Pricing and Provision of Transport Infrastructure with Nonlinear Income Taxation

ANTONIO RUSSO†

Abstract

We study optimal pricing and provision of transportation infrastructure (roads and public transport) in presence of nonlinear income taxes. Individuals are heterogeneous in unobservable earning ability. Each working day requires a commuting trip and daily work hours have diminishing returns. We study both linear and nonlinear transport tariffs. In spite of individuals having separable preferences for goods and leisure, pricing and provision of infrastructure can improve screening of types. Optimal per-trip tariffs depend on the time cost of travel and the effect on labor supply of changes in the amount of working days. Reducing the time cost of journeys facilitates screening. Therefore, redistribution provides an additional motive to raise per-trip car tariffs (thus curbing road congestion) and upgrade infrastructure. We also provide some insights on the usefulness of means-testing for transport tariffs.

JEL Code: R41, H23, H21, H54

Keywords: road pricing; public transport pricing; infrastructure provision; income taxation

*I thank Georges Casamatta, Helmuth Cremer, Bruno De Borger, Jean-Marie Lozachmeur and Emmanuel Thibault for useful comments and suggestions. I also thank audiences in Marseille (LAGV 2011), Montpellier (DMM 2011), Oslo (EEA 2011), Stockholm (ITEA meeting 2011) and Toulouse. All errors are mine.

†Toulouse School of Economics. Email: antonio.russo@tse-fr.eu
1 Introduction

It is often argued that pricing of urban transport networks should make users pay the social cost of their travel. For instance, as road networks suffer from congestion externalities, economic theory suggests that congestion pricing can increase efficiency. This may also have an impact on the distribution of welfare across society. Concerns of a possible regressive effect recently impeded the introduction of road pricing in New York City and Paris.\footnote{In a recent interview, New York State Assemblyman Richard L. Brodsky said he opposed its introduction “for the reason that these schemes put the burden for paying the fees on blue blood and blue collar alike” (see New York Times, “Congestion Pricing: Just Another Regressive Tax?” www.nytimes.com).} Distributional issues also animate debates on infrastructure projects, especially for roads.\footnote{Environmentalist NGO Friends of the Earth recently published a report criticizing the UK government’s plan for spending £5 billion in new road projects, claiming that “road building is regressive spending (as roads are) used far more by richer households” (Friends of the Earth, 2010 p.4).} The economic literature has looked at redistributive concerns in pricing and provision of transportation infrastructure (see Small and Verhoef (2007) for a comprehensive review). However, it has done so (with an important exception discussed below) ignoring the presence of income taxation.\footnote{For road pricing, the literature has also emphasized the importance of revenue redistribution as a determinant of its progressivity (see RAND (2009) for a survey). This implies, however, that other instruments (in particular, income taxes) that the government can use are taken as given. Our work takes a different approach since it considers transport pricing as part of a whole set of instruments that are simultaneously optimized by the government.} This leaves open the question of whether these concerns are actually relevant, as they could possibly be addressed with appropriately designed income taxes. The main objective of this paper is to study such a question.

We consider the problem of a welfare maximizing government that designs both nonlinear income taxes and tariffs for roads and public transportation. We study both the case in which nonlinear tariffs are available and that in which only linear ones can be used.\footnote{Nonlinear pricing is more demanding in technological and informational terms than linear pricing, as it requires observability of individual trip quantities. However, it is becoming increasingly common in transportation services. See, for instance, Wang \textit{et al.} (2011) for a study of nonlinear pricing of tolled roads and Batare and Ivaldi (2011) for public transportation. Cremer and Galvani (2002) study nonlinear pricing by a public firm in the presence of optimal income taxation.} The government also provides infrastructure (e.g. roads or metro lines). Individuals are heterogeneous in (exogenous) earning ability, which is assumed to be private information, as is their
labor supply. Thus, the government faces self-selection constraints that may limit welfare redistribution. To keep the setup as simple as possible, we use a model with only two types of individuals (à la Stiglitz (1982)).

Previous public finance literature has studied how (if at all) a government that can use income taxes should deviate, due to distributional concerns, from correcting market failures (Cremer et al. (1998), Bovenberg and Goulder (2002), Kaplow (2006)). However, it has disregarded two relevant features for transportation, which are central in our analysis. The first is that consumption of transport goods requires travel time. Boadway and Gahvari (2006) and Gahvari (2007) consider time of consumption in an optimal redistributive taxation framework. They do not consider externalities or public goods.\(^5\) Mayeres and Proost (1997) study optimal redistributive taxation in the presence of congestion externalities, but restrict attention to linear taxes. Using such an approach, pricing and provision of transport infrastructure may be a means to compensate for inappropriate tax instruments. Our approach is complementary, since it does not assume restrictions on the design of income taxes (it is constrained only by the available information). A second key feature of our setup is that we explicitly model the relation between travel and labor supply. Individuals can decide the number of days at the workplace (which require commuting) and the (average) length of their working day. Substituting working days for more hours worked per day implies a penalty in terms of productivity. This is due to diminishing returns in daily work hours, as in van Ommeren and Gutierrez-i-Puigarnau (2010).\(^6\) As a consequence, individuals face an important trade-off when deciding on their work (and travel) schedule. On the one hand, an additional working day requires a time-consuming commuting trip. On the other, it allows, for given earned income, to reduce total hours worked. While labor supply plays a central role

\(^5\)Cremer et al. (1998) and Kaplow (2006) studied environmental levies in the presence of nonlinear income taxation. They consider a model where commodities do not require any time for consumption. Moreover, they focus on externalities that do not affect the marginal cost of consuming goods, unlike traffic congestion. An optimal taxation model with time as input for activities and congestion externalities is also studied in De Borger (2011). He focuses, unlike us, on a representative agent framework.

\(^6\)For evidence (at the aggregate level) on diminishing labor productivity with longer work days, see, e.g. Bourlès and Cette (2007).
in models of income taxation, little attention has been dedicated to the impact of transport policies that affect commuting.\(^7\) Parry and Bento (2001) and Van Dender (2003) consider the issue, although only in a setup with homogeneous individuals. Moreover, their model of labor supply is more rigid than ours, allowing only to choose working days (of fixed length). De Borger and Wuyts (2011) study a model with telework but, again, without looking at distributional concerns.\(^8\)

We show, to begin, that pricing and provision of infrastructure have a redistributive role even in the presence of nonlinear income taxation. They can be used to improve screening of types, relaxing the self selection constraints.\(^9\) This is true in spite of the fact that individuals’ preferences are separable in goods and leisure. The reason is that travel requires time (which has the same opportunity cost as labor time in our setup) and that, due to diminishing returns in daily hours, commuting is complementary to labor supply.\(^10\) We find that optimal per-trip charges should be differentiated across individuals of different earning ability. Therefore, when feasible, nonlinear transport tariffs are necessary to implement the second-best allocation (constrained by self-selection). A commuting trip on a given mode may consume more time than it frees up by allowing to reduce average workday length. If so, it is optimal to have high ability individuals pay cheaper per-trip tariffs than low ability ones. If not, the opposite is optimal. This is because, for a given “bundle” of goods and income, the more able have to work less than the others. Therefore, time is more valuable to them. As a consequence, discouraging commuting (via the given mode) may hurt low ability types

\(^{7}\)Although forms of work which do not require physical presence at the workplace are becoming increasingly common, commuting is still a necessity for the large majority of people. For instance, in Belgium an estimated 71% of all person movements in the morning peak in 2000 were for commuting purposes (according to the Mobility Portal of StatBel).

\(^{8}\)We disregard telework for simplicity. However, the model could accommodate it with no impact on the qualitative results. Indeed, assuming working outside the office (e.g., at home) to be less productive than working on-the-job, individuals would face a similar trade-off when deciding on how often they commute.

\(^{9}\)Indeed, in the absence of such constraints, the second welfare theorem implies that optimal prices have to be equal to marginal social costs and optimal provision follows a standard Samuelson’s rule. This is because the government has access to type-specific lump-sum taxes and transfers.

\(^{10}\)In the absence of time costs of travel and of diminishing returns in daily hours, separability would make distributive concerns irrelevant when designing pricing and provision of transport infrastructure. This follows from Atkinson and Stiglitz (1976).
less, in the first case, or more, in the second, than high ability types (at a given bundle), making mimicking less (resp. more) interesting. If the government is constrained to use only linear tariffs, the same trade-off is still a key determinant of the way in which tariffs can be used to relax incentive constraints.

To continue, different value of time for individuals of different ability (again, at a given bundle) implies that a reduction in the time cost of a journey, either because congestion is curbed with higher car tariffs or transportation infrastructure is upgraded, can also improve screening. Thus, a redistribution-minded government should have an additional reason to pursue such policies. This is true for both roads and public transportation. These results are in line with Kreiner and Verdelin (2012), who consider a more general setup but focus on provision of public goods. They neglect pricing activities that make use of them (e.g. car trips for road infrastructure). We investigate optimal provision rules both with optimal and non optimal pricing. We find that, due to induced-demand effects, the extent to which provision of infrastructure improves redistribution may be limited when tariffs are set below marginal costs.\footnote{See also earlier results by Boadway and Keen (1993) and Pirttilä and Tuomala (1997). In their setup, however, the redistributive role of public goods/bads does not survive if individuals have separable preferences for goods and leisure.}

As argued above, transport tariffs that closely follow marginal social costs may be hurtful to the poor. The government may thus want to differentiate them based on income, that is, use introduce means-testing. The suitability of means-testing for urban transportation, as for other government-provided services (e.g., healthcare or schools), is part of the current policy debate (see, e.g. Estupinan et al. (2007)).\footnote{This tariff differentiation already exists, at least to some extent. Discounted or free passes for public transportation are often provided to people qualifying for certain criteria, including income. This is the case, to make an example, of the “Forfaits Solidarité Transport” in the french Ile-de-France region. The issue seems relevant also for road pricing. The planned congestion charging scheme for the city of San Francisco should include a 50% discount granted to low income (Lifeline value) drivers.} In our framework, when individual trip quantities are unobservable, differentiation of (linear) tariffs is incentive-compatible only if conditional on income. Means-testing is, therefore, trivially necessary. Tariff differentiation and means-testing may instead be decoupled when nonlinear tariffs are feasible. Our results support
the optimality of differentiated tariffs for individuals of different earning ability. However, this does not necessarily mean that they should be explicitly conditioned on income. In the last part of the paper, we turn our attention to such a question.\footnote{The results described above are obtained under the assumption that the government uses a general “tax-and-pricing” function, based on income and trip quantities. This means that optimal transport tariffs should, a priori, be conditional on income. Income taxes may also have to be conditional on travel quantities. As an example, consider the income tax deductions of commuting expenditures which is common in many countries, including Germany and Belgium. Means-testing is not required if transport tariffs and income taxes can be implemented separately.} We provide a sufficient condition under which means-testing is not required. Such condition holds when modal split, in the second-best allocation, is such that high income households commute more by car than low income ones and public transport trips have larger time costs than trips by car. We conduct numerical simulations in the final section of the paper. We find only very few counterexamples in which implementability without means-testing fails.

The rest of the paper is organized as follows: Section 2 presents the model. We present optimal tariffs in Section 3. Section 4 looks at infrastructure provision. Section 5 considers implementation and means-testing. Section 6 presents some numerical illustrations of the results. Section 7 concludes. Proofs of all propositions are provided in an Appendix.

2 The model

2.1 Setup

We consider a population composed of two types of individuals $i = 1, 2$. They differ in earning ability (a measure of their productivity at work), identified by the parameter $w_i$, with $w_2 > w_1$. The size of group $i$ is denoted $\pi_i$, with $\sum_{i=1,2} \pi_i = 1$.

There are five goods in the economy: composite consumption $C$ (the numeraire), (peak-hour) trips by car $D$ and public transportation $B$, leisure $x$ and labor supply $L$. The production technology is linear in labor, with constant marginal costs normalized to one, for $C$ and $D$. The production sector is perfectly competitive. The marginal cost of a public transport
trip, sustained by the government (assumed to be the provider of the service), is constant and equal to $c_B$.\footnote{Since public transport is assumed to operate below capacity, the marginal cost of a trip could be zero (see Small and Verhoef (2007)). $c_B$ can be taken to be arbitrarily small.}

A trip by car or public transport has a fixed length and requires $a_j j = D, B$ units of time, for all individuals.\footnote{Transport trips can be seen as activities obtained combining goods and time. We assume a fixed-proportions household production technology, as in Kleven (2004), so our formulation is consistent with that representation.} We assume the time spent consuming $C$ to be a (perfect) substitute for leisure. Thus, contrary to time on travel and at work, it has no opportunity cost.\footnote{see Boadway and Gahvari (2006).}

Individuals face the time constraint

$$a_D D^i + a_B B^i + L^i + x^i \leq 1 \quad i = 1, 2$$

Suffix $i$ stands for individual quantities, which may vary depending on the individual's type. We normalize the time endowment to one (same for all types). To capture road congestion, we assume that $a_D$ is an increasing and convex function of the aggregate amount of car trips. Congestion on public transport is ignored for simplicity.\footnote{Putting it into the model would be an easy extension. It would make the optimal tariff formulae more complicated without adding to the results.} For both networks, trip time costs are functions of the amount of infrastructure provided (i.e. network capacity), denoted $K_j j = D, B$. Therefore $a_D = \varphi^D(\bar{D}, K_D)$ and $a_B = \varphi^B(K_B)$, where $\frac{\partial \varphi^D}{\partial \bar{D}}, \frac{\partial^2 \varphi^D}{\partial \bar{D}^2} > 0$, with $\bar{D} = \sum_{i=1,2} \pi_i D^i$ denoting the total level of congestion, and $\frac{\partial \varphi^D}{\partial K_D}, \frac{\partial \varphi^B}{\partial K_B} < 0$. We assume that $K_j j = D, B$ has a constant unit cost $c_{K_j}$. We assume also that $a_D$ is taken as given when deciding how many car trips to take, which generates a congestion externality.

Individuals choose the amount of labor supply deciding on two key parameters: $N$, the number of working days, and $h$, the amount of hours worked per day on-the-job.\footnote{This should be interpreted as the choice between jobs offering different workdays-hours schedules. There is evidence that labor markets are becoming more flexible in this respect (van Ommeren and Gutierrez-i-Puigarnau (2010)). Moreover, $h$ may also be interpreted as a measure of effort provided (for a given number of hours) per day on the job (e.g. the time the individual focuses on the job's tasks).} Labor supply is thus $L = N \cdot h$. All individuals are assumed to be commuters and to use the transport
network only for this purpose (which we consider a reasonable simplification given our focus on peak-hour travel, when congestion problems are more relevant). A day at the workplace requires a return commuting trip, on one of the two modes.\footnote{We ignore telework for simplicity. However, our model could accommodate for such a form of work without changing the main results. We discuss this in Section 3.} Therefore: $N^i = D^i + B^i$.

Finally, individuals’ income is obtained as

$$I^i = N^i w^i f(h^i) = (D^i + B^i) w^i f(h^i) \quad i = 1, 2$$

where, importantly $f' > 0$, $f'' < 0$. The assumption of decreasing returns in hours per day captures diminishing productivity when staying longer at work or when increasing daily effort.\footnote{We implicitly assume labor is paid at its marginal productivity. While it is conceivable that for low enough values of $h$ the marginal productivity be increasing, we can safely assume that workers will always choose an amount of $h$ such that productivity is diminishing. Otherwise, there would be no reason to split work time across multiple days, given that each day of labor supplied has a fixed commuting cost.} For a given individual of type $w^i$, total labor supply (which is unobservable) can be rewritten as

$$L^i = N^i \cdot h^i = (D^i + B^i) g\left(\frac{I^i}{w^i (D^i + B^i)}\right) \quad \text{where} \quad h^i = g\left(\frac{I^i}{w^i (D^i + B^i)}\right)$$

using the fact that $N^i = D^i + B^i$ and that $f(h^i) = I^i / w^i(\text{Di} + \text{Bi})$. Function $g(.)$ is the inverse of $f(.)$. Therefore $g' > 0$, $g'' > 0$. As may already be understood, diminishing returns generate an important trade-off when deciding on the work and travel schedule. If an individual travels one more day to work, she has to sustain the monetary and time costs of a commuting trip. On the other hand, doing so allows, for given income, to reduce total labor supply. This is because hours per day are reduced and average hourly productivity goes up. We will see below that such a trade-off has important implications for the optimal tariff schemes. Note also that diminishing returns determine (imperfect) complementarity between commuting and total labor supply. If returns were constant, the two would be independent as individuals could substitute working days for daily hours at no cost.
All individuals have the same utility function

\[ U(C, D, B, x) = \Omega(C) + \gamma(D, B) + \phi(x) \]

Note the separability between leisure and goods. In the absence of time costs of travel and of diminishing returns on hours worked per day, this would imply that individuals of different types choosing the same allocation of goods and income have the same marginal rate of substitution between consumption and transport trips. This, in turn, would yield the Atkinson-Stiglitz (1976) result of redundancy of marginal tariffs (except for pigouvian ones). However, in our model this result does not hold. We assume \( \Omega(.) \) and \( \phi(.) \) to be increasing and concave. As for \( \gamma(.) \), it may be increasing or decreasing in \( D \) and \( B \). Transport trips, though necessary for commuting, may provide some utility to the individual (which could be interpreted as an additional purpose of the trip, such as escorting kids to school, i.e. “trip chaining”), or disutility (e.g. stress). In this we follow Parry and Bento (2001) and Van Dender (2003).

The objective of the government is to maximize the social welfare function

\[ W = \sum_{i=1,2} \delta_i U^i \]

where \( \delta^i \) are positive weights, with the normalization \( \sum_{i=1,2} \delta^i = 1 \). We impose no a priori restriction on the instruments it may use, except from the information at its disposal. Assuming individual’s income to be observable, the government has access to a nonlinear income tax schedule \( T(I) \). As for transport trips, we are going to study two alternative scenarios. In the first, for each type \( i = 1, 2 \), individual trip quantities \( D^i \) and \( B^i \) can be observed.\(^{21}\) This allows to design nonlinear transport tariffs. In the second scenario, only

\[^{21}\text{Assuming observability of individual trip quantities is essential for nonlinear tariffs to be feasible. In practice, road and public transport pricing schemes often involve nonlinearities, to various extents. The assumption seems, thus, to be reasonable. Indeed, urban road pricing usually involves the use of electronic systems allowing to track each car’s access to the tolled road. This allows to observe trip quantities. As for public transportation, many cities have adopted the use of smart cards (e.g. the Oyster Card in London or} \]
aggregate trip quantities are observable to the government. Then, linear tariffs are the only feasible instrument. The government also designs network capacity $K_j$.

3 Optimal transport tariffs with nonlinear income taxation

The objective of this section is to present optimal transport tariffs in the presence of nonlinear income taxation, assuming a fixed supply of infrastructure. Infrastructure provision will be considered in Section 4.

3.1 Optimal nonlinear transport tariffs

3.1.1 Government’s maximization problem

When, on top of income $I$, individual consumption of $D$ and $B$ can be observed, the design of nonlinear transport tariffs is essentially akin to that of nonlinear commodity (as well as income) taxes.\footnote{As for $C$, with observable income, if transport trips are observable then individual’s consumption level is observable as well.} We begin by rewriting the utility function of a given type in terms of observable quantities. We also saturate the time constraint and replace for $x$, so

$$U^i = \Omega(C^i) + \gamma(D^i, B^i) + \phi \left( 1 - a_D D^i - a_B B^i - (D^i + B^i) g \left( \frac{I^i}{w^i(D^i + B^i)} \right) \right) \quad i = 1, 2$$

$U^i$ is type-specific since, for a given allocation, it depends on $w_i$. We proceed as if the government directly chose allocations, for each type of individual, of $C, D, B$ and $I$. This follows the Taxation Principle (Stiglitz (1982)). The government’s problem is

$$\max_{\{C^i, D^i, B^i, I^i\}} W$$

the Passe Navigo in Paris) for public transport, which require personal registration and allow to keep track of trips taken.
subject to the budget constraint

\[
\sum_{i=1}^{\sum_i} \left( I^i - C^i - D^i - c_B B^i \right) \geq c_D K_D + c_B K_B + R
\]

(1)

(where \( R \) is an exogenous revenue requirement)\(^{23}\) and, assuming only one self selection constraint is relevant (this is a reasonable assumption in a two-type setup like ours)

\[
U^2 \geq U^{21}
\]

(2)

where

\[
U^{21} = \Omega(C^1) + \gamma \left( D^1, B^1 \right) + \phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) \right) g \left( \frac{I^1}{w^2(D^1 + B^1)} \right)
\]

is the utility of a high ability type mimicking a low ability one. Constraint (2) tells us that the optimal allocations designed by the government have to be such that individuals of high earning ability do not chose the “bundle” (of income, travel quantities and consumption) intended for low ability ones. Note that when mimicking, an high ability type will need to work less while earning the same income and consuming the same amount of \( C, D \) and \( B \) as the type he mimicks. In this framework, mimickers commute to work the same number of days as the mimicked, but provide less hours of work per day.

To implement the optimal allocation, the government sets nonlinear tariffs for the transport network (i.e. road and public transport pricing schedules) as well as nonlinear income taxes. More precisely, the government designs a general general tax function \( \Theta(C, D, B, I) \) based on all observable quantities. This function has four key components: the per trip tariffs \( t_D, t_B \), a marginal tax on income \( t_I \) and a lump-sum tax/transfer \( T \).\(^{24}\) We assume, without

---

\(^{23}\)Any fixed cost of providing transport services incurred by the government can be optimally covered through the “fixed” components of the tariffs or through income taxes. Their presence does not affect the results.

\(^{24}\)They are part of nonlinear schedules, which, a priori, depend on all quantities observed. This means, for instance, that the transport tariff schedules may be dependent on individual income, i.e. they may have to be means-tested. We discuss the need for means-testing in Section 5. We argue there that, under reasonable
loss of generality, that good $C$ is untaxed. The Lagrangian of the government’s problem is

$$\mathcal{L} = W + \mu \left( \sum_{i=1,2} \pi_i (I^i - C^i - D^i - c_B B^i) - \sum_{j=D,B} c_{K_j} K_j - R \right) + \lambda \left( U^2 - U^{21} \right)$$

The first order conditions of this problem are provided in the Appendix. Before continuing, it is useful to illustrate the adjustment in labor supply induced by a marginal change in commuting travel. For a given type, the latter writes as

$$m_i = -g^i \left( \frac{I^i}{w^i(D^i + B^i)} \right) \cdot \left( \frac{I^i}{w^i(D^i + B^i)} \right) + g \left( \frac{I^i}{w^i(D^i + B^i)} \right) \quad i = 1, 2$$

Note that $m_i < 0$ due to convexity of $g(.)$. Given that hours per day at the workplace have diminishing returns, marginally increasing the number of commuting days (i.e. trips) allows the individual, for a given income, to reduce total labor supply. We have also

$$m_{21} = -g^1 \left( \frac{I^1}{w^2(D^1 + B^1)} \right) \cdot \left( \frac{I^1}{w^2(D^1 + B^1)} \right) + g \left( \frac{I^1}{w^2(D^1 + B^1)} \right)$$

as the adjustment for the individual of type 2 mimicking an individual of type 1. It is easy to see that $m_1 < m_{21} < 0$, as $w^2 > w^1$. Individuals of higher ability can, for a given income level, substitute hours worked for days at the workplace suffering smaller productivity losses. Mimicked and mimickers thus face a different trade-off, when deciding on their work schedule, between days at the workplace and workday hours. This has relevant implications for the optimal tariff schemes.

---

assumptions, such practice will not be necessary. The lump-sum tax/transfer may also include the “fixed” components of the transport tariff schemes (e.g. the fixed part of a two part tariff).
3.1.2 Benchmark

As a benchmark, consider the ideal case in which the government observes $w$, $L$, or both. Then $\lambda = 0$. The optimal per-trip tariffs are

$$t_D^1 = t_D^2 = \tau_D \equiv \frac{\partial \varphi^D}{\partial D} \sum_{i=1,2} \pi_i \left( D^i \frac{\phi_i^j}{\Omega^j_i} \right)$$

$\tau_D = \frac{\partial \varphi^D}{\partial D} \sum_{i=1,2} \pi_i \left( D^i \frac{\phi_i^j}{\Omega^j_i} \right)$

Let us begin from tariffs for car trips $t_D$: they should consist simply of a Pigouvian tax. Their only component is $\tau_D$, the marginal external cost of a trip. This is given by the increase in time of journeys (on aggregate) due to additional congestion on the road, weighted by the individuals’ the marginal rate of substitution between leisure and consumption $\frac{\phi_i^j}{\Omega^j_i}$, for $i = 1, 2$. Such ratio provides a measure of the individual’s valuation of time. Tariffs for public transportation $t_B$ should be equal to the marginal cost of providing the trip, $c_B$. Thus, in the presence of optimal income taxation, and if self selection constraints are not relevant, optimal tariff schedules should not deviate from the marginal social cost of a trip. This is because the government can use differentiated lump sum taxes to redistribute welfare and cover the eventual fixed costs of service provision (by the Second Welfare Theorem).

3.1.3 Optimal marginal tariffs with binding self-selection constraints

Consider now the case in which $w$ and $L$ are unobservable and the self selection constraint binds, so $\lambda > 0$. The following holds.

PROPOSITION 1: When nonlinear transport tariffs are feasible, the optimal per-trip tariffs for cars and public transport $t_j^i$ $i = 1, 2$ $j = D, B$ are

---

25 This can also represent the case of unobservable $w$ and $L$ but in which (2) is simply not binding at the optimal allocation.

26 There is a large literature on the value of time in transportation (Jara-Diaz (2008)). Generally, it corresponds to the wage rate corrected for the additional utility (or disutility) of time spent on travel, in monetary terms. In our model, a unit of time at work and on travel have the same opportunity cost in terms of foregone leisure.
\[ t_D^1 = \tau_D + \eta_D + z_D \quad t_D^2 = \tau_D + \eta_D + z_D \quad t_B^1 = c_B + z_B \quad t_B^2 = c_B \]

where

\[ \eta_D = \frac{\lambda}{\mu} \frac{\partial \varphi_D}{\partial \bar{D}} \left( D^1 \Omega_{C}^{21} \left( \frac{\phi_x^1}{\Omega_{C}^{1}} - \frac{\phi_x^{21}}{\Omega_{C}^{21}} \right) \right) \]

\[ z_j = \frac{\lambda \Omega_{C}^{21}}{\mu \pi_1} \left( a_j \left( \frac{\phi_x^1}{\Omega_{C}^{1}} - \frac{\phi_x^{21}}{\Omega_{C}^{21}} \right) + \left( m_1 \frac{\phi_x^1}{\Omega_{C}^{1}} - m_2 \frac{\phi_x^{21}}{\Omega_{C}^{21}} \right) \right) \quad j = D, B \]

For both cars and public transport, marginal tariff formulae are different than in the benchmark case. They contain additional “incentive” terms, whose role is to improve screening of types. This is in spite of leisure-goods separability in preferences. Their presence depends on travel being time consuming and (partially) complementary to labor supply, due to diminishing returns in daily hours. Let us focus first on \( t_D \). The formulae contain, on top of \( \tau_D \), additional terms \( \eta_D \) and \( z_D \). The term \( \eta_D \) accounts for how a reduction in road congestion can foster screening. It is strictly positive. The reason is that a marginal reduction in \( a_D \) is always going to benefit low ability individuals more than high ability ones, at a given allocation of income and goods. The latter need to work less, earning the same income and traveling as much, but being endowed with higher productivity than low ability types. As a consequence, time is less valuable for them at the margin, so \( \frac{\phi_x^1}{\Omega_{C}^{1}} - \frac{\phi_x^{21}}{\Omega_{C}^{21}} > 0 \).

Both \( \tau_D \) and \( \eta_D \) affect the marginal tariffs for both types. Unlike in the benchmark case, the marginal external cost of a trip is not only the “classic” pigouvian one but has to include the extra cost of congestion in making redistribution less effective. The incentive effect of public goods (or bads, as in this case), in the presence of nonlinear income taxation, has been previously analyzed by Boadway and Keen (1993), Pirtilä and Tuomala (1997) and, more recently, Kreiner and Verdelin (2012). However, in the setup of the first two papers, this effect does not survive if individuals have separable preferences for goods and leisure, which we assume. Kreiner and Verdelin pointed out that such an effect exists as long as there is positive correlation between an individual’s ability and her willingness to pay for a
public good, at a given income and consumption bundle. This is indeed the case here since individuals have to allocate time to labor, leisure and travel. By relaxing the time constraint at the individual level, reductions in road congestion (or, as we will see below, improvements in transportation infrastructure) benefit more mimicked than mimickers.

The optimal marginal tariff for high ability types carries two components: \( \tau_D + \eta_D \). It is, therefore, strictly higher than a standard pigouvian tax. Note, also, that \( \eta_D \) is larger the more low ability types commute by car, \( D^1 \). With respect to \( t^2_D \), the per-trip tariff intended for individuals of low ability \( t^1_D \) carries the additional incentive component \( z_D \). This is an additional distortion that the government introduces to low types’ use of cars, in order to improve redistribution. Its sign depends on the trade-off between journey time and time freed by the reduction in total labor supply (for given income) when commuting one more day to the workplace. To be more precise, \( z_D \) depends on how differently it affects, at a given bundle of goods and income, individuals of high and low ability. The sides of such trade-off are represented by the two terms in the expression for \( z_D \). The first is given by the time cost of a trip by car \( a_D \), multiplied by difference of marginal valuation of time for low ability type and high ability (mimicker). For a mimicker time is less valuable, at the margin, than for a low ability individual. As a consequence, the first term contributes positively to \( z_D \). The second term stands for the effect that an additional day at workplace has on total labor supply (at given income). Recall that \( m_1 < m_{21} < 0 \). This, combined with the fact that time is more valuable for mimicked than mimicker, is why this term contributes negatively to \( z_D \). Said differently, it is optimal to further encourage individuals of low ability to use cars, with respect to the high ability types, if and only if an additional car trip allows to save time (in net terms). That is, the labor supply it saves is more relevant than the time cost of the additional journey. This is because, time being more valuable for mimicked than

\[^{27}\text{Previous analyses of transport pricing with redistributive concerns (see, e.g., Mayeres and Proost (1997)) have suggested that (linear) taxes on a given mode should be higher when the mode is used to a large extent by high income individuals. In our model, this does not apply. High income (and ability) types pay, with respect to the pigouvian tax, a “premium” which is higher the more low income types use cars. Moreover, it is not necessarily the case that high ability/income types should be charged more, for a trip, than low ability/income ones.}\]
mimickers, the former will benefit more from the additional commuting trip. It is not easy to say, a priori, which of the two effects above is of greater relevance. We will have more to say about this in the numerical examples provided in Section 6. To fix ideas, consider two extreme cases. Suppose, first, that the time cost of a commuting trip were negligible, so \(a_D \to 0\), and maintain the assumption of diminishing returns on daily work hours, so \(f'' < 0\) (and \(g'' > 0\)). In that case, we have \(z_D < 0\). Suppose, instead, that hours worked per day had constant returns, so \(f'' = g'' = 0\), and the time cost of a car trip were non-negligible, \(a_D > 0\). In that case, \(m_1 = m_{21} = 0\) and \(z_D > 0\). Focusing now on public transport tariffs \(t_B\), marginal tariffs for low ability types \(t_{i_B}\) also carry the component \(z_B\), whose nature is similar to \(z_D\) discussed above. Except, of course, that the relevant time cost of a trip is \(a_B\).

### 3.2 Optimal linear transport tariffs

Let us now consider the case in which individual trip quantities are not observable. Then, the government has to design a “mixed” tax system with nonlinear income taxes and linear tariffs for transportation.

#### 3.2.1 Government’s maximization problem

We proceed, following Cremer et al. (1998), under the assumption that the government designs an optimal revelation mechanism consisting of a set of type-specific before-tax incomes \(I^i\), disposable incomes \(y^i\) (expenditures on consumption and travel) and a vector of transport tariffs \(t = (t_D, t_B)\), which are akin to commodity taxes. Equivalently, the mechanism designs trip prices \(q = (q_D, q_B)\) where \(q_D = 1 + t_D\) and \(q_B = t_B\). Again, without loss of generality, we assume \(C\) is untaxed. The mechanism assigns the bundle \((q, y^i, I^i)\) to an individual that reports type \(i = 1, 2\). The couple \((y^i, I^i)\) is such that \(I^i - T(I^i) = y^i\), where \(T(I)\) is the income tax schedule. Given prices and disposable income, the individual decides consumption and
travel quantities. That is, given \((q, y^i, I^i)\), a type-\(i\) individual solves

\[
\max_{C,D,B} U^i(C, D, B, y, I) \quad i = 1, 2
\]

(note that the utility function \(U^i(C, D, B, y, I)\) is type specific because, at a given allocation, it depends on \(w_i\)) subject to the budget constraint

\[
C + q_D D + q_B B = y
\]

We denote the resulting conditional demand functions as

\[
D^i = D^i(q, y^i, I^i) \quad B^i = B^i(q, y^i, I^i) \quad C^i = C^i(q, y^i, I^i)
\]

again, demands are type specific (given \(q, y^i, I^i\)) since utility depend on \(w_i\). We denote the (type-specific) indirect utility function as \(V^i(q, y^i, I^i) = U^i(D^i, B^i, C^i, y^i, I^i)\). Finally, we define

\[
D^{21} = D^2(q, y^1, I^1) \quad B^{21} = B^2(q, y^1, I^1) \\
C^{21} = C(q, y^1, I^1) \quad V^{21}(q, y^1, I^1) = U^2(D^{21}, B^{21}, C^{21}, y^1, I^1)
\]

as demands and indirect utility function for a mimicker. Once again, given the presence of only two types in our setup, we can safely focus only on cases in which high ability types want to mimic low ability ones. The government’s problem is\(^{28}\)

\[
\max_{q,y^i,I^i,D} \sum_{i=1,2} \delta^i V^i
\]

\(^{28}\)It is convenient to solve this problem assuming that the government also decides on the amount of road congestion \(D\).
subject to the budget constraint

$$\sum_{i=1,2} \pi_i \left(I^i - y^i + t_D D^i + (t_B - c_B) B^i\right) \geq c_D K_D + c_B K_B + R \quad (3)$$

and the self-selection constraint

$$V^2 \geq V^{21} \quad (4)$$

we still denote by $\mu$ and $\lambda$ the Lagrange multipliers for these constraints. The solution to this problem is presented in the Appendix.

### 3.2.2 Benchmark

With no self-selection constraints binding, so $\lambda = 0$, optimal tariffs are $t_D = \frac{\partial \phi^D}{\partial D} \sum_{i=1,2} \pi_i \frac{\phi^i}{\partial y^i}$ and $t_B = c_B$. As in the previous section, they have no redistributive role. 29

### 3.2.3 Optimal transport tariffs with binding self-selection constraints

Consider now the case in which $w$ and $L$ are unobservable and the self selection constraint binds, so $\lambda > 0$. The following holds.

**PROPOSITION 2:** When the government is constrained to use linear transport tariffs, the optimal tariffs $t_j \quad j = D, B$ satisfy the following

$$
\begin{pmatrix}
t_D - \varepsilon \\
t_B - c_B
\end{pmatrix}
= A^{-1} \cdot 
\begin{pmatrix}
\frac{\lambda \partial V^{21}}{\mu} \left(\frac{\partial y}{\partial y} \right) \\
\frac{\lambda \partial V^{21}}{\mu} \left(\frac{\partial y}{\partial y} \right)
\end{pmatrix}
\begin{pmatrix}
(D^1 - D^{21}) \\
(B^1 - B^{21})
\end{pmatrix}
$$

where

$$A = \begin{pmatrix}
\sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial q_d} \chi \\
\sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial q_b} \chi
\end{pmatrix} 
\begin{pmatrix}
\sum_{i=1,2} \pi_i \left(\frac{\partial D^i}{\partial q_d} + \frac{\partial D^i}{\partial \theta_D} \frac{\partial \phi^i}{\partial \theta_D} \right) \\
\sum_{i=1,2} \pi_i \left(\frac{\partial D^i}{\partial q_b} + \frac{\partial D^i}{\partial \theta_D} \frac{\partial \phi^i}{\partial \theta_D} \right)
\end{pmatrix}
$$

29 We here write the value of time as $\phi^i_{\partial V^i/\partial y^i}$, $i = 1, 2$ since this form is more convenient for solving the problem below. With no binding self-selection constraints and nonlinear income taxation, one has $\partial V^i/\partial y^i = \mu = \Omega^i_{C^i}, i = 1, 2$. So the benchmark value of $t_D$ is the same as that of Section 3.1.
\[ \varepsilon = \frac{\partial \varphi^D}{\partial D} \left( \sum_{i=1,2} \pi_i \frac{\phi^D_i D^i}{\partial V^i/\partial y^i} + \frac{\lambda}{\mu} \frac{\partial V^{21}}{\partial y^1} \left( \frac{\phi^1 D^1}{\partial V^1/\partial y^1} - \frac{\phi^{21} D^{21}}{\partial V^{21}/\partial y^1} \right) \right) \]

\[ \chi = \frac{1}{1 - \frac{\partial \varphi^D}{\partial D} \sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial a}} < 1 \]

where \( \tilde{D}^i \) and \( \tilde{B}^i \) denote (aggregate) hicksian demands for, respectively, car and public transport travel. \( \chi \) is a feedback term that stands for the net effect of a change in prices on the demand for car trips, after accounting for the change in road congestion.

The structure of optimal tariffs is affected by two “incentive” terms. Their presence is not novel. Note, though, that they are not zero in spite of separability in preferences between goods and leisure. Once again, this is due to the fact that travel requires time (whose opportunity cost is not the same for individuals of different ability, at a given bundle of goods and income) and to the diminishing returns on daily hours. First, a reduction in road congestion affects the self-selection constraint in a similar way as in the case of nonlinear tariffs (compare the second component of \( \varepsilon \) above and term \( \eta_D \) in Proposition 1). It is not possible to precisely determine what direction the effect takes here. This is because, when tariffs are linear, mimicker and mimicked do not necessarily travel the same amount of times (even if they face the same budget constraint). Nonetheless, except if the mimicker drives many times more than the mimicked, it is reasonable to expect the sign of this term to be positive. The second incentive term (right hand side of the equalities in the proposition) is positive if and only if mimickers use more the given mode (car or public transport) than mimicked. Again, this cannot be immediately determined. To get more insight, we will now present a simplified example with a single travel mode. It will show that the signs of the incentive terms just described may depend crucially on the trade-off between time cost of commuting trips and labor supply reductions when substituting days at the workplace with more hours per day.
A single mode example  Consider the case in which cars are the only travel mode (we could focus on public transport with similar outcomes). Then the optimal tariff is simply

\[ t_D = \varepsilon + \frac{1}{\sum_{i=1,2} \pi_i \frac{\partial \tilde{D}_i}{\partial q} \lambda} \left( \frac{\lambda \partial V^{21}}{\mu \partial q_1} (D^1 - D^{21}) \right) \implies D^{21} \lesssim D^1 \iff t_D \gtrsim \varepsilon\]

In such a simplified setup, the budget constraint of mimicker and mimicked is the same. The only difference between them, at a given \((D_0, C_0)\) couple, is the \(I/w\) ratio (and, given this, the number of daily work hours). Thus, whether a mimicker drives more than a mimicked depends simply on whether her indifference curves in the \((D, C)\) plane are flatter than those of a mimicked. That is

\[ D^{21} \gtrsim D^1 \iff \sigma (D, C; I^1/w^2) \gtrsim \sigma (D, C; I^1/w^1) \]

at a given allocation. The slope of an indifference curve, computed at a given allocation \((D_0, C_0)\), is

\[ \sigma (D_0, C_0; I/w) = -\frac{\partial U/\partial D}{\partial U/\partial C} = -\gamma_D (D_0) + \phi_x \cdot (-a_D - m) \]

where

\[ m = g \left( \frac{I}{wD_0} \right) - g' \left( \frac{I}{wD_0} \right) \cdot \left( \frac{I}{wD_0} \right) < 0 \]

taking the derivative of \(\sigma\) with respect to \(I/w\), one obtains

\[ \frac{\partial \sigma}{\partial (I/w)} = -\phi_{xx} \cdot g' \left( \frac{I}{wD_0} \right) \cdot (-a_D - m) - \phi_x \cdot g'' \left( \frac{I}{wD_0} \right) \cdot \left( \frac{I}{wD_0^2} \right) \]

The sign depends on the trade-off between days at the workplace and commuting trips that drives the difference between marginal tariffs when nonlinear pricing is feasible (see term \(z_j\) in Proposition 1). Indeed, when the time cost of a commuting trip is larger (resp. smaller) than the reduction in labor supply with more day-on-the-job, then \(\frac{\partial \sigma}{\partial (I/w)} < 0\) (resp. \(> 0\)).
Therefore, $D^2_1 < D^1$ (resp. $> D^1$). Since the sign of $\frac{\partial \sigma}{\partial (I/w)}$ is not immediately determined, it is once again useful to look at two extreme cases. Suppose that the time cost of a trip were negligible, so $a_D \to 0$ while $f'' < 0$ (and $g'' > 0$). In that case, $\sigma(D, C; I^1/w^2) < \sigma(D, C; I^1/w^1)$. Then, $t_D < \varepsilon$. Moreover, if a low ability type drives more than a mimicker, the incentive term in $\varepsilon$ is will certainly be positive. Suppose, instead, that hours worked per day had constant returns, so $f'' = g'' = 0$, and the time cost of a car trip were non-negligible, $a_D > 0$. In that case, $\sigma(D, C; I^1/w^2) > \sigma(D, C; I^1/w^1)$ and $t_D > \varepsilon$.

4 Optimal provision of transportation infrastructure

We now look at public provision of transportation infrastructure. The question is what criterion should a redistribution-minded government follow when deciding how much infrastructure to provide, given that it can use optimal income taxation and also design the tariffs to access the infrastructure. The answer varies depending on whether the government can use nonlinear pricing schedules or it is constrained to use only linear tariffs.

4.1 Optimal provision rules when nonlinear tariffs are available

The problem of the government is the same as in Section 3.1, except that it can now choose network capacity for both modes $K_j$. We have the following

PROPOSITION 3: When the government sets optimal nonlinear transport tariffs (as well as income taxes), the optimal rule for provision of infrastructure $K_j$ is

$$- \frac{\partial \phi_j}{\partial K_j} \sum_{i=1,2} \pi_i \phi^i_x Q^i_j + \gamma_j = c_{K_j}, \quad j = D, B$$

where

$$\gamma_j = - \frac{\partial \phi_j}{\partial K_j} \mu Q^1_j \frac{\phi^1_x}{\Omega^1_C} - \frac{\phi^2_{21}}{\Omega^2_{21}} > 0 \quad Q^i_D = D^i \quad Q^1_B = B^i$$
The first term on the left hand side is the sum of reductions in journey times, for given traffic volume. With no (binding) self selection constraints (i.e. if $\lambda = 0$), this has to be optimally equated to marginal capacity cost $c_{K_j}$ (as in a standard Samuelson’s Rule). However, if (2) binds, the government should also take into account the positive term $\gamma_j$. This is the effect that increased capacity has on improving screening of types which is analogous to that of term $\eta_D$ discussed in the previous section. A redistribution-minded government that can rely on nonlinear income taxes should thus stimulate investment in transport infrastructure more than what is prescribed by the standard Samuelson’s rule. This is in line with Kreiner and Verdelin (2012) (see the discussion on the redistributive role of road congestion in Section 3.1). However, their focus is only on provision of a public good and neglects pricing goods (or activities) that make use of it. Indeed, we have obtained the result assuming optimal nonlinear pricing schedules. In the absence of optimal pricing, provision rules may have to be modified. This is what we investigate below.

4.2 Optimal provision rules when only linear tariffs are available

We now focus on the case in which the government is constrained, due to limited information, to use linear tariffs to price access to the transportation infrastructure. Consider now the problem of Section 3.2 and include capacity levels $K_j$ as decision variables. We have the following

PROPOSITION 4: When the government is constrained to use linear transport tariffs (but can use nonlinear income taxes), the optimal rule for provision of road infrastructure $K_D$ is

\[-\frac{\partial \varphi^D}{\partial K_D} \left( \sum_{i=1,2} \pi_i \frac{D^i \phi_x^i}{\partial y^i} + \lambda \frac{\partial V^{21}}{\partial y^1} \left( \frac{\phi_x^1 D^1}{\partial V^1/\partial y^1} - \frac{\phi_x^{21} D^{21}}{\partial V^{21}/\partial y^1} \right) \right) +
\frac{\partial \varphi^D}{\partial K_D} \sum_{i=1,2} \pi_i \left( (t_D - \varepsilon) \frac{\partial D^i}{\partial \alpha_D} \chi + (t_B - c_B) \frac{\partial B^i}{\partial \alpha_D} \chi \right) = c_{K_D}\]
and for public transportation $K_B$ is

$$
- \frac{\partial \varphi^B}{\partial K_B} \left( \sum_{i=1,2} \pi_i \frac{B^i \phi_x^i}{\partial V^i / \partial y^i} + \frac{\lambda}{\mu} \frac{\partial V^{21}}{\partial y_1} \left( \frac{\phi_x^1 B^1}{\partial V^1 / \partial y^1} - \frac{\phi_x^{21} B^{21}}{\partial V^{21} / \partial y^1} \right) \right) + \\
+ \frac{\partial \varphi^B}{\partial K_B} \sum_{i=1,2} \pi_i \left( (t_D - \varepsilon) \frac{\partial \tilde{D}^i}{\partial a_B} \chi + (t_B - c_B) \left( \frac{\partial \tilde{B}^i}{\partial a_B} + \frac{\partial \varphi^D}{\partial D} \cdot \frac{\partial \tilde{D}^i}{\partial a_D} \cdot \chi \right) \right) = c_{K_B}
$$

where

$$
\varepsilon = \frac{\partial \varphi^D}{\partial D} \left( \sum_{i=1,2} \pi_i \frac{\phi_x^i D^i}{\partial V^i / \partial y^i} + \frac{\lambda}{\mu} \frac{\partial V^{21}}{\partial y_1} \left( \frac{\phi_x^1 D^1}{\partial V^1 / \partial y^1} - \frac{\phi_x^{21} D^{21}}{\partial V^{21} / \partial y^1} \right) \right) \quad \chi = \frac{1}{1 - \frac{\partial \varphi^D}{\partial D} \sum_{i=1,2} \pi_i \frac{\partial \tilde{D}^i}{\partial a_D}}
$$

This rule is similar to the one provided in Pirttilä and Tuomala (1997). The gross marginal benefit of providing additional infrastructure is given by the benefit of reducing travel time (first row in the equation above). When redistribution is a concern, this also includes an “indirect” incentive effect on self-selection constraints (second term), whose nature is similar to that of $\gamma_j$ in Proposition 3. However, when the government has to use linear tariffs, there are additional terms affecting optimal provision. First, we have standard tax-base effects, as changes in infrastructure provision modify tariff revenues through “induced-demand” effects. Second, these effects also influence external costs. When the road tariff $t_D$ is lower than $\varepsilon$, the social benefit of an additional car trip is smaller than its private benefit. Therefore, the government may want to limit expansion of road capacity in order to curb additional car trips (see Small and Verhoef (2007, ch.5)). The interesting thing to notice here, stemming from the interaction of transport pricing and income taxation, is that induced demand may limit both the direct benefits of capacity expansion and the “indirect” ones related to redistribution. This is easily seen by rewriting the optimal provision rule for $j = D$ as
\[- \frac{\partial \varphi^D}{\partial K_D} \left( \sum_{i=1,2} \pi_i \left( \frac{\partial V^i}{\partial y^i} \right) + \frac{\lambda}{\mu} \frac{\partial V\!^{21}}{\partial y_1} \left( \frac{\partial V^1}{\partial y^1} - \frac{\partial V^{21}}{\partial y^1} \right) \right) \chi + \frac{\partial \varphi^D}{\partial K_D} \sum_{i=1,2} \pi_i \left( t_D \frac{\partial \tilde{D}^i}{\partial a_D} \chi + (t_B - c_B) \frac{\partial \tilde{B}^i}{\partial a_D} \chi \right) = c_{K_D} \]

Consider, for example, the case in which \( t_D = 0 \) and \( t_B = c_B \). This may be relevant in practice due to political economy constraints on the government’s ability to price transport trips (see the anecdotal evidence reported in the Introduction). Then the (gross) social benefit of providing road capacity \( K_D \), including the redistributive effect of relaxing self-selection constraints, has to be discounted by \( \chi \). There is too large induced demand due to uninternalized external (congestion) costs. Thus, induced demand will diminish the redistributive impact of upgrading road infrastructure (although it will not remove it).\(^{30}\) On the other hand, for public transportation infrastructure, the same concerns may actually increase incentives for investment. As long as the road tariff is not equal to the marginal external cost of a car trip, improving public transport quality has an additional (progressive) redistributive impact.

5 Implementation of optimal tax and tariff schedules

The question we investigate now is whether means-testing is a useful tool for a redistribution-minded government designing both transport tariffs and income taxes. This responds to some questions raised in the policy debate on reforming transport pricing (see the Introduction). The answer is trivial when only linear tariffs are feasible: differentiated tariffs for individuals of different ability can be incentive-compatible only if conditional on income. That is, means-

\(^{30}\) Even if car trips are priced at a “pigouvian” rate (i.e. \( t_D = \frac{\partial \varphi^D}{\partial K_D} \sum_{i=1,2} \pi_i \frac{\partial V^i}{\partial y^i} \)), the incentive term \( \frac{\partial \varphi^D}{\partial K_D} \left( \frac{\partial V\!^{21}}{\partial y_1} \left( \frac{\partial V^1}{\partial y^1} - \frac{\partial V^{21}}{\partial y^1} \right) \right) \) would still have to be discounted by \( \chi \). Of course, here we focus on rules, not quantities. With suboptimal tariffs, optimal road capacity may still have to be higher than when car trips are optimally priced.
testing is necessary. This is why we focus on the case in which the government can use nonlinear tariff schedules. Then, the answer is not straightforward.

In Section 3.1, we have assumed that the government implements the second-best allocation (defined as $A^{SB}$) using a generalized tax-and-tariff function $\Theta(C, D, B, I)$. This means that, a priori, it may have to design tariff schedules for transportation that are differentiated according to income.\(^{31}\) Following Cremer and Gahvari (2002), we are now going to study whether using an income tax function $T(I)$ and a separate transport tariff schedule $P(D, B)$ is enough to implement $A^{SB}$.\(^{32}\) If such a thing is feasible, then transport tariffs do not need to be means-tested. Studying this allows us to further qualify our previous results. Moreover, the question of implementability has practical relevance in itself, as argued above.

We assume, for simplicity, that capacities $K_j$ are fixed. The government looks to implement the second-best allocation

$$A^{SB} = ((C^1, D^1, B^1, I^1); (C^2, D^2, B^2, I^2))$$

that solves the problem presented in Section 3.1, using the functions $T(I)$ and $P(D, B)$. $T^i$ and $P^i$ $i = 1, 2$ denote respectively the payments of income taxes and transport tariffs for individuals of type 1 and 2. Therefore $C^i = I^i - (T^i + P^i)$ $i = 1, 2$. Incentive compatibility of the tax and tariff schedules calls for types 1 and 2 to choose quantities and payments $((T(I^1)); (P(D^1 + B^1)))$ and $((T(I^2)); (P(D^2 + B^2))$ respectively. The increased complexity stems from the fact that individuals have additional possibilities to deviate from the “bundle” designed for them. For instance, they may choose to consume a quantity of trips $D + B$ intended for the other type, while choosing the amount of $I$ intended for them. Or they could choose to mimick the other’s type income, while consuming the “right” amount of

\(^{31}\)Moreover, the income tax schedule may have to depend on commuting trips.

\(^{32}\)The setup of our problem is similar to that of Cremer and Gahvari. Our results are different. The reason is that, even with separable preferences, consumption of transport trips affects the marginal utility of leisure. This makes implementation with separable functions more difficult to achieve, as labor supply (and income) and consumption decisions cannot be separated. In addition, we do not assume any difference in tastes between individuals of different ability. Finally, our problem is of greater complexity due to the presence of two goods that the government has to price.
\(D + B\). Therefore, in order to be implementable through separable payment functions, \(A^{SB}\) has to respect the “standard” incentive compatibility constraint (2), the government’s budget constraint (1), plus four additional incentive constraints ensuring domination of “partial” mimicking strategies (each of them for \(i = 1, 2 \neq i\):

\[
\Omega(I^i - T^i - P^i) + \gamma(D^i, B^i) + \phi(1 - a_D D^i - a_B B^i - (D^i + B^i)) g \left( \frac{I^i}{w^i(D^i + B^i)} \right) \geq \\
\Omega(I^\tilde{i} - T^\tilde{i} - P^i) + \gamma(D^\tilde{i}, B^i) + \phi(1 - a_D D^\tilde{i} - a_B B^i - (D^\tilde{i} + B^i)) g \left( \frac{I^\tilde{i}}{w^i(D^\tilde{i} + B^i)} \right)
\]

(5)

\[
\Omega(I^i - T^i - P^i) + \gamma(D^i, B^i) + \phi(1 - a_D D^i - a_B B^i - (D^i + B^i)) g \left( \frac{I^i}{w^i(D^i + B^i)} \right) \geq \\
\Omega(I^\tilde{i} - T^\tilde{i} - P^\tilde{i}) + \gamma(D^\tilde{i}, B^\tilde{i}) + \phi(1 - a_D D^\tilde{i} - a_B B^\tilde{i} - (D^\tilde{i} + B^\tilde{i})) g \left( \frac{I^\tilde{i}}{w^i(D^\tilde{i} + B^\tilde{i})} \right)
\]

(6)

The first two ensure that an individual of type \(i\) will not, while choosing the number of transport trips intended for his type, choose income level intended for the other type (“partial mimicking” on income). The second set of constraints ensures an individual of type \(i\), while choosing the income intended for his type, will not mimic the other on transport trips. The solution of this problem is provided in the Appendix.

We now provide a sufficient condition under which using separable functions \(T(I)\) and \(P(D, B)\) is enough to implement \(A^{SB}\). As long as the condition holds, means-testing is not required.

PROPOSITION 5: Assume that the government wanted to implement the second best allocation \(A^{SB}\) using a separate payment schedule for income \(T(I)\) and transportation \(P(D, B)\). Then a sufficient condition for \(A^{SB}\) to be implementable is that it satisfies \(D^2 + B^2 \geq D^1 + B^1\) and \(a_D D^1 + a_B B^1 \geq a_D D^2 + a_B B^2\)
The condition requires that high ability/income households travel more, but their total travel time is smaller than for the others. This can be the case if high income households commute more by cars than low income ones, while public transport trips have larger time costs than trips by car.\textsuperscript{33}

In the numerical examples below, the condition given in Proposition 5 generally holds. In fact, even when it fails, we find no counterexample in which implementation with separable functions is unfeasible. We also go one step further. Instead of using functions $T(I)$ and $P(D,B)$, we study whether implementation of $A^{SB}$ can be achieved by complementing the income tax schedule $T(I)$ with \textit{two separate tariff schedules}, $P(D)$ for cars and $Q(B)$ for public transportation. The theoretical problem is similar to the one presented above, but considerably more complex to solve. The volume of conditions to be checked would make treating the problem in an analytical way simply too tedious. This is why we only investigate the issue numerically. The results obtained seem to support the conclusion that implementation is feasible even using fully separable transport tariff functions.

6 Numerical illustration

We present here a numerical example to illustrate the features of the optimal tariff schemes derived above. We are also interested in verifying that conditions for implementability in separable functions, as discussed in Section 5, reasonably hold. In order to focus on these two aspects, we only look at the case of nonlinear tariff schemes and consider fixed network capacities. The examples are based on the following utility and daily productivity functions

\[ U(C, D, B, x) = C^{\frac{3}{2}} + 0.05 \left( D^{\frac{1}{2}} + B^{\frac{1}{2}} \right) + 3x^{\frac{1}{2}} \]

\[ f(h) = h^{\frac{5}{6}} \]

\textsuperscript{33}Empirical evidence suggests that travel (and commuting) tend to be increasing in income. This is particularly true for car travel (Hu and Ruscher (2004), Table 32). A modal split such that high income households travel more by car than low income households is, thus, not unlikely. Moreover, the UK Department for Transport reports a value of time for a commuting trip by car, on average, which is about one third of that of a commuting trip by public transport (DfT (2011), Table 9).
and we assume that $\pi_i = \delta^i = \frac{1}{2}$, $i = 1, 2$. We also use the following function for time of car trips: $a_D = a + 0.00015D$. We are going to describe three scenarios, each characterized by different relative qualities (measured in terms of trip time costs) of cars and public transport. They are obtained by varying the intercept $a$ for car trips, as well as the time cost of public transport trips $a_B$.\footnote{In Scenario 1, we have $a = 0.005$ and $a_B = 0.01$, in Scenario 2 $a = 0.003$ and $a_B = 0.015$. Finally, $a = 0.001$ and $a_B = 0.02$ in Scenario 3.} Recall that individuals’ endowment of time is normalized to one. We fix the monetary cost of a car trip to one and set $c_B = 0.1$ for a public transport trip. In each scenario, $w_2$ is set at 100 and we vary $w_1$ from 50 to 90. This produces differences in earned income, at the second-best allocation, that go from the low ability type earning about 25\% to about 75\% of the (pre-tax) income of the other type. For each scenario, we report individuals’ earned income $I^i$, their amount of travel on each mode and the optimal per-trip tariffs (all computed at the second-best allocation).

Concerning implementability, we refer to Condition $I$ as implementability of the second-best allocation using separable transport tariffs and income taxes, making use, possibly, of a joint payment schedule for cars and public transport. Condition $II$ identifies instead implementability using fully separable transport tariffs (i.e. separate payment schemes for cars and public transport), on top of a separate income tax schedule. For each scenario, we verify whether such conditions hold. Results suggest that implementability can be achieved in many circumstances, even when using three separate payment schedules for cars, public transport and income.

**Scenario 1.** In the first scenario, public transportation is a good alternative to cars. Good enough, in fact, to have both high and low income individuals make it their main commuting mode. This scenario may represent cities in which public transportation is very effective and the primary commuting mode for most of the population. Fitting examples might be European cities like Zurich and Stockholm. We can see that trip quantities are increasing with income, as individuals supply more labor and need to commute increasingly often. Note,
particular, that as her productivity increases, the low ability type works and commutes more, though always less than the high income type. Due to low road congestion, the pigouvian tax $\tau_D$ on car trips is quite small (about 5% of the monetary cost of a car trip). The per-trip tariff $t^D_D$ is strictly higher than that (though by a small amount), while $t^1_D$ is smaller. Low ability types pay the smaller per-trip tariff also on public transport. This is because, at the margin, the cost (in terms of lower daily productivity) of reducing the amount of commuting is larger than the time cost of a journey, on both modes (see the expression for the term $z$ in optimal tariffs of Proposition 1). However, the difference between the marginal tariff intended for high and low types is larger for cars than for public transport.\footnote{The situation in which $w_1 = 50$ is an exception only because $t^1_D$ is constrained to be nonnegative.} This is due to the fact that public transport has higher time costs. Finally, considering implementability of the second-best allocation, the sufficient condition of Proposition 5 fails. Nonetheless, implementability is achievable using fully separable payment functions (i.e. Condition I and II hold), in all cases considered.

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>$a_e=0.005+0.0015D$</th>
<th>$a_0=0.01$</th>
<th>Implementability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W=50$</td>
<td>$l_0=0.005$</td>
<td>$t_0=0.048$</td>
<td>OK</td>
</tr>
<tr>
<td></td>
<td>$D_0=3.650$</td>
<td>$B_0=1.844$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$l_1=3.386$</td>
<td>$D_1=11.886$</td>
<td>$t^2_1=0.048$</td>
</tr>
<tr>
<td></td>
<td>$B_1=0.000$</td>
<td>$t^1_1=0.000$</td>
<td>$t^2_1=0.100$</td>
</tr>
<tr>
<td></td>
<td>$t^3_1=0.031$</td>
<td>$t^4_1=0.031$</td>
<td></td>
</tr>
<tr>
<td>$W=66$</td>
<td>$l_0=0.005$</td>
<td>$t_0=0.053$</td>
<td>OK</td>
</tr>
<tr>
<td></td>
<td>$D_0=3.650$</td>
<td>$B_0=1.844$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$l_1=3.386$</td>
<td>$D_1=11.886$</td>
<td>$t^2_1=0.048$</td>
</tr>
<tr>
<td></td>
<td>$B_1=0.000$</td>
<td>$t^1_1=0.000$</td>
<td>$t^2_1=0.100$</td>
</tr>
<tr>
<td></td>
<td>$t^3_1=0.053$</td>
<td>$t^4_1=0.053$</td>
<td></td>
</tr>
<tr>
<td>$W=75$</td>
<td>$l_0=0.005$</td>
<td>$t_0=0.057$</td>
<td>OK</td>
</tr>
<tr>
<td></td>
<td>$D_0=3.650$</td>
<td>$B_0=1.844$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$l_1=3.386$</td>
<td>$D_1=11.886$</td>
<td>$t^2_1=0.048$</td>
</tr>
<tr>
<td></td>
<td>$B_1=0.000$</td>
<td>$t^1_1=0.000$</td>
<td>$t^2_1=0.100$</td>
</tr>
<tr>
<td></td>
<td>$t^3_1=0.057$</td>
<td>$t^4_1=0.057$</td>
<td></td>
</tr>
<tr>
<td>$W=90$</td>
<td>$l_0=0.005$</td>
<td>$t_0=0.063$</td>
<td>OK</td>
</tr>
<tr>
<td></td>
<td>$D_0=3.650$</td>
<td>$B_0=1.844$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$l_1=3.386$</td>
<td>$D_1=11.886$</td>
<td>$t^2_1=0.048$</td>
</tr>
<tr>
<td></td>
<td>$B_1=0.000$</td>
<td>$t^1_1=0.000$</td>
<td>$t^2_1=0.100$</td>
</tr>
<tr>
<td></td>
<td>$t^3_1=0.063$</td>
<td>$t^4_1=0.063$</td>
<td></td>
</tr>
<tr>
<td>$W=99$</td>
<td>$l_0=0.005$</td>
<td>$t_0=0.068$</td>
<td>OK</td>
</tr>
<tr>
<td></td>
<td>$D_0=3.650$</td>
<td>$B_0=1.844$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$l_1=3.386$</td>
<td>$D_1=11.886$</td>
<td>$t^2_1=0.048$</td>
</tr>
<tr>
<td></td>
<td>$B_1=0.000$</td>
<td>$t^1_1=0.000$</td>
<td>$t^2_1=0.100$</td>
</tr>
<tr>
<td></td>
<td>$t^3_1=0.068$</td>
<td>$t^4_1=0.068$</td>
<td></td>
</tr>
</tbody>
</table>

**Scenario 2.** Compared to Scenario 1, we consider here a situation in which the car, though more expensive, is significantly more attractive than public transportation. As a consequence, modal split is such that public transport is popular only among low income individuals, while the others mostly travel by car (except in the case in which $w_1 = 90$ and earning abilities are very similar). The reason is that low income types work less than the others *in equilibrium* (this is optimal given their lower productivity), are less time constrained and can better cope
with a more time-consuming (but cheaper) travel mode. The higher volume of car trips implies the pigouvian tax $\tau_D$ is at about 4 times higher than in Scenario 1. Once again, optimal per-trip tariffs are smaller when intended for low than for high ability types, with the difference being larger for cars than for public transport. Implementability in separable functions is achievable in all the cases presented. The sufficient condition of Proposition 5 holds, except in case $w_1 = 50$. In that case, however, it is impossible to implement the second-best allocation with separate tariff schedules for cars and public transport (as long as they do not depend on income). Implementation is feasible, instead, if a joint transport tariff scheme (independent of income) is used.

<table>
<thead>
<tr>
<th>Scenario 2</th>
<th></th>
<th>Implementability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_i=50$</td>
<td>$a_{i,0.003+0.0015D}$</td>
<td>$a_{i,0.015}$</td>
</tr>
<tr>
<td>19.99</td>
<td>96.96</td>
<td>0.443</td>
</tr>
<tr>
<td>$W_i=66$</td>
<td>$a_{i,0.004}$</td>
<td>$\tau_c, 0.184$</td>
</tr>
<tr>
<td>41.35</td>
<td>92.67</td>
<td>1.444</td>
</tr>
<tr>
<td>$W_i=75$</td>
<td>$a_{i,0.005}$</td>
<td>$\tau_c, 0.188$</td>
</tr>
<tr>
<td>53.81</td>
<td>90.80</td>
<td>2.607</td>
</tr>
<tr>
<td>$W_i=90$</td>
<td>$a_{i,0.005}$</td>
<td>$\tau_c, 0.231$</td>
</tr>
<tr>
<td>73.93</td>
<td>87.62</td>
<td>5.069</td>
</tr>
</tbody>
</table>

**Scenario 3.** In this scenario, public transport travel is significantly more time consuming than car travel (time cost being more than five times that of a car trip). Cars are thus the preferred mode by both high and low income households, except in the case in which low income ones earn (and work) much less than the others. This scenario seems consistent with the situation of many car-dependent cities. Fitting examples may be American ones such as Atlanta or Los Angeles. Note, however, that low income types commute to a much smaller extent than their high income counterparts. Optimal tariffs follow similar patterns as in Scenario 2, except that the pigouvian tax for cars is larger, given stronger road congestion. As in Scenario 2, the sufficient condition of Proposition 5 holds in all cases presented, except case $w_1 = 50$. Implementability of the second-best allocation is feasible using separate tariffs and income taxes. This is true except when $w_1 = 50$. In that case, a joint tariff schedule for
both transport modes (separate from the income tax schedule) is necessary.

7 Concluding remarks

Our findings suggest that transport pricing can, if properly designed, be used to improve the redistributive capabilities of the tax system. The same is true for provision of infrastructure. As we have seen, this is because individuals of different earning ability have, at the same allocation, different values of time. Another reason is that commuting trips are more valuable to low than to high ability individuals mimicking them. Our results rest, anyway, on some important assumptions. We have assumed that the income tax is optimally designed, which may not always be the case in reality. Yet, we have no reason to believe that the results would not stand even the income tax schedule is suboptimal, as long as it can be flexibly adjusted to account for changes in transportation policy (as in, e.g., Kaplow (2006)). To continue, the focus of our work is on urban transport networks. The government is assumed to be able to control the features of the transport network as well as the income tax schedule. We therefore neglect the presence of multiple government levels (e.g. local and national ones), which may have different powers as well as divergent objectives.\footnote{36 We could consider the presence of an additional part of the population living outside the urban area. Assuming these people do not use its transport network (so that they do not care for $D$ and $B$), fixed residential location, and that tax schedules are flexible enough to be differentiated between people belonging to a given urban agglomeration and those who do not, our results would not change. They would also not change with multiple urban areas and, again, income tax schedules may be differentiated across them.} A more detailed modeling of the interactions between multiple governments looks to be an interesting path to extend.
the study for the future.
Appendix

Proof of Proposition 1

The first order conditions of this problem are

\[
\frac{\partial L}{\partial C_1} = \delta_1 U_1^1 - \pi_1 \mu - \lambda U_1^{21} = 0 \tag{7}
\]

\[
\frac{\partial L}{\partial C_2} = \delta_2 U_2^2 - \pi_2 \mu + \lambda U_2^2 = 0 \tag{8}
\]

\[
\frac{\partial L}{\partial D_1} = (\delta_1 U_1^1 - \lambda U_1^{21}) - \pi_1 \mu - \lambda \pi_1 \frac{\partial \varphi}{\partial D_1} (D_1 \phi_x^{21} - D_2 \phi_x^2) - \pi_1 \sum_{i=1,2} \delta_i D_i \frac{\partial \varphi}{\partial D} \phi_x^i = 0 \tag{9}
\]

\[
\frac{\partial L}{\partial D_2} = U_2^2 (\delta^2 + \lambda) - \pi_2 \mu - \lambda \pi_2 \frac{\partial \varphi}{\partial D} (D_1 \phi_x^{21} - D_2 \phi_x^2) - \pi_2 \sum_{i=1,2} \delta_i D_i \frac{\partial \varphi}{\partial D} \phi_x^i = 0 \tag{10}
\]

\[
\frac{\partial L}{\partial B_1} = \delta_1 U_1^1 - \lambda U_1^{21} - c_B \pi_1 \mu = 0 \tag{11}
\]

\[
\frac{\partial L}{\partial B_2} = U_2^2 (\delta^2 + \lambda) - c_B \pi_2 \mu = 0 \tag{12}
\]

\[
\frac{\partial L}{\partial I_1} = \delta_1 U_1^1 - \lambda U_1^{21} + \pi_1 \mu = 0 \tag{13}
\]

\[
\frac{\partial L}{\partial I_2} = U_2^2 (\delta^2 + \lambda) + \pi_2 \mu = 0 \tag{14}
\]

where subscripts denote partial derivatives, \( U_x^i \equiv \phi_x^i \) is the marginal utility of pure leisure and \( U_j^i \equiv \gamma_j^i - (a_j + m^i) \phi_x^j \). \( j = D, B \) denotes the marginal utility individual \( i = 1, 2 \) derives from a commuting trip \( j = D, B \). This is net of the opportunity cost of trip time, as well as the induced adjustment in labor supply \( m \), at a given income and goods bundle. Take (7), (9) and (11) and rearrange to get to

\[
\frac{U_j^1}{U_C^1} = \frac{\nu_j + \lambda U_j^{21}}{\nu \pi_1} + \frac{\partial \varphi}{\partial D} \frac{\partial \varphi}{\partial Q} \sum_{i=1,2} \delta_i D_i \phi_x^i + \frac{\lambda}{\nu} \frac{\partial \varphi}{\partial D} \frac{\partial \varphi}{\partial Q} (D_1 \phi_x^{21} - D_2 \phi_x^2) \left( 1 + \frac{\lambda U_j^{21}}{\nu \pi_1} \right) \quad j = D, B
\]
where \( \frac{\partial D}{\partial Q^j} = 1 \) if \( j = D \) and 0 otherwise, as public transport trips do not contribute to road congestion. Note that \( v_j = 1 \) if \( j = D \) and \( v_j = c_B \) if \( j = B \). Multiplying both sides by \( 1 + \frac{\mu^2 D}{\pi_1} \) and rearranging we get

\[
\frac{U^1_j}{U^1_C} = v_j + \frac{\lambda U^2_B}{\mu \pi_1} \left( \frac{U^2_B}{U^2_C} - \frac{U^1_B}{U^1_C} \right) + \frac{\partial D}{\partial D} \frac{\partial D}{\partial Q^j} \sum_{i=1,2} \delta^i D_i \phi_x^j \frac{\lambda}{\mu} \frac{\partial D}{\partial D} \frac{\partial D}{\partial Q^j} \left( D^1 \phi_x^{21} - D^2 \phi_x^2 \right) \quad j = D, B
\]

Similarly, using (8), (10) and (12) we get

\[
\frac{U^2_j}{U^2_C} = v_j + \frac{\partial D}{\partial D} \frac{\partial D}{\partial Q^j} \sum_{i=1,2} \delta^i D_i \phi_x^j \frac{\lambda}{\mu} \frac{\partial D}{\partial D} \frac{\partial D}{\partial Q^j} \left( D^1 \phi_x^{21} - D^2 \phi_x^2 \right) \quad j = D, B
\]

In the optimal allocation, we must have

\[
\frac{U^i_D}{U^i_C} = 1 + t^i_D \quad \text{and} \quad \frac{U^i_B}{U^i_C} = t^i_B \quad i = 1, 2
\]

Using these relations, we can obtain the marginal tariff rates \( t^i_j \) provided in the Proposition.

We now focus on \( j = D \) and derive \( \tau_D \) and \( \eta_D \). Rewrite

\[
\frac{\partial \varphi^D}{\partial D} \frac{\partial D}{\partial D} \sum_{i=1,2} \delta^i D_i \phi_x^j \frac{\lambda}{\mu} \frac{\partial D}{\partial D} \frac{\partial D}{\partial Q^j} \left( \frac{\delta^1 D_1 U^1_C}{\mu U^1_C} + \frac{\delta^2 D_2 U^2_C}{\mu U^2_C} \right) \pm \left( \frac{\lambda D_2 U^2_x}{\mu} - \frac{\lambda D_1 U^2_{x1}}{\mu} \right)
\]

now using (7) we have

\[
\frac{\partial \varphi^D}{\partial D} \frac{D_1 U^1_C \delta^1 U^1_C}{\mu U^1_C} = \frac{\partial \varphi^D}{\partial D} \frac{D_1 U^1_C \pi_1}{\mu U^1_C} + \frac{\partial \varphi^D}{\partial D} \frac{D_1 U^1_C \lambda U^2_{x1}}{\mu U^1_C}
\]

and using (8) we have

\[
\frac{\partial \varphi^D}{\partial D} \frac{D_2 U^2_C \delta^2 U^2_C}{\mu U^2_C} = \frac{\partial \varphi^D}{\partial D} \frac{D_2 U^2_C \pi_2}{\mu U^2_C} - \frac{\partial \varphi^D}{\partial D} \frac{D_2 U^2_C \lambda U^2_x}{\mu U^2_C}
\]
so that we can rewrite

\[
\frac{\partial \phi^D}{\partial D} \sum_{i=1,2} \delta^i D_i U^i = \frac{\partial \phi^D}{\partial D} \left( \sum_{i=1,2} \pi_i D_i \frac{U^i}{U_C} + \frac{\lambda D_1 U^1 U^2_2}{\mu U^1_C} - \frac{\lambda D_1 U^2_2}{\mu} + \frac{\lambda}{\mu} (D_1 U^2_2 - D_2 U^1_2) \right)
\]

finally, replacing the above expression in (15) for \( j = D \) and rearranging we have

\[
\frac{U^1_D}{U^1_C} = 1 - \frac{\lambda U^2_1}{\mu \pi_1} \frac{U^1_2}{U^2_1} - \frac{\partial \phi^D}{\partial D} \sum_{i=1,2} \pi_i D_i \frac{U^i}{U_C} + \frac{\partial \phi^D}{\partial D} \frac{\lambda}{\mu} U^2_1 D_1 \left( \frac{U^1_x}{U^1_C} - \frac{U^2_x}{U^2_C} \right)
\]

and

\[
\frac{U^2_D}{U^2_C} = 1 + \frac{\partial \phi^D}{\partial D} \sum_{i=1,2} \pi_i D_i \frac{U^i}{U_C} + \frac{\partial \phi^D}{\partial D} \frac{\lambda}{\mu} U^2_1 D_1 \left( \frac{U^1_x}{U^1_C} - \frac{U^2_x}{U^2_C} \right)
\]

where the terms \( \tau_D \) and \( \eta_D \) as described in the text can be recognized (note that \( U^i_x = \phi^i_x \) \( U^i_C = \Omega^i_C \)). We now focus on \( z_{j,j} = D, B \). We can write

\[
\lambda \frac{U^2_1}{\mu \pi_1} \left( \frac{U^2_2}{U^2_C} - \frac{U^1_2}{U^1_C} \right) = \frac{\lambda}{\mu \pi_1} \left( \left( \gamma^1_j - a_j \phi^1x - m^1 \phi^2x \right) \Omega^1_j - \left( \gamma^1_j - a_j \phi^1x - m^1 \phi^1x \right) \Omega^2_j \right) \quad j = D, B
\]

the right hand side can also be written as

\[
\frac{\lambda}{\mu \pi_1} \left( \Omega^2_1 \left( \frac{\gamma^2_1}{\Omega^1_2} - \gamma^1_1 \right) + a_j \left( \phi_x^1 \Omega^2_1 - \phi_x^2 \right) + \left( \phi_x^1 \Omega^2_1 m_1 - \phi_x^2 m_2 \right) \right) \quad j = D, B
\]

Since \( U(.) \) is such that \( \frac{\gamma^1_1}{\pi^1_C} = \frac{\gamma^2_1}{\pi^2_C} \) (by separability), the expression above becomes

\[
\frac{\lambda \Omega^2_1}{\mu \pi_1} \left( a_j \left( \frac{\phi_x^1}{\Omega^1_2} - \phi_x^2 \right) + \left( \phi_x^1 \frac{m_1}{\Omega^1_2} - \phi_x^2 m_2 \right) \right)
\]

which is \( z_{j,j} = D, B \) in the text.
Optimal income tax rates

Using (13) and (7) we obtain

\[ \frac{U^1_I}{U^1_C} = -1 + \frac{\lambda}{\mu \pi_1} \left( \frac{U^{21}_I}{U^{21}_C} - \frac{U^1_I}{U^1_C} \right) \]

now, using the fact that

\[ U^1_I = -g' \left( \frac{I_1}{w_1 (D_1 + B_1)} \right) \cdot \frac{\phi^1_x}{w_1} U^{21}_I = -g' \left( \frac{I_1}{w_2 (D_1 + B_1)} \right) \cdot \frac{\phi^{21}_x}{w_2} \]

we have

\[ t^I_I = 1 + \frac{U^1_I}{U^1_C} = \frac{\lambda}{\mu \pi_1} \Omega^{21}_C \left( g' \left( \frac{I_1}{w_1 (D_1 + B_1)} \right) \cdot \frac{\phi^1_x}{w_1} - g' \left( \frac{I_1}{w_2 (D_1 + B_1)} \right) \cdot \frac{\phi^{21}_x}{w_2} \right) > 0 \]

while, using (14) and (8), we have \( t^2_I = 1 + \frac{U^2_I}{U^2_C} = 0 \).

Proof of Proposition 2

We solve this problem assuming that the government can directly determine the level of congestion (public bad), denoted \( \bar{D} \). When solving the problem, we have thus an additional equality constraint given by \( \bar{D} = \sum_{i=1,2} \pi_i D^i \). We denote by \( \beta \) the Lagrange multiplier for this constraint. Thus, the Lagrangian is

\[ \mathcal{L} = W + \mu \left( \sum_{i=1,2} \pi_i (I^i - y^i + t_D D^i + (t_B - c_B) B^i) - \sum_{j=D,B} c_{K_j} K_j - R \right) + \lambda (V^2 - V^{21}) + \beta \left( \bar{D} - \sum_{i=1,2} \pi_i D^i \right) \]
The first order condition of this problem are

\[
\frac{\partial L}{\partial q_j} = \delta^1 \frac{\partial V^1}{\partial q_j} + (\delta^2 + \lambda) \frac{\partial V^2}{\partial q_j} - \lambda \frac{\partial V^{21}}{\partial q_j} + \mu \left[ \sum_{i=1,2} \pi_i \left( Q^1_j + (q_D - 1) \frac{\partial D^i}{\partial q_j} + (q_B - c_B) \frac{\partial B^i}{\partial q_j} \right) \right] + 
- \beta \sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial q_j} = 0 \quad j = D, B
\]

\[
\frac{\partial L}{\partial y^i} = \delta^1 \frac{\partial V^1}{\partial y^i} + (\delta^2 + \lambda) \frac{\partial V^2}{\partial y^i} - \lambda \frac{\partial V^{21}}{\partial y^i} + \mu \pi_i \left[ -1 - (q_D - 1) \frac{\partial D^i}{\partial y^i} + (q_B - c_B) \frac{\partial B^i}{\partial y^i} \right] - \beta \pi_i \frac{\partial D^i}{\partial y^i} = 0 \quad i = 1, 2
\]

\[
\frac{\partial L}{\partial I^i} = \delta^1 \frac{\partial V^1}{\partial I^i} + (\delta^2 + \lambda) \frac{\partial V^2}{\partial I^i} - \lambda \frac{\partial V^{21}}{\partial I^i} + \mu \pi_i \left[ 1 - (q_D - 1) \frac{\partial D^i}{\partial I^i} + (q_B - c_B) \frac{\partial B^i}{\partial I^i} \right] - \beta \pi_i \frac{\partial D^i}{\partial I^i} = 0 \quad i = 1, 2
\]

\[
\frac{\partial L}{\partial \bar{D}} = \frac{\partial \phi^D}{\partial D} \left( \delta^1 \frac{\partial V^1}{\partial a_D} + (\delta^2 + \lambda) \frac{\partial V^2}{\partial a_D} - \lambda \frac{\partial V^{21}}{\partial a_D} + \mu \sum_{i=1,2} \pi_i \left[ (q_D - 1) \frac{\partial D^i}{\partial a_D} + (q_B - c_B) \frac{\partial B^i}{\partial a_D} \right] \right) + 
+ \beta \left( 1 - \sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial a_D} \frac{\partial \phi^D}{\partial D} \right) = 0
\]

\[\text{note that}\]

\[
\frac{\partial V^i}{\partial a_D} = -\phi^i_D \quad i = 1, 2 \quad \frac{\partial V^{21}}{\partial a_D} = -\phi^{21}_D
\]

To start, we are going to focus on \(\frac{\partial L}{\partial D}\). Add \(\lambda \frac{\partial V^{21}}{\partial y^i} \left( \frac{\partial V^1}{\partial a_D} / \frac{\partial y^i}{\partial y^1} \right) \frac{\partial \phi^D}{\partial D}\) to both sides and rearrange to get

\[
\frac{\partial \phi^D}{\partial D} \left( \left( \delta^1 \frac{\partial V^1}{\partial y^i} - \lambda \frac{\partial V^{21}}{\partial y^i} \right) \left( \frac{\partial V^1}{\partial a_D} / \frac{\partial y^i}{\partial y^1} \right) + \lambda \frac{\partial V^{21}}{\partial y^i} \left( \frac{\partial V^1}{\partial a_D} / \frac{\partial y^i}{\partial y^1} - \frac{\partial V^1}{\partial a_D} / \frac{\partial y^2}{\partial y^1} \right) \right) + 
+ \frac{\partial \phi^D}{\partial D} \left( (\delta^2 + \lambda) \frac{\partial V^2}{\partial y^2} \left( \frac{\partial V^2}{\partial a_D} / \frac{\partial y^2}{\partial y^2} \right) + \left( \mu \sum_{i=1,2} \pi_i \left[ (q_D - 1) \frac{\partial D^i}{\partial a_D} + (q_B - c_B) \frac{\partial B^i}{\partial a_D} \right] \right) \right) + 
+ \beta \left( 1 - \sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial a_D} \frac{\partial \phi^D}{\partial D} \right) = 0
\]
now substituting $\delta^1 \frac{\partial V^1}{\partial y^1} - \lambda \frac{\partial V^{21}}{\partial y^1}$ and $(\delta^2 + \lambda) \frac{\partial V^2}{\partial y^2}$ from the first order conditions for $\frac{\partial \xi}{\partial y^i}$ above, we obtain, after some rearrangements

$$\frac{\partial \psi_D}{\partial D} \left( \mu \left( \sum_{i=1,2} \pi_i \frac{\partial V^i}{\partial a_D} \frac{\partial V^i}{\partial y^i} \right) + \lambda \frac{\partial V^{21}}{\partial y^1} \left( \frac{\partial V^1}{\partial a_D} \frac{\partial V^1}{\partial y^1} - \frac{\partial V^{21}}{\partial a_D} \frac{\partial V^{21}}{\partial y^1} \right) \right) +$$

$$+ \frac{\partial \psi_D}{\partial D} \left[ \mu \sum_{i=1,2} \pi_i (q_D - 1) \left( \frac{\partial D^i}{\partial a_D} - \frac{\partial D^i}{\partial y^i} \left( \frac{\partial V^i}{\partial a_D} \frac{\partial V^i}{\partial y^i} \right) \right) \right] +$$

$$+ \frac{\partial \psi_D}{\partial D} \left[ \mu \sum_{i=1,2} \pi_i (q_B - c_B) \left( \frac{\partial B^i}{\partial a_D} - \frac{\partial B^i}{\partial y^i} \left( \frac{\partial V^i}{\partial a_D} \frac{\partial V^i}{\partial y^i} \right) \right) \right] +$$

$$+ \beta \frac{\partial \psi_D}{\partial D} \left( \sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial y^i} \left( \frac{\partial V^i}{\partial a_D} \frac{\partial V^i}{\partial y^i} \right) \right) + \beta \left( 1 - \sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial a_D} \frac{\partial \psi_D}{\partial D} \right) = 0$$

To simplify further, we need to use the following Slutsky-type property obtained by Pirttilä and Tuomala (1997)

$$-\pi_i \frac{\partial D^i}{\partial y^i} \left( \frac{\partial V^i}{\partial a_D} \frac{\partial V^i}{\partial y^i} \right) \frac{\partial \psi_D}{\partial D} = \pi_i \left( \frac{\partial \tilde{D}^i}{\partial a_D} - \frac{\partial \tilde{D}^i}{\partial y^i} \right) \frac{\partial \psi_D}{\partial D} \quad i = 1, 2$$

where a tilde denotes hicksian demands. Using these properties, the condition above rewrites as

$$\beta = -\chi \left( \mu \left( \sum_{i=1,2} \pi_i \frac{\partial V^i}{\partial a_D} \frac{\partial V^i}{\partial y^i} \right) + \lambda \frac{\partial V^{21}}{\partial y^1} \left( \frac{\partial V^1}{\partial a_D} \frac{\partial V^1}{\partial y^1} - \frac{\partial V^{21}}{\partial a_D} \frac{\partial V^{21}}{\partial y^1} \right) \right) \frac{\partial \psi_D}{\partial D} +$$

$$-\chi \left( \mu \sum_{i=1,2} \pi_i \left[ (q_D - 1) \frac{\partial \tilde{D}^i}{\partial a_D} + (q_B - c_B) \frac{\partial \tilde{B}^i}{\partial a_D} \right] \right) \frac{\partial \psi_D}{\partial D}$$

where $\chi = \frac{1}{1 - \frac{\partial \psi_D}{\partial D} \sum_{i=1,2} \pi_i \frac{\partial D^i}{\partial a_D}}$. Let us now proceed by multiplying $\frac{\partial \xi}{\partial y^i}$ by $D^i$ for $i = 1, 2$ and adding the resulting expressions to $\frac{\partial \xi}{\partial y^i}$. Then multiply $\frac{\partial \xi}{\partial y^i}$ by $B^i$ for $i = 1, 2$ and add the resulting expressions to $\frac{\partial \xi}{\partial y^i}$. The equations obtained as a result can be simplified making use of Roy’s identity and using the Slutsky equations $\frac{\partial Q^i_j}{\partial q_j} = \frac{\partial Q^i_j}{\partial q_j} - Q_j \frac{\partial V^i_j}{\partial y^i}$, where
\( Q_D^i = D^i \quad Q_B^i = B^i \) and where a tilde denotes hicksian demands. As a result, we obtain

\[
\mu \sum_{i=1,2} \pi_i \left( (q_D - \frac{\beta}{\mu} - 1) \frac{\partial \tilde{D}^i}{\partial q_j} + (q_B - c_B) \frac{\partial \tilde{B}^i}{\partial q_j} \right) = \lambda \frac{\partial V_{21}^i}{\partial y_1^i} (Q_j^i - Q_{j2}^i) \quad j = D, B
\]

Finally, one needs to replace for \( \beta \) as obtained above and rearrange to obtain, from the last two expressions above, the optimal tariffs as expressed in the proposition.

**Proof of Proposition 3**

If the government decides now also on the amount of infrastructure \( K_j \quad j = D, B \), the additional first order condition in the its maximization problem are

\[
\frac{\partial L}{\partial K_j} = -\delta^1 \phi_x^1 \frac{\partial \phi_j^1}{\partial K_j} Q_j^1 - \delta^2 \phi_x^2 \frac{\partial \phi_j^2}{\partial K_j} Q_j^2 - \mu c_{K_j} + \lambda \phi_x^{21} \frac{\partial \phi_j^{21}}{\partial K_j} Q_j^1 - \lambda \phi_x^{21} \frac{\partial \phi_j^{21}}{\partial K_j} Q_j^2 = 0 \quad j = D, B \tag{16}
\]

the other first order conditions as presented in the proof of Proposition 1. We have denoted \( Q_D^i = D^i, Q_B^i = B^i \). Starting from (16), adding and subtracting \( \lambda \frac{\phi_j^{21}}{U_C^1} Q_j \frac{\partial \phi_j^{21}}{\partial U_C^2} \) to the left hand side and using (7) and (8) we get, after some easy rearrangements

\[
\left( -\frac{\partial \phi}{\partial K_j} \sum_{i=1,2} \pi_i \frac{Q_j^i \phi_j^i}{\Omega_C^i} \right) - \frac{\lambda}{\mu} \frac{\partial \phi}{\partial K_j} Q_j^1 \frac{\phi_j^1}{U_C^1} \frac{\partial \phi_j^{21}}{\partial U_C^2} = c_{K_j} \quad j = D, B
\]

the second term on the left hand side being \( \gamma_j \) as presented in the text.

**Proof of Proposition 4**

If the government decides now also on the amount of infrastructure \( K_j \quad j = D, B \), the additional first order condition in the maximization problem are

\[
\frac{\partial L}{\partial K_j} = \frac{\partial \phi_j^i}{\partial K_j} \left( \delta^1 \frac{\partial V_1^i}{\partial a_j} + (\delta^2 + \lambda) \frac{\partial V_2^i}{\partial a_j} - \lambda \frac{\partial V_{21}^i}{\partial a_j} + \mu \sum_{i=1,2} \pi_i \left[ (q_D - 1) \frac{\partial D^i}{\partial a_j} + (q_B - c_B) \frac{\partial B^i}{\partial a_j} \right] \right) +
\]

\[
-\beta \left( \sum_{i=1,2} \pi_i \frac{\partial D^i \partial \phi_j^i}{\partial a_j \partial K_j} \right) = \mu c_{K_j} \quad j = D, B
\]
One needs to follow similar steps as the ones used to rewrite $\frac{\partial C}{\partial D}$ in the proof of Proposition 2 to get

$$\left( \mu \left( \sum_{i=1,2} \pi_i \frac{\partial V_i}{\partial a_j} \right) - \frac{\partial V}{\partial y} \right) \frac{\partial \varphi_j}{\partial K_j} + \left( \mu \sum_{i=1,2} \pi_i \left( (q_D - 1) \frac{\partial \tilde{D}_i}{\partial a_j} + (q_B - c_B) \frac{\partial \tilde{B}_j}{\partial a_j} \right) \right) \frac{\partial \varphi_j}{\partial K_j} + \beta \sum_{i=1,2} \pi_i \frac{\partial \tilde{D}_i}{\partial a_j} \frac{\partial \varphi_j}{\partial K_j} = \mu c K_j, \quad j = D, B$$

replacing for $\beta$ from the proof of proposition 2 and using the definition of $\varepsilon$ given in the text, we obtain the expressions provided in Proposition 4. Replacing the value of $\varepsilon$ in the expression for $j = D$, one can also easily obtain the third condition provided in the text.

**Proof of Proposition 5**

We proceed assuming the following conditions hold at $A^{SB}$: $\frac{I_2}{w^2} > \frac{I_1}{w^2}$, $I^2 - T^2 > I^1 - T^1$, $I^2 > I^1$.  

Payments $P^1$ and $P^2$ are defined as the payments such that

$$\Omega \left( I^2 - T^2 - P^2 \right) + \gamma \left( D^2, B^2 \right) + \phi \left( 1 - a_D D^2 - a_B B^2 - (D^2 + B^2) \right) \left( \frac{I^2}{w^2(D^2 + B^2)} \right) =$$

$$\Omega \left( I^2 - T^2 - P^1 \right) + \gamma \left( D^1, B^1 \right) + \phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) \right) \left( \frac{I^2}{w^2(D^1 + B^1)} \right)$$

that is, $P^2 - P^1$ is the extra payment that needs to be asked to a type 2 individual in order to ensure that she (when choosing the level of income $I^2$) will consume trip quantities $D^2 + B^2$ rather than $D^1 + B^1$.  

37Their meaning is the following: with $\frac{I_2}{w^2} > \frac{I_1}{w^2}$, we assume that the amount of labor supplied by the high ability type is larger than that of the low ability type. We also assume that both the pre-tax and the post-tax income of individuals of high ability is higher than that of low ability types.

38Similarly, $P^1$ should be designed as the payment such that trip quantity $D^1 + B^1$ gives the same utility, to an individual of type 1 choosing to earn income $I^1$, as making no trips at all. However choosing no travel at all would always be a dominated alternative, given their commuting purpose (with no commuting, labor supply would be infinite) even if $P^1$ took away all of the individual’s net income. We thus set $P^1$ arbitrarily. This also means that we can be sure that neither individuals of type 1 nor those of type 2 will prefer zero trips to, respectively, $D^1 + B^1$ and $D^2 + B^2$, as long as $P^1$ and $P^2$ are not unreasonably high.
Proof of validity of (5) for $i = 1$ at $A^{SB}$  Rewrite the left hand side of (17) for $i=2$, using (2) (we know this constraint to be satisfied at equality since, by assumption, it constraint binds at $A^{SB}$). We have the following

$$\Omega \left( I^1 - T^1 - P^1 \right) + \gamma \left( D^1, B^1 \right) + \phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^2(D^1 + B^1)} \right) \right) =$$

$$\Omega \left( I^2 - T^2 - P^1 \right) + \gamma \left( D^1, B^1 \right) + \phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^2}{w^2(D^1 + B^1)} \right) \right)$$

therefore

$$\Omega \left( I^1 - T^1 - P^1 \right) = \Omega \left( I^2 - T^2 - P^1 \right) + \phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^2}{w^2(D^1 + B^1)} \right) \right) +$$

$$-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^2(D^1 + B^1)} \right) \right)$$

So constraint (5) for $i = 1$ is verified if (replacing $\Omega \left( I^1 - T^1 - P^1 \right)$ from the above and rearranging)

$$\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^1(D^1 + B^1)} \right) \right) +$$

$$-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^2}{w^1(D^1 + B^1)} \right) \right) \geq$$

$$\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^2(D^1 + B^1)} \right) \right) +$$

$$-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^2}{w^2(D^1 + B^1)} \right) \right)$$

This is verified by convexity of $g(.)$ and concavity of $\phi(.)$. 

41
Proof of validity of (6) for \( i = 1 \) at \( A^{SB} \). Start from (6) for \( i = 2 \). Using (17), we have

\[
\Omega (I^2 - T^2 - P^1) - \Omega (I^2 - T^2 - P^2) - \gamma (D^2, B^2) + \gamma (D^1, B^1) = \\
\phi \left( 1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^2}{w^2 (D^2 + B^2)} \right) \right) + \\
-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^2}{w^2 (D^1 + B^1)} \right) \right)
\]

Now assume that \( D^2 + B^2 \geq D^1 + B^1 \) and \( a_D D^1 + a_B B^1 \geq a_D D^2 + a_B B^2 \). Then, by \( \frac{I^2}{w^2} > \frac{I^1}{w^1} \) and concavity of \( \phi(.) \), we have

\[
\phi \left( 1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^2}{w^2 (D^2 + B^2)} \right) \right) + \\
-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^2}{w^2 (D^1 + B^1)} \right) \right) > \\
\phi \left( 1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^1}{w^1 (D^2 + B^2)} \right) \right) + \\
-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^1 (D^1 + B^1)} \right) \right)
\]

therefore

\[
\Omega (I^2 - T^2 - P^1) + \gamma (D^1, B^1) - \Omega (I^2 - T^2 - P^2) - \gamma (D^2, B^2) > \\
\phi \left( 1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^1}{w^1 (D^2 + B^2)} \right) \right) + \\
-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^1 (D^1 + B^1)} \right) \right)
\]

Now, by concavity of \( \Omega(.) \) and since \( I^2 - T^2 > I^1 - T^1 \), we can write

\[
\Omega (I^1 - T^1 - P^1) - \Omega (I^1 - T^1 - P^2) > \Omega (I^2 - T^2 - P^1) - \Omega (I^2 - T^2 - P^2)
\]
therefore
\[
\Omega(I^1 - T^1 - P^1) + \gamma(D^1, B^1) - \Omega(I^1 - T^1 - P^2) - \gamma(D^2, B^2) > \\
\phi \left( 1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^1}{w^1(D^2 + B^2)} \right) \right) + \\
-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^1(D^1 + B^1)} \right) \right)
\]

which, rearranged, gives us (6) for \( i = 1 \).

**Validity of (5) for \( i = 2 \) at \( A^{SB} \)** Use (2) to rewrite the left hand side of (5) for \( i = 2 \). We can rearrange to get
\[
\Omega(I^1 - T^1 - P^1) + \gamma(D^1, B^1) - \Omega(I^1 - T^1 - P^2) - \gamma(D^2, B^2) \geq \\
\phi \left( 1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^1}{w^2(D^2 + B^2)} \right) \right) + \\
-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^2(D^1 + B^1)} \right) \right)
\]

Now, by concavity of \( \Omega(.) \) and since \( I^2 - T^2 > I^1 - T^1 \), we have
\[
\Omega(I^1 - T^1 - P^1) - \Omega(I^1 - T^1 - P^2) > \Omega(I^2 - T^2 - P^1) - \Omega(I^2 - T^2 - P^2)
\]

(5) is thus certainly satisfied if
\[
\Omega(I^2 - T^2 - P^1) - \Omega(I^2 - T^2 - P^2) - \gamma(D^2, B^2) + \gamma(D^1, B^1) \geq \\
\phi \left( 1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^1}{w^2(D^2 + B^2)} \right) \right) + \\
-\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^2(D^1 + B^1)} \right) \right)
\]
holds. Using (17) we can replace for the left hand side of the above and rearranging we have

\[ \phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^1}{w^2 (D^1 + B^1)} \right) \right) + \]
\[ -\phi \left( 1 - a_D D^1 - a_B B^1 - (D^1 + B^1) g \left( \frac{I^2}{w^2 (D^1 + B^1)} \right) \right) \geq \]
\[ \phi(1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^1}{w^2 (D^2 + B^2)} \right) ) + \]
\[ -\phi(1 - a_D D^2 - a_B B^2 - (D^2 + B^2) g \left( \frac{I^2}{w^2 (D^2 + B^2)} \right) ) \]

On condition that \( D^2 + B^2 \geq D^1 + B^1 \) and \( a_D D^1 + a_B B^1 \geq a_D D^2 + a_B B^2 \), this is verified. Therefore, (5) holds as well.

References


