Sensitivity Analysis and Optimal Pricing on a Bi-modal Network with Scale Economies*

Jiang Qian YING †

1 Introduction

Sensitivity analysis methods for transport system having a mobile road network and a physically separate transit network is studied. The transit network simply consists of independent lines connecting Origin-Destination pairs and may have economies of scale with the increase of the passengers, conditions under which sensitivity analysis can be properly conducted are investigated. The sensitivity analysis algorithm is applied to the optimal pricing problem (see, e.g., Miyagi and Suzuki, 1996) in a combined network with transit line exhibiting economies of scale.

2 Equilibrium on the mobile network and transit lines

Notations:

- $W = \{rs, \cdots\}$: set of OD pairs
- $\bar{q}_{rs}$: fixed OD demand
- $\hat{q}_{rs}$: demand for using the transit lines
- q_{rs}: demand for using the road network
- p_{rs}: parameter characterizing the performance of the transit line; Typical p_{rs} can be a tax or subsidy for the use of transit
- $\hat{S}_{rs}$: disutility of using the transit lines
- $N = \{i, j, \cdots\}$: set of nodes of the mobile network
- $A = \{ij, \cdots\}$: set of links of the mobile network
- $R_{rs} = \{k, p, \cdots\}$: set of paths connecting rs in the mobile network
- S_{rs}: the expected disutility for OD rs on the road network
- $\bar{S}_{rs}$: the expected disutility for using the transport system
- x_{ij}: link flow, for $ij \in A$
- $t_{ij}(x_{ij}, \epsilon_{ij})$: differentiable cost function of link ij with respect to flow x_{ij}, and parameter ϵ_{ij} (which will be the road toll in the optimal pricing problem).
  It is assumed that t_{ij} is strongly monotone increasing with respect to x_{ij}. For given ϵ_{ij}, the inverse of the cost function is denoted as $x_{ij}(t_{ij}, \epsilon_{ij})$, which is also strongly monotone increasing in t_{ij}
- $(x_{ij})_{t_{ij}}$: partial derivative of x_{ij} with respect to t_{ij}
- $(x_{ij})_{\epsilon_{ij}}$: partial derivative of x_{ij} with respect to ϵ_{ij}
- $\delta^{rs}_{ij,k} = \begin{cases} 1 & \text{if } ij \text{ is a link on path } k; \\ 0 & \text{otherwise.} \end{cases}$
- $\delta^{rs}_{ij,gh} = \begin{cases} 1 & \text{if } ij = gh \in A; \\ 0 & \text{otherwise.} \end{cases}$
- $c^{rs}_k = \sum_{ij \in A} t_{ij} \delta^{rs}_{ij,k}$: the total cost of traveling on a path $k \in R_{rs}$
- Bold letters denote vectors: $\boldsymbol{q}=(q_{rs})_{rs\in W}$, $\boldsymbol{p}=(p_{rs})_{rs\in W}$, $\boldsymbol{x}=(x_{ij})_{ij\in A}$, $\boldsymbol{t}=(t_{ij})_{ij\in A}$, $\boldsymbol{\epsilon}=(\epsilon_{ij})_{ij\in A}$, and so on

We assume that the costs t_{ij} and c^{rs}_k are a kind of generalized costs containing prices, time, and other factors, measured by monetary term. The road network is said to be "congested" in the sense that t_{ij} is strongly monotone increasing with respect to x_{ij}, as is assumed in this paper. In a multinominal logit-based stochastic user equilibrium (SUE), the link flows are

$$x_{ij} = \sum_{rs} q_{rs} \frac{\sum_k \exp(-\theta c^{rs}_k)\delta^{rs}_{ij,k}}{\sum_p \exp(-\theta c^{rs}_p)}, \quad ij \in A. \qquad (1)$$

The disutility S_{rs} for using the road network is

$$S_{rs}(\boldsymbol{c}^{rs}(\boldsymbol{t})) = -\frac{1}{\theta} \ln \sum_k \exp(-\theta c^{rs}_k). \qquad (2)$$

*Keywords: sensitivity analysis, congestion, scale economies, pricing

†Member of JSCE, Associate Professor, Faculty of Regional Studies, Gifu University (Gifu-Shi, 501, Japan. TEL 058-293-3306 FAX 058-293-3118 Email ying@cc.gifu-u.ac.jp)

2.1 Equilibrium Equation

The disutility of using transit $\hat{S}_{rs}$ is a function in p_{rs} and the transit demand $\hat{q}_{rs} = \bar{q}_{rs} - q_{rs}$:

$$\hat{S}_{rs} = \hat{S}_{rs}(p_{rs}, \bar{q}_{rs} - q_{rs}).$$

At equilibrium,

$$q_{rs} = \bar{q}_{rs} \frac{1}{1 + \exp(\alpha(S_{rs} - \hat{S}_{rs}(p_{rs}, \bar{q}_{rs} - q_{rs})))}. \tag{3}$$

Solving this equation for q_{rs}, we can obtain the mobile travel demand in the form

$$q_{rs} = q_{rs}(S_{rs}, p_{rs}). \tag{4}$$

The equilibrium conditions are characterized by the following equations

$$\begin{aligned} & F_{ij}(\boldsymbol{t}, \boldsymbol{\epsilon}, \boldsymbol{p}) \\ = \; & x_{ij}(t_{ij}, \epsilon_{ij}) - \sum_{rs} q_{rs} \frac{\partial S_{rs}}{\partial t_{ij}} \\ = \; & x_{ij}(t_{ij}, \epsilon_{ij}) - \sum_{rs} q_{rs} \frac{\sum_k \exp(-\theta c_k^{rs}) \delta_{ij,k}^{rs}}{\sum_p \exp(-\theta c_p^{rs})} \\ = \; & 0, \quad ij \in A, \end{aligned} \tag{5}$$

or in a vector form

$$\boldsymbol{F}(\boldsymbol{t}, \boldsymbol{\epsilon}, \boldsymbol{p}) = \boldsymbol{0}. \tag{6}$$

2.2 Properties of $\nabla_{\boldsymbol{t}} \boldsymbol{F}$

From (6) is derived the following equation

$$\nabla_{\boldsymbol{t}} \boldsymbol{F} = \mathrm{diag}((x_{ij})_{t_{ij}})_{ij} - \sum_{rs} q_{rs} \nabla_{\boldsymbol{t}}^2 S_{rs}(\boldsymbol{c}^{rs}(\boldsymbol{t}))$$
$$- \sum_{rs} \frac{\partial q_{rs}}{\partial S_{rs}} \left(\nabla_{\boldsymbol{t}} S_{rs}(\boldsymbol{c}^{rs}(\boldsymbol{t}))\right) \left(\nabla_{\boldsymbol{t}} S_{rs}(\boldsymbol{c}^{rs}(\boldsymbol{t}))\right)^T. \tag{7}$$

It can be shown that the first term is positive definite, the second is positive semidefinite. If

$$\frac{\partial q_{rs}}{\partial S_{rs}} < 0, \tag{8}$$

then $\nabla_{\boldsymbol{t}} \boldsymbol{F}$ is a positive definite matrix. If the transit lines are congetsed, $\frac{\partial \hat{S}_{rs}}{\partial \hat{q}_{rs}} > 0$, it can be shown $\frac{\partial q_{rs}}{\partial S_{rs}} < 0$. However, $\frac{\partial q_{rs}}{\partial S_{rs}} < 0$ may fail to hold when the economies of scale prevail. Actually from (3) we have

$$\frac{\partial q_{rs}}{\partial S_{rs}} = - \left[1 - \bar{q}_{rs} \frac{\alpha \exp(\alpha(S_{rs} - \hat{S}_{rs}))}{[1 + \exp(\alpha(S_{rs} - \hat{S}_{rs}))]^2} \frac{\partial \hat{S}_{rs}}{\partial q_{rs}} \right]^{-1}$$
$$\cdot \bar{q}_{rs} \frac{\alpha \exp(\alpha(S_{rs} - \hat{S}_{rs}))}{[1 + \exp(\alpha(S_{rs} - \hat{S}_{rs}))]^2}. \tag{9}$$

$\frac{\partial q_{rs}}{\partial S_{rs}} < 0$ holds if and only if

$$1 - \bar{q}_{rs} \frac{\alpha \exp(\alpha(S_{rs} - \hat{S}_{rs}))}{[1 + \exp(\alpha(S_{rs} - \hat{S}_{rs}))]^2} \frac{\partial \hat{S}_{rs}}{\partial q_{rs}} > 0. \tag{10}$$

In the case that economies of scale prevail in the transit line, we have

$$\frac{\partial \hat{S}_{rs}}{\partial \hat{q}_{rs}} < 0 \quad \text{or} \quad \frac{\partial \hat{S}_{rs}}{\partial q_{rs}} > 0.$$

(10) holds only when the quantity $\frac{\partial \hat{S}_{rs}}{\partial q_{rs}}$ is not too large.

In general, there seems to be no meaningful economic or behavioral explanation of (10). However, for some special class of transit performance functions, implication of (10) can be further explored. Suppose the transit users incur a cost of the following form

$$\hat{C}_{rs} = F/\hat{q}_{rs} + a\hat{q}_{rs} + c. \tag{11}$$

F, a, c are assumed to be positive coefficients. The first term reflects the scale economies of the transit system; the second term reflects a congestion effect when $\hat{q}_{rs}$ grows large.

As it is assumed that the transit has only one line connecting an OD pair rs, the disutility perceived by the transit user is equal to the cost

$$\hat{S}_{rs} = \hat{C}_{rs} = F/\hat{q}_{rs} + a\hat{q}_{rs} + c,$$

(10) is reduced to

$$\frac{\hat{q}_{rs}}{q_{rs}} + a\alpha \frac{\hat{q}_{rs}^2}{\bar{q}_{rs}} > \frac{\alpha F}{\bar{q}_{rs}}.$$

This inequality is satisfied if

$$\frac{\hat{q}_{rs}}{q_{rs}} > \frac{\alpha F}{\bar{q}_{rs}},$$

that is, if the ratio of the transit users to the car drivers is greater than a constant.

When $\boldsymbol{\epsilon}$ and $\boldsymbol{p}$ are given, if $\nabla_{\boldsymbol{t}} \boldsymbol{F}$ is always positive definite, then the solution $\boldsymbol{t}$ for the equation $\boldsymbol{F}(\boldsymbol{t}, \boldsymbol{\epsilon}, \boldsymbol{p}) = \boldsymbol{0}$ are unique, i.e., there is a unique equilibrium.

In the case that $\nabla_{\boldsymbol{t}} \boldsymbol{F}$ is reguler but is not positive definite, there may be multiple solutions for the above equation. In this case, suppose that a solution is specified that describes an equilibrium state, local unique solution still exists around that state, sensitivity analysis can be undertaken by the method developed in Section 3.

There is two cases for which $\nabla_{\boldsymbol{t}}\boldsymbol{F}$ is not regular. The first is the case

$$1-\bar{q}_{rs}\frac{\alpha\exp(\alpha(S_{rs}-\hat{S}_{rs}))}{[1+\exp(\alpha(S_{rs}-\hat{S}_{rs}))]^2}\frac{\partial\hat{S}_{rs}}{\partial q_{rs}}=0.$$

In this case,

$$\frac{\partial q_{rs}(S_{rs},p_{rs})}{\partial S_{rs}}=\infty,$$

$\nabla_{\boldsymbol{t}}\boldsymbol{F}$ is not well-defined. $\frac{\partial q_{rs}}{\partial S_{rs}}=\infty$ occurs at the point in the S_{rs}-q_{rs} plane where the tangent vector of the curve of the multivalued function $q_{rs}(S_{rs},p_{rs})$ (p_{rs} being fixed) turns downward. The second case is that $\nabla_{\boldsymbol{t}}\boldsymbol{F}$ is well-defined but is singular. For both cases, before sensitivity analysis could be discussed, the properties of the equilibrium equations have to be investigated in depth, which will be left for future research.

3 Sensitivity Analysis Method

In the following, with the assumption that the inverse matrix $(\nabla_{\boldsymbol{t}}\boldsymbol{F})^{-1}$ exists, we establish the formulae for computing the derivatives of $\boldsymbol{t}$, $\boldsymbol{x}$ and $\boldsymbol{q}$ with respect to $\boldsymbol{\epsilon}$ and $\boldsymbol{p}$.

3.1 Computing $\frac{\partial\boldsymbol{t}}{\partial\boldsymbol{\epsilon}}$, $\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{\epsilon}}$ and $\frac{\partial\boldsymbol{q}}{\partial\boldsymbol{\epsilon}}$

In the equilibrium equation

$$\boldsymbol{F}(\boldsymbol{t},\boldsymbol{\epsilon},\boldsymbol{p})=\boldsymbol{0}, \tag{12}$$

looking $\boldsymbol{x}$ as mediate (apparent) variables in $\boldsymbol{F}$, $\boldsymbol{p}$ as constants, $\boldsymbol{\epsilon}$ as free variables and and $\boldsymbol{t}$ as functions of $\boldsymbol{\epsilon}$, we have

$$(\boldsymbol{F})_{\boldsymbol{x}}(\boldsymbol{x})_{\boldsymbol{\epsilon}}+\nabla_{\boldsymbol{t}}\boldsymbol{F}(\frac{\partial\boldsymbol{t}}{\partial\boldsymbol{\epsilon}})=\boldsymbol{0}.$$

Once $(\nabla_{\boldsymbol{t}}\boldsymbol{F})^{-1}$ has been computed, the derivatives can be easily obtained as follows.

$$\begin{aligned}(\frac{\partial\boldsymbol{t}}{\partial\boldsymbol{\epsilon}})&=-(\nabla_{\boldsymbol{t}}\boldsymbol{F})^{-1}(\boldsymbol{F})_{\boldsymbol{x}}(\boldsymbol{x})_{\boldsymbol{\epsilon}}\\&=-(\nabla_{\boldsymbol{t}}\boldsymbol{F})^{-1}\mathrm{diag}((x_{ij})_{\epsilon_{ij}})_{ij},\end{aligned} \tag{13}$$

$$(\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{\epsilon}})=\left((x_{ij})_{\epsilon_{ij}}\delta_{ij,gh}+(x_{ij})_{t_{ij}}\frac{\partial t_{ij}}{\partial\epsilon_{gh}}\right)_{ij,gh}, \tag{14}$$

and

$$(\frac{\partial\boldsymbol{q}_{rs}}{\partial\boldsymbol{\epsilon}_{ij}})=\left(\frac{\partial q_{rs}}{\partial S_{rs}}\sum_{gh}\frac{\partial S_{rs}}{\partial t_{gh}}\frac{\partial t_{gh}}{\partial\epsilon_{ij}}\right)_{rs,ij}. \tag{15}$$

3.2 Computing $\frac{\partial\boldsymbol{t}}{\partial\boldsymbol{p}}$, $\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{p}}$ and $\frac{\partial\boldsymbol{q}}{\partial\boldsymbol{p}}$

Looking $\boldsymbol{q}$ as mediate (apparent) variables in $\boldsymbol{F}$, $\boldsymbol{\epsilon}$ as constants, $\boldsymbol{p}$ as free variables and and $\boldsymbol{t}$ as functions of $\boldsymbol{p}$, the following formulae are derived.

$$(\boldsymbol{F})_{\boldsymbol{q}}(\boldsymbol{q})_{\boldsymbol{p}}+\nabla_{\boldsymbol{t}}\boldsymbol{F}(\frac{\partial\boldsymbol{t}}{\partial\boldsymbol{p}})=\boldsymbol{0},$$

$$(\frac{\partial\boldsymbol{t}}{\partial\boldsymbol{p}})=-(\nabla_{\boldsymbol{t}}\boldsymbol{F})^{-1}(\boldsymbol{F})_{\boldsymbol{q}}(\boldsymbol{q})_{\boldsymbol{p}}, \tag{16}$$

where

$$(\boldsymbol{F})_{\boldsymbol{q}}=\left(-\frac{\sum_k\exp(-\theta c_k^{rs})\delta_{ij,k}^{rs}}{\sum_p\exp(-\theta c_p^{rs})}\right)_{ij,rs}$$

and

$$(\boldsymbol{q})_{\boldsymbol{p}}=\ \mathrm{diag}\ \left(\frac{\partial q_{rs}}{\partial p_{rs}}\right)_{rs}.$$

Once $\frac{\partial\boldsymbol{t}}{\partial\boldsymbol{p}}$ is known, we can immediately compute

$$(\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{p}})=\left((x_{ij})_{t_{ij}}\frac{\partial t_{ij}}{\partial p_{rs}}\right)_{ij,rs}, \tag{17}$$

and

$$(\frac{\partial\boldsymbol{q}}{\partial\boldsymbol{p}})=\left(\frac{\partial q_{rs}}{\partial p_{uv}}+\frac{\partial q_{rs}}{\partial S_{rs}}\sum_{gh}\frac{\partial S_{rs}}{\partial t_{gh}}\frac{\partial t_{gh}}{\partial p_{uv}}\right)_{rs,uv}. \tag{18}$$

The derivatives can be computed in a link based fashion (Ying and Miyagi, 2001).

4 Application to Optimal Pricing

The disutility of using in a random manner the combined transportation system consisting of two modes is given by

$$\bar{S}_{rs}=-\frac{1}{\alpha}\ln(\exp(-\alpha S_{rs})+\exp(-\alpha\hat{S}_{rs})). \tag{19}$$

$1/\alpha$ reflects the degree of traveler's preference for variety of modal choice. $-\bar{S}$ is an economically sound utility measure in that its variation is equivalent to that of a generalized Marshallian surplus. (See Williams, 1977).

Pricing Scheme

Let T_{ij} denote the toll imposed on link ij in the mobile network. The road link cost is

$$c_{ij}=t_{ij}+T_{ij}.$$

The disutility of using the mobile network is

$$S_{rs} = -\frac{1}{\theta} \ln(\sum_k \exp(-\theta \sum_{ij} c_{ij}\delta^{rs}_{k,ij})).$$

Let Tax_{rs} denote the tax collected from the transit user. $Tax_{rs} < 0$ means the transit users are actually subsidized. Thus the transit user incurs the following disutility

$$\hat{S}_{rs} = \hat{C}_{rs} + Tax_{rs}.$$

Social Utility

Suppose that the road tolls and transit taxes collected are returned to the traveling community by certain means, a social utility is obtained as

$$SU = -\sum_{rs} \bar{q}_{rs} \cdot \bar{S}_{rs} + \sum_{ij} x_{ij} \cdot T_{ij} + \sum_{rs} \hat{q}_{rs} \cdot Tax_{rs}. \tag{20}$$

It is easily shown that as

$$\theta \longrightarrow \infty \quad \text{and } \alpha \longrightarrow \infty,$$

$$-SU \longrightarrow (\sum_{ij} x_{ij} \cdot t_{ij} + \sum_{rs} \hat{q}_{rs} \cdot \hat{C}_{rs}).$$

This means that when the travelers choose exactly the least cost mode and route, the social utility is equal to the negative of the total transportation cost.

Sensitivity Analysis Based Optimization

For consistence of notations, let

$$\epsilon_{ij} = T_{ij}, \quad p_{rs} = Tax_{rs}.$$

Possible optimal solutions that maximize SU can be obtained as the points where the gradient

$$\nabla_{\boldsymbol{T}} SU = (\cdots, \frac{\partial SU}{\partial p_{rs}}, \cdots, \frac{\partial SU}{\partial \epsilon_{ij}}, \cdots) = \mathbf{0},$$

where $\boldsymbol{T} = (\cdots, p_{rs}, \cdots, \epsilon_{ij}, \cdots)$ denotes the vector of transit tax and mobile link tolls. It is routine to show that $\nabla_{\boldsymbol{T}} SU$ can be computed by the sensitivity analysis method of Section 3. Conventional optimization method, e.g., the steepest decent method, can then be applied to find a pricing scheme that optimizes the social utility.

Computation Experience

The sensitivity analysis optimization method has been implemented on a road network consisting of 12 nodes and 17 links, each link has a BPR performance function. It is assumed that there is a single OD pair and the transit line has the cost function described in Section 2.2. In the example it was shown that the gradient of the social utility with respect to the road tolls and transit tax are 0 when a *marginal cost pricing* scheme is imposed. And, starting from 0 tolls and 0 tax, by applying our sensitivity algorithm and the steepest gradient method, an optimal solution (pricing pattern) could be obtained which is different from the MCP scheme but yields the same transit and network flows and the same optimal social utility.

5 Concluding Remarks

We conculde this paper with some related themes worth further studying.

i. A theoretical study of the properties of equilibria in general multimodal transport system where each mode consists of a general network which may have economies of scale and external diseconomies caused by congestion remains to be conducted. Development of related computational techniques is also an important research topic.

ii. Theoretical and practical aspects of sensitivity analysis methods for such a general combined system need to be investigated.

iii. Dynamic evolution properties of such a system is also interesting.

Acknowledgement

I am grateful to Professor T. Miyagi for helpful discussions regarding the utility and pricing problems.

References

Miyagi, T. and Suzuki, T. (1996) "A Ramsey Price Equilibrium Model for Urban Transit Systems: A Bilevel Programming Approach with Transportation Network Equilibrium Constraints", *Proceedings of the 7th World Conference on Transport Research*, volume **2**, 65-78.

Williams, H. C. W. L. (1977) "On the Formation of Travel Demand Models and Economics Measures of User Benefit," *Environment and Planning* **A 9**, 285-344.

Ying, J. Q. and Miyagi, T. (2001) "Sensitivity Analysis for Stochastic User Equilibrium Network Flows - A Dual Approach," *Transportation Science*, **35**, 124-133.