# Public Private Partnerships in California

**Phase II Report**

**Section III: Analytical Tools**

**July 2012**

**Research Team**
Dr. Genevieve Giuliano, Principal Investigator

Dr. Lisa Schweitzer, Co-Principal Investigator

Kevin Holliday, Project Administrator

Teddy Minch, Research Assistant

**Sol Price School of Public Policy**
**University of Southern California**
**Los Angeles, CA 90089-0626**

## Disclaimer

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, and California Department of Transportation in the interest of information exchange. The US Government and California Department of Transportation assume no liability for the contents or use thereof. The contents do not necessarily reflect the official views or policies of the State of California or the Department of Transportation. This report does not constitute a standard, specification, or regulation.

# Table of contents

## List of tables

# Table of figures

## Acronyms and terms defined

In the following table we outline the acronyms and terms we use in the report

**Table 1: Definition of acronyms and terms**

| Term or Acronym | Definition |
|---|---|
| DB | Design-Build. Contract type in which one private firm is responsible for both design and construction of facility, instead of one firm for each individual phase. |
| DBFO | Design-Build-Finance-Operate. Contract type in which private partner is responsible for all aspects of the project except maintenance. |
| DBFOM | Design-Build-Finance-Operate-Maintain. Contract type in which private partner is responsible for all components of the project from beginning to end. |
| DBOM | Design-Build-Operate-Maintain. Contract type in which private partner is responsible for all aspects of the project except financing. |
| FHWA | US Federal Highway Administration |
| H.M. Treasury | Her Majesty's Treasury, the United States Department of Treasury functional equivalent within the United Kingdom. |
| LCC | Life Cycle Costs. The aggregated cost of all inputs required to design, construct, and maintain an infrastructure asset. |
| MPO | Metropolitan Planning Organization. For example, SCAG in southern California or MTC in the San Francisco Bay Area. |
| NPV | Net Present Value |
| O&M | Operations and Maintenance |
| P3 or P3s | Public-Private Partnership(s) |
| PSC | Public Sector Comparator A hypothetical, risk-adjusted cost estimate for a given project where the asset is delivered exclusively by the public sector through traditional procurement. |
| Public Sponsor | As used in this report, public sponsor refers to any public agency that might propose, build, and/or maintain a transportation facility, e.g., SFMTA, Metro, Caltrans, or Gold Line Phase II Construction Authority. |
| Tour | A series of three or more trip segments – any trip that does not follow an X to Y to X progression. Tours involve, for example, a driver going from home to the coffee shop to work. Such inputs to demand models provide much more layering than can be generated through the traditional four-step model. |
| USDOT | US Department of Transportation |
| VfM | Value for Money. A means of evaluating future cash flows to determine whether a capital project is best-suited for traditional/public procurement or through alternative procurement like a P3. |

# Introduction

In this section of the report, we cover the analytical tools necessary to evaluate and plan for priced facilities, principally toll roads, delivered as P3s. In addition to describing the methods – laying out technical descriptions of the analytical tools required for priced facilities – we review past performance of those tools and offer a few suggestions on improving the entire forecasting process for priced facilities. The report is divided into three sections: 1) analytical tools for demand-side factors of priced facilities (like road pricing methodology and demand/revenue forecasting); 2) analytical tools for supply-side factors of priced facilities (like construction and life cycle cost forecasting); and 3) an assessment of the drivers of forecast inaccuracy on both the demand and supply side, and how the entire forecasting process can be improved for priced facilities.

No evidence in the literature suggests that entirely new analytical tools and methodologies must be developed in order to assess the feasibility of P3s. No literature explicitly addressed P3-specific tools; rather, the literature focused on advanced applications of and modifications to existing tools and approaches. Existing tools, and modifications to them, are employed in assessing P3 feasibility; but the dual importance of risk and sensitivity analysis is amplified for analyses of these types of projects. Because of the increased number of stakeholders involved in the financing and management of P3s as compared to traditionally-procured projects, P3 feasibility studies and forecasts require an increased understanding of the range of potential project outcomes. In this sense P3s do not require entirely new forecasting models or tools so much as they do augmented focus in accurately and meaningfully modeling risk and contingencies, given the increased scrutiny these forecasts will face from financiers and the project finance community at large (Kriger 2005, p. 90).

This report focuses on state-of-the-practice analytical tools, rather than state-of-the-art analytical tools, because state-of-the-art forecasting has yet to be employed consistently or meaningfully for P3 projects on a broad basis. State-of-the-art and state-of-the-practice methodologies represent two very different types of tools. State-of-the-practice tools are those which can be employed "off the shelf," while state-of-the-art tools often require advanced training or individuals with a strong mathematics background. Furthermore, these techniques are often experimental, generally limited in their application given their computational requirements, need to be customized to each application (making them generally un-economical) and have no existing technical support structure, given their newness and complexity.

It is more useful to focus on existing models that agencies actually use, rather than recommending more complex forecasting techniques that appear to take "significant effort" to implement for marginal, if not uncertain, gains in accuracy (NCHRP 2006, p.15). In short, improving the tools agencies already use can more efficiently foster possibilities for future improvements with P3 forecasting, than can attempting to adapt and integrate state-of-the-art methodologies. Complexity is not the same as rigor, and as our interviews indicated, forecasts are only one aspect of how potential investors and managers evaluate a good P3 possibility (see sections V and VI). What appears to matter to P3 partners are good faith efforts to minimize project risks, which includes many factors in addition to competent forecasting.

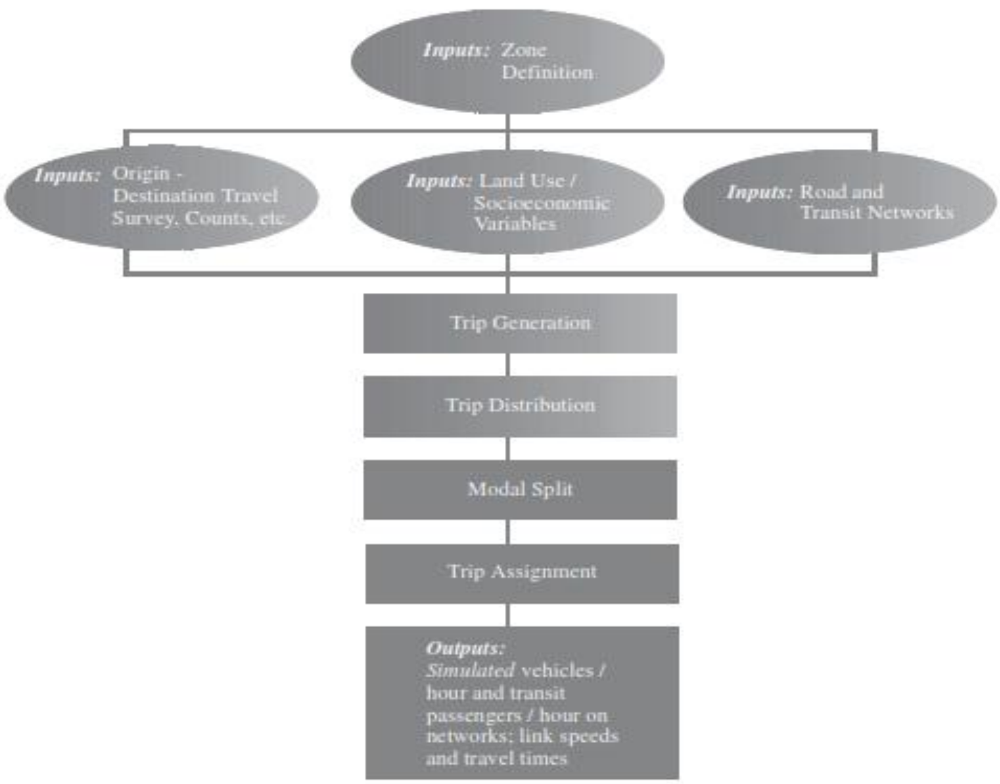
We provide a quick overview and introduction to, and criticisms of, the traditional forecasting methods, primarily for those who are unfamiliar with the methods. From there, we analyze the gaps and extensions needed to answer the some of the pressing informational needs for P3 projects.

## 1) Demand-side factors

### a) The traditional "four-step" model

Traditionally, demand forecasting for transportation facilities has relied on a "four-step model," consisting of distinct phases: 1) trip generation, 2) trip distribution, 3) modal split, and 4) trip assignment (NCHRP 2006, p. 10; Oppenheim 1995, p. 11). The four-step model is usually run for an entire city or planning region, breaking a given geographic area down into traffic analysis zones. Generally conducted by MPOs and other similar regional planning bodies, the traditional four-step model converts zone-specific origin/destination, land use, and transportation network input data into simulated hourly vehicle volumes and corresponding link speeds and travel times. The resulting model outputs are generally in the form of peak-hour volumes and travel times per zone (Oppenheim 1995, p. 11). A chart illustrating the progression of the four-step modeling process (and corresponding outputs) is shown in Figure 1.

**Figure 1: Outline of traditional "four-step" modeling paradigm**




**Source: NCHRP 2006, p. 11**

MPOs have generally employed this forecasting process, because it is easy to adapt for multiple analyses. Given MPOs' focus on environmental regulatory compliance, this forecasting process is also attractive because peak-hour volumes – when placed within the context of overall system capacity – can be converted relatively easily into estimations of both air pollutants and fuel consumption (NCHRP 2006, p. 12). Furthermore, this model allows for an assessment of a given project's impacts (both environmental and transportation-related) within the context of the larger overall transportation system.

## b) Drawbacks to the four-step model

While widely-employed for over 60 years, the four-step model is plagued with a number of shortcomings. Chief among these concerns is the data-intensive nature of the forecasting process, with agencies historically struggling to consistently find accurate, comprehensive travel cost data (Kriger 2005, p. 80). This problem has led to a great deal of inconsistency in demand models over time, both within and among various agencies (NCHRP 2006, p. 13). Additionally, the model itself treats travel choices as independent decisions, despite the fact that they are far from mutually exclusive; as an example, deciding to pursue a given trip (trip generation) may be contingent on using a certain mode (modal split) (Kriger 2005, p.80; NCHRP 2006, p. 13; Oppenheim 1995, p. 18).

The four-step model is generally employed within the aggregate context, based on averages of TAZs, rather than disaggregate analyses which are based on individuals and households and offer more details on travel patterns at the individual project and link-level (NCHRP 2006, p. 13). MPOs, by definition, are focused more on macro-analyses than micro-analyses, particularly given the high cost of conducting a micro-level analysis (e.g., the Census) for an entire planning region. But, with tradeoffs in cost comes an inherent decrease in precision, leading to “questionable and inaccurate” aggregate-level, TAZ-based results (Chung and Goulias 1997, p. 30).

These models tend not have the capacity for feedback, particularly in terms of relating the model’s traffic volume outputs into changes in land use inputs, as well as measuring the impacts of traffic congestion at the general/system level (Chung and Goulias 1997, p. 25; Oppenheim 1995, p. 18). As travel patterns most certainly shape land use decisions – and the resulting distribution of economic and land use activity – lack of feedback in rapid successive iterations within the model cause the single estimate output to remain static and unchanged, omitting an important travel demand driver (Chung and Goulias 1997, p. 26). Modifications to the four-step model that attempt to account for policy and planning changes tend to be overly mechanistic and assume that the simple provision of new supply, particularly of non-motorized modes, will inevitably decrease aggregate travel demand via the other modes rather than account for the possibility of changes in overall demand levels (Schweitzer and Marr 2012).

### c) Common modifications and enhancements to the four-step model

A variety of methods have emerged to address these shortcomings, including activity-based modeling, direct-demand modeling, and time-of-day choice modeling. Activity-based modeling directly addresses the lack of feedback in the four-step model and the assertion that certain decisions in the modeling process are mutually exclusive and successive, rather than simultaneous. The enhancement itself is based on micro-simulation of individual behavior (rather than the same type of macro-analysis the traditional model employs), treating travel as derived from an individual's day-to-day activities (NCHRP 2006, p.13). In this paradigm, the base unit is, rather than a single trip, a "tour" – a sequence of three or more trip segments originating and terminating at home (Shiftan et al. 2003, p. 3). Tour based models differ from traditional trip-based modeling efforts in that trip-based models consist of no more than two segments (X-Y as one segment or X-Y-X as two segments). In this sense, activity-based modeling views transportation as a means to an end (rather than an un-contextualized activity), circumventing some of the problems apparent in the traditional model with regards to mutually exclusive and systematic choice-making. These models are extremely complex, and very difficult to implement.

Direct demand modeling treats all phases of the four-step model as occurring simultaneously, (Oppenheim 1995, p. 17). Such an arrangement measures demand responses more directly and more effectively than the widely-used four-step model. However, such models require a large number of variables, given multiplicative form and the structure of the equations. The data needs for direct demand models are quite large, especially if the aggregate-level analysis is for a sizeable region (Oppenheim 1995, p. 18).

Time-of-day choice modeling divides a given period of time into "slices," making choices a function of the time of day (or day of the week, or season of the year, depending on how time is "sliced"); linking travel behavior choices with various time-of-day segments . Time-of-day choice itself is expressed by: modeled "slices;" modeled days; allocation of trips between peak time and peak "shoulders" (peak spreading); and the actual range of time-of-day values (NCHRP 2006, p. 14; Kriger 2005, p. 88). Generally, peak spreading has been the most common reference, as past findings have suggested peak spreading occurs as a result of overall trip congestion (end-to-end) rather than merely congestion on one link and has a significant impact on congestion and traffic levels (NCHRP 2006, p. 14).

## d) Traditional tools vs. specialized tools for priced facilities

### Demand/revenue forecasting

Despite these improvements and the insertion of various feedback loops and propagation/segmentation mechanisms into the four-step model, the model itself still omits the core component of any toll road demand forecast: pricing. Theoretically, pricing could be incorporated into the four-step model's mode choice segment, which divides traffic based on generalized cost functions; however this output would be a rough demand model, which is very different from modeling revenue. As such, MPOs' application of the four-step model alone is not sufficient for toll road demand and revenue forecasting; long-range transportation and corridor forecasting tools employed by MPOs do not capture factors critical to toll road feasibility (NCHRP 2006, p.12). Toll road forecasting tools require a focus on pricing; while, traditional public sector tools do not. This focus on pricing extends far beyond mere prediction of potential revenue streams to be generated from facility operations. High-level price-based forecasts for toll roads allow for network-wide assessments of the impacts of managed lanes and road pricing, while a concentrated assessment of a specific facility can provide a range of expected toll revenues, given a range of price points, tolling regimes, and macroeconomic conditions (Kriger 2005, p. 90).

In an un-tolled facility, the downside to an overly optimistic forecast is minimal – traffic use on a particular facility does not influence revenue. In a tolled facility, the stakes are much higher as there may be bonds or other types of debt tied to revenues. In order to secure private sector financing, a tolled facility needs consistent, reliable cash flows. Therefore, ratings agencies, financial and consulting firms, and both public sponsors and private firms in the investment have a stake in scrutinizing the forecasts (see sections 5 and 6). Given this level of increased project scrutiny, better forecasts and project analyses make projects more competitive in the market for P3s. Specific examples of these types of models and forecasting tools are presented later in this report.

Toll road forecasts require a significantly different output time horizon than most existing MPO models. While the four-step model outputs peak-hour volumes, toll road demand forecasting requires annual volumes (NCHRP 2006, p.12). Extrapolation of these peak-hour data points to produce annual numbers is an option; however, that method can prompt forecast error to spike dramatically – particularly as one extrapolates data into the future (Armoogum et al. 2009, p. 20). Another major source of uncertainty with such an extrapolation is the underlying assumptions of population and employment growth driving overall facility demand. The resulting extrapolated aggregate annual total is static, in that it does not account for changes in traffic composition, peak-spreading, toll rates, demographic/economic shifts, and/or natural variability of traffic patterns throughout the course of a year (Allen and Schmitt 2005, p. 7). Furthermore, as time of day – and time of year, for that matter – changes, so too does VoT, associated differences in trip purposes, congestion levels across the transportation system, and overall trip distribution (Kriger 2005, p. 91). These factors are of particular importance for toll roads, as they collectively form the basis of a given facility's price point at a given time and overall revenue generation potential. Such factors are noticeably absent from traditional, generalized forecasting methods.

In addition, toll roads can be highly susceptible to shifts in land use patterns. As such, toll road models need to account for potential major shifts in land use patterns, under a variety of land use assumptions. Data requirements for enhanced land use modeling include market-based development trends; data should be as localized and specific to a toll road's corridor as possible (Kriger 2005, p. 91). Toll road models also should incorporate feedback into land use inputs, allowing for a dynamic picture of travel demand that mirrors changes in land use. Among two of the pre-eminent land use and economic modeling tools employed are the Regional Economic, Land Use, and Transportation (RELU-TRAN) algorithm and UrbanSim microsimulation software. RELU-TRAN is a "dynamic general equilibrium" of metropolitan economics and land use, working off of such inputs as available floor space, the market for industries' product, and frequency and distribution of commuter and discretionary travel (Anas and Liu 2007, p.415). UrbanSim, by contrast, employs a series of pricing and utility functions to determine land use and economic activity distribution, all within the context of a regional transportation network (Waddell 2011, p.8).

Traditional models also do not adequately account for truck and commercial traffic levels, a particularly important market segment for toll roads (Allen and Agnello 2003, p. 8). Because trucks and commercial traffic incur a significantly higher cost from traffic delays than passenger automobiles, toll operators can charge a higher rate, thereby generating a significant proportion of overall revenue (Kriger 2005, p. 92). As such, adequately accounting for truck traffic – forecasting truck traffic volumes at a disaggregate level relative to all vehicles, which is not done particularly well within the four-step modeling process – is a key improvement for toll road forecasting models.

Some jurisdictions have developed separate truck models to supplement other automobile forecasting techniques, while others have estimated truck traffic by employing ratios based on observed traffic (Allen and Agnello 2003, p. 1). Such a fix, however, is not necessarily sufficiently accurate, as peak truck volumes often occur at different times than passenger automobile volumes (NCHRP 2006, p. 13). Given that the time-cost relationship for trucks and commercial traffic differs from passenger automobile traffic, extrapolations are not really appropriate logically or mathematically (Kriger 2005, p. 92).

## Risk analysis

Risk management, given all these concerns, is a limited tool, even with good analyses (NCHRP 2006, p. 30; Lemp and Kockelman 2009, p. 6). Past forecast performance and research collectively identify three key factors on which to focus risk and sensitivity analyses as part of demand and revenue forecasting for tolled facilities. Those three key risk factors include:

1) Demographic and socioeconomic inputs;
2) Value of time; and
3) Ramp-up period (Lemp and Kockelman 2009, p. 6; Bain and Wilkins 2002, p. 10; NCHRP 2006, p. 27; Kriger 2005, p. 84).

### *Demographic and social factors*

Two demographic and socioeconomic issues arise: long-term forecasting difficulties and short-term economic fluctuations. With long-range demographic and socioeconomic forecasts, significant evidence exists that these predictions – when implemented by MPOs – often more closely mirror MPOs' planning priorities and aspirations more than actual market trends (George et al. 2003, p. 2; Nunez 2007, p. 45). In order to combat the "wishful thinking" and to base forecasts as much on revealed preferences and market observations, recent tolling demand/revenue forecasts have sought to vary input data age and land use/transportation system across contexts. By doing so, MPOs have been able to incorporate estimates which mirror their planning policies while at the same time generating conservative scenarios, based more heavily on past historical trends (George et al. 2003, pp. 2-4).

How useful refined long-term land use forecasts really are remains relatively unknown, given evidence from – among other cases – California's Orange County Toll Roads (Vollmer Associates 2003, p. 141). (We will profile that case as part of series of in-depth case studies in Section 7.) Orange County's Transportation Corridor Agencies revised their traffic and revenue forecasts in 2003 to better reflect the regional development that had occurred since the forecasts were first prepared. Changes in employment trends (by sector and by location), housing prices, and other variables within 50 identified "focus areas" were incorporated into the land use inputs to adjust the forecasts (NCHRP 2006, p. 25). Over the next two years, the forecasts actually underestimated revenue by an average of 56%, but this underestimation likely stemmed from the fact toll rates increased system-wide in 2004, while the forecasts used the unchanged prior, lower toll rates (TCA 2005, p. 13; NCHRP 2006, p. 25). Either way, the housing market's slowdown and eventual crisis later in the 2000s rendered both models irrelevant. That change reinforces the challenges of calibrating static demand/revenue models to account for long-term exogenous events and macroeconomic shifts (TCA 2012a; TCA 2012b; George et al. 2003, pp. 2-4; Lemp and Kockelman 2009, p. 10; for more on the Orange County case study see section 7).

Turning to near-term economic change, fluctuations in travel demand has often led to over-optimistic financial forecasts for facilities (NCHRP 2006, p. 25; Lemp and Kockelman 2009, p. 7). Many of these near term – and oftentimes initial – demand/revenue forecasting errors occur because "base year" projections and inputs reflect very different economic conditions from a facility's actual first year (Kriger 2005, p. 91). Phrased differently, observed data input as a base year scenario before project implementation can quickly become inaccurate – and very much so, at that – once a project opens following a major economic shift. If local data in the midst of a recession, for example were input as base year parameters, the model would quickly break down if, when the facility opens to traffic, local economic conditions improve. Furthermore, such a model would not adequately account for the natural ebb and flow of business cycles, future changes in zoning, and an accurate picture of future economic activity more broadly.

These errors do not reflect problems with regional forecasting as much as they do troubles with integrating land use and development changes within a given facility's corridor/cachement area (Kriger 2005, p.91). Toll road experiences in Florida and California – where aggregate regional growth was forecasted accurately yet specific toll roads still underperformed – show that successful demand/revenue forecasts are those that focus on past socioeconomic and demographic trends to forecast an accurate *distribution* of economic activity in specific corridors, rather than just aggregate-level development (Vollmer Associates 2003, p. 13).

### *Value of time (VoT) estimation*

Geographically relevant analogues help enlighten VoT differences and potential willingness to pay for a particular P3 facility. Pricing works off of the assumption that a driver's decision to use a given tolled facility, relative to an un-tolled facility, is itself a function of the driver's individual valuation of trip time saved by choosing the tolled facility. As such, VoT varies dramatically, and is very much heterogeneous – by trip purpose, mode, average level of income, and time of day even for the same individual (Lemp and Kockelman 2009, p.8). Prior evidence suggests further variation in VoT according to gender, trip length, and education levels as well (Sullivan 2000, p. 146). VoT also varies widely between regions, and is best calculated from "locally derived data" for specific facilities, rather than at an aggregated or regional level and generalized to a number of facilities (USDOT 1997, p. 5).

Generally, however, for work-specific trips, drivers' hourly wages are an accepted proxy for VoT, but for other personal (non-work) trips, much less agreement exists, and usually VoT is represented as a fraction of drivers' hourly wage rates (NCHRP 2006, p. 27). In demand/revenue forecasts for tolled facilities, these values are aggregated and averaged, which ignores the variation among travelers and trip purposes – especially in terms of changing trip purpose and/or exogenous socioeconomic trends.

Research on tolled facility user preferences on California's SR-91 Express Lanes compared users' stated VoT data with actual choice data, identifying significant differences among the measured stated and revealed characteristics including: trip time of day, flexibility of arrival time, gender, age, household size, occupation, marital status, and education (Small et al. 2005, p. 16).

Those differences bring up some problems for priced facility demand forecasting. Stated preference surveys are a common method for gathering VoT data. By posing hypothetical options to travelers, these surveys serve to quantify how toll rates on a given facility would affect driver behavior; stated VoT can be gleaned from these sorts of surveys. Stated preference surveys consist of three core sections: background information on recent trips within the focus corridor; the actual stated preference experiment; and respondent demographic information (NCHRP 2006, p. 28).

However, downward biases are commonly observed in stated VoT for respondents in an area where tolled facilities do not exist. Because people are accustomed to toll-free access to transportation facilities in these contexts, their VoT is comparatively lower than for individuals in areas where tolled facilities do exist. However, in areas with existing tolled facilities and severe peak-hour congestion, VoT tends to be overestimated, as is the case where electronic toll collection is common (Baez 2004, p. 11). Furthermore, stated VoT values vary widely according to assumptions made on what stated VoT data actually represent. Depending on how stated preferences are ordered based on survey results (is travel comfort/convenience subordinate to travel reliability or safety, e.g.) VoT output can vary dramatically in stated preference (Calfee et al. 2001, p. 699).

Other problems with stated preference surveys can also confound their use in P3 forecasts. Respondents do not possess perfect and complete knowledge – of toll rates, alternatives, and how preferences can shift relatively and absolutely throughout the day, among other items – at the time the survey is administered. Because of these information issues, respondents may not cognitively be able to express their preferences as they would actually on them given a real-world set of conditions. Stated preference survey output, because of these assumptions about the extent of respondents' knowledge, should not be expressed as indicative of a singular VoT data point, but rather as a range of values (Calfee et al. 2001, p. 699; Allen and Schmitt 2005, p. 3).

These findings reinforce the notion that VoT is a highly-specialized calculation for local users, facilities, and contexts. Problems arise in the out-of-context application and/or development of VoT when compared to empirical studies of revealed VoT preferences (Hensher and Goodwin 2003, p. 1). Specific examples of these mis-applications include:

1) trip assignment models simulating route choice while employing VoT associated with empirical studies of mode choice;
2) application of VoT to stated preferences different from those used to initially calculate VoT;
3) relationships between VoT and variables assumed to interact inconsistently with past evidence;
4) VoT calculated from stated-preference surveys (based on short-term/immediate preferences) but applied to long-term models; and
5) applying observed VoT for a given asset to another which is very different in system context and/or physical characteristics (e.g., VoT for a bridge with no substitutes applied to toll facility with readily-existing substitutes) (Hensher and Goodwin 2003, p.1; Spear 2005, p. 18).

In addition to variations according to specific contexts, VoT can fluctuate in accordance with larger socioeconomic trends, as shown in historic analyses of VoT data for specific facilities (Lemp and Kockelman 2009, p. 8). Tolls themselves can also influence VoT over the long. As regions become accustomed to tolled links within the transportation network over time, individual travel patterns and VoT have been shown to shift more towards usage of tolled facilities and a comparatively higher VoT (Burris et al. 2004, p. 84).

Other, closely related issues can also influence demand and willingness-to-pay. Reliability, in this context, relates to the degree of variability in daily travel times and traffic volumes within a given corridor. This variability occurs because of exogenous events as accidents, weather, and construction, among other factors (NCHRP 2006, p. 29). Pricing reliability is critical for all tolled facilities, as the facility is intended to provide a superior travel time relative to its free alternative, but it is especially pivotal in real-time/variably-priced tolled facilities, which price facility access according to traffic volumes relative to the level of service. How travelers value reliability can vary independently of VoT; between this non-correlation and the fact that little data exists to develop a model for how travelers value reliability, incorporating the value of reliability into road pricing models has proven quite difficult (Brownstone and Small 2005, p. 286; Spear 2005,p. 19).

### *Ramp up performance*

A facility's ramp-up period consists of the very first years of facility operations, immediately following its opening to traffic. This period can have exceptionally high, often erratic, growth in traffic, with traffic growth tapering off and stabilizing at lower (closer-to-predicted) levels towards the end of the ramp-up (Kriger 2005, p. 84). Ramp-up periods are import to account for in models because of the equally erratic nature of cash flows – both in terms of magnitude and sequencing. The exact duration of the ramp-up period can vary, depending on the congestion levels on existing free alternatives, the magnitude of expected travel time savings, income levels of expected users, and the newness of the concept of tolling to local drivers (Bain and Wilkins 2002, p. 10).

Modeling ramp-up performance is of particular importance for project financiers, and accordingly, because the probability of default is highest then (Bain and Wilkins 2002, p. 10). This comparatively greater risk of default during ramp up occurs due to uncertain project cash flows, coupled with initial scheduled payments to creditors and other project investors. Past research has identified three primary factors that analysts should consider in thinking about ramp up forecasts:

1) scale of ramp up;
2) duration of ramp up; and

3) extent of catch-up (Bain and Wilkins 2002, p. 10; Flyvbjerg 2005a, p. 526).

The scale of ramp up refers to the magnitude of the difference between actual and forecast traffic (effectively the extent of the initial growth surge followed by a tapering-off. The duration of ramp-up refers to the actual length of the facility's initial operating period – usually from opening day through roughly the first five years (NCHRP 2006, p. 29). The extent of catch-up refers to the probability of facility usage – in the event it is actually lower-than-anticipated during ramp up – catching up with later year forecasts and, in the future, meeting or exceeding demand/revenue projections. Catch-up is, in effect, the extent to which a facility's demand recovers if it was overestimated during ramp-up stages. Catch up volumes and percentages have been empirically observed to be highly significant in determining long-range use performance (Flyvbjerg 2005a, p. 526). Projects that get off to a slow start typically underperform throughout their lifespan.

Generally, past early-year forecasts and facility operating plans have simply neglected ramp-up considerations. However, as more projects have come on line, increased data on project ramp-up performance – and the factors underlying and driving that performance – have been observed. In addition to the tendency of early underperforming projects to continue underperforming well into their maturity, projects have demonstrated an inverse relationship between actual time savings offered and ramp-up length (George et al. 2003, p. 5). Greater observed travel time savings correlate with a shorter ramp-up period, as intuition would suggest. Premium services attract users more quickly. In the case of two Texas area toll roads, the Hardy and Sam Houston Toll Roads, the latter had greater time-savings and performed better financially than the Hardy; the Sam Houston facility was also located in an area with higher average incomes (presumably, the users had a higher VoT too) which may have also contributed to its shorter ramp-up period (NCHRP 2006, p. 30).

## e) Specialized tools for priced facilities

### Demand/revenue forecasting

Five distinct methods for toll road demand and revenue forecasting have emerged in step with renewed interest in managed lanes and road pricing over the past 25 years. These five methods are:

1) Activity-based models
2) Modal-split models;
3) Trip assignment models;
4) Post-processor; and
5) Sketch planning methods.

Of these five different methods, only 2 through 5 appear to be used among public sponsors for toll road demand forecasting; methods 2-5 are very much state-of-the-practice whereas method 1 – activity-based modeling – is state-of-the-art. Only one MPO, Portland, Oregon – has used an activity based model. As such, we focus on the remaining four methods. We could find no consensus in the literature on which of these methods are best suited for forecasting toll road demand and revenue. Which model will work best for a giving P3 will vary with the data availability, constraints, and forecaster experience.

#### *Modal split models*

Modal split models work within the existing four-step model but treat automobile trips on tolled and non-tolled roads as distinct modal choices, with separate modal split functions for both work and non-work trip purposes (Kriger 2005, p. 81; Spear, p. 15). The functions themselves vary according to VoT; such an approach allows "out-of-pocket" costs to be directly and explicitly modeled within each function (Kriger 2005, p. 81). Modal-split models effectively tie drivers' utility – based on VoT calculations – to toll rates and the resultant mode shares, allowing analysts to estimate the aggregate toll revenues, traffic volumes, and cross-mode elasticities (NCHRP 2006, p. 15). Additionally, the logic can be applied to trip distribution modeling as well, largely because of the inherent elasticities built into the modal-split model. With the inclusion of VoT and explicit toll rates, modal-split models can, with minor adjustment, account for the spillover of vehicles from a newly-tolled facility to an un-tolled alternative (Cambridge Systematics 2005, p. 5). The necessary adjustment to the model would classify various VoT by trip purpose and – in conjunction with toll values and knowledge of pre-existing and future potential alternatives – calculate user demand for the tolled facility accordingly.

The modal split approach assumes that VoT estimates are reasonably accurate given both macroeconomic and demographic factors associated with the forecast. This assumption will be discussed in full detail – and within the context of the various types of models discussed in this section – later in the report. Obtaining a reasonably accurate VoT figure across the varieties of trip purposes is especially difficult – yet especially critical – when accounting for non-existing (to-be-built) facilities. The challenges in doing so stem from the need to accurately identify comparable facilities as exemplars and to extrapolate past demand to the non-existing facility. This approach can be particularly troublesome when considering planned facilities for entirely new modes – like high-speed rail in most of the proposed US operating contexts, for example. Hoping to increase the accuracy of forecasting for non-existent services, a number of municipalities – Phoenix, Arizona the largest urban area among them – have incorporated stated preference data into the model in the absence of revealed preference data for the facilities (Nourzad 2004**,** p. 31). These surveys, rather than being developed and administered by the public sponsor, were contracted out to private firms (generally consultancies) which specialize in such types of transportation econometrics and analytics.

In the case of Phoenix, the Mountainland Association of Governments – the Phoenix-area MPO – contracted services out to URS and Urban Analytics, which introduced a step in the modeling process to account for real-time tolling. This new function adjusted per-mile tolls on managed lane links between iterations, maintaining a standard level of service of D/E on tolled facilities; modal split was then adjusted accordingly with toll rates and traffic volumes (VHB 2006, p.7). This addition replaced the general cost measure common in four-step modeling (URS 2004, p. 4). Modal distributions were achieved by including five distinct variables in the model: highway journey time, highway distance, door-to-door transit time and transit fare. Ultimately, the transition made the overall model more sensitive to toll rates, assuming the price of tolls is the key driver in modal split (URS 2004, p.4).

Additionally, the consulting team introduced a time savings term into the overall model – the difference between tolled and non-tolled travel times – and coined it the "reliability" factor (VHB 2006, p.7). This addition further reinforces the importance of comparative time advantage (or disadvantage) of a particular route relative to other options in determining modal distributions. Including two layers in the modal split portion of the model – accounting for both varying toll rates and traffic speed – enables for a more sensitive assessment of modal split, and thus overall revenue potential, than do traditional forecasting methods.

### *Trip assignment models*

Trip assignment models assume trip distribution and modal share – between tolled and non-tolled automobile trips – are static, remaining unchanged without feedback loops within the model. Two general ways of modeling trip diversion in trip assignment have emerged in practice: translation of tolls into time equivalents and basing toll forecasts on diversion curves – the relative likelihood of using a tolled facility given toll levels relative to un-tolled alternatives. For the first method, monetary toll rates are converted into trip time equivalents by employing a VoT range. The VoT themselves are chosen to reflect a variety of trip purposes and traveler socio-economic/demographic factors as well (ORDOT 2009, p. 14). Those equivalent times are incorporated into the model's volume-delay functions, which then allocate trips to various links and paths based on travel time, link capacity, and link congestion (Dehghani et al. 2007, p. 9).

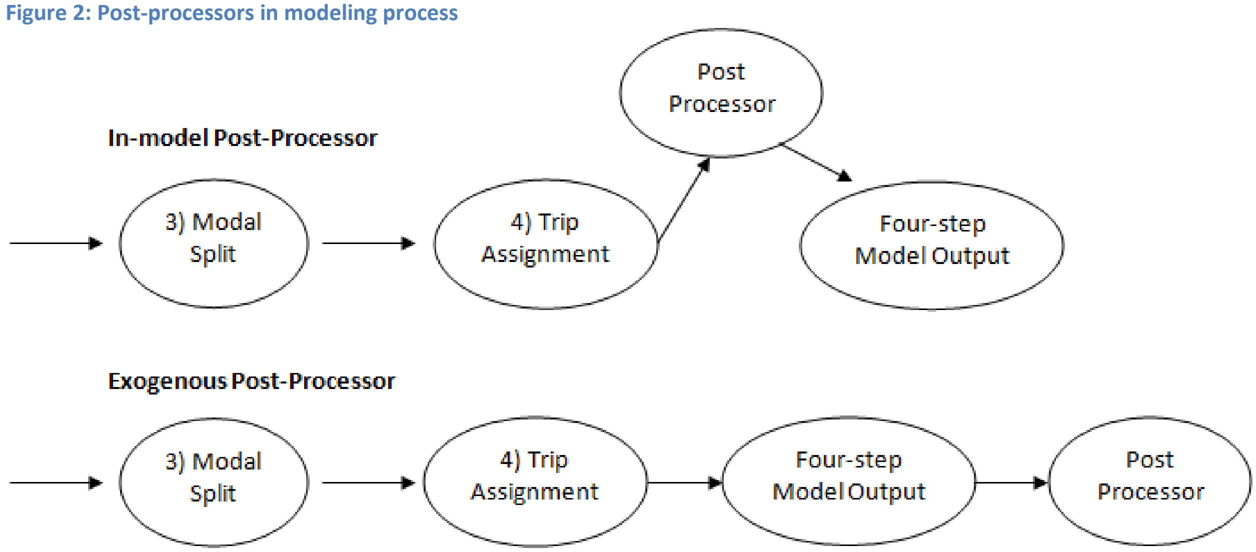
The second method calculates the likelihood of a given traveler using a tolled facility (output as a percentage share of traffic within a system/network) as a function of the relative cost of travel time between tolled and non-tolled routes (Spear 2005, p. 16). This second method generally takes the form of a logit function or "s-curve" (Kriger 2005, p. 81). The slope of the logit function represents demand elasticity of the relative cost or travel time (depending on which calculations are initially specified) for the tolled facility. The function is inversely-proportional to VoT or willingness to pay (again, depending on initial specification) (Spear 2005, p.16).

The process ultimately outputs tolled and non-tolled trip tables, which the analyst can then categorize by a variety of factors or descriptors – trip purpose, income group, vehicle occupancy, trip time, etc. – for assignment within the model (Kriger 2005, p. 82). While more complex and data-intensive than the first method, the diversion models can supplement an existing four-step model without needing to first recalibrate it. However, it would be challenging for MPOs and/or state DOTs to develop their own diversion models given the difficulty and high costs of developing such a model, which is why most existing diversion curve models for states and/or private firms are proprietary (Spear 2005, p. 17). The literature identified but one traffic and revenue study conducted by TxDOT in Austin, Texas which employed diversion curves, in conjunction with stated-preference surveys (NCHRP 2006, p.89; Vollmer Associates 2004, p.37).

### *Post-processor*

Modeling demand as part of a post-processor entails incorporating an independent procedure to divert assigned traffic volumes (volumes following trip assignment) from general purpose lanes to tolled facilities in accordance with excess capacity available in the tolled lanes (Kriger 2005, p. 90). Effectively, post processors feed traffic data – based on VoT and other characteristics – into managed lanes, and the remainder back into general purpose lanes. Post-processors can be employed either within an existing four-step model or exogenously, incorporating the output of the four-step model; both options are diagramed below in Figure 2

Figure 2: Post-processors in modeling process




Washington, DC and San Diego, California have implemented the former approach, incorporating the post-processor as an addition to the four-step model, while Minneapolis/St. Paul, Minnesota has used the latter method, incorporating assigned volumes into FHWA's STEAM[1] model as part of a pricing study for the I-394 MnPass program (Nourzad 2004, p. 26). Washington DC used a post-processor to assess regional freeway pricing; the add-on converts toll rates into equivalent VoT across the freeway network, based on current income stratification data used by the existing model (URS 2004, p. 9). A diversion model then diverts traffic from general purpose lanes to hypothetical managed lanes in accordance with excess capacity in the managed lanes and the pre-defined level of service minimum for the facility.

San Diego, California has also employed a post-processor to divert traffic from congested freeways to managed lanes, however – unlike the Washington, DC post-processor – toll rates are not considered as a part of the model add-on. Diversion is based solely on over-demand on regular lanes and existing capacity on managed lanes, with excess traffic above the volume which would allow a level of service of C diverted back to regular lanes from managed lanes (URS 2004, p.10). Generalized costs

[1] STEAM was developed by FHWA in the 1990s to compute net value of project benefits for major, "regionally important" transportation projects. The most up-to-date version of the model, version 2.02, is available online for free from FHWA. The model can be accessed at: http://www.fhwa.dot.gov/steam/products.htm.

are used for traffic assignment, however toll rates are not included in the modeling process, which is instead based virtually solely on a managed lanes facility's volume to capacity ratio.

While these procedures are generally simple to implement from an operational/mathematical standpoint, they are not particularly sensitive to fluctuations in travel behavior, as evidenced from the lack of toll pricing in the San Diego post-processor (NCHRP 2006, p. 17).

### *Sketch planning methods*

Sketch planning methods are generally "quick response tools" used for project evaluation (Kriger 2005, p. 90; NCHRP 2006, p. 17). FHWA has created and made available a number of such tools, including FHWA's: Spreadsheet Model for Induced Travel Estimation (SMITE)[2]; Spreadsheet Model for Induced Travel Estimation – Managed Lane (SMITE-ML)[3]; and Sketch Planning for Road Use Charge Evaluation Model (Nourzad 2004, p. 28). Most relevant for toll roads forecasting, the Managed Lane FHWA spreadsheet model employs a pivot-point logit function to estimate elasticities for travel demand relative to changes in travel time, tolls, and improved transit service (FHWA 2011a).

Also of particular relevance to toll road demand forecasting is Texas Transportation Institute's spreadsheet-based Toll Viability Screening Tool.[4] The model enables economic assessments of a proposed tolled facility by forecasting a range of values for initial facility demand and outputting a resultant net present value for the facility. This initial range is calculated by incorporating different values (via randomization) for the various input factors, which include daily traffic volumes, toll rates, and assumed diversion. In addition to its net present value output, the model includes an internal risk-assessment tool and sensitivity analysis capability to provide a probability/frequency distribution associated with each set of inputs and net present value outputs (Smith et al. 2004, p. 14). In addition

[2] SMITE available from FHWA at: http://www.fhwa.dot.gov/steam/smite.htm.
[3] SMITE – ML available from FHWA at: http://www.fhwa.dot.gov/steam/smiteml.htm.
[4] TTI's Toll Viability Screening Tool is an adaptation of Palisade Corporation's Decision Tools software suite, which a set of macros based in Microsoft Excel. The TTI tool is an add-in that can only be obtained and operated following installation of the Decision Tools software suite.

to forecasting for greenfield toll facilities, this sort of tool can be applied to forecasting anticipated toll revenues for existing free facilities which are under consideration for tolling.

**Table 2: List of practitioners by model type**

| Region / Agency | 1) Activity-based | 2) Modal Split | 3) Trip Assignment | 4) Post-Processor | 5) Sketch Planning |
|---|---|---|---|---|---|
| Portland, OR | X | | | | |
| Phoenix, AZ | | X | | | |
| Sacramento, CA | | X | | | |
| Minneapolis/St. Paul, MN | | X | | X | |
| Port Authority NY/NJ | | X | | | |
| Austin, TX | | | X | | X |
| Washington, DC | | | | X | |
| San Diego, CA | | | | X | |
| State of Florida DOT | | | | | X |

**Source: NCHRP (2006), pp.15-17**

## Risk analysis

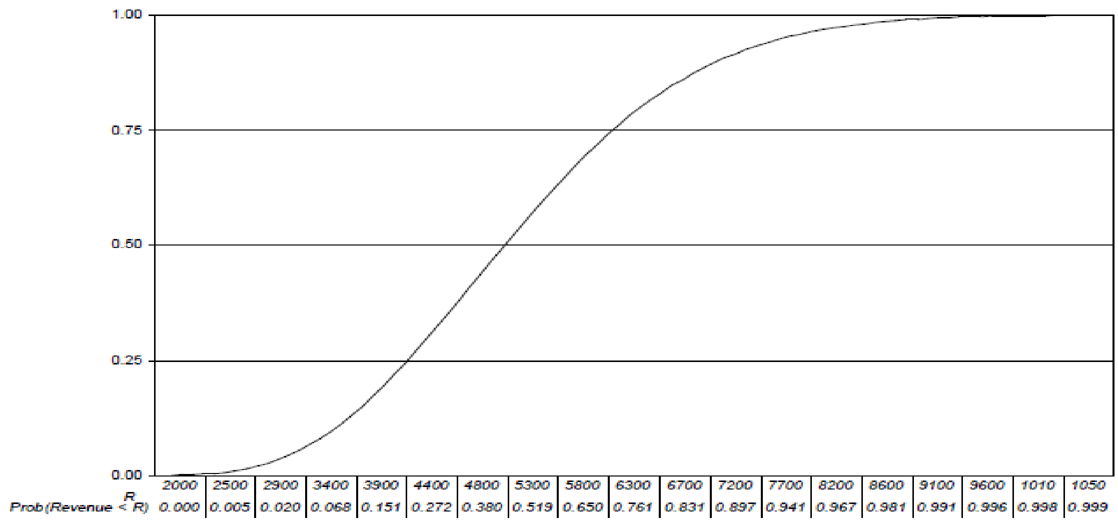
In discussing risk analysis methods, it is important to first draw a distinction between pure risk analysis and sensitivity analysis. Sensitivity analysis is essentially the examination of how different values of an independent variable impact a dependent variable under a given set of assumptions and circumstances. Sensitivity analysis alone is fairly limited in its ability to provide a wide range of outcomes and sufficient insight into project feasibility, as the iterations occur incrementally by individual variables (Lemp and Kockelman 2009, p. 10).

Risk analysis, by contrast, can quantify a much larger range of probable outcomes than can sensitivity analysis (Lemp and Kockelman 2009, p. 10). For tolled facility forecasts, risk analysis refers to assigning probability distributions to inputs – incorporating a range of values for variables integral to forecasting demand – as compared to inputting just one particular value. Probability distributions are also applied to the configuration of a future base year network, when and how that network develops, when and where the adjacent corridor develops, and a range of other "uncertainties in the modeling process and structure" (Kriger 2005, p. 85). Establishing probability distributions for inputs and future conditions better models real-world uncertainties than adjusting fixed input values one by one because factors can change concurrently (NCHRP 2006, p. 31).

The pre-eminent method of generating these probability distributions is through large-scale randomization of changes in parameter values and underlying assumptions, principally through Monte Carlo simulation (Lemp and Kockelman 2009, p. 11). Randomization is superior to pre-determining conditional probabilities and engaging in expected value summation – or similar methods – due to the large variances and major potential for bias inherent in those sorts of calculations (Ben-Akiva 2008, p. 32). Monte Carlo simulation, by contrast, allows for a very large number of scenarios to be simulated very quickly; by not only randomizing the actual input parameter values, but also randomizing which parameters are changed, a simultaneous and randomized “fluctuation” in the model can be achieved (Kriger 2005, p. 93). Through series of these randomized iterations, the changes which created the greatest amount of variation in the model – and the variables to which those changes were tied amidst the iterations – can be identified as key risk drivers for the forecast. Again, Monte Carlo simulation does not produce a singular “output” value so much as it does a probabilistic, risk-adjusted forecast range of output variables.

Bowman et al. (2002) implemented a slightly more advanced variant of stochastic risk modeling where probability distributions were effectively layered for each input parameter, based on potential sources of error. Then, the approach aggregates each parameter and its associated underlying probability distributions, with outcomes calculated by error source. The resulting output was a cumulative distribution function which allowed for the breakdown of results into percentiles, where the percentile value represents the probability of realizing less than that percentile’s associated outcome units on the x-axis (ridership volume, total revenue, etc.) (Bowman et al. 2002, p. 11). An example of this analysis, done for a new transit system in a major Asian City, is shown in Figure 3.

Figure 3: CDF for 2001 revenue of new transit system in major Asian city




Source: Bowman et al. 2002, p. 11

To this point, the risk analysis techniques covered have not been limited to analyzing any particular set of parameters associated with a given level of risk; these techniques are randomized, based on probability distributions, rather than pre-defined risk thresholds. Stress testing, by contrast, evaluates project performance in the face of major exogenous shocks or other extreme events for which probabilities can be calculated (Kriger 2005, p. 93). Generally, existing asset markets lack historical data for how facilities like toll roads have performed in the long-term after experiencing a major event (like the implosion of the US housing market, for example). Stress testing consists of adjusting input parameter values to reflect “exceptional but plausible” spikes in underlying project risk factors, given this lack of longitudinal data (Jones et al. 2004, p. 3). Where randomized/Monte Carlo-based risk analysis analyzes a range of opportunities across the entirety of a project’s aggregated cumulative distribution function, stress-testing focuses on evaluating project performance at specific probabilistic points. Employing both methods allows for a more complete understanding of project risk moving forward, given more “routine” risks as well as those that are generally, while plausible, unexpected.

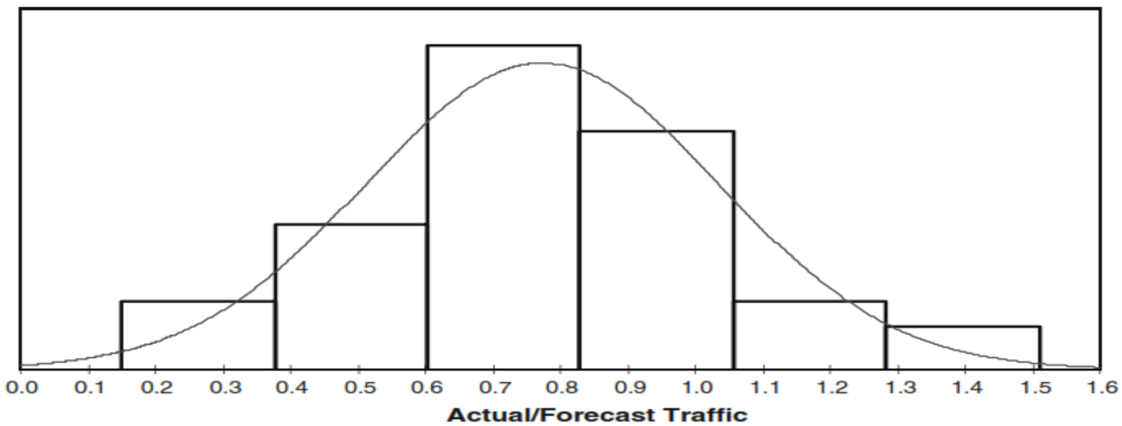
Non-statistical risk analysis techniques have also gained traction over the past decade; however they remain significantly more subjective than empirical and, as such, carry their own risks and challenges in accurate application and implementation (Bain and Wilkins 2002, p. 7). US-based ratings agency Standard and Poor's, for example, developed a traffic risk index informed by case studies of 32 toll roads primarily concentrated in Europe and North America. The index itself is a notational scale running from zero to ten, with zero indicating perfect certainty and ten indicating perfect uncertainty. Projects are generally divided into two clusters: projects of risk zero to five and those of risk five and above.

Project risk is accounted for within the index through the incorporation of ten different project characteristics, including: tolling regime (shadow vs. user-paid tolls); tolling culture/history; forecast horizon; and projected traffic growth (Bain and Wilkins 2002, p. 9). Rather than intending for the Traffic Risk Index to serve as a comprehensive risk identification and management tool, S&P qualified its development as "a starting point for considering toll-project traffic uncertainty in a logical and consistent manner" (Bain and Wilkins 2002, p. 7).

## f) Past performance of demand-side analytical tools

Demand and revenue forecasts for tolled facilities have been jointly plagued by "large errors and considerable optimism bias" (Bain 2009, p.1). In addition, such forecasts have varied widely in their accuracy – there is not an observed bi-modal distribution of forecast performance as much as there has been a very wide gradient of observed forecast accuracy. Investment bank JP Morgan (1997) benchmarked 14 toll road projects, examining the performance of each facility's initial demand/revenue forecast relative to observed early-year performance. Of the projects examined, only one actually met initial projections; three overestimated traffic by up to 25%, four overestimated traffic by up to 30%, and the remaining six projects overestimated traffic by more than 30% - by an average of 42% (Bain 2009, p.2). Ratings agency Standard &Poor's (2005) conducted a similar, albeit larger, study of over 100 global toll roads, comparing original forecasts with observed performance over the course of the four year period from 2002 through 2005. Figure 4: Breakdown of toll road sample (S&P study 2002-2005) shows the results of the study – the distribution of forecast performances according to the ratio of actual to forecast traffic – and the very clear downward bias in forecast performance, as evidenced by the distribution's mean ratio of approximately 0.75.

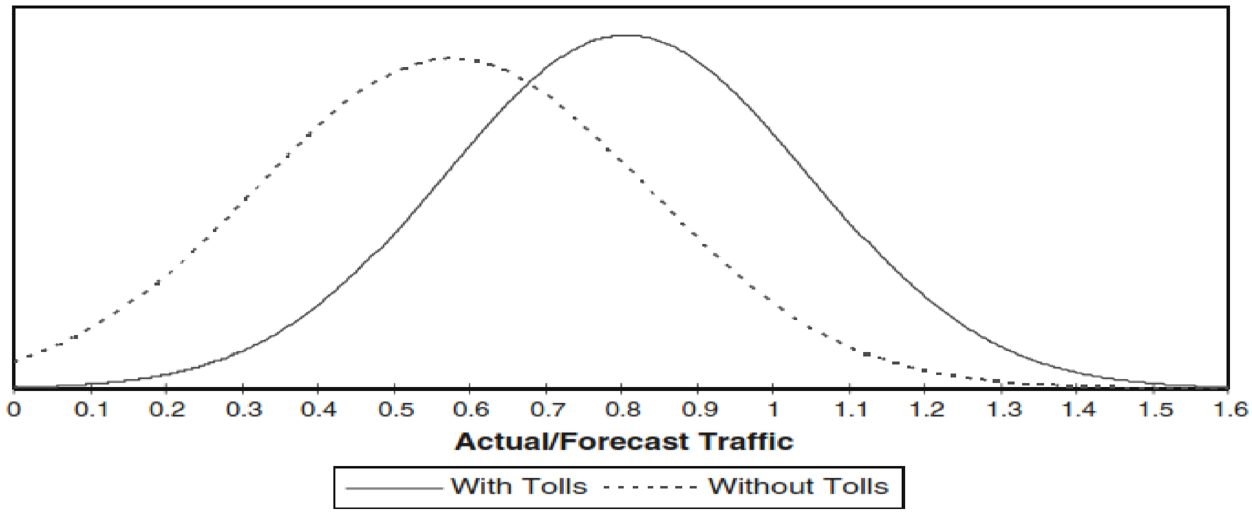
**Figure 4: Breakdown of toll road sample (S&P study 2002-2005)**




**Source: Bain 2009, p.3**

On average, projects overestimated demand/revenue by 23%, with an observed standard deviation of 26%. Such a large standard deviation suggests extreme variability and unpredictability in traffic forecasts, particularly troubling given how important forecasts are to project performance (Parthasarathi and Levinson 2009, p.5). The sample minimum was a forecast which overestimated traffic by 86%, while the sample maximum actually underestimated traffic by 51%. Interestingly, the performance of forecasts in nations with prior tolling experience exhibited a lower over-estimation of traffic (19%) and standard deviation (24%) than forecasts implemented in nations with comparatively less tolling experience, which exhibited a mean 42% overestimation and standard deviation of 26% (Bain 2009, p.5). These results are shown graphically in Figure 5.

Figure 5: Traffic forecast performance in countries with and without prior tolling experience




Source: Bain 2009, p.6

Nonetheless, nations with an extensive history of tolling – and, for that matter, P3s – still suffer from forecast inaccuracy and inconsistency. A report by the Australian Department of Infrastructure and Transport found that demand/revenue forecasts for Australian toll roads "have proven to be highly inaccurate in recent years," as illustrated by the performance of five high profile projects – all of which overestimated traffic levels by an average 45% in the first year of operations (Australia DIT 2011a, p. 4). While some forecasting errors did, in fact, diminish over time, the errors did not altogether disappear – the five assets still averaged a 19% overestimation of traffic six years removed after opening to traffic (Australia DIT 2011a, p. 4).

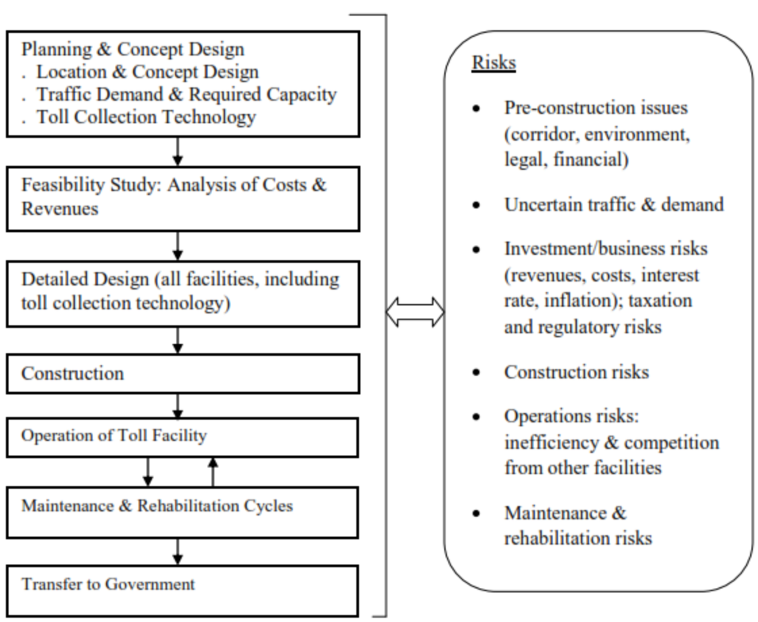
## 2) Supply-side factors

### a) Traditional cost forecasting

Cost projections for traditionally procured projects typically focus on the initial design of a facility, as well as the initial capital outlay to finance the project, rather than considering both up-front costs as well as all other deferred and longer-term costs like ongoing O&M (Miller 2008, p. 20). By contrast, P3s require a life-cycle cost approach that considers all relevant facility costs – from initial delivery through asset hand back. As P3s use alternative financing more regularly than traditionally-procured and financed projects, understanding future cash flows becomes central to evaluating project feasibility. Both the sequencing and value of cash flows directly influence the attractiveness of a P3 to private firms, as both can alter projects' cost-benefit and financial feasibility calculations dramatically (Miller 2008, p. 25). Getting liquidity forecasts right – or, perhaps more realistically, close enough – for priced facilities delivered as P3s is essential to sustain project viability.

Figure 6 breaks down life cycle cost (LCC) components for a toll facility, as well as specific risk factors throughout the project lifecycle, from conception through end-of-contract hand back. Each individual component contributes to the overall financial viability of the project over the course of the facility's lifespan, informing both public and private sponsors of the facility's future spending needs. Toll road LCC components and associated risk factors differ from those of standard, un-tolled facilities because of the fact that a toll road authority – public or private – can assume total responsibility for all aspects of the road and/or the toll system. Toll road agencies manage traffic and road assets according to multiple principles, including profit maximization; thus planning and O&M strategies can differ from free facilities as well. These differences lead to a unique set of LCCs and risks (Khan 2009, p. 3).

**Figure 6: Life cycle costs and risk factors for a toll facility**




**Source: Khan 2009, p. 19**

## b) Specialized tools for priced facilities

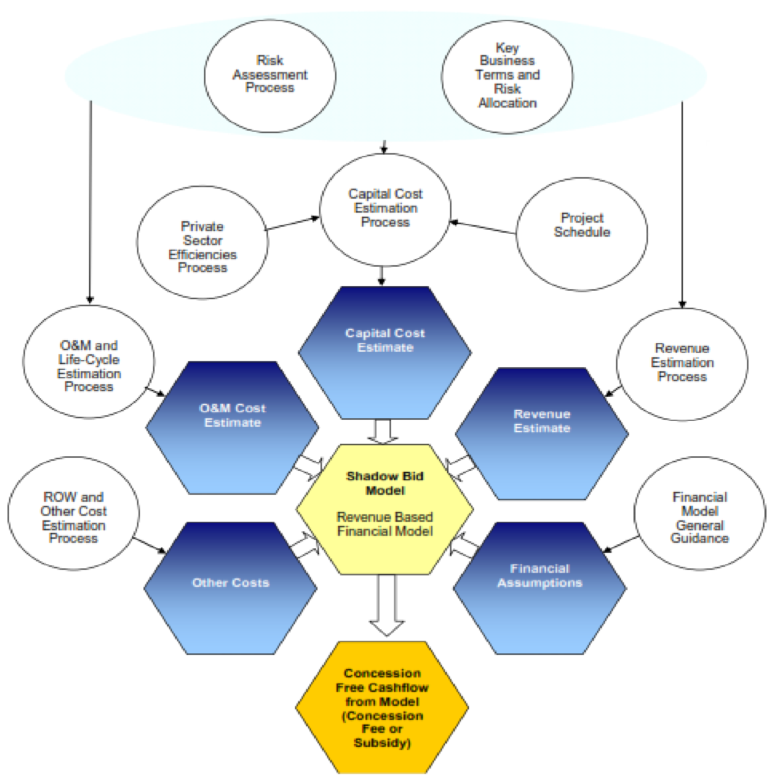
### Life cycle costs

Value for Money (VfM) fiscal analysis consists of a public sector cost estimate (the PSC) and at least one alternatively-procured bid; by comparing costs over the life-cycle of a given project under both delivery types, comparative advantages of alternative project delivery (relative to traditional procurement) can be identified. VfM, effectively a set of NPV calculations, incorporates all future cash flows – both inflows of revenue and outflows for expenditures like O&M – while a) pricing risk and project transaction/financing costs and b) aggregating and discounting total LCC to produce a final value against which to compare other bids (State of Virginia 2011, pp.9-10). The PSC estimates expected LCC to the public agency, and it is the yardstick by which other bids are measured; however, in order to evaluate a variety of bids on a level playing field, competitive neutrality must be ensured in the analysis. Competitive neutrality refers to the "inherent advantages and disadvantages that are available to a government agency pursuing the PSC" but unavailable to private sponsors who would engage in a P3 (Morallos et al. 2009, p. 30). Such advantages that need to be taken into account include public sponsors' exemption from taxes and insurance payments; disadvantages include increased disclosure and compliance requirements for public agencies as compared to private firms.

VfM is a mixed, hybrid quantitative-qualitative process, at least as it has been conducted in US states that have P3s, e.g., Virginia, and in other nations with a strong track record of alternative financing like the UK (State of Virginia 2011, p. 2; Morallos 2008, p. 4). The relevant quantitative factors for which to account include the all the aforementioned components of asset LCC. The qualitative factors impacting VfM generally vary with each agency's own set of priorities and standards; typical factors, however, include: contract quality, relative skills and resources of the project's various sponsors, and stated market interest for a project (Morallos et al. 2009, p. 31).

The VfM assessment itself occurs in two stages – initial and final, with both qualitative and quantitative assessments occurring at each phase. The initial assessment or "shadow bid model" phase occurs during the project development phase in order to effectively determine whether a given project is feasible for alternative procurement, but before private sponsors' bids are solicited (State of Virginia 2011, p. 11). Rather than comparing the PSC against actual bids, during the first of two assessments the PSC is compared to a "shadow bid," a hypothetical estimate of comparable private sector costs – both in terms of capital components and financial/operational risk (PwC 2010, p. 9). At this stage, public agencies conduct a variety of sensitivity analyses, analyzing different types of financial packages and delivery mechanisms. Re-evaluation of the PSC and shadow bid under these different contexts enables a robust analysis of project feasibility across an array of delivery, financing, and funding arrangements (State of Virginia 2011, p. 11; PwC 2010, p. 9; FHWA 2011, p. 13). The valuation processes underlying the initial "shadow bid model" VfM phase are detailed in Figure 7.

**Figure 7: Inputs and outputs for shadow bid model**




**Source: State of Virginia 2011, p. 15**

Should the shadow bid NPV provide a net cost savings compared to the PSC, then the agency conducts a second VfM analysis – this time with actual bids solicited from interested private sponsors. This two-stage evaluation allows for, at multiple points throughout the process, an agency to reconsider the pursuit of a P3, as well as an up-to-date assessment of submitted bids. Projects which incur a large pre-implementation delay –political or initial funding delays can push construction initiation far into the future – or are conceived during times of economic uncertainty may require a third VfM iteration. This third iteration would consist of new solicited bids reflecting the project's most current and up-to-date financial context, be it a rise in borrowing costs, weakened use projects, or an array of other possible factors(State of Virginia 2011, p. 12).

Throughout the existing research, we found little guidance on decision thresholds – at what point calculated cost savings are sufficient to justify alternative procurement. A survey of seven different international VfM practitioners suggested heavily weighting qualitative factors in final VfM assessment as a supplement to NPV. All seven government organizations – from H.M. Treasury in the UK to Partnerships British Columbia – reported that cost is only one factor in VfM final assessment decision-making (Morallos et al. 2009, pp. 28-29). Generally, "affordability calculations" – what the public sponsor of a project can afford in terms of project financing and funding – are done prior to the two-stage VfM; those numbers must then be met before entering into the project procurement process (Morallos et al. 2009, p. 28-29). However, no "hard and fast" rule exists on classifying P3-related procurement savings as sufficiently high or not – such a decision rests with the individual agency.

Choosing a proper discount rate for VfM is essential to accurately reflecting a project's related risks, financing costs, and time horizon as well. Four options emerge in choosing a discounting method:

1) Employing a single discount rate;
2) Explicit project risk valuation;
3) Cash flow risk markup; and
4) Time-declining discounting (Morallos et al. 2009, p. 30; Moore et al. 2004, p. 798).

Employing a single discount rate is the traditional practice of discounted cash flow/NPV calculations. A single rate may be project or sector specific, reflective of risks and costs for the chosen reference class, and could be used for both the PSC and alternative bids (Morallos et al. 2009, p. 30). Explicit project risk valuation takes monetized project risks and incorporates them into all projected cash flows for both the PSC and the alternative bids. Then, it becomes possible to use a risk-free rate for all cash flows, since they have been adjusted for risk already.

Cash flow risk markup is a similar concept, but the risk premiums apply only to the cash flows deemed to possess risk, rather than choosing an overall risk level (be it a single value or a mean or median risk) for all cash flows. Furthermore, the risk premium can either take the form of a monetized premium – as is done in explicit valuation – or it can simply be a higher discount rate to reflect the risk by proxy.

Finally, time-declining discounting applies to projects with intergenerational considerations, or for other projects where there is uncertainty about the choice of just one rate (Moore et al. 2004, p. 797). In this paradigm, there are two chosen rates – a higher rate in the near term (roughly for the first half of the projected cash flows) and a lower rate (often 50% less than the initial rate) for the cash flows further out in time (Moore et al. 2004, p. 798). Such a methodology is not useful for short-term or small scale projects as much as it is for large-scale facilities or for projects where considerable uncertainty about long-range economic performance exists.

In theory, at least, the choice of a discount rate should reflect the government's borrowing rate/cost of capital adjusted for project-related risks. Depending on which project delivery vehicle is under consideration, the basis for the discount rate varies. Ultimately, chosen rates – and methodologies for arriving at those rates – for discounting vary significantly between agencies. Partnerships Victoria employs a risk-free rate of 3% (real) plus a risk premium applied across all cash flows, the exact amount of which is contingent upon which risk range a project fits into – very low, low, or high (Partnerships Victoria 2001, p. 44). H.M. Treasury in the UK employs a risk free discount rate of 3.5%, following the single discount rate methodology, while Partnerships British Columbia employs a discount rate based on the private sector weighted average cost of capital – itself calculated by determining the cost of public debt and adding to it the project's risk premium (Partnerships BC 2005, p. 17; HM Treasury 2006, p. 21). Other governments have opted to employ the discount rate they have typically used for other projects – itself highly variable between different governments.

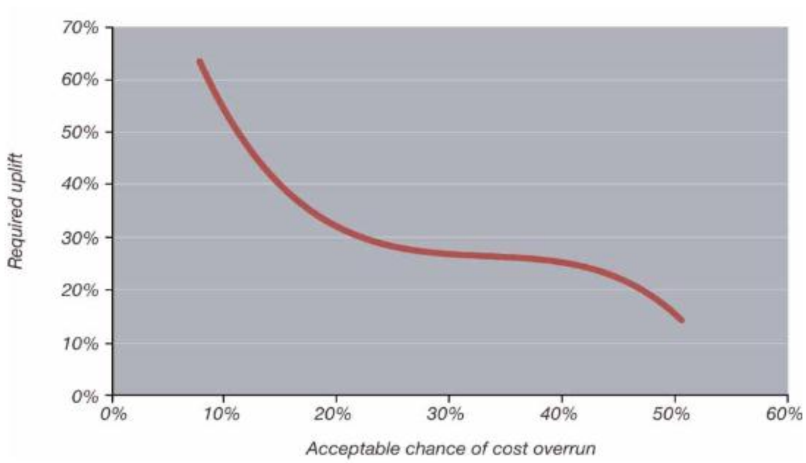
### Construction cost variability

While major project budgets all contain contingencies as per project controls protocol, the pervasiveness and consistency of significant major project construction cost overruns prompts better pre-implementation analysis and risk accounting. Randomization-based risk analysis for construction cost like Monte Carlo simulation – as is recommended for demand/revenue forecasting – is not useful for modeling the distribution of construction costs. Research has indicated (as will be covered in section 2c) that construction cost overruns are far from randomly-distributed or an otherwise stochastic and unpredictable occurrence. Very much to the contrary, there is an empirically observed and highly-consistent trend among construction cost overruns for new transportation facilities – with explicit distributions of average and median overruns by project mode and, in some cases, even geographic location (Flyvbjerg et al. 2004, p. 3; Flyvbjerg 2008, p. 5; Cantarelli 2009, p. 12; Odeck 2004, p. 52).

In lieu of such data, incorporating those modal overrun percentages directly into a project's budget – in addition to whatever contingency may exist already – is one viable option for more realistically accounting for overrun risk and "risk capital" (Flyvbjerg 2008, p. 12; Bialik 2010; Bruzelius et al. 2002, p. 150). Reference class forecasting is one application of this process. Reference class forecasting uses a group of completed, comparable facilities are used to develop "probability distributions for cost overruns for new projects similar in scope and risk to the projects in the reference class" (Flyvbjerg 2008, p. 12). For example, completed light rail systems in Portland, Houston, and

Sacramento, might be analyzed to help Denver predict the potential for cost increase for a potential future system of similar scope. While one project's cost in no way, shape, or form can predict another project's cost, a set of probability distributions of costs – based on a large sample of similar projects – can predict percentage likelihood of overruns. These probability distributions create a set of required "uplifts" – that is, an amount in excess to the allotted budget for a project that would reduce the risk of overrun to a pre-determined percentage. An example of reference class forecasting for a road project is shown in Figure 8, with the required uplifts derived from a probability distribution of similar road projects.

**Figure 8: Reference class forecast for a road project**




**Source: Flyvbjerg 2008, p. 14**

The lower the acceptable overrun risk, the greater the required increase in project budget to ensure an overrun does not occur. In this particular forecast, limiting the risk of a cost overrun to 10 percent, for example, would require an increase of nearly 60 percent in the project budget. For road projects in this particular research sample, a 50 percent chance of an overrun requires a 15 percent increase in project budget, while a 20 percent chance necessitates an increase of 32 percent (Flyvbjerg 2008, p. 16).

## c) Past performance of analytical tools

While alternative procurement and – specifically – the P3 arrangement have both earned a reputation for an increased ability to deliver projects within budgetary constraints, construction cost variability remains a central challenge to overcome, especially for user revenue-dependent facilities like toll roads (PWC 2010, p. 8; Iocobacci 2010, p. 23). Past construction cost forecasts in the transportation sector have generally not performed well. As a class, transportation infrastructure projects have consistently incurred significant construction cost overruns, with little increases in forecast accuracy over time (Flyvbjerg et al. 2004, p. 3; Flyvbjerg 2008, p. 5; Cantarelli 2009, p. 12; Odeck 2004, p. 52).

Of 258 global transportation projects, 90 percent were observed to have incurred cost overruns, with average escalation percentage (relative to initial cost projections) varying by mode – 45% for rail projects, 34% for fixed links (bridges and tunnels), and 20% for road facilities (Flyvbjerg et al. 2004, p. 3). Later research confirmed these overrun percentages and suggested differing rates of variability by mode – road forecasts experienced the least variability with a 30% standard deviation, while fixed links were highly-variable, with a 62% standard deviation (Flyvbjerg 2008, p. 5).

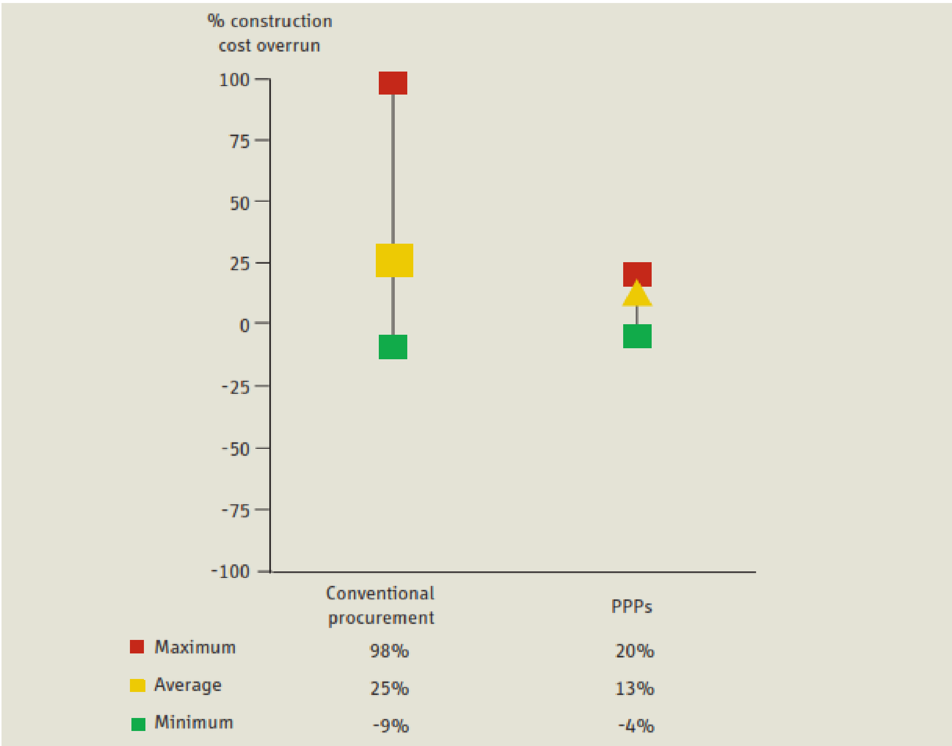
Surveys of large-scale Dutch transportation projects have found evidence of consistent, albeit smaller, overruns in construction cost as well. Compared to the global average overrun percentage (for all modes) of 28%, projects in the Netherlands averaged a 10% escalation, with a significant degree of variability by project type (Cantarelli 2009, p. 12). Fixed links incurred an average 4.5% overrun, road projects 18.5%, and rail projects 7.7% (Cantarelli 2009, p. 12). This survey terminated cost tracking the moment construction began on a facility, as compared to most other cost benchmarking that keep tracking until the project opens. As a result, the overrun percentages found in this particular analysis were significantly lower than in other cost benchmarking research.

Research on construction cost variability for Norwegian toll roads over a three year period in the early 1990s found a wide range of cost outcomes, both in terms of overruns and under-budget performances (Odeck 2004, p. 52). Out of 620 projects, 575 were classified as “very small” or “small,” with budgets less than $3 million or $18 million respectively. The smaller projects experienced the greatest percentage overruns on average – with 7.55% and 10.62% for very small and small projects compared to 2.46% for medium and -2.50% for large projects (Odeck 2004, p. 49). On a project-by-project basis, the overrun percentage ranged from -59% up to 183%, from half the project cost under initial construction budget to an almost triple cost overrun (Odeck 2004, p. 52). Overall, the sample incurred an average 7.88% overrun, with most of that mean driven by comparatively worse performance among smaller projects.

Despite these findings for all transportation infrastructure – and specifically road and toll road – projects as a class, past research and benchmarking have identified significantly better construction cost performance for P3s compared to traditional projects (Duffield et al. 2008, p. 25; Bain 2010, p. 3). A University of Melbourne benchmarking study on behalf of Infrastructure Australia – Australia's governmental infrastructure and P3 advisory bureau – found P3s to reduce construction cost overruns significantly. Of the 67 total projects analyzed – both traditionally-procured and P3 projects – P3s incurred an average cost overrun of just over 4%, compared to an average of 18% for the publically-procured projects (Duffield et al. 2008, p. 25). The relative market value of P3 projects to traditionally-procured projects in the sample was roughly equal, at $4.5 billion and $4.1 billion respectively. However, the aggregated dollar total for construction cost overruns varied by a factor of nearly 12 - $53 million on $4.5 billion of P3s versus $618 million on $4.1 billion of publically procured projects.

Large-sample aggregations of past studies on global construction cost performance for P3s relative to publically-procured projects reveal strikingly similar results. Combining 14 different studies of construction cost performance, P3s have been observed to outperform traditionally-procured projects dramatically – as shown in Figure 9.

Figure 9: Aggregated observed overruns in 14 studies – traditionally-procured facilities vs. P3s




Source: Bain 2010, p. 5

P3s, as observed in these studies, have capped overruns at 20% - as compared to nearly 100% for traditionally-procured projects – and have cut the average overrun roughly in half, from 25% to 13% (Bain 2010, p. 3). There does appear to be a cost advantage associated with P3s stemming from both reduced maximum overruns and increased cost certainty overall; in a 2010 report, PricewaterhouseCoopers, a consultancy, identified this increased cost certainty as an "economically and statistically significant" factor observed to be in P3s' favor (PwC 2010, p. 8).

A publication by AECOM Consult Team on recent US P3s suggests project delays lead to increased uncertainty and cost overrun; costs can increase "significantly" as a result of project delays, particularly as a result of conflict between and among stakeholders (ACT 2007, c.3 p. 72). These types of delays can occur before project construction – during the pre-implementation stage of the project – as well as during the project's construction phase. Pre-implementation opportunity costs are particularly critical for P3s, given that P3s have been shown to take, in a case study of Australian P3s, an average of

15% longer to reach contractual commitment and construction initiation than traditionally-procured projects (Duffield et al. 2008, p. 27). At the same time that P3s have had a longer pre-implementation phase than traditionally -procured projects, P3 cost escalation has also been tied to risks stemming from political and legal delays in the project pre-implementation phase (Bruzelius et al. 2002, p. 145).

Furthermore, road projects (both tolled and free facilities) seem particularly susceptible to pre-implementation cost escalations. In a survey of 28 Dutch road projects, these facilities were characterized by a mean escalation of almost 20% during the pre-implementation phase, as compared to just 6.3% during the construction phase (Cantarelli 2009, p. 10). The pre-implementation cost escalations were also much less predictable and, as such, generated more risk than the construction phase: standard deviation for pre-implementation overruns was found to be 50.9%, as compared to 26.7% for overruns during the construction phase. In a regression analysis, pre-implementation cost escalations were also found to be very tightly correlated with total percentage overrun; this analysis suggests the bulk of a given project's cost overrun as a percentage share of the total overrun occurs during pre-implementation (Cantarelli 2009, p. 11).

Cost escalation risks, however, remain significant during the construction phase. Past observations have found that annual average cost escalation from the decision to build until the commencement of facility operations is 4.64% excluding financing costs, which would push annualized escalations "considerably higher" (Flyvbjerg et al. 2004, p. 16). This risk, to some extent, may be smaller for P3s given that empirical evidence suggests P3s outperform traditionally procured projects from contractual commitment through facility completion (despite underperforming during pre-implementation). Past P3s have incurred an additional 2.6% delay during this period compared to nearly 20% for traditionally-procured projects (Duffield et al. 2008, p. 27).

## 3) Drivers of forecast inaccuracy and how to address errors

Prior research on forecast inaccuracies has divided explanations into three categories: technical explanations, psychological explanations, and political-economic explanations. Technical explanations are the most common, and they explain inaccuracy as the result of mis-specified forecasting models and/or unreliable data used as the basis for forecasts (Flvbjerg 2008, p. 6). Psychological explanations pinpoint "optimism bias" as the key driver in forecast inaccuracies; in this paradigm, "a cognitive predisposition found with most people to judge future events in a more positive light than is warranted by future experience" drives inaccurate forecasts (Flyvbjerg 2008, p. 6). Political-economic explanations, meanwhile, contextualize forecast inaccuracies as the result of strategic misrepresentation stemming from conflicting or perverse intra-organizational incentives (Flyvbjerg 2005b, p.9). Political-economic explanations also assume the presence of optimism bias, asserting the need to make projects appear comparatively more attractive than competitors' projects increases the frequency and impact of optimism bias (Danninger et al. 2005, p.16).

Errors of model misspecification or inadequate data are certainly common; however the strong upward bias in forecasting suggests that optimism bias and strategic misrepresentation are more to blame for inaccuracy than pure technical error alone (Flyvbjerg 2008, p. 6; Odeck 2004, p. 45). Given scarce funding for new facilities, projects compete with each other for what public funding is available. In this sense, all forecasts in the project vetting process may be overly-optimistic. Those who win the competition for public funding tend to be most optimistic forecasters with, in turn, the most troubled projects (Australia DIT 2011b, p.28). This competition (in conjunction with the over-confidence observed as fueling the "planning fallacy" for large-scale projects) creates a strong set of perverse incentives to overstate the benefits of a project while understating costs, in hopes of best positioning a given project to secure scarce funding resources (Kahneman and Tversky 1979a, pp. 282-283; Kahneman and Tversky 1979b, p. 12; Flyvbjerg 2008, p. 8; Odeck 2004, p. 45). In this sense, there is a major disconnect between the incentive of project promoters and empirical observations on the performance of past similar projects (Flvbjerg 2008, p. 8).

A large part of forecast inaccuracy stems from the myriad players involved in the project development process. This consideration is particularly relevant for P3s, given the number of parties involved across both public and private sectors – from governments to financiers – and each participant's competing incentives. The only immediate stakeholders with any sort of incentive, then, to minimize optimism bias and work to make forecasts in any way more accurate are the overarching government agencies that do not necessarily serve as direct project sponsors. By contrast, the parties which are most involved in projects as sponsors on a day-to-day basis – local politicians and transportation analysts/authorities along with individual consultancies oftentimes developing and implementing the forecasts – are those with the least incentive to remedy over-optimistic forecasts (Wachs 1990, p.144; Pickrell 1992, p.158).

Pressures on those managing analytical tools – and thus incentives to misrepresent forecasts – go up during the competitive bidding process, causing demand and revenue forecasts to be inflated across the board, in hopes of gaining competitive advantage over other bidders (Harvey 2011, p.26). In this paradigm, project consortia knowingly generate flawed forecasts "because they have an interest in making errors" and thus winning the contract (Prudhomme 2004, p.34). While private consortia win contracts, public sponsors can capture the benefits associated with projects moving forward; as such, public sponsors "are also quite willing to be misled" (Prudhomme 2004, p.35).

While generally speaking, under-estimations of demand are preferred to overestimations – more conservative forecasts will attain greater financial ratings than aggressive forecasts – wide-ranging performance and unpredictability of forecasts on both extremes can create problems for P3 projects (Fitch 2007, p.5). Liquidity, ultimately, is the issue for tolled P3 facilities. Those who structure the complex financials underlying a P3 transaction for tolled facilities need to build in flexibility and risk, on use upside and downside (Bain 2009, p. 1).

Ultimately, however, fundamental changes to models employed would do little to actually reduce the optimism bias which has been observed to be a core driver of poor performance. Rather, small-scale technical adjustments to models can be made, but major focus should be given on risk analysis and striving to reduce incentives for optimism bias and extreme forecasting error. From the perspective of ratings agencies and potential project financiers, an understanding of project performance under a variety of contexts –and a financial structuring that "provides flexibility and protections" is key to attract private financial interest (Forsgren 2006, p. 4; Fitch 2007, p. 4). An example of the sort of sensitivity and risk analysis employed by ratings agencies is included as an Appendix A for Caltrans' reference; this type of stress testing does not represent advanced facility analytics so much as the *baseline expectation* of analysis private firms require in order to assess an asset's creditworthiness and financial viability (Fitch 2007, pp. 20-21).

Realistically, however, solid risk analysis and an understanding of project contingencies do not make inherently troubled projects any more attractive or feasible. Bad projects should be winnowed out far before forecasting occurs as part of the project evaluation process (Forsgren 2006, p. 4). Improving forecasting performance can contribute to better projects; but it is secondary to eliminating incentives for overly optimistic forecasts and the pursuit of suboptimal, misrepresented projects. Increasing the role of treasury officials – or other government officials in posts related to disbursement of liquidity for project financing purposes – has been cited as a way to "dampen both enthusiasm and potential over-optimism by project proponents from road agencies," as was the mandatory calculation of a commercial project case – omitting social benefits – to supplement the government's project business case (Australia DIT 2011b, p.36).

Generally, re-examining the project tendering stage has strong potential to limit over optimistic forecasts and expectations somewhat; bidders should be allowed to set toll levels in their forecasts – so that capital cost and use risk can be more adequately reflected. Governments could also help the process by being more open with their data. According to the interviews we conducted as part of Section V (Institutional Capacity), most agencies use proprietary forecasting models and only contract out when projects reach requisite levels of complexity. By making these models and data inputs available to bidders, agencies could save bidders money and increase the overall accuracy and applicability of forecasts. After all, bidders will be making additions to existing models and parameters, rather than developing entirely new ones on their own.

One other, albeit more theoretical than practical, consideration for aligning presently mismatched incentives would be a focus on consortia longevity – that is, forcing those who make forecasts be accountable if their forecasts are grossly overblown (Australia DIT 2011b, p.38). Requiring equity participation by all participants is problematic, as are other forms of binding non-exit clauses in a contract; both may scare off potential bidders and are difficult and potentially costly to enforce (Australia DIT 2011b, p.38). Such equity participation requirements and non-exit clauses for forecasters could merely drive up legal costs that private forecasting consultants must pay (rather than actually increasing forecast accuracy). Any sort of increase in recourse or forecast liability will also, inevitably, drive up project costs for private sponsors as it represents a significant transfer of facility use risk from a public sponsor to a private firm.

However, to the extent that the threat of realistic recourse in the event of intentionally and strategically faulty forecasts appears real, over-optimism in the forecasting process could be reduced. Longevity weighting to certain bids – that is, weighting those consortia which appear to be intent on sticking together over the course of the long term positively relative to other bidders – during bid assessment could also help to reduce optimism bias in forecasts.

## Conclusion

In this report, we have: 1) defined the state of practice of MPO forecasting and contrasted it from the analytical tools required for priced facilities; 2) surveyed the suite of analytical tools employed to assess the feasibility of priced facilities; and 3) examined these tools' and methods' past performances, offering explanations for inaccuracy and suggestions to improve the overall forecasting process. We have focused on state-of-the-practice methodologies rather than state-of-the-art methodologies because that the marginal benefits of employing highly complex analytical tools do not appear to be worth the investment for most places doing analysis. Furthermore, this report has underscored the reality of project – and P3 – evaluation: the analytical tools themselves are but one source of inaccuracy in the process.

Incremental technical improvements to the forecasting process can, and should, be made for P3s – given increased demand from private-side project financiers for an understanding of project contingencies and multiple performance scenarios. Priced facilities and P3s do not require an entirely new suite of analytical tools, but rather, the greater adoption of existing private sector financial risk assessment tools. Randomization in risk and sensitivity analysis, the assignment of probability distributions – both normal distributions as well as pre-defined risk profiles where available – and an overarching commitment to ask "what if" can help pin down the risks for all participants in P3s. P3s, because of the incorporation of alternative sources of financing, require a range of values to be derived and accounted for – for an understanding of not just one potential outcome, but a spectrum of potential outcomes.

As flawed as traffic and revenue forecasts are, the technical problems with forecasting appear to be less of an issue than the moral hazard that accompanies project bidding. There is intense inter-consortia competition for funding, artificially inflating use figures while deflating costs. And those who provide faulty forecasts for self-gain are in no way liable for poor forecast outcomes beyond declaring bankruptcy on the project (see section 6). In this paradigm, there is little disincentive to inflate use forecasts while deflating cost projections – the outcome of this game is, of course, that the winning bidder for a project generates the forecast that is politically most useful, but financially least credible, severely damaging the project's long term feasibility.

Moreover, if a project does not perform well in the long-term, the public sector retains residual financial risk; governments will not allow a faltering private toll road to close. Within the context of cost-benefit analysis, social benefits tend to be over-estimated relative to the possible revenue or financial outcomes associated with past projects. Essentially, two different types of forecasts emerge for major projects – the government's business case, which includes social benefits, and a commercial case, which excludes social benefits. Given the incentives on the partners in each sector, however, the benefits are inherently over-estimated.

As a result of the deeply entrenched interests within the project bidding process, adjusting incentives will be neither painless nor immediate, but it appears to be the only real way to reduce forecasting inaccuracy. Increased sensitivity and risk analysis will provide an understanding of potential project outcome ranges but the wide variability of past forecast performance needs to be narrowed and vetted before private sector actors are likely to become willing to invest large amounts of capital. Given that performance is detached from compensation, governments can also work to increase recourse *ex-post*.

# References


AECOM Consult Team (ACT). (2007). *Case Studies of Transportation Public-Private Partnerships in the United States*. A report prepared for Office of Policy and Governmental Affairs, Federal Highway Administration. Arlington, VA: July 7, 2007.

Allen, William G. Jr.; Agnello, Paul. (2003). *Modeling Comercial Vehicle Travel*. Baltimore Metropolitan Council.

Allen, William G. Jr; Schmitt, David. (2005). "Using Your Model Effectively." Presented at *Tenth TRB Transportation Planning Applications Conference*. 25 April 2005.

Anas, Alex; Liu, Yu. (2007). "A Regional Economy, Land Use, and Transportation Model (RELU-TRAN); Formulation, Algorithm Design, and Testing." *Journal of Regional Science*. Volume 47, Number 3, pp.415-455.

Armoogum, Jimmy; Madre, Jean-Loup; and Bussiere, Yves. (2009). "Measuring Uncertainty in Long-Term Travel Demand Forecasting From Demographic Modelling." *IATSS Research*. Volume 33, Number 2, pp. 9-20.

Australia Department of Infrastructure and Transport (DIT). (2011a). *Review of Traffic Forecasting Performance: Toll Roads*. Bureau of Infrastructure, Transport and Regional Economics. June 2011.

Australia Department of Infrastructure and Transport (DIT). (2011b). *An investigation of the causes of over-optimistic patronage forecasts for selected recent toll road projects*. Revised final report prepared by GHD Australia for Department of Infrastructure and Transport. December 8, 2011.

Baez, Gustavo A. (2004). *Investment-Quality Surveys*. Wilbur Smith Associates.

Bain, Robert; Wilkins, Michael. (2002). *Credit Implications of Traffic Risk in Start-Up Toll Facilities*. Standard & Poor's – Infrastructure Finance. September 2002.

Bain, Robert. (2009). *Error and optimism bias in toll road traffic forecasts*. Institute for Transportation Studies, University of Leeds. February 28, 2009.

Bain, Robert. (2010). "Construction risk – What risk?" *Project Finance International*. 10 February 2010, pp. 46-50.

Ben-Akiva, Moshe; Bolduc, D.; Bradley, M. (1993). "Estimation of travel choice models with randomly distributed value of time." *Transportation Research Record: Journal of the Transportation Research Board*. Volume 1413, pp. 88-97.

Ben-Akiva, Moshe. (2008). *Transportation Revenue Forecasting: Theory and Models*. Transportation Systems Analysis: Demand & Economics, Massachusetts Institute of Technology.

Bialik, Carl. (2010). *Where Infrastructure Estimates Come Up Short*. The Wall Street Journal. New York, NY: October 16, 2010. Retrieved November 19, 2010. Available at: http://online.wsj.com/article/SB10001424052748704300604575554121638637724.html.

Bowman, John L.; Bradley, Mark; Shiftan, Yoram; Lawton, T. Keith; Ben-Akiva, Moshe. (1998). *Demonstration of An Activity-Based Model System for Portland*. Published Draft.

Bowman, John L.; Gopinath, Dinesh; Ben-Akiva, Moshe. (2002). *Estimating the probability distribution of a travel demand forecast*. Massachusetts Institute of Technology – Department of Civil and Environmental Engineering.

Brownstone, David; Small, Kenneth A. (2005). “Valuing time and reliability: assessing the evidence from road pricing demonstrations.” *Transportation Research Part A*. Volume 39, pp. 279-293.

Bruzelius, Nils, Bent Flyvbjerg, and Werner Rothengatter. (2002). Big Decisions, big risks: Improving accountability in mega projects. *Transport Policy*. Volume 9, pp. 143-154.

Burris, Mark W.; Konduru, Karun K.; Swenson, Chris R. (2004). “Long-Run Changes in Driver Behavior Due to Variable Tolls.” *Transportation Research Record: Journal of the Transportation Research Board*. Volume 1864, pp. 78-85.

Calfee, John; Winston; Clifford; Stempski, Randolph. (2001). “Econometric Issues in Estimation Consumer Preferences from Stated Preference Data: A Case Study of the Value of Automobile Travel Time.” *The Review of Economics and Statistics*. Volume 83, Number 4, pp. 699-707.

Cambridge Systematics. (2005). *MnPASS System Study Tehcnical Memorandum #3: Travel Demand Forecasting Approach*. Prepared for Minnesota Department of Transportation by Cambridge Systematics, Inc. 3 February 2005.

Cantarelli, Chantal C. (2009). *Cost overruns in Dutch transportation infrastructure projects*. Delft University of Technology. Conference Presentation 19-20 November 2009. Antwerp, Belgium.

Chung, Jin-Hyuk and Goulias, Konstadinos G. (1997). “Travel Demand Forecasting Using Microsimulation: Initial Results from Case Study in Pennsylvania.” *Transportation Research Record*. Volume 1607, pp. 24-30.

Danforth, David R. (2006). *Toll Traffic and Revenue Forecasts*. Wilbur Smith Associates. 21 November 2006.

Danninger, Stephan; Cangiano, Marco; Kyobe, Annette. (2005). *The Political Economy of Revenue-Forecasting Experience from Low-Income Countries*. International Monetary Fund (IMF). IMF Working Paper WP/05/2. January 2005.

Dehghani, Youssef; Adler, Thomas; Doherty, Michael W.; Fox, Randy. (2007). “Development of a New Toll Mode-Choice Modeling System for Florida’s Turnpike Enterprise.” *Transportation Research Record*. Volume 1858 (2003), pp. 9-17.

Duffield, Colin; Raisbeck, Peter; Xu, Ming. (2008). *National PPP Forum - Benchmarking Study, Phase II*. Univeristy of Melbourne, Melbourne Engineering Research Institute (MERIT).Retrieved November 4, 2010. Available at: http://www.infrastructureaustralia.gov.au/files/National_PPP_Forum_Benchmarking_Study_Ph2_dec08.pdf.

Federal Highway Administration (FHWA). (2011a). *Spreadsheet Model for Induced Travel Estimation – Managed Lanes (SMITE-ML)*. 5 April 2011. Retrieved February 4, 2012. Available at: http://www.fhwa.dot.gov/steam/smiteml.htm.

Federal Highway Administration (FHWA). (2011b). *Financial Analysis of Transportation-Related Public Private Partnerships*. Office of Inspector General (OIG). Report Number CR-2011-147. 28 July 2011.

Fitch Ratings. (2007). *Global Toll Road Rating Guidelines*. Fitch Global Infrastructure and Project Finance. March 6, 2007.

Flyvbjerg, Bent; Skamris, Mette K.; Buhl, Soren L. (2004). What Causes Cost Overrun in Transport Infrastructure Projects? *Transport Reviews*. Volume 24, Number 1, pp. 3-18.

Flyvbjerg, Bent. (2005a). "Measuring inaccuracy in travel demand forecasting: methodological considerations regarding ramp up and sampling." *Transportation Research Part A*. Volume 39, pp. 522-530.

Flyvbjerg, Bent. (2005b). *Policy and Planning for Large Infrastructure Projects: Problems, Causes, Cures.* World Bank. World Bank Policy Research Working Paper 3781. December 2005.

Flyvbjerg, Bent. (2008). "Curbing Optimism Bias and Strategic Misrepresentation in Planning: Reference Class Forecasting in Practice." *European Planning Studies*. Volume 16, Issue 1, pp. 3-21.

Forsgren, Kurt. (2006). "Evaluating Deep Future Toll Concessions – Frequently Asked Questions." Presented at *IBTTA 74th Annual Meeting and Exhibition: 19 September 2006, Dallas, TX.*

George, Cherian; Streeter, William; and Trommer, Scott. (2003). *Bliss, Heartburn, and Toll Road Forecasts*. Fitch Ratings Public Finance – Project Finance. 12 November 2003.

Harvey, Mark. (2011). *Traffic forecasting performance of PPP and toll roads*. Australian Department of Infrastructure and Transport. Bureau of Infrastructure, Transport, and Regional Economics.

Hensher, David A. and Goodwin, Phil. (2003). *Using values of travel time savings for toll roads: Avoiding some common errors*. Working Paper. January 2003.

Her Majesty's Treasury (HM Treasury). (2006). *Value for Money Assessment Guide*. United Kingdom – Her Majesty's Treasury. November 2006.

Jones, M.; Hilbers, P.; Skack, G. (2004). *Stress Testing Financial Systems: What to do When the Governor Calls*." Working Paper WP/04/127, International Monetary Fund, Washington DC, July 2004, pp.3-6.

JP Morgan. (1997). "Examining toll road feasibility studies." *Municipal Finance Journal*. Volume 18, Issue 1, pp.1-12.

Kahneman, D.; Tversky, A. (1979a). "Prospect theory: An analysis of decision under risk." *Econometrica*. Volume 47, pp.313-327.

Kahneman, D.; Tversky, A. (1979b). "Intuitive prediction: Biases and corrective procedures" in S. Makridakis and S.C. Wheelwright: *Studies in the Management Sciences: Forecasting*. Amsterdam: North Holland, p.12.

Khan, Ata M. (2009). *Risk Factors in Toll Road Life Cycle Analysis*. Carleton University.

Kriger, David. (2005). *Traffic Revenue Forecasting for Roads and Highways: Concerns, Methods, and a Checklist for Practitioners*. Prepared for "Expert Forum on Road Pricing and Travel Demand Modeling." 14-15 November 2005: Alexandria, VA.

Lam, T.C.; Small, K. (2001). "The value of time and reliability: measurement from a value pricing experiment." *Transportation Research Part E*. Volume 37, pp.231-251.

Lemp, Jason D.; Kockelman, Kara M. (2009). "Understanding and Accommodating Risk and Uncertainty in Toll Road Projects: A Review of the Literature." Presented at *88th Annual Meeting of the Transportation Research Board, January 2009.* Washington D.C.

Masters, T. (1993). *Practical Neural Networks Recipes in C++*. London: Academic Press.

Miller, John B. (2008). *Life Cycle Delivery of Public Infrastructure: Precendent and Opportunities for the Commonwealth*. Pioneer Public Policy Institute. White Paper Number 44. December 2008.

Moore, Mark A.; Boardman, Anthony E.; Vining, Aidan R.; Weimer, David L.; Greenberg, David H. (2004). "'Just Give Me a Number!' Practical Values for the Social Discount Rate." *Journal of Policy Analysis* *and Management – The Journal of the Association for Public Policy Analysis and Management*. Volume 23, Issue 4: pp. 789-812.

Morallos, Dorothy. (2008). "A Review of Value for Money (VfM) Analysis for Comparing Public Private Partnerships to Traditional Procurments." Presented at *Lessons Learned from Public Private Partnerships for Infrastructure*," University of Southern California. 27 March 2008.

Morallos, Dorothy; Amekudzi, Adjo; Ross, Catherine; Meyer, Michael. (2009). "Value for Money Analysis in U.S. Transportation Public-Private Partnerships." *Transportation Research Record: Journal of the Transportation Research Board*. Volume 2115, pp. 27-36.

Mwalwanda, Christopher. (2009). "Long-term Demand Forecasting of Managed Lanes: Challenges in Addressing Key Influential Risk Parameters." Presented at *2nd International Symposium on Freeway and Tollway Operations*. 23 June 2009.

National Cooperative Highway Research Program (NCHRP) Synthesis 364. (2006). *Estimating Toll Road Demand and Revenue*. Transportation Research Board: Washington, D.C.

Nourzad, Firouzeh. (2004). *Estimating Demand for Value Pricing Projects: State of the Practice*. Presented at "AMPO Travel Modeling Subcommittee Meeting." 2 March 2004.

Nunez, Antonio. (2007). *Sources of Errors and Biases in Traffic Forecasts for Toll Road Concessions*. Universite Lumiere Lyon 2. Doctoral Thesis. 5 December 2007.

Odeck, James. (2004). "Cost overruns in road construction – what are their sizes and determinants?" *Transport Policy*. Volume 11, pp. 43-53.

Odeyinka, Henry A.; Lowe; John G.; Kaka, Ammar. (2002). "A construction cost flow risk assessment model." Presented at *18th Annual ARCOM Conference*, University of Northumbria. 2-4 September 2002.

Oppenheim, Norbert. (1995). *Urban Travel Demand Modeling: From Individual Choices to General Equilibrium*. New York, NY: Wiley-Interscience.

Oregon Department of Transportation (ORDOT). (2009). *Tolling White Paper 3: Travel Demand Model Sufficiency*. Prepared for the Oregon Department of Transportation by Parsons Brinckerhoff.

Parthasarathi, Pavithra; Levinson, David. (2009). "Post-Construction Evaluation of Forecast Accuracy." Department of Civil Engineering, University of Minnesota. February 16, 2009.

Partnerships British Columbia (BC). (2005). *Project Report: Achieving Value for Money, Sea-to-Sky Highway Improvement Project*. British Columbia Ministry of Transportation. December 2005.

Partnerships Victoria. (2001). *Public Sector Comparator Technical Note*. Department of Treasury and Finance – State of Victoria, Australia. June 2001.

Pickrell, Don H. (1992). "A Desire Named Streetcar: Fantasy and Fact in Rail Transit Planning." *Journal of the American Planning Association*. Volume 58, Number 2, Spring 1992, pp. 158-176.

PriceWaterhouseCoopers (PwC). (2010). *Public-Private Partnerships: The US Perspective*. Retrieved November 19, 2010. Available at: http://www.pwc.com/us/en/capital-projects-infrastructure/publications/assets/Public_Private_Partnerships.pdf.

Prozzi, Jolanda; Persad, Khali; Flanagan; Kate; Loftus-Otway, Lisa; Porterfield, Beth; Rutzen, Beatriz; Zhao, Mengying; Prozzi, Jorge; Robertson, Chris; and Walton, C. Michael. (2010). *Toll Roads: What We Know About Forecasting Usage and the Characteristics of Texas Users*. Center for Transportation Research, The University of Texas at Austin. January 2010.

Prudhomme, Remy. (2004). "Infrastructure and Development." Presented at *Annual Bank Conference on Development Economics (ABCDE)*. May 3-5, 2004. Washington, DC.

Shiftan, Yoram; Ben-Akiva, Moshe; Proussaloglou, Kimon; de Jong, Gerard; Popuri, Yasasvi; Kasturirangan, Krishnan; and Bekhor, Shlomo. (2003). *Activity-Based Modeling as a Tool for Better Understanding Travel Behaviour*. Presented at the "10th International Conference on Travel Behaviour Research." 10-15 August 2003: Lucerne, Switzerland.

Skitmore, R. Martin; Ng, Thomas S. (2003). "Forecast Models for Actual Construction Time and Cost." *Building and Environment*. Volume 8, Issue 8: pp. 1075-1083.

Small, Kenneth A.; Winston, Clifford; and Yan, Jia. (2005). *Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability*. 19 January 2005. Draft Publication.

Smith, Don R.; Chang-Albitres, Carlos; Stockton, William R.; and Smith, Craig. (2004). *Estimating Revenues Using A Toll Viability Screening Tool*. Texas Transportation Institute, Texas A&M University.

Spear, Bruce D. (2005). *A Summary of the Current State of the Practice in Modeling Road Pricing*. Prepared for "Expert Forum on Road Pricing and Travel Demand Modeling." 14-15 November 2005: Alexandria, VA.

Standard and Poor's (S&P). (2005). *Traffic Forecasting Risk Study 2005: Through Ramp-Up and Beyond*. Standard & Poor's Investor's Services. London, UK.

State of Virginia. (2011). *PPTA Value for Money Guidelines*. Office of Transportation Public Private Partnerships. April 2011.

Sullivan, Edward. (2000). *Continuation Study to Evaluate the Impacts of the SR-91 Value-Priced Express Lanes: Final Report*. Submitted to State of California Department of Transportation (Caltrans) Traffic Operations Program, HOV Systems Branch. Submitted December 2000.

Transportation Corridor Agencies (TCA). (2012a). *Annual Results (133/241/261) Toll Roads*. Retrieved February 4, 2012. Available at: https://www.thetollroads.com/aboutus/investorinformation/transactiontables/fe_annualresults.php.

Transportation Corridor Agencies (TCA). (2012b). *Annual Results (73) Toll Road*. Retrieved February 4, 2012. Available at: https://www.thetollroads.com/aboutus/investorinformation/transactiontables/sj_annualresults.php.

Trommer, Scott. (2006). "Mitigating Toll Road Forecasting Risks." Presented at *Workshop on Public-Private Partnerships in Highways: Institutional, Legal, Financial and Technical Aspects*: April 3, 2006, New York, NY.

United States Department of Transportation (USDOT). (1997). *The Value of Saving Travel Time: Department Guidance for Conducting Economic Evaluations*. 9 April 1997: Washington, D.C.

URS Corporation (URS). (2004). *Estimating Demand for Value Pricing Projects: State of the Practice*. Prepared for North Central Texas Council of Governments by URS Corporation with Urban Analytics, Inc. 26 March 2004. Final Draft Report.

Vanasse Hangen Brustlin, Inc. (VHB). (2006). *Review of Managed Lanes Forecasting Techniques*. Prepared for Metropolitan Washington Council of Governments National Capital Region Transportation Planning Board. 22 September 2006.

Vollmer Associates. (2003). *Draft Final Traffic and Revenue Report*. Prepared for Foothill/Eastern Transportation Corridor by Vollmer Associates LLP. 10 September 2003.

Vollmer Associates. (2004). *183A Project Traffic and Revenue Study: Final Report*. Prepared for Texas Department of Transportation (TxDOT) by Vollmer Associates LLP. 16 December 2004.

Vovsha, P.; Davidson, W.; and Donnelly, R. (2005). "Making the state of the art the state of the practice: advanced modeling techniques for road pricing." *Expert Forum on Road Pricing and Travel Demand Modeling*. Retrieved February 12, 2012. Available at: http://tmip.fhwa.dot.gov/clearinghouse/docs/DOT-OST-P-001-06.

Wachs, Martin. (1990). "Ethics and Advocacy in Forecasting for Public Policy." *Business & Professional Ethics Journal*. Volume 9, Issue 1-2. Papers from the 1990 Conference on Moral Problems in the Professions (Spring-Summer, 1990), pp.141-157.

Waddell, Paul. (2011). *Dynamic Microsimulation: UrbanSim*. TMIP Webinar series on land use forecasting methods. 8 June 2011.

## Appendix A: Fitch Ratings financial analysis guidelines for toll roads

| | Fitch Base Case | Fitch Stress Case |
|---|---|---|
| **Construction** | | |
| **Simple Project** | | |
| Cost | 0%–5% overrun | 5%–10% overrun |
| Schedule | Zero to three-month delay | Three- to six-month delay |
| **Complex Project** | | |
| Cost | 0%–10% overrun | 10%–20% overrun |
| Schedule | Three- to 12-month delay | Six- to 24-month delay |
| **O&M Growth** | | |
| Established | Five- to 10-year historical average, excluding one-time savings, with a minimum of inflation | Base-case assumption plus 1% |
| Start-Up | Adjust initial year base by 0%–10%<br>Inflation plus 1%–2%, excluding start-up and ramp-up costs | Base-case assumption plus 1% |
| **Traffic Growth** | | |
| **Established** | | |
| First 10 Years of Forecast | Five- to 10-year historical average adjusted for asset maturity, capacity constraints, expected demand and peer group profile | Base-case assumption |
| Years 11–30 of Forecast | Tapered reductions based on the above factors down to the low single digits approaching expectations for regional traffic growth | Base-case assumption minus 0%–1% |
| Years 31–50 of Forecast | 0%–1% growth depending on the facility profile | 0% growth |
| **Start-Up** | | |
| Opening Year Base Traffic | Lag in likely economic growth/development between the base/calibration year of the study and planned construction completion of at least five years | Base-case assumption accompanied by a slower ramp-up in economic development |
| | Discount value of time assumption by 25%–50% | Discount value of time assumption by 50%–75% |
| | ETC violation rates up to 10% | ETC violation rates up to 15% |
| | Discount truck traffic levels by 25%–50% on regional roads and 15%–30% on national roads | Discount truck traffic levels by 50%–75% on regional roads and 20%–40% on national roads |
| Years 2–10 Forecast | Underlying economic fundamentals of project based on traffic study and peer group profile | Base-case assumption accompanied by acceleration of competing network facilities |
| | Ramp-down of discount on value of time assumption over 3–7 years | Ramp-down of discount on value of time assumption over 5–9 years |
| | Ramp-down of ETC violation rates to 5% over 3–5 years | Ramp-down of ETC violation rates to 5% over 5–9 years |
| | Increase in truck traffic at 1%–3% above auto traffic volumes | Increase in truck traffic at 0%–2% above auto traffic volumes |
| Opening Year Base Traffic | Lag in likely economic growth/development between the base/calibration year of the study and planned construction completion of at least five years | Base-case assumption accompanied by a slower ramp-up in economic development |
| | Discount value of time assumption by 25%–50% | Discount value of time assumption by 50%–75% |
| | ETC violation rates up to 10% | ETC violation rates up to 15% |
| | Discount truck traffic levels by 25%–50% on regional roads and 15%–30% on national roads | Discount truck traffic levels by 50%–75% on regional roads and 20%–40% on national roads |
| Years 2–10 Forecast | Underlying economic fundamentals of project based on traffic study and peer group profile | Base-case assumption accompanied by acceleration of competing network facilities |
| | Ramp-down of discount on value of time assumption over 3–7 years | Ramp-down of discount on value of time assumption over 5–9 years |
| | Ramp-down of ETC violation rates to 5% over 3–5 years | Ramp-down of ETC violation rates to 5% over 5–9 years |
| | Increase in truck traffic at 1%–3% above auto traffic volumes | Increase in truck traffic at 0%–2% above auto traffic volumes |
| Years 11–30 of Forecast | Tapered reductions for asset maturity, capacity constraints, expected demand and peer group profile approaching expectations for regional traffic growth | Base-case assumption minus 0%–1% |
| Years 31–50 of Forecast | 0%–2% growth depending on the project profile | 0%–1% growth |
| **Toll Rate Increases (Minimal Elasticity Levels)** | | |
| **Established** | | |
| First 10 Years of Forecast | CPI plus 50–200 bps | CPI plus 0–100 bps |
| Years 11–20 of Forecast | CPI plus 0–50 bps | CPI minus 0–50 bps |
| Years 21–30 of Forecast | CPI minus 0–50 bps | CPI minus 50–100 bps |
| Years 31–50 of Forecast | CPI minus 50–100 bps | CPI minus 100–150 bps |
| **Start-Up** | | |
| First 10 Years of Forecast | CPI plus 0–100 bps | CPI plus 0–50 bps |
| Years 11–20 of Forecast | CPI plus 0–50 bps | CPI minus 0–50 bps |
| Years 21–30 of Forecast | CPI minus 0–50 bps | CPI minus 50–100 bps |
| Years 31–50 of Forecast | CPI minus 50–100 bps | CPI minus 100–150 bps |

O&M – Operations and maintenance. bps – Basis points. ETC – Electronic toll collection. CPI – Consumer Price Index. Note: The economic analysis developed using the above methods will be subject to legal and policy constraints. This analysis assumes that all standard protections to minimize construction risk have been incorporated. Simple projects are assumed to be those that do not have difficult construction conditions or significant environmental challenges and do not incorporate significant water crossings, bridges or tunnels. Fitch will use applicable historical local, regional and/or national inflation indices to develop inflation-indexed inflators. The toll increase rates identified are assumed to be minimal elasticity levels in normal, low to moderate inflationary environments that are in the middle of the economic cycle. Higher rates of toll increases in the finance plan will be incorporated with reasonable diversion levels. Deal structures that limit toll increases to inflationary or subinflationary levels may be discounted to a lesser extent in the long run. Fitch will develop alternative scenarios in periods of very high or very low inflation vis-à-vis historical trends.

| | Fitch Base Case | Fitch Stress Case |
|---|---|---|
| **Financial Ratios (x)** | | |
| **User-Pay (Amortizing Structures)** | | |
| Minimum DSCR | 1.30 | 1.00 (including internal liquidity in the near term) |
| Minimum LLCR | 1.50 | 1.25 |
| Minimum PLCR | 1.75 | 1.50 |
| **User-Pay (Long-Dated Negative or Nonamortizing Structures)** | | |
| Minimum DSCR | 1.30 | 1.00 |
| Minimum PLCR | 3.00 | 2.00 |
| **Shadow Toll and Availability Payment (Amortizing Structures)** | | |
| Minimum DSCR | 1.20 | 1.05 |
| Minimum LLCR | 1.30 | 1.20 |
| **Refinancing Risk** | | |
| Interest Rate Assumption | Current rates plus 200 bps | Current rates plus 400 bps |
| **Discount Rates on Future Cash Flows** | | |
| For LLCR/PLCR Calculations | Weighted-average cost of debt | Weighted-average cost of debt |
| For Asset Valuation | Weighted-average cost of capital (equity and debt) | Weighted-average cost of capital (equity and debt) |

DSCR – Debt-service coverage ratio. PLCR – Project life coverage ratio. LLCR – Loan life coverage ratio. Note: Minimum ratios identified are largely applicable to established facilities. Higher ratios may be necessary on projects with meaningful construction and traffic forecasting risks. For negative or nonamortizing debt structures, Fitch will evaluate the minimum PLCR during the planned window of refinancing or beyond 20 years, whichever is shorter. Interest rate assumptions to evaluate debt structures with refinancing risk are assumed in the middle of the economic cycle. Upward and downward adjustments will be made in periods of very low and very high inflation to incorporate then current probabilities for interest rate movements.

**Source: Fitch (2007)**