Title
The Value of "Value Pricing" of Roads: Second-Best Pricing and Product Differentiation

Authors
Small, Kenneth A.
Yan, Jia

Publication Date
1999-11-01

UCI-ITS-WP-99-6

The Value of "Value Pricing" of Roads: Second-Best Pricing and Product Differentiation

UCI-ITS-WP-99-6

Kenneth A. Small
Jia Yan

Department of Economics and
Institute of Transportation Studies
University of California, Irvine
Irvine, California 92697-5100,U.S.A., ksmall@uci.edu

November 1999

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

ABSTRACT:

Some road-pricing demonstrations use an approach called "value pricing", in which travelers can choose between a free but congested roadway and a priced roadway. Recent research has uncovered a potentially serious problem for such demonstrations: in certain models, second-best tolls are far lower than those typically charged, and the welfare gains from profit maximization are small or even negative. That research, however, assumes that all travelers are identical and it therefore neglects the benefits of product differentiation, by which people with different values of time can choose a suitable cost/quality combination. Using a model with two user groups, we find that accounting for heterogeneity in value of time is important in evaluating constrained policies, and improves the relative performance of policies that offer differential prices. Nevertheless, for most of the reasonable range of heterogeneity, second-best pricing produces far fewer benefits than pricing both roadways optimally, and profit-maximizing tolls are so high that overall welfare is reduced from the no-toll baseline.

KEYWORDS: value pricing, congestion pricing, value of time, road pricing, high occupancy/toll lanes

ACKNOWLEDGMENT: The authors are grateful to Richard Arnott, Amihai Glazer, C. Robin Lindsey, Erik Verhoef, three anonymous referees for Transportation Research Board, and seminar participants at Boston College, Georgetown, Harvard, M.I.T., and VATT (Helsinki) for useful discussions.

THE VALUE OF "VALUE PRICING" OF ROADS:
SECOND-BEST PRICING AND PRODUCT DIFFERENTIATION

1. Introduction

Road-pricing concepts have moved to center stage in many transportation planning and policy-making venues around the world. Small and Gómez-Ibanez (1998) describe thirteen significant applications under consideration in nine countries, seven of them implemented as of mid-1997. More projects have been undertaken subsequently, including an innovative no-cash system using combined electronic and video collection technology on a new expressway near Toronto, Ontario, which opened in October 1997. Meanwhile, hardly an issue of the monthly Toll Roads Newsletter goes by without accounts of new pricing proposals by government agencies.

Yet in only one case (Singapore) has congestion pricing been adopted in something like a first-best form: significant time-of-day variations applying to an entire road network. All other applications are limited, such as toll rings with fixed or nearly fixed tolls (Norway), behavioral experiments (Stuttgart), or pricing on a single facility (France, Ontario, California, Texas, Florida). Increasingly, the favored approach is to adopt small-scale "demonstration projects" intended to test and publicize pricing concepts and their associated technologies. This approach is specifically funded in U.S. legislation passed in 1991 and reauthorized in 1998.

Three of the demonstrations currently operating – in Orange County (California), San Diego, and Houston – let travelers choose between two adjacent roadways: one free but congested, the other priced but free-flowing. This scheme is sometimes called "value pricing" because people are given the option to pay for a more highly valued service, much as train or air travelers can purchase a first-class ticket. In these particular examples, the express lanes also serve carpools at zero or at reduced rates, and so are known as "High Occupancy/Toll" (HOT) lanes. (In Houston, furthermore, the value-pricing option is available only to people in two-person carpools.)

Recent research, however, has uncovered a potential problem with value pricing as a demonstration of road pricing. This research examines the nature of "second-best" pricing of two parallel roadways when one is free [Braid (1996), Verhoef et al., (1996), Liu and McDonald (1999)]. An application of these methods by Liu and McDonald [1998], designed to approximate conditions for the Orange County value-pricing demonstration, suggests that in a second-best optimum, the express toll would be far lower than the tolls actually being charged and the express lanes would operate with considerably more congestion than they actually do. Furthermore, Liu and McDonald find that pricing the express lanes lowers welfare compared to leaving them free. In other words, the demonstration cannot be shown, based on their model, to make people better off compared to using the lanes for general traffic. This is obviously a potentially serious weakness in a strategy of using such demonstrations to gain public support for broader pricing schemes.

However, the Liu-McDonald analysis, like the other papers mentioned above, makes the simplifying assumption that all travelers are identical. This assumption obscures the benefits of offering a differentiated product in order to allow people to indulge their varying preferences. To analyze the situation fully, we need a model that includes variation in value of time.

This paper explores the importance of heterogeneity in value of time for value-pricing demonstrations. We extend the Liu-McDonald model to two user groups differing by value of time (after first simplifying the model by considering just one time period). We find that heterogeneity can make a significant difference in evaluating revenue-maximizing and second-best policies. Still, only with quite extreme assumptions can we find positive welfare benefits for private (i.e. revenue-maximizing) ownership of the express lanes compared to making them free. We also examine a policy, adopted in the San Diego demonstration, of setting the toll just high enough to maintain a specified level of service on the express lanes; we find this policy to perform only slightly better than the revenue-maximizing policy on welfare grounds.

A few other papers have addressed heterogeneity in value of time in a two-route problem. Arnott et al. [1992] use a dynamic bottleneck model to investigate first-best pricing in such a context, also with two types of travelers. They show that separating the two user groups on two roadways may be optimal if one group has both higher travel-time and schedule-delay costs than the other. Bradford [1996] shows that in a queue system with multiple servers, a revenue-maximizing system administrator would charge higher tolls, hence offer lower congestion, than is socially optimal. More directly related to our case is Schmanske (1991, 1993), who shows that with heterogeneous users, differential tolls on separate roadways may be superior to a single toll. Verhoef and Small [1999] consider heterogeneity using a continuous value-of-time distribution, calibrated from Dutch stated-preference data, and also account for the possibility that users of the two roadways interact on a congested serial link elsewhere as part of their trips; their conclusions are broadly consistent with those of this paper.

Our analysis does not purport to be a complete assessment of the SR91 or any other actual demonstration projects, which are often constrained by a variety of financial and legal considerations. In particular, we do not treat either incentives for high-occupancy vehicles (HOVs) or capacity costs. Small (1983) and Dahlgren (1998) consider HOV lanes, and Viton (1995) examines the question of when financing highway capacity through private toll collection is viable.

2. The Model

We consider two roadways, A and B, connecting the same origin and destination. Both have the same length L and the same free-flow travel-time $T_f L$. A user of type i (i=1,2) traveling on road r (r=A,B) incurs travel cost c_{ir} which consists of operating cost β plus a time cost $\alpha_i T_r$ per unit distance. The parameter α_i is the value of time, and it is this parameter for which we introduce heterogeneity, by assuming that $\alpha_1 > \alpha_2$. Unit travel time T_r (the inverse of speed) is represented by flow congestion of a standard type, depending on volume-capacity ratio N_r/K_r so that:

$$c_{ir}(N_r) = \beta L + \alpha_i T_f L \left[1 + \gamma \left(N_r / K_r\right)^k\right] \quad i = 1,2; r = A, B \qquad (1)$$

where γ and k are parameters. The congestion-dependent part of cost, $d_{ir} \equiv \alpha_i T_f L \gamma (N_r/K_r)^k$, is what we call *delay cost.*[1] We use values γ=0.15 and k=4, following common practice.[2]

Demand by each group has the linear form

$$N_i(P_i)= a_i - b_i P_i \tag{2}$$

where a_i and b_i are positive parameters and P_i is the "inclusive price", defined as the minimum combination of travel cost plus toll (τ) for this user group:

$$P_i = \underset{r}{Min}\{c_{ir} + \tau_r\} \tag{3}$$

The inverse demand function for user type i is denoted $P_i(N_i)$, and easily solved from (2).

The social welfare function is defined as the area under the inverse demand curve minus total cost:

$$W = \sum_{i=1}^{2} \int_{0}^{N_i} P_i(t)dt - \sum_{i=1}^{2} \sum_{r=A}^{B} N_{ir} c_{ir} \tag{4}$$

where N_{ir} is the number of type-i users on road r. This function is strictly concave in the four variables N_{ir}.

2.1 Types of Solution

The equilibrium conditions are those of Wardrop [1952], stating that users of a given type choose the road or roads that minimize inclusive price, and that those inclusive prices be equalized, for those users, if they use both roads. We assume that if the roads are differentiated it is road A that offers faster travel, so that N_{1A}>0 and N_{2B}>0. This is a substantive restriction if the roads are of unequal capacity. Wardrop's conditions can then be written:

$$c_{1A}(N_A) + \tau_A \le c_{1B}(N_B) + \tau_B \tag{5.a}$$
$$c_{2A}(N_A) + \tau_A \ge c_{2B}(N_B) + \tau_B \tag{5.b}$$
$$N_{1B} \cdot (c_{1A} + \tau_A - c_{1B} - \tau_B) = 0 \tag{5.c}$$
$$N_{2A} \cdot (c_{2B} + \tau_B - c_{2A} - \tau_A) = 0 \tag{5.d}$$
$$N_{1B}, N_{2A} \ge 0 \tag{5.e}$$

[1] This particular functional form has the property that the marginal external cost, i.e. the additional delay cost by a driver on all others, is k times the average delay cost: $MEC_r \equiv \sum_i N_{ir} \partial c_{ir} / \partial N_{ir} = k \cdot \left(\sum_i N_{ir} d_{ir} \right) / N_r$.

[2] See Small [1992], pp,69-72, for a discussion of empirical evidence for this functional form. These particular parameters are known as the Bureau of Public Roads formula.

It is useful to distinguish four possible cases, depending on whether each of (5a) and (5b) is an inequality or an equality.

Case SE: fully separated equilibrium. Both (5a) and (5b) are inequalities, i.e., each group strictly prefers a different roadway. Because we assumed $\alpha_1 > \alpha_2$, these conditions require[3] that road A be more expensive but less congested than road B, i.e., $\tau_A > \tau_B$ and $(N_A/K_A) < (N_B/K_B)$.

Case SE1: partially separated equilibrium with group 1 separated. Group 1 strictly prefers road A, but group 2 is indifferent: that is, (5a) is an inequality but (5b) an equality. Like the fully separated equilibrium, SE1 requires that road A have higher toll but lower travel times. Note it is not impossible that $N_{2A}=0$, if this conditions happens to yield indifference for group 2; but we would expect this only by coincidence.

Case SE2: partially separated equilibrium with group 2 separated. Group 2 strictly prefers road B, but group 1 is indifferent: (5a) is an equality, (5b) an inequality. Again, road A must have a higher toll but is faster. The boundary solution $N_{1B}=0$ can occur; this possibility is in fact relevant because of the second-best optimization process, which may sometimes set the constrained-optimal toll just low enough to retain all type-1 users as toll-road customers. Thus despite the word "separated" in the names of these cases, it is the equality or inequality of costs in (5a-b), not the presence or absence of a given type of user on both roads, that formally distinguishes case SE2 from SE.

Case IE: fully integrated equilibrium. Both groups are indifferent between the two roads; (5a-b) hold with *both* inequalities replaced by equalities. Since the two groups have different values of time, this can occur only if the roads have equal tolls and equal speeds. We assume this equilibrium always applies if no tolls are charged, and it turns out that is the only time it applies.

2.2 Pricing Regimes

We consider five alternative pricing regimes, also called policies.

First-best regime (FB): a public operator charges tolls on both roads that maximize welfare (4). It can be shown that this policy yields conventional marginal-cost pricing on each road.

Second-best regime (SB): the same objective is pursued but subject to the constraint $\tau_B=0$.

Third-best regime (TB): like SB but with an additional constraint designed to guarantee a minimum level of service on the priced roadway, namely[4]

---

[3] Subtracting the second from the first of equations (5) and applying (1) yields $(\alpha_1-\alpha_2)(N_A/K_A)^k < (\alpha_1-\alpha_2)(N_A/K_A)^k$, which (given $\alpha_1 > \alpha_2$ and $k > 0$) implies $N_A/K_A < N_B/K_B$. This in turn implies $c_{2A} < c_{2B}$, so the second of equations (5) is possible only if $\tau_A > \tau_B$

[4] The particular value 0.887 is chosen because it is the maximum volume-capacity ratio for level of service D (Transportation Research Board, 1994, Table 3-1), which is the minimum level of service being sought in the 1999 reauthorization of the San Diego HOT lane.

$$\frac{N_A}{K_A} \leq 0.887 \tag{6}$$

Profit-maximizing regime (PM): τ_A is chosen to maximize revenues subject to the constraint τ_B=0.

No-toll regime (NT): τ_A ,τ_B=0.

The no-toll regime consists of solving (1)-(3) and (5) with equalities in (5a) and (5b); the solution is assumed to be of the integrated equilibrium (IE) type, since there is nothing to distinguish the two roadways from each other. Each of the other regimes calls for maximizing either welfare, as given by (4), or revenues $R=\sum_r \tau_r N_r$, while imposing constraints (5) and, in the case of third-best, constraint (6).

Our solution strategy[5] is first to choose an equilibrium case (SE1, SE2, or SE) to test. We form the relevant Lagrangian, simplifying by taking advantage of the requirement, by (5c-d), that one or both of N_{1B} and N_{2A} be zero, depending on the regime. (Specifically, N_{1B}=0 in regime SE1, N_{2A}=0 in SE2, and both are zero in SE.) We then solve the first-order conditions numerically for N_{ir} and τ_r. Next, we check the non-negativity constraints (5e); if either of them is not satisfied, we impose it as an equality and again solve the first-order conditions. In the case of TB, we also check the level-of-service constraint and, if it is violated, we impose it as an equality and start over. We then check the appropriate inequality (5a or 5b or both) defining the equilibrium type under consideration; if it is violated, we conclude that this equilibrium type cannot exist for this set of parameters. In this manner we generate up to three candidate solutions, (one for each equilibrium type), and we choose the one for which the maximized objective function is largest.

An example is instructive. Consider the SE1 equilibrium for the third-best (TB) policy regime. For this scenario τ_B=0, (5a) holds as an inequality and consequently N_{1B}=0, and (5b) holds as an equality. Therefore equations (3) and (5a-d) simplify to:

$$\tau_A = P_1 - c_{1A} \tag{7a}$$
$$P_1 - c_{1A} = P_2 - c_{2A} \tag{7b}$$
$$P_2 - c_{2B} = 0 \tag{7c}$$
$$P_1 - c_{1B} < 0 \tag{7d}$$

where it is to be remembered that P_i is a function of $(N_{iA}+N_{iB})$ through (2) and c_{ir} is a function of $(N_{1r}+N_{2r})$ through (1). We solve the problem by using ordinary Lagrangian methods to find the values of N_{1A}, N_{2A}, and N_{2B} that maximize (4) subject to equality constraints (7b) and (7c); then τ_A

[5] In the Appendix, we enumerate the full set of possible solutions. For most cases they are not of closed form, so require numerical maximization procedures to find them.

is calculated from (7a). The non-negativity constraint $N_{2A} \geq 0$ is then checked, and (4) is maximized again imposed as an equality if needed. Similarly the level-of-service constraint is checked and imposed if needed. Finally the inequality (7d) is checked to see if the trial solution is valid.

3. Simulation Results

In this section, we design several scenarios to explore the effects of heterogeneity in value of travel time on the efficiency of various pricing policies. We begin with a base scenario that resembles SR-91, the demonstration site in Orange County of California. We then consider alternate demand parameters, first changing the relative sizes of groups 1 and 2, then changing price elasticities. Next we consider a scenario with much heavier traffic. Finally we alter the relative capacity of the two roadways, making road A the larger one. Table 1 presents the parameters used in these scenarios. Except for the unit value of travel time, the cost parameters are the same as in the Liu-McDonald paper.

Table 1. Parameter Values Used in Simulations

Parameter	Base Scenario	Proportional-Demand Scenario	High-Elasticity Scenario	High-Congestion Scenario	Reversed-Capacity Scenario
β (cents/mi.)	6.8	6.8	6.8	6.8	6.8
K_A (veh./hr.)	2000	2000	2000	2000	4000
K_B (veh./hr.)	4000	4000	4000	4000	2000
a_1	5700	3800	7150	6780	5700
a_2	5700	7600	7150	6780	5700

Notes:

1. The following parameters are the same in all scenarios: L=10 miles; $\gamma = 0.15$; $k = 4$; $T_f = 0.9231$

2. Average value of time is defined as: $\left(N_1^{NT}\alpha_1 + N_2^{NT}\alpha_2\right)/\left(N_1^{NT} + N_2^{NT}\right)$ and it is 34.8 cents/min. in all scenarios. N_i^{NT} is the number of type i users in no-toll regime.

3. At each point of value-of-time difference, the slopes of demand functions is chosen to maintain the elasticities of two groups at −0.60 in high-elasticity scenario and at −0.33 in other scenarios and the time difference between roads under PM regime is 15 minutes in high-congestion scenario and 8 minutes in other scenarios except reversed-capacity scenario.

To preserve comparability with Liu and McDonald, we mostly use the same parameters: L=10 miles (16.1 km), β=6.8 cents per vehicle-mile (4.72 cents/veh-km), T_f=65 miles per hour (105 km/hr), and capacities K_A=2000 and K_B=4000 vehicles per hour.

3.1 Base Scenario

In this scenario, we choose the demand parameters so that in the no-toll (NT) regime the price elasticity of demand is -0.33 as in Liu and McDonald, and so that our profit-maximizing (PM) policy produces a toll of about \$2.75 and a travel time differential between routes of about 8 minutes, thereby replicating actual conditions on SR-91 in June 1997 (Sullivan, 1997). This is achieved with an average value of time of 34.38 cents/min. ($20.63/hr.), which is much higher than the value of $6.36 per hour in Liu and McDonald's paper.

Table 2. Results for Base Scenario Under Homogeneity

PRICING REGIME[a]	FB	SB	TB	PM	NT
Type of equilibrium[b]	SE2	SE2	SE2	SE2	IE
Toll[c] — A	389.21	72.61	267.29	275.53	0
Toll — B	389.19	0	0	0	0
Speed[d] — A	49.6	44.8	59.4	60	40
Speed — B	49.6	38.7	33.5	33.3	40
Delay Cost[c]					
1A	97.30	144.21	29.48	26.24	198.30
1B	97.34	216.82	296.77	301.78	198.30
2A	--	--	--	--	198.19
2B	97.28	216.69	296.60	301.60	198.19
Rel. Use[e] –1	0.84	0.99	0.94	0.94	1.00
Rel. Use — 2	0.84	0.99	0.94	0.94	1.00
Elast.[f] — 1	-0.59	-0.34	-0.41	-0.41	-0.33
Elast. — 2	-0.59	-0.34	-0.41	-0.41	-0.33
Welfare Gain per vehicle[g]	61	4	-40	-45	0

Notes:

[a] Pricing regimes: FB=first best; SB=seocnd best; TB=third best; PM=profit maximization; NT=no toll (see Section 2.2)

[b] Types of equilibrium: SE2=partially separated eq., group 2 separated; IE=integrated eq. (see Section 2.1)

[c] All costs (toll, delay cost, welfare gain) are in cents per vehicle. Delay cost is defined as $\alpha_i T_f L\gamma(N_r / K_r)^k$.

[d] Speed is in miles per hour.

[e] Relative use of group is relative to the no-toll regime, i.e. N_i / N_i^{NT}.

[f] Elast. is demand elasticity at usage level in the solution.

[g] Welfare gain is divided by usage in the NT regime, i.e. $\left(W - W^{NT}\right) / N^{NT}$.

The simulation results for homogenous users are shown in Table 2. The pattern of results is the same as in Liu-McDonald's. The welfare gain from second-best pricing (SB) is small, and that from one-route profit-maximizing policy (PM) is negative. The relative efficiency of the second-best compared to the first-best policy[6] is about 6% and that of profit-maximizing policy (PM) is about -74%; these compare to 9% and -50% respectively in Liu-McDonald. In addition, the second-best toll is much lower than the first-best toll, thus it has little effect on total traffic. The first-best toll is about 50 percent higher than the profit-maximizing toll and reduces total traffic by about three times as much. With no toll (NT), speed would be 40 miles per hour.

Now we turn to the effects of product differentiation by examining how the simulation results change when the two groups are assigned different values of travel time. We let α_1 and α_2 diverge by a given amount $\Delta\alpha$. At the same time we alter the slopes of demand functions to keep the

[6] Relative welfare gain is define as RW = $(W^{SB}-W^{NT})/(W^{FB}-W^{NT})$, where W is defined in equation (3) and the superscripts indicate policy regimes.

elasticity of two type users and the weighted average value of travel time (weighted by the number of users of each type in the no-toll regime) in no-toll regime unchanged. Results are shown in Figure 1. At the far left of each of panels, users are homogeneous. At the far right, the two groups' value of time are 2.37 cents/min. and 66.39 cents/min. The partially separated equilibrium SE2 remains optimal for all pricing policies; that is, group 1 users use both roads, which is not surprising because group 1 contains half the population of potential users but the express road contains only a third of the total capacity.

Figure 1a. Toll (Base Scenario)

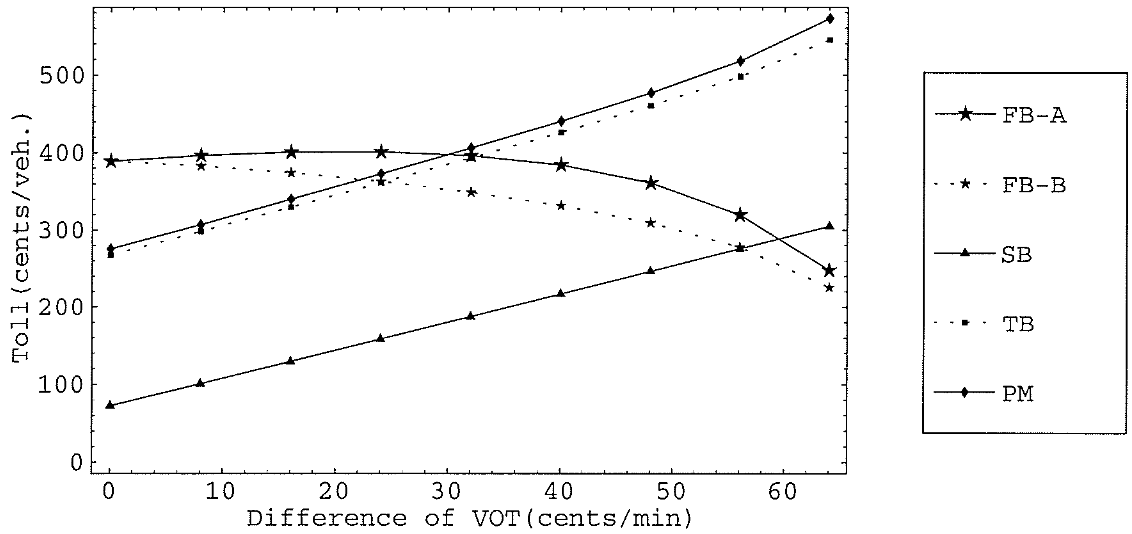


Figure 1a shows the tolls as the function of heterogeneity. In the three constrained pricing policies, the toll rises sharply with the difference in value of time. At the middle of the diagram, the second-best (SB) toll has nearly tripled compared to what it was with identical values of time, although it is still barely half the profit-maximizing (PM) toll. The third-best (TB) toll is nearly identical to that of PM.

The first-best (FB) toll is indeed differentiated, but there is a surprise here: the toll differential gets larger at first but then gets smaller again when heterogeneity is extreme. The reason is that when heterogeneity is large, the marginal benefit of accommodating one more type 1 user is bigger than that of accommodating one more type 2 user. The first-best policy therefore accommodates many more type 1 users than type 2 users on route B: the number of type 1 users increases by about 30% with the increase of heterogeneity, while the number of type 2 decreases by more than 30%. As a result, the difference between average values of travel time on the two routes becomes small.

Figure 1b shows the travel time on both routes under the second-best and profit-maximizing policies, as well as under the no-toll regime. Profit maximization (PM) creates a much greater quality differential between the two roads than does second-best, an indication of exercise of monopoly power on the priced roadway. The third-best regime (not shown) is almost identical to PM.

Figure 1c shows the welfare changes, all relative to no toll (NT). The welfare gains from all the differential-pricing policies are much greater when there is more heterogeneity. The efficiencies of the three constrained regimes also improve when measured as fractions of possible first-best welfare gains: for example, the SB welfare gain increases from 6% to 28% of FB. Even so, the profit-maximizing policy always produces a welfare loss (compared to no toll) and third-best pricing almost always does; and both perform consistently worse than second-best when evaluated according to welfare gain.

Figure 1b. Travel time (Base Scenario)

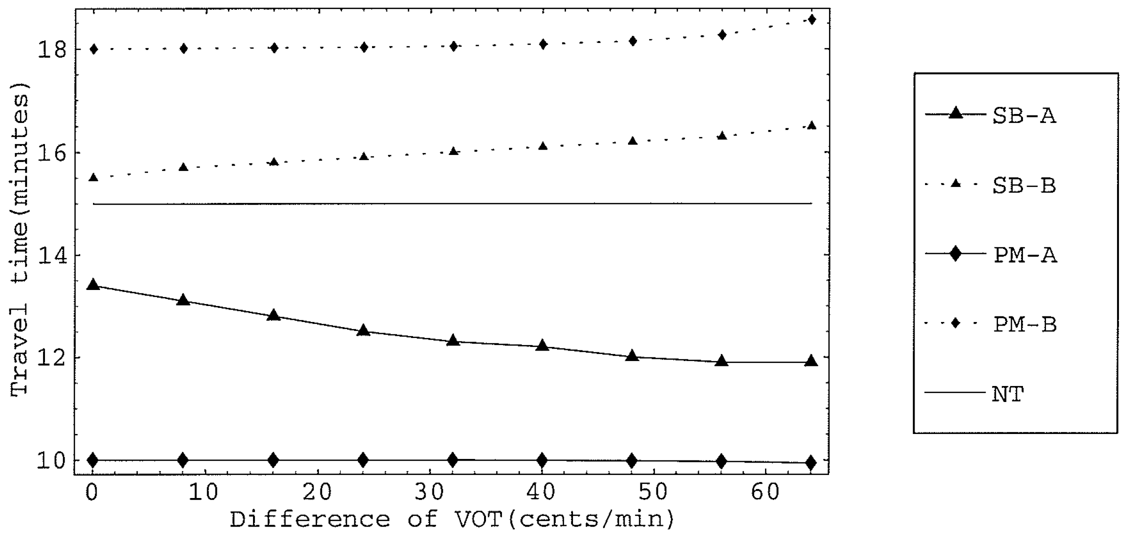
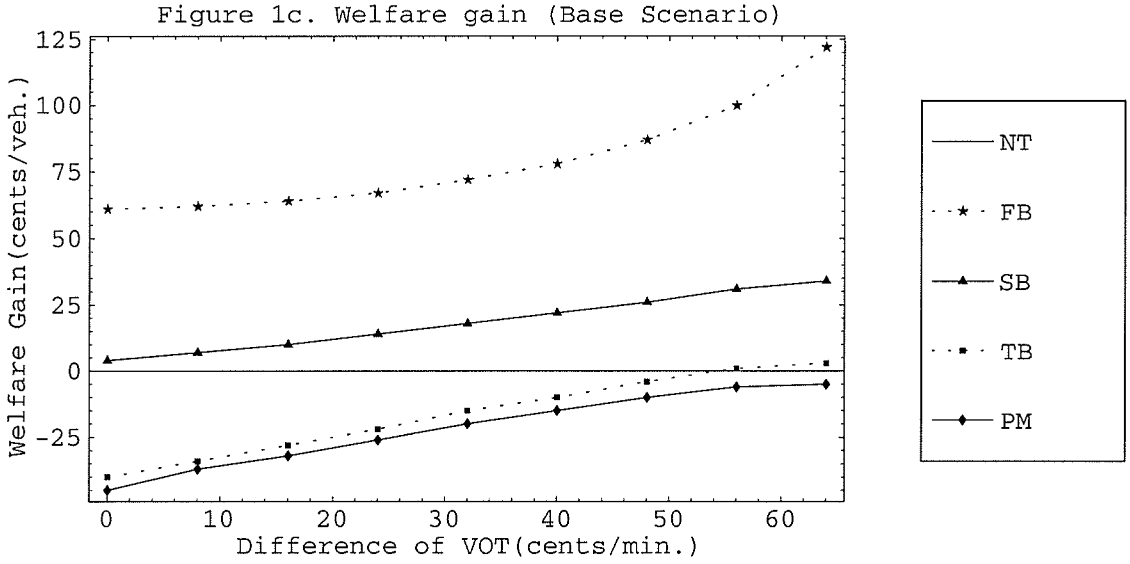


Figure 1c. Welfare gain (Base Scenario)



To check the sensitivity of our results to average value of times, we recalculate the base scenario using half the previous value, i.e. $10.32 per hour, while adjusting intercepts and slopes to maintain price elasticity of -0.33 and a time differential under PM of 8 minutes. The qualitative results do not change.

3.2 Proportional-Demand Scenario

In order to examine cases where product differentiation might be more important, we next consider a scenario where the numbers of users in the two groups are approximately proportional to the capacity of corresponding roadway. We accomplish this by setting the intercepts of the demand functions proportionally to the relative capacities, i.e. $\alpha_1 / \alpha_2 = K_A / K_B = 1/2$, while keeping the total demand under no toll fixed. The slopes of demand functions are also changed to make both types of user have the same elasticity as in the base scenario. Under homogeneity, the value of time is set at the same amount as in base scenario and the results are changed hardly from the base scenario.

Figure 2a. Toll (Proportional-Demanded Scenario)

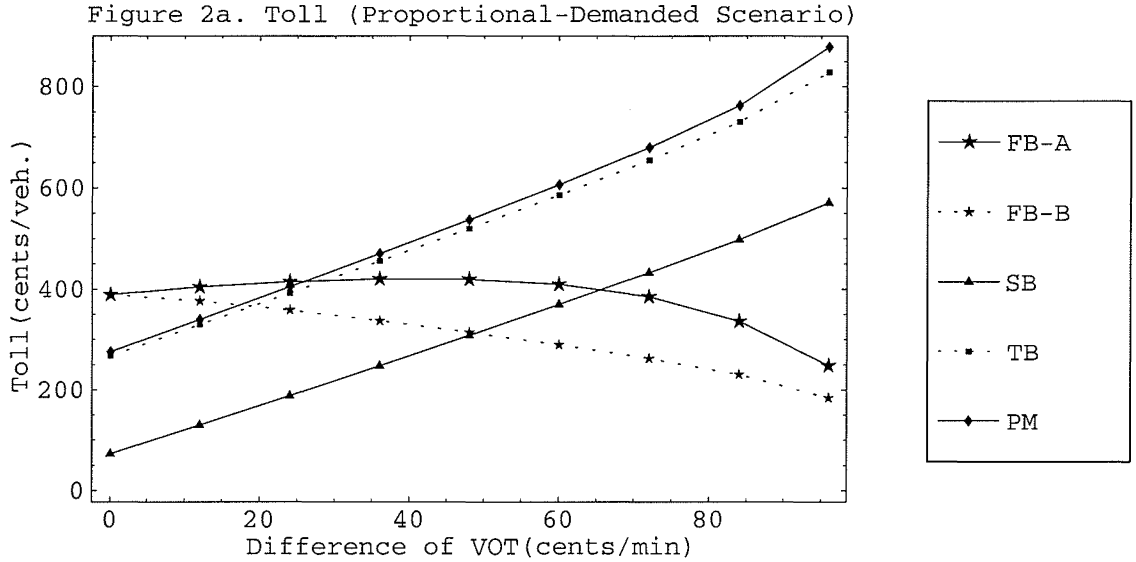
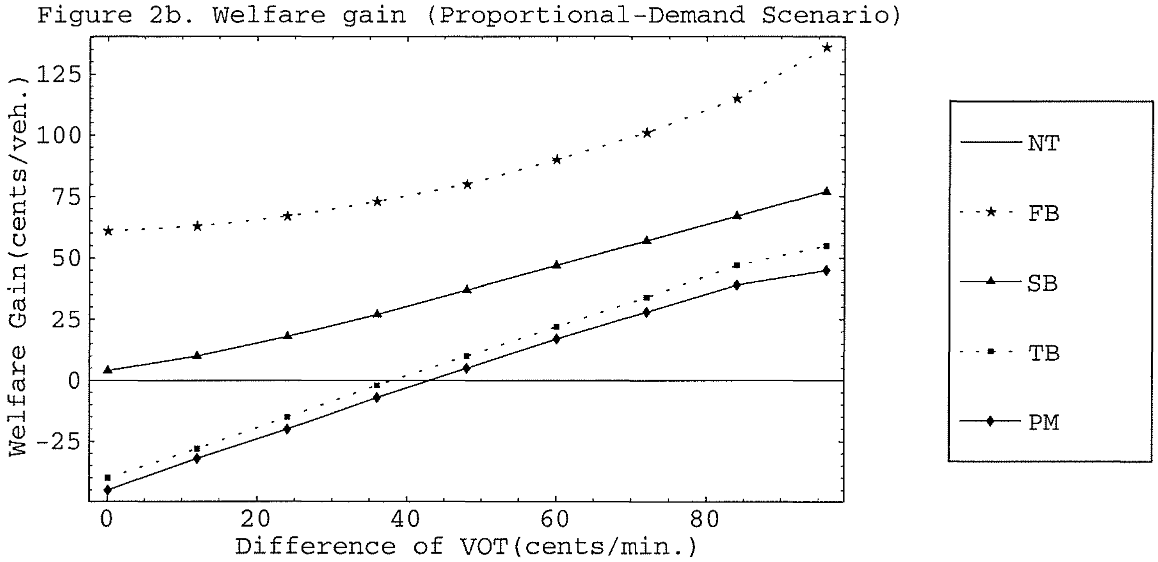


Figure 2b. Welfare gain (Proportional-Demand Scenario)



We introduce the heterogeneity in this scenario by increasing α_1 twice as fast as we decrease α_2. Thus the distribution of values of time becomes not only dispersed but also skewed. The slopes of demand functions are changed as in the base scenario. The results are shown in Figure 2. At the far right of each of the panels, the value of time of type 1 users is 2.37 cents/min., while that of type 2 users is 98.40 cents/min.

Figure 2a shows the change of tolls with value of time difference. The pattern of change is similar to base scenario. Figure 2b shows that the welfare gain from first-best (FB) pricing is almost the same as in the base scenario. But this time the TB and PM policies are considerably improved, generating positive welfare gains under moderate to large heterogeneity. Furthermore, the second-best policy is much more efficient in this scenario, with relative efficiency around 45% with moderate value-of-time differences. The reason for these results is that the differentiated product is better matched to the different user types in this scenario; fewer users are forced into the wrong quality.

The change of travel time under each policy in this scenario is almost the same as the one in base scenario, so is not shown.

3.3 High-Elasticity and High-Congestion Scenarios

Here we first consider a scenario with higher price elasticity of demand, namely -0.60 in the no-toll regime. The weighed average value of time is kept at 34.38 cents/min. Results are shown in Figures 3a and 3b.

Figure 3a shows that the second-best toll is much higher, and the first-best lower, in this scenario. This is well known from previous studies (Verhoef et al., 1996); welfare-maximizing policies are now aimed more at moderating total demand than at distributing demand across the two roads.

Figure 3a. Toll (High-Elasticity Scenario)

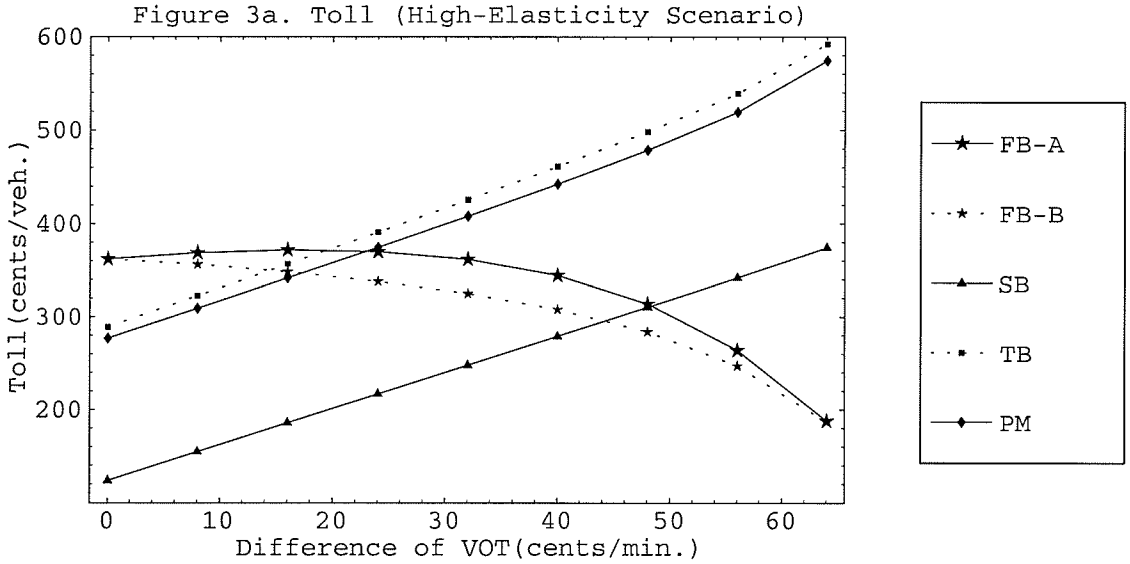


Figure 3b. Welfare gain (High-Elasticity Scenario)

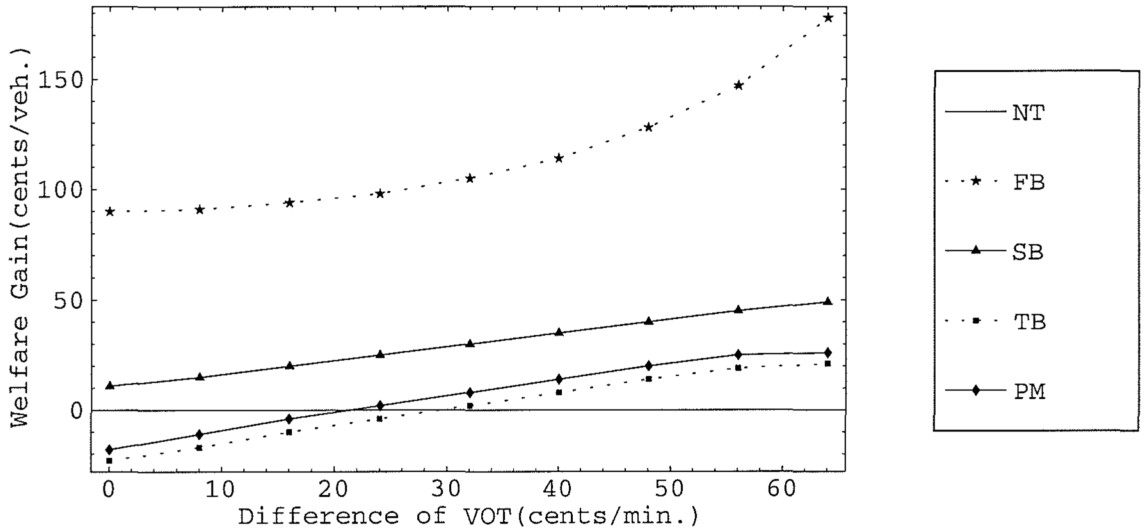


Figure 3b shows that the efficiency of the PM and TB policies is improved significantly. Both of them can generate positive welfare gain when value of time difference is greater than 30 cents/min. SB is not improved, because it emphasizes the toll differential, which is less important now. Thus the gap between SB and the other constrained policies is less, though still there.

Next, we consider a scenario with higher congestion, namely a travel-time differential of 15 minutes under PM. We again accomplish this by changing the intercepts and slopes of the demand functions.

Figure 3c. Toll (High-Congestion Scenario)

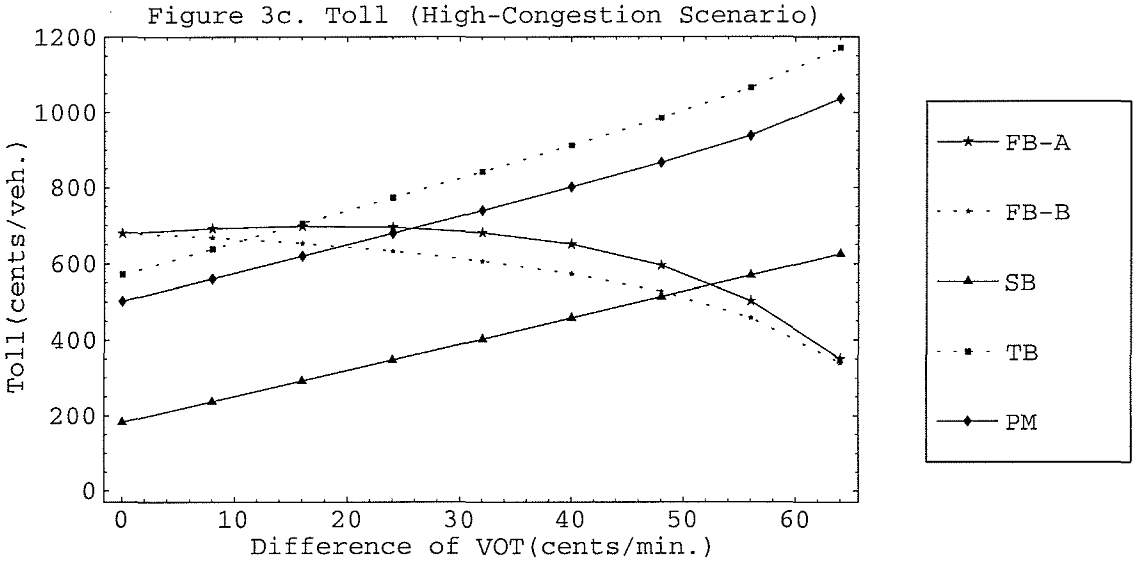
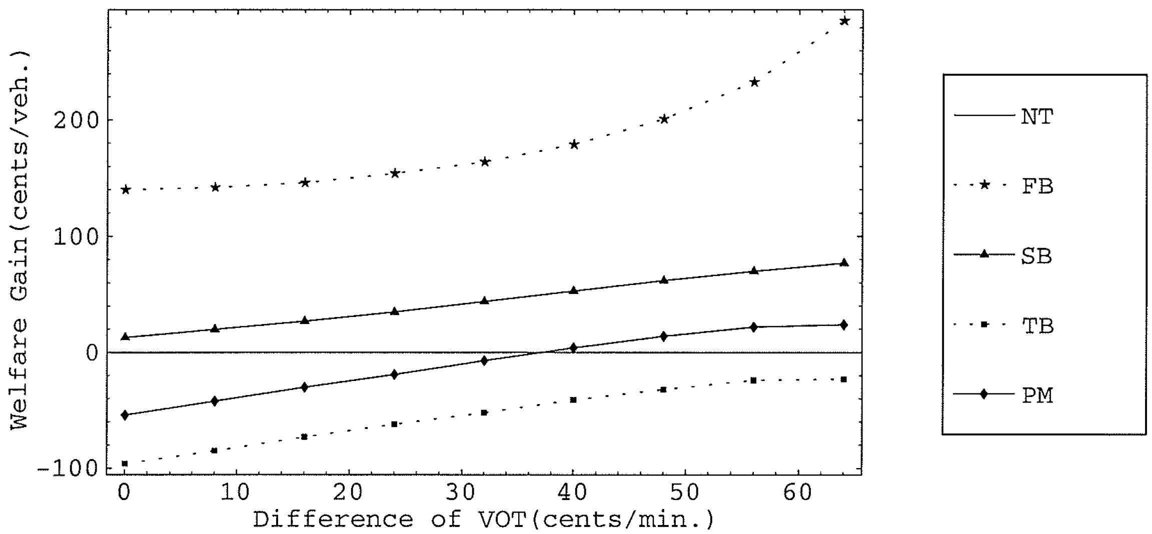


Figure 3d. Welfare gain (High-Congestion Scenario)



The results, shown in Figures 3c and 3d, are mostly similar to the base scenario, but two differences stand out. The TB policy produces a much higher toll than PM because of the heavier traffic; and PM now allows substantial congestion on the toll lanes. The welfare effects in this scenario are similar to those in the high-elasticity scenario.

3.4 Reversed-Capacity Scenario

In order to make a fully separated equilibrium more likely, we tried interchanging the two roadway capacities: 4000 veh/hr for the express lanes and 2000 for the free lanes. All other parameters are as in the base scenario.

Figure 4a. Toll (Reversed-Capacity Scenario)

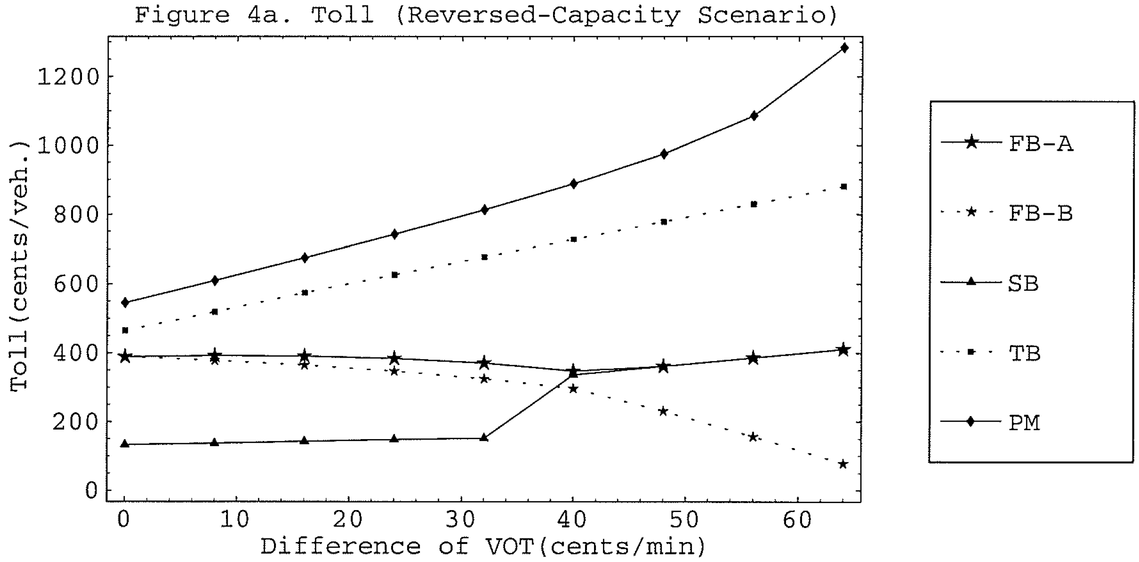
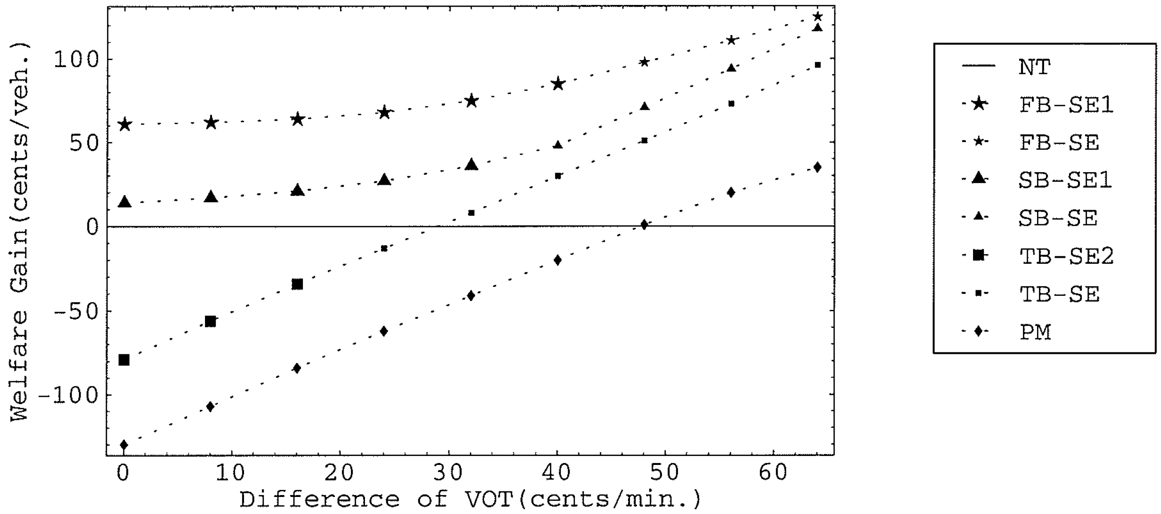


Figure 4b. Welfare gain (Reversed-Capacity Scenario)



Results are shown in Figure 4. The three one-route pricing policies have higher tolls in this scenario because the free roadway is less important as a substitute. SB has a higher welfare gain because it can charge for more capacity. PM and TB generate bigger welfare losses.

We get different equilibrium cases in this scenario. Most interesting, as heterogeneity is increased, user differences simply become too great to be worth accommodating on a shared roadway, and the optimal equilibria tend to become fully separated (SE).

When the value of time difference is extreme large, the welfare gain from SB is very close to that from FB. The relative efficiency of TB policy to FB policy at this point reaches 77%. The efficiency of PM policy is also improved compared with base scenario, and it can produce a positive welfare gain when the value of time difference is high.

4. Conclusion

Our results demonstrate the importance of heterogeneity in value of time for evaluating congestion policies that offer pricing as an option. Generally, the existence of heterogeneity favors such policies because product differentiation then offers a greater advantage: those with high values of time reap more benefits from the high-priced option, while those with low values of time find it all the more important not to be subjected to policies aimed at the average user.

Nevertheless, insisting that one of the products be free imposes quite a large penalty, except when heterogeneity is extreme. In our base scenario and for middling amounts of heterogeneity, a second-best one-route pricing policy achieves only one-fifth to one-half the possible welfare gains of first-best pricing, and uses a toll smaller than even the lower of the two optimally differentiated tolls.

Even more discouraging, policies that maintain nearly congestion-free travel in the priced roadway set the price far higher, and achieve far lower benefits, than second-best pricing. In the majority of cases, the overall benefits from pricing are negative for these policies. Of course, this does not account for the possibility that such policies may be the only way the lanes can be built at all, or the only way they can be opened to general traffic.

From these observations, we draw three conclusions about partial-pricing policies under highly congested conditions. The first two are in accord with studies based on homogeneous users. First, when politics or other considerations dictate that one roadway be free, aggregate costs can be reduced by letting the priced roadway become at least moderately congested; carpooling mandates or privatization goals may prevent this, but they do so at a heavy cost. Second, under many conditions partial pricing policies are inadequate substitutes for more thoroughgoing pricing policies. The third conclusion is that accounting for heterogeneity does improve the performance of partial-pricing policies by creating significant value for product differentiation, especially when the price-elasticities for total demand is high and congestion in the absence of tolls is extreme.

References

Arnott, R., A. de Palma and R. Lindsey (1992) "Route choice with heterogeneous drivers and group-specific congestion costs" *Regional Science and Urban Economics* **22** 71-102.

Bradford, Richard M. (1996) "Pricing, Routing, and Incentive Compatibility in Multiserver Queues," *European Journal of Operational Research* **89**, pp. 226-236.

Braid, Ralph M. (1996) "Peak-load pricing of a transportation route with an unpriced substitute" *Journal of Urban Economics* **40** (179-197).

Brownstone, D., Golob, T. F. and Kazimi, C. (1999) "Modeling non-ignorable attrition and measurement error in panel surveys: an pplication to travel demand modeling," Irvine Economics Paper 99-00-06, University of California at Irvine

Dahlgren, Joy (1998) "High occupancy vehicle lanes: not always more effective than general purpose lanes" *Transportation Research* **32 A**, pp. 99-114

Institute of Transportation Engineers (ITE) Task Force on High-Occupancy/Toll (HOT) Lanes (1998) "High-Occupancy/Toll (HOT) Lanes and Value Pricing: A Preliminary Assessment," *Institute of Traffic Engineers (ITE) Journal* (June), pp. 30-32, 38, 40.

Liu, Louie Nan, and John F. McDonald (1998) "Efficient congestion tolls in the presence of unpriced congestion: a peak and off-peak simulation model" *Journal of Urban Economics* **44**, pp. 352-366.

______ (1999) "Economic efficiency of second-best congestion pricing schemes in urban highway systems" *Transportation Research* **33B** (157-188).

Shmanske, Stephen (1991) "Price discrimination and congestion" *National Tax Journa* **44**, pp. 529-32

______ (1993) "A simulation of price-discrimination tolls" *Journal of Transport Economics and Policy* **27**, pp. 225-35

Small, Kenneth A. (1983) "Bus Priority and Congestion Pricing on Urban Expressways." In T. E. Keeler (ed.), *Research in Transportation Economics*, Vol. 1, (Greenwich, Connecticut: JAI Press), pp. 27-74.

Small, Kenneth A. and José A. Gómez-Ibanez (1998) "Road Pricing for Congestion Management: The Transition from Theory to Policy," in: *Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility*, ed. by K.J. Button and E.T. Verhoef. Cheltenham, UK: Edward Elgar, pp. 213-246.

Sullivan, E. (1998) *Evaluating the impacts of the SR 91 variable-toll express lane facility: Final report*, report to California Department of Transportation. Dept. of Civil and Environmental Engineering, Cal Poly State University, San Luis Obispo, California.

Toll Roads Newsletter, 40 (June 1999). Frederick, Maryland: Toll Roads Newsletter.

Transportation Research Board (1994) *Highway Capacity Manual. Special Report 209*, 3rd edition. Washington: National Research Council.

Verhoef, E.T., P. Nijkamp and P. Rietveld (1996) "Second-best congestion pricing: the case of an untolled alternative" *Journal of Urban Economics* **40** (3) 279-302.

Verhoef, Erik T., and Kenneth A. Small (1999) "Product Differentiation on Roads: Second-Best Congestion Pricing with Heterogeneity under Public and Private Ownership," Irvine Economics Paper 99-00-01, University of California at Irvine.

Viton, Philip A. (1995), "Private roads" *Journal of Urban Economics* **37**, pp. 260-89.

Wardrop, J.G. (1952) "Some Theoretical Aspects of Road Traffic Research" *Proceedings of the Institute of Civil Engineers*, **1**(II), 325-378.

Appendix.

A.1 The general form of the non-linear programming problem and the possible solutions.

We assume that at least some type 1 users use road A and at least some type 2 users use road B. We consider a congested traffic condition, so the toll charged under a policy regime is strictly greater than zero. The general form of the first-best (FB) problem in this paper can therefore be written as:

$$\max W = \int_0^{N_{1A}+N_{1B}} P_1(t)dt + \int_0^{N_{2A}+N_{2B}} P_2(t)dt - \sum_i \sum_r N_{ir}c_{ir}$$

$$s.t.\ h_1 \equiv P_1(N_{1A}+N_{1B}) - c_{1A}(N_{1A}+N_{2A}) - \tau_A = 0 \quad \text{(A.1a)}$$

$$h_2 \equiv P_2(N_{2A}+N_{2B}) - c_{2B}(N_{1B}+N_{2B}) - \tau_B = 0 \quad \text{(A.1b)}$$

$$h_3 \equiv N_{1B} \cdot (P_1 - c_{1B} - \tau_B) = 0 \quad \text{(A.1c)}$$

$$h_4 \equiv N_{2A} \cdot (P_2 - c_{2A} - \tau_A) = 0 \quad \text{(A.1d)}$$

$$g_1 \equiv P_1(N_{1A}+N_{1B}) - c_{1B}(N_{1B}+N_{2B}) - \tau_B \le 0 \quad \text{(A.1e)}$$

$$g_2 \equiv P_2(N \ + N_{2B}) - c_{2A}(N_{1A}+N_{2A}) - \tau_A \le 0 \quad \text{(A.1f)}$$

$$g_3 \equiv -\mathrm{N}_{1\mathrm{B}} \le 0 \quad \text{(A.1g)}$$

$$g_4 \equiv -N_{2A} \le 0 \quad \text{(A.1h)}$$

where $P(\cdot)$ and $c(\cdot)$ are the functions defined by (2) and (1). Certain constraints are added for the SB, TB, and PM policy, and the objective function is replaced by toll revenues in PM policy. Because we assume $N_{1A}, N_{2B} > 0$. (A.1a-b) are the same as (3) of the paper; (A.1c-d) are equivalent to (5c-d); (A.1e-f) to (5a-b); and (A.1g-h) to (5e).

Suppose $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are the Lagrangian multipliers for the first four equality constraint conditions, and $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ are those associated with the inequality constraints. According to the Kuhn-Tucker theorem, the necessary condition for the optimal solution $N^* = (N_{1A}^*, N_{1B}^*, N_{2A}^*, N_{2B}^*)$, $\lambda^* = (\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*)$, $\gamma^* = (\gamma_1^*, \gamma_2^*, \gamma_3^*, \gamma_4^*)$ are:

$$\nabla W(N^*) - \sum_{i=1}^{4} \lambda^*_i \nabla h_i(N^*) - \sum_{j=1}^{4} \gamma_j^* \nabla g_j(N^*) = 0 \quad \text{(A.2a)}$$

$$\gamma_j^* g_j(N^*) = 0,\ j = 1,2,3,4 \quad \text{(A.2b)}$$

$$\gamma_j^* \ge 0,\ j = 1,2,3,4 \quad \text{(A.2c)}$$

$$g_j \le 0,\ j = 1,2,3,4 \quad \text{(A.2d)}$$

If constraints (A.1e) and (A.1f) are binding at the same time, the tolls on both routes must be equal as shown in section 2. This is impossible for SB, TB and PM policy and our numerical results also show that this case is never optimal for FB policy. As a result, the possible solution cases for the programming problem are only three:

1. $\gamma_1^* = 0$, $\gamma_2^* > 0$ (SE1);

In this case, (A.2c)$\Rightarrow g_2 = 0$, i.e., (A.1f) must be binding. This means type 2 users are indifferent for two routes. Then (A.1e) cann't be binding, i.e., type 1 users strictly prefer road A and, (from (A.1c), $N_{1B}^* = 0$.

2. $\gamma_1^* > 0$, $\gamma_2^* = 0$ (SE2);

In this case, constraint (A.1e) is binding and constraint (A.1f) is not binding, and $N_{2A}^* = 0$.

3. $\gamma_1^* = 0$ and $\gamma_2^* = 0$;

In this case, we can only say (from the argument above) that (A.1e) or (A.1f) or both must be non-binding, therefore N_{1B}^* or N_{2A}^* or both must be zero. Considering the following three different solution cases:

3a. (A.1f) is binding and (A.1e) is not. N_{1B}^* is zero in this case (SE1).

3b. (A.1e) is binding and (A.1f) is not. N_{2A}^* is zero in this case (SE2).

3c. Both (A.1e) and (A.1f) are non-binding. N_{1B}^* and N_{2A}^* are both zero (SE).

In the paper, we divide the programming problem into different cases (SE, SE1, SE2) and solve each case under each policy. The above classification shows that the solutions from these cases include all of the possible solutions for the whole problem.

A.2 The derivation of optimal tolls of each equilibrium in each policy

In this section, we show how the general problem simplifies in each policy and equilibrium type (here described as "case"). In each case, we leave the non-negative constraints (A.1g-h) are implicit, as noted in the paper, we check each of them separately and impose it as an equality if required.

A. 2.1 FB Policy

Case SE. Substituting $N_{1B} = 0$ and $N_{2A} = 0$ into the welfare function, the welfare maximizing problem can be written as:

$$\max W = \int_0^{N_{1A}} P_1(t)dt + \int_0^{N_{2B}} P_2(t)dt - N_{1A} \cdot c_{1A}(N_{1A}) - N_{2B} \cdot c_{2B}(N_{2B})$$

The objective function is strictly concave because it equals the sum of four strictly concave functions. Therefore, the solution must be unique. The optimal traffic (N_{1A}^*, N_{2B}^*) in this case can be solved out from the first-order conditions. The corresponding tolls on the two routes are determined by (A.1a-b) and can be shown to be:

$$\tau_A = P_1 - c_{1A} = N_{1A} \cdot c'_{1A}(N_{1A}) \equiv MEC_{1A}$$
$$\tau_B = P_2 - c_{2B} = N_{2B} \cdot c'_{2B}(N_{2B}) \equiv MEC_{2B}$$

The optimal toll on each road is equal to the difference between social and private marginal cost on that road, known as "marginal external cost" MEC, just as in a single-route model.

Case SE1. Substituting $N_{1B} = 0$ into welfare function, we get:

$$\max W = \int_0^{N_{1A}} P_1(t)dt + \int_0^{N_{2A}+N_{2B}} P_2(t)dt - N_{1A} \cdot c_{1A}(N_{1A} + N_{2A}) - N_{2A} \cdot c_{2A}(N_{1A} + N_{2A}) - N_{2B}c_{2B}(N_{2B})$$

This objective function is also strictly concave because it equals the sum of five strictly concave functions. The corresponding tolls are:

$$\tau_A = P_1(N_{1A}) - c_{1A} = N_{1A}c'_{1A}(N_{1A} + N_{2A}) + N_{2A}c'_{2A}(N_{1A} + N_{2A}) \equiv MEC_A = P_2 - c_{2A}$$
$$\tau_B = P_2(N_{2A} + N_{2B}) - c_{2B}(N_{2B}) = N_{2B}c'_{2B}(N_{2B}) \equiv MEC_{2B}$$

The tolls are again the differences between social and private marginal costs on each route. The social cost on route A includes the users of both groups; the social cost on route B includes just the users of group 2. We also check the corner solution of $N_{2A} = 0$ in the simulation study.

Case SE2: Substituting $N_{2A} = 0$ into the welfare function, we get:

$$\max W = \int_0^{N_{1A}+N_{1B}} P_1(t)dt + \int_0^{N_{2B}} P_2(t)dt - N_{1A}c_{1A}(N_{1A}) - N_{1B}c_{1B}(N_{1B} + N_{2B}) - N_{2B}c_{2B}(N_{1B} + N_{2B})$$

Again, the objective function is strictly concave so the so the solution is unique. The tolls to decentralize the optimal traffic allocation in this case are:

$$\tau_A = P_1(N_{1A} + N_{1B}) - c_{1A} = N_{1A}c'_{1A}(N_{1A}) \equiv MEC_{1A}$$
$$\tau_B = P_1(N_{1A} + N_{1B}) - c_{1B} = N_{1B}c'_{1B}(N_{1B} + N_{2B}) + N_{2B}c'_{2B}(N_{1B} + N_{2B}) \equiv MEC_B = P_2 - c_{2B}$$

Here the social cost on route A includes just the users of group 1 and the social cost on route B includes the users of both groups. The corner solution of $N_{1B} = 0$ is also checked in the simulation study.

A. 2.2 SB and TB Policies

Case SE. The welfare maximizing problem under second-best pricing policy for fully separated equilibrium case can be written as:

$$\max W = \int_0^{N_{1A}} P_1(t)dt + \int_0^{N_{2B}} P_2(t)dt - N_{1A}c_{1A}(N_{1A}) - N_{2B}c_{2B}(N_{2B})$$
$$s.t.\ P_2(N_{2B}) = c_{2B}(N_{2B})$$

N_{2B} is determined solely by the constraint and numerical results in the paper show that there is only one positive real solution for N_{2B}. The objective function is a strictly function of N_{1A}, so if this case can occur, the solution is unique. The corresponding toll on route A is:

$$\tau_A = N_{1A}c'_{1A}(N_{1A}) \equiv MEC_{1A}$$

This toll is just the difference of social and private marginal cost on that road, the social cost including just the users of group 1. There are no route spill-overs in fully separated equilibrium: that is, road A is treated just as in the FB policy.

Case SE1. The corresponding Lagrangian is:

$$L = \int_0^{N_{1A}} P_1(t)dt + \int_0^{N_{2A}+N_{2B}} P_2(t)dt - N_{1A}c_{1A}(N_{1A}+N_{2A}) - N_{2A}c_{2A}(N_{1A}+N_{2A}) - N_{2B}c_{2B}(N_{2B})$$
$$-\lambda_1[P_1(N_{1A}) - c_{1A}(N_{1A}+N_{2A}) - P_2 + c_{2A}(N_{1A}+N_{2A})]$$
$$-\lambda_2[P_2(N_{2A}+N_{2B}) - c_{2B}(N_{2B})]$$

where the constraints (A.1a-b) have been rewritten using (A.1f) as an equality in order to eliminate τ_A as a variable. The Lagrangian Multiplier λ_1 represents the "Shadow Price" of not price discriminated on road A, that is, it represents the increase of social welfare that could be achieved by charging type-1 users more than type-2 users, since the latter have a sub-optimally priced substitute (road B).This problem can be solved for N_{1A}, N_{2A}, N_{2B} and λ_1, λ_2. The toll which decentralizes the solution allocation is then determined by (A.1a) as:

$$\tau_A = N_{1A}c'_{1A} + N_{2A}c'_{2A} - \left[\frac{P'_2 N_{2B}c'_{2B} \cdot (P'_1 - c'_{1A} + c'_{2A})}{P'_1 P'_2 - P'_1 c'_{1B} - P'_2 c'_{2B}}\right]$$

The toll on route A equals to marginal external cost plus an adjustment term which depends on the slope of demand function and cost function.

Case SE2. The Lagrangian is:

$$L = W - \lambda_2[P_2(N_{2B}) - c_{2B}(N_{1B}+N_{2B})] - \gamma_1[P_1(N_{1A}+N_{1B}) - c_{1B}(N_{1B}+N_{2B})]$$

where (A.1e) has been used as an equality with Larangian multiplier γ_1 which represents the "shadow price" of not being able to price discriminated on road B.

Again, we solve and use (A.1a) to determine the toll on route A as:

$$\tau_A = N_{1A}c'_{1A} + \left[\frac{(-N_{1B}c'_{1B} - N_{2B}c'_{2B})P'_2P'_1}{P'_1P'_2 - P'_1c'_{2B} - P'_2c'_{1B}}\right]$$

The toll here equals to the marginal congestion cost plus a adjustment term which depends on the slope of demand function as well as costfunction.

It is difficult to judge analytically whether the solution is unique in case SE1 and SE2 of SB policy because of the non-linear form of the constraints. In the simulation study, we use different initial values to show that in these cases no more than one equilibrium solution can be found.

The TB policy is the same as the SB policy except that we add an extra constraint (6), which we check separately rather than including in the Lagrangian.

A. 2.3 PM Policy

The maximizing problem here has the same constraints as the ones in the SB policy. The only different is that the objective function now is:

$$R = (N_{1A})[P_1(N_{1A}) - c_{1A}(N_{1A} + N_{2A})] + N_{2A}[P_2(N_{2A} + N_{2B}) - c_{2a}(N_{1A} + N_{2A})]$$

Case SE. The solution of this case must be unique because the same reason as SE case in SB policy. The toll which maximizes revenue is found to be:

$$\tau_A = N_{1A}[c'_{1A}(N_{1A}) - P'_1]$$

The toll is set at marginal social cost plus a monopolistic mark-up which is inversely related to the demand elasticity of group 1. Equivalently, this equation can be written as $\tau_A + N_{1A}P'_1 = N_{1A}c'_{1A}$, that is, marginal revenue equals marginal cost.

Case SE1. The toll is found to be:

$$\tau_A = N_{1A}c'_{1A} + N_{2A}c'_{2A} - N_{1A}P'_1 + \left[\frac{(N_{2A}P'_2c'_{2B} + N_{1A}P'_1P'_2 - N_{1A}P'_1c'_{2B})(P'_1 - c'_{1A} + c'_{2A})}{P'_1P'_2 - P'_1c'_{2B} - 2(P'_2)^2 + P'_2c'_{2B}}\right]$$

Again the toll equals marginal congestion cost plus a monopolistic mark-up.

Case SE2. The revenue-maximizing toll on route A is:

$$\tau_A = N_{1A}c'_{1A} - N_{1A}P'_1 + \left[\frac{N_{1A}(P'_1)^2 (P'_2 - c'_{2B})}{P'_1(P'_2 - c'_{2B}) - c'_{1B}P'_2} \right]$$

Again, the uniqueness of equilibrium solution for case SE1 and SE2 is proved numerically.