WELFARE CALCULATIONS IN DISCRETE CHOICE SETTINGS: THE ROLE OF ERROR TERM CORRELATION

Yong Zhao
Senior Transportation Analyst
TranSystems Corp
500 West 7th St, Suite 1100
Fort Worth, TX 76102
yzhao@transystems.com
Phone: 817-334-4472
FAX: 817-336-2247

Kara Kockelman
(corresponding author)
Associate Professor of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
Austin, TX 78712-1076
kkockelm@mail.utexas.edu
Phone: 512-471-0210
FAX: 512-475-8744

Anders Karlstrom
Associate Professor of Systems Analysis and Economics
Royal Institute of Technology
SE-100 44 Stockholm Sweden
Phone: +46 8 790 9263
andersk@infra.kth.se

The following paper is a pre-print and the final publication can be found in
Transport Policy 19 (1):76-84, 2012.
Presented at the 87th Annual Meeting of the Transportation Research Board, January 2008

ABSTRACT
A difference in logsum terms (also known as inclusive values) is becoming standard practice for anticipating the welfare impacts of transport policy when choice alternatives are discrete and behavior is (assumed to be) random-utility maximizing. This paper examines the effect of error term correlations in such analyses, recognizing that individual preferences and unobserved attributes influencing choice are unlikely to change much, if at all, across scenarios. Such measures appear reasonably robust to deviations in assumptions of correlation. What is most striking in these results is the substantial variation that emerges across synthetic populations, suggesting that policies that appear welfare-improving (when evaluated with average welfare formulations) may well be welfare-reducing (or vice versa) for a wide variety of actual populations. Another finding is that the synthetic population samples need to be very large for the mean to be realized (in finite populations).

Key Words: Welfare calculation, discrete choice, inclusive value, log-sum, compensating variation, error correlation

INTRODUCTION

The concept of welfare maximization is a key criterion guiding transportation policy decisions and system enhancements. And McFadden's (1978, 1981) logit model for random utility maximization (RUM) provides a microeconomically rigorous foundation.

Of course, the multivariate extreme value model (MEV)[1] is now standard practice in analyzing mode, destination, vehicle purchase and other travel decisions. In particular, the multinomial logit (MNL) is a mainstay of such models. And researchers have recognized the promise of such models in quantifying net benefits via logsum differences (see, e.g., Small and Rosen 1981, Handy and Niemeier 1997, de Jong et al. 2005, and Gupta et al. 2006).

In recent years, travel demand modelers and transportation planners have moved toward the use of logsum differences (such as the user benefit calculation in the FTA's [2005] Summit model and the U.K.'s TUBA model [MacDonald 2006]) in order to more rigorously characterize changes in net social benefits for actual project alternatives. Such measures constitute important tools in project evaluation, including NEPA-required environmental justice considerations (see, e.g., FHWA, 2000). However, they do rest on at least two core assumptions: a constant marginal utility of money and perfectly correlated error terms (between before/after scenarios for each choice alternative). Karlstrom (1998, 2001), Franklin (2006), Cherchi et al. (2005), and Small et al. (2006) have examined the implications of using logsum differences under this first assumption in depth. Using Karlstrom's (1998, 2001) exact methods, Franklin (2006) found that income effects can be sizable when evaluating road pricing projects, especially for those where travel costs loom large (relative to income). Franklin also examined distributional effects closely (via Gini coefficients and Theil's entropy measure). As a complement to such work, this paper examines the impacts of the second assumption, relating to independence of error terms.

Given the limitations of identically and independently distributed GEV Type 1 (Gumbel) error terms, much research has focused on relaxing such assumptions (see, e.g., Swait's [2003] mixture model discussion), in addition to allowing random parameters (see, e.g., Hensher and Greene's [2003] review of such models). While such flexibility can introduce correlation across alternatives and enhance a model's behavioral realism, it tends to complicate the estimation process, as well as any associated welfare analyses. Only recently, have researchers (Cherchi and Polak 2005) examined the welfare implications of random-parameter kernel-logit models and suggested that the GEV error terms may not adequately represent underlying heterogeneity, negatively impacting benefits estimates. And the implications of correlation in alternatives' random error terms, pre- and post-policy implementation, remain unstudied.

---

[1] MEV models are generally known as Generalized Extreme Value (GEV) models in transportation science. However, the term MEV is more accurate, since GEV is used to describe a class of marginal extreme value distributions, as described later in this paper.

It seems quite natural to expect a high degree of correlation in (existing) choice alternatives (such as route, mode and destination) before and after an increase in toll charges, reductions in transit fares, an increase in schedule frequency, the addition of roadway capacity, the addition of a new alternative, and so on. Whatever it is that attracts one individual to a mode, but not others that share his/her attributes, is reflected in these error terms and likely to remain with these individuals after the context shifts. As McFadden observed (1978), a simple difference in logsums follows if all error terms are maintained, in the before/after choice environments (as developed in Karlstrom [1998, 2001]). What are the implications of such an assumption? This paper first discusses the basis for use of a difference in logsums and then examines a variety of policy scenarios for many travelers in a streamlined network in order to get a sense of the bias and mis-prediction, and potential for misplaced evaluations, under a series of error-term correlation assumptions and inclusion of a new route alternative.

BACKGROUND

Welfare evaluations calculate the changes in net social benefits, or consumer surplus, for situations before and after a policy's implementation (e.g., road pricing) or a project's development (e.g., expanded highway). Typically, use of additive, random-utility error terms (ε_{iq}) reflects heterogeneity across travelers (q) and choice alternatives (i) that are unobserved by the analyst, and thus uncontrolled for in the model specification.

$$U_{iq} = V_{iq} + \varepsilon_{iq} = V(x_{iq}, s_q) + \varepsilon_{iq} \quad \forall i \qquad (1)$$

where s_q denotes all traveler-related attributes (such as income and gender, which can be interacted with alternative-specific constants and/or with generic choice attributes).

Assuming that error terms ε_{iq} are iid and follow a Generalized Extreme Value (GEV) Type 1 distribution with zero mode (location parameter) and scale parameter β of 1 (which imply a mean of $\mu = 0 + 0.5772\beta = 0.5772$, and a standard deviation of $\sigma = \pi\beta/\sqrt{6} = 1.283$), the expected maximum utility can be expressed by the following logsum formula:

$$E(Max(U_{iq}, \forall i)) = \ln \sum_i e^{V_{iq}} \qquad (2)$$

More generally, interests focus on the difference between the before and after conditions of a particular policy, and a difference in logsums is applied:

$$\begin{aligned} E\left(Max(U_{iq}^1) - Max(U_{iq}^0)\right) &= E\left(Max(U_{iq}^1)\right) - E\left(Max(U_{iq}^0)\right) \\ &= \ln \sum_i e^{V_{iq}^1} - \ln \sum_i e^{V_{iq}^0} \end{aligned} \qquad (3)$$

Note that equation (3) holds when $\varepsilon_{iq}^1 = \varepsilon_{iq}^0$ and the marginal utility of money is constant (as alluded to in McFadden [1978] and examined at length in Karlstrom [1998, 2001]).

If cost and wealth terms enter the systematic conditional-utility specification in a simple, linear way (producing constant marginal effects), one can normalize this difference in logsums by the marginal utility of money, rendering the amount of money that, in theory, would be needed to compensate individuals for enduring the policy change (Small and Rosen 1981). This "compensating variation" (CV) maintains original levels of utility (see, e.g., Varian 1992) and thus serves as a Hicksian (or compensated) version of consumer surplus. Thanks to the (assumed) lack of income effects, the Marshallian (uncompensated, but observable) and Hicksian demand curves overlap, and *changes* in consumer surplus (CS) are reflected by the CV term (see, e.g., Karlstrom 1998, 2001, and Karlstrom and Dagsvik 2005).

$$E(CV) = \frac{1}{\alpha}\left(\ln \sum_i e^{V_{iq}^1} - \ln \sum_i e^{V_{iq}^0} \right) \tag{4}$$

where $\alpha = -\frac{dV}{dc}$ is the marginal utility of money and assumed to be constant.

Of course, if individuals' unique affinity or distaste for the alternatives exists, due to unobserved, case-specific attributes (like certain routes offering fewer stoplights than another) or unobserved, person-specific preference variations (for example, an added attraction to alternative *i* by person *q*), correlations will emerge in the before and after conditions. These correlations may be perfect in fact (assuming, for example, the before/after periods are not too far apart, and contain the same population, as is standard for such evaluations). As Karlstrom (2001) suggests, the same error terms generally should be used in such welfare calculations, in order to ensure consistency of and comparability across cases. But, of course what if there are arguments to the contrary and/or new alternatives are added to the choice set, as part of the policy being examined? How will logsum differences fare in such cases? Given the size of errors emerging from the use of logsum differences when estimating welfare changes in the presence of income effects (Cherchi et al. 2005 and Franklin 2006), it is worth considering the effects of imperfect error-term correlation in before/after instances.

Unfortunately, an intractable integral is needed to compute the welfare effect of policies and projects in the presence of less-than-perfectly correlated error terms. Here, numerical simulations are used to ascertain the size of such effects under a variety of settings, with perfect and lesser levels of correlation. These scenarios and their results are described in the following sections.

NUMERICAL EXAMPLES

The impacts of assuming no error-term correlation in computing measures of welfare are illustrated here, for a streamlined example of route choice. This 6-node, 11-link network (Figure 1) offers meaningful realism without great complication, across a variety of zone pairs. This skeleton network is a simplified version of a proposed rural toll road network in south Texas. In both cases, trip-making is assumed inelastic between all zone pairs, and travel times and costs are independent of link flows. Of course, in reality, destination choices are impacted by changes in system conditions, and (total) trip generation rates may be as well.

Motorists are facing the travel choices between free roads and toll roads. The random utility functions are specified as follows:

$$U_{free,q} = 0.120 - 0.15T_{free,q} + \varepsilon_{free,q}$$
$$U_{toll,q} = -0.15T_{toll,q} - 0.783C_{toll,q} + \varepsilon_{toll,q} \quad (5)$$

where $T_{free,q}$ and $T_{toll,q}$ are interzonal travel times (in minutes, as shown in Table 1) via the non-tolled and tolled routings for each O-D pair, and $C_{toll,q}$ is the toll charge on the tolled route (in dollars, as shown in Table 2).

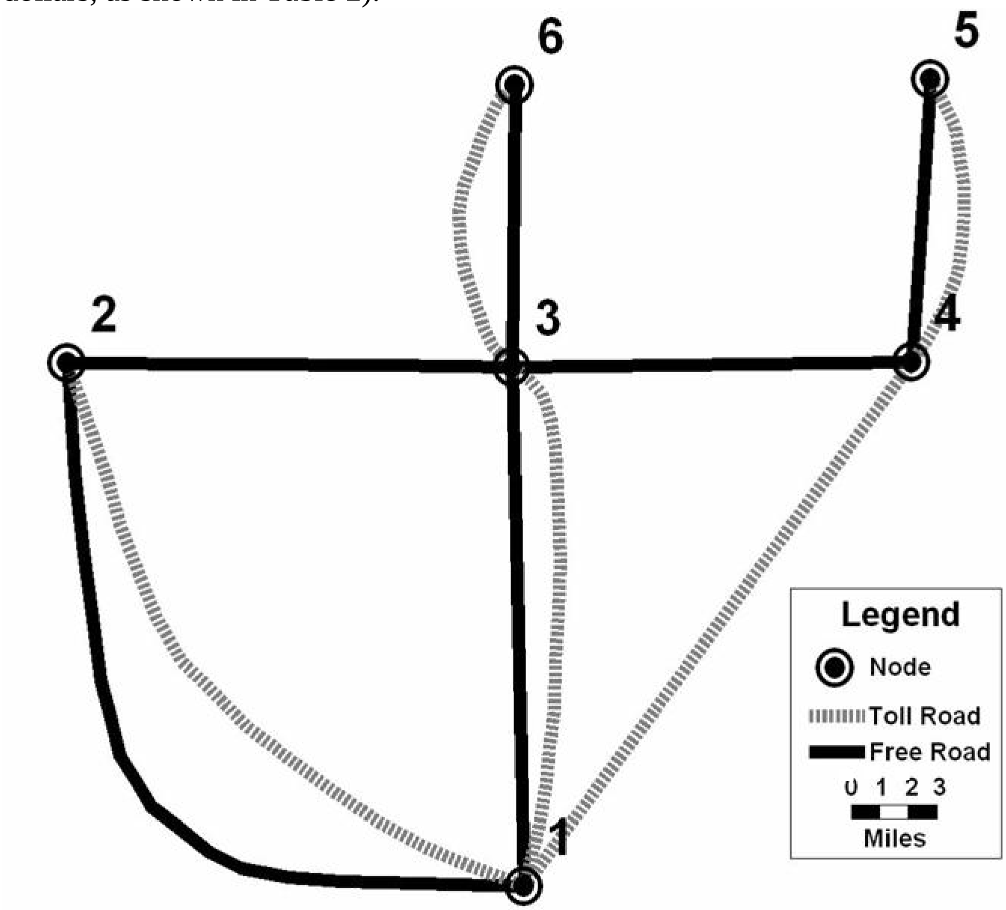

Figure 1. 6-node, 11-link test network (approximately 30 miles across)

Model parameters are based on results of a year-2006 stated preference survey in the Dallas-Fort Worth area (RSG 2006), and equation (5) implies a value of time of \$11.50 per hour. The toll charges are based on a toll rate of \$0.15/mile[2]. For roughly half of the OD pairs, travelers enjoy two alternatives, and each traveler is assumed to select the one which maximizes his/her (latent) utility.

---

[2] 15¢/mile is very similar to what State Highway 121 (in the DFW region) and other new and forthcoming Texas toll roads are/will be charging.

Table 1. Travel times via non-tolled and tolled roads (for all OD pairs)

Time (min)	1		2		3		4		5		6	
	Free	Toll	Free	Toll	Free	Toll	Free	Toll	Free	Toll	Free	Toll
1	-	-	38.91	22.52	24.27	15.69	40.83	20.86	52.59	29.17	33.87	-
2	38.91	22.52	-	-	14.65	-	31.21	-	42.97	-	24.25	-
3	24.27	15.69	14.65	-	-	-	16.56	-	28.32	-	9.60	7.00
4	40.83	20.86	31.21	-	16.56	-	-	-	11.76	8.31	26.16	-
5	52.59	29.17	42.97	-	28.32	-	11.76	8.31	-	-	37.92	-
6	33.87	-	24.25	-	9.60	7.00	26.16	-	37.92	-	-	-

Note: Travel times for unreasonable toll paths are not evaluated here. “Free” implies non-tolled routing.

Table 2. Toll charges on all tolled routes (by OD pair)

Toll Charges	1	2	3	4	5	6
1		$3.70	$2.60	$3.40	$4.70	
2	$3.70					
3	$2.60					$1.00
4	$3.40				$1.30	
5	$4.70			$1.30		
6			$1.00			

The two policy scenarios to be evaluated are as follows: (1) 15% toll rate increases on all tolled routes (without any travel times changes), (2) a $0.05-per-mile congestion pricing scenario (which assumes a $0.05-per-mile toll and 15% travel time reductions on all previously “free” roads). Of course, the network is not being loaded, so these shocks are exogenously presumed here (rather than based on an interplay of roadspace supply and demand). These two scenarios are evaluated for Figure 1’s 6-zone network, first at the level of a single OD pair (from zone 4 to 5, which has 767 one-way travelers) and then for the entire system of users (of 230,892 one-way travelers). Simulated differences in maximum random utilities were compared to the standard CV estimates (based on logsum differences), and these results are provided in the following section.

EMPIRICAL RESULTS

Consider a **scenario where toll charges increase 15%**. Under assumptions of perfect correlation in error terms in the before and after scenarios, CV measures can be estimated using equation (4). Since our interest lies in other, imperfectly correlated cases, large-sample Monte Carlo simulation of correlated Gumbel error terms (using @Risk software [Palisade, 2002]) provides rather precise estimates of the distribution of differences one can expect (over the traveler population space), including an average welfare impact per traveler. Such results were sought here, using R = 1,000,000 draws for each of the travelers using a variety of error-term correlation levels.

Welfare calculations are examined in two fashions: at the level of a single O-D pair (4, 5), and across the entire system of O-D pairings. Based on equation (5), the percentage of toll road users between zones 4 and 5 before and after the toll rate increase will be 35.0% and 31.5%, respectively.

The 1,000,000-draw simulation for this origin-destination pair can be interpreted as though there are 1 million cases of each individual facing the same choice situation, and thus sharing the same observed characteristics. In each simulation, a single, (pseudo-)randomly drawn error term represents each individual's additive random utility, or unobserved preference for one of the two routing alternatives. To ensure proper correlation, the before and after error terms are drawn from a joint multivariate distribution where marginal distributions of each error term are standard GEV Type 1 (0 mode and scale 1), and the correlation matrix between before and after error terms ranges from zero to one, according to the prescribed scenarios.

Table 3 summarizes the average total welfare impact (or average sum of CV estimates) across these 1 million synthetic populations. The average CV value represents the average welfare impact of the new tolling policy, per traveler, and is simply the total CV estimate divided by the number of travelers (767 for this zone pair). Examination of the 1 million CV totals produces precise estimates of the standard deviation and a 90% (5% to 95% data points) confidence interval for individual welfare changes, giving one an idea of how variable populations may be, in terms of policy impacts.

Table 3. Estimates of welfare changes for travel between zones 4 and 5, following an increase in toll rates

Error Term Correlation (Before/After)	Total CV ($)	Index	SD ($)	5% Lower Bound ($)	95% Upper Bound ($)
0.00	-37.91	100.00	1391.2	-2295.5	2221.2
0.25	-38.03	100.32	1251.5	-2069.8	1987.2
0.50	-38.02	100.29	1063.7	-1756.9	1681.4
0.75	-38.63	101.88	783.2	-1308.4	1223.6
1.00	-38.94	102.70	54.5	-117.0	0.0

Notes: Index is the ratio of total welfare change to the zero-correlation case result.

Somewhat surprisingly, as evident in Table 3, 1 million draws are not enough to bridge the gap between zero correlation and perfect correlation. Perfect correlation yields a 2.7% lower estimate of total CV (across all zone pair 4, 5 travelers), as compared to the zero-correlation case. The case of 0.5 correlation produces a 0.29% lower estimate. A key finding of these simulation results is the large variation inherent in CV estimates across synthetic populations. Additionally, as error term correlation increases, the variation falls dramatically; in this case, it remains larger than the average policy impact ($54.5 > $38.94). The decrease in such CV variations also can be seen from the confidence intervals (i.e., the range of 5% and 95% values).

It also should be noted that a total welfare loss of $38.94 per period (case of perfect correlation) across these 767 travelers is equivalent to a cost of 5.0765¢/traveler, per trip. And this result is essentially the same as the difference-in-logsums formula, which yields a cost of 5.0776¢/traveler. Given the accuracy of the simulation for this particular case (i.e., the precision at which preference parameters can be obtained), 1 million draws seem appropriate here, from an applied point of view. In fact, this number of draws is much higher than most simulations and maxes out memory stores rather quickly on standard computers. Higher precision seems unlikely, at this evolutionary stage of world computing.

For the **congestion pricing case**, the previously non-tolled route between zones 4 and 5 now carry a toll of $0.50 and Table 1's "free" travel times are reduced by 15% of their original levels, as seen in Table 4.

Table 4. Congestion pricing on all previously non-tolled routes (by OD pair)

Congestion Tolls	1	2	3	4	5	6
1		$1.47	$0.91	$1.60	$2.09	$1.39
2	$1.47		$0.78	$1.47	$1.96	$1.26
3	$0.91	$0.78		$0.69	$1.18	$0.48
4	$1.60	$1.47	$0.69		$0.50	$1.17
5	$2.09	$1.96	$1.18	$0.50		$1.66
6	$1.39	$1.26	$0.48	$1.17	$1.66	

35.0 and 37.9 percent of travelers are expected to use the tolled routes (between zones 4 and 5) before and after the congestion pricing scheme is implemented, respectively. This is not a significant change, and welfare impacts tend to be low, as summarized in Table 5, using sums of simulated CV values across the 767 travelers. As before, welfare impacts escalate (becoming more negative) as correlation grows, while variations in these estimates (across the 1 million simulated populations) fall off dramatically. Moreover, a simple difference in logsums yields a total impact of $61.74, very close to the perfect correlation simulation average shown in Table 5 – but 1.57% larger than a zero-correlation case. This also suggests that 1 million draws are not enough to bring the gap down to 1% – at least not in this case context.

Table 5. Estimates of welfare changes for travel between zones 4 and 5, following the introduction of congestion pricing

Error Term Correlation (Before/After)	Total CV ($)	Index	SD ($)	5% Lower Bound ($)	95% Upper Bound ($)
0.00	-60.80	100.00	1391.3	-2317.2	2198.2
0.25	-60.70	99.83	1252.8	-2093.6	1965.9
0.50	-60.83	100.05	1065.6	-1780.9	1662.6
0.75	-61.39	100.96	784.9	-1329.5	1208.4
1.00	-61.76	101.57	46.3	-97.1	0.001

Note: Congestion charge is 5¢/mile on previously non-roads roads that were found to be congested in the base case.

Similar to Table 3's tolling results, Table 5's values suggest that different beliefs and assumptions regarding error term correlations result in somewhat similar conclusions.

An interesting extension of this scenario is to consider the **case of an added alternative**. One may choose to keep the existing alternatives while adding a third path that is priced at a rate of 5¢/mile. To reflect growing congestion on the roadways, a future scenario where travel times on the non-tolled route are presumed to have grown by 18 percent makes this case study more meaningful, with travel times on the tolled alternative remaining as before and those on the newly added, priced alternative equaling those that existed previously (on the non-tolled alternative). Though travel times are assumed to rise on the non-tolled existing-road option, and the new route carries a toll, all routes remain competitive, with estimated average shares of

38.4% (non-tolled route), 33.3% (new route), 28.4% (existing tolled route) – of the 767 corridor travelers. In fact, the addition of the alternative has increased estimates of CV (Table 6) – so much so that they are now very positive, though they are associated with high variability. (The SD is now several times higher than the average total CV, at all error levels.) Evidently, the value of new travel alternatives should not be neglected, even if they carry tolls and travel times/costs rise on other routes. Also interesting is the fact that a simple difference in logsums, using 2 alternatives in the base case versus 3 in the new case, is nearly the same ($161.23 and $161.47) as the estimate from 1,000,000 simulations in the perfect-error- correlation case, suggesting that standard logsum differences are likely to be robust to situations of new alternatives. This is reassuring. The difference in CVs between the zero correlation case and the perfect correlation case is only $1, or 0.61% again with 1 million draws. Please note that the error term of the new alternative is assumed to be independent from these error terms of existing alternatives before the new one is introduced.

Table 6. Estimates of welfare changes for travel between zones 4 and 5, following the introduction of new, priced alternative

Error Term Correlation (Before/After)	Total CV ($)	Index	SD ($)	5% Lower Bound ($)	95% Upper Bound ($)
0.00	162.46	100.00	1391.8	-2097.4	2422.1
0.25	162.10	99.78	1289.3	-1925.5	2256.5
0.50	162.04	99.74	1154.3	-1675.1	2061.5
0.75	161.63	99.49	963.2	-1296.1	1796.2
1.00	161.47	99.39	663.5	-243.5	1577.0

Of course, a focus on a single zone pair is a bit limiting. Common practice in welfare evaluation is to present data aggregated over the system's travelers. Assuming the travel demand distribution shown in Table 7 for the **6-zone system**, total welfare effects for the proposed two scenarios can be estimated, as presented in Tables 8 and 9. The welfare estimation used here is based on 1,000,000 simulated draws (per context, individual and alternative).

Table 7. Daily travel demand for the 6-zone system

Daily Trips	1	2	3	4	5	6
1	-	17000	4803	4599	3173	7355
2	17000	-	16904	9503	5708	15263
3	4803	16904	-	4213	2219	12826
4	4599	9503	4213	-	767	7046
5	3173	5708	2219	767	-	4067
6	7355	15263	12826	7046	4067	-

Table 8. Estimates of welfare changes for travel between all zones, following an increase in toll rates

Error Term Correlation (Before/After)	Total CV ($)	Index	SD ($)	5% Lower Bound ($)	95% Upper Bound ($)
0.00	-12313.61	100.00	117903.2	-210373.4	181088.1
0.25	-12440.76	101.03	104215.5	-183630.8	158600.5
0.50	-12533.64	101.79	86838.8	-155748.0	129827.1
0.75	-12432.52	100.97	62759.8	-115781.9	90506.3
1.00	-12451.37	101.12	6609.6	-24914.9	-3042.7

Table 9. Estimates of welfare changes for travel between all zones, following the introduction of congestion pricing

Error Term Correlation (Before/After)	Total CV ($)	Index	SD ($)	5% Lower Bound ($)	95% Upper Bound ($)
0.00	-83,514	100.00	117,682	-276353.5	110188.0
0.25	-83,264	99.70	102,314	-241223.5	96504.8
0.50	-83,180	99.60	88,723	-204595.8	78996.8
0.75	-83,356	99.81	47,736	-152095.0	55071.1
1.00	-83,476	99.95	4,827	-91074.5	-75047.8

Tables 8 and 9 suggest that at the aggregate level, the effects of these policies are somewhat diluted, given the lesser variation in net welfare (Total CV) estimates from zero to perfect correlation cases and the increasing trend is less obvious. This is to be expected, given that 63% of the network's travelers (or 60% of its O-D pairs) do not enjoy an alternative (i.e., have only one route to choose from), in this real-network case examination; so welfare shifts are less likely for those travelers.

Related to this, the variation in net welfare estimates (over the 1,000,000 simulations of all 230,892 travelers) is less, in a relative sense, than for the zone pair 4,5 examination. Effectively, the ratios of SD to Total CV (or coefficient of variation in welfare estimates across synthetic populations) have fallen by factors ranging from 0.05 (case of perfect correlation for congestion pricing scenario) to 9.57 (zero correlation case in toll rate increase scenario). The presence of less variation in network attributes across the larger pool of travelers plays a key role in this result. Nevertheless, the variations in welfare results remain substantial, particularly when correlations are less than perfect. Such indicators of variation suggest that confidence intervals should be given on welfare measures (as presented in the above tables), to indicate the potential for real losses or benefits when the average result or simple estimates suggest otherwise.

Finally, it is useful to consider the estimate of revenues generated in the "congestion pricing" scenario. Across the system's 230,892 users (many of whom rely on the base-case toll roads), the addition of congestion tolls generates $169,891 in revenues (on average). This, of course, fully offsets the estimated welfare loss of roughly $83,500 (in all of the correlation cases, as evident in Table 9). If these revenues could be returned to all travelers in some form (e.g., uniformly distributed travel budgets, as suggested in Kockelman and Kalmanje's (2005) proposed policy of credit-based congestion pricing), net benefits enjoyed directly by travelers would be positive (on average and in general, across nearly all simulated populations in all correlation cases). This presumes that travel times will fall by 15%, demand between all O-D pairs is inelastic, implementation costs are negligible, and so forth. In reality, the net benefits of such policy may be larger or smaller.

Differences in logsums are small with this large number (1 million) of simulation draws. What is striking is the substantial variation that emerges across synthetic populations, suggesting that policies that may appear welfare-improving (when evaluated with average welfare formulations) are in reality welfare-reducing (or vice versa). The variation (or standard deviation, provided in above tables) depends on the nature of error term correlations. As noted earlier, the error terms reflect heterogeneity across travelers and choice alternatives. For nearly all cases, CV variations cover a wide range, suggesting that actual welfare effects (similar to one simulation) may be very different from theoretical expectations (based on the logsum formula).

CONCLUSIONS

As the use of logsum differences to estimate welfare impacts of transport policies grows more popular, it is importation to recognize the assumptions inherent in such estimates, and the potential for variations across populations. The paper explores these issues via a Monte Carlo simulation for a numerical example involving 30 zone pairs and tolled and non-tolled route alternatives. In the cases considered here, and with the large number of simulation draws used, we find that logsum differences appear rather robust to deviations in assumptions of correlation, and to addition of a new choice alternative. However, 1 million simulations is generally not enough to estimate the mean CV within 1% of its true value.

What is most striking in these results is the substantial variation that emerges across synthetic populations, suggesting that policies that may appear welfare-improving (when evaluated with average welfare formulations) to be welfare-reducing (or vice versa). Like any point estimate, welfare measures deserve confidence intervals when presented before policymakers, researchers, and the public. There is enough unobserved heterogeneity in existing populations that we may have little sense of which way a particular policy will move a specific group of individuals, even if the average appears to point in one direction. Basic applications of commercially available software programs like those used here can prove invaluable in illuminating the potential for variation in our estimates.

Finally, in terms of future studies, the relatively simple model specification used here can be extended, to accommodate more complex and realistic functional forms, with more variables, more alternatives, and nonlinear-income effects. It also should prove interesting to determine how large a population and set of simulations must be, in order to provide highly accurate estimates or stable values of welfare changes. And mathematically derived solutions to such problems are always desirable.

REFERENCES

Cherchi, E., Polak, J. and Hyman, G. (2005) On the assessment of user benefits using discrete choice models: empirical evidences of the effect of income and random taste heterogeneity. Under consideration for publication in *Transportation*.

Cherchi, E. and Polak, J. (2005) Assessing User Benefits with Discrete Choice Models: Implications of Specification Errors Under Random Taste Heterogeneity, *Transportation Research Record* 1926: 61-69.

de Jong, G., Pieters, M., Daly, A., Graafland, I., Kroes, E., and Koopmans, C. (2005) Using the logsum as an evaluation measure: Literature and case study. Working paper WR-275-AVV, RAND Europe, Leiden. (Available at http://www.rand.org/pubs/working_papers/2005/RAND_WR275.pdf, June 2007)

Dagsvik, J, and Karlstrom, A. (2005) Compensated Variation in Random Utility Models that Are Nonlinear in Income, *Review of Economic Studies,* 72(1)1, 57-76.

FHWA, (2000). Transportation and Environmental Justice Case Studies: MPO Environmental Justice Report Mid-Ohio Regional Planning Commission, Publication No. FHWA-EP-01-010. USDOT. (Available at http://www.fhwa.dot.gov/environment/ejustice/case/morpc.pdf, July 2007)

Franklin, J. (2006) The Equity Effects of Roadway Tolls: An Application of HicksianWelfare Measures with Income Effects. Paper presented at the 11th International Conference on Travel Behaviour Research. Kyoto, Japan. August 16-20. (Available at http://www.urbansim.org/papers/franklin-2006-iatbr.pdf, June 2007)

FTA, (2005). User's Guide to Summit.

Gupta, S., Kalmanje S., and Kockelman, K.M. (2006) Road Pricing Simulations: Traffic, Land Use and Welfare Impacts for Austin, Texas. *Transportation Planning & Technology* 29 (1), pp. 1-23.

Handy, S. L. and Niemeier, D. A., (1997) Measuring accessibility: an exploration of issues and alternatives, *Environment and Planning A* 29(7) pp.1175-1194.

Hensher, D. A. and Greene, W. H. (2003) The Mixed Logit Model: The State of Practice and Warning for the Unwary. *Transportation*, 30(2), pp. 133-176.

Karlström, A., (1998), Hicksian welfare measures in a nonlinear random utility framework, working paper, Department of Infrastructure and Planning, Royal Institute of Technology, Stockholm, Sweden.

Karlström, A. (2001) Welfare evaluations in non-linear random utility models with income effects. Chapter 22 in in Hensher, D. A. (ed.) *Travel Behaviour: The Leading Edge*. Pergamon, Oxford.

Kockelman, K., and Kalmanje, K. (2005) Credit-Based Congestion Pricing: A Proposed Policy and the Public's Response, *Transportation Research* 39A: 671-690.

MacDonald, M. (2006) Transport Users Benefit Appraisal (TUBA) Guidance. Department forTransport, British. (Available at http://www.dft.gov.uk/pgr/economics/software/tuba/, July 2007)

McFadden, D. (1978) Modelling the choice of residential location. *Spatial Interaction Theory and Residential Location*. North-Holland, Amsterdam.

McFadden, D. (1981) Econometric models of probabilistic choice. *Structural Analysis of Discrete Data: With Econometric Applications*. The MIT Press, Cambridge, Massachusetts.

Palisade (2002) @Risk: Simulation Add-In For Microsoft Excel. Windows Version Release 4.5. Newfield, NY: Palisade Corp.

Resource Systems Group (RSG) (2006) SH 121 Southwest Parkway Stated Preference Travel Survey. Working Report, Texas DOT.

Small, K. A. and Rosen, H. S. (1981) Applied welfare economics with discrete choice models, *Econometrica*, 49, pp. 105-130. (Available at http://www.nber.org/papers/w0319.v5.pdf, June 2007)

Small, K.A., Winston, C., and Yan, J. (2006) Differentiated Road Pricing, Express Lanes, and Carpools: Exploiting Heterogeneous Preferences in Policy Design. Brookings-Wharton Papers on Urban Affairs. Brookings Institution, Washington D.C.

Swait, J. (2003) Flexible Covariance Structures for Categorical Dependent Variables Through Finite Mixtures of Generalized Extreme Value Models. *Journal of Business & Economic Statistics* 21 (1), pp. 80-87.

Varian, H. (1992) *Microeconomic Analysis* (Third Edition). New York: Norton & Company.