Economics Bulletin

Volume 35, Issue 4

A two-route Ramsey pricing problem: second-best congestion pricing with an untolled alternative and a revenue constraint

Seiji Steimetz
California State University, Long Beach

Abstract

This paper derives a second-best congestion toll for when two congestible facilities are perfect substitutes, but only one facility can be priced, and where a minimum amount of toll revenue must be generated. It is shown that the resulting, second-best toll equals a weighted average of the social planner's optimal toll and the profit-maximizing toll of a private monopolist, with the shadow price of the revenue constraint serving as the weight. As the revenue constraint is relaxed, the social planner's toll becomes more efficient. But when the revenue constraint becomes more difficult to satisfy, monopoly pricing becomes more efficient.

I thank Kenneth Small, Luis Rizzi, and Kangoh Lee for their helpful comments and suggestions. This research was supported, in part, by a Research and Scholarly and Creative Activities Award from the California State University, Long Beach.
Citation: Seiji Steimetz, (2015) "A two-route Ramsey pricing problem: second-best congestion pricing with an untolled alternative and a revenue constraint", *Economics Bulletin*, Volume 35, Issue 4, pages 2535-2542
Contact: Seiji Steimetz - seiji.steimetz@csulb.edu.
Submitted: September 14, 2015. **Published:** December 13, 2015.

1. Introduction

Consider a vehicle that enters a highway during a congested commute period. The marginal social cost of that vehicle's trip is the sum of a marginal private cost and a marginal external cost. The private cost includes the vehicle's operating cost and the value of its passengers' time. The external cost is the further delay it imposes on the highway's existing travelers, often referred to as a "congestion externality". Because each entering vehicle does not bear the congestion externality it generates, the traffic volume on the highway will be inefficiently high. A well-established result in the congestion pricing literature is that a road toll equal to the external congestion cost will reduce traffic to an optimal level (Walters 1961). That toll represents a "first-best" toll, which can be interpreted as a Pigouvian tax that bridges the gap between the marginal social and marginal private costs of road use (Small and Verhoef 2007).

In practice, first-best tolls are often infeasible for political or practical reasons. This gives rise to the analysis of "second-best" pricing schemes, where welfare is maximized subject to one or more constraints. One such constraint is when only a portion of a highway can be tolled. For example, several highways in the United States include "express lanes" where motorists can pay for less-congested travel. The highway's other lanes must remain untolled. In the congestion pricing literature, this presents a "two-route pricing problem" where the second-best toll on the express lanes must be determined with the understanding that traffic diverted from those lanes will worsen congestion on the untolled lanes (Verhoef et al. 1996).

Another constraint is when toll revenues must reach some minimal level to help finance public projects. Analyzing those circumstances is where transportation economics overlaps with public economics, because the resulting second-best toll can be interpreted as a form of Ramsey pricing (Sandmo 1975, Oum and Tretheway 1988).

A real-world scenario that has not been analyzed until now is when the above constraints must be simultaneously met. The most prevalent examples worldwide are the express-lane facilities mentioned above, tolled highways that serve the same destinations as untolled highways, and "cordon pricing" schemes where tolls are required along certain routes into urban centers. In each case, tolled and untolled alternatives are available, and some minimum amount of toll revenue must be generated to help finance public projects (Dixit et al. 2010). A more recent example comes from the Los Angeles and Long Beach maritime ports, where trucks entering the ports during peak operating periods must pay a "traffic mitigation fee", but can avoid the fee by entering during off-peak hours. A minimum amount of fee revenue must be generated to finance off-peak port operations and other obligations attached to the program (Steimetz et al. 2008).

Each of those scenarios presents a two-route congestion pricing problem with a Ramsey-type revenue constraint, which has not been previously analyzed. The purpose of this paper is to derive the optimal congestion toll under such circumstances, to reveal its implications, and to demonstrate how it links more familiar results from the literatures on congestion pricing and Ramsey pricing. A particularly compelling finding is that the optimal toll turns out to be a weighted-average of the social planner's (second-best) welfare-maximizing toll and the private monopolist's profit-maximizing toll, where the shadow price of the revenue constraint serves as the weight. A practical implication is that monopoly pricing strategies become more efficient as the revenue requirement becomes increasingly difficult to satisfy. On the other hand, the social planner's pricing strategy becomes more efficient as the revenue constraint is relaxed.

2. Review of Previous Findings

2.1 The Social Planner's Two-Route Pricing Problem

Consider two parallel roads that are perfect substitutes for travelling from an origin to a destination, but tolls may only be charged on one of them. Road users are homogeneous except for differences in their willingness to pay for trips on either road. Let V_T and V_U denote the traffic volumes on the tolled and untolled route, respectively, with $c_T(V_T)$ and $c_U(V_U)$ representing the average user cost on each route. Those average costs are typically interpreted as marginal private costs, with $c_T' > 0$ and $c_U' > 0$ indicating that both routes are congestible. Recalling that the roads are pure substitutes, the overall inverse demand function for the pair of roads is $d(V_T + V_U)$, where $\frac{\partial d}{\partial V_T} = \frac{\partial d}{\partial V_U} \equiv d' < 0$. In equilibrium, the marginal benefit of using either road must equal its "generalized price", P_j, such that $d = P_T = c_T + \tau$ on the tolled route, with τ denoting the toll amount, and $d = P_U = c_U$ on the untolled route.

Following Verhoef et al. (1996), the social planner's second-best toll is derived from maximizing the Lagrangian function:[1]

$$L = \int_0^{V_T+V_U} d(v)dv - c_T(V_T) \cdot V_T - c_U(V_U) \cdot V_U$$

$$+\gamma_T \cdot [c_T(V_T) + \tau - d(V_T + V_U)] + \gamma_U \cdot [c_U(V_U) - d(V_T + V_U)] \quad (1)$$

where γ_T and γ_U are shadow prices for the equilibrium constraints; they measure the marginal welfare gain that would result from a marginal increase in the ability to charge a toll on either road. The second-best outcome yields $\gamma_t = 0$ because the toll on road T is set optimally, and $\gamma_U > 0$ because welfare could increase if a toll were allowed to manage congestion on road U.

Solving the first-order conditions from maximizing (1), Verhoef et al. (1996) show that the second-best toll is:

$$\tau = c_T' V_T - c_U' V_U \frac{-d'}{c_U' - d'} \quad (2)$$

To interpret (2), note in general that the marginal social cost of entering road j is $MSC_j = c_j + c_j' V_j$, which is the sum of the marginal private cost (c_j) and the marginal external congestion cost ($c_j' V_j$). The latter expression shows that the congestion externality is the increased travel cost

---

[1] Throughout this paper, only short-run welfare-maximization problems are analyzed for ease of expression (mainly to avoid cumbersome notation for partial derivatives). Long-run analyses include road capacity costs, whether publically or privately incurred, but the resulting optimal tolls are identical to those of their corresponding short-run analyses. Moreover, only "static congestion" during a single commute period is considered, where it is assumed that tolls do not influence motorists' departure times.

imposed on each user, multiplied by the number of affected users. If tolls could be charged on both roads, then the first-best toll on each road would be $\tau_j = c_j' V_j$. We see from (2), however, that the toll on road T is lower than that first-best toll because traffic diverted from the tolled route spills over onto the untolled route. This is more easily seen if overall demand were perfectly-inelastic such that $d' \to \infty$. In that case, the resulting second-best toll would be

$$\tau = c_T' V_T - c_U' V_U \tag{3}$$

which is the external congestion cost on the tolled route, less the external congestion cost imposed on the untolled route.

2.2 The Private Monopolist's Two-Route Pricing Problem

Suppose that the above, two-route problem were instead faced by a private monopolist operating both routes. A real-world example of that scenario would be the express-lane facility on State Route 91 in Southern California, which was operated by a private franchisee until 2003. The objective then becomes one of finding the short-run, profit-maximizing toll using the Lagrangian function:[2]

$$L = \tau \cdot V_T + \gamma_T \cdot [c_T(V_T) + \tau - d(V_T + V_U)] + \gamma_U \cdot [c_U(V_U) - d(V_T + V_U)] \tag{4}$$

Small and Verhoef (2007) demonstrate that this yields the profit-maximizing toll:

$$\tau = c_T' V_T - V_T d' \frac{c_U'}{c_U' - d'} \tag{5}$$

Comparing (5) to (2), and recalling that $d' < 0$, we see that the monopolist's toll is greater than the social planner's second-best toll, and also greater than the first-best toll ($\tau = c_T' V_T$). The markup over the first-best toll depends on users' price-sensitivity (reflected in d') and the degree of congestion on the untolled route (reflected in c_U'), along with the equilibrium traffic volume on the tolled route.

2.3 The Social Planner's Single-Route Ramsey Pricing Problem

Now consider only a single route, which can be tolled, but must generate short-run toll revenues of at least R to help finance public projects. This represents a specific case of the more general pricing problems analyzed in the Ramsey pricing literatures without externalities (Ramsey 1927, Baumol and Bradford 1970) and with externalities (Sandmo 1975, Oum and Tretheway 1988). Given the revenue constraint, the second-best toll is derived from the Lagrangian:

---

[2] In the long run, this monopolist faces a capacity-cost constraint, which can be viewed as a minimum-revenue constraint. That constraint is omitted from (4), however, to focus on short-run welfare analyses and to avoid confusion when analyzing the effects of changing the shadow price of a revenue constraint. Note that including a capacity-cost constraint would not alter any of the results presented herein.

$$L = \int_0^V d(v)dv - c(V) \cdot V + \lambda \cdot \{[V \cdot [d(V) - c(V)] - R\} \quad (6)$$

noting that the equilibrium constraint, $d = c + \tau$, has been directly incorporated into the objective function. Small and Verhoef (2007) show that the optimal toll in this case is:

$$\tau = c'V - \frac{\lambda}{1+\lambda} d'V \quad (7)$$

where λ is the shadow price of the revenue constraint, and the familiar Ramsey number, $\frac{\lambda}{1+\lambda}$, appears in the markup over the first-best toll. To more clearly see how (7) is related to the general Ramsey-pricing rule, recall that $MSC = c + c'V$ and that P is the generalized price of road use, where in equilibrium $P = d = c_T + \tau$. The well-known Ramsey rule (Baumol and Bradford 1970) emerges when expressed as a markup of the generalized price (including the toll) over marginal social cost:

$$\frac{P - MSC}{P} = \frac{\lambda}{1+\lambda} \cdot \frac{1}{E} \quad (8)$$

where $E \equiv -\frac{1}{d'}\frac{P}{V}$ is the absolute value of the road's demand elasticity. When the revenue constraint is non-binding, the markup over marginal social cost is zero because the first-best toll applies. And when the revenue constraint becomes increasingly-difficult to satisfy, such that $\lambda \rightarrow \infty$, then the private monopolist's profit-maximizing toll is invoked. (In Section 4, the interpretation of λ is discussed in greater detail.)

3. A Two-Route Ramsey Pricing Problem

This paper's main contribution can now be presented. Consider the two parallel routes discussed previously, where only one can be tolled, but now with the additional constraint that a minimum revenue of R must be generated to help finance public projects. An increasingly prominent example is when only some of a highway's lanes are designated as tolled express lanes, and toll revenues are earmarked for public use (beyond covering the highway's capacity costs).

From an analytic standpoint, this amounts to adding a revenue constraint to (1). At first glance, the addition of that constraint may seem subtle. However, since the seminal work of Verhoef et al. (1996), considerable research has been devoted to understanding how optimal taxes and tolls are influenced by such revenue requirements — see, for example, Mayeres and Proost (2001), Parry and Bento (2001), and Proost et al. (2007). Yet, until now, those influences have not been analyzed in the context of a two-route pricing problem. This has left an important gap in the literature because relevant scenarios are becoming increasingly prevalent in the real world, such as the tolled express lanes on several urban highways in the United States, and cordon-pricing schemes in the United Kingdom, Sweden, and Singapore to name a few.

For example, suppose that the revenue from a single toll road is intended to help reduce sales taxes. This will increase the (second-best) optimal toll if the distortion caused by increasing the toll is less than the distortion mitigated by lowering sales taxes. Hence, the revenue requirement affects the optimal toll level. A salient question is: does an analogous result hold for a two-route pricing problem?

To answer that question, the second-best toll from the social planner's perspective is derived by maximizing the Lagrangian function:

$$L = \int_0^{V_T+V_U} d(v)dv - c_T(V_T)\cdot V_T - c_U(V_U)\cdot V_U$$
$$+\gamma_T\cdot[c_T(V_T)+\tau-d(V_T+V_U)]+\gamma_U\cdot[c_U(V_U)-d(V_T+V_U)]$$
$$+\lambda\cdot[\tau\cdot V_T-R] \qquad (9)$$

which mimics (1) but adds the toll-revenue constraint and its associated shadow price, λ. The addition of that revenue constraint is what links the analyses of two-route pricing problems and Ramsey pricing problems.

The first-order conditions with respect to V_T, V_U, and τ are:

$$\frac{\partial L}{\partial V_T} = d - c_T - c_T'V_T + \gamma_T\cdot(c_T'-d') - \gamma_U d' + \lambda\tau = 0$$
$$\frac{\partial L}{\partial V_U} = d - c_U - c_U'V_U - \gamma_T d' + \gamma_U\cdot(c_U'-d') = 0$$
$$\frac{\partial L}{\partial \tau} = \gamma_T + \lambda V_T = 0 \qquad (10)$$

while the first-order conditions for γ_T, γ_U, and λ return the equilibrium constraints on the tolled and untolled routes, along with the revenue constraint.

Note that without the revenue constraint, γ_T would equal zero because the toll could be set optimally (in the second-best sense). From (10), however, we see that $\gamma_T = -\lambda V_T < 0$, indicating that welfare improves with each marginal reduction in the revenue constraint, and $\gamma_T = 0$ only when the revenue constraint is non-binding.

To solve for the optimal toll, begin with the first-order conditions for V_U and τ, along with the equilibrium constraints to obtain:

$$\gamma_U = \frac{c_U'V_U}{c_U'-d'} - \lambda\frac{d'V_T}{c_U'-d'} > 0 \qquad (11)$$

Equation (11) demonstrates that welfare could be improved by allowing tolls to manage congestion on the untolled route, and would be further improved by marginal reductions in the revenue constraint. In fact, Small and Verhoef (2007) show from (1) that $\gamma_U = \frac{c_U'V_U}{c_U'-d'}$. Comparing that result to (11), we see that the welfare loss from the inability to toll route U is exacerbated by the revenue constraint.

Applying (11) to the first-order condition for V_T and solving for τ yields a congestion pricing variant of Ramsey pricing:

$$\tau = \left[\frac{1}{1+\lambda}\right]\left[c_T' V_T - c_U' V_U \frac{-d'}{c_U'-d'}\right] + \left[\frac{\lambda}{1+\lambda}\right]\left[c_T' V_T - V_T d' \frac{c_U'}{c_U'-d'}\right] \tag{12}$$

A striking finding arises when comparing (12) with (2) and (5). We see that the second-best toll is a weighted average of the social planner's optimal two-route pricing scheme and the private monopolist's profit-maximizing scheme. The weight is the shadow price of the revenue constraint, which is related to the social cost of public funds (in a general-equilibrium sense) and is discussed below in further detail. We also see the familiar Ramsey number, $\frac{\lambda}{1+\lambda}$, appearing in the second term on the right-hand-side of (12).

4. Implications

To appreciate the implications of (12), it is important to understand the role played by the revenue constraint's shadow price, λ, which measures the marginal welfare gain from each dollar of toll revenue ($\tau \cdot V_T$) raised. Equivalently, it measures the welfare improvement from reducing the revenue requirement (R) by one dollar. The level of the revenue requirement is dictated by the need for public funds, which might otherwise be raised through taxation on goods or labor. The shadow price is therefore positive ($\lambda > 0$) if raising public funds from toll revenues is less distortionary than raising them from other sources such as a sales tax.[3] If the revenue constraint is non-binding ($\lambda = 0$), it means that there is no welfare gain from generating additional toll revenues (aside from the corrective effects of the toll) because revenues from other sources are no more distortionary.

Accordingly, the optimal toll in (12) depends on the extent to which the revenue constraint is binding. The more distortionary it is to generate revenues from other sources, the more it makes sense to raise revenue from tolls. In the limiting case where $\lambda \to \infty$, we see from (12) that the private monopolist's profit-maximizing toll in (5) is optimal. In that case, the social planner behaves like a monopolist to generate the maximum, feasible amount of toll revenue. On the other hand, if the revenue constraint is non-binding ($\lambda = 0$), then naturally the social planner's second-best toll in (2) is more efficient. More generally, when the burden of financing public projects with toll revenues increases, it reduces the welfare gains from congestion pricing by pushing the toll toward a monopoly price.

To better see how (12) fits within the more traditional Ramsey pricing literature, suppose that only the tolled road were to exist. In that case, $V_U \to 0$ because traffic volume on the untolled

[3] The quantity $1 + \lambda$ can be interpreted as the "marginal cost of public funds", which is the marginal social cost of raising each dollar through taxation (see, for example, Small and Verhoef 2007). Under that interpretation, $\lambda > 0$ means that the marginal cost of public funds exceeds one, implying a relatively distortionary tax system. In contrast, $\lambda = 0$ means that the marginal cost of public funds equals one, implying that the revenues generated by the toll road and the tax system carry the same welfare effects.

road vanishes, and $c_U' \to \infty$ because congestion increases without limit when adding a vehicle to a road with infinitesimal capacity. The second-best toll then reduces to

$$\tau = \frac{1}{1+\lambda} c'V + \frac{\lambda}{1+\lambda}(c'V - d'V) = c'V - \frac{\lambda}{1+\lambda} d'V \tag{13}$$

which is a weighted average of the first-best toll and the monopolist's profit-maximizing toll, with the shadow price of the revenue constraint serving as the weight, and is equivalent to the single-route Ramsey price in (7). Furthermore, applying the conditions $P = d = c + \tau$ and $MSC = c + c'V$ to (13) yields

$$\frac{P - MSC}{P} = \frac{\lambda}{1+\lambda} \cdot \frac{1}{E} \tag{14}$$

as in (8), where E is the absolute value of the road's demand elasticity.

Thus, we see that the two-route Ramsey pricing problem is essentially a generalization of the single-route problem. And its implications are consistent with those of more traditional Ramsey pricing analyses, in the context of a marginal social cost that includes an externality (Oum and Tretheway 1988). If the revenues from corrective pricing mechanisms are intended to replace more distortionary sources of public financing, then those prices depart from their optimal levels and move closer to monopoly levels.

5. Concluding Remarks

This paper derives a second-best, congestion pricing strategy for when two congestible facilities are perfect substitutes, but where only one facility can be priced, and where a minimum amount of toll revenue must be generated. In doing so, it clarifies how such revenue constraints induce departures from optimal pricing schemes toward monopoly pricing schemes. It also clarifies how related findings in the congestion pricing and Ramsey pricing literatures are linked.

References

Baumol, W.J. and D.F. Bradford (1970) "Optimal departures from marginal cost pricing" *American Economic Review* **60**, 265–83.

Dixit, V.V., R.C. Harb, G.W. Harrison, D.M. Marco, M.S. Mard, A.E. Radwan, E.H.H. Rutstrom, and M.P. Schnieder (2010) "Review of congestion pricing experiences" Report to the United States Federal Highway Administration.

Mayeres, I. and S. Proost (2001) "Marginal tax reform, externalities and income distribution" *Journal of Public Economics* **79**, 343–63.

Oum, T.H., and M.W. Tretheway (1988) “Ramsey pricing in the presence of externality costs” *Journal of Transport Economics and Policy* **22**, 307–17.

Parry, I.W.H. and A.M. Bento (2001) “Revenue recycling and the welfare effects of congestion pricing” *Scandinavian Journal of Economics* **103**, 645–71.

Proost, S., B. De Borger, and P. Koskenoja (2007) “Public finance aspects of transport financing and investment” in *Investment and the Use of Tax and Toll Revenues in the Transport Sector* by A. De Palma, R. Lindsey, and S. Proost, Eds., Elsevier JAI: Amsterdam.

Ramsey, F. P. (1927) “A contribution to the theory of taxation” *Economic Journal* **37**, 47–61.

Sandmo, A. (1975) “Optimal taxation in the presence of externalities” *Swedish Journal of Economics* **77**, 86–98.

Small, K.A. and E.T. Verhoef (2007) *The Economics of Urban Transportation*, Routledge: New York.

Steimetz, S.S.C., O. Chan, A. Ng, and S. Fergus (2008) “Evaluating the efficiency of traffic mitigation fees at the San Pedro Bay ports” Report to the METRANS Transportation Center, University of Southern California and California State University, Long Beach.

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

Walters, A.A. (1961) “The theory and measurement of private and social cost of highway congestion” *Econometrica* **29**, 676–699.