainty for each of these features. Another rating agency, Fitch Ratings (George et al. 2003), also suggested several of these same drivers, but added that the use of a regional travel demand model developed for other planning purposes also can cause great error in traffic forecasts. (Such was the case for the San Joaquin toll road forecast in Orange County, California. [George et al. 2003]) This suggests, to some extent, that a comprehensive, regional model may not perform as well as simpler estimation techniques (e.g., OD pair trend analysis), if the regional model lacks appropriate specification for the project scenario. Clearly, there is a great deal of uncertainty in traffic and revenue forecasts of tolled roads stemming from various sources. The next section discusses methods that can be used to measure and evaluate this uncertainty in forecasting models.

METHODS OF ACCOMMODATING RISK IN TDM AND REVENUE ESTIMATION ANALYSES

Accommodating risk and uncertainty in demand and revenue forecasts is an important component of any toll road study. While a single "best" statistical forecast is useful, it lacks the information needed for making long-term financial decisions. With the great number of assumptions, inputs, and estimated parameters entering travel demand models, model outputs can be highly uncertain and inaccurate. Neglecting this uncertainty (or equivalently, assuming determinism) can invite scrutiny from stakeholders, since not all will agree with assumed inputs and parameter values (Duthie 2008). As noted in the previous sections, the magnitude of error in demand forecasts (and, thus, revenue forecasts) can be substantial, and tends to be biased in favor of toll road projects. Even with advances in model designs over the past couple decades, Flyvbjerg's review of the data suggests that forecast accuracy has not improved and may have worsened (Flyvbjerg et al., 2006). Most analysts, policy-makers, and investors agree that it is imperative that modelers quantify forecasting risk in a meaningful way (Rodier 2007), and while the financial community has understood the need to address risk in toll road studies, Kriger et al. (2006) believe that very few practitioners conduct any sort of risk assessment. Some simply verify results by use of "reality checks" (e.g., comparing to older forecasts and using simple intuition to verify whether results seem reasonable) while others use no verification methods at all.

One key component of risk assessment in model outputs lies in explicitly stating all modeling assumptions (Kriger et al., 2006), making the model specification as transparent as possible. If modelers and users understand the implications of alternative assumptions, the uncertainty in the forecasting process will be better understood. Of course, other options for understanding and communicating forecast uncertainty also exist, as discussed here now.

A relatively common and reasonably effective method for accommodating risk in demand and revenue forecasts is the use of sensitivity analyses or "stress tests" (Kriger et al., 2006). Most sensitivity analyses rely on the exploration of a very limited set of different values for key variables, such as a region's or neighborhood's population growth rate, values of travel time, and planned tolls (Kriger et al., 2006). Though such analyses can provide key insights, many practitioners and financial analysts feel that they inadequately reveal the range of possible outcomes (see, e.g., HLB Decision Economics 2003 and Kriger et al. 2006). As their name implies, stress tests seek to understand the outcomes of relatively extreme conditions − generally to anticipate worst- (and best-) case investment scenarios. In this way they help analysts anticipate lower (and upper) bounds on project outcomes, but certainly not a distribution of outcomes, or probability of financial loss.

Model validation studies offer another method for quantifying uncertainty, by examining how well model forecasts match observed data not used in model calibration (Rodier 2007). Such studies measure forecast uncertainty directly from observed data, and thus require data from two points in time: the older data set is used for model estimation and calibration while the newer one is used for validation. It can be impossible to conduct such tests of models developed from recent data, but at least one obtains a sense of the magnitudes of errors that can emerge from transferring behavioral parameters calibrated on old data to current-year contexts. Such validation tests are a valuable complement to sensitivity tests. And such results assist analysts in communicating of the size and relevance of uncertainty to decision makers and the public (Rodier 2007).

Of course, sensitivity testing and model validation studies have their limitations. For example, sensitivity tests are quite constrained, to typically three or four scenarios. In contrast, Monte Carlo simulation techniques more fully explore the range of possible outcomes, by defining and drawing from probability distributions for key inputs. Moreover, such techniques are not new: Ashley (1980) and Lowe et al. (1982) investigated the sensitivity of forecast traffic volumes to model inputs and parameters using Monte Carlo methods; and, more recently, Lam and Tam (1998), Boyce and Bright (2003), Zhou and Kockelman (2002), Beser Hugosson (2005), and de Jong et al. (2007) performed similar analyses. Of course, such techniques also exhibit limitations: They require assumptions of input distributions (and their covariances), when these are often unknown, and generally more sophisticated programming techniques (to ensure rapid run times for testing a high number of scenarios). And as model complexity increases, model run times invariably grow. Under fixed time constraints, an increase in model complexity generally implies fewer Monte Carlo simulations. Nonetheless, these simulation methods can be invaluable for a proper understanding of uncertainty in traffic and revenue forecasts.

Monte Carlo techniques are at the heart of the four-step risk analysis process (RAP) used by HLB Decision Economics (2003). In step 1, HLB defines a “structure and logic” model, in order to forecast traffic and revenue on the basis of an array of inputs and parameters. In step 2, central estimates and probability ranges are assigned to each relevant input and parameter. In step 3, expert opinions regarding the results of step 2 are obtained, and probability ranges and central estimates are revised. In the final step, Monte Carlo simulation techniques are employed, drawing inputs and parameters from their respective probability distributions, and traffic and revenue probability ranges are derived based on the simulation outcomes. (HLB Decision Economics 2003) This approach allows firms like HLB to determine the likelihood that revenue cannot cover the debt service, an important criteria for issuance of debt.

As discussed earlier, Zhao and Kockelman (2002) performed a similar analysis (for a non-tolled case), using a four-step travel demand model for a sub-network of the extensive Dallas-Fort Worth region with 118 variable input and parameter values. Due to typical time constraints on their research, only 100 runs were undertaken (using TransCAD software). Using more streamlined networks or hands-on programming would have allowed for more runs, which can be critical in cases of many uncertain inputs (assuming there is a potential for highly nonlinear model behaviors, for example), since the parameter space grows exponentially with added inputs. In practice, most toll road studies use highly streamlined networks and simplified model specifications (see e.g., Lam and Tam 1998, HLB Decision Economics 2003), to facilitate computation while recognizing that the majority of a region's network is largely irrelevant when evaluating a relatively short new, tolled corridor. Nonetheless, Zhao and Kockelman's analysis provides useful insights into the degree of uncertainty in link- and region-level traffic forecasts. They assigned density functions to the 18 random model parameters (13 in trip generation, 1 in trip distribution, 2 in mode choice, and 2 in assignment) and four major model inputs for each of 25 zones (household counts along with basic, retail, and service job counts). Each of the uncertain parameters and inputs were assumed to follow log-normal distributions with coefficients of variation[5] (CoVs) of 0.3, 0.1, and 0.5. After performing 100 simulation runs (for each of the 3 CoVs), two network links were examined in detail for the case of CoVs equal to 0.3. On both links, flows ranged from around 400 vehicles per hour to over 2000, with CoVs of 0.31 and 0.32. Zhao and Kockelman (2002) also performed a regression analysis of standardized input and parameter values on system-level VMT results. This analysis indicated that inputs and trip generation parameter values were the most important factors in forecasts of total VMT. It seems evident that traffic forecasts can exhibit a great deal of variation and depend greatly on parameter and input assumptions used in model calibration and application. When tolls are present, results could exhibit even greater variation. However, Zhao and Kockelman (2002) observed similar uncertainty levels in model inputs and outputs suggesting that opportunities for errors in one part of the model to offset errors in another can have a dampening effect on overall uncertainty. Thus, adding more uncertain inputs and/or parameters may not amplify forecast uncertainty.

Lam and Tam (1998) also performed a study of uncertainty using Monte Carlo draws in traffic and revenue forecasts for a toll road project connecting Hong Kong to an adjacent region separated by a body of water. No actual travel demand model was used, however, since only one other reasonable route existed between the two regions and a detailed travel study was deemed unnecessary. Instead, trip generation and routing shares were assigned distributions, and allowed to vary across simulation runs in order quantify forecast uncertainty. A total of 10,000 simulations were performed, and overall revenues were found to hit or exceed the base forecast approximately 52% of the time. This is not so surprising, since the base forecast represents a simulation based on the mean values for all 12 unknowns. They also estimated that the standard deviation of forecast revenues rose from just 17% of the mean in the first forecast year to 28% of the mean after 20 years (Lam and Tam, 1998). It is useful to note the smaller coefficients of variation found here, in comparison to Zhao and Kockelman's (2002) study. For instance, the total population and trip generation rates were both assumed to have CoVs of 0.05. Lam and Tam investigated a particular scenario with arguably much less risk. Since their bridge facility enjoyed a clear advantage over competing routes, there was a specific traveler group being serviced, and a single origin-destination pair making up the majority of travel.

As noted earlier, land use is an important determinant of long-run traffic levels, particularly when new highways are provided in largely undeveloped locations. Land use change can be more difficult to predict and more variable in the long run than traffic volumes, and new work is emerging in this area. For example, Sevcikova et al. (2007) recently compared Bayesian Melding techniques and standard sampling approaches to analyze uncertainty in projections of household counts using UrbanSim. And Kockelman et al. (2008) used an antithetic sampling

[5] The coefficient of variation is defined as the ratio of the standard deviation to the mean.

technique to analyze uncertainty in an integrated land use-transportation setting. Methods like these, for sampling thoughtfully and performing estimation rapidly, could potentially aid in obtaining output distributions from complex models relatively quickly.

Consistent with such analyses, the National Federation of Municipal Analysts (NFMA 2005) formally recommends that a range of possible road project and policy outcomes should be explored based on different scenarios (or assumptions), and that varying variables or parameters one at a time is insufficient. By assigning realistic probability distributions to parameter values and inputs, the probability of a given scenario can be understood. The NFMA's (2005) guidelines for traffic and revenue studies include several highlights: a no-build traffic forecast should be produced, a baseline traffic and revenue forecast should be produced, sensitivity analyses should be performed on inputs (including population, employment, and income growth, toll elasticity by consumers, and acceleration of the planned transportation network), and debt service analysis should be performed.

Of course, just as neglecting uncertainty is equivalent to assuming determinism, neglecting covariance in inputs is equivalent to presuming their independence. Thus, it is important to recognize the co-dependence of input distributions due to correlated response under various conditions and as introduced in parameter distributions via the estimation process. For example, economic boom/bust cycles can affect land development and thus population and job growth across zones similarly, along with trip generation rates, vehicle ownership, and income levels. This can result in wider uncertainty bounds than univariate input and parameter distributions would indicate. For example, Zhao and Kockelman (2002) used multivariate distributions for their population and employment input values with +0.30 correlations, but relied on independent distributions for all model parameters.

Another approach is "reference class forecasting," as described by Flyvbjerg et al. (2005). This method essentially relies on past experiences with a sample of similar projects in order to estimate outcome distributions and thus the probability of various events occurring. By comparing the forecasts with past experience, judgments can be made regarding the validity of results. Of course, this is difficult to do without good data on a variety of reasonably comparable projects. But it is a useful strategy when such data exist.

To determine an investment's credit rating, credit agencies and financial analysts use varied approaches to account for revenue forecast risk. For example, Fitch Ratings (George et al. 2003, George et al. 2007) claims to study the key assumptions and inputs of the travel demand model used in creating future forecasts, and then considers a range of possible outcomes associated with each factor in order to develop a "stress" scenario alongside a base scenario (essentially sensitivity testing, but with relatively extreme scenarios). The base case is generally more conservative than the base case developed by the project sponsor, eliminating any evident forecast optimism. The stress case is developed to determine the project's ability to withstand rather severe (but not unreasonable) circumstances in which the ability to pay debt service is stressed. Based on the results of the stress scenario, an investment rating is assigned to the project. For credit analysis of longer-term traffic forecasts, Bain et al. (2006) suggest taking a conservative approach, reducing growth rate expectations and carefully examining future toll schedule increases. They also suggest that long-term growth rates exceeding 1% and toll increases beyond those suggested by reasonable correction for inflation should be viewed with caution. While these techniques simplify uncertainty testing dramatically and help investors understand the real possibility of loss, they do not illuminate the variety (and likelihood) of futures that truly exist, and associated investment risk cannot be fully understood using such methods.

SUMMARY AND RECOMMENDATIONS

As discussed in this paper, a great deal of uncertainty exists in traffic forecasts. Flyvbjerg's analyses (2005 and 2006) suggest that traffic forecast errors exceed 20% roughly half the time across all roadway projects and more than 40% of the time for a quarter of projects. This situation is compounded when traffic forecasts of *tolled* projects are considered, since more unknowns exist. S&P's analysts (Bain and Wilkins 2002, Bain and Plantagie 2003 and 2004) found that, on average, tolled traffic volumes are well below forecasts (on the order of 25% or more) in their first year of operation, suggesting considerable optimism bias, and that this bias does not fade over time. As transportation agencies look more closely at tolling options as a way to fund highway capacity expansion and manage demand, it becomes even more important that models provide reliable traffic forecasts.

Traditionally, travel demand models have been used to provide a single projection of future conditions. Though the models become more sophisticated, the future remains unknown, and model forecasts should be presented as such. It is critical that the uncertainty implicit in travel demand models be communicated to planners and policy makers. Of course, quantifying such uncertainty is not a trivial task. While the sources of misprediction vary, designers and transportation planners have found a number of methods to accommodate forecast uncertainty (or at least quantify it).

Sensitivity testing allows for greater understanding of the magnitudes of uncertainty in the model. By allowing key model inputs and parameters to vary simultaneously, creating multiple possible scenarios, uncertainty in traffic and revenue forecasts can be better bounded. Indeed, this appears to be the most common method for dealing with uncertainty by credit agencies. However, sensitivity testing generally does not provide a probability of particular outcomes occurring. Therefore, it can be difficult for policy makers to truly understand inherent risks. When feasible, comparisons with similar, past projects is a meaningful tool for anticipating potential outcomes.

Monte Carlo simulation may be most appropriate to identify a more comprehensive set of possible futures. By drawing parameters and inputs from reasonable sets of distributions, the probability of particular outcomes can be understood. Of particular importance for projects where financial backing is dependent on toll revenues is the probability that toll revenues will cover debt service, and whether additional revenues will remain (over and above debt service). Moreover, since most toll road studies use rather streamlined model systems, computing time is typically not an issue. Thus, the recommended best practice for dealing with uncertainty in toll road projects is the use of Monte Carlo simulation. Sensitivity testing is valuable in some cases where simulation may be too computationally expensive, though more thoughtful sampling methods, such as Bayesian melding and antithetic sampling, can reduce such computational burden in many cases.

ACKNOWLEDGEMENTS

The authors thank several anonymous reviewers, Mr. Peter Vovsha of Parsons Brinckerhoff, Inc. (PB), Robert Bain at the University of Leeds, and Jennifer Duthie for their useful suggestions and Ms. Annette Perrone for editorial assistance. We also want to thank the National Cooperative Highway Research Program (NCHRP) for financially supporting this study under Project HR 08-57, titled "Improved Framework and Tools for Highway Pricing Decisions."

REFERENCES

Ashley, D.J. (1980) Uncertainty in the Context of Highway Appraisal. *Transportation*, 9, 249-267.

Bain, R. and M. Wilkins (2002) Infrastructure Finance: Traffic Risk in Start-Up Toll Facilities, Standard & Poor's, McGraw-Hill International (UK) Ltd., September 2002.

Bain, R. and J.W. Plantagie (2003) Traffic Forecasting Risk: Study Update 2003, Standard & Poor's, McGraw-Hill International (UK) Ltd., November 2003.

Bain, R. and J.W. Plantagie (2004) Traffic Forecasting Risk: Study Update 2004, Standard & Poor's, McGraw-Hill International (UK) Ltd., October 2004.

Bain, R. and L. Polakovic (2005) Traffic Forecasting Risk Study Update 2005: Through Ramp-Up and Beyond, Standard & Poor's, McGraw-Hill International (UK) Ltd., August 2005.

Bain, R., K. Forsgren, and P.B. Calder (2006) Credit FAQ: Assessing the Credit Quality of Highly Leveraged Deep-Future Toll-Road Concessions, Standard & Poor's, McGraw-Hill International (UK) Ltd., February 2006.

Beser Hugosson, M. (2005) Quantifying Uncertainties in a National Forecasting Model. *Transportation Research Part A*, 39 (6), 531-547.

Boyce, A.M., and M.J. Bright (2003) Reducing or Managing the Forecasting Risk in Privately-Financed Projects. Paper presented at the European Transport Conference, Strasbourg, France.

Clay, M.J. and R.A. Johnston (2006) Multivariate Uncertainty Analysis of an Integrated Land Use and Transportation Model: MEPLAN. *Transportation Research Part D*, 11, 191-203.

de Jong, G., A. Daly, M. Pieters, S. Miller, R. Plasmeijer, and F. Hofman (2007) Uncertainty in Traffic Forecasts: Literature Review and New Results for The Netherlands. *Transportation*, 34, 375-395.

Duthie, J., A. Voruganti, K. Kockelman, and S.T. Waller (2008) Uncertainty Analysis and its Impacts on Decision-Making in an Integrated Transportation and Gravity-Based Land Use Model. Paper submitted for presentation to the 88th Annual Meeting of the Transportation Research Board, January 2009, Washington, D.C., and for publication in *Transportation Research Record*.

Duthie, J. (2008) Implications of Uncertain Future Network Performance on Satisfying Environmental Justice and Tolling. Doctoral dissertation, Department of Civil, Architectural, and Environmental Engineering, University of Texas at Austin.

Flyvbjerg, B., M.K. Skamris Holm, and S.L. Buhl (2005) How (In)accurate Are Demand Forecasts in Public Works Projects: The Case of Transportation. *Journal of the American Planning Association*, 71 (2), 131-146.

Flyvbjerg, B., M.K. Skamris Holm, and S.L. Buhl (2006) Inaccuracy in Traffic Forecasts. *Transport Reviews*, 26 (1), 1-24.

George, C., W. Streeter, and S. Trommer (2003) Bliss, Heartburn, and Toll Road Forecasts. Project Finance Special Report, Fitch Ratings, November 2003.

George, C., S. Trommer, M. McDermott, G. Zurita, C. Lewis, L. Monnier, W. Streeter, E. Lopez, and C. Fuenalida (2007) Global Toll Road Rating Guidelines. Criteria Report, Fitch Ratings, March 2007.

HLB Decision Economics, Inc. (2003) S.R. 125 South Toll Road Project: Risk Analysis of Traffic and Toll Revenue Forecasts. Draft Final Report, Washington, D.C.

HLB Decision Economics, Inc. (2004) Toll Road Project Risk Analysis. Presentation for MBIA Insurance Corporation.

Kockelman, K.M., J. Duthie, S.K. Kakaraparthi, B. Zhou, A. Anjomani, and S. Marepally (2008) An Examination of Land Use Models, Emphasizing UrbanSim, TELUM, and Suitability Analysis. Research Report 0-5667-1, Texas Department of Transportation, Center for Transportation Research, The University of Texas at Austin, Austin, TX.

Kriger, D., S. Shiu, and S. Naylor (2006) *Estimating Toll Road Demand and Revenue: A Synthesis of Highway Practice*. NCHRP Synthesis 364, Transportation Research Board, Washington, D.C.

Krishnamurthy, S. and K.M. Kockelman (2003) Propagation of Uncertainty in Transportation-Land Use Models: Investigation of DRAM-EMPAL and UTPP Predictions in Austin, Texas. *Transportation Research Record No. 1831*, 219-229.

Lam, W.H.K. and M.L. Tam (1998) Risk Analysis of Traffic and Revenue Forecasts for Road Investment Projects. *Journal of Infrastructure Systems*, 4 (1), 19-27.

Land Transport New Zealand (2006) Risk Analysis. Appendix A13 of 'Economic Evaluation Manual: Volume 1,' Land Transport New Zealand. http://www.ltsa.govt.nz/funding/economic-evaluation-manual/eem1-1.pdf.

Lowe, S., D. Morrell, G. Copley (1982) Uncertainties in Highway Appraisal: The Development of Systematic Sensitivity Testing. Paper presented at the Planning and Transport, Research and Computation (PTRC) Summer Annual Meeting (SAM), University of Warwick, UK.

NFMA (2005) Recommended Best Practices in Disclosure for Toll Road Financings. National Federation of Municipal Analysts. Retrieved July 2008 from http://www.nfma.org/disclosure/rbp_toll_road.pdf.

Ortiz, I.N., J.N. Buxbaum, and R. Little (2008) Protecting the Public Interest: The Role of Long-Term Concession Agreements for Providing Transportation Infrastructure. Proceedings of the 87th Annual Meeting of the Transportation Research Board, Washington, D.C., January 13-17.

Perez, B.G. and G.C. Sciara (2003) A Guide for HOT Lane Development. Report prepared by Parsons Brinckerhoff with the Texas Transportation Institute (TTI) for the Federal Highway Administration (FHWA), Publication Number FHWA-OP-03-009.

Poole, R.W. (2007) Tolling and Public-Private Partnerships in Texas: Separating Myth from Fact. Reason Foundation Working Paper, May, 2007. Retrieved August 2007 from www.reason.org/TX_toll_roads_working_paper.pdf.

Pradhan, A. and K.M. Kockelman (2002) Uncertainty Propagation in an Integrated Land Use-Transportation Modeling Framework: Output Variation via UrbanSIM. *Transportation Research Record No. 1805*, 128-135.

Rodier, C.J. (2003) Verifying the Accuracy of Regional Models Used in Transportation and Air Quality Planning. Mineta Transportation Institute (MTI) Report 02-03, College of Business, San Jose State University. Retrieved June 2008 from http://transweb.sjsu.edu/mtiportal/research/publications/documents/02-03.pdf.

Rodier, C.J. (2005) Verifying the Accuracy of Land Use Models Used in Transportation and Air Quality Planning: A Case Study in the Sacramento, California Region. Mineta Transportation Institute (MTI) Report 05-02, College of Business, San Jose State University. Retrieved June 2008 from http://transweb.sjsu.edu/mtiportal/research/publications/documents/05-02.pdf.

Rodier, C.J. (2007) Beyond Uncertainty: Modeling Transportation, Land Use, and Air Quality in Planning. Mineta Transportation Institute (MTI) Report 07-01, College of Business, San Jose State University. Retrieved June 2008 from http://transweb.sjsu.edu/mtiportal/research/publications/documents/07-01.pdf.

Rodier, C.J. and R.A. Johnston (2002) Uncertain Socioeconomic Projections Used in Travel Demand Model and Emissions Models: Could Plausible Errors Result in Air Quality Nonconformity?" *Transportation Research Part A*, 36, 613-631.

Sevcikova, H., A.E. Raftery, and P.A. Waddell (2007) Assessing Uncertainty in Urban Simulations Using Bayesian Melding. *Transportation Research Part B*, 41, 652-669.

Vassallo, J.M. and Baeza, M.A. (2007) Why Traffic Forecasts in PPP Contracts Are Often Overestimated? Research Paper, Final Draft, EIB University Research Sponsorship Programme. European Investment Bank, Luxembourg.

Zhao, Y. and K.M. Kockelman (2002) The Propagation of Uncertainty through Travel Demand Models: An Exploratory Analysis. *Annals of Regional Science*, 36 (1), 145-163.