Economics of Transportation 3 (2014) 257–269

Contents lists available at ScienceDirect

Economics of Transportation

journal homepage: www.elsevier.com/locate/ecotra

Second-best urban tolling with distributive concerns

A. Diaz [a,b,*], S. Proost [b,c]

[a] IFP Energies Nouvelles, 1 et 4 Avenue de Bois-Preau, 92852 Rueil Malmaison Cedex, France
[b] Center of Economics Studies, KULeuven, Naamsestraat 69, 3000 Leuven, Belgium
[c] Centre for Transport Studies, KTH, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 11 January 2015
Received in revised form
25 February 2015
Accepted 25 February 2015

Keywords:
Tax reform
Congestion pricing
Urban tolls

ABSTRACT

This paper analyzes the optimal urban congestion toll in a second-best setting where only one road in a network can be tolled. Both heterogeneity in labor productivity and income distribution concerns are considered. The optimal toll balances two types of considerations. The first consideration is the correction of the congestion externality on the tolled road given the distortion on the non-tolled roads, while the second is the equity consideration that takes into account which income group uses the tolled road and how toll revenues are spent. Both separating and pooling equilibria are analyzed for two alternative uses of toll revenues: poll transfers and labor-tax cuts. Using numerical simulations, we show that equity concerns can lead a government to prefer inefficient toll levels and recycling via poll transfers rather than via labor tax reductions.

1. Introduction

Transport economists advocate road pricing as an efficient instrument to regulate the use of road infrastructures. Imposing a road toll that reflects marginal external congestion costs makes consumers use the road up to the point where marginal social costs equalize marginal social benefits. Optimal road pricing therefore ensures that the only trips made are those that bring the highest benefits to society. This is only true, however, as long as tolling is analyzed in a first-best framework. Additional conditions, e.g. not being able to toll all roads in a network, pre-existing distortions on the labor market, or equity concerns complicate the optimal design of urban congestion tolls.

The related literature is mainly focused on the interaction of road taxes with taxes on labor income (see Mayeres and Proost, 1997; Parry and Oates, 2000; Parry and Bento, 2001; Van Dender, 2003; Parry and Small, 2005; De Borger, 2009). The issue can be summarized as follows. Road taxes have a positive welfare impact by reducing congestion externalities. At the same time, however, they have a negative impact since an increase in commuting costs discourages labor supply. Which effect (externality reduction or reduced labor supply) prevails has become a central question in transport economics. Parry and Bento (2001) showed that the welfare impact of a road tax differs according to the use of the tax revenues. Using road tax revenues to reduce taxes on labor increases social welfare because reduced congestion and reduced labor taxes compensate workers for the congestion toll. Other revenue uses, such as poll transfers, do not compensate the negative labor supply impact and reduce welfare. On the other hand, Mayeres and Proost (1997, 2001) demonstrated that as long as equity objectives are relevant, obtaining significant welfare gains from recycling tax revenues requires a careful balance of several options. They show that imposing a tax on congestion externalities may need a reconfiguration of all taxes, and that a reduction of labor taxes is not necessarily the best option.[1]

This paper contributes to this line of research by analyzing the importance of revenue allocation when heterogeneous drivers use a congested network. We wonder whether taking into account differences across road users and redistribution objectives for transport policy can change the welfare effect implied by the recycling scheme.

Our approach is close to that of Parry and Bento (2001) but we add two dimensions to their model. First, instead of a choice between a congested road and uncongested public transit, we model two congested transport options. They can be both roads or

* Corresponding author at: Center of Economics Studies, KULeuven, Naamsestraat 69, 3000 Leuven, Belgium.
E-mail addresses: Andrea.DiazRincon@kuleuven.be (A. Diaz), Stef.Proost@econ.kuleuven.be (S. Proost).

[1] Proost and Van Regemorter (1995) apply this idea to a macro-economic disequilibrium framework.

http://dx.doi.org/10.1016/j.ecotra.2015.02.003
2212-0122/

one of them can be public transit. Allowing congestion on the untolled alternative is particularly interesting because it implies that the toll not only brings efficiency gains in the transport market but also efficiency losses in the form of increased congestion in the rest of the network (see e.g. Rouwendal and Verhoef, 2004). Second, Parry and Bento consider homogeneous consumers without paying attention to income distribution issues. However, we know that at the origin of labor taxes there is often the income distribution objective. With this in mind, we model labor-force heterogeneity in the form of differences in labor productivity between two types of workers (low- and high-income). Differences in productivity imply differences in values of time. This in turn determines the sorting of commuters over the tolled and the untolled route (see e.g. Small and Yan, 2001; De Palma and Lindsey, 2004). Tolling the faster route will tend to attract the most productive commuters. Therefore, the tax can be imposed on high-income consumers and can be used either to redistribute resources to low-income consumers or to obtain additional gains by lowering labor taxes for all commuters.

Our analysis shows that the optimal toll differs from the Pigouvian tax. The toll can be lower or higher than the marginal external cost on the tolled road. The magnitude of the deviation depends on several aspects: the equity concerns, who uses the tolled road, who benefits from redistribution and how easily consumers switch to other alternatives. A numerical exercise provides two significant insights. First, when accounting for heterogeneity, tolling off those that are least able to pay for the toll can be welfare improving, on the condition that the revenue recycling scheme benefits them. Consequently, if income distribution concerns seek to favor the least productive workers, the policymaker would prefer to recycle toll revenues through poll transfers. Second, assumptions about the relationship between the tolling policy and congestion in the rest of the network determine the effects of the recycling scheme on labor supply. Neglecting congestion on alternative routes may result in an overestimation of benefits from the tolling policy.

This paper is organized as follows. In Section 2, we develop an analytical model and analyze the problem with homogeneous households. In Section 3, we introduce heterogeneity in labor productivity and define four different equilibria of road use. In Section 4, we analyze the social planner's problem and derive the optimal tolling rules for the different equilibria and for two ways of recycling the toll revenues: poll transfers and labor tax cuts. In Section 5, we present a numerical illustration. In the last section we conclude.

2. The household's problem: road choice

We start with a simple model—in the spirit of Parry and Bento (2001)—of a representative household whose utility function depends on aggregate consumption of market goods (X, whose price is normalized to one), leisure time (t_L), and the disutility of commuting[2] ($\Psi(\cdot)$). The household owns a car and uses it to commute to work by either one of the two parallel congested roads (routes U and T) that connect residential areas to workplaces. Total number of worked days in a period (D) is the sum of the number of days the household commutes by road U and by road T (D_U and D_T).

$$\mathcal{U}(X, t_L, D_U, D_T) = U(X, t_L) + \Psi(D_U, D_T). \tag{1}$$

A congestion toll (τ) related to distance (d) is applied on one of the two roads (route T), while the other (route U) remains untolled. Households choose which route to use to commute to work, U or T. The household faces the following budget constraint:

$$X + gd_U D_U + (g+\tau)\, d_T D_T \le \varepsilon w(1-\tau_w)(D_U + D_T) + G. \tag{2}$$

The right-hand side of (2) corresponds to total household's income composed of work income and a head subsidy (G). Work income in a period is the product of the daily net wage and the number of days worked in the period, where ε is labor productivity per day, w is the gross daily wage per unit produced and τ_w is a tax levied on labor income. We assume that households are homogeneous in all respects except that they exhibit different exogenous levels of labor productivity. Thus, for the same level of labor supplied, high-productivity households get a higher income than low-productivity households.

The left-hand side of (2) corresponds to household expenditures on aggregate consumption and commuting. Each day of work requires a commuting round trip that involves time and monetary costs. When commuting by the untolled road, only fuel costs are relevant.[3] g represents fuel price per kilometer, $g = c_g(1+\tau_g)$, where c_g is the resource fuel cost (which takes into account vehicle fuel efficiency) and τ_g the fuel tax. Commuting by the tolled road implies paying for the fuel consumption plus the toll. However, this road allows faster trips, while the untolled road requires more time and higher fuel consumption due to a longer distance: $d_U = \beta d_T$ with $\beta > 1$.

Households also face a time constraint:

$$\bar{t} = D_U + D_T + t_U d_U D_U + t_T d_T D_T + t_L. \tag{3}$$

The household's time endowment during a period ($\bar{t}$) equates the sum of labor supplied, commuting time and leisure time. t_U and t_T are two different functions of time per unit of distance (e.g. the inverse of the speed—h/km). Households choose how many days to work in a period (hours of work per day are fixed[4]), and how to commute to work. By choosing the optimal number of workdays (D_T and D_U) in a period, households indirectly set total income and total leisure time during the period.

The first-order conditions of maximizing utility (1) subject to (2) and (3) are (see Appendix B for detailed derivations)

$$\varepsilon w(1-\tau_w) - \frac{U_{t_L}}{U_X} = gd_U + t_U d_U \frac{U_{t_L}}{U_X} - \frac{\Psi_{D_U}}{U_X}, \tag{4}$$

$$\varepsilon w(1-\tau_w) - \frac{U_{t_L}}{U_X} = (g+\tau)\, d_T + t_T d_T \frac{U_{t_L}}{U_X} - \frac{\Psi_{D_T}}{U_X}. \tag{5}$$

These expressions equate the private benefit from an extra day of work (daily net wage minus the value of daily leisure time foregone by working) with the generalized private cost of commuting (monetary and time costs). The monetary cost of transport consists of the fuel consumption charge in the case of commuting by the untolled road (4), whereas it consists of the fuel consumption plus the toll when commuting by the tolled road (5).

As a result of considering time as a resource, we get the monetary value of time for each household (U_{t_L}/U_X). This is the ratio between the Lagrange multiplier of the time constraint and the Lagrange multiplier of the income constraint (see Appendix B). The value of spending time in transport[5] (value of transport time, VTT) is represented in (4) and (5) by the value of time foregone by

[2] The separability of the utility function implies that the amount of labor supplied is independent of the road choice.

[3] We consider that costs such as maintenance, insurance, vehicle ownership taxes, etc., are constant, since they do not vary with the level of congestion.

[4] This is a typical assumption in the related literature (see e.g. Parry and Bento, 2001; Van Dender, 2003). However, it can be argued that hours per day can also be chosen. By using a labor supply model that allows for optimal choice of both daily work hours and number of workdays, Gutierrez-i-Puigarnau and van Ommeran (2009) show that commuting costs can increase daily hours worked. However, they find that the effect on total labor supply is ambiguous.

[5] For a detailed explanation of travel time valuation, see Small and Verhoef (2007) and Jara-Diaz (2000).