KU LEUVEN

CENTER FOR ECONOMIC STUDIES

DISCUSSION PAPER SERIES
DPS14.33
DECEMBER 2014

Second-best urban tolling with distributive concerns

Andrea DIAZ and Stef PROOST

Energy, Transport and Environment

Faculty of Economics
and Business

Second-best urban tolling with distributive concerns

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

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

Abstract

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

Keywords: Tax reform, congestion pricing, urban tolls

1. Introduction

Transport economists advocate road pricing as an efficient instrument to regulate the use of road infrastructures. Imposing a road toll that reflects marginal external congestion costs makes consumers use the road up to the

*Corresponding author
Email addresses: Andrea.DiazRincon@kuleuven.be (A. Diaz), Stef.Proost@econ.kuleuven.be (S. Proost)

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

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

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

Our approach is close to that of Parry and Bento (2001) but we add two dimensions to their model. First, instead of a choice between a congested road and uncongested public transit, we model two congested transport op-

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

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

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

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

2. The household's problem: road choice

We start with a simple model–in the spirit of Parry and Bento (2001)– of a representative household whose utility function depends on aggregate consumption of market goods (X, whose price is normalized to one), leisure time (t_L), and the number of days devoted to work (D).

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

The representative household owns a car and uses it to commute to work by taking either one of the two parallel congested roads (of given capacities) that connect residential areas to workplaces. A congestion toll (τ) related to distance (d) is applied on one of the two roads (route T), while the other (route U) remains untolled.

Households choose which route to use to commute to work, U or T. Total number of worked days in a period (D) is the sum of the number of days the household commutes by the untolled road D_U and by the tolled road D_T. Thus, the budget constraint is:

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

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

The left-hand side of (2) corresponds to household expenditures on aggregate consumption and commuting. Each day of work requires a commuting round trip that involves time and monetary costs. When commuting by the untolled road, only fuel costs are relevant[2]. g represents fuel price per kilometer, $g = c_g(1 + \tau_g)$, where c_g is the resource fuel cost (which takes into account vehicle fuel efficiency) and τ_g the fuel tax. Commuting by the tolled

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

road implies paying for the fuel consumption plus the toll. However, this road allows faster trips, while the untolled road requires more time and higher fuel consumption due to a longer distance: $d_U = \beta d_T$ with $\beta > 1$.

Households also face a time constraint:

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

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

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

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

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

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

As a result of considering time as a resource, we get the monetary value of time for each household (U_{t_L}/U_X). This is the ratio between the Lagrange multiplier of the time constraint and the Lagrange multiplier of the income constraint (see Appendix A). The value of spending time in transport[3] (Value of Transport Time, VTT) is represented in (4) and (5) by the value of time foregone by commuting minus the marginal disutility of commuting ($VTT = t_R d_R (U_{t_L}/U_X) - (C_{D_R}/U_X)$, where $R = U, T$).

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

Equating (4) and (5) yields the Wardrop equilibrium condition[4] in which the two roads have equal generalized prices[5]:

$$\tau = g(\beta - 1) + (\beta t_U - t_T)\frac{U_{t_L}}{U_X} + \frac{1}{d_T}\left(\frac{C_{D_T}}{U_X} - \frac{C_{D_U}}{U_X}\right). \qquad (6)$$

This expression indicates that households are indifferent to taking either of the two roads when the toll imposed on road T equals the extra-cost of commuting by road U. That is, the extra-gasoline and the extra travel time costs plus the difference between the marginal disutility of commuting by each road[6].

A household's individual decision depends on its own value of time (U_{t_L}/U_X), which also determines its willingness to pay for a trip. The opportunity cost of time indirectly depends on labor productivity. As high-productivity households will normally get higher wages, they should exhibit higher values of leisure time, whereas low-productivity households exhibit lower values, i.e. $U^h_{t_L}/U_X > U^\ell_{t_L}/U_X$ where h and ℓ indicate highly- and less-productive households, respectively. Thus, a sufficiently high toll should make high-productivity households stay on the tolled road and therefore save high-valued time. In contrast, as low-productivity households have lower budgets, they should be more sensitive to monetary cost and should prefer taking the untolled road in order to save money.

We finally define the differentiable demand functions for each road $D^*_U = D_U(\tau_g, \tau_w, \tau, t_U, w, \varepsilon), D^*_T = D_T(\tau_g, \tau_w, \tau, t_T, w, \varepsilon)$. Assuming that they exist allows us to get the household's indirect utility function $v(\tau_g, \tau_w, \tau, t_T, t_U, w, \varepsilon, G)$ as a function of exogenous parameters (see Appendix A).

3. Use of the congested roads by heterogeneous households

In the presence of congestion, households take into account their own travel cost but not the external cost imposed on other users ($\partial t_T/\partial D_T > 0$,

[4]Wardrop principle: *"For a given origin-destination pair of substitute roads, all used routes should have equal average cost and there should be no unused routes with lower costs"* (Small and Verhoef, 2007).

[5]We assume we can exclude corner solutions where only one of the two roads is used.

[6]If the marginal disutility of commuting is the same by the two roads $C_{D_T}/U_X = C_{D_U}/U_X$, the right hand side of condition (6) is reduced to the extra-gasoline and extra-time costs: $\tau = p_g g(\beta - 1) + (\beta t_U - t_T)U_{t_L}/U_X$.

and $\partial t_U/\partial D_U > 0$). First-best pricing calls for tolling both roads at their marginal external costs. However, we are interested in analyzing the second-best configuration where only one of the two roads can be tolled. In this section we study the user equilibrium. The properties of the user equilibrium will be instrumental in the derivation of the optimal taxes in the next section.

We assume that the economy is populated by n^h highly productive and n^ℓ less productive households (such that $n^h + n^\ell = N$). Both kinds of households independently choose the number of trips they make in a period D_R^i ($i = \ell, h$ and $R = U, T$). As households differ in their willingness to pay for commuting, differentiating them according to the road used may be useful. As a start, we may expect consumers with higher values of time to take road T and consumers with lower values of time to take road U[7].

Let us assume for a moment that there is no specific preference for a road. That is, $C_{D_T^\ell}/U_X = C_{D_U^\ell}/U_X$ and $C_{D_T^h}/U_X = C_{D_U^h}/U_X$. From the right-hand side of equations (4) and (5) we can compare the generalized cost of commuting by each road per type of household:

$$\underbrace{\tau + g + t_T\left(\sum n^i D_T^i\right)\frac{U_{t_L}^h}{U_X}}_{c_T^h} \leq \underbrace{\beta\left(g + t_U\left(\sum n^i D_U^i\right)\frac{U_{t_L}^h}{U_X}\right)}_{c_U^h}, \tag{7}$$

$$\underbrace{\tau + g + t_T\left(\sum n^i D_T^i\right)\frac{U_{t_L}^\ell}{U_X}}_{c_T^\ell} \geq \underbrace{\beta\left(g + t_U\left(\sum n^i D_U^i\right)\frac{U_{t_L}^\ell}{U_X}\right)}_{c_U^\ell}. \tag{8}$$

These conditions compare the total generalized cost of commuting by T (left hand-side) with the cost of commuting by U (right hand-side) for a high-productivity household (7) and for a low-productivity household (8). When a household takes the decision to commute by one of the roads, it already knows the cost of time it will face: total time required multiplied by its own value of time[8]. Time required (per unit of distance) by each road is

[7] As road T attracts those consumers willing to pay more for faster commuting.

[8] It implicitly assumes consumers are informed about current traffic congestion conditions on both roads, by for example, electronic bulletin boards or services such as traffic forecast, and of course, by their own experience.

an increasing function[9] of total traffic volume. $U_{t_L}^h/U_X$ and $U_{t_L}^\ell/U_X$ represent the value of time for high- and low-productivity households, respectively.

From (7) and (8) we establish four different equilibriums of use of the roads by the households, similar to those established in Small and Yan (2001).

3.1. Separating equilibrium

This is the case where high-income households commute only by T ($D_U^h = 0$) and low-income households only by U ($D_T^\ell = 0$). This requires equations (7) and (8) to hold both with inequality (i.e. $c_T^h < c_U^h$ and $c_T^\ell > c_U^\ell$).

3.2. Partially separating equilibrium with low-income groups separated

In this case high-income households commute by T and U, and low-income households commute only by U ($D_T^\ell = 0$). This requires equation (7) to hold with equality and (8) with inequality (i.e. $c_T^h = c_U^h$ and $c_T^\ell > c_U^\ell$).

3.3. Partially separating equilibrium with high-income groups separated

In this case high-income households commute by T ($D_U^h = 0$), and low-income households commute by T and U. This requires equation (7) to hold with inequality and (8) with equality (i.e. $c_T^h < c_U^h$ and $c_T^\ell = c_U^\ell$).

3.4. Pooling equilibrium

In this case both kinds of households commute by T and U. If both equations hold with equality, both types of households will be indifferent towards taking either of the two roads.

4. The social planner's problem: the optimal toll

The government raises revenues to finance public goods F and a head subsidy G, using three tax instruments: fuel taxes (τ_g), tolls (τ) and wage taxes (τ_w). We assume equal labor tax rates for both types of households[10].

[9] $t_R(\sum n^i D_R^i)$ in (7) and (8) denotes the time function for each one of the roads.
[10] This assumption is relaxed in the numerical illustration.

The government maximizes social welfare[11] $\mathbb{W} = \sum_i n^i \theta^i \upsilon^i(\tau_g, \tau_w, \tau, t_U, t_T, w, \varepsilon^i, G)$, subject to the following budget constraint:

$$w\tau_w \sum_i \sum_R n^i \varepsilon^i D_R^i + \tau_g c_g \sum_i \sum_R n^i d_R D_R^i + \tau \sum_i n^i d_T D_T^i = F + NG. \quad (9)$$

Each household chooses the optimal number of commuting trips (D_R^i) that maximizes its individual utility. The budget constraint (9) varies as a function of the use of the roads by the households. Thus, each equilibrium implies a different budget constraint. In what follows, we analyze the optimal toll level for the four possible equilibriums studied in the previous section. For each case, toll revenues are returned to the individuals either through poll transfers or through labor-tax cuts.

4.1. Separating equilibrium

In this equilibrium high-income households take only road T and low-income households take only road U. Although this case might not seem realistic, it is useful as a benchmark that allows comparison with the more complex cases. The government's budget constraint (9) becomes:

$$w\tau_w(n^h\varepsilon^h D_T^h + n^\ell \varepsilon^\ell D_U^\ell) + \tau_g c_g d_T(n^h D_T^h + \beta n^\ell D_U^\ell) + \tau d_T n^h D_T^h = F + NG. \quad (10)$$

By assumption in the separating equilibrium, the h-group continues to use only the tolled road but reduces the number of trips made on this road as the toll increases. On the other hand, the ℓ-group keeps the number of trips on U fixed.

4.1.1. Toll revenues used to finance head transfers

We derive the optimal congestion tax (τ_{pt}, where pt stands for poll transfers) that maximizes social welfare when revenues are returned to households as poll transfers.

$$\tau_{pt} = \frac{1}{1-\phi_{pt}\vartheta\xi} \underbrace{\frac{\partial t_T}{\partial D_T^h}\frac{U_{t_L}^h}{U_X^h} D_T^h}_{mecc} - \frac{1-\phi_{pt}\vartheta}{1-\phi_{pt}\vartheta\xi}\Big(\underbrace{\tau_g c_g + \tau_w \tfrac{w\varepsilon^h}{d_T}}_{other\ taxes\ per\ trip} \Big), \quad (11)$$

[11]This is a purely utilitarian social welfare function where increases or decreases in individual utilities translate into identical changes in social utility. Assumptions on the concavity of the utility function allow for the differentiation of the social marginal value of one unit of income over individuals. Aversion to income inequality is introduced via θ, the social weight given by the government to each kind of household (with $\sum_i \theta^i = 1$).

where $\phi_{pt} = \frac{n^\ell}{N}$, $\vartheta = \left(1 - \frac{U^\ell_X}{U^h_X}\frac{\theta^\ell}{\theta^h}\right)$ and $\xi = 1 + \frac{1}{\epsilon^\tau_{D^h_T}}$, with $\epsilon^\tau_{D^h_T}$ the elasticity of demand of high-income households for the tolled road (see Appendix B.1.1).

The optimal congestion toll has two main components, the marginal external congestion cost (*mecc*) and the *other taxes* levied per trip. The *mecc* measures the increase in traffic-time cost to all road users caused by an extra trip per period. In equation (11), it is represented by the product of: the increase in commute time from an additional trip ($\partial t_T / \partial D^h_T$), the value of time of the commuter ($U^h_{t_L}/U^h_X$), and the number of trips made per period D^h_T. Other taxes per trip appear in (11) as the complementary relationship between work-related trips and the labor market ensures that all taxes (per kilometer) levied per day of work serve to tackle the externality caused by each day of work, namely congestion. Thus, for example, if the sum of the fuel and the labor tax exceeds the *mecc*, rather than taxing road T commuters, the government should subsidize them. Equation (11) therefore suggests an optimal combination of the toll, the fuel tax and the labor tax, rather than a unique optimal toll level.

Each term in (11) is multiplied by a factor that depends on the government distribution concerns (ϑ). The governmental distribution concerns depend on the ratio of the marginal utility of income of both types of consumers (U^ℓ_X/U^h_X), and the relative social weight given to a unit of utility of a poor individual with respect to a rich individual (θ^ℓ/θ^h). Normally, $U^\ell_X/U^h_X > 1$ as the marginal utility of income declines with the level of income. Similarly, $\theta^\ell/\theta^h > 1$ when the decision maker attaches a higher weight to low-productivity consumers.

If there was no difference between the groups ($U^\ell_X = U^h_X$) and the government attached the same weight to both of them ($\theta^\ell = \theta^h$), the toll should equal the difference between the *mecc* and the sum of the *other taxes*[12]. When this is not the case, however, the revenue raising effect implied by the price elasticity of the tolled road ($\epsilon^\tau_{D^h_T}$) plays an important role. If the weight attached by the government to the poor is higher and the demand for the tolled road is inelastic (elastic), the toll should be higher (lower) than the difference between the *mecc* and the sum of the *other taxes*.

[12]Therefore (11) contains Parry and Bento's (2001) result ($\tau = mecc$) as a special case: homogeneous consumers, no redistribution concerns and no other taxes levied per trip.

This suggests that a greater concern for the welfare of the poor leads to the use of the toll as an instrument to redistribute income when the demand for the tolled road is inelastic (in this equilibrium road T is used only by the rich). However, when demand is elastic the use of the toll as an instrument to redistribute income is limited, as every euro of tax revenues then has a high efficiency cost.

4.1.2. Toll revenues used to cut labor taxes

Following the same procedure, we obtain the optimal toll (τ_{lt}, where lt stands for labor tax cuts) when the incremental toll revenues are used to cut labor tax rates:

$$\tau_{lt} = \frac{1}{1-\phi_{lt}\vartheta\xi}\frac{\partial t_T}{\partial D_T^h}\frac{U_{t_L}^h}{U_X^h}D_T^h - \frac{1-\phi_{lt}\vartheta}{1-\phi_{lt}\vartheta\xi}(\tau_g c_g + \tau_w \tfrac{w\varepsilon^h}{d_T}), \tag{12}$$

where $\phi_{lt} = \frac{n^\ell \varepsilon^\ell D_U^\ell}{n^h \varepsilon^h D_T^h + n^\ell \varepsilon^\ell D_U^\ell}$ (see Appendix B.1.2). This expression differs from (11) in that ϕ_{lt} takes into account the proportion of labor supplied by low-productivity households. Labor productivity enters in the toll rule, so that redistributing income through the labor tax implies that what drives the toll level is the proportion of labor supplied by low-productivity consumers (ϕ_{lt}) rather than their proportion in the economy (ϕ_{pt}).

4.2. Partially separating equilibrium with low-income groups separated

In this equilibrium high-income households take road T and road U, and low-income households take only road U. As the roads are substitutes, we assume that if the toll increases, high-income consumers reduce the number of trips they make by road T and increase the number of trips they make by road U. To keep things simple we assume that low-income users do not change their number of trips by road U as the toll increases[13]. Thus, the government's budget constraint becomes:

$$\begin{aligned} w\tau_w(n^h\varepsilon^h(D_T^h + D_U^h) + n^\ell\varepsilon^\ell D_U^\ell) + \tau_g c_g d_T(n^h(D_T^h + \beta D_U^h) + \beta n^\ell D_U^\ell) \\ + \tau d_T n^h D_T^h = F + NG. \end{aligned} \tag{13}$$

[13] This assumption is relaxed in the numerical illustration

4.2.1. Toll revenues used to finance head transfers

When the incremental toll revenues are used to finance lump-sum transfers we get (see Appendix B.2.1):

$$
\begin{aligned}
\tau_{pt} =& \frac{1}{1-\phi_{pt}\vartheta\xi}\frac{U^h_{t_L}}{U^h_X}\left(\frac{\partial t_T}{\partial D^h_T}D^h_T + \frac{\partial t_U}{\partial D^h_U}D^h_U\beta D^h_{TU}\right) \\
&- \frac{1-\phi_{pt}\vartheta}{1-\phi_{pt}\vartheta\xi}\Big(\tau_g c_g(1+\beta D^h_{TU}) + \tau_w \tfrac{w\varepsilon^h}{d_T}(1+D^h_{TU})\Big),
\end{aligned} \tag{14}
$$

Because in this case high-income commuters have the possibility to exchange trips on road T for trips on road U as the toll increases, we get the term $D^h_{TU} = \partial D^h_U/\partial D^h_T < 0$, which gives the number of trips added to U per trip removed from T[14]. Although equation (14) has the same structure as (11), it incorporates the marginal external congestion cost caused on road U by the fraction of trips moved from T to U. This is a typical second best result: mitigate the distortion on one market only to the extent that it does not aggravate the distortion on the other market (Small and Verhoef, 2007, p. 140). As before, (14) implies that the optimal toll should be set as a fraction of the difference between the *mecc* and other taxes per trip.

4.2.2. Toll revenues used to cut labor taxes

The optimal toll when the incremental toll revenues are used to cut labor taxes is given by (see Appendix B.2.2):

$$
\begin{aligned}
\tau_{lt} =& \frac{1}{1-\phi_{lt}\vartheta\xi}\frac{U^h_{t_L}}{U^h_X}\left(\frac{\partial t_T}{\partial D^h_T}D^h_T + \frac{\partial t_U}{\partial D^h_U}D^h_U\beta D^h_{TU}\right) \\
&- \frac{1-\phi_{lt}\vartheta}{1-\phi_{lt}\vartheta\xi}\Big(\tau_g c_g(1+\beta D^h_{TU}) + \tau_w \tfrac{w\varepsilon^h}{d_T}(1+D^h_{TU})\Big),
\end{aligned} \tag{15}
$$

with $\phi_{lt} = \frac{n^\ell\varepsilon^\ell D^\ell_U}{n^h\varepsilon^h(D^h_T+D^h_U)+n^\ell\varepsilon^\ell D^\ell_U}$. This equation has the same structure as (14) and contains the externality-correction term. Again, the only difference between (14) and (15) is ϕ_{lt}, which takes into account the proportion of labor supplied by low-productivity households as in (12).

Equations (14) and (15) imply therefore a toll level lower than that implied by (11) and (12), respectively, as the former includes a *mecc* reduced by the effect of traffic diversion.

[14] This trade-off between roads ($\partial D^h_U/\partial D^h_T$) affects the *mecc* and the second part of (14) since revenues collected from other taxes also depend on the road used.

4.3. Partially separating equilibrium with high-income groups separated

In this equilibrium high-income households take road T, and low-income households take both roads. As before, we assume that if the toll increases, low-productivity consumers reduce the number of trips they make by road T and increase the number of trips they make by road U. In addition, we assume that high-productivity consumers reduce their number of trips by road T only as a result of the toll increase. However, they do not move to road U[15]. The government's budget constraint becomes:

$$w\tau_w(n^h\varepsilon^h D_T^h + n^\ell\varepsilon^\ell(D_T^\ell + D_U^\ell)) + \tau_g c_g d_T(n^h D_T^h + n^\ell(D_T^\ell + \beta D_U^\ell)) \\ + \tau d_T(n^h D_T^h + n^\ell D_T^\ell) = F + NG. \tag{16}$$

The optimal toll, if revenues are used to make poll transfers, is as follows:

$$\begin{aligned}\tau_{pt} = \frac{1}{1-\phi_{pt}\vartheta-\zeta}&\left[E^\tau_{D_T^h}\frac{U^h_{t_L}}{U^h_X}\frac{\partial t_T}{\partial D_T^h}D_T^h\right.\\ &\left.+\frac{\theta^\ell}{\theta^h}E^\tau_{D_T^\ell}\frac{U^\ell_{t_L}}{U^h_X}\left(\frac{\partial t_T}{\partial D_T^\ell}D_T^\ell+\frac{\partial t_U}{\partial D_U^\ell}D_U^\ell\beta D_{TU}^\ell\right)\right]\\ -\frac{1-\phi_{pt}\vartheta}{1-\phi_{pt}\vartheta-\zeta}&\left[\tau_g c_g\left(E^\tau_{D_T^h}+E^\tau_{D_T^\ell}(1+\beta D_{TU}^\ell)\right)\right.\\ &\left.+\tau_w\tfrac{w}{d_T}\left(E^\tau_{D_T^h}\varepsilon^h+E^\tau_{D_T^\ell}\varepsilon^\ell(1+D_{TU}^\ell)\right)\right],\end{aligned} \tag{17}$$

where $\zeta = \frac{\sum_i n^i D_T^i + \vartheta\ell D_T^\ell}{\sum_i n^i D_T^i \epsilon^\tau_{D_T^i}}$ and $E^\tau_{D_T^i} = \frac{n^i D_T^i \epsilon^\tau_{D_T^i}}{\sum_i n^i D_T^i \epsilon^\tau_{D_T^i}}$ (see Appendix B.3).

Although (17) is more complex than previous equations, we can identify the same structure. The optimal toll should be set as a fraction of the difference between the externality-correction term and the level of other taxes per trip. The externality-correction term here consists of three terms: the *mecc* imposed on road T by both kinds of households and the *mecc* imposed on road U by low-income households.

In this case, the value of time and price elasticity of demand of both types of consumers appears in the equation as they both take the tolled road. Each term in this expression is weighted by a factor ($E^\tau_{D_T^i}$) that depends on the

[15]This assumption is relaxed in the numerical illustration

price elasticity of each type of household, as the response of consumers to toll increases depends on their price elasticity. As indicated by Small and Verhoef (2007, p. 145), when tolls cannot be differentiated among user groups, the second-best toll depends on a weighted average (by the price sensitivity of demand) of the marginal external costs for the different groups.

In this section and the next one we only derive the expression for the optimal toll when revenues are used to finance poll transfers. The complexity of the expressions makes that deriving the second case does not add any further insight other than in (17).

4.4. Pooling equilibrium

In this equilibrium both kinds of households take both roads. Assume that if the toll increases, both groups reduce the number of trips they make by road T and increase the number of trips they make by road U. Moreover, assume that road users reduce the number of trips they make on this road only as a result of the toll increase. The government's budget constraint is:

$$\begin{aligned} w\tau_w(n^h\varepsilon^h(D_T^h + D_U^h) + n^\ell\varepsilon^\ell(D_T^\ell + D_U^\ell)) + \tau_g c_g d_T(n^h(D_T^h + \beta D_U^h) \\ + n^\ell(D_T^\ell + \beta D_U^\ell)) + \tau d_T(n^h D_T^h + n^\ell D_T^\ell) = F + NG. \end{aligned} \tag{18}$$

The optimal toll when the incremental toll revenues are used to make poll transfers is as follows (see Appendix B.4):

$$\begin{aligned} \tau_{pt} = \frac{1}{1-\phi_{pt}\vartheta-\zeta} & \left[E^\tau_{D_T^h} \frac{U^h_{t_L}}{U^h_X} \left(\frac{\partial t_T}{\partial D_T^h} D_T^h + \frac{\partial t_U}{\partial D_U^h} D_U^h \beta D_{TU}^h \right) \right. \\ & \left. + \frac{\theta^\ell}{\theta^h} E^\tau_{D_T^\ell} \frac{U^\ell_{t_L}}{U^h_X} \left(\frac{\partial t_T}{\partial D_T^\ell} D_T^\ell + \frac{\partial t_U}{\partial D_U^\ell} D_U^\ell \beta D_{TU}^\ell \right) \right] \\ - \frac{1-\phi_{pt}\vartheta}{1-\phi_{pt}\vartheta-\zeta} & \left[\tau_g c_g \left(E^\tau_{D_T^h}(1+\beta D_{TU}^h) + E^\tau_{D_T^\ell}(1+\beta D_{TU}^\ell) \right) \right. \\ & \left. + \tau_w \tfrac{w}{d_T} \left(E^\tau_{D_T^h}\varepsilon^h(1+D_{TU}^h) + E^\tau_{D_T^\ell}\varepsilon^\ell(1+D_{TU}^\ell) \right) \right], \end{aligned} \tag{19}$$

Equation (19) is the more general equation derived in this analysis. It takes into account the use of the roads by both kinds of consumers. Interpretation is as explained before.

4.5. The optimal tolling rule: summary

The tolling rules derived in this Section are optimal deviations of the Pigouvian tax. They are in line with second-best pricing literature. For instance, they show that when only a single road in a network can be tolled, the congestion externality induced on substitute roads by traffic diversion should be taken into account. Similarly, when the toll cannot be differentiated among heterogeneous drivers, the value of the externality should be a weighted average of the externality for all the drivers using the tolled road.

A significant new insight can also be drawn from these tolling rules. A tolling policy can be used, to some extent, for redistributive purposes. The congestion toll can in fact act as an instrument to redistribute income given that the reduction of congestion is valued proportionally more by the people with a high value of time. Clearly, the price-elasticity of demand plays an important role in enhancing or limiting the potential for redistribution.

The redistributive role of the tolling policy can be easily seen when the toll allows either sorting road users between the tolled and the free alternatives or tolling off those that are least able to pay (Equilibriums 4.1 and 4.2, respectively). The degree of heterogeneity across road users determines how easily one can switch from one regime to another. Unfortunately, the use of the toll for redistributive purposes seems more difficult to identify when both high- and low-income groups use the tolled alternative (Equilibriums 4.3 and 4.4). Similarly, the analytical expressions derived in this Section give only limited insight about which way of revenue use is the most suitable in terms of redistribution. A numerical illustration is therefore useful.

5. Numerical illustration

This section presents the results of a numerical simulation[16] of a road network such as described in Section 2. Although this exercise is merely for illustrative purposes, we calibrate the model with French data in order to be coherent and give a realistic flavor to the illustration. Data for the labor and transport markets are taken from the *National Institute of Statistics and Economics Studies -INSEE* (Fesseau et al. (2009) and Baccani et al. (2007), respectively). Sections 5.1 and 5.2 present the parameter values used

[16]The algorithm was written in GAMS.

to calibrate the model and the simulation results for the base case. Next, Sections 5.3 and 5.4 examine the effect of changes in some key parameters.

5.1. Calibration

We choose an urban area of 500.000 inhabitants[17] where the average distance of a daily (round) commuting trip is 50 Km. The slope of the congestion function is such that the free-flow speed (60 Km/h) is reduced to one-third in peak hours (this implies a highly congested commuting traffic). Travel time increases linearly with increasing traffic volume. Both roads exhibit the same congestion functions, but the secondary network (i.e. the untolled road) is 1.5 times longer than the tolled road.

We define a household's utility function[18] separable in two terms, the utility of consumption/leisure and the disutility of traveling:

$$\mathcal{U}(X, t_L, D_U, D_T) = (\alpha_X X^{\frac{\sigma-1}{\sigma}} + (1-\alpha_X) t_L^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} + \alpha_C(D_U + D_T). \quad (1')$$

We choose $\sigma = 1.52$ to be consistent with values of consumption/leisure elasticity of the related literature[19]. α_X is chosen to imply (on average) around 200 days of work per year. We set $\alpha_C = -1$ and give no particular weight to any of the roads, so that the marginal disutility of traveling for any of the routes is the same[20]. In other words, preferences with respect to the use of the two roads are symmetric, such that households do not systematically prefer one road to the other.

There are two groups of workers that differ only in their labor productivity. The labor productivity of high-productivity households is around four times that of low-productivity households[21]. There is a higher proportion of

[17]With a labor force rate of 70% and an unemployment rate of 15%.

[18]Similar functions are used in Parry and Bento (2001) and Van Dender (2003).

[19]See Parry and Bento (2001) p.658 for a discussion of empirical evidence of this parameter.

[20]This implies that the roads are perfect substitutes from the consumer perspective. This reflects the consumer taste and has no relation with the characteristics of the roads.

[21]According to Fesseau et al. (2009) the best-off households have five times as much disposable income as the most modest. However this is considering total disposable income, without distinguishing between the source of income. When excluding returns on financial investments and property income, so that labor income is better accounted for, the difference in productivity decreases.

low-productivity workers in the economy (65%). We assume wage tax rates of 22% and 30% for low- and high-productive households[22], respectively, and 8 hours of work per day. The gasoline tax is 235% of producer prices (E.C., 2009, p. 11) and the vehicle fuel efficiency is 10 litres per 100 Km.

The constraints of this maximization problem are those described in (2) and (3). Thus, each household individually chooses, with perfect knowledge of the travel conditions on the network ($t_R(\sum n^i D^i_R)$), the route and the number of commuting trips. When a toll is imposed, toll extra-revenues are recycled in two ways: poll transfers and labor tax reductions.

5.2. Base-case results

We first concentrate on changes in the transport market. When there is no toll (the *no-toll equilibrium–NTE*), 69% of total traffic is concentrated on road T. This makes T highly congested. The average speed of a trip on this road is around 20 Km/h, whereas on U it is 31 Km/h. Figure 1 depicts the number of trips that the representative household of each type makes in a period, as a function of the toll (€/V-Km), in the case of labor-tax cuts[23].

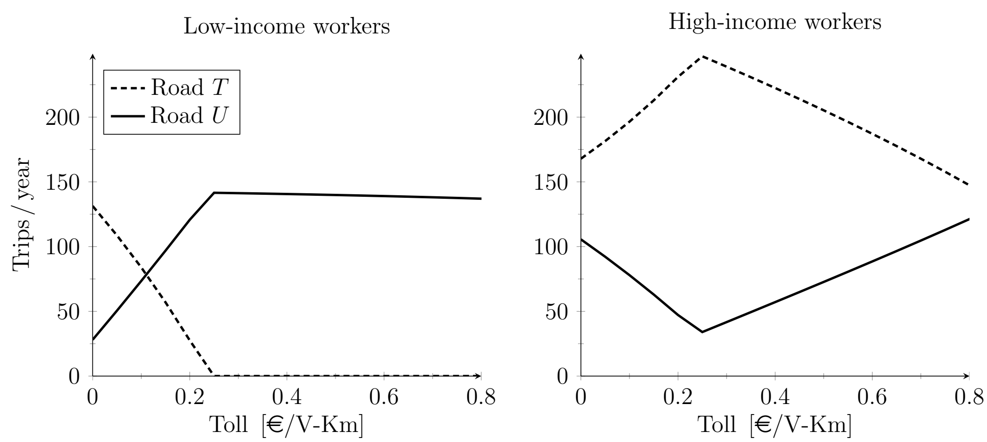

Figure 1: Trips per household in the base case (labor-tax recycling)

[22]This corresponds to the average rate of social contributions in Fesseau et al. (2009).
[23]Results from poll transfers are very similar in terms of road use.

Workers react differently to the toll. At the *NTE* low- and high-income consumers commute mainly by road T. Although this road is highly congested, it is shorter and allows less fuel-consuming trips. When the toll is imposed, low-income consumers reduce the number of trips on T and go to U. But high-income consumers react differently: they exchange trips on U by trips on T. As they can pay for the toll, they can take advantage of the reduction of congestion on T resulting from low-income commuters leaving this road[24]. This is true until the point where the toll approaches 0.25 €/V-Km.

For low toll values ($0 \leq \tau < 0.25$) we are in the *Pooling Equilibrium* where both kinds of workers commute by both roads. For higher toll values ($\tau \geq 0.25$) we are in the *Partially Separating Equilibrium with low-income groups separated*. In this case, low-income consumers are priced off road T[25], given that the cost of commuting by T exceeds the cost of commuting by U. High-income consumers, on the other hand, start switching to road U since paying a higher toll level no longer compensates the gain in time.

Given that road T allows faster and less fuel-consuming trips than the alternative, imposing a toll helps to reduce congestion on this road. At $\tau = 0.25$ (the level that allows the *separating regime* under labor-tax cuts), for instance, the reduction of traffic on T compared with traffic at the *NTE* is around 15% (the average speed rises from 20 to 26 Km/h). But this reduction comes at the expense of an increase of traffic on U of 35% compared with the *NTE* (the average speed falls from 31 to 24 Km/h). As a consequence, the reduction of congestion does not benefit all commuters in the same way. Given that low-productive workers use only road U at $\tau = 0.25$, their commuting time (per commuter) at this point increases by 12% whereas the commuting time for highly-productive workers falls by 14% (compared with the *NTE*).

Now we turn to the effects on the labor market. Reduced congestion induces a positive feedback effect, that mitigates the negative impact of the toll, but this holds only for highly productive workers (left panel Figure 2). The impact of the toll on the less productive workers is negative. Given the large losses on the less productive labor market, the impact on aggregate labor supply ends up negative (right panel Figure 2).

---

[24] The VOT for low- and high-income consumers are 3.89 and 18.43 €/h, respectively.

[25] Specifically, low-income consumers are tolled off road T at $\tau = 0.23$ for poll transfers and $\tau = 0.25$ for labor-tax cuts.

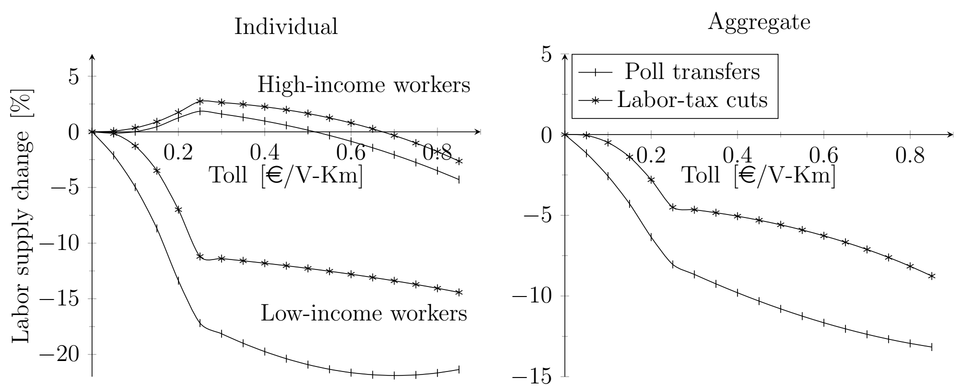

Figure 2: Labor supply in the base case: per type of household (left) and aggregate (right)

This differs from Parry and Bento (2001). Our results suggest that, when labor-force heterogeneity and congestion on the untolled alternative are accounted for, using toll revenues to reduce distortionary labor taxes does encourage labor force participation but only among the most productive workers (those able to pay for the toll). Similar results are found for poll transfers. This means that, both types of revenue allocation can discourage labor force participation at the margin. Our results still show, however, that labor supply would decline more when toll revenues are used to make poll transfers instead of labor tax cuts.

Welfare effects of both policies are depicted in Figure 3 (the vertical axis shows the change in individual welfare, in monetary terms, compared with welfare at the NTE). There is a clear difference between the two scenarios of revenue use across the income groups. Low-income consumers benefit in the case of head transfers. Recycling via labor-tax cuts is welfare reducing for them. On the other hand, high-income consumers benefit from both measures but the welfare gains are (slightly) higher when revenues are redistributed through labor-tax cuts. This is because, given the same percentage-point reduction of the labor tax for both groups, the resulting head subsidy is lower (higher) than the tax rebate that a high-income (low-income) household gets. Labor-tax cuts benefit the *h-households* given that, besides the gains from reduced congestion, they receive the greatest part of the revenues that are to be redistributed through this scheme. The *ℓ-households*, on the other hand, are not compensated enough for the increase of congestion they experience. This means that benefits and costs of recycling through this scheme are to some extent distributed in an *inequitable* way.

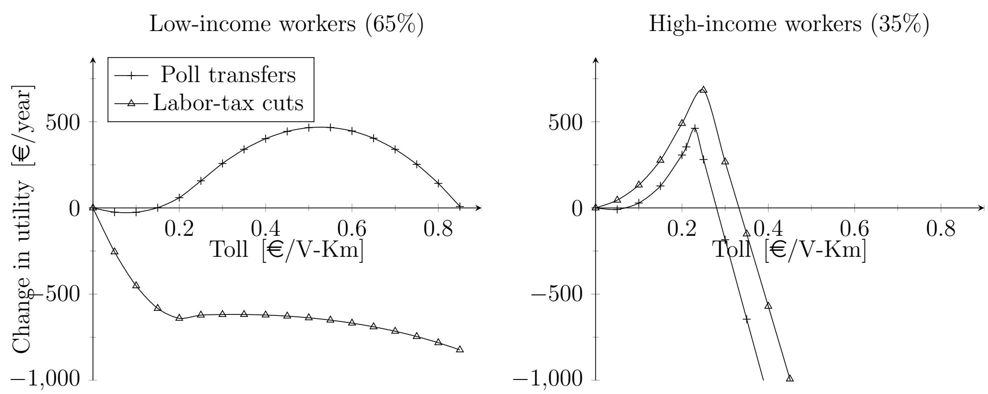

Figure 3: Welfare per type of household in the base case

It is worth noting an additional result from Figure 3. High-income workers get the maximum welfare gain at the point where low-income workers are tolled off road T (at $\tau = 0.23$ for poll transfers and at $\tau = 0.25$ for labor-tax cuts). Surprisingly, low-income workers also get the highest benefits when they are tolled off (of course, only in the case of poll transfers). In fact, in this case they do not pay for the toll but get the transfers from the high-income group. This would imply that product differentiation is beneficial, even for lower income groups, provided the right allocation of toll revenues.

This is consistent with Small and Yan (2001) in the sense that, there is a welfare gain when heterogeneity is accounted for. However, given that we consider labor markets, revenue allocation and redistribution[26], we find that the efficiency gain and the impact on both types of users depend on the way toll revenues are spent.

Even if this result contrasts with that of Parry and Bento (2001), it seems to be in line with Proost and Van Regemorter (1995) who show that the choice of the recycling scheme depends on preexisting conditions in the labor market[27].

[26]Small and Yan (2001) do not deal with the use of the toll revenues nor the effects of redistribution among users. Their focus is instead on the efficiency of the second-best

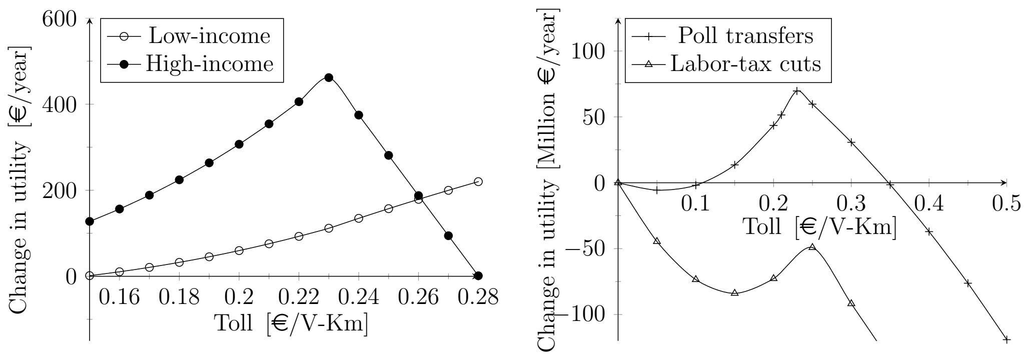

Figure 4: Welfare gains per household with poll-transfer recycling (left) and total welfare for both recycling policies (right)

Under the Pareto principle only recycling through poll transfers represents an improvement over the *NTE*. When acceptability requires the consent of both types of households, only this tax reform would be approved. Under this recycling regime, any toll level in the range from 0.15 to 0.28 would make all households better off (left panel Figure 4). The choice of the toll level depends on the government distribution concerns. Toll values close to the *separating-regime* level (e.g. 0.23) favor the highly productive workers (without making the ℓ-group worse off), and toll values closer to the higher limit (e.g. 0.28) favor the less productive workers (without making the h-group worse off). This implies that a toll level higher than the one that achieves efficiency in the transport market can be justified in regard to redistribution and acceptability concerns. Total welfare is depicted in the right panel of Figure 4. Welfare effects of labor-tax recycling are negative whereas poll-transfer recycling can be considered as *potential-Pareto* welfare improving in the range from 0.1 to 0.35 €/V-Km[28].

one-route pricing policy compared with the first-best result.

[27]Proost and Van Regemorter (1995) use an AGE-model for Belgium to study the effects of a carbon-energy tax by comparing the same two kinds of revenue recycling measures as analyzed here.

[28]Note that the *potential-Pareto* set (*Kaldor-Hicks criterion*) contains the *strict-Pareto* subset identified in the left panel of Figure 4.

5.3. The importance of congestibility of the untolled alternative

We explore here one of the key assumptions of our model: congestion on the untolled alternative. In this case, this road's congestion function is replaced by a constant time-per-unit-of-distance function. This means that, independent of the travel speed, an extra trip added to this road does not increase the user's time cost. We call this a *non-congestible* alternative. We calibrate the time function to exhibit the same speed as road T at the NTE[29] (16 Km/h). The rest of the parameters remain unchanged, except for α_X that, as in the base case, is chosen to imply on average 200 days of work per year at the NTE.

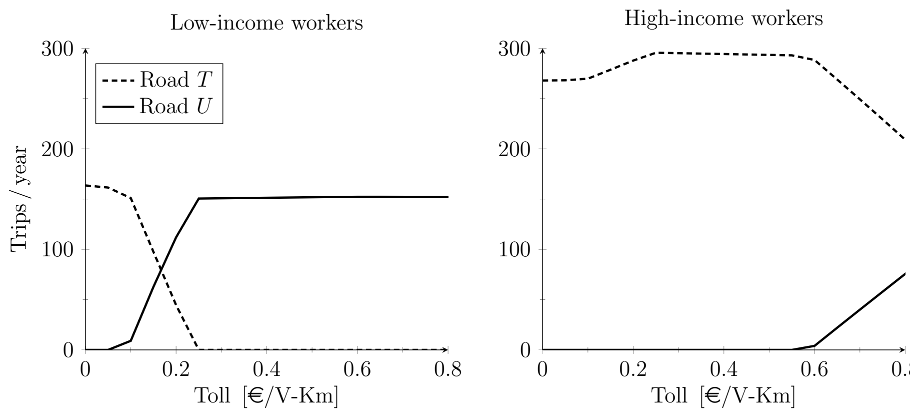

Figure 5: Trips per household with a *non-congestible* alternative (labor-tax recycling)

Travel patterns change significantly (Figure 5), not because the fixed-time cost of U, but rather because both roads exhibit a high time cost at the NTE. At this point, consumers only use road T. As the toll increases high-income households increase the use of this road and the low-income ones switch to road U.

Not surprisingly, the negative impact on labor supply is reduced now that road U can accommodate the traffic removed from T without any congestion effect (Figure 6). When toll revenues are recycled through labor-tax cuts,

[29] Think of this as a road with infinite capacity but whose travel speed is somehow limited to a given level. We take road T speed as a reference in order to be coherent with the previous section.

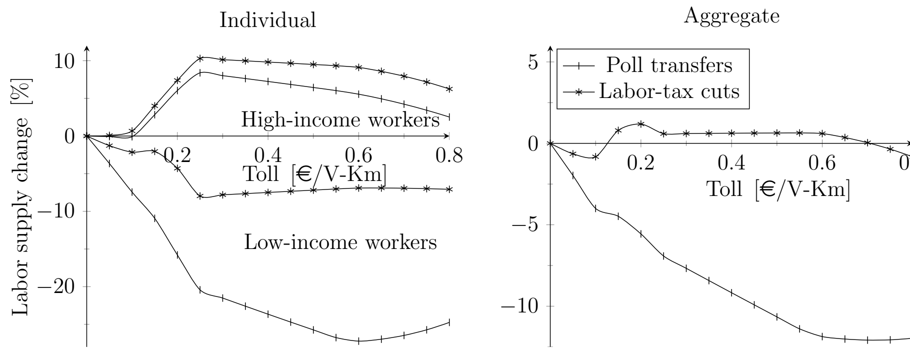

Figure 6: Labor supply with a *non-congestible* alternative: per type of household (left) and aggregate (right)

the net impact on aggregate labor supply is positive (right panel Figure 6). On the other hand, if toll revenues are recycled through poll transfers, the effects on aggregate labor supply are negative (but significantly lower than those in Figure 2) given the labor supply reduction by the less skilled. This means that neglecting congestion on the untolled alternative may result in an overestimation of the labor-supply gains from recycling, regardless of the recycling scheme.

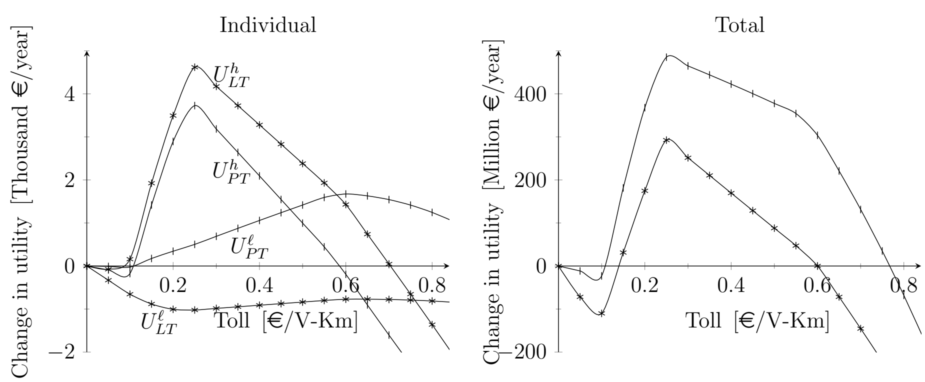

Figure 7: Change in utility for both recycling policies with a *non-congestible* alternative: per type of household (left) and total (right)

Welfare effects of both policies are similar to those in Figure 3. Recycling through labor-tax cuts remains welfare reducing for the low-income group and welfare improving for the high-income group. A separating equilibrium is beneficial for both groups. The main difference with the base-case result is that eliminating congestion on the alternative road increases the potential welfare gains from each regime. The labor-tax recycling scheme becomes (*potential-Pareto*) welfare improving for a large range of toll levels (right panel Figure 7). These results are clearly driven by the assumptions on the substitute for the tolled road. A close example is illustrated in Basso and Silva (2014). They show that congestion pricing is welfare improving for low-income groups whenever the substitute mode–in their case the public transport system–is optimized to accommodate the demand priced off the tolled road.

5.4. The relative size of the income groups

We briefly consider the sensitivity of results to the relative size of the income groups. Our results show that the composition of the economy plays an important role in the efficiency and welfare effects of the tolling policy.

We vary the size of the groups to imply a share for the low-income group of 5 to 95%. We find that the higher the proportion of low-income households in the economy, the higher the reduction of congestion on road T. Clearly, the more commuters that are willing to leave the tolled road whenever a toll is imposed, the more traffic that can be removed from T. The reduction of travel time ranges from around 4%, when the share of the ℓ*-households* is 5%, to around 48%, when the share of the ℓ*-households* is 95%.

Interestingly, the pattern of households that benefit from the tolling policy changes with the variation of these shares. The tolling policy is beneficial[30] for both types of households when the share of the ℓ*-households* ranges between 45 and 75%. For lower shares ($\ell < 45\%$) only the ℓ*-households* benefit from the tolling policy, and for higher shares ($\ell > 75\%$) only the *h-households* benefit[31]. In other words, the policy is welfare improving for both groups when the group with the lowest value of time represents more than a certain proportion of the population. However, if one of the groups

[30]It still holds that the low-income households only benefit from poll transfers and that the high-income ones benefit from both recycling regimes.

[31]These ranges are for the set of parameters of the base case. Of course, changes of some of those parameters could shift the limits of the ranges, but that should not change the main insight.

is the large majority, the tolling policy is welfare reducing for that group. Basically, if the majority of housheholds are low-income, toll revenues (paid by those who keep using T) are not enough to compensate this group for being diverted from T. And, if the majority of households are high-income, the reduction of congestion on T is too low to result in any benefit to this group.

The variation of the group shares also affects labor supply. This is directly related to the efficiency of the instrument in terms of reducing congestion. The higher the share of the low-income group, the more the high-income group will increase its labor supply and benefit from the pricing of one of the roads.

6. Conclusion

This paper considered the introduction of road pricing in an economy with low and high productive workers and where only one of two congested links can be tolled. Revenues can be recycled via poll transfers or labor-tax cuts. We show that the introduction of a toll recycled via lower labor taxes may benefit only the more productive workers. The main reason is that the less productive workers, who are tolled off the fast route, end up on the more congested untolled route and are insufficiently compensated by the labor tax reduction. For this reason recycling via a head subsidy may make road pricing more acceptable. Of course, whenever the untolled route is not subject to congestion, road pricing becomes a much more efficient instrument and it is much more likely that labor tax recycling becomes a more acceptable instrument.

Our results are relevant for all situations where the transport network cannot be tolled completely and where there are large differences in worker productivity. In many developing and developed countries these two conditions are present. The untolled alternative can be back roads or public transport. Some countries (e.g. France) even require the presence of an untolled alternative for every road that is tolled. We show that this is not a guarantee for a Pareto-improvement.

This paper can be extended in several ways. One can consider more complex networks, consider explicitly two modes rather than two links, introduce other local externalities along the two roads, consider leisure trips in addition, etc. However, this is unlikely to change our main insights.

7. Acknowledgements

We would like to thank useful comments of S. Tchung-Ming, F. Dunkerley, A. Ayong Le Kama, L. De Maack, and A. Decoster, as well as participants of the ITEA Conference 2013 and the seminar FLM-CEC N°61. Financial support of IFP Energies Nouvelles and EIB, administered via the EIBURS Program, is gratefully acknowledged.

8. Bibliography

Baccani, B., Semecurbe, F., Thomas, G., 2007. Les deplacements domicile-travail amplifies par la periurbanisation. INSEE PREMIERE 1129, INSEE.

De Borger, B., 2009. Commuting, congestion tolls and the structure of the labour market: Optimal congestion pricing in a wage bargaining model. Regional Science and Urban Economics 39 (4), 434 – 448.

E.C., 2009. Evolution of oil and petroleum product prices and taxation levels during the year 2008 in the european union. Tech. rep., Market Observatory for Energy. European Commission.

Fesseau, M., Bellamy, V., Raynaud, E., 2009. Inequality between households in the national accounts. INSEE PREMIERE 1265A, INSEE.

Jara-Diaz, S. R., 2000. Allocation and valuation of travel time savings. Handbooks in Transport 1, 303–319.

Mayeres, I., Proost, S., 1997. Optimal tax and public investment rules for congestion type of externalities. Scandinavian Journal of Economics 99 (2), 261–79.

Mayeres, I., Proost, S., 2001. Can we use transport accounts for pricing policy and distributional analysis? Energy, Transport and Environment Working Papers Series ETE0123, Katholieke Universiteit Leuven, Centrum voor Economische Studien.

Parry, I. W. H., Bento, A., 2001. Revenue recycling and the welfare effects of road pricing. Scandinavian Journal of Economics 103 (4), 645–71.

Parry, I. W. H., Oates, W. E., 2000. Policy analysis in the presence of distorting taxes. Journal of Policy Analysis and Management 19, 603–613.

Parry, I. W. H., Small, K. A., 2005. Does Britain or the United States have the right gasoline tax? American Economic Review 95 (4), 1276–1289.

Proost, S., Van Regemorter, D., 1995. The double dividend and the role of inequality aversion and macroeconomic regimes. International Tax and Public Finance 2 (2), 207–219.

Rouwendal, J., Verhoef, E. T., 2004. Second-best pricing for imperfect substitutes in urban networks. Research in Transportation Economics 9 (1), 27–60.

Small, K. A., Verhoef, E. T., 2007. The economics of urban transportation. Routledge.

Small, K. A., Yan, J., 2001. The value of "value pricing" of roads: Second-best pricing and product differentiation. Journal of Urban Economics 49 (2), 310–336.

Van Dender, K., 2003. Transport taxes with multiple trip purposes. Scandinavian Journal of Economics 105 (2), 295–310.

Appendix A. The household's problem

The household's problem defined by (1), (2) and (3) can be solved by maximizing the following Lagrangian function:

$$\mathcal{L} = \mathcal{U}(X, t_L) + C(D_U, D_T) - \lambda_c \left[X + g d_U D_U + (g+\tau) d_T D_T \right. \\ \left. -\varepsilon w (1-\tau_w)(D_U + D_T) - G\right] + \mu_c \left[\bar{t} - D_U(1 + t_U \beta d_T) - D_T(1 + t_T d_T) - t_L\right],$$

where the Lagrangian multiplier related to the income constraint (λ_c) is the *marginal utility of income*, and the Lagrangian multiplier related to the time constraint (μ_c) is the *resource value of time*. For $X > 0$, $D_U > 0$, $D_T > 0$ and $t_L > 0$, the system of first-order conditions can be written as:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial X} &= U_X - \lambda_c = 0 \qquad \Rightarrow \lambda_c = U_X \\ \frac{\partial \mathcal{L}}{\partial t_L} &= U_{t_L} - \mu_c = 0 \qquad \Rightarrow \mu_c = U_{t_L} \\ \frac{\partial \mathcal{L}}{\partial D_U} &= C_{D_U} - \lambda_c \left[g \beta d_T - \varepsilon w (1-\tau_w)\right] - \mu_c (1 + t_U \beta d_T) = 0 \\ \frac{\partial \mathcal{L}}{\partial D_T} &= C_{D_T} - \lambda_c \left[(g+\tau) d_T - \varepsilon w (1-\tau_w)\right] - \mu_c (1 + t_T d_T) = 0 \end{aligned} \tag{A.1}$$

Using these conditions and both budget constraints, we obtain the demand functions for X^*, D_U^*, D_T^*, and t_L^*. Replacing these functions in the utility gives the indirect utility function $\upsilon(\tau_g, \tau_w, \tau, t_T, t_U, w, \varepsilon, G)$ which enables rewriting the household's problem as:

$$\mathcal{L} = \upsilon(\tau_g, \tau_w, \tau, t_T, t_U, w, \varepsilon, G) + \lambda_c \left[X + g d_U D_U + (g+\tau) d_T D_T \right.$$
$$\left. -\varepsilon w(1-\tau_w)(D_U + D_T) - G\right] - \mu_c \left[\bar{t} - D_U(1 + t_U \beta d_T) - D_T(1 + t_T d_T) - t_L\right].$$

F.O.C.:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \tau} &= \upsilon_\tau + \lambda_c d_T D_T = 0 && \Rightarrow \upsilon_\tau = -U_X d_T D_T \\
\frac{\partial \mathcal{L}}{\partial \tau_g} &= \upsilon_{\tau_g} + \lambda_c c_g d_T(\beta D_U + D_T) = 0 && \Rightarrow \upsilon_{\tau_g} = -U_X c_g d_T(\beta D_U + D_T) \\
\frac{\partial \mathcal{L}}{\partial \tau_w} &= \upsilon_{\tau_w} + \lambda_c \varepsilon w(D_U + D_T) = 0 && \Rightarrow \upsilon_{\tau_w} = -U_X \varepsilon w(D_U + D_T) \\
\frac{\partial \mathcal{L}}{\partial t_U} &= \upsilon_{t_U} + \mu_c d_U D_U = 0 && \Rightarrow \upsilon_{t_U} = -U_{t_L} d_U D_U \\
\frac{\partial \mathcal{L}}{\partial t_T} &= \upsilon_{t_T} + \mu_c d_T D_T = 0 && \Rightarrow \upsilon_{t_T} = -U_{t_L} d_T D_T \\
\frac{\partial \mathcal{L}}{\partial G} &= \upsilon_G - \lambda_c = 0 && \Rightarrow \upsilon_G = U_X
\end{aligned} \tag{A.2}$$

Note that the marginal disutility of the toll increase (υ_τ) is the *marginal utility of income* (U_X) multiplied by the optimal number of trips (D_T). Similarly, the marginal disutility of an increase of travel time on road T (υ_{t_T}) is the *the resource value of time* (U_{t_L}) multiplied by the optimal number of trips.

Appendix B. The social's planner problem

Appendix B.1. Separated equilibrium

By assumption, in this case, $D_U^h = 0$ and $D_T^\ell = 0$.

Appendix B.1.1. Poll transfers

Differentiating the social welfare function with respect to τ, when $d\tau$ affects dG, gives:

$$\frac{d\mathbb{W}}{d\tau} = \theta^h n^h \left(\upsilon_\tau^h + \upsilon_{t_T}^h \frac{\partial t_T}{\partial D_T^h} \frac{dD_T^h}{d\tau} + \upsilon_G^h \frac{dG}{d\tau} \right) + \theta^\ell n^\ell \upsilon_G^\ell \frac{dG}{d\tau} \tag{B.1}$$

with $\partial t_T / \partial D_T^h > 0$, $dD_T^h / d\tau < 0$, and $dG/d\tau > 0$. Replacing A.2 into B.1 gives:

$$\frac{d\mathbb{W}}{d\tau} = -\theta^h n^h d_T D_T^h \left(U_X^h + U_{t_L}^h \frac{\partial t_T}{\partial D_T^h} \frac{dD_T^h}{d\tau} \right) + \left(\theta^h n^h U_X^h + \theta^\ell n^\ell U_X^\ell \right) \frac{dG}{d\tau} \tag{B.2}$$

Differentiating (10) with respect to τ gives the change in the transfer (dG) associated to a change in the toll ($d\tau$):

$$\frac{dG}{d\tau} = \frac{n^h d_T}{N}\left[\left(\tau + \tau_g c_g + \tau_w \frac{w\varepsilon^h}{d_T}\right)\frac{dD_T^h}{d\tau} + D_T^h\right] \tag{B.3}$$

Inserting B.3 into B.2 and dividing by U_X^h, we have:

$$\begin{aligned}\frac{d\mathbb{W}/d\tau}{U_X^h} =& \left(\frac{\theta^h n^h}{N} + \frac{\theta^\ell n^\ell}{N}\frac{U_X^\ell}{U_X^h}\right)\left(\tau + \tau_g c_g + \tau_w \frac{w\varepsilon^h}{d_T}\right)\frac{dD_T^h}{d\tau} \\ &+ \left(\frac{\theta^h n^h}{N} + \frac{\theta^\ell n^\ell}{N}\frac{U_X^\ell}{U_X^h} - \theta^h\right) D_T^h - \theta^h \frac{U_{t_L}^h}{U_X^h} D_T^h \frac{\partial t_T}{\partial D_T^h}\frac{dD_T^h}{d\tau}\end{aligned} \tag{B.4}$$

Setting $\frac{d\mathbb{W}/d\tau}{U_X^h} = 0$, and defining the elasticity of demand of high-income consumers for the tolled road as $\epsilon_{D_T^h}^{\tau} = \frac{dD_T^h}{d\tau}\frac{\tau}{D_T^h}$, we get (11).

Appendix B.1.2. Labor-tax cuts

The welfare impact when incremental toll revenues are used to cut the labor-tax is:

$$\frac{d\mathbb{W}}{d\tau} = \theta^h n^h\left(\upsilon_\tau^h + \upsilon_{t_T}^h \frac{\partial t_T}{\partial D_T^h}\frac{dD_T^h}{d\tau} + \upsilon_{\tau_w}^h \frac{d\tau_w}{d\tau}\right) + \theta^\ell n^\ell \upsilon_{\tau_w}^\ell \frac{d\tau_w}{d\tau} \tag{B.5}$$

With $\frac{d\tau_w}{d\tau} < 0$. Replacing A.2 into B.5 we have:

$$\begin{aligned}\frac{d\mathbb{W}}{d\tau} =& -\theta^h n^h d_T D_T^h\left(U_X^h + U_{t_L}^h \frac{\partial t_T}{\partial D_T^h}\frac{dD_T^h}{d\tau}\right) \\ &- \left(\theta^h n^h \varepsilon^h D_T^h U_X^h + \theta^\ell n^\ell \varepsilon^\ell D_U^\ell U_X^\ell\right) w\frac{d\tau_w}{d\tau}\end{aligned} \tag{B.6}$$

Differentiating (10) with respect to τ and solving for $w\frac{d\tau_w}{d\tau}$ gives:

$$w\frac{d\tau_w}{d\tau} = \frac{-n^h d_T}{n^h\varepsilon^h D_T^h + n^\ell \varepsilon^\ell D_U^\ell}\left[\left(\tau + \tau_g c_g + \tau_w\frac{w\varepsilon^h}{d_T}\right)\frac{dD_T^h}{d\tau} + D_T^h\right] \tag{B.7}$$

Inserting B.7 into B.6, dividing by U_X^h, setting $\frac{d\mathbb{W}/d\tau}{U_X^h} = 0$, and using $\epsilon_{D_T^h}^{\tau}$ we get (12).

Appendix B.2. Partially separated equilibrium with ℓ-groups separated

Here by assumption $D_T^\ell = 0$.

Appendix B.2.1. Poll transfers

$$\frac{d\mathbb{W}}{d\tau} = \theta^h n^h\left(\upsilon_\tau^h + \upsilon_{t_T}^h \frac{\partial t_T}{\partial D_T^h}\frac{dD_T^h}{d\tau} + \upsilon_{t_U}^h \frac{\partial t_U}{\partial D_U^h}\frac{\partial D_U^h}{\partial D_T^h}\frac{dD_U^h}{d\tau} + \upsilon_G^h\frac{dG}{d\tau}\right) + \theta^\ell n^\ell \upsilon_G^\ell \frac{dG}{d\tau} \tag{B.8}$$

with $\frac{\partial D_U^h}{\partial D_T^h} < 0$. Replacing A.2 into B.8 gives:

$$\frac{d\mathbb{W}}{d\tau} = -\theta^h n^h \left(U_X^h d_T D_T^h + U_{t_L}^h d_T D_T^h \frac{\partial t_T}{\partial D_T^h} \frac{dD_T^h}{d\tau} + U_{t_L}^h \beta d_T D_U^h \frac{\partial t_U}{\partial D_U^h} \frac{\partial D_U^h}{\partial D_T^h} \frac{dD_T^h}{d\tau} - U_X^h \frac{dG}{d\tau} \right) + \theta^\ell n^\ell U_G^\ell \frac{dG}{d\tau} \tag{B.9}$$

Differentiating (13) with respect to τ and solving for $\frac{dG}{d\tau}$ gives:

$$\frac{dG}{d\tau} = \frac{n^h d_T}{N} \left[\left(\tau + \tau_g c_g \left(1 + \beta \frac{\partial D_U^h}{\partial D_T^h} \right) + \tau_w \frac{w\varepsilon^h}{d_T} \left(1 + \frac{\partial D_U^h}{\partial D_T^h} \right) \right) \frac{dD_T^h}{d\tau} + D_T^h \right] \tag{B.10}$$

Inserting B.10 into B.9, dividing by U_X^h, setting $\frac{d\mathbb{W}/d\tau}{U_X^h} = 0$, and using $\epsilon_{D_T^h}^\tau$ we get (14).

Appendix B.2.2. Labor-tax cuts

$$\frac{d\mathbb{W}}{d\tau} = \theta^h n^h \left(\upsilon_\tau^h + \upsilon_{t_T}^h \frac{\partial t_T}{\partial D_T^h} \frac{dD_T^h}{d\tau} + \upsilon_{t_U}^h \frac{\partial t_U}{\partial D_U^h} \frac{\partial D_U^h}{\partial D_T^h} \frac{dD_U^h}{d\tau} + \upsilon_{\tau_w}^h \frac{d\tau_w}{d\tau} \right) + \theta^\ell n^\ell \upsilon_{\tau_w}^\ell \frac{d\tau_w}{d\tau} \tag{B.11}$$

Replacing A.2 into B.11 gives:

$$\frac{d\mathbb{W}}{d\tau} = -\theta^h n^h \left(U_X^h d_T D_T^h + U_{t_L}^h d_T D_T^h \frac{\partial t_T}{\partial D_T^h} \frac{dD_T^h}{d\tau} + U_{t_L}^h \beta d_T D_U^h \frac{\partial t_U}{\partial D_U^h} \frac{\partial D_U^h}{\partial D_T^h} \frac{dD_T^h}{d\tau} - U_X^h \varepsilon^h W (D_T^h + D_U^h) \frac{d\tau_w}{d\tau} \right) - \theta^\ell n^\ell U_X^\ell w \varepsilon^\ell D_U^\ell \frac{d\tau_w}{d\tau} \tag{B.12}$$

Differentiating (13) with respect to τ and solving for $W \frac{d\tau_w}{d\tau}$ gives:

$$w \frac{d\tau_w}{d\tau} = -\frac{n^h d_T}{n^h \varepsilon^h (D_T^h + D_U^h) + n^\ell \varepsilon^\ell D_U^\ell} \left[\left(\tau + \tau_g c_g \left(1 + \beta \frac{\partial D_U^h}{\partial D_T^h} \right) + \tau_w \frac{w\varepsilon^h}{d_T} \left(1 + \frac{\partial D_U^h}{\partial D_T^h} \right) \right) \frac{dD_T^h}{d\tau} + D_T^h \right] \tag{B.13}$$

Inserting B.13 into B.12, dividing by U_X^h, setting $\frac{d\mathbb{W}/d\tau}{U_X^h} = 0$, and using $\epsilon_{D_T^h}^\tau$ we get (15).

Appendix B.3. Partially separated equilibrium with h-groups separated and revenues recycled via poll transfers

Here by assumption $D_T^h = 0$.

$$\frac{d\mathbb{W}}{d\tau} = \theta^h n^h \left(\upsilon_\tau^h + \upsilon_{t_T}^h \frac{\partial t_T}{\partial D_T^h} \frac{dD_T^h}{d\tau} + \upsilon_G^h \frac{dG}{d\tau} \right) + \theta^\ell n^\ell \left(\upsilon_\tau^\ell + \upsilon_{t_T}^\ell \frac{\partial t_T}{\partial D_T^\ell} \frac{dD_T^\ell}{d\tau} + \upsilon_{t_U}^\ell \frac{\partial t_U}{\partial D_U^\ell} \frac{\partial D_U^\ell}{\partial D_T^\ell} \frac{dD_T^\ell}{d\tau} + \upsilon_G^\ell \frac{dG}{d\tau} \right) \tag{B.14}$$

Replacing A.2 into B.14 gives:

$$\frac{d\mathbb{W}}{d\tau} = -\left(\theta^h n^h U_X^h D_T^h + \theta^\ell n^\ell U_X^\ell D_T^\ell\right) d_T - \left(\theta^h n^h U_{t_L}^h D_T^h \frac{\partial t_T}{\partial D_T^h}\frac{dD_T^h}{d\tau} + \theta^\ell n^\ell U_{t_L}^\ell D_T^\ell \frac{\partial t_T}{\partial D_T^\ell}\frac{dD_T^\ell}{d\tau}\right) d_T$$
$$- \theta^\ell n^\ell U_{t_L}^\ell \beta d_T D_U^\ell \frac{\partial t_U}{\partial D_U^\ell}\frac{\partial D_U^\ell}{\partial D_T^\ell}\frac{dD_T^\ell}{d\tau} + \left(\theta^h n^h U_X^h + \theta^\ell n^\ell U_X^\ell\right)\frac{dG}{d\tau} \tag{B.15}$$

Differentiating (16) with respect to τ and solving for $\frac{dG}{d\tau}$ gives:

$$\frac{dG}{d\tau} = \frac{d_T}{N}\left[\tau\left(n^h \frac{dD_T^h}{d\tau} + n^\ell \frac{dD_T^\ell}{d\tau}\right) + \tau_g c_g \left(n^h \frac{dD_T^h}{d\tau} + n^\ell\left(1 + \beta\frac{\partial D_U^\ell}{\partial D_T^\ell}\right)\frac{dD_T^\ell}{d\tau}\right)\right.$$
$$\left. + \tau_w \frac{w}{d_T}\left(n^h \varepsilon^h \frac{dD_T^h}{d\tau} + n^h \varepsilon^h \left(1 + \frac{\partial D_U^\ell}{\partial D_T^\ell}\right)\frac{dD_T^\ell}{d\tau}\right) + n^h D_T^h + n^\ell D_T^\ell\right] \tag{B.16}$$

Inserting B.16 into B.15, dividing by U_X^h, setting $\frac{d\mathbb{W}/d\tau}{U_X^h} = 0$, and using $\epsilon^\tau_{D_T^i}$ we get (17).

Appendix B.4. Pooling equilibrium and revenues recycled via poll transfers

$$\frac{d\mathbb{W}}{d\tau} = \theta^h n^h \left(\upsilon_\tau^h + \upsilon_{t_T}^h \frac{\partial t_T}{\partial D_T^h}\frac{dD_T^h}{d\tau} + \upsilon_{t_U}^h \frac{\partial t_U}{\partial D_U^h}\frac{\partial D_U^h}{\partial D_T^h}\frac{dD_T^h}{d\tau} + \upsilon_G^h \frac{dG}{d\tau}\right)$$
$$+ \theta^\ell n^\ell \left(\upsilon_\tau^\ell + \upsilon_{t_T}^\ell \frac{\partial t_T}{\partial D_T^\ell}\frac{dD_T^\ell}{d\tau} + \upsilon_{t_U}^\ell \frac{\partial t_U}{\partial D_U^\ell}\frac{\partial D_U^\ell}{\partial D_T^\ell}\frac{dD_T^\ell}{d\tau} + \upsilon_G^\ell \frac{dG}{d\tau}\right) \tag{B.17}$$

Replacing A.2 into B.17 gives:

$$\frac{d\mathbb{W}}{d\tau} = -\left(\theta^h n^h U_X^h D_T^h + \theta^\ell n^\ell U_X^\ell D_T^\ell\right) d_T - \left(\theta^h n^h U_{t_L}^h D_T^h \frac{\partial t_T}{\partial D_T^h}\frac{dD_T^h}{d\tau} + \theta^\ell n^\ell U_{t_L}^\ell D_T^\ell \frac{\partial t_T}{\partial D_T^\ell}\frac{dD_T^\ell}{d\tau}\right) d_T$$
$$- \left(\theta^h n^h U_{t_L}^h D_U^h \frac{\partial t_U}{\partial D_U^h}\frac{\partial D_U^h}{\partial D_T^h}\frac{dD_T^h}{d\tau} + \theta^\ell n^\ell U_{t_L}^\ell D_U^\ell \frac{\partial t_U}{\partial D_U^\ell}\frac{\partial D_U^\ell}{\partial D_T^\ell}\frac{dD_T^\ell}{d\tau}\right)\beta d_T + \left(\theta^h n^h U_X^h + \theta^\ell n^\ell U_X^\ell\right)\frac{dG}{d\tau} \tag{B.18}$$

Differentiating (18) with respect to τ and solving for $\frac{dG}{d\tau}$ gives:

$$\frac{dG}{d\tau} = \frac{d_T}{N}\left[\tau\left(n^h \frac{dD_T^h}{d\tau} + n^\ell \frac{dD_T^\ell}{d\tau}\right) + \tau_g c_g \left(n^h\left(1 + \beta\frac{\partial D_U^h}{\partial D_T^h}\right)\frac{dD_T^h}{d\tau} + n^\ell\left(1 + \beta\frac{\partial D_U^\ell}{\partial D_T^\ell}\right)\frac{dD_T^\ell}{d\tau}\right)\right.$$
$$\left. + \tau_w \frac{w}{d_T}\left(n^h \varepsilon^h \left(1 + \frac{\partial D_U^h}{\partial D_T^h}\right)\frac{dD_T^h}{d\tau} + n^h \varepsilon^h \left(1 + \frac{\partial D_U^\ell}{\partial D_T^\ell}\right)\frac{dD_T^\ell}{d\tau}\right) + n^h D_T^h + n^\ell D_T^\ell\right] \tag{B.19}$$

Inserting B.19 into B.18, dividing by U_X^h, setting $\frac{d\mathbb{W}/d\tau}{U_X^h} = 0$, and using $\epsilon^\tau_{D_T^i}$ we get (19).