Tradable mobility permits in a monocentric city with pre-existing labor taxation: a general equilibrium perspective

Diego Candia[1]
Erik T. Verhoef[1]

[1] Vrije Universiteit Amsterdam

Tradable mobility permits in a monocentric city with pre-existing labor taxation: a general equilibrium perspective

Diego Candia[*, a, b], Erik Verhoef [a, b]

[*]*Corresponding author. Email address: d.a.candiariquelme@vu.nl*
[a]*Department of Spatial Economics, VU Amsterdam, De Boelelaan 1105, Amsterdam 1081 HV, the Netherlands*
[b]*Tinbergen Institute, Gustav Mahlerplein 117, Amsterdam 1082 MS, The Netherlands*

Abstract

This paper considers tradable mobility permit schemes in a monocentric city with a distorting labor tax. Three schemes are analyzed, that differ by the (spatial) allocation of permits to households. Numerical results show that the scheme with permits allocated in proportion to labor supply reaches about 99.9% of the first-best welfare, being the only welfare-increasing scheme for higher levels of labor tax, and being more efficient than the second-best tax for levels of the labor tax of 20% or higher. This is due to the welfare gains of incentivizing labor supply, which has an effect similar to the revenue recycling effect of conventional road pricing, but that can be better optimized by the social planner by exploiting spatial differentiation. When permits are allocated on the basis of households' residential distance from the CBD, a spatially differentiated component leads to higher welfare levels than with the third scheme considered, which is the allocation in proportion to households' kilometers traveled.

1. Introduction

Tradable mobility permits are a relatively recent, quantity-based instrument for managing urban congestion. Having their origin in environmental economics (Dales, 1968), tradable permit schemes have also been studied for a while in urban transportation (e.g. Verhoef et al., 1997; Goddard, 1997) as an alternative to road pricing. Under a tradable permit scheme, a social planner allocates permits that are required to perform certain activity (e.g., to travel one kilometer), and sets a targeted quantity (e.g., total kilometers traveled). Permits can be freely traded, which results in an equilibrium price determined by the market, that is meant to affect behavior in a way comparable to a conventional road price.

While tradable permits and road pricing are equivalent in terms of economic efficiency in the most basic setting (Weitzman, 1974), in practice a number of differences are relevant. First, a tradable permit scheme does not require a financial transfer from users to the social planner (as is the case under a road pricing scheme) which likely will increase its social and political acceptability, even

though it is true that the recycling of road pricing revenues could counter the financial transfer of that kind of schemes. In terms of effectiveness, a permit scheme is better targeted than road pricing for direct quantity control when there is uncertainty in the response of agents to pricing (Fan and Jiang, 2013), or on demand and costs structures in general (Verhoef et al., 1997).

On the other hand, the implementation of tradable permit schemes would need the creation of a market where permits can be traded between users, which could lead to non-negligible transaction costs, limiting their efficiency to some extent. Nie (2012) showed that even small market transaction costs could cause the market diverting from the targeted equilibrium. However, recent technological trends in vehicles connectivity and mobile internet could lead to transaction costs close to zero. Recently, Brands et al. (2020) designed and implemented a lab-in-the-field tradable permit market where online trading made transaction costs small enough to prevent these from being a significant distorting element. In this paper, we assume the existence of that kind of market design, and therefore will ignore transaction costs.

Another relevant difference is the outcome in presence of uncertainty (Weitzman, 1974). de Palma et al. (2018) compared both instruments in a simple road network with uncertainty in supply and demand. The authors found that when both instruments are not day-to-day adaptive, permits perform better on average than pricing in terms of total welfare. Shirmohammadi et al. (2013) studied the price fluctuations of tradable permits under uncertainty, finding a high volatility that can lead to a price unacceptably high.

One important characteristic of tradable permit schemes is that they would offer the possibility to address redistributional aspects of congestion policies more directly, through focusing the allocation of permits to certain groups of the population, whereas, in the case of road pricing, revenues are usually recycled through public investments or general tax measures, which are difficult to focus. However, the presence of the so-called revenue recycling effect of road pricing, achieved by using the revenues collected to finance a reduction of a distortionary tax, could be a key determinant for the welfare result of a pricing policy, and it may even determine whether such a policy increases or decreases welfare, as shown by Parry and Bento (2001) and Tikoudis et al. (2015). On the other hand, while tradable permits do not raise revenues, the fact that the social planner is in charge of allocating permits to users could provide powerful incentives to increase welfare.

The welfare effects of both congestion policies and of recycling measures, and their optimal design, will be strongly differentiated over urban space. Also, it is to be expected that strong interrelations between transport, labor and housing markets will affect the eventual outcomes of these instruments. Against this background, the aim of this paper is to analyze the spatial general equilibrium effects of different tradable permit schemes when there is a pre-existing labor tax. For that, we use the monocentric spatial general equilibrium model developed by Verhoef and Nijkamp (2002) and Verhoef (2005). In our paper, one permit is needed for each kilometer traveled, and permits can be freely traded between households. The equilibrium permit price is determined endogenously through the permit market equilibrium. Each tradable permit scheme is defined by a second-best optimized targeted quantity (aggregate kilometers traveled), and by the initial allocation of permits to households. We consider three different allocations, characterized by the spatial variable that determines the number of permits that each household receives. For that, we consider three variables that seem natural: i) the household's residential location, given by its distance from the CBD; ii) the household's labor supply (number of days worked); and iii) the household's kilometers traveled, which is a combination of labor supply and distance from the CBD.

Following Tikoudis et al. (2015), we consider a pre-existing distorting labor tax, which is in place to finance an exogenous public revenue constraint, associated with the supply of an unmodeled public good. There is a road externality, which is congestion caused by drivers commuting to work. Therefore, the tradable permit schemes interact with these two distortions in the economy. Tikoudis et al. (2015) showed how the spatial differentiation of optimal taxation with a public revenue constraint is linked with two forces that interact, generating patterns of the optimal tax that can be non-monotonic over space: the *Mirrlees rule*, which indicates that taxes should be higher where the marginal utility of income is lower; and the *Ramsey rule*, which indicates that taxes should be higher where labor supply is less elastic with respect to labor tax.

Furthermore, Tikoudis et al. (2015) showed that a pre-existing labor tax implies that different possible effects of a congestion instrument become relevant: the welfare gains from a marginal reduction of congestion, i.e. the *Pigouvian effect*; the welfare gains from using road tax revenues to finance a labor tax cut, diminishing the distortion in the market i.e., the *revenue recycling effect*; and the welfare losses from diminishing real wage in an already tax-distorted labor market, i.e., the *tax interaction effect*. We expect to see the same kinds of phenomena interacting with the tradable permit schemes considered in this paper.

Tradable permits have been previously studied with different perspectives, for example, the general network approach started by Yang and Wang (2011) and continued by Wang et al. (2012); Nie (2012); Wu et al. (2012); He et al. (2013), among others; or the bottleneck approach, e.g.: Zhang et al. (2011), Tian et al. (2013); Xiao et al. (2013); Nie (2015); Akamatsu and Wada (2017). To the best of our knowledge, this is the first article to study the effects of tradable mobility permit schemes in a spatial general equilibrium model.

The remainder of this article is organized as follows. Section 2 presents the analytical model, where the tradable permit schemes are described and the analytical properties are discussed. Section 3 describes and discusses the calibration of the base case and the optimal taxation profile. Section 4 presents the numerical results and analyses of tradable permit schemes. Finally, Section 5 summarizes and concludes.

2. The analytical spatial equilibrium model

The model[1] considers a linear monocentric city, with a continuous urban space denoted by $z \in (0, \bar{z})$, where $\bar{z}$ is the endogenous city boundary and $z = 0$ is the location of the spaceless CBD. The city is considered as closed, which means that the population size N is exogenous and fixed.

Identical utility-maximizing households decide where to locate, considering their composite good consumption, labor supply, leisure and housing-space consumption. A representative competitive firm, located at the CBD, produces the composite good using labor supply as the only input. Households commute from their residential location to their work (located at the CBD), using a single congested road. The only mode available for commuting is private car, the only purpose for traveling is commuting and the number of hours on a working day is fixed, which means that diminishing the number of commuting trips is equivalent to diminishing labor supply, and that the kilometers traveled by one household are equal to its labor supply multiplied by its residential distance from the CBD.

---

[1] The model presented in this section has been previously described in Verhoef (2005) and Tikoudis et al. (2015), with slight changes of notation. The previous framework developed in those papers to study first-best and second-best pricing is expanded to incorporate tradable permit schemes (Section 2.5).

The different agents of this model and their interactions are detailed in the remainder of this section.

2.1 Households

There are N households in the closed city. Households are treated as a continuum of entities, each solving a utility maximization problem, where the utility function at a given residential location z has the consumption of the composite good $y(z)$, housing space $s(z)$, and leisure time $l(z)$ as its arguments. We assume a CES utility function, with exogenous parameters δ_y, δ_s, δ_l and ρ. Thus, the utility at a given location z is:

$$U(z) = [(\delta_y y(z))^\rho + (\delta_s s(z))^\rho + (\delta_l l(z))^\rho]^{\frac{1}{\rho}} \tag{1}$$

The time constraint (Eq. 2) states that households spend their total time endowment T on commuting $t_c(z)$, working $t_w(z)$ and leisure $l(z)$.

$$T = t_c(z) + t_w(z) + l(z) \tag{2}$$

We assume that the working day has a fixed duration $\bar{t}_w$, so that the number of commuting trips varies in direct proportion with labor supply. Normalizing $\bar{t}_w = 1$, and considering that $t_c(z) = d_w(z)t_t(z)$, where $t_t(z)$ is the commuting time of one trip and $d_w(z)$ are the number of days worked, Eq. (2) yields:

$$T = d_w(z)(t_t(z) + 1) + l(z) \iff d_w(z) = \frac{T - l(z)}{1 + t_t(z)} \tag{3}$$

The budget constraint (Eq. 4) is expressed in terms of the net wage rate $\bar{w}(z)$, which is the gross wage w paid by the firm minus the labor tax τ_w and the monetary costs of commuting, which are defined in Section 2.5. The market price of the composite good is p, and the price per unit of land at location z is $r(z)$. The income of households includes the lump-sum transfers of aggregate excess land rents R, and of revenue obtained with the pre-existing labor tax G [2], which represents the public revenue constraint; plus the transfer $B(z)$ that results from the allocation of permits. All these transfers are described in Sections 2.4 and 2.5.

$$\bar{w}(z)\frac{T - l(z)}{1 + t_t(z)} + \frac{R}{N} + \frac{G}{N} + B(z) = py(z) + r(z)s(z) \tag{4}$$

It is useful to define $M(z)$ as the maximum income that can be realized when leisure time is chosen to be zero, therefore:

$$M(z) = \frac{R}{N} + \frac{G}{N} + B(z) + \bar{w}(z)\frac{T}{1 + t_t(z)} \tag{5}$$

Maximizing Eq. (1) conditional on z and subject to Eq. (4) yields the Marshallian demand functions for the composite good, housing space and leisure time (Eqs. 6 - 8), where we have defined $\chi = \rho/(\rho - 1)$ and where σ is the elasticity of substitution: $\sigma = 1/(1 - \rho)$.

$$y^*(z) = M(z)\frac{\left(p/\delta_y^\rho\right)^{-\sigma}}{(p/\delta_y)^\chi + (r(z)/\delta_s)^\chi + \left[\frac{\bar{w}(z)}{\delta_l(1 + t_t(z))}\right]^\chi} \tag{6}$$

---

[2] The lump-sum recycling of labor tax revenues of course looks like a silly overall policy. We consider it to simultaneously have an initial tax, to avoid complications from optimizing the size of the public good and from market failures stemming from its public good character, and to avoid wealth from leaking out of our urban economy.

$$s^*(z) = M(z) \frac{\left(r(z)/\delta_s^\rho\right)^{-\sigma}}{(p/\delta_y)^\chi + (r(z)/\delta_s)^\chi + \left[\frac{\bar{w}(z)}{\delta_l(1+t_t(z))}\right]^\chi} \qquad (7)$$

$$l^*(z) = M(z) \frac{\left[\frac{\bar{w}(z)}{\delta_l^\rho(1+t_t(z))}\right]^{-\sigma}}{(p/\delta_y)^\chi + (r(z)/\delta_s)^\chi + \left[\frac{\bar{w}(z)}{\delta_l(1+t_t(z))}\right]^\chi} \qquad (8)$$

In addition, from Eq. (7) we obtain an expression for the local population density $n(z)$:

$$n(z) = \frac{1}{s^*(z)} \qquad (9)$$

Substituting Eqs. (6) – (8) into the utility function (Eq. 1) we obtain the indirect utility function:

$$V(z) = M(z) \left[(p/\delta_y)^\chi + (r(z)/\delta_s)^\chi + \left[\frac{\bar{w}(z)}{\delta_l(1+t_t(z))}\right]^\chi \right]^{-1/\chi} \qquad (10)$$

In addition, the assumption of exogenous population yields:

$$N = \int_0^{\bar{z}} n(z)dz \qquad (11)$$

2.2 Firms

There is a representative competitive firm located at the CBD. The firm produces the total amount of composite good Q under constant returns to scale, with aggregate labor supply L as the only input:

$$Q = AL \qquad (12)$$

where A is the marginal productivity of labor and:

$$L = \int_0^{\bar{z}} n(z)d_w(z)dz \qquad (13)$$

Finally, from the zero profit condition, we obtain:

$$w = Ap \qquad (14)$$

2.3 Commuting time and congestion

Commuting for any given location takes place through a single road. Commuting time $t_t(z)$ depends on the distance from the CBD considering free-flow travel time and congestion. First, let consider the travel time per unit of distance $\bar{t}_t(\varepsilon)$ at a location ε along the road, which is equal to the sum of the free-flow travel time per unit of distance t_0 and the congestion delay caused by the aggregate traffic flow at ε, which consists in all the trips made by households located between ε and the city boundary $\bar{z}$. We assume a linear relation between aggregate traffic flow and $\bar{t}_t(\varepsilon)$, therefore:

$$\bar{t}_t(\varepsilon) = t_0 + t_1 \int_\varepsilon^{\bar{z}} n(\zeta)d_w(\zeta)d\zeta \qquad (15)$$

where t_1 is an exogenous sensitivity parameter. Integrating Eq. (15) over the interval $(0,z)$ yields the time per trip for a household located at z:

$$t_t(z) = t_0 z + t_1 \left(\int_0^z \varepsilon n(\varepsilon) d_w(\varepsilon) d\varepsilon + \int_z^{\bar{z}} z\, n(\varepsilon) d_w(\varepsilon) d\varepsilon \right) = t_0 z + t_1 \int_0^{\bar{z}} min\{z, \varepsilon\} n(\varepsilon) d_w(\varepsilon) d\varepsilon \quad (16)$$

Eq. (16) states that the effect of traffic flow on the travel time of a household located at z depends on whether that traffic is originated at a location closer or further from the CBD. Equivalently, an additional trip (hence an extra unit of labor supply) originated at z increases the travel time of a commuter located at $\xi \leq z$ by ξt_1, and the travel time of a commuter located at $\xi \geq z$ by $z\, t_1$. Considering that $n(\xi) d_w(\xi)$ commuters will be affected at location ξ, and that the shadow price of travel time at ξ is $\bar{w}(\xi)/(1 + t_t(\xi))$, we obtain the marginal external congestion cost generated by a household located at z:

$$mecc(z) = t_1 \int_0^{\bar{z}} min\{z, \xi\} n(\xi) d_w(\xi) \frac{\bar{w}(\xi)}{1 + t_t(\xi)} d\xi \quad (17)$$

Finally, we define the aggregate kilometers traveled *KM*:

$$KM = \int_0^{\bar{z}} n(z) z\, d_w(z) dz \quad (18)$$

the time lost due to congestion of a household located at z, $t_{lost}(z)$:

$$t_{lost}(z) = d_w(z) t_1 \int_0^{\bar{z}} min\{z, \varepsilon\} n(\varepsilon) d_w(\varepsilon) d\varepsilon = t_c(z) - d_w(z) z\, t_0 \quad (19)$$

and its correspondent aggregate variable:

$$T_{lost} = \int_0^{\bar{z}} n(z) t_{lost}(z) dz = T_c - t_0 KM \quad (20)$$

where $T_c = \int_0^{\bar{z}} n(z) t_c(z) dz$.

2.4 Public budget and financial transfers

In this subsection, the financial transfers that each household receives from the social planner are described. These transfers are the aggregate excess land rents R/N, and recycling of public tax revenue G/N. Besides, for each method of allocation of permits, a term equivalent to a labor-supply-independent transfer appears. This term will be defined in Section 2.5.

2.4.1 Aggregate excess land rents

The behavior of the land suppliers is not modeled explicitly. Instead, we assume that the social planner acquires land from an absentee landlord at a price equal to the agricultural land rent, r_A (representing the opportunity cost of land), and sells every unit of land in a price equal to the willingness to pay for it, returning the total surplus (i.e. the aggregate excess land rents) to the population in a lump-sum manner.

This guarantees that the city cannot expand without cost, that there is no additional failure in the land market and that the excess rents remain within the economy. Therefore, the aggregate excess land rents are:

$$R = \int_0^{\bar{z}} (r(z) - r_A)\, dz \quad (21)$$

6

2.4.2 *Public revenue*

We consider that there is a public revenue constraint, i.e. there is a fixed revenue that needs to be raised from the economy. This can be seen as the necessity of indirectly financing some public good (although not explicitly included in the model). This formulation aims to represent that pre-exsiting labor taxes cannot be reduced without replacing them with revenue from another source. This fixed revenue G is returned in a lump-sum manner to households. G is defined by:

$$G = \int_0^{\bar{z}} n^b(z) d_w^b(z) \tau_w^b dz \tag{22}$$

where the superscript b denote the variables of the base case, i.e., the case with a labor tax of τ_w^b and no additional taxes or policies.

When a tradable permit scheme is introduced, labor supply varies due to the changes in the net wage rate and non-labor income induced by the scheme. Therefore, the labor tax needs to be adjusted to meet the public revenue constraint. If L^t is the aggregated labor supply after the introduction of a tradable permit scheme, the new labor tax τ_w^t is:

$$\tau_w^t = \frac{G}{L^t} \tag{23}$$

2.5 *Tradable permit schemes*

In the tradable permit schemes studied in this paper, a commuter needs to spend one permit per each kilometer traveled. Therefore, the Pigouvian effect of the schemes is equivalent to that of a flat, space-independent kilometer charge. We assume that permits can be freely traded between drivers (households), that every desired transaction takes place with no transaction costs, and that the market always clears, assumptions that we in fact make for all other markets in this general-equilibrium framework.

The social planner is in charge of designing the tradable permits, aiming to maximize welfare. For that, the planner decides the targeted quantity and the allocation of permits to households. The targeted quantity naturally corresponds to aggregate kilometers traveled given the nature of the permits. For the initial allocation of permits to households, we consider three different methods, differentiated by the spatial variable that determines the number of permits that a household receives: i) household's residential distance from the CBD; ii) household's labor supply, which is determined by number of days worked (this is equivalent to number of commuting trips); and iii) household's kilometers traveled, which is the household's residential distance from the CBD multiplied by its labor supply. These variables are chosen because we consider them to be the most natural options to determine the allocation of permits in a practical implementation.

Denoting the allocation of permits to households as $\theta(f(z))$, where $f(z)$ is one of the three relevant variables listed above, the assumption that the market always clears yield:

$$\int_0^{\bar{z}} n(z) \theta(f(z)) dz = KM \tag{24}$$

For each of the three allocation methods, we consider that the number of permits allocated to a household depends linearly on the relevant variable, while allowing for a fixed term. This means that

two parameters define the allocation: the intercept α and the slope of the linear relation β, therefore $\theta(z) = \alpha + \beta f(z)$. From Eq. (24) it is possible to obtain α as a function of β and KM for each different $f(z)$. Hence, the two variables to be optimized to maximize welfare are KM (the targeted quantity) and β, which defines the differentiation of the allocation of permits.

The endogenous equilibrium price of one permit is denoted as π. Thus, each kilometer traveled by a commuter has a cost of π: the cost of buying a permit per each kilometer that exceeds his/her allocation, or the opportunity cost of not selling a permit and using it instead.

The three tradable permit schemes are described in the remainder of this subsection.

2.5.1 *Allocation by distance from the CBD*

We impose a non-negative allocation of permits, and we denote as z_0 the location of the city where the allocation of permits becomes zero:

$$\alpha + \beta z_0 = 0 \tag{25}$$

which is relevant if $z_0 \in (0, \bar{z})$. For $\beta > 0$, the allocation of permits is defined by:

$$\theta(z) = \begin{cases} \alpha + \beta z, & \text{if } z \geq \hat{z} \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

where:

$$\hat{z} = Max\{0, z_0\} \tag{27}$$

Finally, from Eqs. (24) and (26) we obtain:

$$\alpha = \left(KM - \beta \int_{\hat{z}}^{\bar{z}} z\, n(z) dz \right) \left(\int_{\hat{z}}^{\bar{z}} n(z) dz \right)^{-1} \tag{28}$$

For $\beta < 0$, the allocation of permits is defined by:

$$\theta(z) = \begin{cases} \alpha + \beta z, & \text{if } z \leq \hat{z} \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

where:

$$\hat{z} = min\{z_0, \bar{z}\} \tag{30}$$

From Eqs. (24) and (29) we obtain:

$$\alpha = \left(KM - \beta \int_{0}^{\hat{z}} z\, n(z) dz \right) \left(\int_{0}^{\hat{z}} n(z) dz \right)^{-1} \tag{31}$$

Let consider a household that receives a positive number of permits. Then, the budget constraint (Eq. 4) becomes:

$$(w - \tau_w)\frac{T - l(z)}{1 + t_t(z)} + \frac{R}{N} + \frac{G}{N} + \underbrace{\pi\left(\alpha + \beta z - z d_w(z)\right)}_{\text{Extra residual income from permit market}} = py(z) + r(z)s(z) \tag{32}$$

8

where $zd_w(z)$ are the kilometers traveled by the household. It can be seen that if a household uses less permits (i.e. travels less kilometers) than its initial allocation, it obtains an extra residual income from the permit market; and when a household's use of permits exceeds its initial allocation, the household has a net cost of buying permits. Rearranging and using Eq. (3) we obtain:

$$\bar{w}(z)\frac{T-l(z)}{1+t_t(z)} + \frac{R}{N} + \frac{G}{N} + B(z) = py(z) + r(z)s(z) \tag{33}$$

where:

$$\bar{w}(z) = w - \tau_w - \pi z \tag{34}$$

and:

$$B(z) = \pi(\alpha + \beta z) \tag{35}$$

Therefore, $\bar{w}(z)$ is the net wage rate. The term πz present in the net wage rate represents the effect of the permits inducing a cost per kilometer traveled, although it is not a classic tax. Therefore, we denote as *total equivalent tax* the difference between the gross and net wages: $\tau_w + \pi z$. Consequently, this scheme has both a Pigouvian and a distortionary effect, because it increases the cost of commuting, thus affecting the net wage rate. The Pigouvian effect is equivalent to that of a flat kilometer road charge. With this scheme is not possible to use revenues to diminish the labor tax, therefore the revenue recycling effect is not present.

The term $B(z)$ is equivalent to a labor-supply-independent income transfer. The expected sign and magnitude of the optimal slope β are not straightforward to determine, due to the different spatial effects that arise. Our numerical results indeed reveal that the sign of β may vary depending on the initial circumstances. As Tikoudis et al. (2015) showed, the optimal total tax profile (that is, road tax plus labor tax) in presence of a public revenue constraint is the sum of a Pigouvian component plus an extra tax necessary to meet the constraint, which is differentiated across space. This is the result of the Ramsey and Mirrlees forces, which add a spatial component to the optimization of that extra tax, affected by the spatial variation of the marginal utility of income and of the elasticity of labor supply.

For this permit scheme, due to the flat price per kilometer imposed and the absence of revenue recycling effect, we expect that households will be facing a total tax that differs from the second-best optimum. The magnitude of the welfare losses generated by that difference and the effect of using the financial transfer to alleviate them will also be influenced by the spatial forces mentioned above. In addition, the interpretation of $B(z)$ is further complicated by the fact that it has a direct effect on only one margin of behaviour: location choice. The impact on labor supply and commuting behaviour, where the distortions primarily arise, is only indirect: via the induced change of location. The impossibility of fully identifying analytically the different effects makes it required to employ numerical results in the assessment of the sign and magnitude of β.

Finally, we define the *equivalent non-labor income*, as the households' income that comes from sources different from labor, and accounting for the labor-supply-independent transfer that appears from the allocation of permits. It follows that, under this scheme, the equivalent non-labor income is $R/N + G/N + B(z)$, and therefore, is differentiated across space.

2.5.2 *Allocation by household's labor supply*

The allocation of permits on the basis of labor supply is defined by:

$$\theta(z) = \alpha + \beta d_w(z) \tag{36}$$

From Eqs.(11), (24) and (36) we obtain:

$$\alpha = \frac{KM - \beta L}{N} \tag{37}$$

and re-writing the budget constraint (analogous to Eqs. 32 and 33) we obtain:

$$\bar{w}(z) = w - \tau_w - \pi z + \pi \beta \tag{38}$$

and:

$$B(z) = B = \pi \alpha \tag{39}$$

where it is straightforward to see that this scheme implies a *equivalent non-labor income* that is constant across space. In addition, the term B (Eq. 39) can be equivalent to a transfer or to a tax, depending on the sign of α.

From Eq. (38), it follows that the *total equivalent tax* is $\tau_w + \pi z - \pi \beta$. Hence, this permit scheme has a Pigouvian and a distortionary effect, just as the scheme by distance described in Section 2.5.1, but in addition it has a revenue recycling effect (if $\beta > 0$), due to the incentive to labor supply provided by the term $\beta \pi$. This captures the fact that, if $\beta > 0$, an increase in labor supply by a household directly increases the number of permits that the household receives.

By rearranging the last two terms of Eq. (38), we obtain the term $-\pi(z - \beta)$, which means that an equivalent way of seeing this scheme is considering that, for those locations $z > \beta$, the scheme translates into an additional tax, and for $z < \beta$ it translates into a tax cut. Hence, the parameter β represents the location of the city where this scheme goes from having a dominating revenue recycling effect to having a dominating Pigouvian effect (besides the tax interaction effect).

It is important to note that if $\alpha = 0$, from Eq. (37) we obtain $\beta = KM/L$, which implies $\bar{w}(z) = w - (\tau_w - (KM/L)\pi) - \pi z$. Thus, the net wage rate of such a permit scheme and of a flat kilometer road charge are equivalent (this is shown in Appendix A). As a consequence, a non-zero value of α implies a deviation of the equivalent labor tax cut (the term $\beta \pi$) from its level under a flat kilometer pricing. Therefore, allowing for $\alpha \neq 0$ means that this scheme will result in a utility at least equal to that of a flat kilometer pricing scheme.

2.5.3 *Allocation by kilometers traveled*

The allocation of permits in proportion to the household's kilometers traveled is defined by:

$$\theta(z) = \alpha + \beta z d_w(z) \tag{40}$$

From Eqs.(24), (18) and (40) we obtain:

$$\alpha = \frac{KM(1 - \beta)}{N} \tag{41}$$

and re-writing the budget constraint (analogous to Eqs. 32 and 33) we obtain:

$$\bar{w}(z) = w - \tau_w - z\pi(1 - \beta) \tag{42}$$

and

$$B(z) = B = \alpha \pi \tag{43}$$

where we can see that the *equivalent non-labor income* effect of this scheme is constant across space.

The fact that permits are allocated on the basis of kilometers traveled, which are the household's labor supply multiplied by its residential distance, could lead to the wrong intuition that this scheme provides an incentive to increase labor supply, just as the scheme described in Section 2.5.2. But from Eq. (42), it follows that the *total equivalent tax* of this scheme is $\tau_w + z\pi(1-\beta)$. As $\beta < 1$ (otherwise each household would receive more permits than its kilometers traveled, so the equilibrium price would be zero), the net effect of this scheme on the wage rate is equivalent to that of a flat kilometer charge, with a price per kilometer equal to $\pi(1-\beta)$, thus not providing incentive to increase labor supply. This is because with this scheme, a marginal increase in household's labor supply implies a marginal monetary increase of $z\beta\pi$ due to the increase in permits received, and, at the same time, a marginal increase in commuting costs of $z\pi$.

Let define $\hat{\pi} = \pi(1-\beta)$. From Eqs. (41) - (43), it is obtained that the effect of this scheme, and consequently the optimal scheme with allocation by kilometers traveled, are completely determined by $\hat{\pi}$. This means that for each given β, it is possible to obtain a price per kilometer π that leads to the maximum welfare given the nature of the scheme, and vice-versa. Thus, it is not possible to identify a unique optimal allocation of permits. This is the consequence of allocating permits linearly on kilometers traveled, where if a household travels an additional kilometer, it receives β more permits, and at the same time having a flat kilometer price of permits, being both values decided by the social planner (the price is endogenous but determined through the setting of the targeted aggregate kilometers traveled). An alternative way of seeing this is to look at the extra revenue (or cost) derived from the permit market: $\pi(\theta(z) - zd_w(z))$. Considering Eqs. (40) and (41), the extra revenue is $\pi(1-\beta)(KM/N - zd_w(z))$, showing again that the effect of this scheme depends on $\hat{\pi}$.

From Eq. (42), it follows that $\hat{\pi}$ is the price per kilometer that the scheme imposes, and from Eqs. (41) and (43) we obtain that the equivalent financial transfer is $\hat{\pi}(KM/N)$. Therefore, this scheme is equivalent to a flat kilometer pricing with the revenues returned in a lump-sum manner (the equivalence is shown in Appendix A). As shown by Tikoudis et al. (2015), the absence of a revenue recycling effect implies that a flat kilometer road pricing is likely not to be welfare-increasing for relatively high levels of labor tax, thus, the same is expected for this permit scheme.

Finally, we note that when $\alpha = 0$, β must be equal to 1 in order to meet the equilibrium condition of Eq. (24), therefore, as each household would receive the same number of permits that it chooses to use, the permit scheme would have no effect. This makes the presence of the intercept α crucial for the efficiency of the scheme.

2.6 Equilibrium

In equilibrium, the spatial derivative of $V(z)$ must be equal to zero for all households. Using Eq. (5) and defining $D(z) = -\frac{d}{dz}\frac{\bar{w}(z)}{1+t_t(z)}$ yields a differential equation for $r(z)$:

$$\frac{dr(z)}{dz} = r(z)^{1-\chi}\left\{ \left(\frac{\delta_s}{\delta_l}\right)^\chi \left(\frac{\bar{w}(z)}{1+t_t(z)}\right)^{\chi-1} D(z) + \frac{\left(\frac{dB(z)}{dz} - D(z)T\right)\delta_s^\chi}{M(z)} + \right.$$
$$\left. (p/\delta_y)^\chi + (r(z)/\delta_s)^\chi + \left[\frac{\bar{w}(z)}{1+t_t(z)}\right]^\chi \right\}$$

(44)

with terminal condition:

$$r(\bar{z}) = r_A \tag{45}$$

where r_A is the exogenous agricultural rent.

The exogenous parameters are the utility parameters δ_y, δ_s, δ_f and ρ, the total time endowment T, the total population N, the marginal productivity of labour A, the travel time parameters t_0 and t_1 and the agricultural rent r_A. The price of the composite good p is used as a numeraire and is set equal to 1.

There is one additional condition that states that the monetary value transferred to the absentee landlord must be equal to the value of the difference between production and consumption:

$$p(Q - Y) = \bar{z}r_A \tag{46}$$

where we defined the aggregate demand for composite good Y:

$$Y = \int_0^{\bar{z}} n(z)y^*(z)dz \tag{47}$$

We do not need to include Eq. (46) in the system of equations, since it is a direct result of combining the zero profit condition of firms (Eq. 14) and the budget constraint (Eq. 4).

Therefore, the following unknowns determine the equilibrium: the spatial functions V, M, $\bar{w}$, y, s, l, r, n and B, and the variables w, π, τ_w, $\bar{z}$, Q, L, Y, R and KM. Since there are no insightful closed-form solutions, we numerically solve the equilibrium conditions described in this section. The results are shown and analysed in Section 3 for the base case and road pricing benchmarks, and in Section 4 for the tradable permit schemes.

3. Base equilibria and optimal road charges

3.1 Base equilibria: numerical calibration and description

We use the calibrated parameters obtained by Verhoef (2005) for studying road pricing schemes without a pre-existing labor tax. These parameters are $A = 1$, $r_A = 1$, $T = 1$, $N = 1000$, the travel time parameters $(t_0, t_1) = (3x10^{-3}, 5x10^{-5})$, and the utility parameters $(\delta_y, \delta_s, \delta_f) = (25, 64.9275, 1.77225)$.

The numerical model was written in Mathematica 12.1. It finds the spatial equilibrium by using a repeated nested approach, in which the various markets of interest are successively brought into equilibrium while keeping other prices fixed, until convergence is reached (the convergence criterion used in the different loops was set at $1x10^{-7}$ for relative changes in key variables between successive iterations within each loop). This procedure follows Verhoef (2005).

In order to analyse the tradable permit schemes in more detail, we will focus on two cases defined by the initial level of labor tax. One with $\tau_w = 0.15$, which we denote as the low labor tax case, and the other with $\tau_w = 0.35$, the high labor tax case.

In the low labor tax case, there is an increase in utility[3] of 0.13% with respect to the free-market case, which means that the Pigouvian effect of the decrease in labor supply induced by the labor tax is

---

[3] The terms utility and welfare are interchangeable in this formulation, and we use both throughout the paper

12

higher than its distortionary effect[4]. In the high labor tax case, there is a decrease in utility of 0.15%, therefore, the distortionary effect of the labor tax exceeds its Pigouvian effect.

In the low labor tax base case, households work on average 22.5% of their time endowment. If we consider a week of 112 hours (16 hours per day, which excludes 8 hours per day that are assigned to activities not modeled, such as sleep) households work 25.2 hours per week on average. This average encompasses holidays, part-time workers and other situations that could diminish the average labor supply.

Households spend on average 7.7% of their time in commuting trips, which means around 8.6 hours per week. The total time lost due to congestion is 5.39% (around 6 hours per week per household on average). The utility of a case with the same level of labor tax and without congestion ($t_1 = 0$) is 6.15% higher than the base case, which translates into an average equivalent variation (EV) of 22.5%[5]. Labor income is on average 74.5% of total income (the reminder is excess land rent and recycling of revenues from labor tax) and households spend on average 57.9% of their total income in goods consumption and 42.1% on housing related consumption.

For a higher labor tax base case, a 20.9% of the time endowment of households is spent on labor, 6.6% on commuting and 72.5% on leisure. Labor income is 58.8% of total income and households spend on average 58.9 % of their total income in goods and 41.1% in housing related consumption. The total time lost due to congestion is 4.5% and the increase in utility of a case without congestion is 5.68%, with an EV of 24.7%.

3.2 Optimal road charges

We consider two relevant charges as benchmarks for the efficiency of the tradable permit schemes: the unconstrained first-best and a second-best tax constrained by the necessity of collecting the public revenue.

3.2.1 *Unconstrained first-best charge*

The unconstrained first-best charge drops the tax revenue constraint and consequently consists in the complete elimination of the pre-existing labor tax, and simultaneously the implementation of a Pigouvian road charge (i.e, equal to the marginal external congestion cost per location, Eq. 17), with the revenue recycled as a lump-sum transfer. We use it as a benchmark for the model, but not as a viable policy alternative as we treat the revenue constraint as 'hard'.

The increase in utility with respect to the base case, for the low and high labor tax cases are 0.065% (EV = 0.25%) and 0.347% (EV = 1.5%), respectively. The larger increase of the latter case is due to the fact that in the first-best charge the labor tax is always eliminated (hence the resulting utility does not vary with the level of labor tax), therefore the variation in utility depends entirely on the base case utility, which is lower for the higher level of labor tax.

---

[4] In fact, the optimal labor tax in absence of additional policies is $\tau_w = 0.162$.

[5] The equivalent variation of a policy represents the amount of income that a household needs in order to reach a variation in utility equivalent to the variation induced by the policy. Throughout this paper, we express the average equivalent variation (EV) as a percentage of labor income $\bar{w}d_w(z)$ in the base case.

3.2.2 *Approximate second-best tax*

It is also relevant to find the second-best tax given the public revenue constraint. We consider that the collected revenues are used to finance a labor tax cut, as Parry and Bento (2001) and Tikoudis et al. (2015) did, showing the existence of a weak double dividend of road charges; that is, using tax revenues to finance a labor tax cut is more efficient than recycling revenues through a lump-sum transfer. The revenue collected, G_p is:

$$G_p = \int_0^{\bar{z}} n(z) d_w(z) \tau(z) dz \tag{48}$$

where $\tau(z)$ is the tax charged to a commuter located at z. Hence, the resulting labor tax is:

$$\tau_w = \frac{G - G_p}{L} \tag{49}$$

For obtaining the second-best tax, which can vary over space and unfortunately cannot be solved analytically, we used a quadratic interpolating polynomial to make a numerical approximation of the optimal profile, where we maximize utility as a function of the coefficients of the polynomial using a search grid method. The consideration of a quadratic polynomial allows to capture possible non-monotonic effects. Higher-degree polynomial were also tested, with non significant improvements in utility.

For the low labor tax case, the second-best tax turns out to be almost equal to a Pigouvian charge (the increase in utility of the second-best tax with respect to a Pigouvian charge is lower than our convergence criterion). This is because the revenues raised by the Pigouvian charge allows to cut the labor tax up until 0.018% of the wage rate, making its distortionary effect almost negligible.

For the high labor tax case, a Pigouvian tax generates a welfare loss of 0.016% compared with the base case, whereas the approximate second-best tax generates an increase in utility of 0.026% (EV = 0.11%).

In Fig. 1 it is shown that the approximate second-best tax is non-monotonic across space. This is the result of the interaction between the two forces that determines optimization of the tax profile: the Ramsey component, which states that taxes should be higher where elasticity of labor supply is lower, and the Mirrlees component, which states that taxes should be higher where the marginal utility of income is lower. Therefore, the drop of the tax close to the city fringe is the result of the higher marginal utility of income there, as well as the higher elasticity of labor supply, as can be seen in Appendix B for the base case. A complete analysis and discussion of the Ramsey and Mirrlees forces and the resulting optimal tax in a monocentric city can be found in Tikoudis et al. (2015).
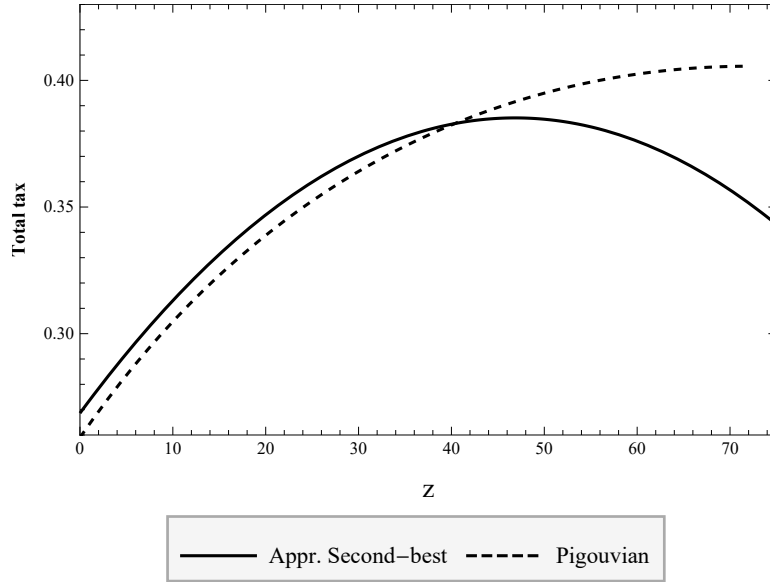
Fig. 1. Total tax per location for the approximate second-best and Pigouvian charges. High labor tax case.

4. Tradable permit schemes: numerical results and policy analyses

4.1 Resulting utility and optimal schemes

For obtaining the optimum tradable permit schemes, as no closed-form solutions exist, we use heuristic search grid methods. For the allocations by distance and labor supply, we obtained the optimal pair (β, π) by constructing a 3D interpolation of the resulting utility as a function of these two variables. It is important to note that we can optimize the permit price π even though it is endogenous: it reflects the price that would result from the true policy variable chosen by the social planner: KM. Therefore, this procedure is equivalent to determine the optimal pair (β, KM), but more efficient in terms of computing time. An analogous procedure using a 2D interpolation was used to find the optimum $\hat{\pi}$ for the allocation by kilometers traveled.

All interpolated functions resulted to be concave. This allows to obtain the optimal points in each case, which were verified by varying the starting points of the search method, by making the grid denser at the region close to the predicted optimum, and by numerically computing derivatives at points in that region. All optimal points were robust to these verifications.

In Fig. 2 we show the utility of each tradable permit scheme and of some benchmarks as a function of the initial level of pre-existing labor tax. The higher efficiency of the allocation by labor supply is evident: its utility is 99.99% of the first-best utility, and the variation of utility as a function of the initial labor tax is lower than 10^{-5}. When compared with the second-best tax, this permit scheme is more efficient starting around a labor tax of 20%. For lower labor taxes, the second-best tax has the decisive advantage of spatial differentiation of the price per kilometer. For higher labor taxes, the permit scheme wins because the more sophisticated permit allocation scheme brings increasingly important welfare gains. In addition, the allocation by labor supply is the only welfare-increasing permit scheme for levels of labor tax around 27% or higher, which is a consequence of its unique property of providing a direct incentive to increase labor supply.

As shown in Fig. 3, the high efficiency of the permits allocated by labor supply is acquired by the adjustment of the parameter β, while the variation in equilibrium price remains very small, with a
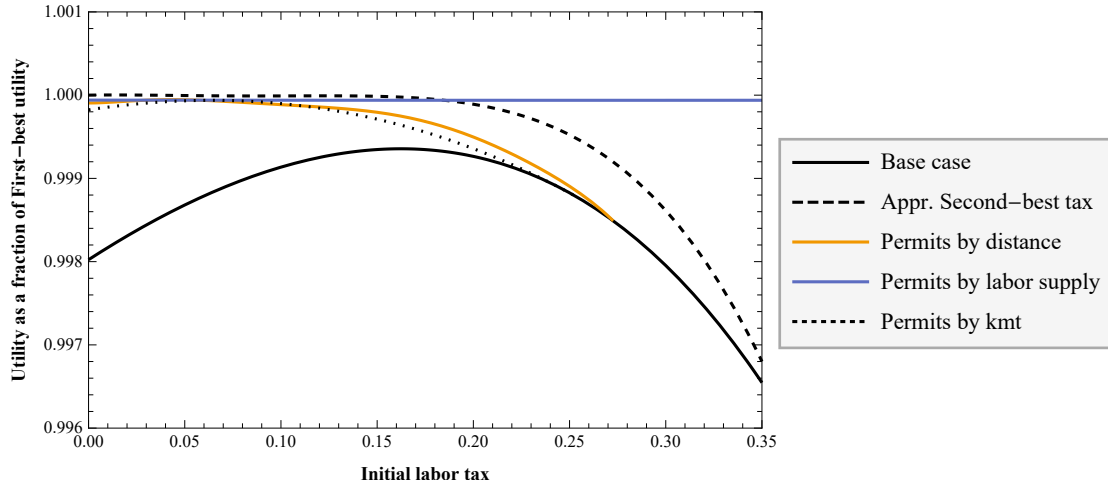
15

Fig. 2. Utility of different schemes by level of initial labor tax.

maximum difference of 3% between the highest and lowest price per kilometer. This corresponds with a maximum difference of around 0.5% of the gross wage (for the household located at the city fringe). Therefore, when varying the level of labor tax, the Pigouvian effect of the scheme remains almost constant, while the revenue recycling effect is optimized to maximize the efficiency of the policy.
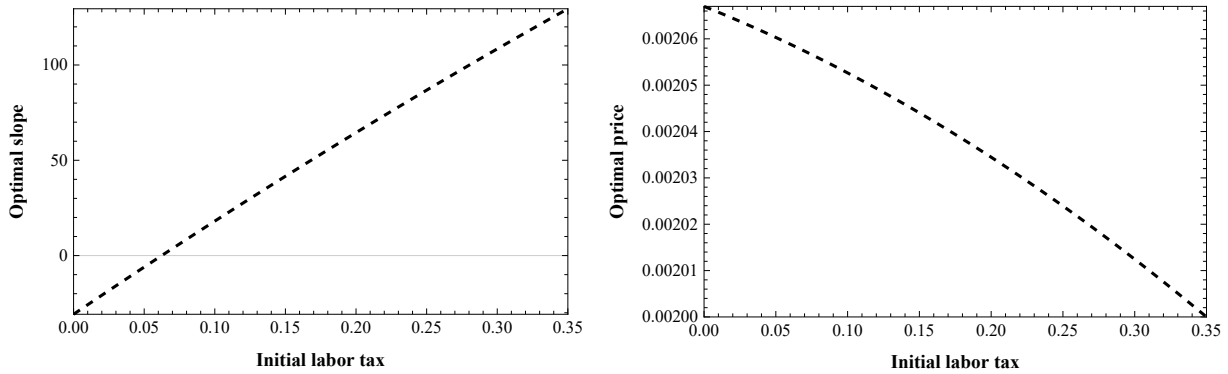

Fig. 3. Optimal slope and price of permits allocated by labor supply as a function of initial labor tax.

In the left panel of Fig. 3 it can be seen that the optimal slope β is negative for lower levels of labor tax, which means that the scheme has a dominating Pigouvian effect over the entire city, as discussed in Section 2.5.2. This is consistent with the fact that when labor tax is low, the revenue recycling effect is less important for the optimization of the schemes.

In addition, a negative value of β at those levels of labor tax, in combination with the price per kilometer obtained, means that, when compared with an optimal taxation, households closer to the CBD are overtaxed and households further from the CBD are undertaxed, as shown in Fig 4, where the equivalent tax of the permit scheme is compared with a first-best charge for a zero labor tax. The permit scheme is not able to reproduce exactly the optimal taxation, due to being constrained to a flat price per kilometer. Therefore, which households are overtaxed as a result of the implementation of the permit scheme is relevant for its optimality, that is determined by all the forces in place in this economy: the Ramsey and Mirrlees forces, as well as all the general-equilibrium effects of the taxation. Thus, the optimization of the spatial pattern of the overtaxation results in a combination of values that, for lower levels of labor tax, results in a negative β. In fact, the same resulting spatial

16

pattern of over and undertaxation is obtained for higher levels of labor taxes (as shown in Fig. 9 in Section 4.2, and Fig. 10 in Section 4.3), but in that cases requiring positive values of β.
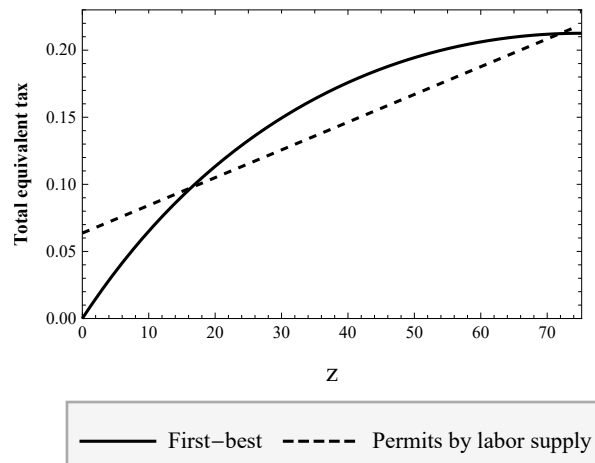

Fig. 4. Total equivalent tax per location of tradable permits when allocated by labor supply compared with first-best taxation for a zero labor tax.

By combining Eqs. (37) and (39) it is obtained that for the permits allocated by labor supply, the parameter β has a negative effect on B, which could make a higher β less desirable. However, this negative effect is compensated by the higher public revenue transfer when labor tax increases, and also higher excess land rents, such as the total equivalent non-labor income (that is, $B + G/N + R/N$) has a small variation with labor tax (the maximum difference is 0.3%).

In the case of permits allocated by distance from the CBD, the absence of revenue recycling effect implies that the scheme is not welfare-increasing (the optimal price is zero) for high levels of labor tax. The optimal β and price by level of labor tax are shown in Fig. 5, where it can be seen that both decrease with labor tax and that the optimal price falls to zero for a labor tax of around 0.27. The decrease in the optimal price is a natural consequence of the absence of a labor tax cut, which causes the taxation of road use to have a negative impact on welfare.
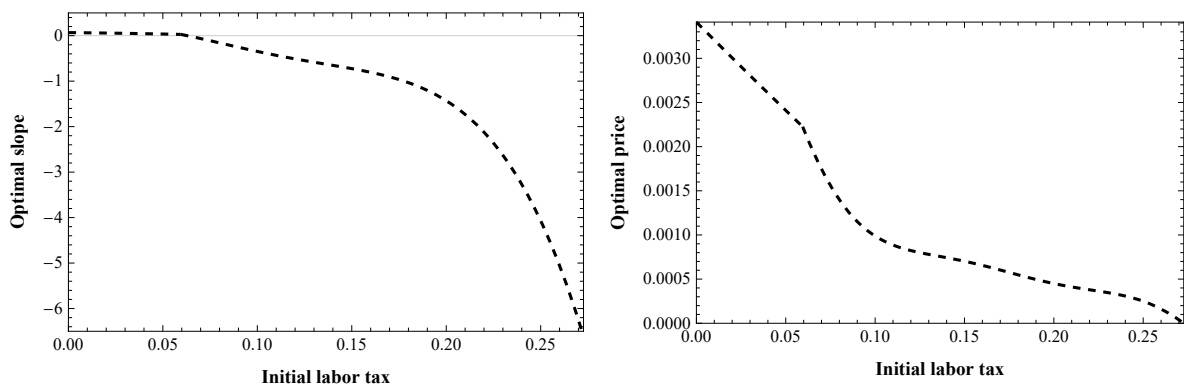

Fig. 5. Optimal slope and price of permits allocated by distance as a function of initial labor tax.

The decrease in the optimal β, which is the slope of the allocation as a function of distance from the CBD, even changing its sign (its highest value is 0.064, when labor tax is zero) indicates a relation between the optimization of β and the pattern of overtaxation compared with the optimal charge, similar to the relation previously explained for the scheme with permits allocated by labor supply, but with a different margin of behavior affected, which is in this case is location. In Fig.6, we compare

the optimal and the total equivalent tax obtained with this permit scheme, for an initial value of labor tax equal to 0 in the left panel (where the optimal tax is equal to a Pigouvian tax), which results in a positive β; and for a labor tax of 0.15 in the right panel (where the optimal tax is the approximate second-best tax), which results in a negative β. We can see that households are facing a higher or lower charge than the optimum depending on the level of labor tax. When there is a zero labor tax, households further from the CBD face a higher charge than in the optimum. Therefore, the optimal slope β is positive, acting as a financial compensation by allocating more permits to those households: it is a way to bring budget to households whose marginal utility of income has increased due to the overtaxation. The same happens for households closer to the CBD when labor tax is 0.15, and to a greater extent because the overtaxation is higher in this case, which means that the optimal β is negative and larger in magnitude.
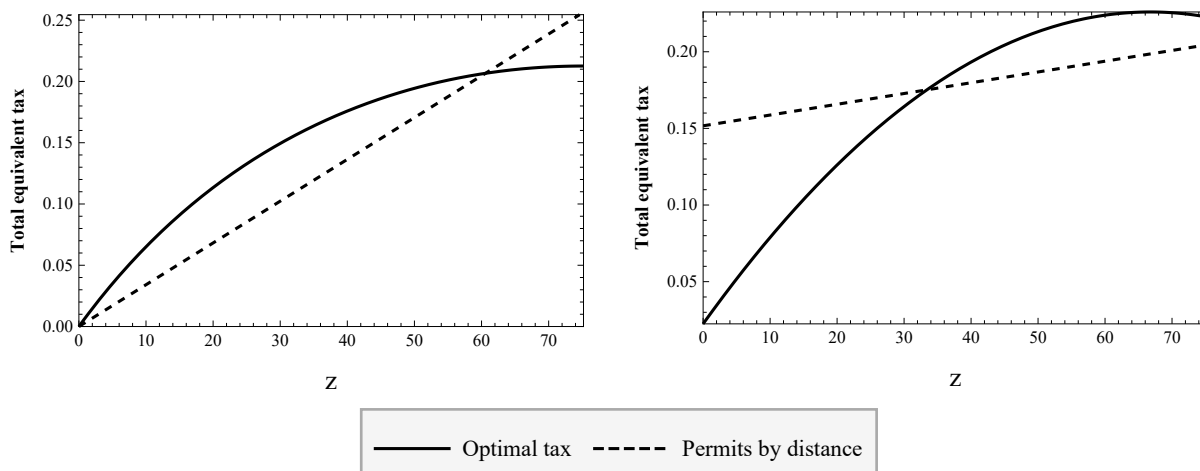

Fig. 6. Total equivalent tax per location of tradable permits when allocated by distance from the CDB compared with optimal taxation. Left: Zero labor tax. Right: Labor tax 0.15.

Finally, another insightful result is the difference in utility between permits allocated by distance versus by kilometers traveled. Both schemes do not have a revenue recycling effect, which results in an optimal price decreasing with labor tax (the optimal price as a function of labor tax with permits allocated by kilometers traveled is shown in Appendix C). The fact that permits allocated by distance reach a higher utility is explained by the spatial differentiation of the labor-supply-independent transfer, that is not present when permits are allocated by kilometers traveled, as shown in Sections 2.5.1 and 2.5.3.

Now, in order to provide more detailed insights about the effect of the permit schemes, we focus on two cases characterized by the initial level of labor tax: the low labor tax case (15%) and the high labor tax case (35%).

4.2 Low labor tax case (15%)

Fig. 7 shows for the two allocations of permits by distance and by labor supply the resulting spatial allocations of permits, where it is important to note that both are shown as a function of distance from the CBD, even though the allocation by labor supply is not directly determined by this variable. The resulting parameters for the allocation by distance are $\alpha = 27.0$ and $\beta = -0.7249$. The combination of these two values implies that the spatially differentiated transfer associated with the allocation of permits (defined in Eq. 35) decreases with distance, becoming zero at $\hat{z} = 37.3$. This means it would be optimal to allocate 'negative' permits (that is, to apply a tax) to households located further from the CBD. This is the consequence of the necessity of a higher allocation to households closer to the

18

CBD (as we explained in Section 4.1) and the assumed linear relation between number of permits and distance.
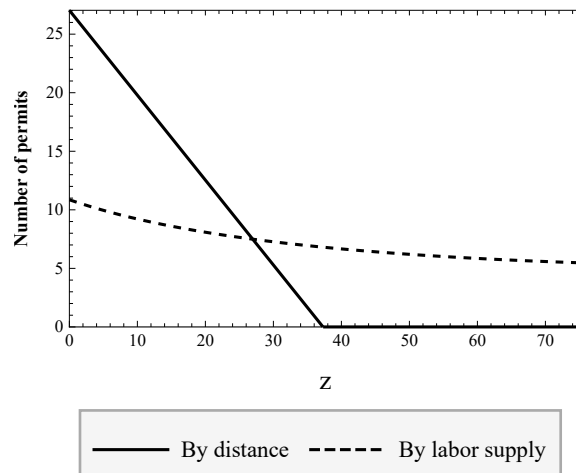

Fig. 7. Number of permits received by households per location, for optimized allocations by distance from the CBD and by labor supply. Low labor tax case.

The optimized variables for the allocation by labor supply are $\alpha = -2.03$ and $\beta = 41.60$. From Eq. (39), we obtain that the lump-sum transfer is negative as a result of the high value of β (as shown by Eq. 37). This highlights the importance of the revenue recycling effect in this city: as described in Section 2.5.2, β represents the location of the city where this scheme goes from having a dominant revenue recycling effect to having a dominant Pigouvian effect, therefore, a higher β implies that the revenue recycling effect is needed in a larger portion of the city.
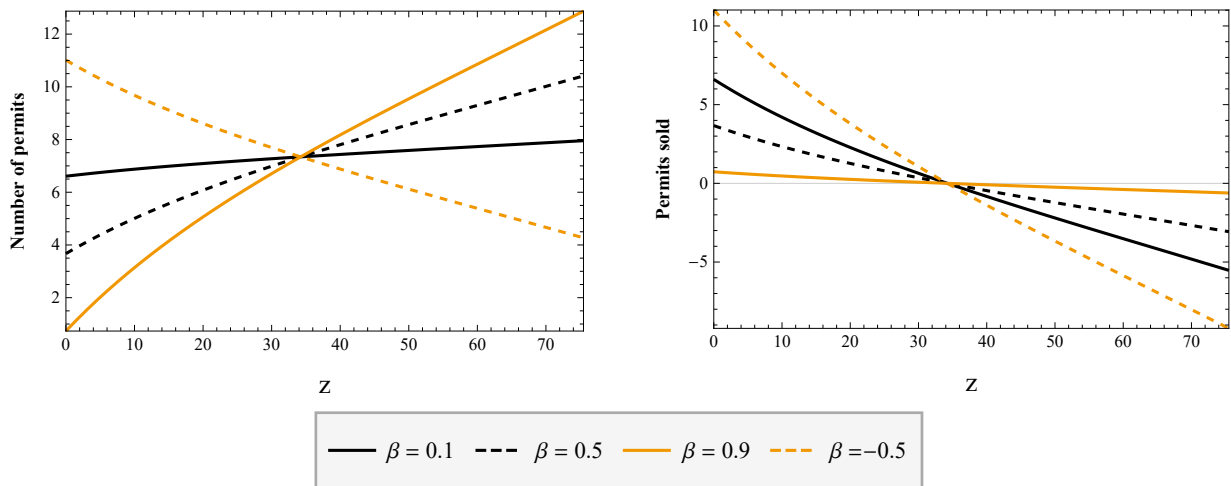

Fig. 8. Equivalent allocations of permits by kilometers traveled for different values of β. Left: Number of permits per location. Right: Resulting permit market.

In the scheme with permits allocated by kilometers traveled, the optimum is defined by $\hat{\pi} = 0.0011$. Equivalent optimal allocations by kilometers traveled are shown in the left panel of Fig. 8, for different values of β. All the allocations shown in Fig. 8 result in the same equilibrium: when varying β, the equilibrium permit price is adjusted to keep $\hat{\pi}$ at its optimal level. For comparison, when $\beta = 0.9$, the price of each permit is 0.0106, and when $\beta = -0.5$, the price is 0.0007. As explained in Section 2.5.3, this is due to having a flat kilometer price of permits and at the same time allocating permits linearly on kilometers traveled.

19

The equivalent effect of the different allocations by kilometers traveled is also reflected in the permit market. In the right panel of Fig. 8 we show the number of permits sold at each location for different values of β. This, therefore, gives the difference between the permits received and the kilometers traveled by a household at that location (a negative value indicates that the household needs to buy permits). It can be seen that households that sell permits are the same for the different allocations (the location of the city where households sell zero permits is determined by the equation $zd_w(z) = KM/N$). As explained is Section 2.5.3, the resulting optimal price for each level of β varies in order to keep the revenue (or cost) that a household obtains from the permit market constant for that household.

In Table 1 we show the charges related with each policy, the resulting variables and welfare measures, where we include a flat kilometer charge with labor tax cut (described in Appendix A) as an additional benchmark. All equilibrium prices per kilometer are different from zero, and therefore, all policies studied yield an increase in utility. The policies that acquire a higher increase in utility are the ones that have a revenue recycling effect (the second-best tax, the flat kilometer charge and the permits allocated by labor supply). The presence of a revenue recycling effect allows these policies to have a higher charge per kilometer, therefore having a higher Pigouvian effect while diminishing the distortionary effect through recycling (or by incentivizing labor, in the case of the permit scheme).

In addition, it can be seen that all policies acquire a somewhat modest EV, with the EV of the permit schemes ranging from 0.13% to 0.21%. This is partly due to the assumed congestion technology, where the linear congestion function depresses the net efficiency gains of pricing, and the static nature excludes departure time adjustments as an important behavioural margin for the optimization of congestion. But it is also because the base case is already a good improvement when compared with the free-market case, as shown in Fig. 2.

In order to determine the relative efficiency of the schemes, we employ two indicators: ω and $\hat{\omega}$, which are the fraction of the EV of a scheme compared with the EV of the unrestricted first-best and second-best taxes, respectively. Policies that consider a labor tax cut acquire relative efficiencies of about 90% and higher. The permit schemes without revenue recycling effect also acquire a significant efficiency, confirming that the absence of this effect is less damaging at this lower level of labor tax.

As we discussed in Section 4.1, the schemes with permits allocated by distance and by kilometers traveled differ in the labor-supply-independent income transfer. We find a difference in relative efficiency of both schemes, going from $\hat{\omega} = 57\%$ in the permits allocated by kilometers traveled to $\hat{\omega} = 72\%$ when permits are allocated by distance, therefore the fact that the transfer is differentiated over space is relevant for increasing the efficiency of the scheme with permits by distance at this level of labor tax.

In Fig. 9, we show the total equivalent tax and equivalent non-labor income of the different policies, which are defined in Sections 2.5.1 - 2.5.3. In the right panel of Fig. 9, it can be seen that the scheme with permits allocated by distance is the only one that allows for a differentiation of the equivalent non-labor income across space.

As we discussed in Section 3.2.2, at this level of labor tax, the second-best charge is almost equal to a Pigouvian tax, therefore very close to zero for those households closer to the CBD. Thus, the flat kilometer charge and the permit scheme with allocation by labor supply are able to reach a higher efficiency because they diminish the equivalent taxation (when compared to the other permit schemes)

Table 1

Charges, results and welfare measures. Low labor tax case

	Benchmarks			Permit schemes by allocation		
	Base case	Second-best tax	Flat km charge	Distance	Labor supply	KM traveled
Charges						
Equivalent charge per km	0	See text	0.0019	0.0007	0.0020	0.0011
Labor tax τ_w	0.15	0.0225	0.0892	0.1517	0.0652[b]	0.1524
Results						
KM	7727	0.9449	0.9432	0.9498	0.9479	0.9503
$\bar{z}$	78.01	0.9643	0.9561	0.9657	0.9567	0.9676
Commuting time (t_c)	0.077[a]	0.9360	0.9375	0.9410	0.9456	0.9398
t_{lost}	0.054[a]	0.9322	0.9351	0.9372	0.9446	0.9353
Labor supply (t_w)	0.225[a]	0.9945	0.9930	0.9890	0.9986	0.9842
Leisure (l)	0.698[a]	1.0089	1.0092	1.0101	1.0065	1.0118
Y	147.2	1.0105	1.0126	1.0013	1.0209	0.9931
R	30.61	1.3541	1.3939	1.2504	1.4580	1.1596
Welfare measures						
U	0.781439	1.00064	1.00057	1.00046	1.00059	1.00037
EV[c]		0.0023	0.0021	0.0017	0.0021	0.0013
ω[d]		0.9787	0.8753	0.7073	0.906	0.5586
$\hat{\omega}$[e]		1	0.8944	0.7227	0.9258	0.5707

Note: U, KM, $\bar{z}$, t_c, t_{lost}, t_w, l, Y and R are expressed as fractions of the base case result.

[a] Averages as a fraction of total time.

[b] It includes the recycling induced by the term $\pi\beta$. See Eq. (38).

[c] Equivalent variation as a fraction of labor income in the base case.

[d] Relative efficiency with respect to first-best charge.

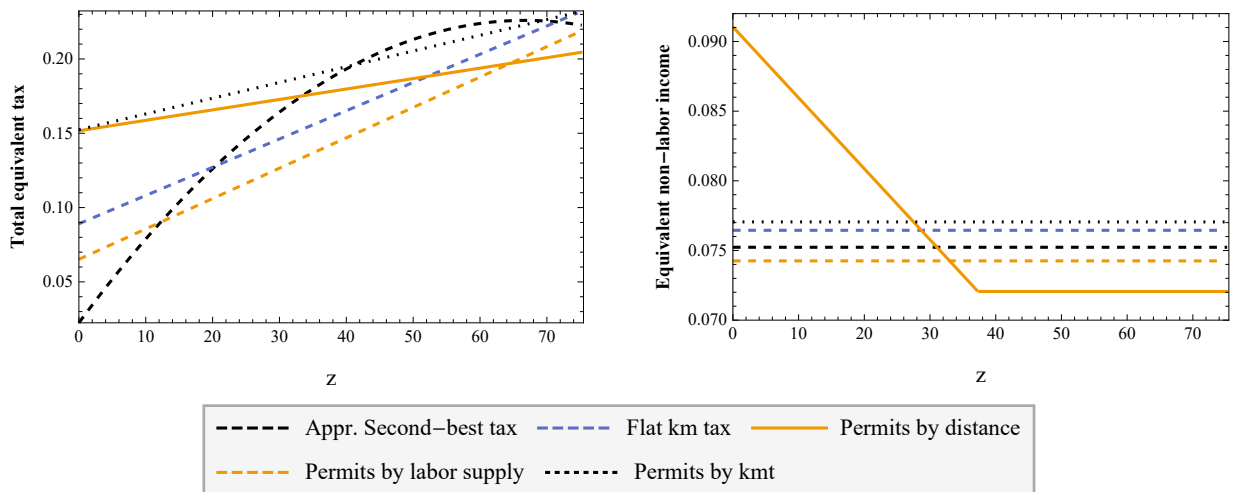[e] Relative efficiency with respect to approximate second-best tax.

Fig. 9. Financial effects of different schemes per location. Left: Total equivalent tax. Right: Equivalent non-labor income.

for households closer to the CBD, making it closer to its Pigouvian level, while at the same time increasing the charge per kilometer, as can be seen in the left panel of Fig. 5.

It is important to remind that the scheme with permits allocated by labor supply and the flat kilometer charge are equivalent only when $\alpha = 0$, as shown in Section 2.5.2. Thus, the resulting $\alpha < 0$ of the permit scheme (which explains the differences between these policies in both panels of Fig. 9) implies that it is optimal to have a higher revenue recycling effect than the flat kilometer charge has (or equivalently, that the optimal slope of the allocation of permits is larger than KM/L) and at the same time to have a slight increase in the road charge per kilometer (as shown in Table 1), therefore diminishing the distortionary effect of the total tax profile while increasing its Pigouvian effect.

An important result is that the scheme with permits allocated by labor supply does not result in a decrease in congestion when compared with the flat kilometer charge, even though congestion falls compared to the base case. In fact, this permit scheme reaches a higher utility even when is the scheme with the lowest decrease in commuting time and time lost due to congestion, but its higher labor supply and consumption compensate these time-related losses.

This suggests that, when restricting to a flat kilometer price (thus, when the road tax is imperfectly related with the road externality), it may be better to have a slightly higher congestion in order to have higher consumption and labor supply. The interpretation of this result must be careful. First, in this model there are only two distorted markets: the labor market and the congested road, and both are simultaneously optimized in order to maximize welfare. Therefore, a decrease in congestion is not always the desirable outcome. Second, the only externality considered in this paper is congestion, while other travel-related externalities, such as pollution and road safety are not included, which restricts the potential benefits of diminishing kilometers traveled. Third, as we already mentioned, the assumed linear congestion technology and the static nature of the model also diminish the impact of congestion. And fourth, focusing on aggregates in drawing policy conclusions is risky, as also spatial variation of impacts matters for eventual welfare effects.

Finally, we highlight the increase in rents that all policies induce. This is a natural result of having a charge that increases with the distance from the CBD, which leads to those locations closer to the CBD to be relatively more attractive, and to the city to shrink (as the values of $\bar{z}$ in Table 1 show). This also explains the fact that the policies with a higher utility induce a higher increase in rents, because these are the ones that are able to reduce the taxation at locations closer to the CBD, therefore increasing more their attractiveness. In line with the general equilibrium nature of the analysis, the recycling of excess land rents prevents higher rents from reducing welfare in the city.

4.3 High labor tax case (35%)

When considering an initial labor tax of 0.35, the only permit scheme that is welfare-increasing is the allocation by labor supply (as we showed in Fig. 2), due to the absence of a revenue recycling effect of the other schemes. The optimal scheme is defined by $\beta = 124.6$ and $\pi = 0.0020$. The optimal allocation and the revenues from the permit market per location are shown in Appendix D.

In Table 2, we show the results for this level of labor tax. The high efficiency of the scheme with permits allocated by labor supply is because it turns out to be very similar to a first-best charge, as shown in Fig. 10. This is explained by the fact that the revenue recycling effect of this scheme depends on two variables that are optimized by the social planner: the price per kilometer and the slope of the allocation of permits. This allows the total equivalent tax to be adjusted better than in

road pricing schemes, where the revenue recycling effect depends only on the price per kilometer. This is evident in Fig. 10 by looking at the second-best charge, where the labor tax reduction is not enough to have levels of taxation similar to the first-best. The resulting labor tax of the permit scheme is 0.0579 (considering the incentive provided by the allocation of permits), while the resulting labor tax of the second-best charge is 0.2686, therefore, the permit scheme has a much lower distortionary effect. It is important to remind that under this scheme, the public revenue constraint still holds. In fact, households are facing a labor tax of 0.326 if the incentive provided by the allocation of permits is not considered. This labor tax is lower than its initial level due to the increase in labor supply induced by the permit scheme.

Table 2
Charges, results and welfare measures. High labor tax case

	Base case	First-best	Permits by labor supply
Charges			
Equivalent charge per km	0	$mecc(z)$	0.0020
Labor tax τ_w	0.35	0	0.0579[b]
Results			
KM	6929	1.0582	1.0559
$\bar{z}$	74.42	1.0010	1.0014
Commuting time (t_c)	0.066[a]	1.1013	1.1036
t_{lost}	0.045[a]	1.1211	1.1255
Labor supply (t_w)	0.209[a]	1.0744	1.0751
Leisure (l)	0.673[a]	0.9693	0.9689
Y	135.0	1.1098	1.1157
R	20.70	2.0843	2.1970
Welfare measures			
U	0.779249	1.0035	1.0034
EV[c]		0.0150	0.0148
ω[d]		1	0.9817

Note: U, KM, $\bar{z}$, t_c, t_{lost}, t_w, l, Y and R are expressed as fractions of the base case result.

[a] Averages as a fraction of total time.
[b] It includes the recycling induced by the term $\pi\beta$. See Eq. (38).
[c] Equivalent variation as a fraction of labor income in the base case.
[d] Relative efficiency with respect to first-best charge.

The importance of a high revenue recycling effect is also demonstrated by the resulting value of $\bar{z}$, shown in Table 2, that is lower than β, which implies that the revenue recycling effect is needed over the entire city.
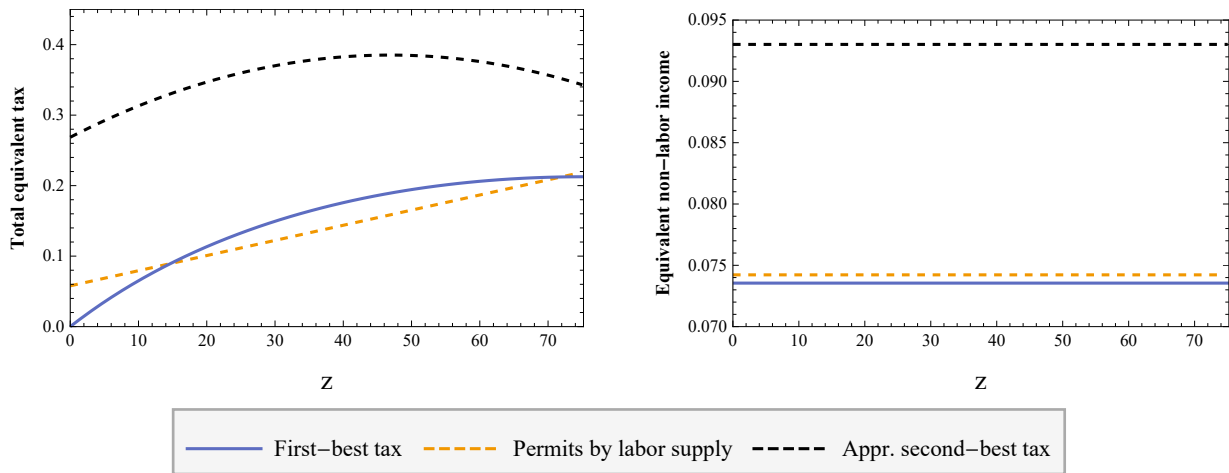
23

Fig. 10. Equivalent total tax and non-labor income per location of allocation by labor supply compared with first-best and second-best charges.

5. Synthesis and conclusions

We have incorporated different tradable permit schemes in a monocentric city with two distorted markets, which are a congested road and the labor market, that has a distortionary pre-existing labor tax to finance a publicly provided good.

We studied three tradable permit schemes, that differ in the variable that defines their allocation: i) household's residential distance from the CBD; ii) household's labor supply (number of days worked or, equivalently, number of commuting trips); and iii) household's kilometers traveled. For each allocation method, we assumed a linear relation between the relevant variable and the permits received by a household. Therefore, in each case, the allocation is determined by one parameter, which is the slope of the linear function, while the intercept is determined through the market clearance condition.

The analytical properties of the different allocations show that, as one permit is needed to travel one kilometer, the permit schemes induce a cost per kilometer equivalent to that of a flat kilometer road pricing. In addition, the scheme with permits allocated by labor supply has an important property that can enhance its efficiency: when the slope of the linear allocation has a positive sign, the scheme has an effect equivalent to those road pricing schemes that recycle revenues through a labor tax cut, the so-called revenue recycling effect, due to the incentive to increase labor supply. Moreover, this scheme is at least as efficient as a flat kilometer road pricing, being able to reach a higher welfare because the extent of the revenue recycling effect depends not only on the price per kilometer, but also on the slope of the allocation of permits, which is directly decided by the planner.

Regarding the other two schemes, the analytical properties show that both do not have a direct revenue recycling effect, therefore, their efficiency will be undermined at higher levels of labor tax. The scheme with permits allocated by kilometers traveled is equivalent to a flat kilometer charge with revenues returned as a lump-sum transfer.

On the other hand, when permits are allocated by distance, there is an effect equivalent to a spatially differentiated financial transfer to households, and its optimization depends on the slope of the allocation, therefore being a policy decision. This leads to higher levels of utility than with the allocation by kilometers traveled. Still, it cannot compensate the absence of a revenue recycling effect, due to only affecting directly one margin of behavior, which is the location decision.

The relevance of the revenue recycling effect is confirmed by the numerical results. The scheme with permits allocated by labor supply results to be the most efficient, with a welfare of 99.99% of the first-best charge for all levels of labor tax, being the only welfare-increasing permit scheme when labor tax is around 27% or higher, and being more efficient than second-best taxation starting from levels of labor tax of around 20%. This is because the Pigouvian effect of this scheme is kept constant when labor tax varies, while the allocation of permits varies in order to reach the revenue recycling effect needed to maximize welfare given the flat price per kilometer induced by the scheme.

Furthermore, we found that both schemes without a revenue recycling effect are welfare increasing at lower levels of labor tax. As expected, between them the highest utility is reached with permits allocated by distance, which is explained by the spatially differentiated transfer that they have, which allows to a significant increase in relative efficiency by giving budget to those households that face an overtaxation compared with the optimum, and therefore have a high marginal utility of income.

Our results highlight the importance of a careful design of tradable permit schemes, especially in already distorted markets. First, we showed that a differentiation in the allocation of permits is relevant even when considering homogeneous households, due to the differences that arises when considering a spatial setting.

Second, we showed that it is preferable to design the allocation of permits in order to provide incentives to the market activity that is distorted, which in this model is labor supply. Because the distortion is spatially differentiated in its severeness, also a spatial differentiation of the allocation of permits is in order to optimize the impacts of the policy in an urban economy. The fact that the social planner can decide the extent of the incentive to that activity, hence deciding the revenue recycling effect, is a key element of tradable permit schemes that is not present in traditional forms of road pricing, but that is also ignored in conventional partial-equilibrium studies of tradable mobility permits.

Third, we showed how allocations that depend on variables that are not the distorted activity could be also welfare increasing when the distortion is not very high, and that a spatially differentiated transfer that affects the location decisions is relevant to increase efficiency by compensating those households facing a higher charge than the optimum and that thus obtain a higher marginal utility of income. However, the higher the present distortion, the higher the need to address it directly by the design of the permit schemes, and therefore, the less efficient are allocations that do not address it.

All of this provides evidence that taking in account complex interactions between the different markets present in the urban economy is needed to correctly design tradable permit schemes and to compare their potential results with traditional road pricing. Finally, considering aspects not included in this model, such as potential adjustments of scheduling decisions, households that are heterogeneous in preferences and skills, or trips with purposes other than commuting is a future research agenda that would help to shed light about the efficiency and the equilibrium effects of tradable permit schemes.

Acknowledgments

Declaration of interest

None.

Appendix A. Flat kilometer road charge

Under a flat kilometer road charge, the social planner decides the price per kilometer π in order to maximize welfare. The revenue collected, G_p, is:

$$G_p = \int_0^{\bar{z}} n(z)d_w(z)\pi z dz \tag{A.1}$$

When revenues are used to finance a labor tax cut, the resulting labor tax τ_w^p is:

$$\tau_w^p = \frac{G - G_p}{L} = \tau_w^b - \frac{G_p}{L} \tag{A.2}$$

where τ_w^b is the initial level of labor tax. From Eq. A.1, considering that $n(z)zd_w(z) = KM$, we obtain $G_p = \pi KM$. Hence, the resulting labor tax is:

$$\tau_w^p = \tau_w^b - \frac{\pi KM}{L} \tag{A.3}$$

therefore, the net wage rate is:

$$\bar{w} = w - \tau_w^p - \pi z = w - \left(\tau_w^b - \frac{\pi KM}{L} \right) - \pi z \tag{A.4}$$

As mentioned in Section 2.5.2, when a tradable permit scheme with allocation by labor supply is implemented, the net wage rate is $w - (\tau_w^b - \pi\beta) - \pi z$. If $\alpha = 0$, the using Eq. 37 we obtain $\beta = KM/L$, therefore, both schemes are equivalent in this case.

When revenues collected with a flat kilometer road pricing are returned as a lump-sum transfer to each household, G_{ls}/N, we obtain:

$$\frac{G_{ls}}{N} = \frac{\pi KM}{N} \tag{A.5}$$

where it is clear that in this case the flat kilometer road charge is equivalent to the scheme with permits allocated by kilometers traveled, as mentioned in Section 2.5.3.

Appendix B. Marginal utility of income and elasticity of labor supply

The derivative of labor supply with respect to the net wage rate required for obtaining the elasticity was approximated numerically at each location by computing a new equilibrium where the net wage rate was increased by $\varepsilon = 10^{-5}$ for all z. Then, $\partial t_w(z)/\partial \bar{w}(z) = (t_w^1(z) - t_w^0(z))/\varepsilon$.
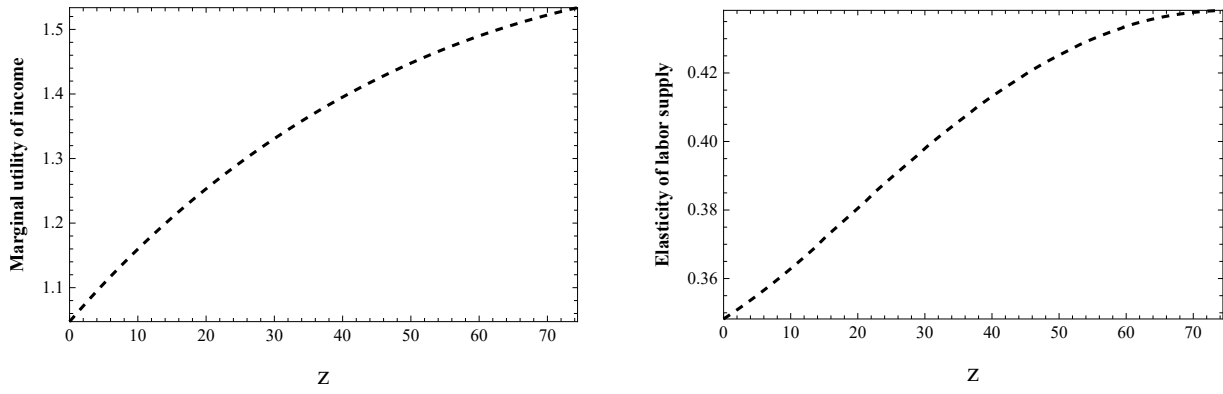
Fig. B.1. Left: Marginal utility of income per location. Right: Elasticity of labor supply per location. High labor tax base case.

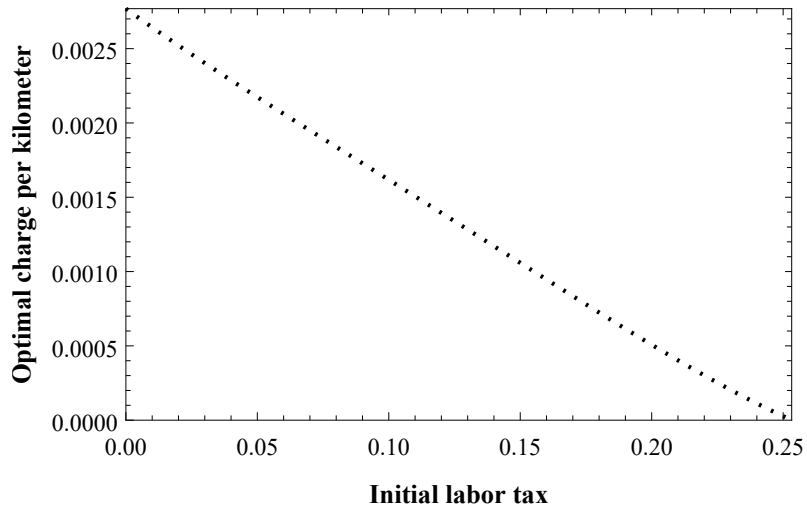## Appendix C. Optimal price when permits are allocated by kilometers traveled

Fig. C.1. Optimal price when permits are allocated by kilometers traveled as a function of labor tax.

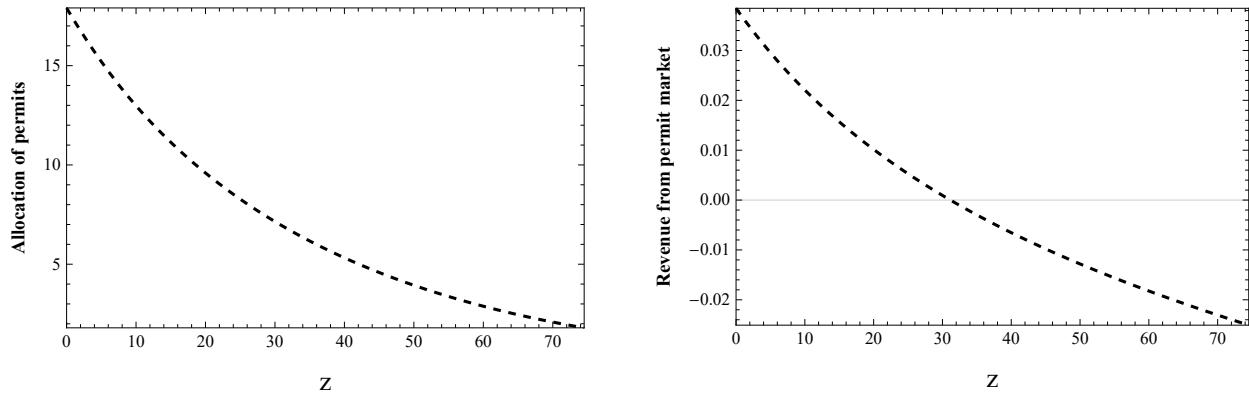## Appendix D. Permits by labor supply, high labor tax case

Fig. D.1. Optimal allocation and resulting permit market per location with permits allocated by labor supply. High labor tax case.

References

Akamatsu, T. and Wada, K. (2017). Tradable network permits: A new scheme for the most efficient use of network capacity. *Transportation Research Part C: Emerging Technologies* 79, 178–195.

Brands, D. K., Verhoef, E. T., Knockaert, J., and Koster, P. R. (2020). Tradable permits to manage urban mobility: market design and experimental implementation. *Transportation Research Part A: Policy and Practice* 137, 34–46.

Dales, J. (1968). Land, water, and ownership. *The Canadian Journal of Economics/Revue canadienne d'Economique* 1 (4), 791–804.

de Palma, A., Proost, S., Seshadri, R., and Ben-Akiva, M. (2018). Congestion tolling-dollars versus tokens: A comparative analysis. *Transportation Research Part B: Methodological* 108, 261–280.

Fan, W. and Jiang, X. (2013). Tradable mobility permits in roadway capacity allocation: Review and appraisal. *Transport Policy* 30, 132–142.

Goddard, H. (1997). Using tradeable permits to achieve sustainability in the world's large cities: policy design issues and efficiency conditions for controlling vehicle emissions, congestion and urban decentralization with an application to Mexico City. *Environmental and Resource Economics* 10 (1), 63–99.

He, F., Yin, Y., Shirmohammadi, N., and Nie, Y. M. (2013). Tradable credit schemes on networks with mixed equilibrium behaviors. *Transportation Research Part B: Methodological* 57, 47–65.

Nie, Y. M. (2012). Transaction costs and tradable mobility credits. *Transportation Research Part B: Methodological* 46 (1), 189–203.

Nie, Y. M. (2015). A new tradable credit scheme for the morning commute problem. *Networks and Spatial Economics* 15 (3), 719–741.

Parry, I. and Bento, A. (2001). Revenue recycling and the welfare effects of road pricing. *Scandinavian Journal of Economics* 103 (4), 645–671.

Shirmohammadi, N., Zangui, M., Yin, Y., and Nie, Y. (2013). Analysis and design of tradable credit schemes under uncertainty. *Transportation Research Record* 2333 (1), 27–36.

Tian, L.-J., Yang, H., and Huang, H.-J. (2013). Tradable credit schemes for managing bottleneck congestion and modal split with heterogeneous users. *Transportation Research Part E: Logistics and Transportation Review* 54, 1–13.

Tikoudis, I., Verhoef, E. T., and Van Ommeren, J. N. (2015). On revenue recycling and the welfare effects of second-best congestion pricing in a monocentric city. *Journal of Urban Economics* 89, 32–47.

Verhoef, E. (2005). Second-best congestion pricing schemes in the monocentric city. *Journal of Urban Economics* 58 (3), 367–388.

Verhoef, E. and Nijkamp, P. (2002). Externalities in urban sustainability: environmental versus localization-type agglomeration externalities in a general spatial equilibrium model of a single-sector monocentric industrial city. *Ecological Economics* 40 (2), 157–179.

Verhoef, E., Nijkamp, P., and Rietveld, P. (1997). Tradeable permits: their potential in the regulation of road transport externalities. *Environment and Planning B: Planning and Design* 24 (4), 527–548.

Wang, X., Yang, H., Zhu, D., and Li, C. (2012). Tradable travel credits for congestion management with heterogeneous users. *Transportation Research Part E: Logistics and Transportation Review* 48 (2), 426–437.

Weitzman, M. (1974). Prices vs. quantities. *The review of economic studies* 41 (4), 477–491.

Wu, D., Yin, Y., Lawphongpanich, S., and Yang, H. (2012). Design of more equitable congestion pricing and tradable credit schemes for multimodal transportation networks. *Transportation Research Part B: Methodological* 46 (9), 1273–1287.

Xiao, F., Qian, Z. S., and Zhang, H. M. (2013). Managing bottleneck congestion with tradable credits. *Transportation Research Part B: Methodological* 56, 1–14.

Yang, H. and Wang, X. (2011). Managing network mobility with tradable credits. *Transportation Research Part B: Methodological* 45 (3), 580–594.

Zhang, X., Yang, H., and Huang, H.-J. (2011). Improving travel efficiency by parking permits distribution and trading. *Transportation Research Part B: Methodological* 45 (7), 1018–1034.