

A congestion-pricing problem with a polycentric region and multi-class users: A continuum modeling approach

H.W. Ho^a, S.C. Wong^b and A. Sumalee^{a*}

^a*Department of Civil & Structural Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, P.R. China.*

^b*Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, P.R. China.*

(Received XXX)

Consider a region of arbitrary shape with multiple cities competing for multi-class users that are distributed continuously over the region. Within this region, the road network is represented as a continuum and users patronize in a two-dimensional continuum transportation system to their chosen city. A logit-type distribution function is specified to model the probabilistic destination choices made by the different classes of users. In this paper, two different congestion-pricing models for this multi-class and multi-city continuum transportation system are studied. The first model focused on utility maximization, which determines the optimal toll rates that maximize the total utility of the system, while the second model is a cordon-based congestion-pricing model that offers a sub-optimal but more practical tolling strategy. Both models are solved by finite element method and a promising Newtonian-based solution algorithm. A numerical example is given to show the effectiveness of the mathematical program and solution algorithm.

Keywords: continuum transportation system; multiple user classes; national road pricing; traffic equilibrium; logit-type distribution function; finite element method

1. Introduction

Congestion pricing can be used to reduce traffic congestion and raise revenue for the funding of transportation improvements. The principle underlying congestion pricing is that the choice behavior of users should be regulated with a view to approaching a system-optimal travel pattern which maximizes the total benefit of the system as a whole by imposing tolls on users. This is known as the first-best pricing principle and can be implemented by introducing a toll that is equal to the user externality, or the difference between the marginal social cost and the marginal private cost, for each link, such that the user-optimal network traffic flow condition maximizes the total benefit of the system as a whole.

Owing to the advantages of congestion pricing and the need to regulate traffic at the planning/policy level, government agencies, practitioners and academia are always interested in formulating national congestion-pricing schemes, which will affect the entire road system of the region/country concerned, in a way designed to optimize the performance of the transportation system (in terms of yardsticks such as travel delay or pollution). The idea of

* Correspondence to: A. Sumalee, Department of Civil & Structural Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, P.R. China. E-mail: ceasumal@polyu.edu.hk

1 national road user charging, or region-wide road user charges, has been studied in a number
2 of countries such as the United Kingdom (Department for Transport, 2004; Glaister and
3 Graham, 2005; Cottingham et al., 2007) and, in a European context, in relation to the
4 Eurovignette system (Commission of the European Community, 1999).

5 In the literature, the applications of congestion-pricing models have been focused on the
6 discrete network (Beckmann, 1965; Dafermos and Sparrow, 1971; Smith, 1979; Yang and
7 Huang, 1998, 2004; Gentile et al., 2005; Zhang et al., 2011), in which the network is modeled
8 as a set of nodes and links. Although this discrete network approach is relatively simple and
9 has been well-developed in the literature, it may not be suitable for national congestion-
10 pricing studies due to the scale of the issues to be tackled and the amount of data required to
11 set up such a model.

12 In terms of substitutes for the traditional discrete modeling approach, the continuum
13 modeling approach (Blumenfeld, 1977; Sasaki et al., 1990; Wong, 1998; Wong et al., 1998;
14 Ho et al., 2006, 2007a), in which the road network is approximated as a continuum such that
15 users are free to choose their routes in a two-dimensional space (Ho and Wong, 2006), offers
16 an alternative strategy for tackling traffic equilibrium problems. The fundamental assumption
17 made is that differences in modeling characteristics, such as travel cost and demand pattern,
18 as between adjacent areas within a network are relatively small when compared to the degree
19 of variation over the entire network. Hence, the characteristics of a network, such as flow
20 intensity, demand, and travel cost, can be represented by smooth mathematical functions
21 (Vaughan, 1987).

22 Other similar approaches have also been adopted in modeling a regional traffic system at
23 the macro level, e.g. the area-speed flow model (May et al., 2000) and macroscopic
24 fundamental diagrams (Geroliminis and Daganzo, 2008), which operate at a highly
25 aggregated level (e.g. across a whole city) without considering any directional properties of
26 network congestion. These approaches and the continuum approach are particularly suitable
27 for situations in which we are more interested in the general trends and patterns of the
28 distribution and travel choices of users, and in changes in such variables in response to policy
29 changes that affect the transportation system at the macroscopic level, rather than the detailed
30 level. Also, the models based on these macroscopic approaches can be set up using a small
31 amount of data (compared to network models) to provide an efficient and theoretically sound
32 approach for various types of regional transportation studies. Compared with the area-speed
33 flow model and macroscopic fundamental diagrams, the continuum approach, due to its
34 strength in the directional representation of network congestion, can be used to provide a
35 more detailed spatial analysis of a large-scale region and estimate geographical responses to a
36 national road user charging scheme that can supplement the findings of traditional economic
37 analysis (e.g., Glaister and Graham, 2005; Graham et al., 2009).

38 For congestion-pricing applications, Ho et al. (2005) developed a continuum model for
39 cities with a single CBD and a single class of users. Three different cases, the first-best case,
40 the initial-cordon case and the second-best case, were considered in the study and compared
41 with a no-toll scenario. It was found that while the cordon-based charging scheme would
42 achieve a level of social welfare comparable to that realized using the distance-based
43 charging scheme of the first-best case, the cordon-based charging scheme was much easier to
44 implement and more likely to be socially and politically acceptable (Sumalee et al., 2005).

45 The current study provides new insights into both congestion pricing at the macroscopic
46 level and the effectiveness of cordon charging. In any transportation system, users can be
47 classified into different classes by purpose of trip, vehicle type, income level, etc. Each user
48 class may have its own criteria for making trips and selecting particular routes, and may be
49 charged at differing rates. In addition, a common feature of national (or regional) road pricing
50 schemes is a vehicle-type-based charging structure, i.e., a framework in which private cars

1 and heavy good vehicles (HGVs) are subject to different charges. Thus, it is reasonable to
 2 incorporate the concept of multiple user classes into congestion-pricing problems to facilitate
 3 more realistic estimates of flow patterns and more efficient charging mechanisms.

4 A number of discrete modeling approach studies have incorporated multiple user classes
 5 into the congestion-pricing problem (Bellei et al., 2002; Chen et al., 2004; Huang, 2000, 2004;
 6 Yang and Huang, 2004) and the optimization of cordon locations (Zhang and Yang, 2004;
 7 Sumalee, 2007). Taking a different approach, Ho et al. (2007b) introduced a multi-class
 8 congestion model for a continuum transportation system in a city with a single CBD. In that
 9 study, a social welfare maximization model was set up and solved. It was proven that for a
 10 particular type of transportation cost function, there is an anonymous toll rate that enables all
 11 classes of users to attain a welfare-maximized travel pattern. Based on this anonymous toll
 12 pattern, Ho et al. (2007b) suggested a cordon-based charging scheme that strives to achieve a
 13 balance between simplicity in implementation and congestion-pricing effectiveness.

14 By considering the strength of the continuum modeling approach in macroscopic studies,
 15 the current study formulates a regional congestion-pricing problem with multiple user classes
 16 and competing cities. For this regional model, we extend the original multi-class congestion-
 17 pricing model (Ho et al., 2007b) to a more generalized asymmetrical cost function, which
 18 leads to a discriminatory toll pattern. Moreover, a logit-type demand distribution function is
 19 specified to model the probabilistic destination choice behavior of these multi-class users in
 20 such a polycentric region.

21 The remainder of this paper is organized as follows. Section 2 defines the notation and
 22 specific equations that are used in this study. Section 3 introduces the formulation of the
 23 multi-class congestion-pricing problem. The cordon-based congestion-charging scheme for
 24 this multi-class multi-destination problem is discussed in section 4. The relationships between
 25 the proposed models and the algorithms presented in sections 3 and 4 are shown in Figure 1.
 26 A numerical example that demonstrates the effectiveness of the proposed models is given in
 27 Section 5. We conclude the paper in Section 6.

30 2. Definitions and notation

31 Consider a study region of an arbitrary shape with multiple cities competing for different
 32 classes of users who are continuously distributed over the study region (Figure 2). Each user
 33 within the study region makes a choice of city based on his utility gain in choosing that city,
 34 the total cost of patronizing that city and the toll incurred along the path to that city. The
 35 users will travel from their respective demand locations to the chosen city along the least
 36 costly route. Let M denote the set of vehicle types, such as motorcycles, passenger cars,
 37 trucks, etc., used to identify different classes of users in this study. Let the study region be
 38 represented by Ω and the boundary by Γ . Further, let the locations of cities be $O_n, \forall n \in N$,
 39 where N is the set of cities within the region. To avoid singularity among the cities, it is
 40 assumed that each of the cities is of finite size and enclosed by a boundary, Γ_{cn} . The travel
 41 cost per unit distance of travel at location $(x, y) \in \Omega$ for class m users is denoted by $c_m(x, y)$,
 42 which is location dependent and has the following functional relationship with traffic flows at
 43 that location:

$$44 \quad c_m(x, y) = a_m(x, y) + \sum_{r \in M} \sum_{s \in N} b_{mr}(x, y) |\mathbf{f}_{rs}(x, y)|, \quad \forall (x, y) \in \Omega, m \in M, \quad (1)$$

45 where $a_m(x, y)$ and $b_{mr}(x, y)$, which are strictly positive scalar functions of the cost-flow
 46 relationship that reflects the local characteristics of the streets at location $(x, y) \in \Omega$, are
 47 respectively the free-flow travel cost of class m users and the congestion-related parameter of

1 class r users on class m . As $b_{mr}(x, y)$ is not necessarily equal to $b_{rm}(x, y)$, the asymmetric
 2 congestion effect between different vehicle types can be modeled.
 3 $\mathbf{f}_{mn}(x, y) = (f_{xmn}(x, y), f_{ymn}(x, y))$ is the flow vector of class m users who are heading to city
 4 n , and $f_{xmn}(x, y)$ and $f_{ymn}(x, y)$ are the corresponding flow in the x and y directions,
 5 respectively. The flow vector $\mathbf{f}_{mn}(x, y)$ indicates the class m users' movement direction at
 6 location $(x, y) \in \Omega$ when heading to city n in the two-dimensional plane, and

$$7 \quad |\mathbf{f}_{mn}(x, y)| = \sqrt{f_{xmn}^2(x, y) + f_{ymn}^2(x, y)} \quad (2)$$

8 is the corresponding flow intensity which measures in a unit time the number of class m users
 9 who cross a small segment of unit width perpendicular to the flow direction when heading to
 10 city n . Equation (1) represents an isotropic cost function because it depends only on flow
 11 intensity and not on flow direction. The congestion component includes the flow intensities
 12 for all classes of users because all users share the same road space. Each user creates an
 13 additional delay cost to all other users as he makes the road more congested.

14 For each class of users, they have the choice of traveling to any of the cities within the
 15 modeling region, as well as the choice of non-travel. In order to demonstrate the flexibility of
 16 the proposed model, the stochastic choice of travel, which is the most general case, is adopted
 17 for the model formulation. The probabilities of these choices being made depend on the
 18 utility gain, the total transportation cost and the total toll paid to travel from the user's
 19 demand location to each of the cities, and are governed by a logit-type distribution. The
 20 demand distributed to any of these choices (including the choice of all cities and the choice of
 21 non-travel) can be expressed as:

$$22 \quad q_{m\tilde{n}}(x, y) = q_m(x, y) \frac{\exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}})]}{\sum_s \exp[\zeta_m (U_{ms} - \bar{u}_{ms})]}, \quad \forall (x, y) \in \Omega, \tilde{n}, s \in \tilde{N}, m \in M, \quad (3)$$

23 where $\tilde{n} \in \tilde{N}$ and \tilde{N} is the enlarged set of N that includes the choice of non-travel; $q_{m\tilde{n}}(x, y)$
 24 is the demand of class m users at location $(x, y) \in \Omega$ who make a choice \tilde{n} ; $q_m(x, y)$ is the
 25 fixed total demand of class m users; $U_{m\tilde{n}}$ is the utility gain for class m users making a choice
 26 \tilde{n} ; $\bar{u}_{m\tilde{n}}(x, y)$ is the total travel cost (including the toll, if any) for a class m user at his
 27 demand location $(x, y) \in \Omega$ of making a choice \tilde{n} ($\bar{u}_{m\tilde{n}}(x, y) = 0$ if the user chooses not to
 28 travel); and ζ_m is a scalar sensitivity parameter of the distribution function for class m users.
 29 A large value of ζ_m indicates that users are more sensitive to the utility gain in making their
 30 choice. The system reduces to a deterministic choice model if the value of ζ_m is very large.
 31 For each combination of user class and city, the following flow conservation equation should
 32 be satisfied:

$$33 \quad \nabla \cdot \mathbf{f}_{mn}(x, y) - q_{mn}(x, y) = 0, \quad \forall (x, y) \in \Omega, n \in N, m \in M. \quad (4)$$

34 Assuming no flow across the boundary of the study area, we have the following boundary
 35 condition:

$$36 \quad \mathbf{f}_{mn}(x, y) \cdot \mathbf{n}(x, y) = 0, \quad \forall (x, y) \in \Gamma, n \in N, m \in M. \quad (5)$$

37 However, it is not difficult to extend the model to represent a given demand pattern that
 38 enters or leaves the region at the boundary. In such a case, we need only replace the above
 39 constraint with $\mathbf{f}_{mn}(x, y) \cdot \mathbf{n}(x, y) = g_{mn}(x, y)$, where $g_{mn}(x, y)$ is the demand function and

$\mathbf{n}(x, y)$ is the unit normal vector pointing away from the modeled region at the boundary point $(x, y) \in \Gamma$. Please be noted that for reducing the notational burden, the argument (x, y) for various variables are suppressed in the subsequence sections.

For the continuum modeling approach, as discussed in previously paragraphs, is capable of providing a detailed analysis of the traffic conditions with relatively small number of control variables. Such characteristics of the continuum model make it more suitable than the discrete network in national-scale models. One of the major concerns in using continuum model for national-scale model is its suitability in modeling the sparse rural area network. This issue can be addressed in two ways. First, for areas that are totally not accessible by traffic (e.g. lakes and mountains), we can set a boundary that restricts traffic from entering these areas in the continuum network (Ho et al., 2005). Second, for areas with sparse rural roads with low capacities, which normally will not be used for the inter-regional travel, additional travel cost penalty can be added to the unit transportation cost function (Equation 1) to increase the travel cost across these areas. As a result, traffic will be diverted to the inter-regional highways with a comparatively low and reasonable travel cost.

3. Utility maximization (first-best solution)

This section aims to find the optimal toll rate pattern such that the utility of the system as a whole is maximized. To find this optimal toll rate, the direct utility function of the representative traveler for this multi-class continuum transportation system with multiple cities is first derived (Appendix A). Based on this utility function, the total utility of the system is found and maximized in section 3.1 to ascertain the optimal toll rate. Section 3.2 gives the algorithm used to solve the utility maximization problem. The relationship between the relevant models is illustrated in Figure 1.

3.1. Formulation of the utility maximization model

The multi-class congestion-pricing problem with a logit-type distribution is formulated as a utility maximization problem (details could be found in Appendix A). The problem can be formulated as the following mathematical program for maximizing the total system utility, $\tilde{U}(\mathbf{f}, \mathbf{q})$:

$$\text{Maximize}_{\mathbf{f}, \mathbf{q}} \tilde{U}(\mathbf{f}, \mathbf{q}) = \iint_{\Omega} \sum_m \sum_{\tilde{n}} q_{m\tilde{n}} U_{m\tilde{n}} + \sum_m \frac{1}{\zeta_m} q_m \ln q_m - \sum_m \sum_{\tilde{n}} \frac{1}{\zeta_m} q_{m\tilde{n}} \ln q_{m\tilde{n}} - \sum_m \sum_n c_m |\mathbf{f}_{mn}| \, d\Omega \quad (6a)$$

$$\text{subject to} \quad \nabla \mathbf{f}_{mn} - q_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Omega, \quad (6b)$$

$$q_m - \sum_{\tilde{n}} q_{m\tilde{n}} = 0 \quad \forall m \in M, \tilde{n} \in \tilde{N}, (x, y) \in \Omega, \quad (6c)$$

$$\mathbf{f}_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Gamma. \quad (6d)$$

Consider the Lagrangian for the above maximization problem:

$$L = \iint_{\Omega} \sum_m \sum_{\tilde{n}} q_{m\tilde{n}} U_{m\tilde{n}} + \sum_m \frac{1}{\zeta_m} q_m \ln q_m - \sum_m \sum_{\tilde{n}} \frac{1}{\zeta_m} q_{m\tilde{n}} \ln q_{m\tilde{n}} - \sum_m \sum_n c_m |\mathbf{f}_{mn}| + \sum_m \sum_n \bar{u}_{mn} (\nabla \mathbf{f}_{mn} - q_{mn}) + \sum_m \nu_m \left(q_m - \sum_{\tilde{n}} q_{m\tilde{n}} \right) \, d\Omega + \sum_m \sum_n \int_{\Gamma} \mathbf{f}_{mn} \cdot \mathbf{w}_{mn} \, d\Gamma$$

1 As it could be proven (in the later part of this section) that the Lagrangian multipliers of
 2 constraints (6b) are equivalent to the corresponding total travel costs at the optimal solution
 3 of the maximization problem (6), \bar{u}_{mn} , which defines the total travel cost for class m users
 4 making a choice n , is used to denote the Lagrangian multiplier of Equation (6b) in order to
 5 reduce the notational burden. Applying the variational principle to the Lagrangian, we have:

$$\begin{aligned} \delta L = & \iint_{\Omega} \sum_m \sum_{\tilde{n}} U_{m\tilde{n}} \delta q_{m\tilde{n}} - \sum_m \sum_{\tilde{n}} \frac{1}{\zeta_m} (\ln q_{m\tilde{n}} + 1) \delta q_{m\tilde{n}} - \sum_m \sum_n a_m \frac{\mathbf{f}_{mn} \cdot \delta \mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} \\ & - \sum_m \sum_n \sum_o \sum_p (b_{mo} + b_{on}) \frac{\mathbf{f}_{mn} \cdot \delta \mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} + \sum_m \sum_n (\nabla \mathbf{f}_{mn} - q_{mn}) \delta \bar{u}_{mn} \\ & + \sum_m \sum_n \bar{u}_{mn} \nabla \delta \mathbf{f}_{mn} - \sum_m \sum_n \bar{u}_{mn} \delta q_{mn} + \sum_m \left(q_m - \sum_{\tilde{n}} q_{m\tilde{n}} \right) \delta v_m + \sum_m \sum_{\tilde{n}} v_m \delta q_{m\tilde{n}} \, d\Omega \\ & + \sum_m \sum_n \int_{\Gamma} \mathbf{f}_{mn} \cdot \delta \mathbf{w}_{mn} \, d\Gamma + \sum_m \sum_n \int_{\Gamma} \mathbf{w}_{mn} \cdot \delta \mathbf{f}_{mn} \, d\Gamma \end{aligned}$$

7
 8 Rearranging and applying the divergence theorem, we have:

$$\begin{aligned} \delta L = & \iint_{\Omega} \sum_m \sum_{\tilde{n}} \left(U_{m\tilde{n}} - \frac{1}{\zeta_m} \ln q_{m\tilde{n}} - \frac{1}{\zeta_m} - \bar{u}_{m\tilde{n}} + v_m \right) \delta q_{m\tilde{n}} + \sum_m \sum_n (\nabla \mathbf{f}_{mn} - q_{mn}) \delta \bar{u}_{mn} \\ & + \sum_m \left(q_m - \sum_{\tilde{n}} q_{m\tilde{n}} \right) \delta v_m - \sum_m \sum_n \left[\left(a_m + \sum_o \sum_p (b_{mo} + b_{op}) \frac{|\mathbf{f}_{op}|}{|\mathbf{f}_{mn}|} \right) \frac{\mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} + \nabla \bar{u}_{mn} \right] \cdot \delta \mathbf{f}_{mn} \, d\Omega \\ & + \sum_m \sum_n \int_{\Gamma_{cn}} \bar{u}_{mn} \mathbf{n}_c \cdot \delta \mathbf{f}_{mn} \, d\Gamma + \sum_m \sum_n \int_{\Gamma} \mathbf{f}_{mn} \cdot \delta \mathbf{w}_{mn} \, d\Gamma + \sum_m \sum_n \int_{\Gamma} (\mathbf{w}_{mn} + u_{mn} \mathbf{n}) \cdot \delta \mathbf{f}_{mn} \, d\Gamma \end{aligned}$$

10 Given that $\delta q_{m\tilde{n}}$, $\delta \bar{u}_{mn}$ and δv_m are arbitrary functions in Ω , $\delta \mathbf{f}_{mn}$ is an arbitrary function in
 11 $\Omega \cup \Gamma_c$, $\delta \mathbf{w}_{mn}$ is an arbitrary function in Γ , and $\delta \mathbf{f}_{mn}$ vanishes in Γ , at $\delta L = 0$, it follows that

$$12 \quad U_{m\tilde{n}} - \frac{1}{\zeta_m} \ln q_{m\tilde{n}} - \frac{1}{\zeta_m} - \bar{u}_{m\tilde{n}} + v_m = 0 \quad \forall m \in M, \tilde{n} \in \tilde{N}, (x, y) \in \Omega \quad (7)$$

$$13 \quad \left(a_m + \sum_o \sum_p (b_{mo} + b_{op}) \frac{|\mathbf{f}_{op}|}{|\mathbf{f}_{mn}|} \right) \frac{\mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} + \nabla \bar{u}_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Omega \quad (8)$$

$$14 \quad \nabla \mathbf{f}_{mn} - q_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Omega \quad (9)$$

$$15 \quad q_m - \sum_{\tilde{n}} q_{m\tilde{n}} = 0 \quad \forall m \in M, \tilde{n} \in \tilde{N}, (x, y) \in \Omega \quad (10)$$

$$16 \quad \bar{u}_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Gamma_{cn} \quad (11)$$

$$17 \quad \mathbf{f}_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Gamma \quad (12)$$

18 From Equation (8), we can observe that the flow vector is directly opposite to the gradient of the
 19 Lagrange multiplier \bar{u}_{mn} , i.e.

$$20 \quad -\mathbf{f}_{mn} // \nabla \bar{u}_{mn} \quad \text{wherever } \mathbf{f}_{mn} \neq 0, \quad (13)$$

21 where ‘//’ means that the two vectors are in the same direction and

$$|\nabla \bar{u}_{mn}| = a_m + \sum_o \sum_p b_{mo} |\mathbf{f}_{op}| + \sum_o \sum_p b_{om} |\mathbf{f}_{op}| \quad (14)$$

Now, let us specify a modified cost function

$$\bar{c}_m = a_m + \sum_o \sum_p b_{mo} |\mathbf{f}_{op}| + \sum_o \sum_p b_{om} |\mathbf{f}_{op}| \quad (15)$$

within the modeled region. For any route p used by a class m user to travel from the demand location (H) to city n (O_n), if we integrate the modified cost along this path, the total cost incurred by the user can be obtained as

$$\bar{C}_{mnp} = \int_p \bar{c}_m ds = \int_p \bar{c}_m \frac{\mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} \cdot d\mathbf{s} = -\int_p \nabla \bar{u}_{mn} \cdot d\mathbf{s} = -\bar{u}_{mn}(O_n) + \bar{u}_{mn}(H) = \bar{u}_{mn}(H) \quad (16)$$

using equations (8, 11 and 15) and based on the fact that $\mathbf{f}_{mn} / |\mathbf{f}_{mn}|$ is a unit vector that is parallel to $d\mathbf{s}$ along the path. Therefore, the total modified cost is independent of the path used and $\bar{C}_{mnp} = \bar{u}_{mn}(H)$. In contrast, for any unused path p' between a user's demand location H in Ω and the chosen city n , the total cost incurred by the user is

$$\bar{C}_{mnp'} = \int_{p'} \bar{c}_m ds \geq \int_{p'} \bar{c}_m \frac{\mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} \cdot d\mathbf{s} = -\int_{p'} \nabla \bar{u}_{mn} \cdot d\mathbf{s} = -\bar{u}_{mn}(O_n) + \bar{u}_{mn}(H) = \bar{u}_{mn}(H) \quad (17)$$

using equations (8, 11 and 15). Hence, $\bar{C}_{mnp'} \geq \bar{u}_{mn}(H)$. The inequality in the above derivation is due to the fact that for some segments along the path p' , the normal vectors $\mathbf{f}_{mn} / |\mathbf{f}_{mn}|$ and $d\mathbf{s}$ are not parallel and hence $ds > (\mathbf{f}_{mn} / |\mathbf{f}_{mn}|) \cdot d\mathbf{s}$ for these segments of path p' . Therefore, for any unused path, the total modified cost is greater than or equal to that of any of the used paths. In this way, the model guarantees that the user will choose his/her route in the region in a user-optimal manner with respect to the modified cost functions. Moreover, the Lagrangian multiplier $\bar{u}_{mn}(x, y)$ can be interpreted as the minimum total modified cost incurred by class m users in traveling from a location $(x, y) \in \Omega$ to city n .

Comparing the modified cost function defined in Equation (15) with the travel cost function defined in Equation (1), it can be observed that the modified cost function consists of the following two components:

$a_m + \sum_o \sum_p b_{mo} |\mathbf{f}_{op}|$ = the actual cost per unit distance experienced by a class m user traveling at $(x, y) \in \Omega$;

$\sum_o \sum_p b_{om} |\mathbf{f}_{op}|$ = the user externality, which is the cost per unit distance that a marginal class m user imposes on other classes of users that are already traveling at $(x, y) \in \Omega$.

This means that if the travel cost of class m users is modified by charging a toll rate of

$$\tau_m = \sum_o \sum_p b_{om} |\mathbf{f}_{op}|, \quad \forall o \in M, p \in N, (x, y) \in \Omega, \quad (18)$$

per unit distance of travel in the modeled region, the flow pattern corresponding to utility maximization will be in user equilibrium. In contrast to Ho et al. (2007b), which a single anonymous toll rate is found for all classes of users, Equation (18) shows that (due to the consideration of asymmetrical cost function in this paper) discriminatory toll rates for

1 different classes of users are necessary for achieving the first-best solution. With the
 2 advances made in intelligent transportation systems, charging for this component is now
 3 feasible using global positioning system (GPS) technology, in which every unit of movement
 4 at location (x, y) in the modeled region can be charged at an amount equivalent to $\tau_m(x, y)$,
 5 as shown in Equation (18). Rearranging Equation (7), we have:

$$6 \quad q_{m\tilde{n}} = \exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}} + v_m) - 1] = \exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}})] \exp[\zeta_m v_m - 1] \quad (19)$$

7 Substituting Equation (19) into Equation (10), we have:

$$8 \quad q_m - \exp[\zeta_m v_m - 1] \sum_{s \in \tilde{N}} \exp[\zeta_m (U_{ms} - \bar{u}_{ms})] = 0 \quad (20)$$

$$\exp[\zeta_m v_m - 1] = \frac{q_m}{\sum_{s \in \tilde{N}} \exp[\zeta_m (U_{ms} - \bar{u}_{ms})]}$$

9 Substituting Equation (20) back into Equation (19), we have:

$$10 \quad q_{m\tilde{n}} = q_m \frac{\exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}})]}{\sum_{s \in \tilde{N}} \exp[\zeta_m (U_{ms} - \bar{u}_{ms})]} \quad (21)$$

11 As $q_{m\tilde{n}}$ is only dependent on the exponential functions and a positively defined quantity q_m ,
 12 the non-negativity constraint of $q_{m\tilde{n}}$, Equation (A-14c), will be satisfied. Equation (21) shows
 13 that there is a stochastic distribution of demand for users among their choices, \tilde{N} .
 14

15 **3.2. A solution algorithm for the utility maximization model**

16 Similar to Wong (1998), the finite element method (FEM) is used to approximate the
 17 continuum nature of the utility maximization problem stated in Equation (6) (Zienkiewicz
 18 and Taylor, 1989). Expanding the Lagrangian (L) of this maximization problem with Taylor
 19 series and neglecting the higher order term, we have:

$$20 \quad \Psi_{k+1} = \Psi_k - \mathbf{H}(\Psi_k)^{-1} \mathbf{R}(\Psi_k) \quad (22)$$

21 where $\Psi_k = (\mathbf{f}_{mn}, \bar{u}_{mn}, \forall m \in M, n \in N)$ is the solution vector at iteration k ; $\mathbf{R}(\Psi_k)$ and
 22 $\mathbf{H}(\Psi_k)$ are respectively the residual vector and Hessian matrix of the Lagrangian in iteration k .
 23 Based on this iterative equation and the golden section method (Sheffi, 1985) for determining
 24 the step size ν , the solution procedure for this utility maximization problem can be set as
 25 follows.
 26

27 Solution Procedure A

28 Step A1: Find an initial solution Ψ_0 . Set $k = 0$.

29 Step A2: Evaluate $\mathbf{R}(\Psi_k)$ and $\mathbf{H}(\Psi_k)$.

30 Step A3: Apply the golden section method to determine the step size ν^* which minimizes
 31 the norm of the residual vector $\left| \mathbf{R}(\Psi_k - \nu \mathbf{H}(\Psi_k)^{-1} \mathbf{R}(\Psi_k)) \right|$. Then, set

$$32 \quad \Psi_{k+1} = \Psi_k - \nu^* \mathbf{H}(\Psi_k)^{-1} \mathbf{R}(\Psi_k).$$

33 Step A4: If the relative error $\left| \Psi_{k+1} - \Psi_k \right| / \left| \Psi_k \right|$, is less than an acceptable error ϵ , then
 34 terminate, and Ψ_{k+1} is the solution.

1 Step A5: If the relative error exceeds the acceptable error ε then replace Ψ_k with Ψ_{k+1} . Set k
 2 $= k + 1$ and go to Step A2.

4. The cordon-based congestion-pricing problem

5 If a distance-based charging technology (e.g. a GPS-based charging system) is not available,
 6 the first-best scheme cannot be practically implemented. Moreover, it is not socially
 7 acceptable to change the charging rate continuously for those traveling in the modeled region.
 8 It is more practical to formulate a multi-layer cordon-based charging scheme in which each
 9 layer is charged a fixed toll. In this scheme, the whole modeled region is divided into a small
 10 number of sub-regions that are separated by well-defined cordons. In section 4.1, the cordons
 11 and the corresponding cordon tolls are defined on the basis of the optimal toll rate found by
 12 solving the utility maximization problem given in section 3.1. For the set of given cordons
 13 and cordon tolls, section 4.2 develops a cordon-based congestion pricing model which is
 14 solved using the solution algorithm introduced in section 4.3 to account for the user behavior
 15 in the presence of such a charging scheme. Taking into account the given cordons and the
 16 reactions of users, section 4.4 formulates an optimization model that allows for the optimal
 17 cordon toll, i.e. the toll that maximizes total system utility, to be ascertained. The
 18 relationships of these models and algorithms with the model introduced in the previous
 19 section are summarized in Figure 1.

20

4.1. Definition of cordons and cordon tolls

22 For the multi-layer cordon-based charging scheme, we need to determine the total toll that a
 23 user has to pay when traveling from his/her home location to the city boundary. Although the
 24 minimum total modified cost, $\bar{u}_{mn}(x, y)$, can be obtained by solving mathematical program
 25 (6), the total travel cost and toll cannot be readily ascertained from the solution. To separate
 26 the total travel cost, $u_{mn}(x, y)$, and the toll paid from the function $\bar{u}_{mn}(x, y)$, we define a toll
 27 function, $T_{mn}(x, y)$, which measures the total toll that a class m user has to pay when
 28 traveling from location (x, y) to the boundary of city n . Clearly, we have $\nabla T_{mn} = \tau_m^* (\mathbf{f}_{mn}^* / |\mathbf{f}_{mn}^*|)$,
 29 where $\nabla T_{mn} = (\partial T_{mn} / \partial x, \partial T_{mn} / \partial y)$ is the gradient of the function $T_{mn}(x, y)$. \mathbf{f}_{mn}^* and τ_m^* are
 30 respectively the solutions of the flow vector and toll rate, which are determined by solving
 31 mathematical program (6), and $\mathbf{f}_{mn}^* / |\mathbf{f}_{mn}^*|$ is the normal vector of the optimal flow pattern. To
 32 solve this partial differential equation, we reformulate it as a least square problem as follows:

$$33 \quad \text{Minimize}_{\mathbf{T}} \quad Z(\mathbf{T}) = \iint_{\Omega} \sum_m \sum_n \left| \nabla T_{mn} + \tau_m^* \frac{\mathbf{f}_{mn}^*}{|\mathbf{f}_{mn}^*|} \right|^2 d\Omega \quad (23a)$$

$$34 \quad T_{mn} = 0, \quad \forall m \in M, \forall n \in N, \forall (x, y) \in \Gamma_{cn} \quad (23b)$$

35 This is a post-analysis of the optimal solution obtained from mathematical problem (6).
 36 Based on the total toll $T_{mn}(x, y)$ determined from this post-analysis, we can (for each
 37 combination of m and n) construct $J_{mn}+1$ iso-toll contours, $\mathbf{W}_{mn} = (W_{mn}^0, W_{mn}^1, \dots, W_{mn}^J)$ with
 38 $\bar{\mathbf{T}}_{mn} = (T_{mn}^0, T_{mn}^1, \dots, T_{mn}^J)$ as the corresponding toll levels of the constructed cordons, in the
 39 modeled region. Note that $W_{mn}^0 = \Gamma_{cn}$, which is also the boundary of city n , has a toll level
 40 $T_{mn}^0 = 0$. For the two consecutive iso-toll contours of levels T_{mn}^j and T_{mn}^{j+1} (with $T_{mn}^{j+1} > T_{mn}^j$

1 and $j \in [0, J_{mn}]$), which are considered to be the charging cordons, the area embraced by
 2 these two contours forms a layer in which any class m user who patronizes city n and enters
 3 this layer is charged a fixed toll of $\hat{\tau}_{mn}^{j+1} = T_{mn}^{j+1} - T_{mn}^j$. This means that $\hat{\tau}_{mn}^{j+1}$ is the fixed toll
 4 charged on all class m users, who are heading to city n , as they cross the W_{mn}^{j+1} cordon.
 5 Although this cordon-based charging scheme is more realistic than a point-wise charging
 6 scheme, it remains impractical given that not all users have to be charged whenever they pass
 7 through a cordon. For example, for a class 1 user traveling to city 1, based on this scheme, he
 8 will not be charged when he passes the W_{12}^1 cordon, as this cordon is set up for this class of
 9 users when they travel to city 2. Thus, to improve the practicality of the charging scheme, a
 10 new set of J' cordons, $\Lambda = (\Lambda^1, \Lambda^2, \dots, \Lambda^{J'})$, which are applicable to all classes of users, are
 11 selected from the set of previously established cordons, $\mathbf{W} = (\mathbf{W}_{mn}, \forall m \in M, n \in N)$. Let $\tilde{\tau}_m^i$
 12 be the toll for class m users who cross the Λ^i cordon and is defined based on the total toll
 13 pattern $\mathbf{T} = (\bar{\mathbf{T}}_{mn}, \forall m \in M, n \in N)$. The set of fixed tolls $\tilde{\boldsymbol{\tau}} = (\tilde{\tau}_m^j, \forall m \in M, j \in J')$ and the
 14 corresponding cordon $\Lambda = (\Lambda^1, \Lambda^2, \dots, \Lambda^{J'})$ then define the cordon-based charging scheme.

15

16 **4.2. Formulation of a cordon-based congestion-charging model**

17 Based on the cordon set Λ and the corresponding charging scheme $\tilde{\boldsymbol{\tau}}$ defined in the
 18 previous section, the cordon-based congestion-pricing model, which gives users' responses to
 19 such cordon set and charging scheme, is set up. Compared with the cordon-based model
 20 introduced in Ho et al. (2005), the cordon-based model introduced in this paper is more
 21 complicated, as the cordon-toll charged in this model, while path-dependent, is not location-
 22 dependent. Thus, a fixed-point model is introduced to solve this path-dependent cordon-toll
 23 problem. Two sub-models, named the toll determination sub-model and the flow pattern
 24 determination sub-model, are included in this fixed-point problem. In the toll determination
 25 sub-model, the paths chosen by system users are traced graphically from their home locations
 26 to their chosen cities based on the given flow pattern. Based on these traced paths, the total
 27 toll (\tilde{T}_{mn}) for the system users is determined by adding together the fixed tolls ($\tilde{\tau}_m^j$) for all the
 28 cordons crossed along the users' paths. This sub-model can be represented in the following
 29 abstract form:

$$30 \quad \tilde{\mathbf{T}} = P(\mathbf{F}) \quad (24)$$

31 where $\tilde{\mathbf{T}} = (\tilde{T}_{mn}, \forall m \in M, n \in N)$ is the toll calculated on the basis of this cordon-based
 32 charging scheme, and $\mathbf{F} = (\mathbf{f}_{mn}, \forall m \in M, n \in N)$ is the corresponding flow pattern. On the
 33 other side of the fixed-point problem, the flow pattern determination sub-model is aimed at
 34 finding the user equilibrium flow pattern \mathbf{F} for a given toll pattern $\tilde{\mathbf{T}}$. This user equilibrium
 35 flow pattern can be obtained by solving the following set of differential equations:

$$36 \quad \left(a_m + \sum_o \sum_p b_{mo} |\mathbf{f}_{op}| \right) \frac{\mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} + \nabla u_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Omega \quad (25a)$$

$$37 \quad q_m \frac{\exp[\zeta_m (U_{mn} - u_{mn} - \tilde{T}_{mn})]}{\sum_{\tilde{n}} \exp[\zeta_m (U_{m\tilde{n}} - u_{m\tilde{n}} - \tilde{T}_{m\tilde{n}})]} - \nabla \mathbf{f}_{mn} = 0 \quad \forall m \in M, n \in N, \tilde{n} \in \tilde{N}, (x, y) \in \Omega \quad (25b)$$

$$38 \quad \mathbf{f}_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Gamma \quad (25c)$$

$$1 \quad u_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Gamma_{cn} \quad (25d)$$

2 where u_{mn} , as defined in Section 4.1, is the total travel cost (excluding tolls); $\tilde{T}_{m\tilde{n}}$ is the total
 3 toll that a class m user has to pay for making choice \tilde{n} and is kept constant in this sub-model.
 4 $\tilde{T}_{m\tilde{n}}$ is equal to \tilde{T}_{mn} , which is the solution of the toll determination sub-model, for users
 5 choose to travel to any of the cities. Note that $\tilde{T}_{m\tilde{n}}$ is equal to zero where users choose not to
 6 travel. The fraction in Equation (25b) is the modified demand distribution function based on
 7 this total toll ($\tilde{T}_{m\tilde{n}}$). Similar to Equation (16) and (17), the total cost incurred by a class m
 8 user for any used route p or unused route p' in traveling from the demand location (H) to city
 9 n (O_n) can be defined as:

$$10 \quad \tilde{C}_{mnp} = \int_p c_m ds = \int_p c_m \frac{\mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} \cdot \mathbf{ds} = -\int_p \nabla u_{mn} \cdot \mathbf{ds} = u_{mn}(H)$$

$$11 \quad \tilde{C}_{mnp'} = \int_{p'} c_m ds \geq \int_{p'} c_m \frac{\mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} \cdot \mathbf{ds} = -\int_{p'} \nabla u_{mn} \cdot \mathbf{ds} = u_{mn}(H)$$

12 Therefore, $\tilde{C}_{mnp'} \geq \tilde{C}_{mnp}$, indicating that this set of differential equations satisfies the user
 13 equilibrium condition in path selection. Similar to the toll determination sub-model, this flow
 14 pattern determination sub-model can be represented in the following abstract form:

$$15 \quad \mathbf{F} = S(\tilde{\mathbf{T}}) \quad (26)$$

16 Combining Equation (24) and (26), the cordon-based congestion-charging problem can be
 17 formulated as the following fixed-point problem:

$$18 \quad \mathbf{F} = S(P(\mathbf{F})) \quad (27)$$

19 Solving the fixed-point problem (27), the flow pattern \mathbf{F} that satisfies the user
 20 equilibrium condition (the flow pattern determination sub-model) and takes into account the
 21 cordon toll charged (the toll determination sub-model) can be found. Using this flow pattern
 22 \mathbf{F} and the total no-toll travel cost u_{mn} , which is found from the flow pattern, the solution Φ for
 23 this cordon-based congestion-pricing problem can be determined.
 24

25 **4.3. A solution algorithm for the cordon-based congestion-charging model**

26 In the toll determination sub-model, the total toll ($\tilde{\mathbf{T}}$) can be found by tracing the chosen path,
 27 which is defined by the flow pattern (\mathbf{F}), from the demand location to the destination. For the
 28 flow pattern determination sub-model, the finite element method (FEM) is used to
 29 approximate the continuous variables in the modeled region (Zienkiewicz and Taylor, 1989).
 30 As there is no explicit objective function for this multi-class cordon-based congestion-
 31 charging problem, the mixed finite element procedure used in Wong (1998) cannot be applied
 32 directly. Hence, we adopt the Galerkin formulation of the weighted residual technique
 33 (Cheung et al., 1996; Zienkiewicz and Taylor, 1989) in which differential equations (25a) and
 34 (25b) are transformed into the following equivalent integral expressions:

$$35 \quad \iint_{\Omega} \left[\left(a_m + \sum_o \sum_p b_{mo} |\mathbf{f}_{op}| \right) \frac{\mathbf{f}_{mn}}{|\mathbf{f}_{mn}|} + \nabla u_{mn} \right] \psi(x, y) d\Omega = 0 \quad \forall m \in M, n \in N, \psi(x, y) \quad (28)$$

$$1 \quad \iint_{\Omega} \left[q_m \frac{\exp[\zeta_m(U_{mn} - u_{mn} - \tilde{T}_{mn})]}{\sum_{\tilde{n}} \exp[\zeta_m(U_{m\tilde{n}} - u_{m\tilde{n}} - \tilde{T}_{m\tilde{n}})]} - \nabla \mathbf{f}_{mn} \right] \psi(x, y) d\Omega = 0 \quad \forall m \in M, n \in N, \tilde{n} \in \tilde{N}, \psi(x, y)$$

$$2$$

$$3 \quad (29)$$

4 where $\psi(x, y)$ is the trial (or weight) function in the weighted residual technique. Boundary
5 conditions (25c) and (25d) are enforced by taking a zero weight function (Cheung et al.,
6 1996). In the Galerkin formulation, the local interpolation function of the finite element is
7 used as the trial function. The modeling area is first discretized into a finite element mesh
8 (see Figure 3 for an example) in which the Galerkin formulation is applied at the element
9 level. The governing equations at a particular finite element node s are given as follows:

$$10 \quad \left\{ \mathbf{r}_{mn}^s(\Phi) \right\} = \begin{Bmatrix} r_{mn1}^s \\ r_{mn2}^s \\ r_{mn3}^s \end{Bmatrix} = \begin{Bmatrix} \sum_{e \in E_s} \iint_{\Omega_e} \left[\left(a_m + \sum_o \sum_p b_{mo} |\mathbf{f}_{op}| \right) \frac{f_{xmn}}{|\mathbf{f}_{mn}|} + \frac{\partial u_{mn}}{\partial x} \right] N_s(x, y) d\Omega \\ \sum_{e \in E_s} \iint_{\Omega_e} \left[\left(a_m + \sum_o \sum_p b_{mo} |\mathbf{f}_{op}| \right) \frac{f_{ymn}}{|\mathbf{f}_{mn}|} + \frac{\partial u_{mn}}{\partial y} \right] N_s(x, y) d\Omega \\ \sum_{e \in E_s} \iint_{\Omega_e} \left[q_m \frac{\exp[\zeta_m(U_{mn} - u_{mn} - \tilde{T}_{mn})]}{\sum_{\tilde{n} \in N} \exp[\zeta_m(U_{m\tilde{n}} - u_{m\tilde{n}} - \tilde{T}_{m\tilde{n}})]} - \nabla \mathbf{f}_{mn} \right] N_s(x, y) d\Omega \end{Bmatrix} \quad \forall m, n$$

$$11 \quad (30)$$

12 where Ω_e denotes the domain of the finite element e ; E_s is the set of finite elements that
13 connects node s ; $N_s(x, y)$ is the local interpolation function of the finite element that
14 corresponds to node s ; \mathbf{r}_{mn}^s is the nodal residual vector for class m users traveling to city n at
15 node s , which represents how far governing equations (25a) and (25b) are satisfied locally
16 around node s ; and $\Phi = (\mathbf{F}, \mathbf{T}, \mathbf{u})$, where $\mathbf{u} = (u_{mn}, \forall m \in M, n \in N)$, is the solution vector of
17 the problem. For the global satisfaction of governing equations (25), we require that
18 $\mathbf{R}(\Phi) = \text{Col}(\mathbf{r}_{mn}^s(\Phi)) = \mathbf{0}$. For this flow pattern determination sub-model, we apply the typical
19 Newton-Raphson algorithm with a line search to solve the problem, in which we derive the
20 iterative equation

$$21 \quad \Phi_{k+1} = \Phi_k - \lambda \mathbf{J}(\Phi_k)^{-1} \mathbf{R}(\Phi_k) \quad (31)$$

22 where $\mathbf{J}(\Phi_k)$ is the Jacobian matrix of vector $\mathbf{R}(\Phi_k)$ in iteration k and λ is the step size (to
23 be determined using the golden section method). By adding a step for solving the toll
24 determination sub-model in the Newton-Raphson algorithm adopted for the flow pattern
25 determination sub-model, both sub-models, thus the fixed-point problem, could be solved
26 simultaneously. The solution procedure for this fixed-point cordon-based congestion-pricing
27 problem can be set as follows:

29 Solution Procedure B

30 Step B1: Find an initial solution Φ_0 . Set $k = 0$.

31 Step B2: With the solution Φ_k , solve the total toll $\tilde{\mathbf{T}}_k$ using the toll determination sub-
32 model (24)

- 1 Step B3: Evaluate $\mathbf{R}(\Phi_k)$ and $\mathbf{J}(\Phi_k)$.
- 2 Step B4: Apply the golden section method to determine the step size λ^* which minimizes
- 3 the norm of the residual vector $\left| \mathbf{R}(\Phi_k - \lambda \mathbf{J}(\Phi_k)^{-1} \mathbf{R}(\Phi_k)) \right|$. Then, set
- 4 $\Phi_{k+1} = \Phi_k - \lambda^* \mathbf{J}(\Phi_k)^{-1} \mathbf{R}(\Phi_k)$.
- 5 Step B5: If the relative error $|\Phi_{k+1} - \Phi_k| / |\Phi_k|$, is less than an acceptable error ε , then
- 6 terminate, and Φ_{k+1} is the solution.
- 7 Step B6: If the relative error exceeds the acceptable error ε then replace Φ_k with Φ_{k+1} . Set k
- 8 $= k + 1$ and go to Step B2.
- 9 The no-toll scenario considered in the numerical example, which is a special case of the
- 10 cordon-based congestion-charging model with $\tilde{T}_{mn} = 0$, will also adopt the above solution
- 11 algorithm but skip the toll determination step (Step B2).
- 12

13 4.4. The second-best solution

14 When setting up a cordon-based congestion-pricing scheme, two critical issues that have to

15 be addressed are the location of the cordons and their corresponding toll levels. Section 4.2

16 introduces a method for fixing the location of charging cordons from the first-best solution.

17 However, the fixed tolls always undercharge users because they are bounded by the first-best

18 level of toll charges. To fine tune the cordon-based charging levels, given the cordons set

19 used in the above procedure remain unchanged, we can derive the second-best congestion

20 charging scheme by further optimizing the toll values using the following bi-level

21 mathematical program. The upper-level subprogram is designed to maximize the total utility

22 of the system.

$$23 \quad \text{Maximize}_{\tilde{\tau}} \quad \tilde{U}(\tilde{\tau}) = \iint_{\Omega} \sum_m \sum_{\tilde{n}} q_{m\tilde{n}}^*(\tilde{\tau}) U_{m\tilde{n}} + \sum_m \frac{1}{\zeta_m} q_m \ln q_m -$$

$$\sum_m \sum_{\tilde{n}} \frac{1}{\zeta_m} q_{m\tilde{n}}^*(\tilde{\tau}) \ln q_{m\tilde{n}}^*(\tilde{\tau}) - \sum_m \sum_n c_m^*(\tilde{\tau}) \mathbf{f}_{mn}^*(\tilde{\tau}) \, d\Omega \quad (32a)$$

$$24 \quad \text{subject to} \quad \tau_m^j > 0, \quad \forall m \in M, j \in J' \quad (32b)$$

25 where the quantities $q_{m\tilde{n}}^*(\tilde{\tau})$, $c_m^*(\tilde{\tau})$ and $\mathbf{f}_{mn}^*(\tilde{\tau})$ are determined by solving the set of

26 differential equations (25) at the lower level, in which the set of fixed tolls $\tilde{\tau}$ is passed from

27 the upper level to the lower level subprogram. Given that in practice only very few cordons

28 are introduced or the number of fixed toll variables is very small, the upper level subprogram

29 can easily be solved using the pattern search method (Hookes and Jeeves, 1961). Note that

30 the resultant second-best congestion charging scheme is optimized in the context of the

31 chosen cordon locations. A better second-best scheme can be obtained by joint optimization

32 of the cordon locations and toll levels, which is a useful extension of the present methodology

33 and is worthy for exploration in a future study.

34

35

36 5. A numerical example

37 We now present a numerical example to illustrate the proposed congestion-pricing model

38 with multiple user classes within a continuum system. The Chinese region modeled is shown

39 in Figure 2. The region spans about 350 km from east to west and 250 km from north to south,

40 and has two cities competing for users of two different vehicle classes. Note that the vehicle

41 classes used in this example is for demonstration purpose, and thus they are not referred to

1 any real categories of vehicles. For the sake of simplicity, it is assumed that the users are
 2 uniformly distributed over this region. In a real-life application, the demand surface may vary
 3 over space according to the zonal population density and other demographic and socio-
 4 economic characteristics of the population specified by the modelers. The demand for class 1
 5 and 2 users is respectively taken as 50 users/km² and 60 users/km², while the corresponding
 6 demand distribution functions for the different destination choices are specified as:

$$7 \quad \text{Class 1 non-travel: } q_{10} = 50 \frac{\exp[0.008(100)]}{\exp[0.008(100)] + \exp[0.008(200 - \bar{u}_{11})] + \exp[0.008(310 - \bar{u}_{12})]}$$

$$8 \quad \text{Class 1 to city 1: } q_{11} = 50 \frac{\exp[0.008(200 - \bar{u}_{11})]}{\exp[0.008(100)] + \exp[0.008(200 - \bar{u}_{11})] + \exp[0.008(310 - \bar{u}_{12})]}$$

$$9 \quad \text{Class 1 to city 2: } q_{12} = 50 \frac{\exp[0.008(310 - \bar{u}_{12})]}{\exp[0.008(100)] + \exp[0.008(200 - \bar{u}_{11})] + \exp[0.008(310 - \bar{u}_{12})]}$$

$$10 \quad \text{Class 2 non-travel: } q_{20} = 60 \frac{\exp[0.006(150)]}{\exp[0.006(150)] + \exp[0.006(260 - \bar{u}_{21})] + \exp[0.006(380 - \bar{u}_{22})]}$$

$$11 \quad \text{Class 2 to city 1: } q_{21} = 60 \frac{\exp[0.006(260 - \bar{u}_{21})]}{\exp[0.006(150)] + \exp[0.006(260 - \bar{u}_{21})] + \exp[0.006(380 - \bar{u}_{22})]}$$

$$12 \quad \text{Class 2 to city 2: } q_{22} = 60 \frac{\exp[0.006(380 - \bar{u}_{22})]}{\exp[0.006(150)] + \exp[0.006(260 - \bar{u}_{21})] + \exp[0.006(380 - \bar{u}_{22})]}$$

13 The demand function is dependent on the total travel cost, \bar{u}_{mn} , incurred in traveling
 14 from the demand locations continuously dispersed on the 2-dimensional plane to a common
 15 and compact destination in the city. The unit transportation cost function is specified as:

$$16 \quad \text{Class 1 user: } c_1(x, y) = 0.07v_a(x, y) + v_b(x, y)[0.00004(|\mathbf{f}_{11}| + |\mathbf{f}_{12}|) + 0.00005(|\mathbf{f}_{21}| + |\mathbf{f}_{22}|)].$$

$$17 \quad \text{Class 2 user: } c_2(x, y) = 0.08v_a(x, y) + v_b(x, y)[0.00006(|\mathbf{f}_{11}| + |\mathbf{f}_{12}|) + 0.00007(|\mathbf{f}_{21}| + |\mathbf{f}_{22}|)].$$

18 where $c_m(x, y)$ is measured in CNY (Chinese yuan) per kilometer.
 19 $v_a(x, y) = 1.10 - 0.00025\bar{d}(x, y)$ and $v_b(x, y) = 1.20 - 0.00035\bar{d}(x, y)$ are the factors that
 20 account for the variation in the location-dependent parameters of the unit transportation cost
 21 function and $\bar{d}(x, y)$ is the mean distance from the two cities to locations (x, y) . These factors
 22 increase when the mean distance from the cities decreases, which reveals the network
 23 characteristic that junctions are more closely spaced as they are nearer to the cities. Hence,
 24 the parameters of the unit transportation cost function increase. The cost function is flow-
 25 dependent and explicitly takes into account the fact that traffic congestion will lead to a
 26 higher travel cost through the opportunity cost of waiting in queues or the payment of a
 27 congestion toll. Given the above definition of the problem, the study region is discretized into
 28 a finite element mesh (see Figure 3), and the continuum approach is adopted to solve the
 29 traffic equilibrium problems in no-toll, first-best, initial-cordon and second-best scenarios.

30 First, let us start with the no-toll scenario, which is also known as the user optimal
 31 model. In this scenario, all users choose their optimal routes over the continuum with the
 32 lowest transportation cost (no toll). The resultant traffic flow pattern for class 1 users
 33 traveling to city 1 is displayed in Figure 4. In this paper, only figures for class 1 and city 1 are
 34 included, as the solution patterns for the other combinations of user classes and cities are
 35 quite similar. Certain indicators associated with this no-toll scenario are calculated and listed

1 in Table 1. For this no-toll scenario, we find that city 2 has a higher market share (51% for
 2 class 1 users and 45% for class 2 users) as city 2 has a higher utility gain level, U_{m2} , (310
 3 CNY for class 1 users and 380 CNY for class 2 users) as compared to city 1, U_{m1} , and non-
 4 travel option, U_{m0} . The total system cost is 506.10 million CNY and the total system utility is
 5 about 1,356.39 million CNY in this no-toll scenario. The resultant traffic flow intensity for
 6 class 1 users traveling to city 1 is displayed in Figure 5. The flow intensity increases rapidly
 7 as traffic approaches city 1, which is also true for city 2 and class 2 users, resulting in a steep
 8 increase in the total travel cost incurred by users. Figure 5 shows that in the eastern part of
 9 the modeled region (around city 2), the flow intensities around the northern and southern
 10 boundaries of the region are higher than that in the vicinity of city 2. This indicates that some
 11 of the users, especially those located to the east of city 2, are diverted away from the vicinity
 12 of city 2 while heading to city 1. Such behavior can be explained by the interactions of
 13 different users within the region. As the vicinity of city 2 is congested with the users who
 14 choose to patronize city 2, the delay due to congestion is very high. Therefore, those who
 15 choose to use city 1 try to avoid the congested area near city 2 and take a longer but less
 16 congested route.

17 Figure 6 shows the contour of the total travel cost of class 1 users, where the boundary
 18 of city 1 represents the contour of zero total travel cost for class 1 users traveling to city 1. By
 19 comparing the spacing between contours in Figure 6, it can be seen that the change of total
 20 travel cost with respect to the change of demand location is more rapid in the proximity of
 21 city 1 and 2 than in the other locations. Such rapid change of total travel cost in these
 22 proximities is resulted from the high flow intensities (i.e. heavy congestion) in those areas.
 23 Turning to the proximity of city 2, Figures 5 and 6 show that while the flow intensity is
 24 relatively low, the travel cost contour is highly distorted. This is because in the region of low
 25 flow intensity of class 1 users traveling to city 1, the effect of round-off and convergence
 26 error of the flow vector (\mathbf{f}_{11}) is relatively large. These errors, which affect the direction of the
 27 cost gradient (Equation 8), will be amplified by the high travel cost, which due to the
 28 congestion of this area with the traffic heading to city 2, and causes the irregularities of the
 29 iso-cost contours. Despite the above non-smooth (distorted) iso-cost contours, Figure 6 gives
 30 a reasonable pattern of total travel cost as the highest travel cost (over 160 CYN) is located
 31 around city 2 for the high congestion delay in that area.

32 Traffic congestion is particularly severe in the no-toll scenario, especially as traffic
 33 approaches the cities. As there is no congestion toll, all road users spend their time waiting in
 34 queues rather than using it fruitfully for other purposes. To maximize the total utility of the
 35 modeled system, a congestion pricing scheme should be introduced. In this numerical
 36 example, the continuum modeling approach is applied to find the first-best solution of utility
 37 maximization (using the mathematical program expressed in (6)). Following the imposition
 38 of congestion tolls, a portion of the congestion externality is internalized by road users. Given
 39 the regional configuration, the congestion toll for each class of users is defined by the toll rate
 40 defined in Equation (18). The results of the first-best solution are presented in Table 1. In the
 41 first-best scenario, the toll rate, τ_m , varies continuously over space according to location and
 42 traffic volume. Geographically, the toll rate can be shown by the iso-toll rate contours and the
 43 contours for all class 1 users (for travel to both city 1 and city 2) are shown in Figure 7. As
 44 traffic intensity rises sharply near the cities, the iso-toll contours become very densely packed
 45 in these locations. The maximum toll rates for class 1 and class 2 users, which are
 46 respectively 5.38 and 6.42 CNY per km (Table 1), occur around city 2, as this is the most
 47 congested area (and is even more congested than the proximity of city 1). For road users
 48 departing from any location (x, y) within the region to a city, the total toll paid T_{mn} can be
 49 calculated by solving problem (23). The iso-toll contours of this total toll pattern for class 1
 50 users traveling to city 1 are displayed in Figure 8. In Figure 8, it can be seen that the total toll,

1 T_{mn} , is a location-dependent variable and that the highest total toll is paid by those living
 2 farthest away from the city because they travel for a longer distance within the region.

3 As discussed in section 4.1, the charging cordons are defined on the basis of the iso-
 4 contours of the total toll pattern T_{mn} from the first-best scenario. For the sake of simplicity,
 5 we choose two iso-contours from the total toll pattern for class 1 users as the charging
 6 cordons for the initial-cordon and second-best scenarios. The exact locations of these
 7 charging cordons are shown in Figure 9. As these cordons are the iso-contours of the total
 8 toll pattern for class 1 users, the tolls charged to class 1 users at these cordons, which are 30
 9 CNY for cordon 1 and 45 CNY for cordon 2, can be calculated easily. For class 2 users,
 10 although these cordons are not exactly the same as their total toll iso-contours, they should be
 11 very similar, as the toll rates for both class 1 and class 2 users are dependent on the flow
 12 intensity at any particular location. Thus, based on the first-best solution, we take 35 CNY
 13 and 55 CNY as the toll charges for class 2 users at cordons 1 and 2, respectively. In this
 14 numerical example, it is assumed that only the inbound direction of the cordons is charged
 15 with the designed toll. Based on this cordon-based charging scheme from the first-best
 16 scenario, a new user equilibrium solution, which forms the initial-cordon scenario, can be
 17 obtained. The results are summarized in Table 1. Although the improvement in total system
 18 utility of the initial-cordon scheme (57.90%) is not as high as that of the first-best solution
 19 (100%), it still leads to an improvement in total system utility when compared to the no-toll
 20 scenario. In Table 1, it can be seen that while the percentage of non-travel cases and the total
 21 utility of the initial-cordon scheme are both higher than in the no-toll scenario, they are lower
 22 than in the first-best case. This indicates that although the initial-cordon charging scheme is
 23 both an efficient way of alleviating congestion and improves the total utility of the system, it
 24 is not as effective as the point-wise charging scheme (the first-best scenario). Similar
 25 conclusion could also be drawn by comparing the total system utility of the initial-cordon
 26 scenario (1.38 billion CNY) to that of the first-best scenario (1.40 billion CNY).

27 In the initial-cordon scheme, as the system users are only charged at the cordons and the
 28 toll is the same as in the first-best solution, this does not maximize the total utility of the
 29 system as a whole. In other words, this user equilibrium solution should not be the second-
 30 best solution under the constraint of cordon-based charging at the selected cordons. Hence, it
 31 is reasonable to find the second-best solution by adjusting the cordon toll levels. The pattern
 32 search method introduced by Hookes and Jeeves (1961) is thus adopted to find the
 33 combination of cordon tolls that maximizes total utility under this cordon-based charging
 34 scheme. It is found that when total utility is maximized, the charges at cordon 1 are,
 35 respectively, 42 CNY and 43 CNY for class 1 and class 2 users, while at cordon 2, the
 36 charges are 67 CNY and 63 CNY. This forms the second-best scenario of this numerical
 37 example involving charges at cordons 1 and 2. Table 1 gives the detailed numerical results
 38 for this second-best scenario.

39 Compared to the initial-cordon scenario, the second-best scenario is associated with an
 40 increase in the total system utility of the system of 6.42 million CNY (to a level of 1,386.39
 41 million CNY). The system cost falls slightly from 440.85 million CNY to 420.05 million
 42 CNY, and total toll revenue rises from 180.25 million CNY to 227.53 million CNY. Similar
 43 to the comparison of the initial-cordon scenario with the no-toll and first-best scenarios, the
 44 results of the second-best charging scenario indicate that while it is superior to the initial-
 45 cordon scenario in alleviating congestion, it is not as good as the first-best scenario. Also, by
 46 considering the improvement in total system utility, it could be seen that the second-best
 47 scenario (73.65%) is better than the initial-cordon scenario (57.9%) but is inferior to the first-
 48 best scenario (100%). Figure 10 shows the pattern of total tolls paid by class 1 users traveling
 49 to city 1 in the second-best charging scenario. Comparing Figure 10 with Figure 9, it can be
 50 seen that all class 1 users that are not located within cordon 1 have to pay a 42 CNY toll,

1 which is the toll at cordon 1 for class 1 users, when they travel from their home locations to
 2 city 1. This indicates that these users attempt to avoid entering cordon 2, which requires an
 3 extra toll of 67 CNY, when they devise their paths. While from some locations (such as to the
 4 east of cordon 2), the path is shorter if the user chooses to pass through cordon 2, this saving
 5 in travel distance cannot compensate for the extra toll incurred when passing through cordon
 6 2 and the long delay experienced due to the heavy congestion near city 2. Thus, users are
 7 likely to forego the advantage of a shorter route and detour to another path that, while longer,
 8 is nevertheless toll-free and less congested. This demonstrates the usefulness of using a more
 9 controllable and effective cordon-based congestion charging scheme, in contrast to the less
 10 predictable congestion delay experienced by road users, to detour traffic away from a
 11 congested area.

12 Due to the difference in geographical locations and choices of destination, the impact of
 13 congestion-pricing varies over the modeling region. In this study, the impact of congestion-
 14 pricing is measured as the difference of user surplus, which is the logsum of user surpluses of
 15 all travel choices ($\ln \sum_n \exp(U_{mn} - \bar{u}_{mn})$), of each class of users before and after the
 16 implementation of scheme. Figure 11 and 12 show the spatial variation of percentage change
 17 of user surplus for class 1 users after the implementation of the first-best and second-best
 18 congestion-pricing schemes, respectively. In Figure 11, it can be seen that the user surplus of
 19 class 1 users decreases (≤ 0) for the whole modeling region after the first-best congestion
 20 pricing scheme is implemented. This is because under the first-best congestion pricing
 21 scheme, all the users, who choose to travel, have to pay an extra distance-based toll that
 22 cannot be compensated for by the travel time reduction from the less congested network. Due
 23 to the shorter travel distance (less paid toll in the first-best charging scheme), the percentage
 24 decrease in user surplus reduces for the locations that are closer to either of the destinations
 25 (City 1 and City 2). In contrast, the highest reduction of user surplus of class 1 users is at the
 26 locations that are between the two destinations (i.e. locations that are not close to either of the
 27 destinations). On the other hand, the second-best congestion pricing scheme gives a different
 28 pattern of percentage change of user surplus as compared to the first-best scenario (Figure 12).
 29 In general, the reduction of user surplus for class 1 users in the second-best scenario is less
 30 than that in the first-best scenario. This is because the cordon toll that users pay in the
 31 second-best scenario is generally less than the distance-based toll in the first-best scenario. In
 32 Figure 12, it can be seen that there is an increase in the user surplus for class 1 users located
 33 within the charging cordons. Such increase comes from the fact that users within the charging
 34 cordons enjoy a less congested, which due to the implementation of congestion pricing, and
 35 toll free travel environment. In Figure 12, it can be observed that the location for the highest
 36 reduction in user surplus is closer to city 1. Based on the percentage change of user surplus
 37 shown in Figure 11 and 12, it is expected that after the implementation congestion pricing,
 38 class 1 users tend to live closer to the destinations (first-best scheme) or inside the charging
 39 cordons (second-best scheme). As a result, the rent/property value in those areas increases.

40 41 **6. Conclusions**

42 We have proposed a congestion-pricing model for a continuum transportation system with
 43 multiple cities and vehicle types (user classes). In this study, two different models, namely
 44 the utility maximization model and the cordon-based congestion-pricing model, have been
 45 considered and solved under the assumptions of the continuum modeling approach. For the
 46 utility maximization model, which is also known as the first-best solution for this congestion-
 47 pricing problem, it is found that different toll rate patterns for different user class and
 48 destination combinations maximize the utility of travel patterns within the system. Based on
 49 this first-best solution, an initial-cordon scheme has been constructed to locate the cordons

1 and determine the corresponding set of fixed tolls which are based purely on user class. Using
2 this initial-cordon scheme, a bi-level programming problem has also been formulated and
3 solved by way of a Hookes and Jeeves pattern search algorithm to identify the set of cordon
4 toll levels for different user classes, thereby maximizing total utility and leading to the
5 second-best scenario. In all of these problems, the continuum models have been solved using
6 the efficient finite element method. A numerical example has been given to compare and
7 contrast the efficiency of the utility maximization model and the cordon-based congestion-
8 pricing model. It has been found that the second-best charging scheme, in spite of the simple
9 charging method it incorporates, delivers total system utility comparable to that of the first-
10 best scenario.

11 The use of a continuum approach can help to intuitively identify the level of congestion
12 and the external cost throughout the region, thereby obviating the need for detailed network
13 modeling work and exhaustive evaluations at the early stage of planning. Moreover, this type
14 of intuitive approach allows traffic planners to select with ease one or more toll cordon(s)
15 over space and evaluate the impact of cordon toll charges on the resultant total utility and toll
16 revenue. This will provide insightful information that can be used to inform the future
17 formulation of a detailed cordon scheme within a discrete network.

18 19 20 **Acknowledgements**

21 The work described in this paper was supported by grants from the Research Grants Council of the
22 Hong Kong Special Administrative Region, China (Project No. HKU7126/04E, HKU7183/08E and
23 POLYU5261/07E) and post-doctoral fellowship from The Hong Kong Polytechnic University.
24
25

1 **Appendix A**

2
3 This appendix gives the proof of the direct utility of the representative traveler for a multi-class
4 congestion-pricing problem with logit-type demand distribution. Based on this direct utility, the
5 mathematical program that maximizes the total system utility of this problem is derived. For the
6 congestion-pricing problem with a single user class and destination (Ho et al., 2005), the problem is
7 formulated as a maximization of the total system benefit. For that single user class and destination
8 case, the system benefit can be defined explicitly as the difference between the user benefit and the
9 system cost. However, due to the introduction of the random term in the total perceived cost, the total
10 system benefit cannot be explicitly defined. To set up the optimization problem for this logit-type
11 demand distribution congestion-pricing problem, the idea of maximizing the utility of a representative
12 traveler (R.T.), which was introduced in Oppenheim (1995), is adopted. Consider a class m user
13 making a choice \tilde{n} from among a set of cities and the choice of non-travel. His perceived utility (or
14 indirect utility) can be specified as:

$$15 \quad V_{m\tilde{n}} = U_{m\tilde{n}} - \bar{u}_{m\tilde{n}} + \varepsilon_{m\tilde{n}}, \quad (\text{A-1})$$

16 where $U_{m\tilde{n}}$ is the utility gain of this class m user making a choice \tilde{n} ; $\bar{u}_{m\tilde{n}}$ is the total travel cost
17 (including the toll, if any) of that user making this choice; $\varepsilon_{m\tilde{n}}$ is the i.i.d. Gumbel distributed random
18 term. Based on this indirect utility, the corresponding expected utility (\bar{V}_m) for class m users is
19 defined by:

$$20 \quad \bar{V}_m = \frac{1}{\zeta_m} \ln \sum_{\tilde{n}} \exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}})]. \quad (\text{A-2})$$

21 Using this expected utility function, the corresponding direct utility can be found by minimizing
22 the total expected utility of all system users (Oppenheim, 1995):

$$23 \quad \text{Minimize}_{\bar{\mathbf{u}}} \bar{U}(\bar{\mathbf{u}}) = \iint_{\Omega} \sum_m \frac{q_m}{\zeta_m} \ln \sum_{\tilde{n}} \exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}})] d\Omega, \quad (\text{A-3a})$$

$$24 \quad \text{subject to} \quad \sum_{\tilde{n}} \bar{u}_{m\tilde{n}} q_{m\tilde{n}} + C_m - B_m = 0 \quad \forall m \in M, (x, y) \in \Omega, \quad (\text{A-3b})$$

25 where $\bar{\mathbf{u}} = \text{Col}(\bar{u}_{m\tilde{n}}, m \in M, \tilde{n} \in \tilde{N})$; C_m is the total expenditure on non-travel related items for all
26 class m users; Constraint (A-3b) is the budgetary constraint with B_m being the total budget of class m
27 users. Consider:

$$28 \quad \sum_{\tilde{n}} q_{m\tilde{n}} - q_m = 0 \quad \forall m \in M, (x, y) \in \Omega. \quad (\text{A-4})$$

29 This is the definition of the demand q_m and will also be used in finding the direct utility. Defining
30 $\tilde{u}_{m\tilde{n}} = \exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}})]$, we have

$$31 \quad \bar{u}_{m\tilde{n}} = U_{m\tilde{n}} - \frac{1}{\zeta_m} \ln \tilde{u}_{m\tilde{n}}, \quad (\text{A-5})$$

32 and Equation (A-3a) becomes

$$33 \quad \bar{U}(\tilde{\mathbf{u}}) = \iint_{\Omega} \sum_m \frac{q_m}{\zeta_m} \ln \sum_{\tilde{n}} \tilde{u}_{m\tilde{n}} d\Omega. \quad (\text{A-6})$$

34 As a natural logarithm is an increasing function, problem (A-3) can be rewritten as

$$35 \quad \text{Minimize}_{\tilde{\mathbf{u}}} \bar{U}(\tilde{\mathbf{u}}) = \iint_{\Omega} \sum_m \sum_{\tilde{n}} \frac{q_m}{\zeta_m} \tilde{u}_{m\tilde{n}} d\Omega, \quad (\text{A-7a})$$

$$1 \quad \text{subject to} \quad \sum_{\tilde{n}} q_{m\tilde{n}} \left(U_{m\tilde{n}} - \frac{1}{\zeta_m} \ln \tilde{u}_{m\tilde{n}} \right) + C_m - B_m = 0 \quad \forall m \in M, (x, y) \in \Omega, \quad (\text{A-7b})$$

2 where $\tilde{\mathbf{u}} = \text{Col}(\tilde{u}_{m\tilde{n}}, m \in M, \tilde{n} \in \tilde{N})$. Due to i) the identical set of feasible solution for problem (A-3)
3 and (A-7) and; ii) the one-to-one mapping and strictly increasing property of natural logarithm,
4 solution that minimize problem (A-7) will also minimize problem (A-3), and vice versa. Consider the
5 Lagrangian of the above minimization problem:

$$6 \quad L = \iint_{\Omega} \sum_m \sum_{\tilde{n}} \frac{q_m}{\zeta_m} \tilde{u}_{m\tilde{n}} - \sum_m \nu_m \left[\sum_{\tilde{n}} q_{m\tilde{n}} \left(U_{m\tilde{n}} - \frac{1}{\zeta_m} \ln \tilde{u}_{m\tilde{n}} \right) + C_m - B_m \right] d\Omega,$$

7 where ν_m is the Lagrangian multipliers for constraints (A-7b). Considering the first-order conditions of
8 the above Lagrangian, we have

$$9 \quad \frac{q_m}{\zeta_m} + \frac{\nu_m q_{m\tilde{n}}}{\zeta_m \tilde{u}_{m\tilde{n}}} = 0 \quad \forall m \in M, \tilde{n} \in \tilde{N}, (x, y) \in \Omega, \quad (\text{A-8})$$

$$10 \quad \sum_{\tilde{n}} q_{m\tilde{n}} \left(U_{m\tilde{n}} - \frac{1}{\zeta_m} \ln \tilde{u}_{m\tilde{n}} \right) + C_m - B_m = 0 \quad \forall m \in M, (x, y) \in \Omega, \quad (\text{A-9})$$

11 From Equation (A-8), we have

$$12 \quad \tilde{u}_{m\tilde{n}} = -\frac{\nu_m q_{m\tilde{n}}}{q_m}. \quad (\text{A-10})$$

13 Substituting Equation (A-10) into Equation (A-9), using Equation (A-4) and after rearranging we have:

$$14 \quad \sum_{\tilde{n}} q_{m\tilde{n}} U_{m\tilde{n}} + \frac{1}{\zeta_m} q_m \ln q_m - \frac{1}{\zeta_m} \sum_{\tilde{n}} q_{m\tilde{n}} \ln q_{m\tilde{n}} + (C_m - B_m) = \frac{q_m}{\zeta_m} \ln \sum_{\tilde{n}} \tilde{u}_{m\tilde{n}}.$$

15 Thus, using the definition of $\tilde{u}_{m\tilde{n}}$, the following equation holds at the optimal point of the
16 minimization problem (A-7),

$$17 \quad \frac{q_m}{\zeta_m} \ln \sum_{\tilde{n}} \exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}})] = \sum_{\tilde{n}} q_{m\tilde{n}} U_{m\tilde{n}} + \frac{1}{\zeta_m} q_m \ln q_m - \frac{1}{\zeta_m} \sum_{\tilde{n}} q_{m\tilde{n}} \ln q_{m\tilde{n}} + (C_m - B_m). \quad (\text{A-11})$$

18 Summing up Equation (A-11) for all classes of users, we get

$$19 \quad \sum_m \frac{q_m}{\zeta_m} \ln \sum_{\tilde{n}} \exp[\zeta_m (U_{m\tilde{n}} - \bar{u}_{m\tilde{n}})] = \sum_m \sum_{\tilde{n}} q_{m\tilde{n}} U_{m\tilde{n}} + \sum_m \frac{1}{\zeta_m} q_m \ln q_m - \sum_m \sum_{\tilde{n}} \frac{1}{\zeta_m} q_{m\tilde{n}} \ln q_{m\tilde{n}} + \sum_m (C_m - B_m). \quad (\text{A-12})$$

20 As problem (A-3) is equivalent to problem (A-7) and Equation (A-12) holds at the optimal solution,
21 the right-hand side of Equation (A-12) gives the direct utility of the representative traveler. To solve
22 for the distributed demand, $q_{m\tilde{n}}$, this direct utility function is maximized by the following problem:

$$23 \quad \text{Maximize}_{\mathbf{q}} \bar{U}(\mathbf{q}) = \iint_{\Omega} \sum_m \sum_{\tilde{n}} q_{m\tilde{n}} U_{m\tilde{n}} + \sum_m \frac{1}{\zeta_m} q_m \ln q_m - \sum_m \sum_{\tilde{n}} \frac{1}{\zeta_m} q_{m\tilde{n}} \ln q_{m\tilde{n}} + \sum_m (C_m - B_m) d\Omega \quad (\text{A-13a})$$

$$24 \quad \text{subject to} \quad \sum_{\tilde{n}} q_{m\tilde{n}} - q_m = 0 \quad \forall m \in M, (x, y) \in \Omega \quad (\text{A-13b})$$

$$26 \quad \sum_{\tilde{n}} \bar{u}_{m\tilde{n}} q_{m\tilde{n}} + C_m - B_m = 0 \quad \forall m \in M, (x, y) \in \Omega \quad (\text{A-13c})$$

$$1 \quad q_{m\tilde{n}} \geq 0 \quad \forall m \in M, \tilde{n} \in \tilde{N}, (x, y) \in \Omega, \quad (\text{A-13d})$$

2 where $\mathbf{q} = \text{Col}(q_{m\tilde{n}}, m \in M, \tilde{n} \in \tilde{N})$. By i) considering the fact that the total travel cost $\bar{u}_{m\tilde{n}}$ vanishes for
 3 the choice of “non-travel”; and ii) substituting Equation (A-13c) into (A-13a), we have:

$$4 \quad \text{Maximize}_{\bar{\mathbf{u}}, \mathbf{q}} \bar{U}(\bar{\mathbf{u}}, \mathbf{q}) = \iint_{\Omega} \sum_m \sum_{\tilde{n}} q_{m\tilde{n}} U_{m\tilde{n}} + \sum_m \frac{1}{\zeta_m} q_m \ln q_m - \sum_m \sum_{\tilde{n}} \frac{1}{\zeta_m} q_{m\tilde{n}} \ln q_{m\tilde{n}} - \sum_m \sum_n \bar{u}_{mn} q_{mn} \, d\Omega$$

5 (A-14a)

$$6 \quad \text{subject to} \quad \sum_{\tilde{n}} q_{m\tilde{n}} - q_m = 0 \quad \forall m \in M, (x, y) \in \Omega \quad (\text{A-14b})$$

$$7 \quad q_{m\tilde{n}} \geq 0 \quad \forall m \in M, \tilde{n} \in \tilde{N}, (x, y) \in \Omega, \quad (\text{A-14c})$$

8 For Equation (A-14a), the first term represents the weighted utility gain of making the relevant
 9 choice, using demand as the weight; the second and third term represents the utility gain by the R.T.
 10 due to the presence of the diversity of choice (Yang, 1999; Oppenheim, 1995); the sum of these three
 11 terms defines the expected direct utility gain by the R.T. for giving the set of choices. The fourth term
 12 of Equation (A-14a) is the total travel cost for all system users, which includes any tolls paid by such
 13 users. Considering this term, \bar{u}_{mn} includes some transferable costs, which are the tolls paid by the
 14 users, that are beneficial to the system but not to the users. Thus, maximizing the total direct utility for
 15 the R.T. within the system, as represented by Equation (A-14a), will not necessarily maximize the
 16 total utility of the system. To define the total system utility for maximization, the term
 17 $\iint_{\Omega} \sum_m \sum_n \bar{u}_{mn} q_{mn} \, d\Omega$ in Equation (A-14a) should be replaced by the total system cost,

18 $\iint_{\Omega} \sum_m \sum_n u_{mn} q_{mn} \, d\Omega$, where u_{mn} is the total travel cost (excluding tolls) for class m users patronizing
 19 city n . To make this problem easier to solve, the following replacement is made:

$$20 \quad \iint_{\Omega} \sum_m \sum_n u_{mn} q_{mn} \, d\Omega = \iint_{\Omega} \sum_m \sum_n c_m |\mathbf{f}_{mn}| \, d\Omega \quad (\text{A-15})$$

21 Equation (A-15) can be easily proven by a post-analysis least square problem based on the
 22 results (i.e. flow pattern and total travel cost) of the optimal solution. A complete proof can be found
 23 in Appendix B. By having the replacement in Equation (A-15), two additional constraints related to
 24 the flow vector \mathbf{f}_{mn} are introduced. Based on the above modifications, the mathematical program for
 25 maximizing the total system utility is defined as:

$$26 \quad \text{Maximize}_{\mathbf{f}, \mathbf{q}} \tilde{U}(\mathbf{f}, \mathbf{q}) = \iint_{\Omega} \sum_m \sum_{\tilde{n}} q_{m\tilde{n}} U_{m\tilde{n}} + \sum_m \frac{1}{\zeta_m} q_m \ln q_m - \sum_m \sum_{\tilde{n}} \frac{1}{\zeta_m} q_{m\tilde{n}} \ln q_{m\tilde{n}} - \sum_m \sum_n c_m |\mathbf{f}_{mn}| \, d\Omega$$

27 (A-16a)

$$28 \quad \text{subject to} \quad \nabla \mathbf{f}_{mn} - q_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Omega, \quad (\text{A-16b})$$

$$29 \quad q_m - \sum_{\tilde{n}} q_{m\tilde{n}} = 0 \quad \forall m \in M, \tilde{n} \in \tilde{N}, (x, y) \in \Omega, \quad (\text{A-16c})$$

$$30 \quad \mathbf{f}_{mn} = 0 \quad \forall m \in M, n \in N, (x, y) \in \Gamma. \quad (\text{A-16d})$$

31 As compare to problem (A-14), problem (A-16) had included flow vector as one of its decision
 32 variables. Thus, Equation (A-16b) and (A-16d) are added in problem (A-16) to respectively govern
 33 the conservation and boundary condition of these flow variables. In this problem, Constraint (A-14c)
 34 is removed, as it can be proven that this constraint will automatically satisfy at the optimal point.
 35

1 **Appendix B**

2

3 This appendix gives the proof of Equation (A-15) that is used in the formulation of the utility
4 maximization model discussed in section 3.1. Considering the L.H.S. of Equation (A-15) and
5 equations (9), (11) and (12), it can be seen that:

$$\begin{aligned}
 \iint_{\Omega} \sum_m \sum_n u_{mn}^* q_{mn}^* \, d\Omega &= \iint_{\Omega} \sum_m \sum_n u_{mn}^* \nabla \mathbf{f}_{mn}^* \, d\Omega \\
 &= \iint_{\Omega} \sum_m \sum_n \left[\nabla (u_{mn}^* \mathbf{f}_{mn}^*) - \mathbf{f}_{mn}^* \nabla u_{mn}^* \right] \, d\Omega \\
 &= \int_{\Gamma} \sum_m \sum_n u_{mn}^* \mathbf{f}_{mn}^* \cdot \mathbf{n} \, d\Gamma + \int_{\Gamma_c} \sum_m \sum_n u_{mn}^* \mathbf{f}_{mn}^* \cdot \mathbf{n} \, d\Gamma - \iint_{\Omega} \sum_m \sum_n \mathbf{f}_{mn}^* \nabla u_{mn}^* \, d\Omega \\
 &= - \iint_{\Omega} \sum_m \sum_n \mathbf{f}_{mn}^* \nabla u_{mn}^* \, d\Omega
 \end{aligned} \tag{B-1}$$

7 where the superscript indicates that the variables are at the solution of problem (6). Considering the
8 least square problem (23) for the total toll T_{mn} , a similar model for the total transportation cost (no toll)
9 u_{mn} can be found using the following post-analysis least square problem:

$$\text{Minimize}_{\mathbf{u}} \quad Z(\mathbf{u}) = \iint_{\Omega} \sum_m \sum_n \left| \nabla u_{mn} + c_m^* \frac{\mathbf{f}_{mn}^*}{|\mathbf{f}_{mn}^*|} \right|^2 \, d\Omega \tag{B-2a}$$

$$\text{subject to} \quad u_{mn} = 0, \quad \forall (x, y) \in \Gamma_c, \quad \forall m \in M, \quad \forall n \in N \tag{B-2b}$$

12 Therefore, at the optimal point of the above problem (B-2), we have:

$$\nabla u_{mn}^* = -c_m^* \frac{\mathbf{f}_{mn}^*}{|\mathbf{f}_{mn}^*|} \tag{B-3}$$

14 Putting Equation (B-3) into Equation (B-1), we have:

$$\begin{aligned}
 \iint_{\Omega} \sum_m \sum_n u_{mn}^* q_{mn}^* \, d\Omega &= \iint_{\Omega} \sum_m \sum_n \mathbf{f}_{mn}^* \cdot \left(c_m^* \frac{\mathbf{f}_{mn}^*}{|\mathbf{f}_{mn}^*|} \right) \, d\Omega \\
 &= \iint_{\Omega} \sum_m \sum_n c_m^* |\mathbf{f}_{mn}^*| \, d\Omega
 \end{aligned}$$

16 Therefore, Equation (A-15) holds at the optimal point of the minimization problem, thereby
17 completing the proof.

18

1 **References**

- 2
- 3 Beckmann, M.J., 1965. On optimal tolls for highways, tunnels and bridges. *In*: L.C. Edie, R.
4 Herman and R. Rothery, eds. *Vehicular Traffic Science*. New York: American Elsevier,
5 331-341.
- 6 Bellei, G., Gentile, G. and Papola, N., 2002. Network pricing optimization in multi-user and
7 multimodal context with elastic demand. *Transportation Research Part B*, 36, 779-798.
- 8 Blumenfeld, D.E., 1977. Modeling the joint distribution of home and workplace locations in a
9 city. *Transportation Science*, 11(4), 307-377.
- 10 Chen, Mei, Bernstein, David H. and Spasovic Laxar N., 2004. Toll-design problem stochastic
11 route choice. *Environment and Planning B*, 31, 731-742.
- 12 Cheung, Y.K., Lo, S.H. and Leung, A.Y.T., 1996. *Finite Element Implementation*. England:
13 Blackwell Science.
- 14 Commission of the European Community, 1999. Directive 1999/62/EC of the European
15 Parliament and of the Council of 17 June 1999 on the Charging of Heavy Goods
16 Vehicles for the Use of Certain Infrastructure, European Commission.
- 17 Cottingham, D.N., Beresford, A.R. and Harle, R.K., 2007. Survey of technologies for the
18 implementation of national-scale road user charging. *Transport Reviews*, 27(4), 499-523.
- 19 Dafermos, S.C. and Sparrow, F.T., 1971. Optimal resource allocation and toll patterns in
20 user-optimized transport networks. *Journal of Transport Economics and Policy*, 5, 198-
21 200.
- 22 Department for Transport, 2004. *Feasibility study of road pricing in UK – Full report*.
23 Department for Transport.
- 24 Gentile, G., Papola, N. and Persia, L., 2005. Advanced pricing and rationing policy for large
25 scale multimodal networks. *Transportation Research Part A*, 39, 612-631.
- 26 Geroliminis, N. and Daganzo C.F., 2008. Existence of urban-scale macroscopic fundamental
27 diagrams: Some experimental findings. *Transportation Research Part B*, 42, 759-550
- 28 Glaister, S. and Graham, D.J., 2005. An evaluation of national road user charging in England.
29 *Transportation Research Part A*, 39, 632-650.
- 30 Graham, D.J., Glaister, S., Quddus, M. and Wadud, Z., 2009. Testing for the distributional
31 effects of national road user charging. *International Journal of Sustainable*
32 *Transportation*, 3(1), 18-38.
- 33 Ho, H.W. and Wong, S.C., 2006. A review of the two-dimensional continuum modeling
34 approach to transportation problems. *Journal of Transportation Systems Engineering and*
35 *Information Technology*, 6, 53-72.
- 36 Ho, H.W., Wong, S.C., Yang, H. and Loo, B.P.Y., 2005. Cordon-based congestion pricing in
37 a continuum traffic equilibrium system. *Transportation Research Part A*, 39, 813-834.
- 38 Ho, H.W., Wong, S.C. and Loo, B.P.Y., 2006. Combined distribution and assignment model
39 for a continuum traffic equilibrium problem with multiple user classes. *Transportation*
40 *Research Part B*, 40, 633-650.
- 41 Ho, H.W., Wong, S.C. and Hau, T.D., 2007a. Existence and uniqueness of a solution for the
42 multi-class user equilibrium problem in a continuum transportation system.
43 *Transportmetrica*, 3, 107-117.
- 44 Ho, H.W., Wong, S.C. and Hau, T.D., 2007b A multi-class congestion pricing problem in a
45 continuum transportation system. *Journal of the Eastern Asia Society for Transportation*
46 *Studies*, 7, 238-253.
- 47 Hookes, R. and Jeeves, T.A., 1961. Direct search solution of numerical and statistical
48 problems. *J. Assoc. Comp. Machinery*, 8, 212-229.
- 49 Huang, H.J., 2000. Fares and tolls in a competitive system with transit and highways: the case
50 with two groups of commuters. *Transportation Research Part E*, 36, 267-284.

- 1 Huang, H.J., 2004. Pricing and logit-based mode choice models of a transit and highway
2 system with elastic demand. *European Journal of Operational Research*, 140, 562-570.
- 3 May, A.D., Shepherd, S.P. and Bates, J.J., 2000. Supply curves for urban road networks.
4 *Journal of Transport Economics and Policy*, 34, 261-290.
- 5 Oppenheim, N., 1995. *Urban Travel Demand Modeling: From Individual Choices to General*
6 *Equilibrium*. Canada: John Wiley & Sons, Inc..
- 7 Sasaki, T., Iida, Y. and Yang, H., 1990. User-equilibrium traffic assignment by continuum
8 approximation of network flow. *Proceedings of 11th International Symposium on*
9 *Transportation and Traffic Theory*, 233-252, Japan, Yokohama, July 1990.
- 10 Sheffi, Y., 1985. *Urban Transportation Networks*. U.S.A.: Prentice-Hall.
- 11 Smith, M.J., 1979. The marginal cost pricing of a transportation network. *Transportation*
12 *Research*, 13, 237-242.
- 13 Sumalee, A., 2007. Multi-concentric optimal charging cordon design. *Transportmetrica*, 3,
14 41-71.
- 15 Sumalee A., May A.D. and Shepherd, S.P., 2005. Comparison of judgmental and optimal
16 road pricing cordons. *Journal of Transport Policy*, 12(5), 384-390.
- 17 Vaughan, R.J., 1987. *Urban Spatial Traffic Patterns*. London: Pion.
- 18 Wong, S.C., 1998. Multi-commodity traffic assignment by continuum approximation of
19 network flow with variable demand. *Transportation Research Part B*, 32, 567-581.
- 20 Wong, S.C., Lee, C.K. and Tong, C.O., 1998. Finite element solution for the continuum
21 traffic equilibrium problems. *International Journal for Numerical Methods in*
22 *Engineering*, 43, 1253-1273.
- 23 Yang, H., 1999. System optimum, stochastic user equilibrium, and optimal link tolls.
24 *Transportation Science*, 33(4), 354-360.
- 25 Yang, H. and Huang, H.J., 1998. Principle of marginal-cost pricing: how does it work in a
26 general road network? *Transportation Research Part A*, 32, 45-54.
- 27 Yang, H. and Huang, H.J., 2004. The multiclass, multicriteria traffic network equilibrium and
28 system optimum problem. *Transportation Research Part B*, 38, 1-15.
- 29 Zhang, X. and Yang, H., 2004. The optimal cordon-based network congestion pricing
30 problem. *Transportation Research Part B*, 38, 517-537.
- 31 Zhang, Xiaoning, Zhang, H.M., Huang, Hai-Jun, Sun, Lijun and Tang, Tie-Qiao, 2011.
32 Competitive, cooperative and Stackelberg congestion pricing for multiple regions in
33 transportation networks. *Transportmetrica*, 7(4), 297-320.
- 34 Zienkiewicz, O.C. and Taylor, R.L., 1989. *The Finite Element Method*. International
35 Editions: McGraw-Hill.
- 36

1	List of tables and figures
2	
3	Table 1 Comparison of results for the no-toll, first-best, initial-cordon, and second-best
4	scenarios
5	Figure 1 Framework for the current study
6	Figure 2 The modeled region
7	Figure 3 Finite element discretization of the study region
8	Figure 4 Flow pattern of class 1 users traveling to city 1 in the no-toll scenario
9	Figure 5 Flow intensity of class 1 users traveling to city 1 in the no-toll scenario
10	Figure 6 Total travel cost for class 1 users traveling to city 1 in the no-toll scenario (CNY)
11	Figure 7 Contour plot of the distance-based iso-toll rate for class 1 users in the first-best
12	scenario (CNY/km)
13	Figure 8 Contour plot of the total toll paid by class 1 users traveling to city 1 in the first-best
14	scenario (CNY)
15	Figure 9 The two charging cordons
16	Figure 10 Total toll paid by class 1 users traveling to city 1 in the second-best charging
17	scenario
18	Figure 11 Percentage change of user surplus for class 1 users as first-best congestion-pricing
19	scheme is implemented
20	Figure 12 Percentage change of user surplus for class 1 users as second-best congestion-
21	pricing scheme is implemented
22	
23	
24	

Table 1. Comparison of results for the no-toll, first-best, initial-cordon, and second-best scenarios.

	No-toll scenario	First-best scenario	Initial-cordon scenario	Second-best scenario
Demand of class 1 users to city 1 ('000 veh/h)	797 (27%)	771 (26%)	790 (27%)	796 (27%)
Demand of class 1 users to city 2 ('000 veh/h)	1,535 (51%)	1,305 (44%)	1,410 (47%)	1,334 (45%)
Non-travelling class 1 users ('000 veh/h)	644 (22%)	900 (30%)	777 (26%)	847 (28%)
Demand of class 2 users to city 1 ('000 veh/h)	992 (28%)	935 (26%)	975 (27%)	961 (27%)
Demand of class 2 users to city 2 ('000 veh/h)	1,622 (45%)	1,390 (39%)	1,484 (42%)	1,476 (41%)
Non-travelling class 2 users ('000 veh/h)	958 (27%)	1,246 (35%)	1,113 (31%)	1,134 (32%)
Maximum toll rate for class 1 users (CNY/km)	-----	5.38	-----	-----
Maximum toll rate for class 2 users (CNY/km)	-----	6.42	-----	-----
Toll at cordon 1 for class 1 users $\tilde{\tau}_1^1$ (CNY)	-----	-----	30.00	42.00
Toll at cordon 1 for class 2 users $\tilde{\tau}_2^1$ (CNY)	-----	-----	35.00	43.00
Toll at cordon 2 for class 1 users $\tilde{\tau}_1^2$ (CNY)	-----	-----	45.00	67.00
Toll at cordon 2 for class 2 users $\tilde{\tau}_2^2$ (CNY)	-----	-----	55.00	63.00
Total tolls received (million CNY)	-----	320.69	180.26	227.53
Total system cost (million CNY)	506.10	352.96	440.85	420.05
Total system utility, \tilde{U} (million CNY)	1,356.39	1,397.12	1,379.97	1,386.39
Improvement in total system utility (%) (Relative to the improve of the first-best scenario)	-----	100.00	57.90	73.65

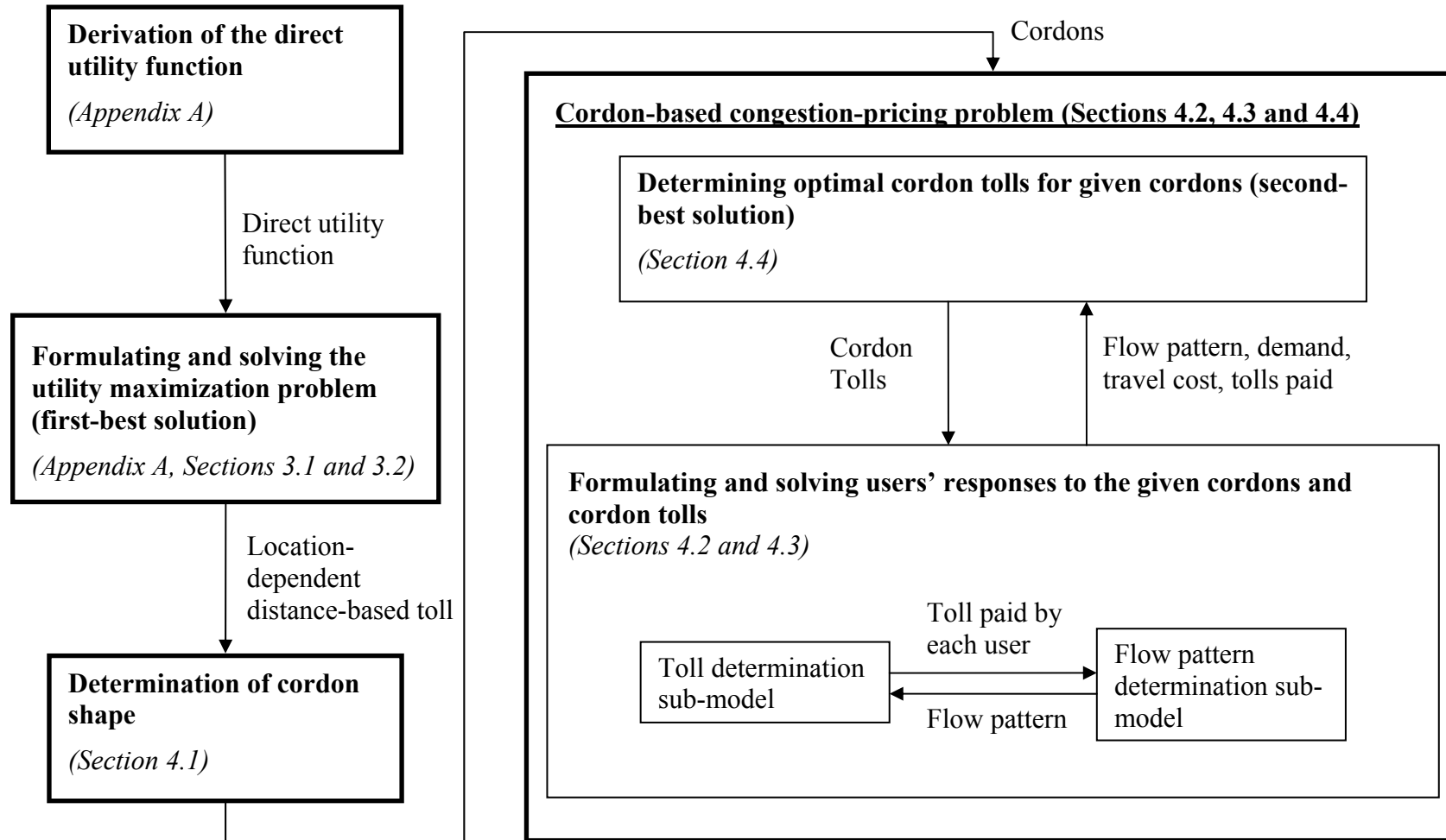


Figure 1. Framework for the current study.

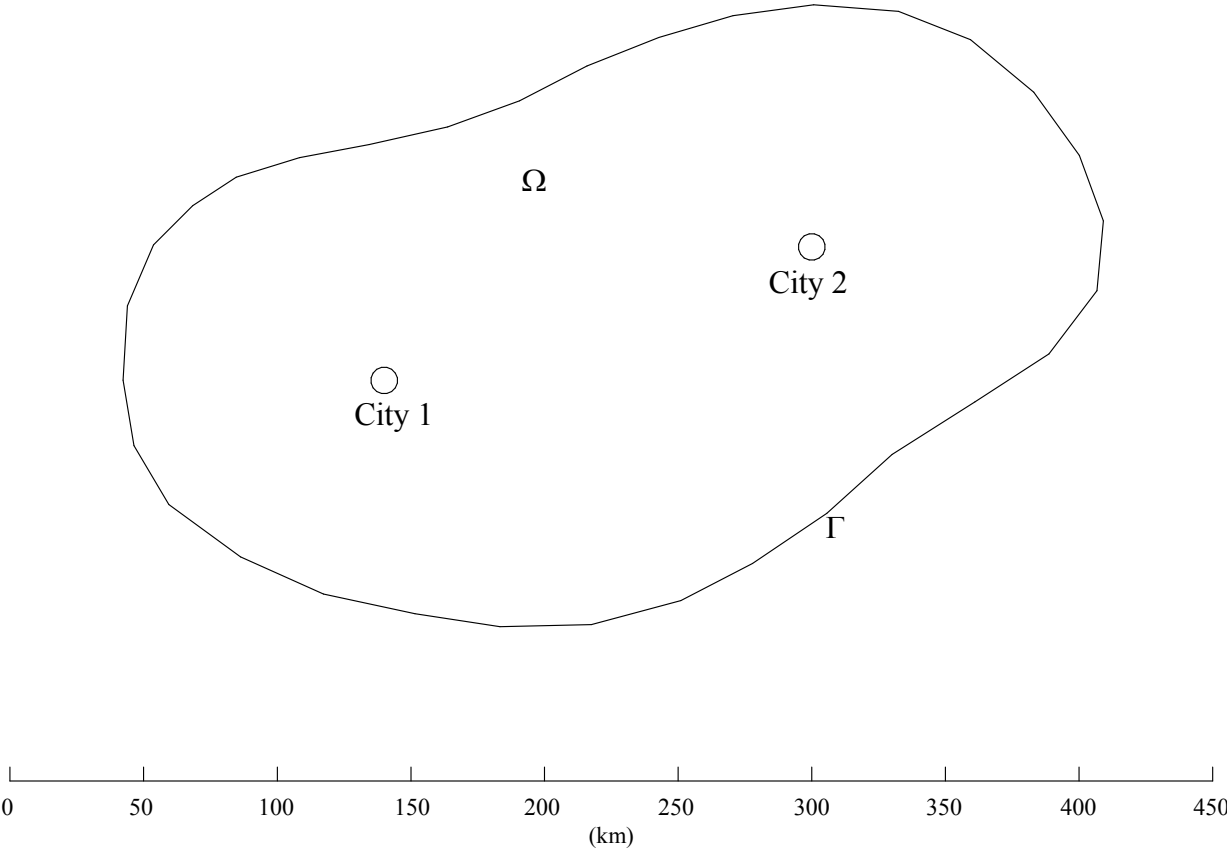


Figure 2. The modeled region.

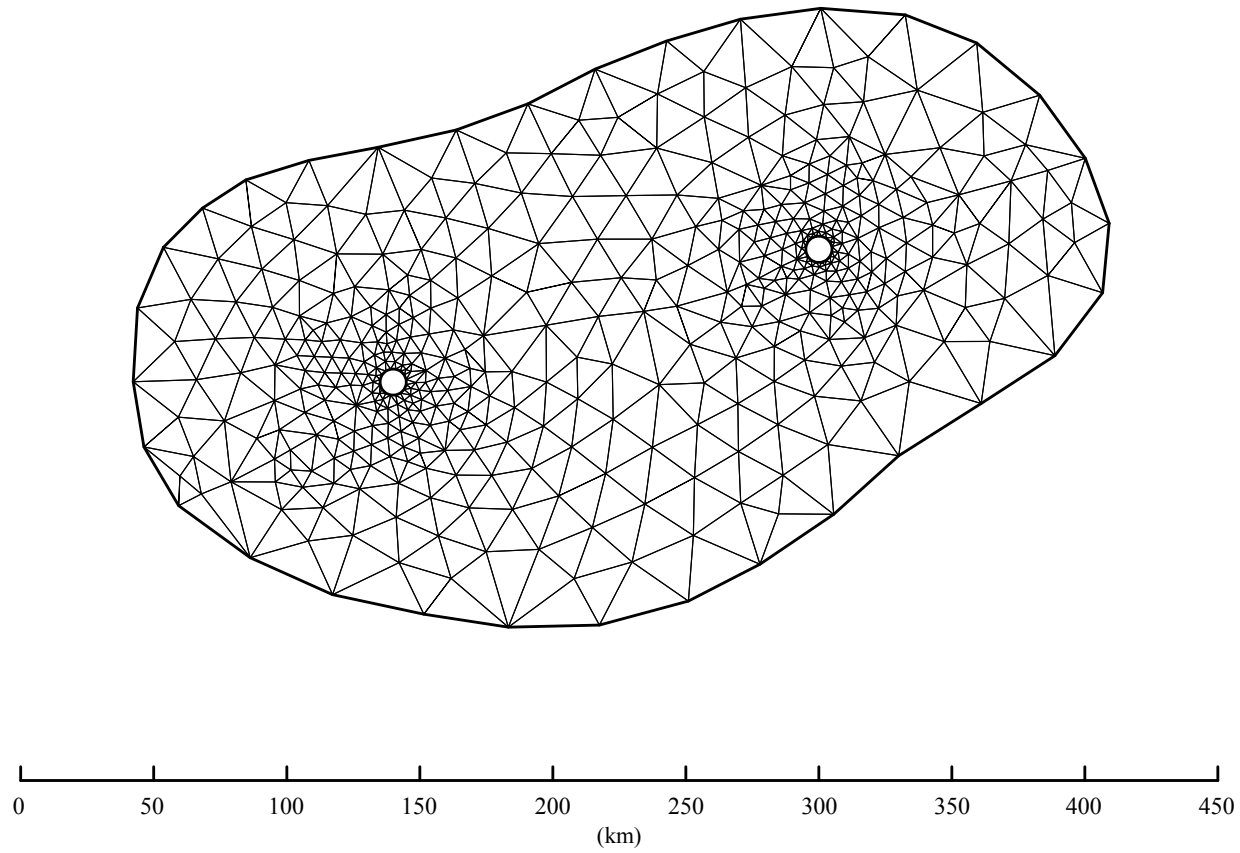


Figure 3. Finite element discretization of the study region.

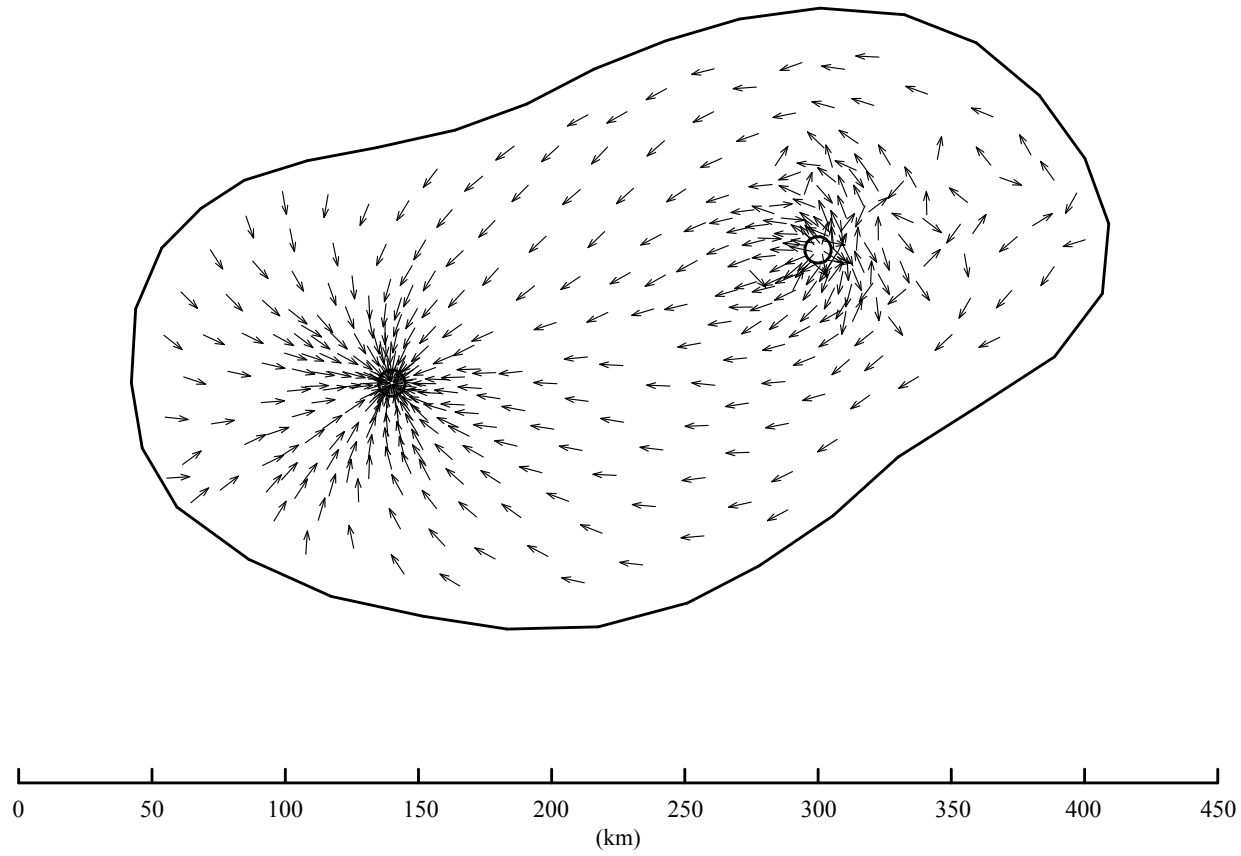


Figure 4. Flow pattern of class 1 users traveling to city 1 in the no-toll scenario.

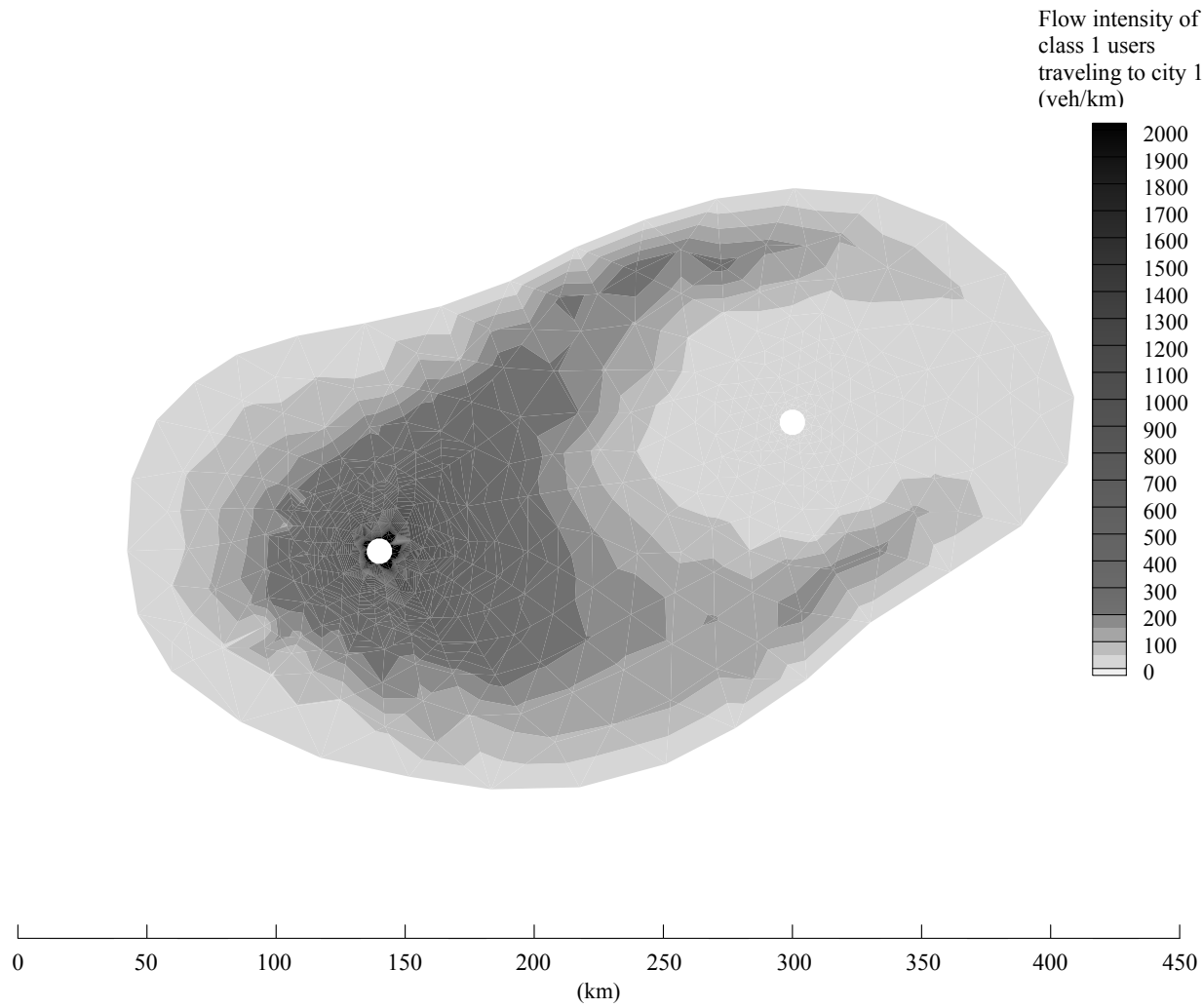


Figure 5. Flow intensity of class 1 users traveling to city 1 in the no-toll scenario.

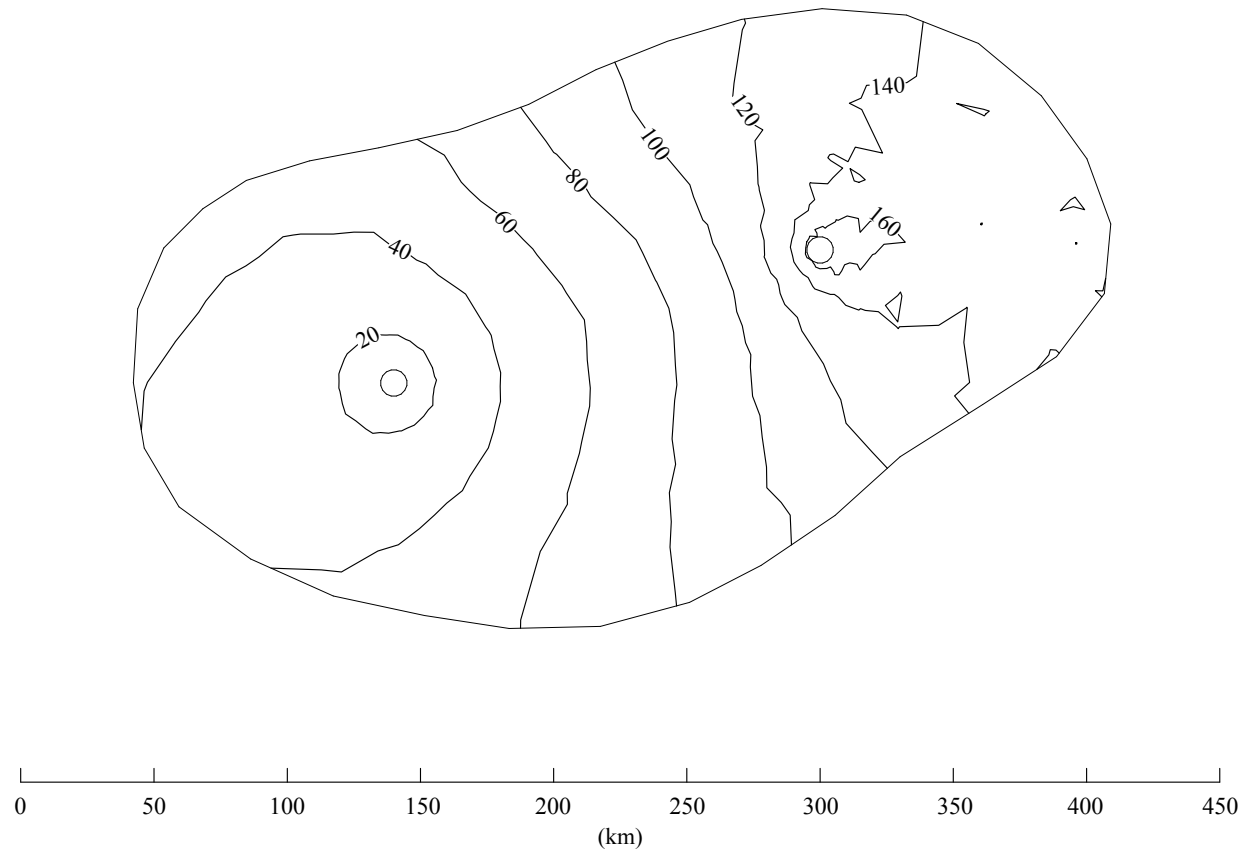


Figure 6. Total travel cost for class 1 users traveling to city 1 in the no-toll scenario (CNY).

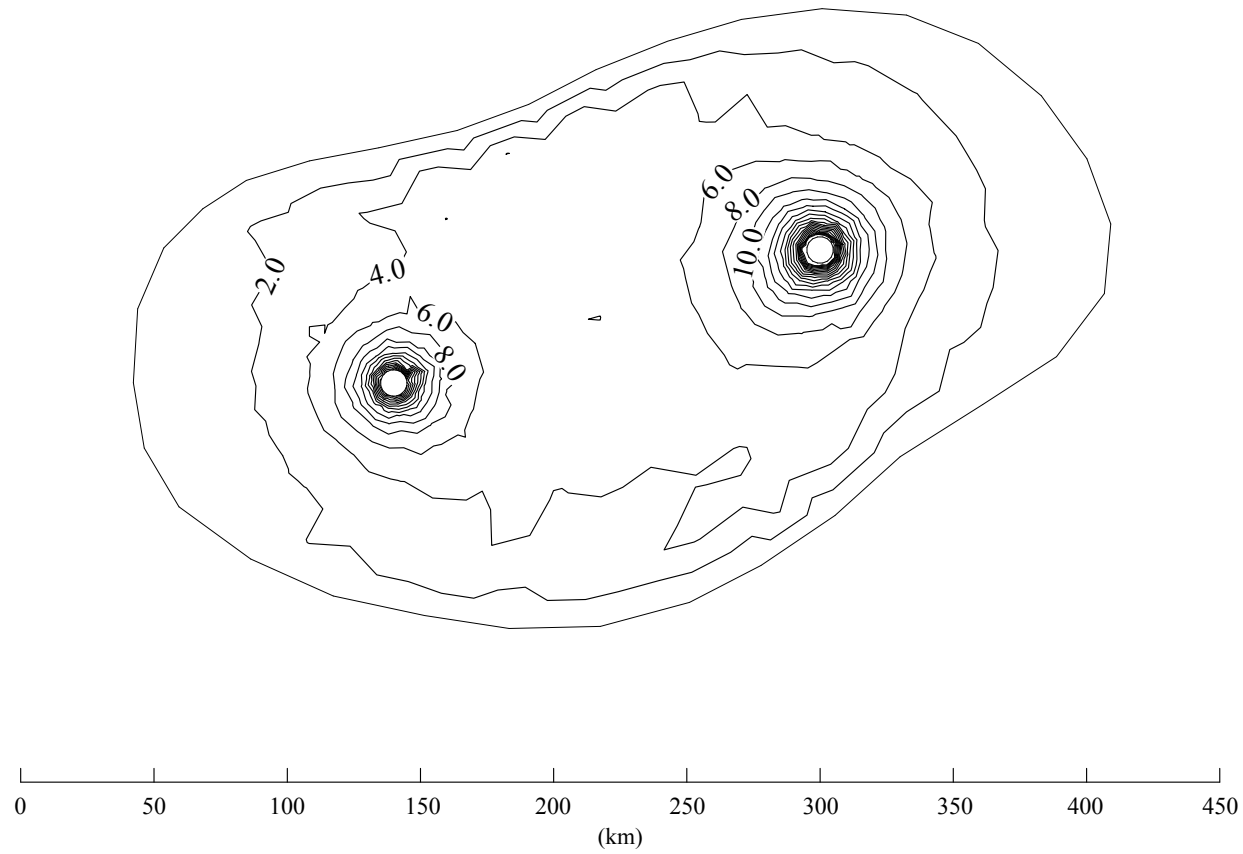


Figure 7. Contour plot of the distance-based iso-toll rate for class 1 users in the first-best scenario (CNY/km).

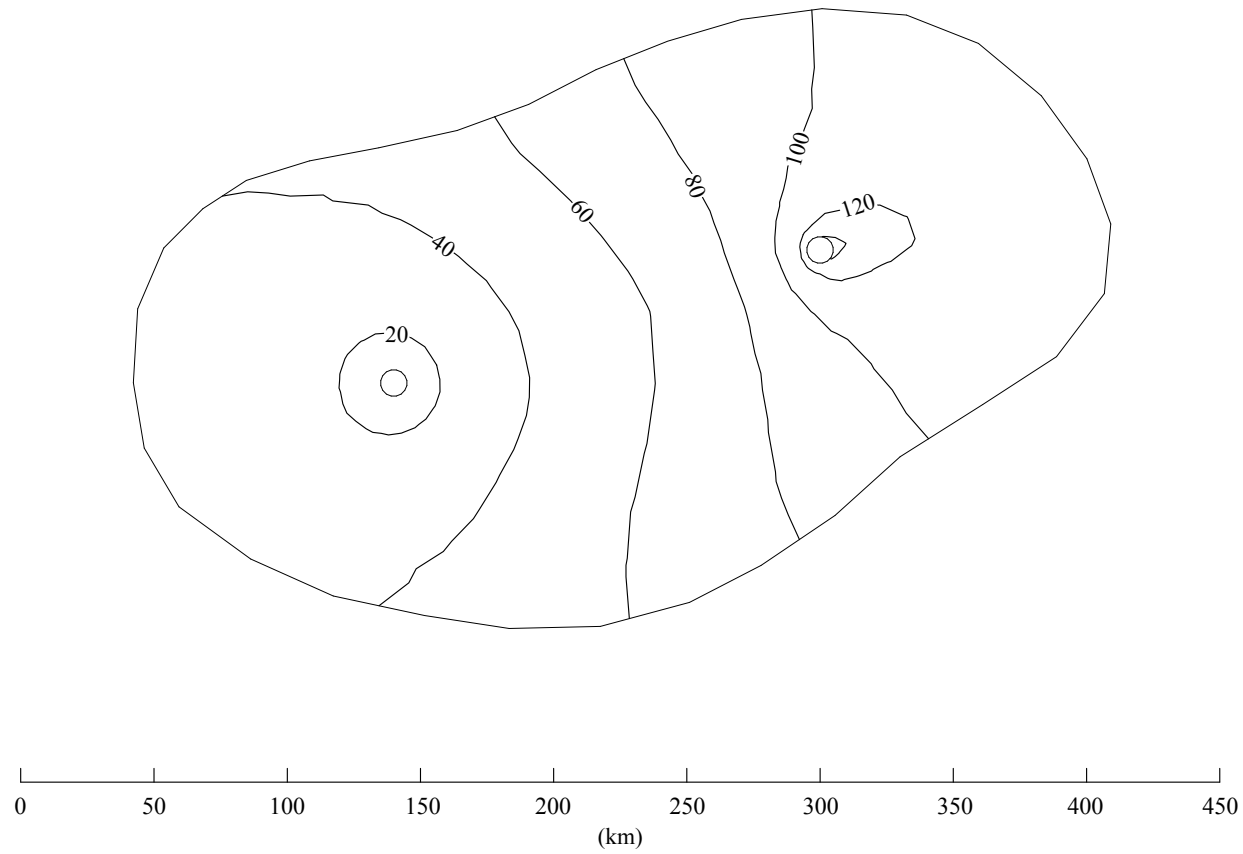


Figure 8. Contour plot of the total toll paid by class 1 users traveling to city 1 in the first-best scenario (CNY).

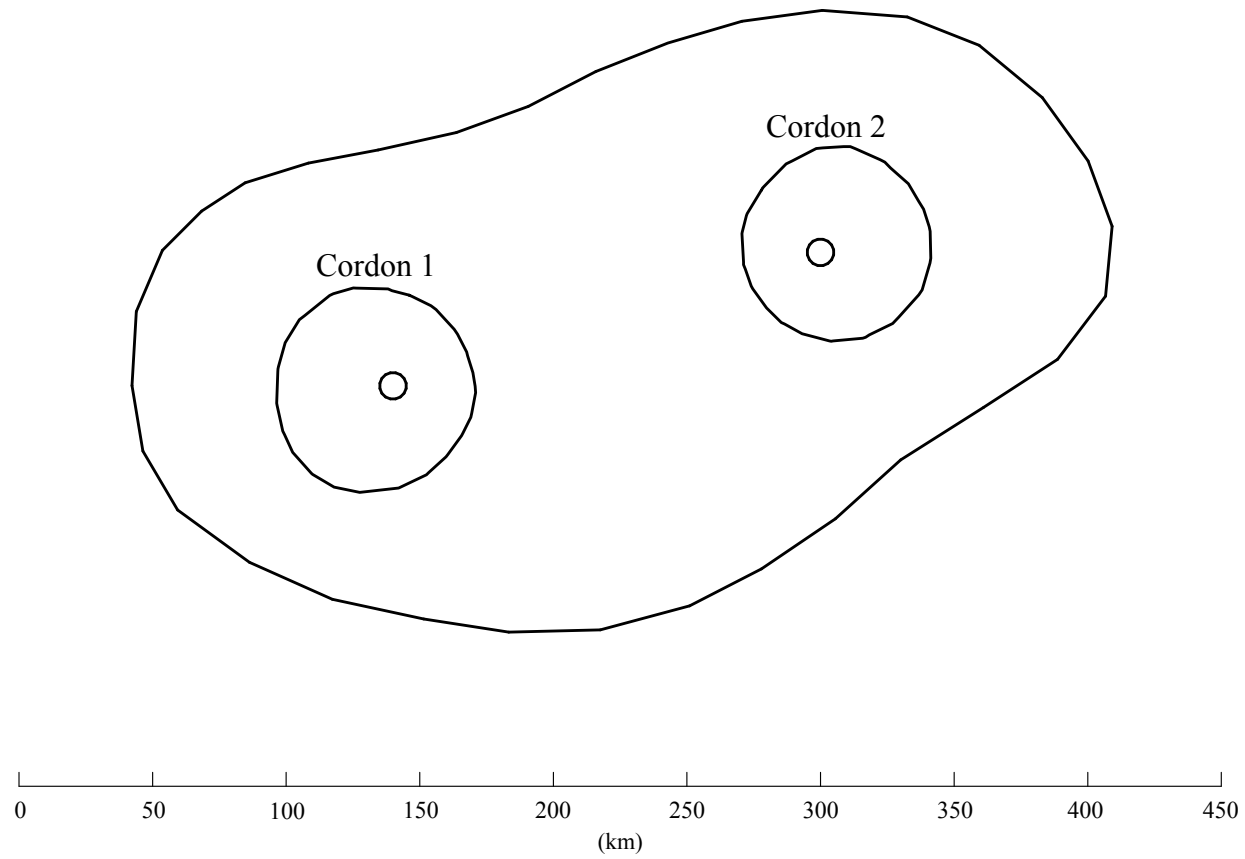


Figure 9. The two charging cordons.

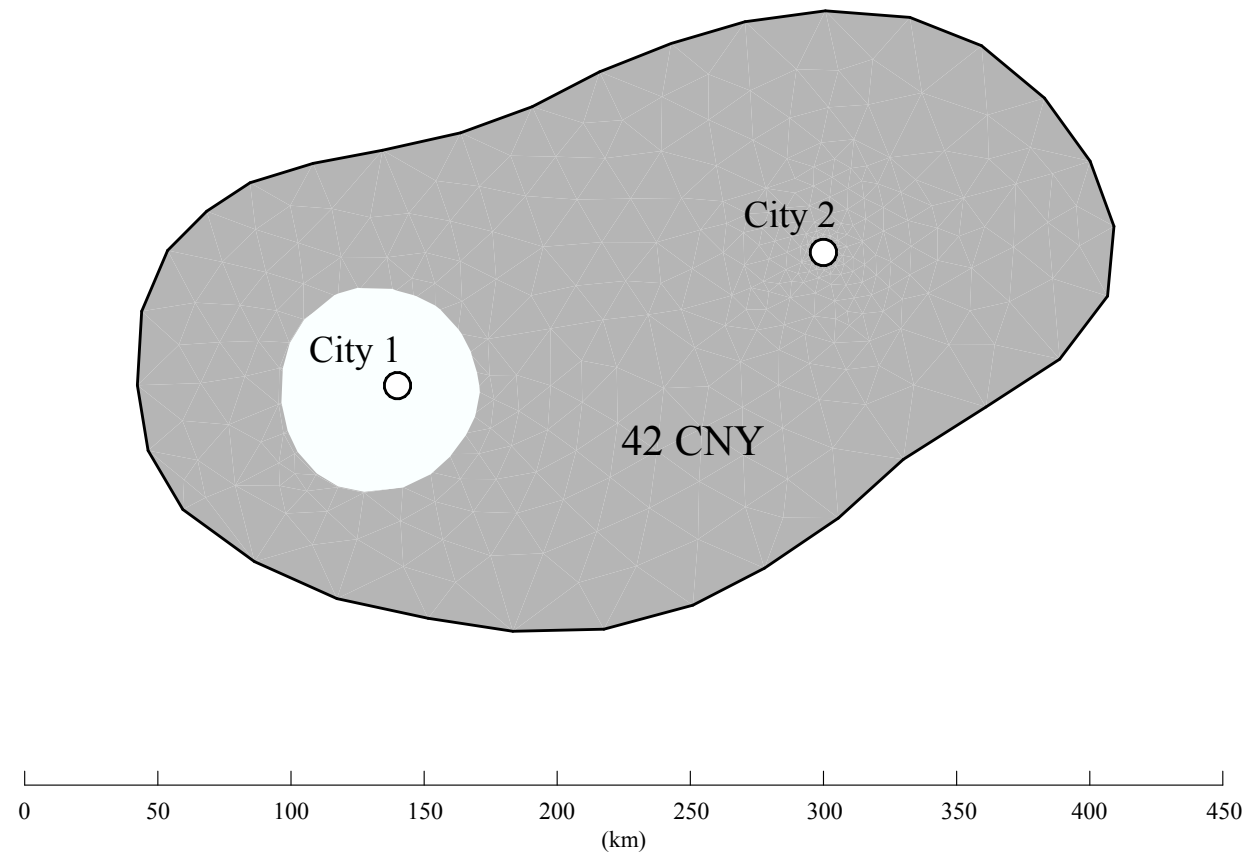


Figure 10. Total toll paid by class 1 users traveling to city 1 in the second-best charging scenario.

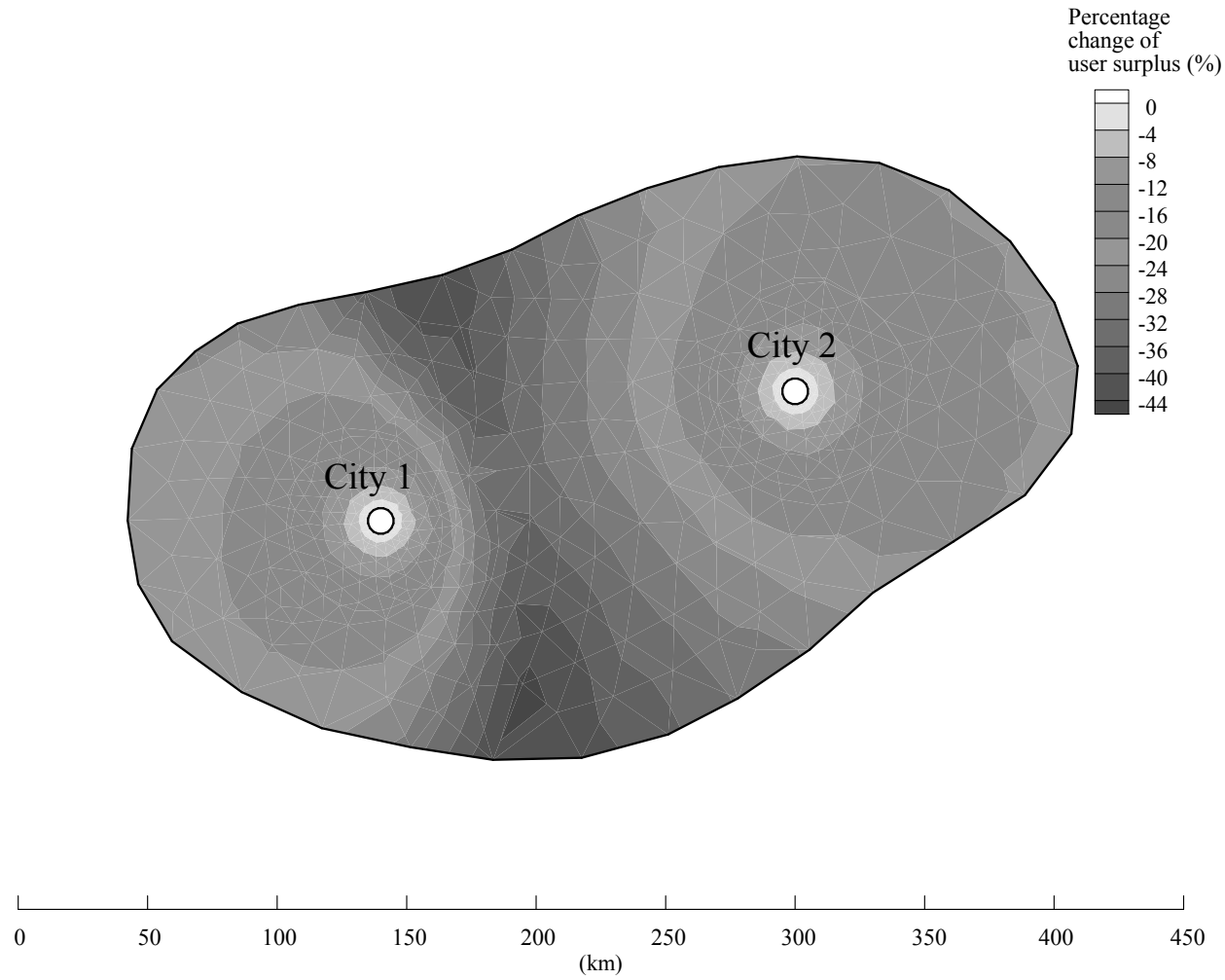


Figure 11 Percentage change of user surplus for class 1 users as first-best congestion-pricing scheme is implemented

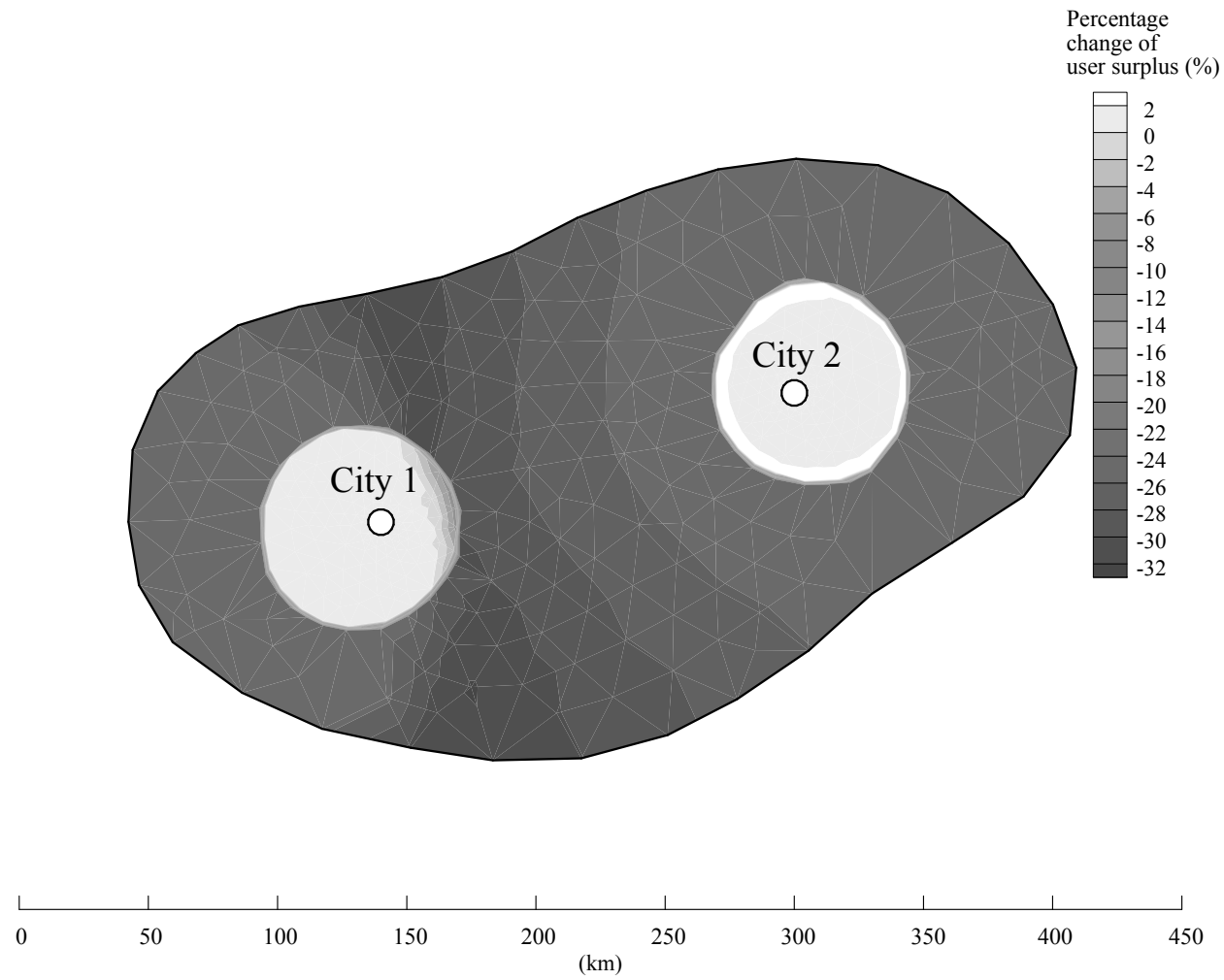


Figure 12 Percentage change of user surplus for class 1 users as second-best congestion-pricing scheme is implemented