

The Case against Road Congestion Pricing: Full Internalization of Congestion Externalities Is Not Enough

by

CRISTIAN ANGELO GUEVARA

Universidad de los Andes, Chile
Facultad de Ingeniería y Ciencias Aplicadas
San Carlos de Apoquindo, 2200
Las Condes, Santiago Chile
Tel: 56-2-412-9621
Fax: 56-2-412-9642
aguevara@uandes.cl

ABSTRACT

Road congestion pricing is sustained on the Pigouvian principle that making agents pay for their external costs will always lead the system to its optimum. This public policy enjoys an undisputed appeal in academia as the ultimate tool to alleviate congestion. The purpose of this paper is to question this consolidated belief. First, we make a critical review of the main arguments that have been raised against road congestion pricing in previous literature. Then, we use an example, in which drivers differ in their values of time, to show that first-best Pigouvian tolls (ones that fully eliminate the divergence between private and social cost) can prevent latent opportunities, and may result in a social loss when compared to selfish equilibrium. This result can be seen as an extension of Coase's critique to Pigouvian theory. Afterwards, we argue that, even if information on latent opportunities were available, the sophistication of the modeling tools required to solve such systems would prevent any practical determination of optimal tolls. Finally, we show that, even if such modeling tools would be available, the optimization problem to be solved is extremely challenging, and their results can be misleading. Extending the previous example, we show that the problem is not convex, that local optima are first-best Pigouvian equilibriums, but some may be inferior to the selfish equilibrium. The paper finishes with a call for a broader view in the design of policies to alleviate congestion.

Key Words: Road Congestion Pricing, Coase, Pigou, Heterogeneous Values of Time

1 Introduction

The fundamental problem of economics is the allocation of scarce resources. Pigou (1920) stated that this allocation is inefficient when there is a divergence between the cost perceived by each agent (private cost) and the cost caused by each agent to the whole society (social cost). To correct this divergence, Pigou proposed the application of a tax equivalent to difference between private and social cost. Vickerey (1955, 1963), Walters (1961) and others extended Pigouvian theory to the field of transportation. The idea in this particular field is to make drivers account for their marginal external congestion costs by imposing an equivalent toll. This toll is termed a Pigouvian toll.

Road congestion pricing enjoys an undisputed appeal in academia as the ultimate tool to alleviate congestion. However, this public policy has been scarcely applied in practice. Only Singapore, London and Stockholm use it nowadays. Explanations for this discrepancy between what seems theoretically suitable and what is applied in practice usually range from lack of political leadership (King et al., 2007) to limitations in explaining the true benefits of this policy to the public (Ison, 2005). Consequently, many efforts have been devoted to devise artifacts to enhance road congestion pricing popularity (see, e.g Jones, 1998; Verhoef and Small, 2004; Daganzo and Garcia, 1998). However, despite academic enthusiasm toward road congestion pricing, it can be shown that the skepticism of drivers and authorities toward road congestion pricing has very logical roots and, in many cases, is aligned with the interest of the whole society. In Section 2 we make a critical review the main arguments that have been raised against road congestion pricing in recent literature.

Then, in Section 3, we go a step beyond on the critique of road congestion pricing. We show that the basic principle upon which this public policy is based is not always valid. It is not always true that making drivers pay for their external costs will necessarily result in an increase of social welfare. We use an example, in which drivers differ in their values of time, to show that first-best Pigouvian tolls (ones that fully eliminate the divergence between private and social cost) can prevent latent opportunities, resulting then in a social loss when compared to selfish equilibrium. This result can be seen of an extension, to the urban transportation field, of Coase's (1960) critique to Pigouvian theory.

Then, in Section 4, we argue that a full account of latent opportunities in an urban context would require knowing not only the congestion externality (which is relatively easy to monitor in the streets), but also the opportunity cost of each individual. This would imply being able to model, for example, the potential set of activity participation and scheduling choice-sets that are specific to each person. These types of modeling tools are still under development and remain full of challenges. Before such modeling capabilities are available, the case against road congestion pricing as the ultimate tool to alleviate congestion will be a strong one. Afterwards we show that, even if suitable microscopic integrated models of the urban system are available, the optimization problem to be solved would be extremely challenging and their results may be misleading. We first show that the problem is generally not convex, troubling substantially its resolution. Moreover, we show that local optima are first-best Pigouvian equilibriums (in the sense that there is no divergence

between private and social cost) but may be inferior to selfish equilibrium. This may be misleading for the planner who, facing the fact that there is no divergence between private and social costs may have a false sense of assurance (sustained in the widely accepted Pigouvian tradition) that the current tolling scheme is socially superior.

In Section 5 we analyze the impact of this research in the design of policies to alleviate congestion and, finally, in Section 6 we summarize our findings, make a call for a broader view in the design of policies to alleviate congestion, and suggest future lines of research in this area.

2 The Case against Road Congestion Pricing in Recent Literature

Within the classical Pigouvian approach to understand road congestion pricing, there are some strong arguments that explain why users and authorities are usually not very enthusiastic with the idea of road congestion pricing. This section reviews the main arguments that have been raised against this policy tool in the recent literature.

Consider first the drivers' point of view. It can be shown that when the value of time is homogenous all users are worse off after the application of road congestion pricing (see, e.g. Hau, 2005). Only when the tolls collected are recycled, some users may become better off. This fact is sometimes neglected in the description of road congestion pricing, and have even led to some confusion, such as to affirm that the tolls collected could be burnt and the benefits from road congestion pricing would not be lost (see, e.g. Rouwendal and Verhoef, 2006, pp. 107).

But even after revenue recycling, road congestion pricing might be unpopular. Which group of drivers gets potentially benefited from the application of road congestion pricing will depend on the specific revenue recycling scheme considered in each case. De Borger and Proost (2010) showed that the users that may be objectively or subjectively benefited from road congestion pricing are usually a minority.

Furthermore, road congestion pricing might be regressive. When the value of time is heterogeneous, drivers with a high valuation may result directly benefited from road pricing before revenue recycling, but in that case drivers with low valuation are generally even more harmed (Nie and Liu, 2010). Since high valuations of travel time are usually associated to higher incomes, it can be said that road congestion pricing is regressive.

Given the facts that most drivers get negatively affected by congestion pricing and, that this policy may be regressive, driver's rejection of this public policy seems to have very logical roots. This explains why citizens from Hong Kong, Edinburgh, Manchester, Birmingham, and New York, rejected road congestion pricing projects when correctly consulted about it.

Authorities on the other hand also have some strong arguments to be skeptic about road congestion pricing. To begin with, if the majority of the users often have a negative assessment of road congestion pricing, most authorities might not be willing risk their

political capital in a policy that would prevent them from being reelected. But even if the authority cares more about economic efficiency than votes, their enthusiasm toward road congestion pricing is questionable for several reasons.

First, a natural question to ask is how large are the potential benefits that can be obtained from road congestion pricing. This problem is known as the “price of anarchy”. It has been shown that the benefits from road congestion pricing are generally bounded and in some cases could be very limited (Roughgarden, 2005). Therefore, there are many other policies for alleviating congestion that may outperform road congestion pricing. For example, Basso et al. (2011) showed that dedicated bus lanes may not only be more efficient than road congestion pricing, but also its impact in consumer surplus may make this alternative policy easier to implement, from a political point of view. Given the negative perception of road congestion pricing among users, and its potentially small benefits compared to more popular policies, moving away from road congestion pricing seems a very logical choice for an authority on where to invest an always scarce political capital. Even a successful implementation of road congestion pricing could end up being, from authority’s point of view, a pyrrhic victory.

Second, toll collecting costs may be large and should be discounted from any benefit gathered from road congestion pricing. This fact is sometimes forgotten in the description of this policy in transportation literature. Interestingly, according to Prud’homme and Bocarejo (2005), when toll collecting costs are discounted from the potential benefits attained by London’s road congestion pricing project, the net result becomes a social loss.

Third, proactive authorities may try to revert the unpopularity of road congestion pricing by implementing creative revenue recycling schemes such as those explored by Goodwin (1989) or by Small (1992). However, the authority should be careful on this action. It has been shown that inappropriate revenue recycling schemes, designed solely to increase public acceptability of road congestion pricing, may induce welfare losses elsewhere that may be larger than any gain attained from road congestion pricing (Lindsley and Verhoef, 2001). Furthermore, Parry and Bento (2005) showed that, when the labor market is distorted by an income tax (which is the case in most societies), returning collected tolls by a lump sum may end up causing a social detriment derived from a reduction in productivity.

Fourth, the implementation of ideal or first-best congestion pricing is almost impossible in practice, since it should involve tolling each and every link of a complicated network. Therefore, the problem needs to be addressed using a second-best approach, like cordon pricing. The problem is that these simplifications may cause a significant reduction of the potential welfare gains obtained from road congestion pricing (see, eg. Zhang and Yang, 2004; Santos, 2004, Verhoef, 2005), and may easily result in a net social loss.

These are some of the principal issues that have been raised against road congestion pricing within the Pigouvian framework in recent literature. This shows that skepticism of users and authorities toward this policy, not only has very logical roots, but is likely to be aligned with social interest in several plausible scenarios. In what follows, we go a step beyond in making the case against road congestion pricing.

3 Full Internalization of Congestion Externalities May Cause a Social Loss

In this section we use an example to show that first-best Pigouvian tolls, ones that fully correct for any divergence between private and social congestion costs, may cause a social loss. We show that first-best Pigouvian tolls may preclude latent opportunities that, under a selfish equilibrium, would lead the system to a socially superior equilibrium.

Consider the problem described in Figure 1, where rational drivers are assigned to a static, non-atomic and time-independent network, with inelastic demand. All drivers travel alone in the same type of car from origin (O) to destination (D). Two links connect O to D. Link 1 has a fixed travel time of 105[m] and link 2 has the variable travel time function shown in Figure 1, where time (t) is in minutes [m] and the flow (q) is in drivers per hour [d/h]. There are two types of drivers: Commuters (C) and Leisure (L) travelers. There is a flow of $q_C=1$ [d/h] C drivers and of $q_L=3$ [d/h] L drivers. C drivers have a high valuation of travel time $VT_C=4$ [\$/m] and L drivers have a low valuation $VT_L=1$ [\$/m].

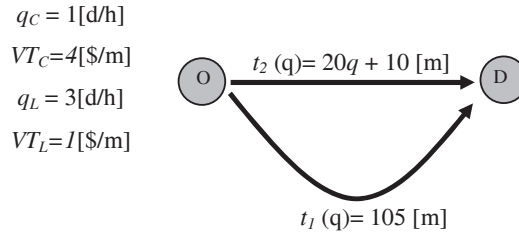


Figure 1: Network Setting Original Case

The determination of user's equilibrium (Sheffi, 1985) resulting from the selfish or spontaneous behavior of agents in the problem described in Figure 1, can be achieved using a dominant strategy approach. Note that if the entire 4 [d/h] use link 2, travel time on that link would be 90 [m]. Since 90 [m] is below the fixed travel time of link 1, no agent would be willing to move from link 2 to link 1, what means that this flow assignment constitutes an equilibrium. We will term this the User's Equilibrium 1 (UE1), which results are summarized in Table 1.

Table 1: Users' Equilibrium 1 (UE1)

Link	q_C [d/h]	q_L [d/h]	Time [m]
1	0	0	105
2	1	3	90

Consider now the problem of achieving efficiency in the problem portrayed in Figure 1. To correctly answer this question we first need to define a measure of welfare to compare this UE1 with other potential scenarios. To simplify the setting, we will assume that the differences in value of time across user classes are only due to differences in the marginal utility of time. In other words, we will assume that the marginal utility of income is

equal for C and L drivers. This assumption is somehow reasonable to sustain when classes are differentiated by trip purposes, as in the example described in Figure 1. Under this assumption, maximizing social welfare in the problem described in Figure 1 would be equivalent to minimize the following social cost function:

$$SC = \sum_{l \in \text{Links}} (VT_C q_{C_l} + VT_L q_{L_l}) t_l.$$

Other assumptions on the nature of the heterogeneity of travel time would lead to more complicated social cost functions (see e.g the discussion in Galvez and Jara-Diaz, 1998) that may also trigger additional difficulties. For example, if the differences in value of time are only due to differences in income (the marginal utility of time is the same across classes), the social cost function to be minimized would correspond to the total time consumed. However, in that case Pigouvian tolls cannot be anonymous, making road congestion pricing completely impractical (Yang and Hang, 2008). Given that the issues that we want to illustrate can occur with any definition of the social cost function used, we will stick on this paper to the assumption that heterogeneity of travel time is only due to differences in the marginal utility of time.

Given this definition of the social cost function, the total social cost for User's Equilibrium 1 (SC_{UE1}) described in Table 1, can be calculated as shown in Eq. (1).

$$SC_{UE1} = \left(\underbrace{0}_{q_C} * \underbrace{4}_{VT_C} + \underbrace{0}_{q_L} * \underbrace{1}_{VT_L} \right) \underbrace{105}_{t_1} + \left(\underbrace{1}_{q_C} * \underbrace{4}_{VT_C} + \underbrace{3}_{q_L} * \underbrace{1}_{VT_L} \right) \underbrace{90}_{t_2} = 630[\$/h] \quad (1)$$

Consider now the problem of determining the social optimum for the problem described in Figure 1. Yang and Huang (2004) show that first order conditions of this problem imply that the optimal network equilibrium is attained when, for each link, each user of each class faces a toll that is equivalent to their external congestion cost. This toll is equal to the product of travel time externality and the average value of time of the users traversing that link. For the example deployed in Figure 1, the optimal toll for link i will therefore correspond to the following expression, evaluated at the optimal flows:

$$Toll_i = \underbrace{q_i \frac{\partial t_i(q_i)}{\partial q_i}}_{\text{Time Externality}} \underbrace{\frac{(VT_C q_{C_i} + VT_L q_{L_i})}{q_i}}_{\text{Average Value of Time}}. \quad (2)$$

Yang and Huang's (2004) result implies that, despite the heterogeneity in users' preferences, the Pigouvian toll is anonymous and determined solely based on the externality caused. This is fully consistent with the Pigouvian tradition, and with what is widely accepted by academia in the problem of road congestion pricing with heterogeneous value of time (see, e.g. Small and Verhoef, 2007; Dial, 1999a,b; Florian, 1998; Yang et al., 2002; and Yang and Zhang, 2002).

Yang and Huang (2004 pp. 7) go a step beyond and affirm that the problem of determining optimal flows with heterogeneous values of time is convex. This would imply that optimal tolls may be determined using any algorithm for multi-class network equilibrium, provided link marginal cost functions are used. For the problem described in Figure 1, this would imply adding Eq. (2) to the cost function of each link, considering the flows q as variables.

Although Yang and Huang's (2004) approach to determine the social optimum does work for the particular problem described in Figure 1, their statement on the convexity of multiple-class assignment with variable tolls is incorrect for a more general case. Interestingly, convexity is a critical assumption in the derivation of the theory of externalities (see, e.g. Mas-Collel, 1995 Ch. 11). The failure of the convexity assumption when the value of time is heterogeneous is in the root of the issues we want to highlight about road congestion pricing in this paper.

The question about the convexity of multiple-class assignment with variable tolls was originally raised by Dial (1999). Milchtaich (2006) showed that this problem is convex, in terms of link's flows, when link's costs differ by class but in a constant that is independent of the flows. As it can be seen in Eq. (2), this assumption is broken in Yang and Huang's (2004) model. We provide in Section 4 a counter-example to show this fact and also that some local optima, although being Pigouvian equilibriums, might be inferior to selfish equilibrium.

Solving a non-convex nonlinear problem is generally very complex but the simple example described in Figure 1 can be solved by inspection. Note that any L driver moved from link 2 to link 1 would imply a reduction of 20 [m] in the travel time experienced by C drivers, a reduction that would worth $20[\text{m}] \cdot 4[\$/\text{m}] = \80 . But also, each L driver moved from link 2 to link 1 will experience an increase in travel time ranging from 15 [m] (when all L drivers use link 2), to 65 [m] (when the last L driver moved link 2). Since $VT_L = 1$ [\$/m], the net effect of moving L drivers from link 2 to link 1 will always be positive. In turn, any move of C drivers from link 2 to link 1 would cause a reduction in social welfare. Therefore, the minimum total social cost will be attained by moving all drivers with a low valuation to link 1, the link with the largest value of time. Note that this result is equivalent to those obtained for the problem of price discrimination in a monopoly (Varian, 1985).

The minimum total social cost flow assignment for the problem described in Figure 1 should satisfy the optimality conditions derived by Yang and Huang's (2004), and should be therefore sustainable with the tolls shown in Eq. (2), evaluated at those optimal flows. Given that travel time in link 1 is fixed, Eq. (2) implies that the toll in that link should be zero. In turn, the toll for link 2 will correspond to $20[\text{m}] \cdot 1[\text{d/h}] \cdot 4[\$/\text{m}] = \$80$. It can be immediately noted that this toll imposes equilibrium for the problem described in Figure 1 because, under it, no driver would be willing to change route. We will call this equilibrium the Pigouvian Equilibrium 1 (PE1), which is summarized in Table 2.

Table 2: Pigouvian Equilibrium 1 (PE1)

Link	q_C [d/h]	q_L [d/h]	Time [m]	Toll [\$]
1	0	3	105	0
2	1	0	30	80

The total social cost in this case SC_{PE1} , can be calculated as shown in Eq. (3)

$$SC_{PE1} = \left(\underbrace{0 * \frac{4}{VT_C} + 3 * \frac{1}{VT_L}}_{t_1} \right) 105 + \left(\underbrace{1 * \frac{4}{VT_C} + 0 * \frac{1}{VT_L}}_{t_2} \right) * 30 = 435[\$/h]. \quad (3)$$

Comparing UE1 (Eq. 1) with PE1 (Eq. 3) it can be noted that there is a substantial reduction of about 30% in social cost. In UE1 there was a divergence between private and social cost, there was an external cost that was not internalized by the agents. In turn, in PE1, this divergence between social and private costs was corrected by means of a toll and, fully consistent with the Pigouvian theory, the new equilibrium resulted in a reduced social cost, in an increase of efficiency.

Although PE1 equilibrium produced an increase in social welfare compared to UE1, there are some arguments against it within the Pigouvian theory that worth highlighting. First, note that although the price of anarchy is arguably large, if toll collecting costs are above 195[\$/h] the benefits from road congestion pricing would vanish. Also, any other policy able to produce a reduction larger than 195[\$/h] (minus respective implementation costs) would outperform road congestion pricing. Note also that the consumer surplus of the majority of the drivers is also reduced when comparing UE1 with PE1. 1[d/h] reduced their travel time from 90[m] to 30[m], but 3[d/h] increased it from 90[m] to 105[m]. Therefore, if a poll on approval of road congestion pricing is implemented, the policy would suffer a clear loss.

However, even if collecting costs are zero and the authority cares more on economic efficiency than votes, an important argument can be raised against the convenience of implementing the tolls associated with PE1. To explain this critique, consider that the problem described in Figure 1 is not the true problem, that there is a latent opportunity for C drivers that is not accounted on it. This latent opportunity can correspond, for example, to the ability of re-scheduling the activity pattern. Consider for example that C drivers are able to negotiate working from home or arriving later at work to avoid the large travel times experienced. This re-scheduling opportunity is said to be latent in the sense of being apparent to C drivers only when facing the large travel times that result from UE1 for a while. Only under those inconvenient conditions C drivers might be impelled to make an arrangement with their employers that, otherwise, would be unknown to all agents, particularly to the researcher who builds the model to give advice to the authority.

This latent opportunity is not free. For the problem described in Figure 1, consider that the cost of this latent opportunity can be measured in an equivalent welfare variation (for C drivers) that worth \$220 or, equivalently, 55[m]. For example, C drivers may be able to negotiate to work at an un-congested time of the day when there are no L drivers and,

therefore, travel time experienced by them would only be 30[m]. However, given their loss in productivity, C drivers may also have to accept a reduction in salary of \$100 if they decide to travel at the un-congested time of day. Given that $VT_C=4$, the combination of 30[m] and \$100 of salary reduction, is equivalent to a toll of \$220 or a travel time of 55 [m] for C drivers.

The problem considering this latent opportunity can be represented by the network described Figure 2. In this case, link 3, with a travel time of 55, represents the equivalent welfare variation associated with the latent opportunity. Also, the topology considered allows precluding L drivers from accessing the latent opportunity link, by making C drivers to travel from O' to D and L drivers from O to D .

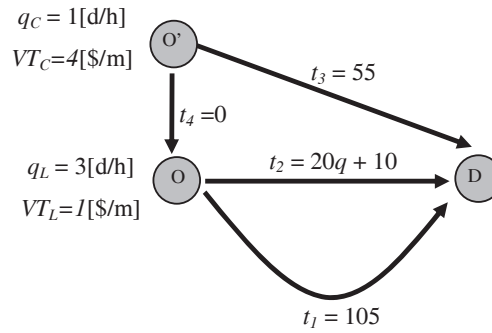


Figure 2: Network Setting with Latent Opportunity

Users' Equilibrium for the problem described in Figure 2 (UE2) can be found using a dominant strategy analysis. If all L and C drivers use link 2, travel time on that link would be 90 [m], which is still below the 105 [m] of link 1. This means that all L drivers will use link 2 in equilibrium. Then, if all L drivers use link 2, the travel time on that link would be 70 [m], which is above the travel time of link 3. This means that C drivers will necessarily use link 3. Consequently, in UE2, 3 L [d/h] use link 2, spending 70 [m]; and 1 C [d/h] use link 3 spending 55 [m] on it. This is summarized in Table 3.

Table 3: Users' Equilibrium 2 (UE2)

Link	q_C [d/h]	q_L [d/h]	Time [m]	Toll [\$]
1	0	0	105	0
2	0	3	70	0
3	1	0	55	0

The total social cost with User's Equilibrium 2 (SC_{UE2}) is calculated then as the sum and product of travel times and respective values of time, as shown in Eq. (4).

$$SC_{UE2} = \left(\underset{q_C}{0} * \underset{VT_C}{4} + \underset{q_L}{0} * \underset{VT_L}{1} \right) \underset{t_1}{105} + \left(\underset{q_C}{0} * \underset{VT_C}{4} + \underset{q_L}{3} * \underset{VT_L}{1} \right) \underset{t_2}{70} + \left(\underset{q_C}{3} * \underset{VT_C}{4} + \underset{q_L}{0} * \underset{VT_L}{1} \right) \underset{t_3}{55} = 430[\$/h] \quad (4)$$

Note that, $SC_{UE2} < SC_{PE1}$. This means that the selfish equilibrium represented by UE2, where private and social costs differ, is indeed superior to the equilibrium PE1 where the divergence between private and social costs is zero. Interestingly, note also that under PE1, the latent opportunity is no longer attractive for C since the generalized travel time for link 2 is $50 = 30[m] + 80[\$/4[\$/m]]$, which is below the 55 [m] of link 3. This means that the application of tolls that fully correct for the divergence between private and social cost precluded the use of the latent opportunity by C drivers, leading the system to a social loss, when compared with the selfish equilibrium.

To illustrate the implications of this result, consider a different setting for the same problem. Consider that C drivers are originally the only ones using the system. Under this condition all C drivers would be assigned to link 2 spending a total of 30 [m] traveling. Then, some leisure project (e.g. a cinema) is about to be opened at D , attracting the flow L drivers. C drivers are worried because the leisure project will result in a substantial increase of their travel time from 30 [m] to 90 [m].

A diligent authority (possibly after some riots from of C drivers) is worried about the loss in efficiency that would result from the external cost produced by the L drivers onto the C drivers. Seeking the best economical advice decides to apply road congestion pricing. The best modelers available built a model network for the problem at hand. However, since C drivers currently spend only 30 [m], any activity re-scheduling opportunity would be latent. C drivers are honestly unaware of latent opportunities because only facing delays as large as 90 [m] might be impelled to negotiate alternatives that take 55 [m]. Consequently, the modeler would only be able to develop a model like that on Figure 1, what would finally result in advising tolls that would drive the system to PE1 when the new leisure project begins to function.

On the other hand, a negligent authority, one that simply ignores or maybe does not have the political courage to implement road congestion pricing, leave C drivers by themselves. Then, when the leisure project is implemented, C drivers experience the large externality described in UE1 for a while. Facing the fact that the authority will not intervene making all agents perceive the full cost of their actions, C drivers are eventually able to re-schedule their activity pattern by negotiating an alternative work plan. Later, the system evolves toward UE2. Although C drivers are indeed worse in UE2 compared with PE1, the fact is that, for this example, it would be socially preferable to have a negligent authority.

This result may seem paradoxical from the point of view of Pigouvian tradition. In UE2 there is a divergence between private and social cost, there are some external cost that are not internalized by the agents. In turn, in PE1, this divergence between social and private costs is fully corrected by means of a toll but, rather surprisingly, UE2 is socially superior to PE1. Notably, PE1 would preclude the manifestation of the latent opportunity, since it would result completely unattractive to C drivers. This counter-example can be seen as an extension of Coase (1960, pp. 32) critique to Pigouvian theory. Coase's (1960) shows that, when intervening in a system, we have to consider that any change that may seem to be an improvement in some area, may lead to a worsening in others and that the net effect may be negative. The divergences between private and social costs show that it might be possible

to improve the system, but the correction of this difference with a tax is not necessarily the best solution to the problem. The final call advocated by Coase is that, instead of concentrating in the divergences between social and private cost, we should pay attention to the total effect of the interventions on a particular system.

These results obviously depend on the particular setting used. For many other settings, the application of Pigouvian tolls will lead to equilibriums that imply an increase in welfare. For example, if value of time of C drivers is 2 [\$/m] instead of 4 [\$/m], the counterexample will not occur. This counterexample will also not occur if the travel time on link 1 is 55 [m] or if it is 1000 [m]. The problem is that, in some cases, it might be more costly or cheaper for the agents that cause relatively larger externality (the L drivers), to pay for it, than for those who bear relatively more externality (the C drivers), to deal with it. This shows that, in the design of tolls to alleviate congestion it is necessary to look not only at the external congestion costs, but also at the opportunity costs of each agent in the system.

In this section we showed that when the true model is Figure 2, but the researcher used Figure 1 to represent the system, a full internalization of congestion externalities may result in a social detriment when compared with the selfish equilibrium. In the next section we analyze the huge challenges involved in modeling the availability of latent opportunities in a practical model to determine optimal tolls in a real transportation network. We also show that, even if such modeling capabilities are available, the optimization problem to be solved is not only difficult but may lead to misleading results. We show that some local optima, despite being Pigouvian equilibriums where all agents perceive the full social costs of their actions, may be socially inferior when compared to the selfish equilibrium.

4 Global Optimum Tolls Considering Latent Opportunities

Consider now the problem of determining globally optimum tolls in the problem deployed in Figure 2, where the latent opportunities are included in the model. It should be noted first that, in the most general case, to address such a problem in a real city we would need to know and be able to forecast, for example, the activity participation and scheduling choices of each and every individual in a city. This includes modeling, for example, the social networks and character of each and every agent in the city, to be able to depict their true ability to re-schedule activities. Although the theoretical framework to acquire such modeling capabilities has been studied for some years (see. e.g Bowman, 1998), practical applications of this framework are still limited.

Moreover, models that try to integrate activity participation and scheduling, with land use and mode choice modeling, require even more sophisticated modeling skills. This type of models are known as microscopic integrated urban models, from which UrbanSim (Waddell et al., 2008), ILUTE (Miller et al., 2004), ILLUMAS (Strauch et al., 2005) and iTEAM (Almeida et al., 2009), are the main examples. However, this type of decision support tools are still under development and are full of practical and theoretical challenges (Guevara, 2010). Before such modeling capabilities are available, we would only be able to represent the reality by a model like Figure 1, whereas the true model should be that of

Figure 2. As explained in Section 3, this model misspecification may lead to mistaken advice on road congestion pricing schemes.

However, even considering that microscopic integrated urban models are fully robust and reliable, there are still additional challenges in resolving an optimization problem such as the one described in Figure 2. As stated before, multi-class assignment is not convex when the tolls depend on the flows and, therefore, traditional Newton methods, such as the ones described by Yang and Hang's (2004), cannot be used to determine optimal tolls. Instead, a different procedure can be formulated. Given that multi-class assignment with fixed tolls is convex in terms of the flows by link (Milchtaich, 2006), a full search of all possible tolls by link, determining the respective flow assignment and social cost, can be used to determine the system's optimum.

A full search of all possible tolls by link is impractical in a real network. However, this becomes simple enough for the network described in Figure 2. Note first that, given that optimality conditions shown in Eq. (2) must be fulfilled, only link 2 should have a toll in optimum. Therefore, the search for optimal flows can be achieved by determining the multi-class assignment for all possible tolls of link 2. This exercise is summarized by Figure 3, where is shown the total social cost that results for different values of the toll of link 2.

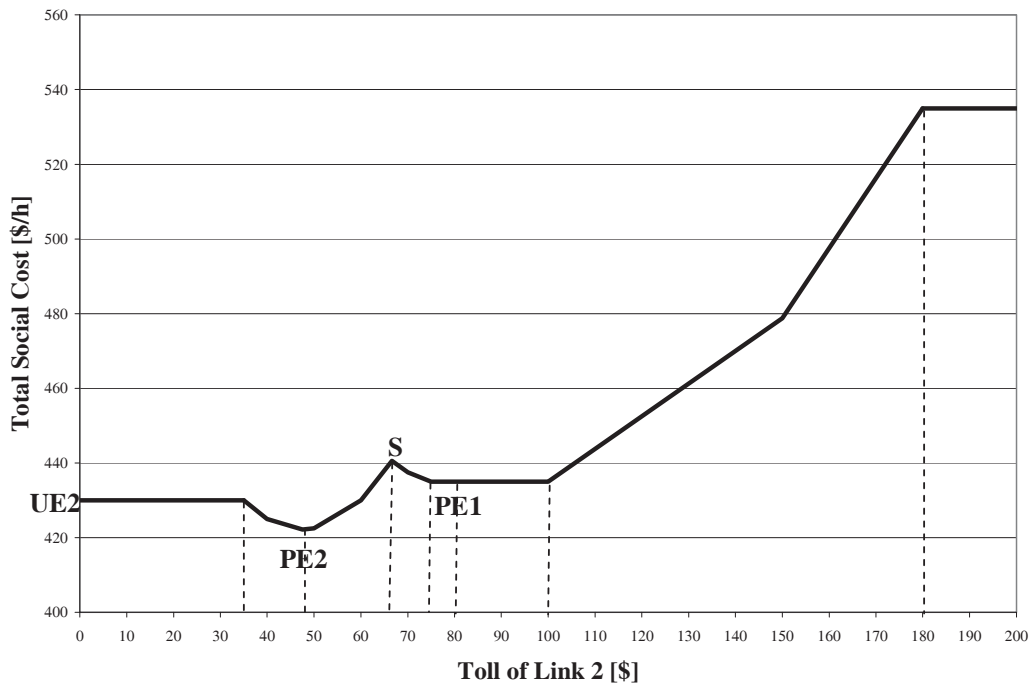


Figure 3: Total Social Cost for Various Tolls of Link 2 Including Latent Opportunities

For tolls of 180 and above, nobody uses link 2 and therefore the total social cost in that range is $4[\$/m] \cdot 55[m] \cdot 1[q/h] + 1[\$/m] \cdot 105[m] \cdot 3[q/h] = 535 [\$/h]$, as shown in Figure 3. For tolls in the range of 180 and 100, link 2 is attractive only for C users and social cost experience a decrease from 535 $[\$/h]$, when no C drivers use link 2, to 435 $[\$/h]$ when all C drivers use link 2. Then, for tolls between 100 and 75, all 1 $[q/h]$ C drivers but no L drivers use link 2. Note within this range is the toll of 80 that corresponds to the PE1 equilibrium described in Table 2, which also implies a social cost of 435 $[\$/h]$. Then, from tolls between 75 and 66.6, still all C drivers use link2, but some L drivers also begin using it and social cost begins to rise. Therefore, PE1 is a local optimum for the problem described in Figure 2.

Interestingly, the toll of 66.6 corresponds to a singularity (S in Figure 3) where there are infinitum equilibriums. Any combination of feasible flows of C and L drivers that sum up to 1.42 will be a solution with the same total social cost of 440.6 $[\$/h]$. Although the assignment for this toll is not unique in terms of flows by class, it is unique in terms of the flows by link and, therefore, in terms of total social cost. For tolls between 66.6 and 35, all C drivers move to link 3 and there is a convex section of the objective function. The minimum in this section occurs for a toll of 47.5, which implies the flows shown in Table 3 and the corresponding to total social cost of shown in Eq. (5).

Table 3: Pigouvian Equilibrium 2 (PE2)

Link	q_C [d/h]	q_L [d/h]	Time [m]	Toll [\\$]
1	0	0.625	105	0
2	0	2.375	57.5	47.5
3	1	0	55	0

$$\begin{aligned}
 SC_{PE2} = & \left(\underbrace{0}_{q_C} * \underbrace{4}_{VT_C} + \underbrace{0.625}_{q_L} * \underbrace{1}_{VT_L} \right) \underbrace{105}_{t_1} + \left(\underbrace{0}_{q_C} * \underbrace{4}_{VT_C} + \underbrace{2.375}_{q_L} * \underbrace{1}_{VT_L} \right) \underbrace{57.5}_{t_2} + \\
 & + \left(\underbrace{1}_{q_C} * \underbrace{4}_{VT_C} + \underbrace{0}_{q_L} * \underbrace{1}_{VT_L} \right) \underbrace{55}_{t_3} = 422 [\$/h]
 \end{aligned} \tag{5}$$

Figure 3 shows that toll of 47.5 produces the global minimum social cost for the problem described in Figure 2. Concordant with the result of Yang and Hang (2004), this optimum satisfies the Pigouvian conditions shown in Eq. (2). We will call this the Pigouvian Equilibrium 2 (PE2). Finally, for tolls between 35 and 0, the equilibrium is the same as that attained with UE2 shown in Table 2 and Eq. (4).

Comparing the net effect of the application of PE2 in Eq. (5) with UE2 in Eq. (4), it is clear that the intervention of the authority by imposing a Pigouvian tax produced an increase in social welfare of 8 $[\$/h]$. In UE2 there was a divergence between private and social cost, there was an external cost that was not internalized by the agents. In turn, in PE2, this divergence between social and private costs was corrected by means of a toll and, fully consistent with the Pigouvian theory, the new equilibrium resulted in a reduced social cost, in an increase of efficiency.

Although the determination of optimal tolls for the problem described in Figure 2 was feasible with the full search procedure described, it should be clear that finding the optimum in a real network would be almost impossible. In reality, even it would be possible to model all latent opportunities, arriving at local optimum like PE1 is very plausible. This result may be misleading for the researcher who, facing the fact that there is no divergence between private and social costs in PE1, may have a false sense of assurance (sustained in the widely accepted Pigouvian tradition) that the tolling scheme is socially superior to a selfish equilibrium.

But even if PE2 could be obtained, there are some critiques to this policy that can be raised. First, note that the price of anarchy in this example is quite small. The application of first-best Pigouvian tolls produced a reduction in social costs that was below 2% comparing UE2 with PE2. This means that many other policies capable of producing a small reduction of congestion may outperform road congestion pricing. Furthermore, if the costs that the society have to incur in the collection of the Pigouvian tolls are larger than 8[\$/h], the net effect of the measure would be a social loss. On top of this, the application of the Pigouvian toll made most drivers worse off and was regressive. Comparing UE2 with PE2, it can be noted that C drivers are indifferent, but L driver are worse off. In PE1 L drivers spend an equivalent of 105 [m], and in UE2 they were spending only 70 [m]. These facts show that, even if we would be able to model the latent opportunities and to solve the non-convex optimization problem, the skepticism of users and authorities toward road congestion pricing is not only logical, but also in many cases may be aligned with social interests.

Finally, it worth noting that Engelson and Lindberg (2006) claim to have provided a counterexample that seem to refute Yang and Huang's (2004) statement on the convexity, in the same sense that Figure 3 does it in this paper. However, Engelson and Lindberg (2006, Section 5) show using an example, is that multiclass-assignment is not convex in terms of the flows by classes, a statement that does has no implications at all for Yang and Huang (2004) model.

5 Impact on the Design of Policies to Alleviate Congestion

Tolls are powerful tools to shape drivers behavior. However, this research showed that the design of tolls that would actually led to a social improvement is more complicated than currently understood by transportation academia. It is not true that a full internalization of congestion externalities would necessarily lead to system's optimum. Pigouvian tolls (marginal social cost pricing) will work if and only if the problem is convex and, when the value of time is heterogeneous, transportation systems are not convex. Pigouvian tolls in non-convex problems may led the system to a local optimum and, depending on the relationship between the externality and the opportunities (alternatives) available to each agent, this policy may preclude latent opportunities, making these local optimums inferior to selfish equilibrium.

This research showed that the modeling complexity associated with the determination of optimal tolls is larger than usually understood. Before, to implement tolls like the ones suggested by Yang and Huang (2004) it would only be necessary to monitor streets and to

develop some type of surveys to recover users' composition by link. In turn, the counterexample deployed in Section 4 showed that the calculation of appropriate road congestion pricing tolls would require, in general, not only knowing the marginal external congestion costs, but also the opportunity cost of each agent, the true set of all latent opportunities. In the most general case, we would need to know and be able to forecast, for example, the activity participation and scheduling of each and every individual in a city. Before such modeling tools become available, the risk of implementation of road congestion pricing would be arguably too high.

But, even if the development of suitable microscopic integrated urban system models is successful, there is still an additional challenge to be resolved. The determination of the optimal tolls in a complex urban system would require the resolution of an ad-hoc optimization problem where, on the one hand, the authority is trying to determine the optimal set of tolls for the city and, on the other hand, the citizens respond to authority's actions by making choices. Such optimization problem would be nonlinear and non-convex and, therefore, algorithms to solve it would have to be very sophisticated. We showed that it is possible to solve this problem, for the simple example described in Figure 2, by performing a full search of all possible tolls in the system. However, the implementation of such a procedure in practice seems impossible. In other words, the risk of ending up determining a set of tolls that will cause a social detriment in a real application can be unmanageable.

The issues raised in this paper against road congestion pricing are not exclusive for the problem described in Figure 2. For example, even if activity participation and scheduling is homogenous, Pigouvian tolls may cause a social loss. Consider that, instead of the topology shown in Figure 2, all drivers depart from O and, instead of links 3 and 4, there is a new link 5 to reach D. Consider that link 5 represents an alternative transportation mode that is marginally priced at \$160 and for which travel time is 15 [m]. Despite link 5 is topologically available to both types of drivers, the differences in value of time would make this mode unattractive to L drivers, but fully equivalent to the former link 3 for C drivers. Therefore, the same results attained with the problem described in Figure 2, would be reached for this alternative problem. This last example is a possible explanation for a result obtained by Guevara et al. (2008). In that paper the authors used the software ESTRAUS (De Cea et al., 2000) to study the impact of calculating first-best Pigouvian tolls in the city of Santiago, Chile. ESTRAUS is a four-stage transportation model that was calibrated using real data, works at the level of zones, and where the problems of trip distribution, modal split and trip assignment are solved simultaneously. Within this context, Guevara et al. (2008) found not only that dedicated bus lanes can outperform first-best Pigouvian tolls, but also that Pigouvian tolls may result in a social detriment. Therefore, Guevara et al. (2008) could be seen as an example of the issues studied in this paper, but with real data.

Furthermore, it is even not indispensable to have different values of time to have that first-best Pigouvian may cause a social detriment. Milchtaich (2006) showed that multi-class assignment is not convex when costs' differences by class are not independent of the flows. Therefore, the same issues illustrated here would arise when drivers differ in their external costs by, for example, using cars of different sizes or by driving at different speeds. The analysis of these alternative settings is left for further research.

Before appropriate modeling and optimization tools are developed, two lessons can be learned from this research regarding the design of public policies to alleviate congestion. First of all, a blind commitment with congestion pricing is not recommendable. Other public policies like dedicated bus lanes or traffic management should be carefully assessed with available modeling tools before a formal decision toward congestion pricing is made. Second, if in some cases congestion pricing seems to be the unique alternative available, the counterexample described in Figure 2 suggests that heterogeneity of values of time might make this policy problematic. This means on the one hand that, when users' value of time is more homogenous, pricing is more likely to succeed and, on the other hand, that when users' value of time is less homogenous, palliative measures should be considered to alleviate the impact of road congestion pricing over users with a low valuation. It should be remarked however that the palliative measures should be performed in this case not to improve peoples' acceptability of road pricing nor to improve equity, but to be able to achieve an increase in social welfare with road pricing at all. The palliative measure could include subsidizing or improvement of alternatives that are available to users with low value of time (such as reducing travel time in link 1 in the example shown in Figure 2) or subsidizing the users with a low valuation of time so they have chance to access better latent opportunities (link 3 in Figure 2). The analysis of these and other palliative measures are left for further research.

6 Conclusion

The main purpose of this research was to question the consolidated belief that Pigouvian road congestion pricing is the ultimate tool to alleviate congestion. To do so, we first reviewed the main arguments that have been raised against road congestion pricing within the Pigouvian framework. Then, extending Coase's (1960) critique to Pigouvian theory, we showed that first-best road congestion pricing may cause a social detriment. It was shown that, in order to increase social welfare by imposing tolls, it is essential to account not only for the external congestion costs, but also for the latent opportunities that are available to each agent. In doing so, it would first be necessary to develop suitable microscopic integrated models of the urban system, decision support tools that are still under development.

Furthermore, a counter-example is used to show, for the first time, that multi-class assignment with heterogeneous value of time is not convex in terms of the flows by link. This implies that, even if suitable modeling and optimization tools were available, social optimum cannot be guaranteed by making all drivers internalize their full external cost. Because of the non-convexity, Pigouvian equilibrium could be local optimums, and even inferior to a situation in which external costs are not internalized at all. Formally, full internalization of external costs is a necessary but not an enough condition for achieving social optimum in multi-class assignment.

Following Coase's (1960) insights, this article is a call for a change in perspective in the design of policies to alleviate road congestion. When analyzing policies, transportation specialists should focus on the net effect for the whole society and not just on the correction of the distortion between social and private cost. Before sophisticated enough modeling

tools are available, a blind commitment toward road congestion pricing is unadvisable. There is a choice in the design of policies to alleviate congestion and it should be recalled that, even if truly optimal tolls could be obtained, there are some alternative policies that may outperform road congestion pricing and may be easier to implement.

Regarding future lines of research in this area, it would first be interesting to develop more counterexamples to explore the conditions that may make Pigouvian tolls socially detrimental. That type of studies could be based on traffic simulation or in the development of the formal mathematical conditions required for the counterexample to hold. Such analysis should be useful to shed more light on which cases road pricing might or might not be a suitable tool to alleviate congestion. The type of models that might be considered on that analysis should include cases where vehicles have different external costs, where public transportation services coexist with cars, and where some level of cross congestion is allowed, such as in Guevara (2008). Another line of research corresponds to the determination of the most suitable palliative measures for drivers with a low valuation of time that would make road pricing an appropriate tool to alleviate congestion when users are heterogeneous. Finally, a full analysis of the impact of the findings of this research in a model representing the equilibrium of a real city should also be of interest.

References

- Almeida, A., M. Ben-Akiva, F. Pereira, A. Ghauche, C. Guevara, S. Niza and C. Zegras (2009), "A Framework for Integrated Modeling of Urban Systems," presented at the 45th **ISOCARP Conference**, Porto, Portugal.
- Basso, L., Guevara, C., Gschwender, A. and M. Fuster, (2011), "Congestion Pricing, Transit Subsidies and Dedicated Bus Lanes: Efficient and Practical Solutions to Congestion," **Transport Policy**, Forthcoming.
- Bowman, J. (1998), "The Day Activity Schedule Approach to Travel Demand Analysis," Ph.D. Thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.
- Coase, R. (1960), "The problem of Social Cost," **Journal of Law and Economics** 3, 1-44.
- Daganzo, C. and R. Garcia (2000), "A Pareto Improving Strategy for the Time-Dependent Morning Commute Problem," **Transportation Science**, 34(3), 303-311.
- De Borger, B. and S. Proost (2010) "A Political Economy Model of Road Pricing," Working Paper 2010014, University of Antwerp, Faculty of Applied Economics.
- De Cea, J., Fernández, J. and A. Soto (2000), "ESTRAUS: Un Modelo de Equilibrio Simultáneo para Analizar Impactos y Apoyar la Evaluación Social de Planes Estratégicos de Transporte Urbano," **XI Congreso Panamericano de Ingeniería de Tránsito y Transporte**, Gramado, Brasil.
- Dial, R. (1999a) "Network-optimized Road Pricing. Part 1: A Parable and a Model," **Transportation Science** 47, 54-64.
- Dial, R. (1999b), "Network-optimized Road Pricing. Part 2: Algorithms and Examples," **Transportation Science** 47, 327-336.
- Florian, M. (1998), "Network Equilibrium Models for Analyzing Toll Highways," in: **Proceedings of the International Conference on Transportation into the Next Millenium**. Singapore.

- Gálvez, T. and S. R. Jara-Díaz (1998), "On the Social Valuation of Travel Time Savings," **International Journal of Transport Economics**, 25 (2), 205-219.
- Goodwin, P. (1989). "The Rule of Three: A Possible Solution to the Political Problem of Competing Objectives for Road Pricing." **Traffic Engineering and Control**, 30(10), 495-497.
- Guevara, C. (2010), "Endogeneity and Sampling of Alternatives in Spatial Choice Models," Ph.D. Thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.
- Guevara, C. (2008), "Sobre la Tarificación por Congestión en Sistemas Mixtos", **Ingeniería de Transporte**, 13(3), 12-17.
- Guevara, C., Valdes, E. and J. Dávila (2008), "Urban Transportation System Optimum: The Effect of Accounting for the Interaction of Buses and Cars" presented at the **European Transport Conference**, Leeuwenhorst, The Netherlands.
- Hau, T. (2005), "Economic Fundamentals of Road Pricing: A Diagrammatic Analysis, Part I—Fundamentals," **Transportmetrica**, 1(2), 81-117.
- Ison, S. (2005), "Congestion Charging: Selling the Concept," **Proceedings of the Institution of Civil Engineers: Civil Engineering**, 158 (5), 19-25.
- Jones, P. (1998), "Urban Road Pricing: Public Acceptability and Barriers to Implementation," in: Button and Verhoef, eds. **Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility**, , 263-284. Edward Elgar: Cheltenham, UK.
- King, D., M. Manvillea and D. Shoup (2007), "The Political Calculus of Congestion Pricing," **Transport Policy**, 14 (2), 111-123.
- Lindsey, C., and E. Verhoef, (2001), "Traffic Congestion and Congestion Pricing" in: Hensher, y Button, eds. **Handbook of Transport systems and Traffic control**, 77 - 105.
- Mas-Colell, A. (1995), **Microeconomic Theory**, New York : Oxford University Press.
- Milchtaich, I. (2005), "Topological Conditions for Uniqueness of Equilibrium in Networks," **Mathematics of Operations Research**, 30 (1), 225-244.
- Miller, E., J. Hunt, J. Abraham and P. Salvini (2004), "Microsimulating Urban Systems," **Computers, Environments and Urban Systems**, 28, 9-44.
- Nie, Y. and Liu. Y. (2010) "Existence of self-financing and Pareto-improving congestion pricing: impact of value of time distribution," **Transportation Research Part A**44, 39-51.
- Parry, I. and Bento, A., (2001). "Revenue Recycling and the Welfare Effects of Road Pricing", **Scandinavian Journal of Economics**, 103(4), 645-71.
- Pigou, A. (1920), **The economics of Welfare**, McMillian, London.
- Prud'homme and Bocarejo (2005), "The London Congestion Charge: A Tentative Economic Appraisal," **Transport Policy** 12(3), 279-287
- Roughgarden, T. (2005), **Selfish Routing and the Price of Anarchy**, MIT press, Cambridge, MA.
- Rouwendal, J. and E. Verhoef (2006), "Basic Economic Principles of Road Pricing: From Theory to Applications," **Transport Policy**, 13 (2), 106-114.
- Santos, G. (2004), "Urban Congestion Charging: A Second-best Alternative," **Journal of Transport Economics and Policy**, 38, 345-369.

- Sheffi, Y. (1985), **Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods**, Prentice Hall, Englewood Cliffs, New Jersey.
- Small, K. (1992). "Using the Revenues from Congestion Pricing," **Transportation**, 19(4), 359-381.
- Small, K. and E. Verhoef (2007), **The Economics of Urban Transportation**, London. Routledge.
- Strauch, D., R. Moeckel, M. Wegener, J. Gräfe, H. Muhlhans, G. Rindsfuser and K. Beckmann (2005), "Linking Transport and Land Use Planning: The Microscopic Dynamic Simulation Model ILUMASS," in Atkinson, Foody, Darby and Wu eds. **Geodynamics**, CRC Press, Boca Raton, Florida, 295-311.
- Verhoef, E (2005), "Second-best Congestion Pricing Schemes in the Monocentric City," **Journal of Urban Economics** 58, 367-388.
- Verhoef, E. and Small, K. (2004), "Product Differentiation on Roads: Constrained Congestion Pricing with Heterogenous Users," **Journal of Transport, Economics and Policy**, 38 (1), 127-156.
- Vickerey, W. (1963), "Pricing in Urban and Suburban Transport," **American Economic Review**, Papers and Proceedings 53 (2), 452-465.
- Vickery, W. (1955), "Some Implications of Marginal Cost Pricing for Public Utilities," **American Economic Review**, 45, 605-620.
- Waddell, P., L. Wang and X. Liu (2008), "UrbanSim: An Evolving Planning Support System for Evolving Communities," in Brail eds. **Planning Support Systems for Cities and Regions**, Lincoln Institute for Land Policy, Cambridge, MA, 103-138.
- Walters, A. (1961), "The Theory and Measurement of Private and Social Cost of Highway Congestion," **Econometrica**, 29(4), 676-697.
- Yang, H. and H. Huang (2004), "The Multi-class, Multi-criteria Traffic Network Equilibrium," **Transportation Research Part B** 38 (2004), 1-15.
- Zhang, X. and H. Yang (2004), "The Optimal Cordon-based Network Congestion Pricing Problem," **Transportation Research Part B** 38, 517-537.