Which Level of the Travel Demand Hierarchy Should be Used for Benefit Measure in Transport Appraisal?

2007/05/01

Takuya MARUYAMA

Postdoctoral fellow
Japan Society for Promotion of Science

The University of Texas at Austin, 6.9 ECJ, Austin, TX 78712, USA
E-mail: maruyama@ut.t.u-tokyo.ac.jp

Abstract

Transport markets have a unique character wherein demand can be expressed as series of aggregated levels such as generation, OD, mode, route and link. Thus one can have a question, "Which level of aggregation should be used for benefit estimation of transport project?" For this question, this paper proves theoretically that benefit measures for each level of travel demand have same value if consistent travel demand model and price quantities are used. We validate the property in any kind of model that is consistent with random utility theory, a model with deterministic route choice, and a model with any OD-level demand functions. Finally we verify the applicability of this theoretical property empirically with Nested Logit based network model developed by Maruyama *et al*. (2003).

Key Words: *benefit estimation*, *network model*, *OD*, *path*, *link*

*) An earlier version of this paper was published in Japanese as Maruyama (2006).

1. Introduction

Transport market has a unique character that demand quantity can be expressed as several aggregated level or demand hierarchy such as generation, OD, mode, route and link. Thus one can have a question, "Which level of aggregation should be used for benefit estimation of transport project?" Traditionally transport planners have used OD based "Rule-of-Half" (RoH) formula for benefit estimation. However, recent papers by an economist (Kidokoro, 2004, 2006) imply that link/route level measures would be more appropriate measures. In Japan, there is a heated academic-debate on the question "OD vs Route" between economists and engineers. Therefore, this paper aims to clarify the confusion on the issue by both theoretical and empirical analysis.

At first, this paper proves theoretically that benefit measures for each level of travel demand expressed with integral term have same value if consistent travel demand model and price quantities are used. We validate the property in any kind of model that is consistent with random utility theory, a model with deterministic route choice, and a model with any OD-level demand functions. The proof for deterministic route choice uses the division of integral interval. In addition, we prove that link level and route level measures always have the same value in any model with link-additive route cost. Please note that you may think it is natural that each benefit measure has consistent value if consistent model is used, but the original contribution by this paper is to show that even in *link*-level measure has the same value even in the deterministic/stochastic choice situation.

However, we also show that, if we consider the issue practically, such theoretically proof is not applicable in any case. In practical evaluation of user benefit, we use the RoH instead of exact value of integral of demand function. The RoH is based upon linear approximation of demand functions and it is a good approximation of exact value when price changes are small. If we apply the RoH in OD-level, the error of approximation may be small. However, if we use the formula in link/route-level, the error would be larger than that in OD-level.

Finally we verify these properties with a Nested Logit based network model. If we use generation-level RoH, the approximation error is at most 0.03%. OD-level RoH gives 2% error at most. However, error by link/route-level RoH is higher than 60%. We can see that the error is expanded in lower hierarchy.

Through these theoretical and empirical investigations, we have the concrete answer; "As long as the consistent travel demand model and price quantities are used, theoretical benefit measures for each demand hierarchy have same value even in *link* level. However, OD level or higher level demand hierarchy is appropriate measure because the approximation error by RoH for link/route level may be so large."

2. Equivalence of Theoretical Benefit Measures for Each Level of Travel Demand

For easier understanding, we prove the equivalence property in order of following models.

(1) Nested logit model
(2) Any kind of model that is consistent with random utility theory
(3) A model with deterministic route choice
(4) A model with any OD-level demand functions

2.1 Nested Logit model

Let's consider the 4-level nested logit model for user choice of trip-making, destination, mode, and route. In this model, demand quantities can be expressed as following (Maruyama *et al*. 2002, 2003, Oppenheim, 1995).

[link] $$x_a^m = \sum_{rs,m,k} \delta_{a,k}^{m,rs} f_{m,k}^{rs} \tag{1a}$$

[route] $$f_{m,k}^{rs} = \frac{\exp[-\theta_1^m c_{km}^{rs}]}{\sum_k \exp[-\theta_1^m c_{km}^{rs}]} q_m^{rs} \tag{1b}$$

[mode] $$q_m^{rs} = \frac{\exp[-\theta_2^m (C_s^m + S_{rs}^m)]}{\sum_s \exp[-\theta_2^m (C_s^m + S_{rs}^m)]} O_r^m \tag{1c}$$

[destination] $$O_r^m = \frac{\exp[-\theta_3 (C_m^r + S_m^r)]}{\sum_m \exp[-\theta_3 (C_m^r + S_m^r)]} O_r \tag{1d}$$

[generation] $$O_r = \frac{\exp[-\theta_4 (C_r + S_r)]}{1 + \exp[-\theta_4 (C_r + S_r)]} N_r \tag{1e}$$

where monetary-based generalized travel cost of route k on mode m between OD pair rs is defined as

$$c_{km}^{rs} = \sum_{m,a} \delta_{a,k}^{m,rs} t_a^m \; . \tag{2a}$$

Expected minimum cost for each hierarchy are:

$$S_{rs}^m = -\frac{1}{\theta_1^m} \ln \sum_k \exp[-\theta_1^m c_{km}^{rs}] \tag{2b}$$

$$S_m^r = -\frac{1}{\theta_2^m} \ln \sum_s \exp[-\theta_2^m (C_s^m + S_{rs}^m)] \tag{2c}$$

$$S_r = -\frac{1}{\theta_3} \ln \sum_m \exp[-\theta_3 (C_m^r + S_m^r)] \tag{2d}$$

where x_a^m is traffic flow of link a on mode m, t_a^m is travel cost of link a on mode m, $\delta_{a,k}^{m,rs}$ is 1 if link a on mode m is on route k between OD-pair rs and 0 otherwise. We let cost be equal to negative of utility in this paper. $f_{m,k}^{rs}$ is route flow of route k on mode m between OD-pair rs, q_m^{rs} is modal OD flow by mode m between OD-pair rs, O_r^m is generated traffic from zone r by mode m, O_r is generated traffic from zone r, N_r is population index for trip-generation on zone r. θ_1^m, θ_2^m, θ_3, θ_4 are parameters, and C_s^m, C_m^r, C_r are generic dis-utility for each travel hierarchy.

If we define expected maximum utility based on origin zone as

$$\tilde{S}_r = \frac{1}{\theta_4} \ln\{1 + \exp[-\theta_4 (C_r + S_r)]\} \tag{2e}$$

We have following equations.

$$\frac{\partial \tilde{S}_r}{\partial S_r} = -\frac{O_r}{N_r}, \frac{\partial S_r}{\partial S_m^r} = \frac{O_r^m}{O_r}, \frac{\partial S_m^r}{\partial S_{rs}^m} = \frac{q_m^{rs}}{O_r^m}, \frac{\partial S_{rs}^m}{\partial c_{km}^{rs}} = \frac{f_{m,k}^{rs}}{q_m^{rs}} \tag{3}$$

In this paper subscript "w" means "with" and "wo" means "without". Then, the following consumer surplus is the exact measure for user benefit.

$$UB \equiv \sum_r \int_{\tilde{S}_r^{wo}}^{\tilde{S}_r^{w}} N_r \, d\tilde{S}_r = \sum_r N_r (\tilde{S}_r^w - \tilde{S}_r^{wo}) \tag{4a}$$

We define benefit measures for each level of the hierarchy as follows.

$$UB_{GEN} \equiv -\sum_r \int_{S_r^{wo}}^{S_r^{w}} O_r \, dS_r \tag{4b}$$

$$UB_{MODE} \equiv -\sum_{r,m} \int_{S_m^{r,wo}}^{S_m^{r,w}} O_r^m \, dS_m^r \tag{4c}$$

$$UB_{OD} \equiv -\sum_{r,s,m} \int_{S_{rs}^{m,wo}}^{S_{rs}^{m,w}} q_m^{rs}\, dS_{rs}^m \tag{4d}$$

$$UB_{PATH} \equiv -\sum_{r,s,m,k} \int_{c_{km}^{rs,wo}}^{c_{km}^{rs,w}} f_{m,k}^{rs}\, dc_{km}^{rs} \tag{4e}$$

$$UB_{LINK} \equiv -\sum_{m,a} \int_{t_a^{m,wo}}^{t_a^{m,w}} x_a^m\, dt_a^m \tag{4f}$$

Then we have following proposition.

Theorem 1
In nested logit model, theoretical benefit measures for each demand level have equivalent value (equivalent property):

$$UB = UB_{GEN} = UB_{MODE} = UB_{OD} = UB_{PATH} = UB_{LINK} \tag{5}$$

Proof
By eqn. (3), we have following.

$$UB = \sum_r \int_{S_r^{wo}}^{S_r^w} N_r \frac{\partial \tilde{S}_r}{\partial S_r} dS_r = \sum_r \int_{S_r^{wo}}^{S_r^w} N_r \left(-\frac{O_r}{N_r} \right) dS_r = -\sum_r \int_{S_r^{wo}}^{S_r^w} O_r dS_r \left(\equiv UB_{GEN} \right)$$

$$= -\sum_{r,m} \int_{S_m^{r,wo}}^{S_m^{r,w}} O_r \frac{\partial S_r}{\partial S_m^r} dS_m^r = -\sum_{r,m} \int_{S_m^{r,wo}}^{S_m^{r,w}} O_r \frac{O_r^m}{O_r} dS_m^r = -\sum_{r,m} \int_{S_m^{r,wo}}^{S_m^{r,w}} O_r^m dS_m^r \left(\equiv UB_{MODE} \right)$$

$$= -\sum_{m,rs} \int_{S_{rs}^{m,wo}}^{S_{rs}^{m,w}} O_r^m \frac{\partial S_m^r}{\partial S_{rs}^m} dS_{rs}^m = -\sum_{m,rs} \int_{S_{rs}^{m,wo}}^{S_{rs}^{m,w}} O_r^m \frac{q_m^{rs}}{O_r^m} dS_{rs}^m = -\sum_{m,rs} \int_{S_{rs}^{m,wo}}^{S_{rs}^{r,w}} q_m^{rs}\, dS_{rs}^m \left(\equiv UB_{OD} \right)$$

$$= -\sum_{m,rs,k} \int_{c_{km}^{rs,wo}}^{c_{km}^{rs,w}} q_m^{rs} \frac{\partial S_{rs}^m}{\partial c_{km}^{rs}} dc_{km}^{rs} = -\sum_{m,rs,k} \int_{c_{km}^{rs,wo}}^{c_{km}^{rs,w}} q_m^{rs} \frac{f_{m,k}^{rs}}{q_m^{rs}} dc_{km}^{rs} = -\sum_{m,rs,k} \int_{c_{km}^{rs,wo}}^{c_{km}^{rs,w}} f_{m,k}^{rs} dc_{km}^{rs} (\equiv UB_{PATH})$$

$$= -\sum_{m,rs,k} \int_{t_a^{m,wo}}^{t_a^{m,w}} f_{m,k}^{rs} \sum_a \delta_{a,k}^{m,rs} dt_a^m = \sum_{m,rs,k} \int_{t_a^{m,wo}}^{t_a^{m,w}} \sum_a f_{m,k}^{rs} \delta_{a,k}^{m,rs} dt_a^m = -\sum_{m,a} \int_{t_a^{m,wo}}^{t_a^{m,w}} x_a^m dt_a^m (\equiv UB_{LINK})$$

This completes the proof.

One may also notice that

$$UB_{PATH} = UB_{LINK} \tag{6}$$

is satisfied in any model with eqns (1a) and (2a). In other words, route-based and link-based benefit measures have the same value in any travel model with additive route cost.
Please note that above proof process is not valid if any level of nested logit model is deterministic choice. For example, if route choice is not stochastic but deterministic, then $\frac{\partial S_{rs}^m}{\partial c_{km}^{rs}} = \frac{f_{m,k}^{rs}}{q_m^{rs}}$ in eqn. (3) is not valid. However, as shown below, one can prove eqn. (5) in another way.

2.2 Any model that is consistent with random utility theory

Equation (3) is a special case of well-known property that partial differentiation of expected maximum utility by systematic utility is equivalent to choice probability. This property is valid not only in nested logit model but any model that is consistent with random utility theory of discrete choice. So the equivalence property of benefit calculation at any level of the demand hierarchy is expected to hold in any model that is consistent with random utility theory.

Generally, in discrete choice models based on random utility theory, the probability that alternative

i is chosen among choice set C by decision maker is given by

$$\begin{aligned} P(i \mid C) &= \Pr[U_i \geq U_j, \forall j \neq i, j \in C] \\ &= \Pr[V_i + \varepsilon_i \geq V_j + \varepsilon_j, \forall j \neq i, j \in C] \end{aligned} \tag{7}$$

where U_i is the random utility function for alternative i, and V_i, ε_i are deterministic and random components of this utility function, respectively.

The aforementioned property that the derivative of the expected maximum utility with respect to the systematic component of the utility of any alternative is equal to that alternative's choice probability (Williams, 1977) can be described as,

$$\frac{\partial}{\partial V_i} E\left[\max_{i \in C} U_i\right] = P(i \mid C). \tag{8}$$

Then, consider that choice set C is divided into N subsets $\tilde{C}_1, \cdots, \tilde{C}_n, \cdots, \tilde{C}_N$. In the case of 2.1's nested logit model, the subsets are those with identical modes or destinations. Then, the following equation, with marginal probability and conditional probabilities, will hold.

$$P(i \mid C) = P(i \mid \tilde{C}_n) P(\tilde{C}_n \mid C) \tag{9}$$

Here,

$$P(\tilde{C}_n \mid C) = \sum_{i \in \tilde{C}_n} P(i \mid C_n). \tag{10}$$

If we define the conditional expected maximum utility for this subset as

$$\tilde{S}_n = E\left[\max_{i \in \tilde{C}_n} U_i\right] = E\left[\max_{i \in \tilde{C}_n} (V_i + \varepsilon_i)\right] \tag{11}$$

then, we have following formula that is similar to eqn. (8).

$$\frac{\partial \tilde{S}_n(\mathbf{V})}{\partial V_i} = P(i \mid \tilde{C}_n), \quad \forall i \in \tilde{C}_n \tag{12}$$

Also, the total differentiation of the conditional expected maximum utility is

$$d\tilde{S}_n(\mathbf{V}) = \sum_{i \in \tilde{C}_n} \frac{\partial \tilde{S}_n(\mathbf{V})}{\partial V_i} dV_i = \sum_{i \in \tilde{C}_n} P(i \mid \tilde{C}_n) dV_i \tag{13}$$

If we view the choice model as an individual's demand curve, then user benefit is described as

$$UB = \frac{1}{\mu} \sum_{i \in C} \int_{V_i^{wo}}^{V_i^{w}} P(i \mid C) dV_i \tag{14}$$

where μ is marginal utility of income (usually coefficient of travel cost). Then with eqns. (9) and (13), we have

$$\begin{aligned} UB &= \frac{1}{\mu} \sum_{n} \sum_{i \in \tilde{C}_n} \int_{V_i^{wo}}^{V_i^{w}} P(i \mid \tilde{C}_n) P(\tilde{C}_n \mid C) dV_i \\ &= \frac{1}{\mu} \sum_{n} \int_{\tilde{S}_n^{wo}}^{\tilde{S}_n^{w}} P(\tilde{C}_n \mid C) d\tilde{S}_n \end{aligned} \tag{15}$$

where $P(\tilde{C}_n \mid C)$ is demand aggregated at a specific a level, and $\tilde{S}_n / \mu$ is a price index that is consistent with this level. Then, we have proved that benefit measures for demand hierarchy with any aggregation level have equivalent value if the consistent price index is used. This proof is based on individual disaggregate choice, but you can easily understand that the result can be applied to aggregate models if you assume homogeneous individuals. Furthermore you can also see that link-level value has also equivalent value, which is an aggregation of individual behavior.

It depends on the assumption of error term whether we can calculate easily the value $\tilde{S}_n$ defined in eqn. (11). If we use logit/nested logit type model, then $\tilde{S}_n$ can be easily calculated as log-sum variable (e.g. eqn. (2b-d)). Analytical expression of $\tilde{S}_n$ is still possible for GEV type models (McFadden, 1978; Choi and Moon, 1997). On the other hand, Probit model and Mixed Logit model will require numerical calculation with simulations, and then analytical treatment will not be so easy.

2.3 A model with deterministic route choice

Deterministic route choice can alternatively be called as perfect substitution in route-level demand. At first, we assume the upper level choice except for route choice is described by Nested Logit model of eqn (1). You may think that the proof of **2.1**. can be directly applied even in this case if we regard deterministic route choice as a special case of stochastic route choice with $\theta_1^m \to \infty$. However, this is incorrect because stochastic route choice with $\theta_1^m \to \infty$ is merely a special case of general deterministic route choice. To show this let's consider the following example. If there is multiple minimum cost route on network, the deterministic route choice can assign flow in any of the minimum route. This is so-called All or Nothing assignment. On the other hand, stochastic route choice with $\theta_1^m \to \infty$ will assign equal flow on each minimum route. The restriction of assignment by stochastic route choice with $\theta_1^m \to \infty$ is severer than that of deterministic choice. Therefore general deterministic route choice can not always be a special case of stochastic route choice and we need another method of proof for deterministic route choice.

In this case eqn. (1b) is rewritten as

$$c_{km}^{rs} - c_{rs}^{m} \geq 0, f_{mk}^{rs}(c_{km}^{rs} - c_{rs}^{m}) = 0 \tag{16}$$

and eqn. (2b) is rewritten as

$$c_{rs}^{m} = \min_{k}(c_{km}^{rs}) \tag{17}$$

where c_{rs}^{m} minimum cost between OD-pair *rs* on mode *m*. Then OD-based benefit measures are

$$UB_{OD} \equiv -\sum_{r,s,m} \int_{c_{rs}^{m,wo}}^{c_{rs}^{m,w}} q_m^{rs}\, dc_{rs}^{m} \tag{18}$$

With these settings, we can easily verify that the following equations will hold among eqn(5).

$$UB = UB_{GEN} = UB_{MODE} = UB_{OD}, UB_{PATH} = UB_{LINK} \tag{19}$$

Then the last formula we have to prove is $UB_{OD} = UB_{PATH}$.

Theorem 2
In a model with deterministic route choice, theoretical benefit measures for OD level and path level have equivalent value: $UB_{OD} = UB_{PATH}$

Proof
In order to simplify the notation, we omit the notation *rs* and *m* below. We define the choice set of all route on network as K, and the set can be divided into following 4 set.

K_1 : subset of routes that are used in neither with nor without scenario.
K_2: subset of routes that are used in with scenario but are not used in without scenario.
K_3: subset of routes that are used in without scenario but are not used in with scenario.
K_4: subset of routes that are used in both with and without scenario.

Then eqn. (16) is described as,

$$c_k^w - c^w \geq 0, f_k^w = 0, c_k^{wo} - c^{wo} \geq 0, f_k^{wo} = 0 \quad if\ k \in K_1 \tag{20a}$$
$$c_k^w - c^w = 0, f_k^w > 0, c_k^{wo} - c^{wo} \geq 0, f_k^{wo} = 0 \quad if\ k \in K_2 \tag{20b}$$

$$c_k^w - c^w \geq 0, f_k^w = 0, c_k^{wo} - c^{wo} = 0, f_k^{wo} > 0 \quad if\ k \in K_3 \tag{20c}$$
$$c_k^w - c^w = 0, f_k^w > 0, c_k^{wo} - c^{wo} = 0, f_k^{wo} > 0 \quad if\ k \in K_4 \tag{20d}$$

Here, we define

$$UB_{PATH} = \sum_{i=1}^{4} UB_{PATH}(i) \tag{21a}$$

$$UB_{PATH}(i) \equiv -\sum_{k \in K_i} \int_{c_k^{wo}}^{c_k^w} f_k \ dc_k, i = 1,2,3,4 \tag{21b}$$

Then we can divide integral interval as follows.

$$UB_{PATH}(1) = -\sum_{k \in K_1} \int_{c_k^{wo}}^{c_k^w} f_k \ dc_k = 0 \tag{22a}$$

$$\begin{aligned} UB_{PATH}(2) &= \underline{-\sum_{k \in K_2} \int_{c_k^{wo}}^{c^{wo}} f_k \ dc_k} - \sum_{k \in K_2} \int_{c^{wo}}^{c^w} f_k \ dc_k \\ &= -\sum_{k \in K_2} \int_{c^{wo}}^{c^w} f_k \ dc \end{aligned} \tag{22b}$$

$$\begin{aligned} UB_{PATH}(3) &= -\sum_{k \in K_3} \int_{c^{wo}}^{c^w} f_k \ dc_k \underline{- \sum_{k \in K_3} \int_{c^w}^{c_k^w} f_k \ dc_k} \\ &= -\sum_{k \in K_3} \int_{c^{wo}}^{c^w} f_k \ dc \end{aligned} \tag{22c}$$

$$UB_{PATH}(4) = -\sum_{k \in K_4} \int_{c^{wo}}^{c^w} f_k \ dc \tag{22d}$$

Here underlined term is zero because of eqn (20). Therefore, we have

$$\begin{aligned} UB_{PATH} &= \sum_{i=1}^{4} UB_{PATH}(i) = -\int_{c^{wo}}^{c^w} \sum_k f_k \ dc \\ &= -\int_{c^{wo}}^{c^w} q \, dc \ = UB_{OD} \end{aligned} \tag{23}$$

Then, we have proved that route-based benefit measure and OD-based benefit measure have same value even in the deterministic route choice case.

We can easily see that this equivalent property will hold even in the case where upper choice except route choice is described by any model that consistent with random utility theory besides Nested Logit model.

2.4 A model with any OD-level demand functions

Consider the case demand function is set in OD-level. Beckman model is one example (Beckmann *et al*. 1956). Doubly-constrained trip distribution/ assignment model is another example (Sheffi, 1985).

Then, it'll be said to be consistent that the demand function is defined as OD-based minimum travel cost if route choice is deterministic, or the demand function is defined as OD-based expected minimum cost if route choice is stochastic. For simplicity we omit the notation *m* here.

In the case of deterministic route choice, if we define demand function $D_{rs}(.)$ between OD-pair *rs*,

$$q_{rs} = D_{rs}(\ldots, c_{rs}, \ldots) = D_{rs}(\mathbf{c}) \tag{24}$$

then OD-based benefit measure is

$$UB_{OD} \equiv -\sum_{r,s} \int_{c_{rs}^{wo}}^{c_{rs}^{w}} D_{rs}(\mathbf{c})\, dc_{rs} = -\sum_{r,s} \int_{c_{rs}^{wo}}^{c_{rs}^{w}} q_{rs}\, dc_{rs} \tag{25}$$

and we can prove $UB_{OD} = UB_{PATH}$ as **2.3**. In the case of stochastic route choice if we define demand function as

$$q_{rs} = D_{rs}(\ldots, S_{rs}, \ldots) = D_{rs}(\mathbf{S}) \tag{26}$$

then OD-based benefit measure is

$$UB_{OD} \equiv -\sum_{r,s} \int_{S_{rs}^{wo}}^{S_{rs}^{w}} D_{rs}(\mathbf{S})\, dS_{rs} = -\sum_{r,s} \int_{S_{rs}^{wo}}^{S_{rs}^{w}} q_{rs}\, dS_{rs} \tag{27}$$

and we can prove $UB_{OD} = UB_{PATH}$ as **2.2**.

This chapter proves theoretically that benefit measures for each level of travel demand have same value if internally consistent travel demand model and price quantities are used. You may think it is natural that each benefit measure has consistent value if consistent model is used, but the original contribution by this paper is to show that even in *link*-level measure has the same value even in the deterministic/stochastic choice situation. To my knowledge there is no paper which describes this property explicitly before.

3. Practical Method to Calculate Benefit for Each Demand Level

This chapter considers the practical method to calculate benefits for each demand level. For the easier understanding, we consider the Nested Logit model in 2.1.

3.1 Applying Rule-of-Half

Practically the integral-based value in eqn (4) are approximated with following Rule-of-Half (RoH) formula.

$$\widehat{UB}_{GEN} \equiv \frac{1}{2} \sum_{r} (O_r^{w} + O_r^{wo})(S_r^{wo} - S_r^{w}) \tag{28a}$$

$$\widehat{UB}_{MODE} \equiv \frac{1}{2} \sum_{r,m} (O_r^{m,w} + O_r^{m,wo})(S_m^{r,wo} - S_m^{r,w}) \tag{28b}$$

$$\widehat{UB}_{OD} \equiv \frac{1}{2} \sum_{r,s,m} (q_m^{rs,w} + q_m^{rs,wo})(S_{rs}^{m,wo} - S_{rs}^{m,w}) \tag{28c}$$

$$\widehat{UB}_{PATH} \equiv \frac{1}{2} \sum_{r,s,m,k} (f_{m,k}^{rs,w} + f_{m,k}^{rs,wo})(c_{km}^{rs,wo} - c_{km}^{rs,w}) \tag{28d}$$

$$\widehat{UB}_{LINK} \equiv \frac{1}{2} \sum_{m,a} (x_a^{m,w} + x_a^{m,wo})(t_a^{m,wo} - t_a^{m,w}) \tag{28e}$$

These equations will have some approximation error because they are approximation of eqn (4). Before investigating the approximation error, we show the relationship among these equations.
At first, in any model, following equations are hold.

$$\widehat{UB}_{PATH} = \widehat{UB}_{LINK} \tag{29}$$

The proof is given below with omitting the notation *m* here.

$$
\begin{aligned}
\widehat{UB}_{LINK} &\equiv \frac{1}{2}\sum_a (x_a^w + x_a^{wo})(t_a^{wo} - t_a^w) = \frac{1}{2}\sum_a \sum_{rs,k} (\delta_{rs,k}^a f_{rs}^{k,w} + \delta_{rs,k}^a f_{rs}^{k,wo})(t_a^{wo} - t_a^w) \\
&= \frac{1}{2}\sum_a \sum_{rs,k} \delta_{rs,k}^a (f_{rs}^{k,w} + f_{rs}^{k,wo})(t_a^{wo} - t_a^w) = \frac{1}{2}\sum_{rs,k} (f_{rs}^{k,w} + f_{rs}^{k,wo})(\sum_a \delta_{rs,k}^a t_a^{wo} - \sum_a \delta_{rs,k}^a t_a^w) \\
&= \frac{1}{2}\sum_{rs,k} (f_{rs}^{k,w} + f_{rs}^{k,wo})(c_k^{rs,wo} - c_k^{rs,w}) = \widehat{UB}_{PATH}
\end{aligned}
$$

If route choice is deterministic, the following relationship will hold.

$$
\begin{aligned}
\widehat{UB}_{OD} - \widehat{UB}_{PATH} &= \frac{1}{2}(q^{wo} + q^{wo})(c^{wo} - c^w) - \frac{1}{2}\sum_k (f_k^w + f_k^{wo})(c_k^{wo} - c_k^w) \\
&= \frac{1}{2}\sum_k (f_k^w + f_k^{wo})(c^{wo} - c^w) - \frac{1}{2}\sum_k (f_k^w + f_k^{wo})(c_k^{wo} - c_k^w) \\
&= \frac{1}{2}\sum_k \left[(f_k^w + f_k^{wo})(c^{wo} - c^w) - (f_k^w + f_k^{wo})(c_k^{wo} - c_k^w) \right] \\
&= \frac{1}{2}\sum_k \left[(c_k^w - c^w) - (c_k^{wo} - c^{wo}) \right] (f_k^w + f_k^{wo}) \\
&= \frac{1}{2}\sum_k \left[\underline{(c_k^w - c^w) f_k^w} - (c_k^{wo} - c^{wo}) f_k^w + (c_k^w - c^w) f_k^{wo} - \underline{(c_k^{wo} - c^{wo}) f_k^{wo}} \right] \\
&= \frac{1}{2}\sum_k \left[(c_k^w - c^w) f_k^{wo} - (c_k^{wo} - c^{wo}) f_k^w \right] \equiv \frac{1}{2}\sum_k A_k
\end{aligned}
$$

Where underlined term is zero because of deterministic route choice condition eqn.(16). Then with the route choice set (K_1,K_2,K_3,K_4) defined in 2.3, we have following formula.

$$
\begin{aligned}
A_k &\equiv (c_k^w - c^w) f_k^{wo} - (c_k^{wo} - c^{wo}) f_k^w \\
&= \begin{cases} 0 \quad if\ k \in K_1, K_4 \\ -(c_k^{wo} - c^{wo}) f_k^w \le 0 \quad if\ k \in K_2 \\ (c_k^w - c^w) f_k^{wo} \ge 0 \quad if\ k \in K_3 \end{cases}
\end{aligned}
$$

Therefore generally we can conclude $\widehat{UB}_{OD} \neq \widehat{UB}_{PATH}$. In addition, if there are many route flows for choice set K_2 then $\widehat{UB}_{OD} < \widehat{UB}_{PATH}$ will be hold. If there are many route flows for choice set K_3 then $\widehat{UB}_{OD} > \widehat{UB}_{PATH}$ will be hold. Generally the transport projects are intended to improve the traffic situation and number of used route choice alternative will increase, so there will be many route flows for choice set K_2, and $\widehat{UB}_{OD} < \widehat{UB}_{PATH}$ will hold.

Therefore in Deterministic User Equilibrium (DUE) model applying RoH at link/path level doesn't necessarily produces the exact values. However, in the DUE model with fixed demand we can calculate the exact value using another formula based on link-based values. The formula is the difference of total travel cost. The proof is shown below.

Theorem 3
In Deterministic User Equilibrium model with fixed OD-level demand, the user benefit can be calculated as the difference of total travel cost: $\widehat{UB}_{OD} = TC$

Proof
Notation *m* is omitted here. If we assume fixed demand $q_{rs}^w = q_{rs}^{wo}, \forall rs$ in the following,

$$
\widehat{UB}_{OD} \equiv \frac{1}{2}\sum_{r,s} (q_{rs}^w + q_{rs}^{wo})(c_{rs}^{wo} - c_{rs}^w)
$$

we have

$$\widehat{UB}_{OD} = \sum_{r,s} q_{rs}^{wo}(c_{rs}^{wo} - c_{rs}^{w}) = \sum_{r,s} q_{rs}^{wo} c_{rs}^{wo} - \sum_{r,s} q_{rs}^{w} c_{rs}^{w}$$
$$= \sum_{r,s}\left(\sum_k f_{rs,k}^{wo}\right) c_{rs}^{wo} - \sum_{r,s}\left(\sum_k f_{rs,k}^{w}\right) c_{rs}^{w} \tag{30}$$

On the other hand, the total travel cost TC is given by;

$$TC = \sum_a t_a^{wo} x_a^{wo} - \sum_a t_a^{w} x_a^{w}$$
$$= \sum_a t_a^{wo}\left(\sum_{r,s,k} f_{rs,k}^{wo}\delta_{rs,k}^{a,wo}\right) - \sum_a t_a^{w}\left(\sum_{r,s,k} f_{rs,k}^{w}\delta_{rs,k}^{a,w}\right)$$
$$= \sum_{r,s}\sum_k f_{rs,k}^{wo}\left(\sum_a t_a^{wo}\delta_{rs,k}^{a,wo}\right) - \sum_{r,s}\sum_k f_{rs,k}^{w}\left(\sum_a t_a^{w}\delta_{rs,k}^{a,w}\right) \tag{31}$$
$$= \sum_{r,s}\sum_k f_{rs,k}^{wo} c_{rs,k}^{wo} - \sum_{r,s}\sum_k f_{rs,k}^{w} c_{rs,k}^{w}$$

Then eqn (30) – eqn (31) yields.

$$\widehat{UB}_{OD} - TC = \sum_{r,s}\left(\sum_k f_{rs,k}^{wo}\right) c_{rs}^{wo} - \sum_{r,s}\left(\sum_k f_{rs,k}^{w}\right) c_{rs}^{w} -$$
$$\left[\sum_{r,s}\sum_k f_{rs,k}^{wo} c_{rs,k}^{wo} - \sum_{r,s}\sum_k f_{rs,k}^{w} c_{rs,k}^{w}\right] \tag{32}$$
$$= \sum_{r,s}\sum_k f_{rs,k}^{wo}\left(c_{rs}^{wo} - c_{rs,k}^{wo}\right) - \sum_{r,s}\sum_k f_{rs,k}^{w}\left(c_{rs}^{w} - c_{rs,k}^{w}\right)$$

Here we have following user equilibrium conditions.

$$f_{rs,k}^{wo}\left(c_{rs}^{wo} - c_{rs,k}^{wo}\right) = 0 \; and \; f_{rs,k}^{w}\left(c_{rs}^{w} - c_{rs,k}^{w}\right) = 0, \quad \forall r,s,k \tag{33}$$

Thus $\widehat{UB}_{OD} = TC$. This completes the proof.

It can be easily understand that there will be generally no approximation error if we measure the value at the level of fixed demand. In addition, if the demand function is convex downward and project is intended to improve the situation, the RoH will overestimate the exact value. Special attention is needed for logit based demand function. In this case the logit model is S shaped, and the demand function is convex downward and overestimation will occur if choice probability is less than 0.5. However if choice probability is more than 0.5, the demand function will be convex upward and underestimation may occur. If the number of choice set is numerous such as destination choice and route choice situation, every choice probability may less than 0.5 and there will be the overestimation only.

Generally speaking benefit measures based on RoH for each demand level is not equivalent. In the chapter 3, it was the point that we don’t use the RoH but use the integral-based exact value.

3.2 Solution by Numerical Integral

The RoH is known to break down in the following situations.

i) large generalized cost changes

ii) the introduction of new modes

For this problem Nellthorp and Hyman (2001) proposed a numerical integral based method that divides the integral interval and apply the RoH for each divided interval. Explicitly, for an aggregation level i, travel demand q_i, and consistent generalized cost measure c_i, we have following formula.

$$\widehat{UB} = \frac{1}{2}\sum_i \sum_{\omega \in \Omega} (q_i^{\omega} + q_i^{\omega+1})(c_i^{\omega} - c_i^{\omega+1}) \tag{34}$$

where $\omega \in \Omega = \{wo, 1, 2, ..., w-1\}$ is the label which denote the interpolating point for divided integral interval and $w-1$ means nearest interpolating point for *with* scenario.

Using this formula, the integral-based value by eqn (4) can be evaluated approximately. If the number of division is sufficient, the approximation error is expected to be small.
Nellthorp and Hyman (2001) provided the method based on numerical integral only for OD-level evaluation. There is the following problem if we apply the method to path/link level based measure of eqn (4e), (4f).

The problem is how to set the prohibited price if there are several new links. If the number of new link is one, then implement the assignment procedure several times with increasing the cost of new link gradually, and prohibited price is determined as the price which yields the zero traffic flow on the link. Then divide the integral interval between the prohibited price and the price at with scenario. However, if there are several new links, such as construction of new ring road, the checking the prohibited price for each new link is burdensome. In addition, the prohibited price for each new link can not be set uniquely.

If we accept the idea that we calculate one set of prohibited prices for each new link, the following method can be proposed. Beginning with "with" scenario network, implement the assignment several time with adding equally small travel cost on each new link, and record the result of link flow and link cost. If the every new link has zero link flow, then terminate the iteration and regard the link flow and link cost calculated until then as the interpolating point and calculate the value by eqn (34).

We can easily verify that benefit measures for link level and path level have same value even with the numerical integral.

4. Empirical Investigation

This chapter validates the equivalence property using the method shown in chapter 3 with actual data.

We pick up the same project as Maruyama et al. (2002, 2003) used as a case study. The project involves construction of new links that constitute a part of outer ring road for Tokyo Metropolitan Area. Those model parameters and data settings are all used there (Maruyama et al. 2002, 2003). In this large scale application, explicit treatment of path variable is infeasible. However, we have shown that path-level value is same as link-level value in exact integral formula, RoH and numerical integral method, thus we can see the link-level value instead.

We try to investigate the precision of numerical integral by varying the number of integral divisions. Specifically in Numerical Integral (I) we implement the assignment several times with adding the 0, 5, 10, 50 (min) additional cost on the new links. We verify that using 50 (min) additional cost yield zero traffic on each new links. In this case the integral interval is divided up to three. Numerical Integral (II) use the additional cost of 0, 1, 3, 6, 10, 15, 20, 25, 30, 40, 50 (min) and integral interval is divided up to ten.

Table -1 Empirical Comparison of Estimates of Benefit for Each Level of Travel Demand
(Unit: 10^3 Japanese Yen)

Time period	Exact Value Eq.(4a)	Generation level		Mode level		OD level				Link/Path level			
		Rule of Half Eq.(28 a)	Error	Rule of Half Eq.(28b)	Error	Rule of Half Eq.(28c)	Error	Numerical Integral (I) of Eq. (4d)	Error	Numerical Integral (I) of Eq. (4f)	Error	Numerical Integral (II) of Eq. (4f)	Error
	A	B	(B–A)/A	C	(C–A)/A	D	(D–A)/A	E	(E–A)/A	F	(F–A)/A	G	(G–A)/A
0	327	327	0.000%	327	0.000%	328	0.344%	327	0.104%	415	27%	342	4%
1	260	260	0.000%	260	0.000%	261	0.464%	260	0.135%	378	45%	292	12%
2	244	244	0.000%	244	0.000%	246	0.422%	245	0.126%	324	33%	254	4%
3	999	999	0.000%	999	0.000%	1,003	0.404%	1,000	0.115%	1,271	27%	1,072	7%
4	1,526	1,526	0.000%	1,526	0.000%	1,535	0.538%	1,529	0.154%	1,864	22%	1,565	3%
5	5,005	5,005	0.006%	5,005	0.006%	5,027	0.441%	5,013	0.155%	6,052	21%	5,474	9%
6	14,251	14,251	0.002%	14,251	0.002%	14,368	0.820%	14,279	0.197%	21,555	51%	15,103	6%
7	22,766	22,768	0.010%	22,768	0.010%	23,208	1.941%	22,872	0.467%	36,919	62%	23,549	3%
8	20,425	20,431	0.030%	20,431	0.030%	20,717	1.426%	20,486	0.300%	30,999	52%	20,962	3%
9	21,221	21,223	0.011%	21,223	0.012%	21,433	0.999%	21,269	0.228%	32,401	53%	22,572	6%
10	19,795	19,796	0.006%	19,797	0.011%	19,969	0.879%	19,835	0.200%	29,589	49%	21,524	9%
11	20,734	20,736	0.006%	20,736	0.010%	20,884	0.721%	20,769	0.169%	28,389	37%	20,959	1%
12	14,554	14,554	0.003%	14,555	0.006%	14,638	0.578%	14,576	0.149%	19,806	36%	14,986	3%
13	22,569	22,570	0.005%	22,571	0.009%	22,748	0.790%	22,610	0.181%	32,093	42%	23,336	3%
14	19,988	19,989	0.005%	19,990	0.009%	20,135	0.737%	20,022	0.171%	28,559	43%	21,004	5%
15	21,985	21,986	0.004%	21,987	0.007%	22,138	0.693%	22,020	0.159%	29,793	36%	22,112	1%
16	22,681	22,681	0.000%	22,681	-0.003%	22,823	0.626%	22,712	0.137%	31,231	38%	23,097	2%
17	28,405	28,403	-0.007%	28,401	-0.014%	28,566	0.567%	28,439	0.121%	39,246	38%	30,022	6%
18	20,568	20,565	-0.013%	20,564	-0.019%	20,696	0.622%	20,599	0.152%	28,229	37%	22,275	8%
19	8,478	8,479	0.009%	8,479	0.009%	8,508	0.357%	8,494	0.190%	10,633	25%	8,943	5%
20	5,319	5,319	0.007%	5,319	0.007%	5,339	0.383%	5,332	0.239%	6,474	22%	5,468	3%
21	3,058	3,059	0.007%	3,059	0.007%	3,071	0.410%	3,067	0.283%	4,169	36%	3,218	5%
22	1,403	1,403	0.001%	1,403	0.001%	1,408	0.393%	1,407	0.312%	2,085	49%	1,487	6%
23	778	778	0.002%	778	0.002%	780	0.364%	780	0.302%	1,299	67%	831	7%
Day total	297,339	297,353	0.0047%	297,354	0.0049%	299,828	0.837%	297,943	0.203%	423,772	43%	297,355	4%

Note 1) Value for column A and D is same as Table 8 in Maruyama et al. (2003)

Note 2) Numerical Integral (I) divide integral interval up to three and Numerical Integral (II) divide integral interval up to ten. Please refer to text for details.

The results for each time period are shown in Table 1. Compared to the exact formula, approximation error of the RoH formula for generation/ mode level is up to 0.03% and this is extremely small. The approximation error of RoH formula for OD level is up to 2% and this is sufficiently small. The error is made much smaller if Numerical Integral (I) is applied. In addition, as mentioned earlier the error by RoH is roughly overestimation in each time period. The error tends to be large in the crowded time period. In addition, the error tends to increase as the demand hierarchy goes down.

The value by link/path level has big error of 20~60% even if Numerical Integral (I) is used. If Numerical Integral (II) is used, the error decreases but it is not within practical range yet. In off-peak time period (0~5 a.m.), we recalculated numerical integral by dividing integral interval by 1 min, and the error is around -0.1% ~ 2.91%. This procedure need numerous computing times. The RoH for link/path level is not calculated because the calculating prohibited price is burdensome but the value will have larger error than the value by Numerical Integral (I). Thus RoH for link/path level won't be practical.

In order to investigate the large error in link level benefit measure, we plot relationship between link flow and link cost in a link that come from the procedure with integral interval of 0.5 (min) in Figure 2. When the additional cost is zero, the link cost is about 3.5 (min) and link flow is about 365(vehicle). Additional cost increases by 0.5(min) and plot the relevant link flow. If the additional

cost exceed 11.5(min), the link flow is zero. So the prohibited price for this link is around 15 (3.5+11.5) (min). Figure 2 shows this procedure.
This Figure 2 is figure of the general equilibrium demand curve for this link. To author's knowledge this is the first to figure the link-based general equilibrium demand curve. From this figure the general equilibrium demand curve for this link is far from line and roughly convex curve.
If we approximate the curve with line between *with* and *without* scenario and estimate the benefit by triangular area, the estimated benefit will be overestimated by more than two times. Even if we use numerical integral, there will be a lot of approximation error without sufficient integral divisions.

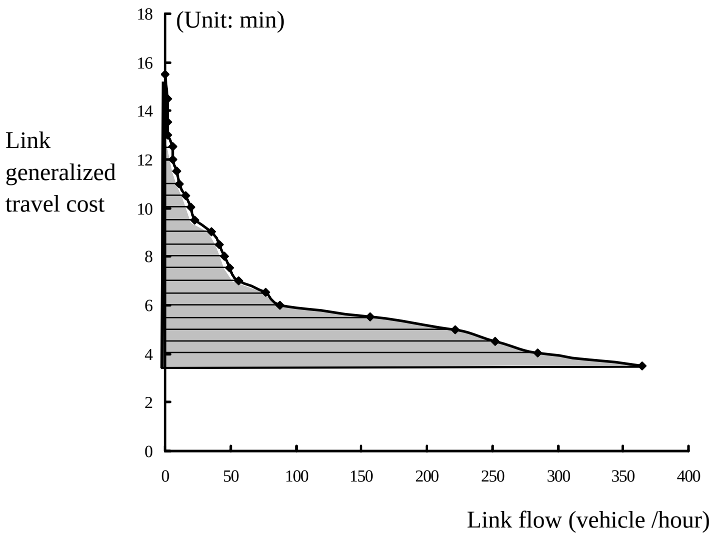

Figure 2 An example of Estimate of Link-based General Equilibrium Demand Curve and User Benefit by Numerical Integral

Last chapter verify the equivalence of benefit measures for each level of travel demand theoretically. On the other hand, this chapter verifies the equivalence property in practical situations. In this calculation, benefit measures for generation/ mode/OD level are very close to the exact value even using the RoH formula. However, if we apply the RoH formula to link level, the approximation error will be extremely large, and this error can be reduced in numerical integral method only if the sufficient interpolating points are set as integral divisions. The implication that the approximation error is large in applying RoH in link/path level may be specific to the setting of this calculation which considers congestion in metropolitan level. However, because the number of such project is large, estimation of link/path level is not recommended at least in such cases.

5. Discussion

The question raised by this research is which is recommended to be used in benefit estimation, link/path or OD level. For this question theoretically an answer can be made that any level will OK, because the benefit measure is equivalent in internal consistent travel demand model. However, theoretical value is based on integral and this value is not always easy to be calculated. Therefore, an approximation is needed for the integral value for a hierarchy. Such kind of discussion has never made in existing literature.

Generally transport project are described as change of generalized cost in link-level price. If the internal consistent travel demand model is given, a change of generalized cost in OD level that is consistent with link-level change can be defined. These change rates of generalized cost can be very large in link-level measures but it can be made smaller in OD-level measures because of aggregation. Generally the RoH is based on the liner approximation of integral, thus the approximation error will be made large if the change of generalized cost is large. Therefore approximation error of link-level RoH is large than that of OD-level RoH. This point is strongly supported by the calculation shown in last chapter. As you can see in Figure 2, the approximation error of general equilibrium demand curve in link-level will be extremely large.

6. Conclusion

This paper shows the following point.
(1) Travel demand can be expressed in various aggregation levels such as generation/mode/OD/link and benefit measures for each level are equivalent in theoretical integral based value if the internal consistent demand forecasting model and consistent price index are used. We validate the property in a model that is consistent with random utility theory (e.g. Nested Logit), a model with deterministic route choice, and a model with any OD-level demand functions.
(2) Path-based benefit measures and link-based benefit measures have the same value in any model with additive route cost. This hold not only in exact value expressed in the integral but also in Rule of Half (RoH) formula.
(3) In deterministic route choice context, we show RoH for OD level is different from that for link/path level. In addition we clarify the tendency of the difference property.
(4) The equivalence property is validated in an empirical situation that develops Nested Logit model in large metropolitan area and shown that even RoH formula has very close value to the exact value in generation/mode/OD level. However, applying the RoH to link/path level yields very large approximation error.
(5) Practically the OD-level estimation is more recommended than the link/path level estimation because the approximation error of the former is smaller than the latter.

Acknowledgement
The author thanks Shin-ichi Inoue, Yoshitsugu Kanemoto, Yukihiro Kidokoro, Takayuki Ueda and Kara Kockelman for their helpful suggestions and discussions for earlier versions of this paper. However, and as usual, all remaining errors are my sole responsibility.

Reference

Beckmann, M., McGuire, C. B. and Winsten, C. B.(1956) Studies in the Economics of Transportation, Cowles Commission Monograph, Yale University Press, New Haven, CN,.
Choi, K.-H. and Moon, C.-G. (1997) Generalized extreme value model and additively separable generator function, *Journal of Econometrics*, Vol. 76, No. 1/2, pp. 129–140.
Maruyama, T., Harata, N., and Ohta, K. (2002) Estimation of user benefits of road investment considering induced traffic with combined network equilibrium model in Tokyo area, *Proceedings of European Transport Conference* 2002.
Maruyama, T., Harata, N., and Ohta, K. (2003) Estimation of user benefits of road investment considering induced traffic in congested area, JSCE, *Journal of Infrastructure Planning and Management*, No. 744/IV-61, pp. 123-137 (in Japanese).
Maruyama, T. (2006) Equivalence of benefit measures for each level of travel demand, *Doboku Gakkai Ronbunshuu D* [JSCE, Journal of Infrastructure Planning and Management], Vol. 62, No. 3, pp.460-473 (in Japanese). (http://www.jstage.jst.go.jp/article/jscejd/62/3/62_460/_article)
McFadden, D. (1981) Econometric models of probabilistic choice, in Manski, C.F. and McFadden, D. (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 198-272, MIT Press: Cambridge.

Nellthorp, J. and Hyman, G. (2001) Alternatives to the rule of a half in matrix-based appraisal, *Proceedings of European Transport Conference*.
Kidokoro, Y. (2004) Cost-benefit analysis for transport networks: Theory and application, *Journal of Transport Economics and Policy*, Vol. 38, No. 2, pp. 275 – 307.
Kidokoro, Y. (2006) Benefit estimation of transport projects –A representative consumer approach–, *Transportation Research Part B*, Vol. 40, No. 7, pp.521-542.
Oppenheim, N. (1995) *Urban Travel Demand Modeling: From Individual Choices to General Equilibrium*, John Wiley & Sons, N. Y.
Small, K. A. and Rosen, H. S. (1981) Applied welfare economics with discrete choice models, *Econometrica*, Vol. 49, No. 1, pp. 105-130.
Sheffi, Y.(1985) *Urban Transportation Networks: Equilibrium analysis with mathematical programming methods*, Prentice-Hall, Englewood Cliffs, N. J.
Williams, H.C.W.L. (1977) On the formulation of travel demand models and economic evaluation measures of user benefit, *Environment and Planning A*, Vol. 9, No. 3, pp. 285-344.