Stochastic Dominance Approach to Evaluate Optimism Bias in Truck Toll Forecasts

Sen Gupta, Rajorshi and Vadali, Sharada R

Texas Transportation Institute, Texas Transportation Institute

2007

Stochastic Dominance Approach to Evaluate Optimism Bias in Truck Toll Forecasts

Rajorshi Sen Gupta and Sharada R. Vadali

Optimism bias is a consistent feature associated with truck toll forecasts, à la Standard & Poor's and the NCHRP synthesis reports. Given the persistent problem, two major sources of this bias are explored. In particular, the ignorance of operating cost as a demand-side factor and lack of attention to user heterogeneity are found to contribute to this bias. To address it, stochastic dominance analysis is used to assess the risk associated with toll revenue forecasts. For a hypothetical corridor, it is shown that ignorance of operating cost savings can lead to upward bias in the threshold value of time distribution. Furthermore, dominance analysis demonstrates that there is greater risk associated with the revenue forecast when demand heterogeneity is factored in. The approach presented is general and can be applied to all toll forecasts and is not restricted to trucks.

Surface transportation involves complex interaction between key players including direct users, such as the trucking industry, as well as indirect user groups, including shippers and consumers. Feasibility analyses of proposed policies (say, construction of a toll road) therefore require a thorough understanding of how these different stakeholders view them. In this paper the trucking industry is being considered, given its importance to the overall supply chain. This research is motivated by a few observations.

First, in the event of inadequate resources to fund surface transportation, Texas and several other states are exploring alternative avenues. Options such as funding highways through direct user fees and particularly truck toll lanes are therefore gaining momentum.

Second, it is important to quantify the sources of optimism bias, which is a rather consistent feature of toll road forecasts, a la Standard & Poor's (*1–3*) and the NCHRP synthesis reports (*4*). One reason behind optimism bias is overestimated truck usage. This bias could result from a number of factors including, but not limited to, the divergence between actual (ex post) and perceived level of benefits, the differences in short- and long-term responses, industry constraints, routing strategies adopted by the companies, and nonrecognition of user heterogeneity. It is worthwhile to examine these reasons because the response of the trucking industry is a crucial factor behind feasibility analysis of proposed toll roads.

Texas Transportation Institute, Texas A&M University, Suite 402F, CE/TTI Tower, 199 Spence Street, College Station, TX 77843-3135. Corresponding author: S. R. Vadali, s-vadali@ttimail.tamu.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2066, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 98–105.
DOI: 10.3141/2066-11

Third, the behavioral responses to tolls need to be understood in light of the "critical issues" for the trucking industry as enumerated in *Top Industry Issues* (*5*). With the significant rise in different components of operating costs (such as fuel and insurance), the demand for toll roads ought to be reexamined to obtain reliable revenue estimates.

BACKGROUND AND RESEARCH OBJECTIVES

At the strategic level of planning, project evaluation is a crucial first step. Better decision making involves some key issues from the perspective of bondholders, practitioners, and researchers. For instance, in regard to a proposed toll facility one should have reliable information on expected toll revenue and costs over a specified planning horizon. At the initial stage of planning, however, not all of this information is available. That leads to a considerable amount of uncertainty in regard to decision making. Existing models of revenue and traffic forecast try to incorporate uncertainty through conservative assumptions on model inputs and varying them one at a time (*4*). From a practical standpoint, it is better to obtain probabilistic forecasts by including risk explicitly into the model. Obtaining simulated distributions is a methodological advancement over sensitivity analysis, whereby it is possible to consider numerous "what if" scenarios to examine the impact of uncertainty on the variables of interest. Although the superiority of simulation has been pointed out in the literature (*4*), actual applications of simulation with respect to truck toll studies have been rare. One exception is Kawamura (*6*), who applies stochastic simulation in calculating dollar benefits accrued by trucks from congestion pricing. However, even simulated distributions require some objective criterion to compare the risk associated with different alternatives. Stochastic dominance is one such approach, whereby simulated distributions can be ranked objectively in regard to the underlying risk.

Economic and financial feasibility analysis on toll lanes, as done in the literature, typically involves estimation of agency costs and user benefits in regard to travel time and vehicle operating costs savings [see, for instance, Veras et al. (*7*)]. Although analyses of these kinds are conducted at the aggregate level, it is worthwhile to investigate the issue of optimism bias by considering it as a demand-side problem because demand uncertainty is transmitted to both revenue and traffic forecasts (*4*). This, in turn, requires careful attention to user heterogeneity. For a highly fragmented industry of this kind, the predictive capability of any study would depend on recognizing the underlying demand characteristics and industry segmentation. It is therefore useful to control for the demand-side variation in regard to operational and financial characteristics (*8*). The existing literature does reflect on the importance of user heterogeneity. For instance, Yang et al. (*9*) examined the impact of user diversity on the profitability and social welfare aspects of new toll roads. Kenneth (*10*) and Kawamura (*6*) illustrated the importance of heterogeneity in the value of time (VOT) in the context of value and congestion pricing. Finally, Verhoef and Small (*11*) investigated the properties of various types of public and private pricing, with user heterogeneity represented by a continuum of VOT.

Although user heterogeneity is acknowledged as a crucial factor, its link to the risk associated with revenue yield is not well documented. It is important to understand and incorporate the demand uncertainty from various sources. The toll rate, travel time savings, and user heterogeneity (usage patterns and within-group variation) can each affect the demand for toll roads. Also, the importance of operating cost in relation to the toll costs as a demand-side factor has not received its due attention. This is particularly important when a tolled route significantly affects generalized travel costs in addition to time savings.

The principal objectives of this paper are thus threefold. First, the demand-side effect of operating cost on VOT is demonstrated. Second, the consequences of ignoring user heterogeneity (emanating from operating cost differentials and demand structure) on the risk associated with toll revenue forecasts are examined. Finally, a better method of quantifying the risk associated with toll revenue forecasts is presented. In doing so, stochastic dominance analysis is adopted as a tool to aid decision making at the strategic level.

LITERATURE ON TRUCK DIVERSION

What is the evidence on responsiveness of the trucking industry to changes in toll rates? In one of the earlier studies, McKnight et al. (*12*) estimated toll elasticities for Triborough Bridge and Tunnel Authority bridges and tunnels in New York and found elasticity of demand varying between [–0.61, –0.29] for medium and [–0.93, –0.27] for heavy trucks. One exception to this was the Verrazano Narrows Bridge, in which case heavy trucks exhibited complete price inelasticity as a result of the lack of "good substitute routes."

For the Central Texas Turnpike System, probabilities of trucks selecting the toll road were obtained by using market share–diversion estimates using time–cost trade-off (*13*). The highly sensitive nature of the trucking industry can be seen from the elasticities, which are within the range of [–0.698,–0.681], depending on the time of day considered. Likewise, the effect of toll rates on truck volumes is characterized by an elasticity of –0.593 for the Ohio Turnpike (*14*). [The elasticities have been computed from the traffic volumes reported in URS Corporation and Vollmer Associates, L.L.P. (*13*) and Taft (*14*).]

Veras et al. (*15*), however, found that in the context of New York, 70% of the carriers did not exhibit any change in behavior after the 2001 Port Authority of New York and New Jersey toll increase. "Customer requirements" is cited to be the underlying reason behind this phenomenon. It is also attributable to the existing tolling culture, whereby it is easier to pass toll increases on to the customers. This is however not applicable for states like Texas, in which in the absence of such a culture, contractual constraints are perceived to be more severe.

At this point the state of the practice methodology to model diversion will be reviewed. Taking utility maximization as the basic postulate, the propensity to adopt a tolled road is obtained as a function of time–cost differentials between a tolled route and an alternative free route. Typically estimated as a logit model, this leads to the so-called S-shaped diversion curves [for instance NCHRP Synthesis 364 (*4*), URS Corporation and Vollmer Associates (*13*), and Reebie Associates (*16*)]. Several issues concerning this approach deserve attention. This method is often applied to obtain demand parameters at the aggregate level. The Central Texas Turnpike System Report, for instance, obtains toll revenue forecasts by estimating a VOT parameter for the entire trucking industry (*13*). Although toll rates are varied by number of axles, heterogeneity in regard to willingness to pay remains unheeded despite evidence from the literature that responsiveness might vary across truck categories.

Another qualification that merits attention is the issue of bounded rationality (*17*), as a result of which a user would tend to stay on the same route unless the alternative is a great deal better. Although most of the economic models assume rationality on the part of decision makers [for instance Reebie Associates (*16*), Yalcin et al. (*18*), and Knorring et al. (*19*)], that might not be an innocuous assumption. Users could be bounded rational for various reasons such as complex and unpredictable toll structures (*20*) or even irrational bias (*14, 21*). Consequently, they might adopt second-best instantaneous decisions. Bounded rationality can also lead to lags in behavioral adaptation (*22*), leading to short- and long-term variation in responses. All of these factors might give rise to additional demand-side risk, especially for greenfield projects. Another important implication of bounded rationality is that it forces stated choice experiments to incorporate a limited set of variables, typically time savings and toll costs. Models derived from these surveys presume perfect knowledge and short-term preferences of the decision maker (*4*). This, in turn, could affect the route diversion estimates and revenue forecasts. Bounded rationality, combined with the evidence on response variability across user groups, does suggest that it is important to recognize which user groups comprise the gross truck volumes for a candidate toll road.

METHODOLOGY

A methodological framework is developed here for decision making at the strategic and planning level. This involves three distinct steps. First, a simple model is presented to incorporate the effect of vehicle operating cost savings on the demand side along the lines of Adkins et al. (*23*). Next, simulated distributions of the uncertain variables (VOT and toll revenue) are obtained. Finally, stochastic dominance analysis is used to assess the risk associated with these variables. The simulation study is based on a hypothetical corridor, much in line with Veras et al. (*7*).

Cost–benefit analysis is taken to be the basic behavioral postulate for the decision maker. In regard to benefits for the trucking industry, the economic significance of toll lanes includes productivity gains through time saved and potential changes in operating costs in addition to improved safety and reliability. Carriers therefore need to consider vehicle operating costs in addition to time savings when making route choices (*1, 4*). To help fix ideas, consider a hypothetical tolled route (of length D mi) that saves time (Δt) as well as operating cost (Δc_i) for the ith user group, however at a toll cost equal to τ. It is assumed that cost savings include only the distance-dependent portion. Thus user i, characterized by value of time v_i, would be willing to take the tolled route only when the following condition holds good:

$$v_i \cdot \Delta t + \Delta c_i \cdot D \geq \tau \cdot D \tag{1}$$

In equilibrium, when this holds with equality, the threshold value point for user i is given by Equation 2. [Equality is taken as a working assumption to obtain an analytical form of the threshold value.

Also, this threshold value is needed for evaluating the integral in Equation 3. Because for a continuous distribution, the probability at a single point is zero, evaluating the integral from $v^{\text{threshold}}$ or some $v^{\text{threshold}} + \epsilon$ would give the same probability.]

$$v_i = \frac{(\tau - \Delta c_i) \cdot D}{\Delta t} \tag{2}$$

Thus, for a toll road that involves operating cost savings, the threshold value of time and consequently demand for toll roads would depend on the cost savings. It is worth taking a closer look at Equation 2. It incorporates the difference between the state of the practice and the proposed method in imputing VOT. Going by tradition, if operating cost changes are ignored, VOT is simply a ratio of toll cost and time savings. Therefore ignorance of operating costs and cost heterogeneity can lead to biased estimates of VOT. To the extent that the tolled route leads to operational cost savings (dis-savings), this bias will be in the upward (downward) direction. This leads to the first proposition of the paper: Proposition 1. Ignorance of operating cost savings of the user group would lead to an upward bias in the implied (boundary) value of time distribution. For a scenario in which operating costs are likely to be higher on the tolled route, this bias will be in the downward direction.

The probability that value of time of a user sampled at random will exceed a prespecified threshold level ($v_i^{\text{threshold}}$) is given by Equation 3, which also captures the probability of diversion to the tolled route (*6, 8*). The symbol ~ is reflective of the fact that the underlying variable is stochastic.

$$\text{Prob}\left(\tilde{v}_i \geq v_i^{\text{threshold}}\right) = \int_{v_i^{\text{threshold}}}^{\infty} f\left(\tilde{v}_i\right) d\tilde{v}_i \tag{3}$$

Thus, as pointed out by Hensher and Goodwin, the use of mean VOT to approximate a skewed distribution would yield a biased number of users adopting the tolled route (*8*). For this hypothetical segment, the expected toll revenue (R) can now be expressed as

$$\tilde{R} = \sum_{i=1}^{n} N_i \cdot \text{Prob}\left(\tilde{v}_i \geq v_i^{\text{threshold}}\right) \cdot \tilde{\tau} \cdot D \tag{4}$$

where N_i denotes the truck volumes in the ith user group. The scaled revenue for a total of N ($=\Sigma N_i$) trucks can thus be expressed as

$$\frac{\tilde{R}}{N} = \sum_{i=1}^{n} \left(\frac{N_i}{N}\right) \cdot \text{Prob}\left(\tilde{v}_i \geq v_i^{\text{threshold}}\right) \cdot \tilde{\tau} \cdot D \tag{5}$$

Denoting the proportion of trucks in the ith group by θ_i, Equation 5 can be rewritten as

$$\frac{\tilde{R}}{N} = \sum_{i=1}^{n} \theta_i \cdot \text{Prob}\left(\tilde{v}_i \geq v_i^{\text{threshold}}\right) \cdot \tilde{\tau} \cdot D \tag{6}$$

Given the pedagogical purpose of this paper, the number of splits is restricted to private and for hire without any loss of generality. In principle, the splits would depend on the categorization of corridor users, say, by number of axles and should also include owner–operators. However, because of the dearth of VOT parameters for the latter, they are not being considered here. Thus for the assumed splits, Equation 6 becomes

$$\frac{\tilde{R}}{N} = \left[\theta_{\text{private}} \cdot \text{Prob}\left(\tilde{v}_{\text{private}} \geq v_{\text{private}}^{\text{threshold}}\right) + \theta_{\text{for hire}} \cdot \text{Prob}\left(\tilde{v}_{\text{for hire}} \geq v_{\text{for hire}}^{\text{threshold}}\right)\right] \cdot \tilde{\tau} \cdot D \tag{7}$$

This equation will be applied while simulating the (scaled) toll revenue. In this formulation the revenue equation accommodates response variability across user groups. [Equation 7 is directly comparable to the general equation for estimating truck toll revenue, where

$$\begin{aligned}\text{truck\_toll\_revenue}_t = {} & \text{average\_daily\_volume}_t * \text{truck\%}_t \\ & * \text{assumed\_diversion\_rate}_t * \text{toll\_rate}_t \\ & * \text{distance} * \text{annual\_revenue\_days}\end{aligned}$$

As a methodological advancement, this method directly incorporates the effect of (*a*) demand characteristics, (*b*) segmentation by user groups, and (*c*) usage propensities. A periodwise treatment of this equation could allow one to differentiate other factors such as short- and longer-term variations in usage.]

In contrast, if user heterogeneity were ignored, then the expected toll revenue from all trucks would be given by

$$\frac{\tilde{R}}{N} = \text{Prob}\left(\tilde{v} \geq v^{\text{threshold}}\right) \cdot \tilde{\tau} \cdot D \tag{8}$$

Given these basic formulations in the background, a simulation-based study is presented in the ensuing section.

SIMULATION AND STOCHASTIC DOMINANCE ANALYSIS

As noted earlier, the financial viability of a proposed toll road relies on a host of factors. Although improved revenue forecasts do rely on better data to capture user heterogeneity, this information may not be readily available in early stages. Simulation becomes particularly useful in this context. The risk associated with the simulated distributions of uncertain variables can be further analyzed by using stochastic dominance analysis. Typically, there are two criteria to choose from—first and second order stochastic dominance (FSD and SSD, respectively). Stated simply, when one distribution yields unambiguously higher returns and realizations, then it is said to first order stochastically dominate the other distribution. However, if a distribution is less risky, then it is said to dominate the other distribution in a second order sense.

The policy implications of dominance analysis are noteworthy. Although higher realizations are preferable to risk-averse decision makers, they would also prefer a distribution that is less risky. It is this ranking that can provide valuable insight while feasibility analysis is conducted. For instance, it might be the case that even in the presence of much uncertainty in the input variables, the risk associated with toll revenue is within acceptable bounds to have it financed. Alternatively a huge variability in toll revenue would indicate the possibility of optimism bias. A priori, without making a quantitative assessment of risk, the decision maker cannot foretell which one will occur. It is in this sense that simulation provides analytically superior results in comparison with simple ceteris paribus sensitivity analyses.

The two dominance criteria are used to rank the risk associated with the state of practice vis-à-vis the proposed method outlined in this paper. Specifically, two sets of exercises are taken up: Exercise 1—VOT distributions are simulated from Equation 2 by incorporating (*a*) only time savings and (*b*) both time and operating cost savings.

The resultant simulated probability distributions are then ranked by using the appropriate criterion (FSD or SSD) to assess the degree of risk associated with the threshold VOT distributions. Exercise 2—Probability distributions of toll revenue are simulated (*a*) with and (*b*) without considering user heterogeneity by using Equations 7 and 8, respectively. In particular, the risk associated with toll revenue is compared by assuming (*a*) a single VOT distribution for the entire trucking industry and (*b*) separate VOT distributions for the assumed splits (private and for hire). The basic steps for obtaining the simulated distributions are enumerated below, followed by a detailed description of the model inputs.

Step 1. Given that accurate prediction of toll revenue is an extremely daunting task, the most important step is to isolate the uncertain factors captured in Equation 2 (for threshold VOT) and Equations 7 and 8 (for revenue).

Step 2. Workable assumptions are made with respect to each of the uncertain variables. These are the key input variables of the model.

Step 3. Key output variables (KOV in the parlance of simulation) are obtained as a function of the uncertain input variables.

Step 4. Simulation stage. The KOVs are simulated for a thousand iterations, using the Latin hypercube sampling technique (by virtue of using stratified random sampling, Latin hypercube ensures additional accuracy at the tails in comparison with the Monte Carlo technique).

Step 5. Probability distributions of the KOVs are obtained. By using the appropriate decision criterion (FSD or SSD), a ranking is developed over these probability distributions.

With these fundamental principles in the background, the focus will now be on the variables that are relevant to truck diversion.

DESCRIPTION OF VARIABLES AND INPUT DATA

For the two sets of exercises outlined above, the working assumptions with respect to the key input variables are consolidated in Tables 1 and 2, respectively. Following the literature, VOT is specified as log normal (*6, 24*), given that it is nonnegative. Although it is best to obtain localized VOT distributions for different segments of the trucking industry, one might resort to "imported parameters" from other comparable studies (*1*). This would entail, however, a source of optimism bias. Nevertheless, for the exploratory purpose of this paper, the parameters are obtained from Kawamura (*6*). In regard to demand variation with respect to user group, the level of disaggregation considered here involves two kinds of splits: for hire and private.

Travel time savings (TTS) are also subject to uncertainty. As shown in Table 1, there are two different scenarios contingent on how TTS are modeled. The variation can be captured by assuming simple uniform distribution (Scenario 1) or triangular distribution (Scenario 2). The motivation behind using two different scenarios is to check the robustness of the rankings, discussed at length in the following section.

TABLE 1 Key Input Variables to Simulate Value of Time Distribution

Variable	Scenario 1	Scenario 2
Toll rate ($ per mile)	Uniform (0.10, 0.50)	Uniform (0.10, 0.50)
Travel time savings (minutes)	Uniform (5, 25)	Triangle (5, 10, 25)
Operating cost savings ($ per mile)	Uniform (0.10, 0.25)	Uniform (0.10, 0.25)
Distance (miles)	50	50

TABLE 2 Key Input Variables to Simulate Toll Revenue Distribution

Variable	Scenario 1	Scenario 2
Toll rate ($ per mile)	Uniform (0.10, 0.50)	Uniform (0.10, 0.50)
Travel time savings (minutes)	Uniform (5, 25)	Triangle (5, 10, 25)
Operating cost savings ($ per mile)	Uniform (0.10, 0.25)	Uniform (0.10, 0.25)
Distance (miles)	50	50
For hire private split	$\theta_{private}$ ~ uniform (0, 1) $\theta_{for\ hire} = 1 — \theta_{private}$	$\theta_{private}$ ~ uniform (0, 1) $\theta_{for\ hire} = 1 — \theta_{private}$
VOT	For hire ($v_{for\ hire}$) ~ log normal (−1.187, 0.924) private ($v_{private}$) ~ log normal (−1.78, 1.053) industry (v) ~ log normal (−1.467, 1.027)	

SOURCE: Kawamura (*6*).

[As an alternative, TTS was also modeled as a Gray, Richardson, Klose, and Schumann (GRKS) distribution, which is a two-piece normal distribution with 50% weight below and above the midvalue and 2.5% less than the minimum and 2.5% above the maximum. This distribution is potentially useful when the minimum as well as maximum values are uncertain (*25*). The qualitative results, and in particular the dominance rankings, were similar to those reported in the paper and therefore are not repeated for the sake of brevity.]

The use of triangular distribution can be justified on the following grounds. First, TTS typically follow finite range continuous distribution. Second, at the strategic level of planning, when limited information and data are available, it is difficult to specify the true underlying distribution of the travel time saved. The triangular distribution is intuitively appealing to the planner–analyst because it requires just three pieces of information—the minimum, modal, and maximum values—to derive an approximate distribution (*26*).

The heterogeneity in regard to vehicle operating cost can be captured from the existing literature (*27–29*). However, because these documents provide data at levels, it remains an important task for the analyst to make reasonable assumptions with respect to the cost savings. Given the exploratory purpose of this paper, heterogeneity is introduced by drawing random numbers from an assumed range of 10 to 25 cents per mile (the authors thank one of the reviewers for suggesting the range of variability).

Toll rates are assumed to follow a uniform distribution within the range of parameters specific to the trucking industry (*30*). Finally, the distance of the hypothetical corridor is taken to be 50 mi.

RESULTS

For the first set of exercises enumerated above, Equation 2 is used in generating VOT. Both the state of the practice and the proposed methodology (of incorporating time as well as operating cost savings) are implemented. The cumulative distribution functions (CDFs) as shown in Figure 1 provide a succinct way of depicting and comparing the risk associated with these distributions. Figures 1*a* and 1*b* correspond to the two different scenarios pertaining to TTS. From the CDFs it can be seen that exclusion of operating cost leads to higher realization of VOT at each probability.

Stoplight charts, as shown in Figure 2, are insightful in comparing two risky distributions. [Two target values (lower and upper target) need to be specified while obtaining the stoplight charts. For the subjective thresholds, these charts give the probabilities of (*a*) exceeding the upper target (green), (*b*) being less than the lower target (red),

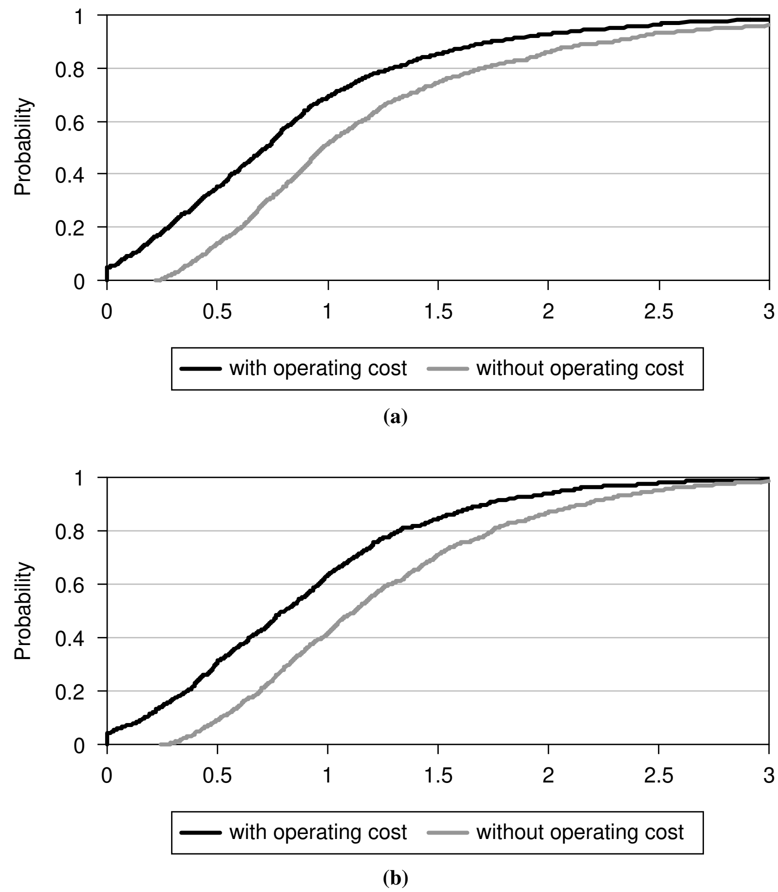


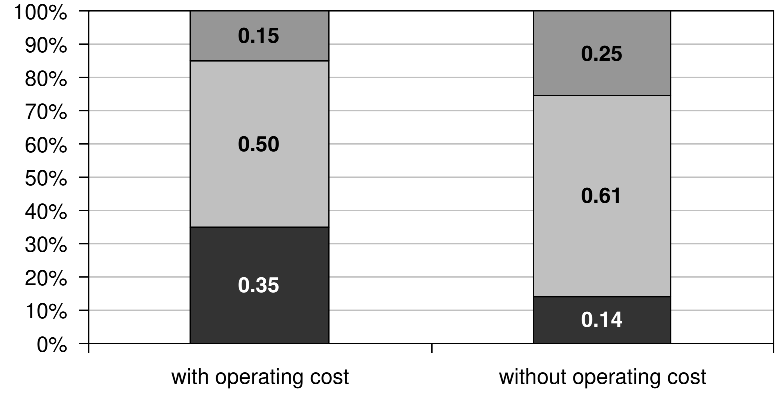
FIGURE 1 Cumulative distribution functions of VOT: (*a*) Scenario 1, travel time saved~uniform distribution, and (*b*) Scenario 2, travel time saved~triangular distribution.

and (*c*) observing values between the targets (yellow). A detailed description is available in Richardson (*25*).]

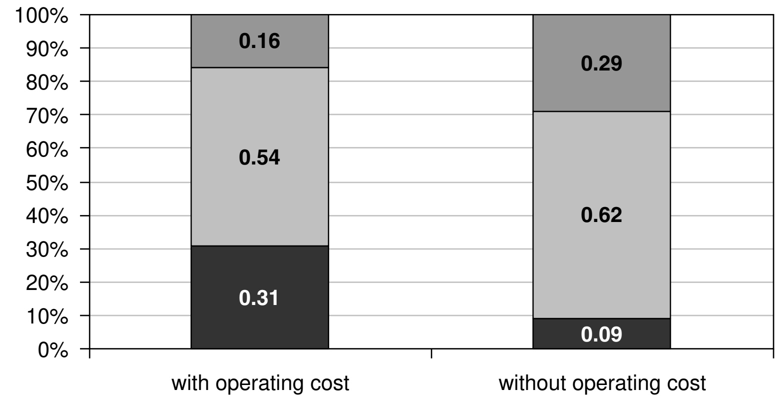
The two panels correspond to the two scenarios with respect to TTS. From Figure 2*a* it can be seen that by incorporating the cost savings, there is a 15% probability of exceeding a prespecified upper threshold (1.5) and a 35% chance of being below the lower threshold (0.5). In contrast, ignorance of operating cost savings leads to a 25% probability of exceeding the higher threshold, and a 14% chance of falling short of the lower threshold. Thus the state of practice is likely to give higher probability to higher values and at the same time lower probability to lower values. In other words, it yields a VOT distribution that tends to underestimate the revenue risk. This result remains valid for Scenario 2, which is depicted in Figure 2*b*.

Although cumulative distribution functions and stoplight charts are useful graphical tools, for analytical purposes an objective criterion should be used to choose between the two risky distributions. Dominance analysis indicates that VOT distribution obtained by ignoring operating cost differentials first order stochastically dominates the distribution in which operating cost has been accounted for. These findings lead to this paper's second proposition, Proposition 2: The value of time distribution obtained by ignoring operating cost would first order stochastically dominate the distribution when operating cost savings are accounted for. In other words, the state of the practice methodology would tend to underestimate the true risk.

For the second exercise, Equations 7 and 8 are used to generate the (scaled) probability distributions of toll revenue, which is the KOV now. The underlying assumptions are summarized in Table 2. In line with the previous exercise, revenue is first generated under the distributional assumptions of the input variables. The resultant probability distribution functions are shown in Figure 3, with the two panels corresponding to the two scenarios for TTS. The horizontal axes relate to scaled revenue, which can be multiplied by actual traffic volume to obtain the aggregate revenue. The probability distribution and the ranking would however remain unaffected by this rescaling. The stoplight charts associated with the probability distributions are presented in Figure 4. Technically speaking, the revenue distribution without user heterogeneity second order stochastically dominates the distribution obtained by incorporating user heterogeneity. Irrespective of the two scenarios considered, the conclusion remains robust that ignorance of heterogeneity is tantamount to lowering of risk associated with expected revenue distribution. This in turn, explains the other source of optimism bias stemming from usage patterns. This leads to the third and final proposition of this paper, Proposition 3: Ignorance of the underlying heterogeneity of the potential user groups (in regard to operating cost and value of time parameters) would lead to lower risk associated with projected toll revenue. With value of time varying across different users, adopting a "single" value of time parameter for the entire industry would lead to optimism bias in truck toll forecasts.

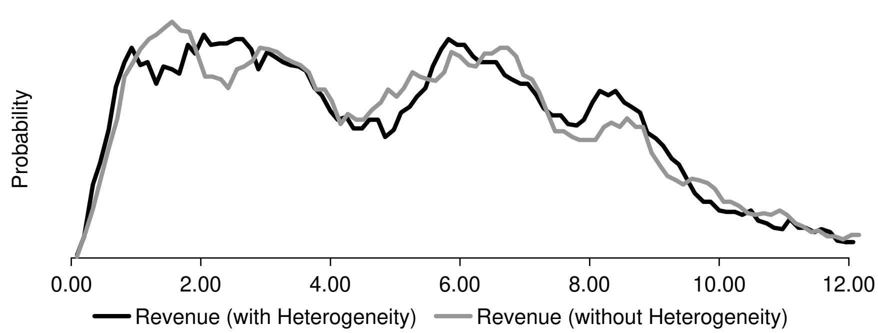


(a)

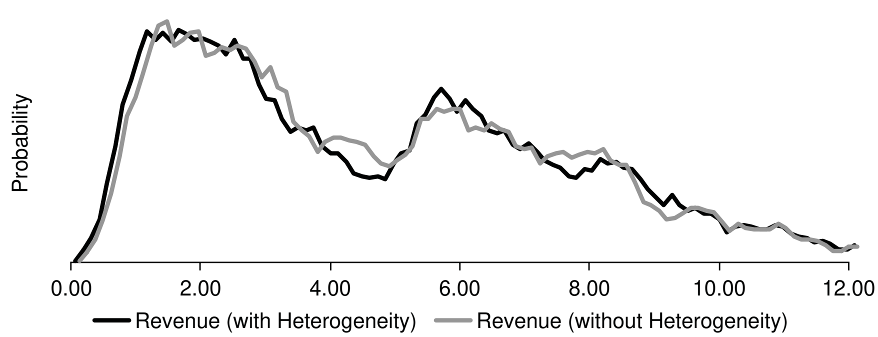


(b)

FIGURE 2 Stoplight charts for simulated VOT distributions (using Lower Cutoff 0.5 and Upper Cutoff 1.5): (*a*) Scenario 1, travel time saved~uniform distribution, and (*b*) Scenario 2, travel time saved~triangular distribution.



(a)



(b)

FIGURE 3 Probability distribution functions of revenue: (*a*) Scenario 1, TTS~uniform distribution, and (*b*) Scenario 2, TTS~triangular distribution.

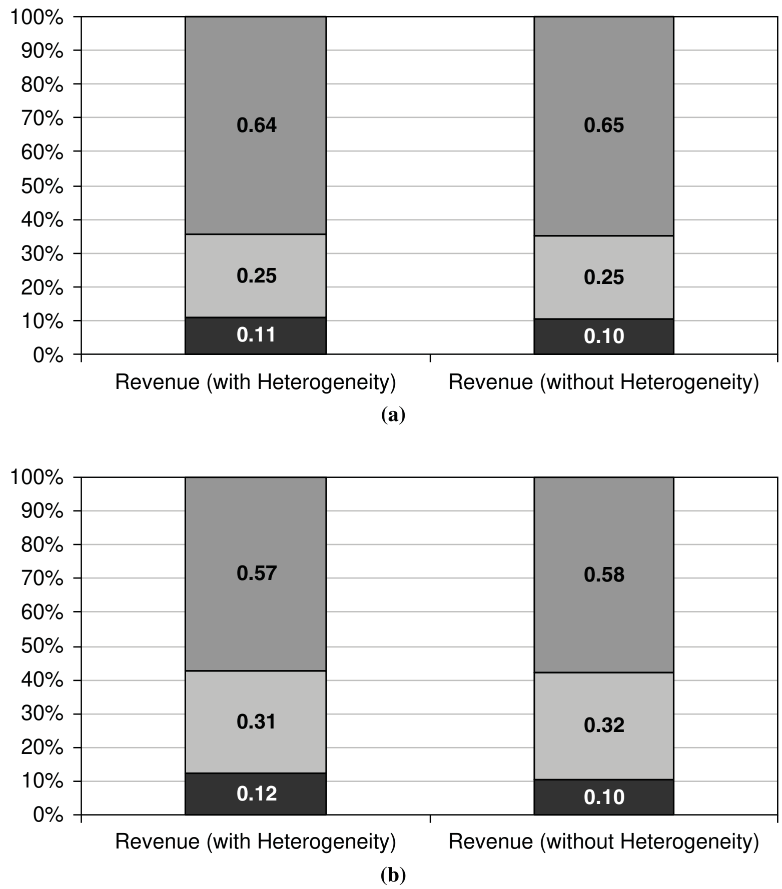


FIGURE 4 Stoplight charts for simulated revenue distributions (using Lower Cutoff 1 and Upper Cutoff 3): (*a*) Scenario 1, TTS~uniform distribution, and (*b*) Scenario 2, TTS~triangular distribution.

DISCUSSION OF RESULTS

There are several important messages that can be gleaned from this paper. First, stochastic dominance analysis can be particularly insightful in conducting feasibility studies associated with toll revenue forecasting. Consistent with the exploratory objectives of this paper, an application of this method was demonstrated in assessing the risk associated with truck toll forecasts. The approach outlined here could very well be applied to a real-life scenario and might be expanded to include other variables affecting demand for toll roads. For instance, several parameters enumerated in Tables 1 and 2 can be taken as close approximations of highly congested I-35 and its alternative tolled route SH-130. Exhaustive analysis would however require localized parameters with respect to the splits, VOT, and operating cost savings of the potential user groups. Using "industrywide averages" should be avoided, as suggested by Hensher and Goodwin (*8*) and Kawamura (*6*).

Second, from a planning perspective, because truck-related projects are also subject to a high degree of optimism bias (*1*), this research would imply that the candidate corridors be thoroughly examined for various freight movement types (truckload–less than truckload), ownership of users, and the distribution of owner operators. Also, given that cargo can be a significant source of variation in route decision (*15*) and that operating costs do vary by cargo type (*28*), it is prudent to include them in the analysis. The Reebie study, for instance, found significant diversion for coal and indicated that extremely bulky commodities with specific production–consumption locations would tend to be relatively inelastic and exhibit "all or nothing" type diversion (*16*). Note, however, that the propensity of diversion by commodity type is an empirical question and requires further research. Currently, the only known approaches to accomplish these goals are to obtain the parameters from (*a*) nationally available data sources such as REEBIE Transearch, FHWA Freight Analysis Framework, and Census Vehicle Inventory and Use Survey, (*b*) commercial vehicle surveys, and (*c*) local corridor studies. National data suffer from limitations in that they can at best provide county-level, first-cut assessments of user heterogeneity for some aspects of freight movement. When possible, localized truck origin–destination studies are superior alternatives.

Third, even though the simulations were conducted with a limited number of user groups (driven by knowledge of VOT parameters), the results could be generalized to accommodate more user groups as long as enough information exists with respect to their willingness to pay and other variables such as operating costs. The typical scenario assumed for this paper is a case in which the existing roadway is characterized by a mix of users.

Fourth, the bias in toll revenue emanating from the two factors explored in this paper might get dampened under certain circumstances. Travel time reliability, for instance, sometimes becomes an overriding factor behind route decision. This may or may not be correlated with VOT of the user (*4*). The methodology outlined in this paper can accommodate this bias through a reduction in the variance of TTS. Other empirical characteristics, such as a higher proportion of trucks carrying time-sensitive goods, can also reduce the bias.

This only corroborates the point made earlier that understanding usage patterns and traffic composition in regard to user groups can go a long way in improving the forecasts.

CONCLUSIONS

This paper explores two crucial factors behind optimism bias associated with truck toll forecasts. It is demonstrated that ignorance of operating costs might lead to upward bias in the implied VOT distribution. Also, by using stochastic dominance analysis, it is shown that ignorance of heterogeneity among freight user groups can lead to a less risky distribution associated with projected toll revenue. User heterogeneity is incorporated through (*a*) specific composition of users and (*b*) variation in preferences–demand parameters. This research is therefore important for proper assessment of the uncertainty associated with a proposed toll road. Although stochastic dominance is applied to analyze the implications for threshold VOT and of heterogeneous users on revenue risk given stochastic tolls, travel time, and costs, the analysis is only illustrative of the opportunities it offers to accommodate all factors that could belong to the revenue equation and risk in general. The particular strengths of this approach might be exactly in those situations in which the key input variables are interdependent. To strengthen and improve the analysis, further research is needed along several avenues, which include but are not limited to (*a*) better understanding of variations in preferences by cargo attributes and (*b*) incorporation of decision makers' risk preferences. Although the approach has been applied to truck toll forecasts, the approach discussed in this paper is general and can be applied to all toll forecasts.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided by the Southwest University Transportation Center. The authors are thankful to two anonymous referees and James Richardson for helpful comments on the earlier draft.

REFERENCES

1. Bain, R., and M. Wilkins. *Credit Implications of Traffic Risk in Start-Up Toll Facilities.* Standard & Poor's, New York, 2002.
2. Bain, R., and L. Polakovic. *Traffic Forecasting Risk: Study Update 2005: Through Ramp-up and Beyond.* Standard & Poor's, New York, 2005.
3. Bain, R. *Traffic Forecasting Risk Study Update 2004—Through Ramp-up and Beyond.* Standard & Poor's, London, 2005.
4. *NCHRP Synthesis of Highway Practice 364: Estimating Toll Road Demand and Revenue.* Transportation Research Board of the National Academics, Washington, D.C., 2006.
5. American Trucking Research Institute. *Top Industry Issues,* 2005. atri-online.org/research/results/economicanalysis/index.htm. Accessed July 1, 2007.
6. Kawamura, K. Perceived Benefits of Congestion Pricing for Trucks. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1833,* Transportation Research Board of the National Academics, Washington, D.C., 2003, pp. 59–65.
7. Holguin-Veras, J., D. Sackey, S. Hussain, and V. Ochieng. Economic and Financial Feasibility of Truck Toll Lanes. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1833,* Transportation Research Board of the National Academics, Washington, D.C., 2003, pp. 66–72.
8. Hensher, D. A., and P. Goodwin. Using Values of Travel Time Savings for Toll Roads: Avoiding Some Common Errors. *Transport Policy,* Vol. 11, 2004, pp. 171–181.
9. Yang, H., W. H. Tang, W. M. Cheung, and Q. Meng. Profitability and Welfare Gain of Private Toll Roads in a Network with Heterogeneous Users. *Transportation Research,* Part A, Vol. 36, 2002, pp. 537–554.
10. Kenneth, D. B. *Principles of Transportation Economics.* Addison Wesley Longman, Inc., 1997, pp. 20–24.
11. Verhoef, E. T., and K. A. Small. Product Differentiation on Roads: Second-Best Congestion Pricing with Heterogeneity Under Public and Private Ownership. *Journal of Transport Economics and Policy,* Vol. 38, 2004, pp. 127–156.
12. McKnight, C. E., I. Hirschman, J. R. Pucher, J. Berechman, R. E. Paaswell, J. A. Hernandez, and J. Gamill. *Optimal Toll Strategies for the Triborough Bridge and Tunnel Authority.* Final Report. 1992. www.utrc2.org/research/assets/51/tolls1.pdf. Accessed Sept. 30, 2007.
13. URS Corporation and Vollmer Associates, L.L.P. *Central Texas Turnpike System: Traffic and Revenue Forecast Update.* 2005. www.centraltexasturnpike.org. Accessed July 1, 2007.
14. Taft, B. *Northern Ohio Freight Strategy: Recommendations to Improve Traffic Safety and Congestion.* www.dot.state.oh.us/news/2004/10-11-04Gov.htm. Accessed July 1, 2007.
15. Veras, J. H., Q. Wang, N. Xu., K. Ozbay, M. Cetin, and J. Polimeni. The Impacts of Time of Day Pricing on the Behavior of Freight Traffic in a Congested Urban Area: Implications to Road Pricing. *Transportation Research,* Part A, Vol. 40, 2006, pp. 744–766.
16. Reebie Associates. *The Impact of Tolls on Freight Movement for I-81 in Virginia.* Final Report. Department of Rail and Public Transportation, Virginia, 2004.
17. Simon, H. A. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics,* Vol. 69, 1955, pp. 99–118.
18. Yalcin, A., J. Hashiuchi, and S. Mizokami. A Study on Expressway Toll Pricing Based on Results of Social Experiments. Presented at 85th Annual Meeting of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C., 2006.
19. Knorring, J. H., R. He, and A. L. Kornhauser. Analysis of Route Choice Decisions by Long-Haul Truck Drivers. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1923,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 46–60.
20. Bonsall, P., J. Shires, J. Maule, B. Matthews, and J. Beale. Responses to Complex Pricing Signals: Theory, Evidence and Implications for Road Pricing. *Transportation Research,* Part A, Vol. 41, No. 7, Aug. 2007, pp. 672–683.
21. Van Zyl, N. J. W., and M. Raza. In Search of the Value of Time: From South Africa to India. In *Travel Survey Methods: Quality and Future Directions* (P. Stopher and C. Stecher, eds.), Pergamon Press, New York, 2006, pp. 457–484.
22. Gifford, J. L., and C. Checherita. Bounded Rationality and Transportation Behavior: Lessons for Public Policy. Presented at 86th Annual Meeting of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C., 2007.
23. Adkins, G. A., A. W. Ward, and W. F. McFarland. *NCHRP Report 33: Values of Time Savings of Commercial Vehicles,* HRB, National Research Council, Washington, D.C., 1967.
24. Smalkoski, B., and D. Levinson. *Value of Time for Commercial Vehicle Operators.* Working Papers 200501. University of Minnesota, Nexus Research Group, Minneapolis, 2005.
25. Richardson, J. W. *Simulation for Applied Risk Management.* Department of Agricultural Economics, Agricultural and Food Policy Center, Texas A&M University, College Station, 2006.
26. Williams, T. M. Practical Use of Distributions in Network Analysis. *Journal of the Operational Research Society,* Vol. 43, 1992, pp. 265–270.
27. Levinson, D., M. Corbett, and M. Hashami. *Operating Costs for Trucks.* Working Series Number 000024. University of Minnesota, Nexus Research Group, Minneapolis, 2004.
28. Barnes, G., and P. Langworthy. *The Per-Mile Costs of Operating Automobiles and Trucks.* Minnesota Department of Transportation, St. Paul, 2003.
29. *American Trucking Trends.* American Trucking Association, Washington, D.C., 2005–2006.
30. Veras, J. H., M. Cetin, and S. Xia. A Comparative Analysis of US Toll Policy. *Transportation Research,* Part A, Vol. 40, 2006, pp. 852–871.

The Freight Transportation Economics and Regulation Committee sponsored publication of this paper.