TRANSPORTATION PROJECT OUTCOMES UNDER UNCERTAINTY: AN EXAMINATION OF BENEFIT-COST RATIOS AND OTHER IMPACTS

Daniel J. Fagnant
The University of Texas at Austin – 6.9, E. Cockrell Jr. Hall
Austin, TX 78712-1076
danfagnant@hotmail.com

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

The following is a pre-print and the final publication can be found in the *Transportation Research Record,* No. 2303, 89-93, 2012.

Key Words: Transportation project evaluation, sensitivity testing, uncertainty analysis, travel behavior modeling, transportation planning

ABSTRACT

Budget constraints and competing opportunities demand thoughtful project evaluation before investment. Significant uncertainty surrounds travel choices, demographic futures, project costs, and model parameters. The impacts of this uncertainty are explored by conducting hundreds of sensitivity test runs across 28 random parameter sets to evaluate highway capacity expansion and tolling project scenarios in Austin, Texas. The effects of different parameter sets on project benefit-cost ratios, crash counts, emissions, traffic volumes, and tolling revenues are examined in detail. Linear regression results show that link capacity, link-performance parameters – and their covariation – are key to results, followed by the elasticity of demand, trip growth rates and values of travel time.

INTRODUCTION

As a consequence of global recession, governments around the world are trimming budgets (Economist 2011a, 2011b and 2011c). With U.S. gas taxes stagnant and transportation construction prices rising 59% between 2000 and 2009 among six representative states (WSDOT 2011), transportation professionals must determine and pursue the most socially beneficial and budget-sensitive projects possible, under tight funding constraints. Kockelman et al.'s (2010) Project Evaluation Toolkit or PET allows users to quickly and with minimal input ascertain trip tables for abstracted networks (Xie et al. 2010), anticipate demand shifts under different network scenarios, and generate a host of project-evaluation metrics for side-by-side comparison. PET anticipates emissions, crashes, traveler welfare, and network reliability impacts, relative to Base-Case network conditions, and relies on user specification of project costs to estimate long-term performance metrics (like internal rates of return, benefit-cost ratios, and net present values).

PET also enables sensitivity testing of project impacts, by allowing users to randomize 28 sets of parameters (including values of travel time, link performance parameters, demand elasticities, and regional growth rates, among others). Sensitivity testing allows basic project assumptions to exhibit a degree of uncertainty and vary over the course of multiple trial runs, producing a distribution for possible outcomes and giving analysts a sense of risks and rewards across project alternatives.

An appreciation of the likely distributions of project outcomes is essential to wise decision making, since actual outcomes can be much different than those expected. Standard & Poor's Bain and Plantagie (2004) describe much of this problem, noting that estimates for some project types have not only been inaccurate, but biased overall. In this light, transportation planners, policymakers and investors may opt for a project with more certain benefits, rather than one with slightly greater predicted benefits, but also a greater degree of uncertainty (with some outcomes which may be particularly bad). Additionally, agencies may wish to package projects with a significant degree of risk as a public-private partnership (Bain and Plantagie 2007).

This paper examines the potential impacts of upgrading an arterial street to a freeway or a tollway in Austin, Texas. The nature of performance distributions is examined in greater detail by varying 28 sets of parameters – one set at a time and in combination – in order to better understand the impacts of multiple degrees of uncertainty across project scenario alternatives. Simulation results suggest how much uncertainty exists in model predictions, with regression results identifying key assumptions and inputs.

BACKGROUND

In 2003 Flyvbjerg et al. published an important cost-overrun study of 258 major public works across 20 nations, emphasizing road projects (167 of the 258 cases) with the rest comprising rail, bridges and tunnels. Two years later, a follow-up study (Flyvbjerg et al. 2005) focused on travel demand forecasts, for 210 major rail and roadway projects. Both studies concluded that cost and traffic estimates are highly uncertain, and regularly much different from actual values. For example, road project costs averaged 20.4% higher than projected costs – with a standard deviation of 29.9%, and half of all road projects had overstated demand by more than 20 percent, with a quarter of estimates overstating demand by at least 40 percent. Rail-related biases were even more dramatic, with 72% of all projects overstating ridership by 70 percent or more. Such results highlight a substantial underlying degree of uncertainty in forecasting traffic flows (as well as bias).

In their review of this literature, Lemp and Kockelman (2009b) noted that predicted traffic volumes exceed actual volumes by over 30% in half of the hundreds of cases examined. Even when correcting for optimism bias, uncertainty of traffic volumes and revenues remains substantial, suggesting that analyst assumptions are far from perfect. To address at least the variance in potential project outcomes, Lemp and Kockelman recommended that project evaluations be conducted using Monte Carlo or related simulations to "provide probability distributions of future traffic and revenue." (2009b, p 1). This is consistent with the practice of many others, including Ševčíková et al.'s (2007) model projecting future households by traffic analysis zone, Gregor's (2009) GreenSTEP emissions model which uses Monte Carlo sampling to generate distributions and Wang's (2008) application estimating uncertainty impacts using a freight mode choice model. While such sampling addresses the variability underlying key sources of project uncertainty, it does not address issues of model misspecification and bias.

The methods employed here use processes similar to those applied by Zhao and Kockelman (2002), Pradhan and Kockelman (2002) and Krishnamurthy and Kockelman (2003) in their investigations of uncertainty propagation through a standard four-step travel demand model and integrated transport-land use models. For example, Krishnamurthy and Kockelman varied 95 parameters and two demographic inputs over 200 simulations. After generating model predictions, they identified key inputs by regressing important outputs on the sets of variable inputs. Results were most strongly impacted by changes in the link performance function parameters, and shares of peak versus off-peak traffic (Krishnamurthy and Kockelman [2003]). This investigation builds on these three previous works by looking at the value of specific projects, rather than how urban-system model parameters affect total system travel distances and land use patterns. PET produces a wide variety of parameter-dependent project-related impacts, including emissions, crashes, traveler welfare, tolling revenues, and benefit-cost ratios.

TOOLKIT DESCRIPTION

The Project Evaluation Toolkit (PET) acts as a stand-alone tool for transportation project impact assessment. PET is intended for use upstream of the NEPA process, allowing planners to quickly evaluate a number of potential project variations before selecting the most appealing candidate(s) for more detailed demand modeling (and more detailed networks). PET uses an Excel interface with a C++-coded travel demand model that accommodates hundreds of network links and relies on user-entered link volumes and attributes, plus hundreds of parameter values to predict changes in travel patterns, emissions, crash counts, and other impacts (versus a base [e.g., no-build] case).

PET simultaneously serves several needs not currently met by any other single model. For example, its data requirements and run times (less than an hour to evaluate three scenarios, versus the Base-Case) are less cumbersome than required by regional planning models. PET allows up to five traveler classes for assignment and infers trip tables from link counts, closely mimicking traffic shifts on a complete network following network changes (Xie et al. 2010). PET holistically evaluates full-network impacts, unlike other sketch-planning tools that lack embedded travel demand models and focus on corridors. However, PET faces network-size limitations (with run times growing exponentially with network links) and neglects land use details (typically used for trip generation and attraction computations).

PET's travel demand model operates by first estimating a base trip table from coded-link volumes using a maximum entropy methodology (Xie et al. 2010). Next, PET performs an iterative process to equilibrate travel times, costs, and flows, beginning with the application of elastic demand functions governing all origin-destination (O-D) pairs:

$$x_{ij,d}^{k} = x_{ij,d}^{b,k} \left(\frac{g_{ij,d}^{k}}{g_{ij,d}^{b,k}} \right)^{\eta} \tag{1}$$

where $x_{ij,d}^{k}$ is the traffic volume of traveler class *k* (in vehicles per time period) between origin *i* and destination *j* during time-of-day period *d*, $g_{ij,d}^{k}$ represents the generalized cost (linearly combining time and money) of class *k* individuals traveling between origin *i* and destination *j* during time-of-day period *d*, and the superscript *b* denotes the Base-Case scenario traffic volumes or path costs. The η term represents the elasticity of trip-making and is set to -0.69 based on weighted elasticities observed from application of regional travel demand models to the complete Austin network (Lemp and Kockelman 2009a). This function estimates changes in trip making for each user class, based on general travel cost changes between each O-D pair.

After application of the elastic demand function, an incremental logit model (Ben-Akiva and Lerman 1985) is used to estimate changes in travel mode (e.g., SOV, HOV2, HOV3 or transit). For the heavy-truck driver class, the probability of choosing the heavy-truck mode is 1.0 so the mode split step is effectively ignored for these users. For other traveler classes, mode-split probabilities depend on user type (work-related [non-commute] travel, commuters and travelers with other non-work purposes), with each user type possessing distinct values of travel time and reliability. Their mode splits take the form:

$$P_{ij,m}^{k} = \frac{P_{ij,m}^{b,k} e^{-\lambda_m \Delta g_{ij,m}^{k}}}{\sum_m P_{ij,m}^{b,k} e^{-\lambda_m \Delta g_{ij,m}^{k}}} \tag{2}$$

In this model, $P_{ij,m}^{k}$ represents the probability that a traveler of type *k* originating at origin *i* and traveling to destination *j* will choose mode *m*; and $\Delta g_{ij,m}^{k}$ represents changes in generalized travel costs (as defined earlier). The mode-choice model requires a single mode scale parameter (λ_m) to reflect the generalized cost term's coefficient in the associated systematic utility function (Ben-Akiva and Lerman 1985).

The model then estimates changes in up to five time-of-day splits using a similar incremental logit model (with an associated time of day scale parameter, λ_t). Finally, the demand model relies on the Floyd-Warshall algorithm for shortest-path user equilibrium traffic assignment (Floyd 1962). This four-step iterative process (of elastic demand, mode and time-of-day choice, and network assignment, across multiple traveler classes) continues until equilibrium is reached using the method of successive averages. For more travel demand modeling details, see Kockelman et al. (2010).

Once the demand model reaches convergence, traveler welfare impacts (consisting of changes in monetized travel times and operating costs plus any surplus from new travelers) are estimated for each O-D pair (*ij*), traveler class (*k*), and time of day (*d*) using the rule-of-half (RoH) (Geurs et al. 2010):

$$\Delta CS_{ij,d}^{k} \cong \frac{1}{2}\left(w_{ij}^{b,k}x_{ij,d}^{b,k} + w_{ij}^{k}x_{ij,d}^{k}\right)\left(g_{ij,d}^{b,k} - g_{ij,d}^{k}\right) + w_{ij}^{b,k}x_{ij,d}^{b,k}\left(g_{ij,d}^{b,k} - g_{ij,d}^{k}\right) \qquad (3)$$

where x's represent each O-D pair's flow rate (before and after the network or policy change: x^b and x), g is generalized travel cost, and w is vehicle occupancy rate. This formulation accounts for benefits to new travelers who may be adding new trips due to reduced travel costs, as well as benefits to travelers who were already traveling from a given origin to a given destination, and see their travel costs fall. Preliminary testing was conducted based on Lemp and Kockelman's (2009a) demand model specifications to find that the RoH results very closely track (<5%) nested and standard logsums, provided that no major network changes are made or new alternatives are added (such as a subway mode).

All flows and welfare estimates are then imported to the Excel component, which estimates changes in network reliability, crashes, emissions, toll revenues, and fuel consumption. Reliability is estimated as a link-level travel time variance, using the following formula:

$$r_a = r_a^0\left(1 + \sigma\left(\gamma + \frac{v_a}{c_a}\right)^{\tau}\right) \qquad (4)$$

where r_a^0 is the free-flow travel time variance of link a, and σ, γ and τ are parameters estimated using traffic data obtained from freeway segments in Atlanta, Los Angeles, Seattle and Minneapolis (Margiotta, 2009). Using nonlinear least-squares regression, parameters were estimated to define the relationship between travel time variance and hourly volume-capacity ratios, with resulting values of $\sigma = 2.3$, $\gamma = 0.7$, and $\tau = 8.4$ (with an R^2_{adj} of 0.408) (Kockelman et al. 2010).

PET uses safety performance functions from Bonneson and Pratt's (2009) Roadway Safety Handbook to predict the total number of fatal plus injurious crashes on each directional link in the PET networks. Fatal and injurious count *shares* or splits, along with extrapolations of property damage only (PDO) crash counts, are then estimated from Texas crash data sets, (TxDOT 2009). Emissions estimates employ lookup tables generated using EPA's MOBILE 6.2, for 13 distinct species based on vehicle-fleet age and type distributions, ambient temperatures (summer and winter), model-estimated speeds, analysis year, and road facility type (freeway, arterial, ramp, etc.). Local calibration factors may be used to scale up or down crash counts and emissions volumes (due to local area crash histories, atmospheric variations, vehicle technologies, and so forth). For example, a 1.1 local crash calibration factor indicates 10% more crashes are expected than using default formulae. Summary measures are provided in the form of benefit-cost ratios, net present values, internal rates of return and payback periods (for each alternative policy or project, versus the Base-Case [no-build scenario]). PET's sensitivity testing module provides distributions on these, and many other model outputs, as illustrated in this work.

CASE STUDY

For this investigation, two scenarios were examined converting a four-lane arterial to a four-lane freeway or tollway. A 5% discount rate was assumed which is lower than the 7% required by the

OMB for federal projects, but is on the high end of the 3 to 5% discount rates typically used for state transportation projects (FHWA 2007). Additionally, a 20-year design life was assumed along with a 1% annual growth rate in Base-Case trip rates between all O-D pairs, though PET has the ability to account for pair-specific growth rates. The 1% growth rate is lower than the estimated regional population growth (Robinson 2008) but close to or higher than the expected growth rate for zip codes in which the most congested roadways lie. Figure 1 illustrates the case study location on the 194 link Austin regional network.

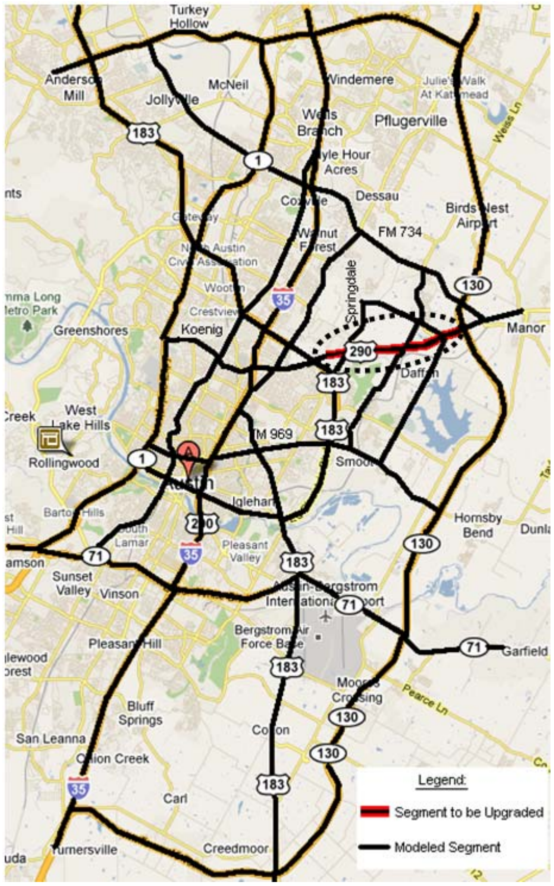

Figure 1: Case Study Location

Both scenarios included upgrading the existing four-lane segment (two through lanes in each direction) from an arterial to a grade-separated freeway or tollway, while retaining the two lanes in each direction configuration. Two-way (total) capacity was estimated as increasing from 3080 vehicles per hour (vph) to 7640 vph, as well as eliminating seven intersections between US 290 and minor streets. The first scenario (Freeway) was modeled as a non-tolled freeway, and the second scenario (Tollway) with fares at $0.20 per mile for SOVs (similar to Austin's US 130 [TxDOT 2011]), $0.10 per mile for HOVs (2 or more persons), no toll for transit users, and $0.60 per mile for heavy-trucks.

Initial project costs were estimated at $71.8 million for the Freeway scenario and $80.5 million for the Tollway scenario, based on an estimated construction cost of $3.2 million per freeway lane-mile plus another $760,000 per directional mile for installation of toll collection infrastructure and 10 percent for design costs, as per recent Texas projects (TxDOT 2008). A $30 million repaving project was also assumed needed 10 years after the initial-year in the Base-Case scenario. Annual maintenance and operations costs were estimated at $410,000 for the Freeway scenario plus another $1.13 million for the two tolling scenarios, based on recent Texas estimates (TxDOT 2008). Fagnant et al. (2011a) previously examined a similar case study, and changes in PET's specification have resulted in somewhat different B/C ratios and other outputs.

Due to rising input prices and the nature of this expansion project, the true tollway construction cost may be closer to $7 million per lane-mile (as confirmed by Austin tollway expert Burford [2011]). A project recently was bid in the same location with a larger footprint (6 managed lanes + 6 frontage lanes) with a per-mile project bid cost indicating this new estimate (though costs could be lower than $7 million per lane-mile due to the lack of new right-of-way acquisition). This variation of the base assumptions results in an approximate doubling of project costs and roughly a halving of these benefit-cost (B/C) ratios, as estimated below (parenthetically).

Both alternative scenarios showed favorable B/C ratios. The Freeway scenario was most favorable from the public's perspective, with a 14:1 B/C ratio, while the Tollway enjoys a respectable B/C ratio of 6.5:1. (These ratios are 7.7:1 and 3.5:1, respectively, under the higher construction cost assumption, of $7 million per lane-mile.) The main reason for the Freeway alternative's strong performance lies in its superior traveler welfare impacts, as shown in Table 1:

Table 1: Present Value of Capacity Expansion Scenario Impacts (in $Millions)

		Freeway	Tollway
Initial-Year	Total Impacts	**$32.0 M**	**$12.4 M**
	Traveler Welfare	$23.8	$5.0
	Reliability	$7.0	$6.3
	Crashes	$0.7	$0.7
	Emissions	$0.5	$0.4
Design-Year	Total Impacts	**$132.9 M**	**$99.6 M**
	Traveler Welfare	$76.6	$49.1
	Reliability	$51.8	$47.2
	Crashes	$1.4	$1.3
	Emissions	$3.1	$2.0

In the Freeway scenario, travelers gain a mobility benefit without having to pay an extra fee, as in the tolling scenarios. However, this carries an implicit cost since the Freeway scenario must be financed through tax revenues. Conversely, the Tollway scenario is not only self-financing, but likely revenue generating with an estimated 23% internal rate of return (or 11% under the higher-construction-cost assumption) - to tolling authorities, rather than to society at large.

While these estimates are high, and the projects may seem unusually attractive (from an engineering accounting standpoint), Austin tollway expert Burford (2011) feels that PET's revenue projections appear reasonable. Transportation planners and policy makers may prefer the Tollway scenario, since it offers a mechanism to quickly recover invested funds. This project may be much more profitable than existing tollways in Austin, due to its lack of parallel (non-tolled) frontage roads and assumption of no additional right-of-way requirements (which are likely required, due to state laws that mandate provision of a "free" alternative to new tolled routes).

Crashes and emissions also require further consideration, though outweighed by traveler welfare and travel time reliability benefits when monetized. Over the 20-year design life, the projects are projected to avoid 480-530 injurious crashes, 6 or 7 of which are expected to be fatal. Most emissions types are forecasted to fall in the initial-year and all are lower in the design-year. In particular network-wide emissions of hydrocarbons, butadiene, formaldehyde and acrolein all fall by over 0.9% when comparing the Freeway scenario's design-year with the Base-Case scenario. This is particularly impressive when considering that the improved links handle only 1.45% of total system traffic. For both crashes and emissions, the Freeway scenario is preferred, though the Tollway is still beneficial. The major reason for this is that some vehicles in the Tollway scenario chose longer and slower routes along arterials, thus increasing emissions and crash risks.

Average daily speeds on the upgraded segment increased in both scenarios relative to the Base-Case scenario, showing a 23 mph (55 vs. 32 mph) difference in the initial-year and 31 mph (54 vs. 23 mph) by the design-year. US 290 Traffic volumes are predicted to increase in both scenarios versus the Base-Case, with 160 and 275 vpd in the first year growing to 930 and 1100 by the design-year for the Tollway and Freeway scenarios, respectively.

PARAMETER VARIATION

Twenty-eight parameter sets were then varied during sensitivity analysis in order to determine the impact of parameter variation on outcomes. All random draws originate from lognormal distributions, where the corresponding/underlying normal random variable's standard deviation varies, as per user specification, and is centered at zero. These draws result in lognormal distributions with means centered approximately at 1, with reported coefficients of variation (CoV), where CoV equals the distribution's standard deviation divided by the absolute value of its mean. Variations were conducted by drawing an independent random value for each parameter set. (For example, all user classes' values of travel time have a single, shared draw for a given iteration, so all move up or down together, to help ensure some necessary correlation.) This random draw was then applied to the Base Case and all alternative scenarios for that iteration, and to the initial and design-life years (with interpolation of project impacts in intermediate years, to moderate computational burdens). Hundreds of iterations were run, for hundreds of evaluations across all scenarios (each versus the corresponding Base Case scenario).

Two sets of runs were conducted with three hundred iterations each, the first run containing a low degree of uncertainty (0.10 or 10.0% CoV for all parameter sets) and the second a higher degree of uncertainty for most parameters (10.0% CoV for three parameter sets with a relatively strong degree of certainty, 30.7% CoV for most parameters, and 53.3% CoV for four parameter sets with a high degree of uncertainty). These lognormal CoVs correspond to draws from the underlying (normal) random variables centered at 0 with standard deviations of 0.1, 0.3 and 0.5.

Table 2 shows which parameters were varied and the CoV for each set of runs, as well as the default average parameter values. For a full listing of these and other default-parameter value sources, please see the PET Guidebook (Fagnant et al. 2011b).

Table 2: Sensitivity Testing Parameters and Assumed Variations

Parameter	Low CoV	High CoV	Mean Values Used
Value of Travel Time	10.0%	30.7%	\$5 to \$50 per hour
Value of Reliability	10.0%	30.7%	\$5 to \$50 per hour of travel-time std. dev.
Vehicle Operating Costs	10.0%	30.7%	\$0.20 to \$0.50 per mile
Crash Costs	10.0%	30.7%	\$7500 (PDO) to \$1.13M (Fatal)
Emissions Values	10.0%	53.3%	For 5 species, varies widely
Link Capacities	10.0%	10.0%	Varies based on indiv. hwy link
Link Performance Params. (α & β) for BPR Formula	10.0%	10.0%	Varies based on link class
Free-flow Speeds	10.0%	10.0%	Varies based on link class
Reliability Parameters (σ & τ)	10.0%	53.3%	2.3, 8.4
Local Crash Rate Calibration Factor	10.0%	30.7%	1.0
Emissions Rate Calibration Factor	10.0%	30.7%	1.0
Mode Scale Parameter	10.0%	53.3%	1.0
Time-of-day Scale Parameter	10.0%	53.3%	0.1
Ambient Temperatures	10.0%	30.7%	76 (April-Oct), 56 (Nov-March) degrees Fahrenheit
Average Vehicle Occupancies	10.0%	30.7%	Averages 1.6 across all modes
User Class Share: Heavy-Truck Driver (very high VOT)	10.0%	30.7%	5%
User Class Share: Work Related (high VOT)	10.0%	30.7%	10%
User Class Share: Commuter (high VOT)	10.0%	30.7%	20%
User Class Share: Non-Work Related (low VOT)	10.0%	30.7%	65%
Mode Probability: SOV	10.0%	30.7%	35.9%
Mode Probability: HOV2	10.0%	30.7%	33.3%
Mode Probability: HOV3	10.0%	30.7%	29.6%
Mode Probability: Transit	10.0%	30.7%	0.12%
Annual Trip Growth Rates (over time)	10.0%	30.7%	1% Annually
Demand Elasticity (for O-D pairs)	10.0%	30.7%	-0.69
Initial Project Costs	10.0%	30.7%	\$71.8M - \$80.5M
Maint. & Operat. Costs	10.0%	30.7%	\$409,000 - \$1.13M

Note: User class shares and mode split shares must sum to one, so sets of drawn values were normalized (after mean-one draws were multiplied by base shares, and heavy-truck shares were removed from consideration).

ANALYSIS OF SENSITIVITY TEST RESULTS

Since benefit-cost (B/C) ratios drive many projects decisions, this output was examined first. B/C variations were dramatic, suggesting that input uncertainty can easily make or break a project. 15% of the 600 runs had B/C ratios below -100, and 17% had B/C ratios greater than 100 in both scenarios. However, B/C output distribution of was very similar for both sets of sensitivity test parameters (i.e., both high and low CoV values), likely due to the invariance of CoV (held at 10 percent) in the BPR link-performance parameters (α, β, and link capacities, c). These sets of parameters are found to be key to impact assessment, since they regulate the estimated traffic speed on each traveled link (s_a) via the Bureau of Public Roads link performance function (TRB 2000), as follows:

$$s_a = \frac{s_a^0}{1+\alpha\left(\frac{v_a}{c_a}\right)^{\beta}} \qquad (5)$$

where s_a^0 is the link's free-flow speed (obtained from Cambridge Systematics [2008]), v_a/c_a is the link's volume-capacity ratio, and α and β are behavioral parameters.

Given their similar results, the low and high variation (CoV) sets of runs were combined for further evaluation. A histogram of the combined B/C ratios shows a very wide distribution of values, with a compact center, as shown in Figure 2.

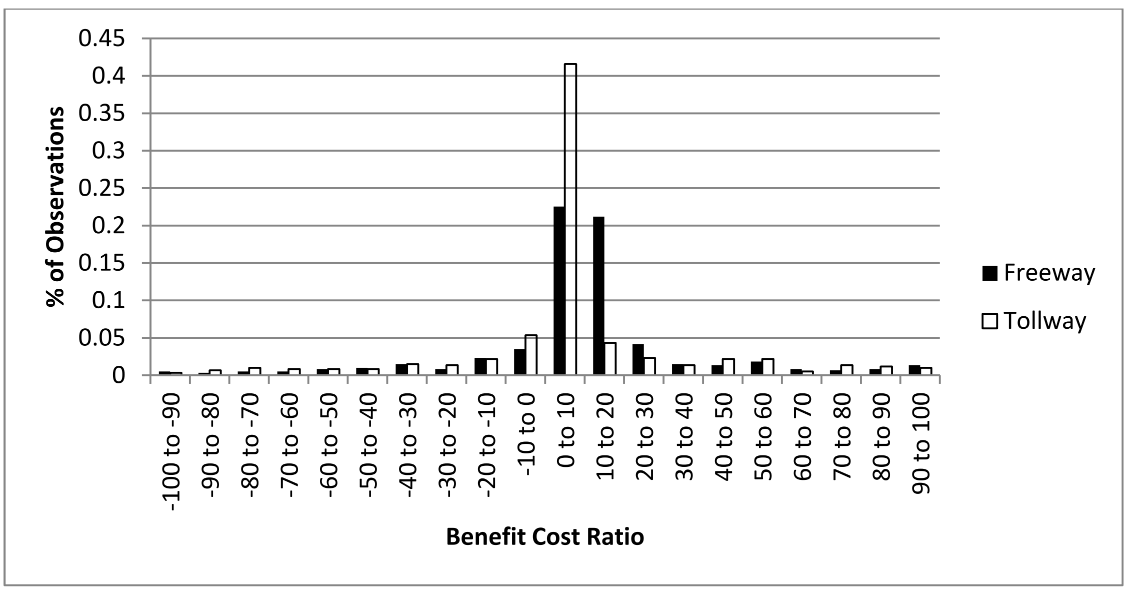

Figure 2: B/C Ratios (with values beyond +/- 100 not shown)

Shares of B/C ratios were similar across both scenarios, with 54% (Freeway) and 56% (Tollway) of outcomes falling in the -20 to +30 band of reasonable B/C ratios, 21% (Freeway) and 19% (Tollway) falling below -20 and 25% (both scenarios) lying above +30. In other words, there was much more variation in performance measures across test runs than across project alternatives. Nevertheless, important differences across project alternatives can be observed near the median values. The median B/C value for the Freeway scenario was 10.3, compared to 4.0

for the Tollway scenario, as apparent in Figure 2's distribution spikes. In fact, in 63 percent of alternative comparisons, the Freeway scenario bested its competitors (Base-Case and Tollway scenarios) while the Tollway scenario was preferred in just 22 percent of trials. In the remaining instances, both alternatives showed B/C ratios less than 1.0, indicating a Base-Case (no-build) preference. This shows how complex transportation networks can have unpredictable consequences (similar to Braess' Paradox), and how improving travel for some travelers (even at zero cost) may negatively impact others, particularly when modeling elastic demand under congested conditions.

Also of note, the *median* B/C ratios across both alternative scenarios were less than the B/C ratios estimated at mean parameter values. When PET was run without parameter variation, the scenarios yielded favorable B/C ratios of 14.1 and 6.2 for the Freeway and Tollway scenarios, respectively. In both instances, the 14.1 and 6.2 values fell around the 62nd percentile of the sensitivity-test outcomes, suggesting that false certainty in model parameter values can mask potential project downsides.

One clear factor in extreme B/C cases is a dramatic increase (or decrease) in total VMT versus the Base-Case scenario. In instances with B/C ratios lower than -100, VMT averaged a 24% design-year decrease vs. the Base-Case scenario, compared to a 51% VMT increase in instances where the V/C ratio was greater than 100 and an average VMT decrease of 1.7% for all other instances. Initial-year comparisons show similar patterns, though to a much smaller degree (1.4% average decrease vs. 4.8% average increase). Alternative scenarios' design-year VMTs grew in almost all sampled runs, though sometimes at a lower rate than the corresponding Base-Case scenario. Large VMT changes also coincided with dramatic changes in traveler welfare and reliability. More VMT ties to higher welfare estimates for induced travelers (thanks to the Rule of Half), but can congest roadways, resulting in negative reliability impacts and often resulting in negative welfare impacts for existing travelers. Therefore, depending on the specific nature of the VMT increase, it can quickly lead to either much higher or lower overall welfare values.

Since each scenario is distinct (e.g., some are weak proposals and others strong), there is no guarantee inputs will impact outputs similarly across scenarios. Therefore, regression analyses were conducted separately for each scenario (using stepwise deletion and addition of input values as covariates, with a p-value cutoff of 0.05). The best fits were found using the natural log of the absolute value of the simulated B/C ratios. Other specifications were investigated, using B/C ratios directly or attaching a sign to their logarithm (to reflect the original ratio's sign), but these performed poorly (with R^2 values less than 0.11). This is likely due to extreme B/C values or outliers (causing non-linear impacts for extreme outcomes) and common factors that contributed to both positive and negative outliers. These regression results are shown for B/C ratios in Table 3.

One important limitation of using $Y = \ln(|B/C|)$ is that it fails to predict whether a particular, random setting will result in a win ($B/C > 1$) or a loss ($B/C < 1$). In the presence of extreme (and unlikely) outcomes, it remains important to determine which input factors influence the direction and sign of project impacts, in addition to magnitude. Beyond B/C values, crash counts, emissions estimates, link volumes, toll revenues and other PET outputs exhibited similar outlying values, with most outliers emerging in the design-year (rather than in the initial-year, which is unaffected by the trip growth rate factor). To address the issue of outcome sign, standard linear regressions were performed (using untransformed outputs – e.g., Y = B/C) on the middle 50% of initial-year values (in the B/C case) and middle 90% of initial-year outcomes (for other outputs), by simply discarding the top and bottom 25 or 5% of points, in order to eliminate the disproportionate impact of outliers. Such results are also shown, for the B/C values, in the final columns of Table 3, and in Tables 4 and 5 for other impacts.

Table 3: B/C Ratios Regression Model Estimates for Freeway and Tollway Scenarios

	y = ln(|B/C Ratio|)		y = B/C (50 % truncated sample)	
	Freeway	Tollway	Freeway	Tollway
Constant	2.408	1.879	38.522	20.889
Value of Travel Time	**2.881**	**2.552**	***7.390***	***8.408***
Value of Reliability	0.665			
Vehicle Operating Costs	-1.436	-0.938	-5.737	
Emissions Values			1.904	1.294
Link Capacities	**-14.946**	**-17.950**	***-31.156***	***-39.275***
Link Performance Params.	**7.914**	**10.130**	***10.508***	***18.561***
Free Flow Speeds				-7.278
Reliability Parameters			3.749	1.620
Emissions Rate Calibration Factor	-0.596			
Time of Day Scale Parameter	0.435	0.886	-1.599	
User Class Share: Heavy-Truck Driver (High VOT)	**2.346**	**2.089**	**5.198**	
User Class Share: Work Related (High VOT)				2.487
User Class Share: Non-Work Related (Low VOT)	-1.311	-1.020	-3.739	
Mode Probability: SOV		-0.576		
Mode Probability: HOV2	0.761			
Mode Probability: Transit			-3.338	-3.169
Annual Trip Growth Rate	**3.207**	**3.853**		5.994
Demand Elasticity	**2.741**	**3.137**		
Initial Project Costs	-1.306	-1.027	***-8.351***	***-1.752***
N_{obs}	600	600	300	300
R^2	0.655	0.732	0.403	0.438
R^2_{Adj}	0.647	0.727	0.380	0.419

Several significant findings emerge from Table 3's parameter estimates. First, the signs on estimated parameters are the same using transformed and untransformed B/C values, in the two datasets (n=600 vs. n=300). Similarly, the most important factors in the first model are also key in the second. The results suggest that, while networks that congest more quickly (due to link-performance parameter value shifts), lower initial costs, and higher values of travel time, trip growth and demand elasticity tend to produce more extreme B/C values, most lead to positive B/C results, on average.

Such results are mostly intuitive, and encouraging. In less extreme input-set cases, α and β increases and constraints on system capacity appear to benefit travelers greatly. Capacity reductions make travel speeds more responsive to demand levels, thus enhancing the value of the two scenarios' capacity increases. The importance of these parameters is consistent with Krishnamurthy and Kockelman's (2003) propagation of uncertainty tests (in land use-transportation model applications for Austin). Additionally, when the outcome results in a high negative cost, it makes sense that further constriction of system capacity and increases in α and β can make a bad situation worse. In the most extreme cases, low capacity and high α and β values resulted in instances where system VMT was nearly 8 times larger or smaller in the two expanded-capacity scenarios than in the same-iteration Base-Case scenario, thereby generating the unlikely results.

As noted earlier, the importance of these parameters also explains why the B/C distributions for the high- and low-variation (Table 2) sets of runs were so similar. Capacity values and the other two link performance parameters (α and β) were modeled with a single 10% CoV in both sets of sensitivity testing runs. While other parameters were allowed to vary more (in the high-variation runs), capacity and link performance parameters remain the driving force behind B/C outcomes. They clearly dominate results, suggesting that link-performance assumptions deserve careful generation and treatment.

Though their parameter values are not quite as large, sizable increases in VOTTs and the share of heavy-trucks (which effectively diminishes link capacities) also improved B/C ratios (Table 3). Interestingly, the values of traffic growth and demand elasticity appear to have greater impact on the size – rather than sign – of the B/C outcomes. All parameters with Table 3 coefficients exceeding the project-cost coefficient are practically most important. Initial project costs comprise 90% or more of these two scenario's project lifecycle costs and so serve as a useful reference point: essentially, a doubling of initial costs should reduce the B/C ratio's magnitude by about 50 percent.

The impact of parameter variation on other key impacts was also evaluated. These assessed impacts included the impact of variation changes on crashes, emissions, traffic volumes on the impacted segment and system-wide tolling revenues. Even with a benefit-cost ratio in hand, each of these key measures is likely still independently important to decision makers attempting to discern which alternative scenario to fund, if either. Crashes in this evaluation were monetized, using crash valuations as noted by Blincoe et al. (2002) inflated to current (2010) values. However, non-economic "soft" crash components (such as the value of life and pain and suffering) were not monetized and should therefore be independently evaluated. Changes five emissions species (Hydrocarbons, Nitrous Oxide, Carbon Monoxide, Particulate Matter < 2.5 μm and Particulate Matter < 10 μm) were also monetized using EU data (Mailbach et al. 2008). These emissions values may be important to cities seeking to meet air quality attainment goals and the "monetary emissions benefits" output provides a framework for users measure broad impacts across all five monetized species. Table 4 details the initial-year regression outputs for number of crashes and emissions costs:

Table 4: Impacts on Initial-Year Crashes & Emissions Costs (mid 90%)

	Initial-Year Crashes (Fatal & Injurious)		Initial-Year Emissions ($M)	
	Freeway	Tollway	Freeway	Tollway
Constant	-5.63	-7.23	-4.99	-3.93
Value of Travel Time	-1.94	-3.63		-0.25
Vehicle Operating Costs			0.20	
Emissions Values			-0.40	-0.39
Link Capacities	***10.97***	***11.92***	***4.90***	***4.59***
Link Performance Params.	***-4.94***	***-4.83***	***-1.30***	***-1.15***
Free Flow Speeds			*1.35*	*1.11*
Local Crash Calibration Factor	***-16.86***	***-15.11***		
Emissions Rate Calibration Factor			-0.54	-0.39
User Class Share: Heavy-Truck Driver (High VOT)		-2.01	-0.23	-0.24
User Class Share: Non-Work Related (Low VOT)	1.82	1.74		
Mode Probability: HOV 3+		1.91	0.22	
N_{obs}	540	540	540	540
R^2	0.563	0.479	0.522	0.451
R^2_{Adj}	0.558	0.473	0.515	0.443

Several fundamental inferences may be made from Table 4. As with the B/C ratio results, capacity and the link performance parameters were influential in predicting crash and emissions changes vs. the Base-Case scenario. As link capacities increase, more people travel in the alternative scenarios than the Base-Case, resulting in more crashes and lower emissions savings (or higher costs). One major difference between the crash and emissions models, however, is that the local crash calibration factor is more influential in predicting crashes than the emissions rate calibration factor for predicting emissions costs. This indicates that estimated crash predictions are much more stable than emissions costs, since a 10% increase in either value should result in a 10% respective increase in crashes or emissions.

In addition to reviewing crash and emissions impacts, traffic volumes and tolling revenues were also analyzed. Transportation agencies spending money to improve a facility want to know how much it will be used. Furthermore, tolling revenues were not included in the overall benefit-cost ratio as collected tolls were assumed to be a benefit-neutral transfer payment from individual travelers to society (Kockelman et al. 2010). However, transportation agencies (or private enterprise) using PET will undoubtedly be interested in revenues as this may be the mechanism used to pay for the project. Additionally, PET's tolling output is structured such that system revenues are reported rather than just for the improved link. This is particularly key in this case study as the project could impact collected tolls on an adjacent priced facility.

As with crashes and emissions, linear regression models were run for initial-year upgraded segment traffic volumes in both scenarios, though only for the changes in tolling revenues for the Tollway scenario, as shown in Table 5:

Table 5: Impacts on Traffic Volume & Tolling Revenues (mid 90%)

	Initial-Year US 290 AADT		Initial-Year Tolls ($M)
	Freeway	Tollway	Tollway
Constant	1039	716	8.06
Value of Travel Time	72	***1118***	***1.43***
Vehicle Operating Costs	-51	***1638***	***2.03***
Link Capacities	***-1591***	***-2488***	***-1.44***
Link Performance Params.	***722***	***662***	
Free Flow Speeds		***-1396***	
Mode Scale Parameter	23	-195	0.56
Time of Day Scale Parameter			0.40
Average Vehicle Occupancy	56		0.60
User Class Share: Heavy-Truck Driver (High VOT)	84	**246**	***2.93***
User Class Share: Work Related (High VOT)	32		
User Class Share: Commuter (Mid VOTT)	-34		
User Class Share: Non-Work Related (Low VOT)	-62	***-505***	-0.71
Mode Probability: SOV		-253	0.76
Mode Probability: HOV 2	44		
Mode Probability: Transit		196	0.76
Demand Elasticity	-34		
N_{obs}	540	540	540
R^2	0.811	0.540	0.284
R^2_{Adj}	0.806	0.531	0.271

Again, capacity and the link performance parameters were found to be highly influential in estimating traffic volumes. As capacity increases, travelers have less incentive to use the improved link, as opposed to other routes. However, for the Tollway scenario, numerous other inputs had substantial impact. Values of travel time (VOTTs) and operating costs appear critical in determining traveler route choices and revenues. When either fall, travelers choose alternative non-tolled routes. Also, user shares substantially impacted revenues and traffic volumes on the Tollway. The Heavy-Truck Driver user class has the highest VOTT (and pays the highest fare), while the Non-Work Related user class has the lowest VOTT. Therefore, any increases in the proportion of the Heavy-Truck Driver user class or decreases in the Non-Work Related user class resulted in more travelers on US 290 and more tolling revenues.

Finally, a series of log-linear ordinary least squares regression estimates was conducted on the natural-log transformed (ln|y|) values of design-year crashes, emissions, traffic volumes and tolling revenues, as shown in Table 6. As with the B/C ratio estimates, the best fits were found using this transformation, though other specifications were investigated. Again, attaching a sign outside the transformation resulted in weak fit statistics, likely due to non-linear impacts for the more extreme outcomes and common factors that contributed to both positive and negative outliers.

Table 6: Estimating the Variation in Design-Year Impacts

	Crashes	Emissions	US 290 AADT	Tolls
	Freeway/Tollway	Freeway/Tollway	Freeway/Tollway	Tollway
Constant	0.84	13.24	6.33	16.209
Value of Travel Time	0.74	1.49	0.65	
Vehicle Operating Costs	-0.72	-1.20	-0.78 (Tollway)	
Link Capacities	***-7.34***	***-15.58***	***-9.33***	***-2.158***
Link Performance Params.	***4.95***	***9.82***	***5.20***	***0.821***
Local Crash Calibration Factor	0.84			
Emissions Rate Calibration Factor		1.03		
Mode Scale Parameter				0.156
Time of Day Scale Parameter		0.53	0.26 (Tollway)	
Ambient Temperatures		-0.55		
Average Vehicle Occupancy	0.43 (Tollway)			0.297
User Class Share: Heavy-Truck Driver (High VOT)	1.08 (Freeway) 0.65 (Tollway)	1.68	0.85	
User Class Share: Work Related (High VOT)			0.49	
User Class Share: Non-Work Related (Low VOT)	-0.45	-1.01	-0.62 (Freeway)	
Mode Probability: HOV 2	0.33	0.90	0.30	
Mode Probability: HOV 3+			0.35	
Annual Trip Growth Rate	***2.01***	***3.52***	***1.89***	***0.713***
Demand Elasticity	***2.36***	***2.82***	***1.44***	***0.86***
N_{obs}	1200	1200	1200	600
R^2	0.511	0.640	0.511	0.192
R^2_{Adj}	0.506	0.636	0.506	0.183

Note: Table shows OLS regression results of the natural log of crash counts, emissions tons, AADT, and tolls revenues on variable inputs.

As noted earlier, the purpose of this investigation was to determine which factors most influence the design-year output levels. Unsurprisingly, capacity and the link performance parameters dominated the outcomes' magnitude in all cases, with outputs rising as capacity becomes constrained. The next parameter sets exhibiting important impacts are consistent across all alternatives: annual trip growth rate and demand elasticity. In all four models, these two parameters had a greater impact than any other, excluding capacity and the link performance parameters. This makes sense, since the impacts of a trip growth rate will be compounded over time and demand elasticity will regulate the additional number of trips that occur as travel costs change, both crucial elements in the Freeway and Tollway scenarios.

CONCLUSION

This paper conducted a thorough investigation into the impacts of parameter uncertainty on highway project outcomes. Twenty-eight parameter variations and their effects on benefit-cost ratios, crashes, emissions, traffic volumes and tolling revenues were examined in detail. From this evaluation it quickly became clear that if analysts underestimate capacity or overestimate link performance parameters the benefit-cost ratio may quickly become unreasonable. Another crucial finding showed that the median B/C ratio was significantly lower than the B/C ratio when no variation was assumed. This is particularly important as these results would lead analysts to expect lower probable benefits from these scenarios than the no-variation case would suggest.

Even when omitting extreme outcome variations, capacity and the link performance parameters had greater impact on B/C ratios, crashes, emissions, traffic volumes, and tolling revenues than any other examined inputs. B/C ratios were strongly depended on VOTTs in both scenarios, though the value of time (and operating costs) impacted the actual use and collected revenues of the improved facility for the Tollway much more heavily than in the Freeway scenario. Crashes were found insensitive to congestion relative to emissions, though both were impacted. Finally, the magnitude of final design-year outcomes was strongly influenced by travel growth rate and demand elasticity parameters.

In summary, this paper illustrated potential applications of PET and provides detailed analysis conducted of its outputs using sensitivity testing and input variation. Transportation planners may employ similar methods, using the Toolkit to produce a range of likely outcomes, rather than a single point estimate. This paper details which parameter variations tend to cause the greatest variations in impacts under two scenarios, though ultimate results will depend on the nature of any future project under consideration. These methods and findings should enhance the ability of decision makers to allocate limited transportation funding resources while providing the most beneficial outcomes for society at large.

ACKNOWLEDGEMENTS

This paper was made possible through the diligent support of Duncan Stewart and TxDOT research projects 0-6235 and 0-6487. Special thanks to Annette Perrone for her keen editing and logistical assistance, and to TRB's anonymous reviewers for their suggestions. Additional thanks go to Chi Xie, who was instrumental in the formulation and coding of PET's travel demand model.

REFERENCES

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

Bain, Robert and J.W. Plantagie (2007) Infrastructure Finance: The Anatomy of Construction Risk: Lessons from a Millennium of PPP Experience, Standard & Poor's, McGraw-Hill International (UK).

Ben-Akiva, M. and S. Lerman. (1985) Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, MA.

Blincoe, L.; Seay, A.; Zaloshnja, E; Miller, T.; Romano, E.; Luchter, S.; and Spicer, R. (2002) The Economic Impact of Motor Vehicle Crashes 2000. National Highway Traffic Safety Administration.

Bonneson, James A. and Michael P. Pratt (2009) Road Safety Design Workbook. Texas Transportation Institute. Prepared for the Texas Department of Transportation. College Station, Texas.

Burford, Wes (2011) Personal communication by phone. Director of Engineering, Central Texas Regional Mobility Authority. August 2.

Cambridge Systematics, Inc., Dowling Associates, Inc., System Metrics Group, Texas Transportation Institute (2008) Cost Effective Performance Measures for Travel Time Delay, Variation, and Reliability. NCHRP Report 618. Washington D.C.

Economist (2011a) A Budget for the Long Term. April 4. Online at: http://www.economist.com/blogs/freeexchange/2011/04/american_government_debt.

Economist (2011b) British Austerity and the Price of Black Swan Insurance. February 3. Online at: http://www.economist.com/blogs/freeexchange/2011/02/budget_cuts.

Economist (2011c) No Gold in the State. May 21. Online at: http://www.economist.com/node/13702838?story_id=E1_TPSDNRPR.

Fagnant, Daniel, Kara Kockelman and Chi Xie (2011a) Anticipating Roadway Expansion and Tolling Impacts: A Toolkit for Abstracted Networks. *Proceedings of the 90th Annual Meeting of the Transportation Research Board* and under review for publication in the *Journal of Transportation Engineering*.

Fagnant, Daniel, Kara Kockelman and Chi Xie (2011b) Project Evaluation Toolkit: A Sketch Planning Tool for Evaluating Highway Transportation Projects. Texas Department of Transportation, Center for Transportation Research, UT Austin.

Federal Highway Administration (2007) Asset Management – Evaluation and Economic Investment. Online at http://www.fhwa.dot.gov/infrastructure/asstmgmt/primer03.cfm.

Floyd, R. W. (1962) Algorithm 97: Shortest Path. Communications of the ACM 5(6), 345.

Flyvbjerg, Bent, Mette K. Skamris Holm, and Soren Buhl (2003) How Common and How Large are Cost Overruns in Transport Infrastructure Projects? *Transport Reviews*, 23(1), 71-78.

Flyvbjerg, Bent, Mette K. Skamris Holm, and Soren Buhl (2005) How (In)accurate are Demand Forecasts in Public Works Projects? *Journal of the American Planning Association*, 71(2), 131-146.

Gregor, B. (2009) GreenSTEP Model Documentation. Oregon Department of Transportation, Salem.

Geurs, Karst, Barry Zondag, Gerard de Jong and Michiel de Bok (2010) Accessibility Appraisal of Land-Use/Transport Policy Strategies: More than just Adding up Travel-Time Savings. *Transportation Research Part D: Transport and Environment*, 15 (7) 382-393.

Kockelman, Kara, Chi Xie, Dan Fagnant, Tammy Thompson, Elena McDonald-Buller and Travis Waller (2010) Comprehensive Evaluation of Transportation Projects: A Toolkit for Sketch Planning. Research Report 0-6235-1, Texas Department of Transportation, Center for Transportation Research, University of Texas at Austin.

Krishnamurthy, Sriram and Kara Kockelman (2003) Propagation of Uncertainty in Transportation Land-Use Models: An Investigation of DRAM-EMPAL and UTPP Predictions in Austin, Texas. *Transportation Research Record*, 1831, 219-229.

Lemp and Kockelman (2009a) Anticipating Welfare Impacts via Travel Demand Forecasting Models: Comparison of Aggregate and Activity-Based Approaches for the Austin, Texas Region. *Transportation Research Record*, 2133, 11-22.

Lemp and Kockelman (2009b) Understanding & Accommodating Risk & Uncertainty in Toll Road Projects: A Review of the Literature. *Transportation Research Record*, 2132, 106-112.

Mailbach, M. et al. (2008) Handbook on Estimation of External Costs in the Transport Sector, CE Delft (www.ce.nl).

Margiotta, J. (2009) Private communication by e-mail exchange. October 7, 2009.

Pradhan and Kockelman (2002) Uncertainty Propagation in an Integrated Land-Use Transport Modeling Framework: Output Variation via UrbanSim. *Transportation Research Record*, 1805, 128-135.

Robinson, Ryan (2008) City of Austin Population and Households Forecast by ZIP Code. City of Austin.

Ševčíková, Hana, Adrian Raftery and Paul Waddell (2007) Assessing Uncertainty in Urban Simulations using Bayesian Melding. *Transportation Research Part B: Methodology*, 41 (6) 652-659.

Texas Department of Transportation (2008) IH 35E Managed Lanes From IH 635 – To: US 380 Preliminary Financial Feasibility Study (PFFS) Final Draft.

Texas Department of Transportation (2009) Rural and Urban Crashes and Injuries by Severity 2008.

Texas Department of Transportation (2011) Texas Tollways – Austin Area Toll Roads. Online at http://www.texastollways.com/austintollroads/english/rates.htm.

Transportation Research Board (2000) Highway Capacity Manual 2000. Washington, D.C.

Wang, Min (2008) Uncertain Analysis of Inventory Theoretical Model for Freight Mode Choice. International Conference on Intelligent Computation Technology and Automation.

Washington State DOT (2011) Construction Cost Indices. Online at: http://www.wsdot.wa.gov/biz/construction/CostIndex/pdf/CostIndexData.pdf.

Xie, C., Kockelman, K. and Waller, S.T. (2010) Maximum Entropy Method for Subnetwork Origin-Destination Trip Matrix Estimation. *Transportation Research Record*, 2196, 111-119.

Zhao, Yong and Kara Kockelman (2002) Propagation of Uncertainty through Travel Demand Models. *Annals of Regional Science*, 36 (1) 145-163.