



THEMA

théorie économique,  
modélisation et applications

THEMA Working Paper n°2020-14  
CY Cergy Paris Université, France

# Impacts of Metering-Based Dynamic Priority Schemes

R. Lamotte, A. de Palma, N. Geroliminis



December 2020

# Impacts of Metering-Based Dynamic Priority Schemes

R. Lamotte<sup>1</sup>, A. de Palma<sup>2</sup>, and N. Geroliminis<sup>1</sup>

<sup>1</sup>*School of Architecture, Civil and Environmental Engineering, EPFL, Switzerland.*

<sup>2</sup>*CY Cergy Paris Université, THEMA, France.*

December 30, 2020

## Abstract

Several works published over the last two decades have shown for a stylized set-up with homogeneous users that metering-based priority (MBP) schemes may generate Pareto improving departure time adjustments similar to those induced by congestion pricing, but without any financial transaction. We investigate whether MBP (i) still generates significant savings and (ii) remains Pareto-improving, with various sources of heterogeneity (in schedule flexibility, desired arrival time, and capacity usage). We consider two types of schemes: one where the priority status is allocated randomly (R-MBP) and another (HOV-MBP), which only prioritizes users with small capacity usages (e.g. carpoolers). We find that the relative total cost savings of R-MBP decrease with heterogeneity in flexibility, but may increase with heterogeneity in desired arrival time. It fails however to be Pareto-improving, as non-prioritized users are almost systematically worse-off. HOV-MBP circumvents this issue by generating an ordering effect and a modal shift, which both contribute to a better distribution of benefits among users. Under favorable circumstances, they may even restore a Pareto improvement. Overall, MBP appears as a realistic way to alleviate congestion, scoring well both in terms of efficiency and social acceptability.

*Keywords:* priority, metering, departure time choice, bottleneck, congestion

*JEL codes:* L92, R41

## 1 Introduction

The type of dynamic priority scheme considered here relies on demand metering, which refers to the action of regulating the inflow to an infrastructure. In usual applications such as ramp metering (Papamichail et al. 2010, Geroliminis, Srivastava, and Michalopoulos 2011), perimeter control (Ramezani, Haddad, and Geroliminis 2015) and passenger inflow control (Xu et al. 2014), the purpose of metering is to prevent capacity losses or safety hazards at critical infrastructures by holding queues in locations where they do not pose any risk. There, metering typically applies to all users, without discrimination. In this paper however, metering is combined with bypasses, such that some “priority users” can overtake the queues. This paper focuses on the additional benefits of such bypasses.

When individuals do not choose their time of departure time (i.e. it is considered as given), priority has an *ordering effect* and a *modal shift effect*. The first is a simple result from queueing theory. When users with different values of time and service times wait in a queue, it is socially optimal to reorganize the queue such that users with a larger ratio “value of time/service time” pass first. This exact discipline would be impractical in traffic applications, but a good approximation consists in prioritizing vehicles based on their occupancy. This *ordering effect* can already bring a significant improvement over the traditional First-In, First-Out (FIFO) discipline. Yet, these priority schemes also have a second benefit. They encourage users to shift towards prioritized modes, which are also more socially efficient. We refer to this as the *modal shift effect*. These two effects motivate for instance the pre-signals that let buses “jump the queue” at signalized intersections (Wu and Hounsell 1998, Guler and Menendez 2014).

A third effect appears with departure time choice: the priority status protects selected users against the negative externalities of non-priority ones. We refer to this as the *compartmentalization effect*.<sup>1</sup> In the literature that documents it, a transport authority selects the priority vehicles either based on an arbitrary characteristic like the license plate (Daganzo and Garcia 2000, Fosgerau 2011, Knockaert, Verhoef, and Rouwendal 2016), or based on the on-ramp at which users enter a highway (Lago and Daganzo 2007, Shen and Zhang 2010). These vehicles are then granted a privileged access to the bottleneck at the center of the peak period, either via temporary reserved lanes, or via a time- and group-dependent pricing (third-degree price differentiation). These works largely focus on the case of a homogeneous population (same vehicle occupancy and same schedule preferences), such that there is no ordering or modal shift effect. Yet, they show that when demand is inelastic, the resulting temporal segregation is Pareto-improving *regardless of the proportion that is prioritized*, and can achieve the same total cost savings as the best two-valued time-dependent constant congestion toll (Fosgerau 2011). This is in sharp contrast with conventional static reserved lanes, which improve welfare only under very specific conditions, and mildly so.<sup>2</sup>

This paper investigates whether these promising results can (i) be extended to cases with heterogeneity in schedule preferences, and (ii) be combined with the modal shift and ordering effects, when users have different capacity usages. The combination of all three effects can have desirable consequences as we will show. The remainder of this paper is structured as follows. Section 2 reviews more in detail the literature dealing with the compartmentalization effect. Section 3 introduces our modeling assumptions and discusses practical set-ups for implementations. Section 4 focuses on the compartmentalization effect with a random selection of priority vehicles. Although this is not necessarily a recommended policy, it isolates the compartmentalization effect from the other two, thus simplifying the study of priority schemes with heterogeneous schedule preferences. Section 5 then adds successively the ordering and modal shift effects and concludes with a case-study comparing the relative significance of these effects and comparing metering-based priority with conventional static reserved lanes. The mathematical derivations that are non essential to the flow of the text are relegated to appendices.

## 2 The compartmentalization effect in the literature

From a purely economic viewpoint, the most natural way to spread the demand over the time of day is to enforce a time-dependent toll, such as the one proposed by Vickrey (1969). Yet, public opposition to congestion pricing and its regressive effects have motivated researchers to investigate other alternatives, such as tradable credits (see de Palma et al. (2018) and the review by Grant-Muller and Xu (2014)), reservations (Lamotte, de Palma, and Geroliminis 2017), and priority. To the best of our knowledge, the idea of metering-based priority first appeared in a paper considering departure time choice in the context of a highway merge (Lago and Daganzo 2007). Lago and Daganzo show the counter-intuitive result that the total congestion cost can be alleviated by reducing the capacity of one the two roads upstream of the merge. As a side comment, they suggest that this could be taken to the extreme by applying metering.

Shen and Zhang (2010) explored this idea more in depth, for a freeway with several uncapacitated on-ramps and a unique bottleneck at the end. The authors compare various ramp-metering plans and show that large reductions of the total cost (up to 25 % with two on-ramps) can be achieved by prioritizing the on-ramps with respect to each other. Fosgerau (2011) further formalized this concept as follows: “*A set of travelers is assigned to a priority group. Not all travelers can be given priority. A more than proportional share of capacity is reserved for the priority group. When the reserved capacity is not used, it is available for the nonprioritized travelers.*” Although Fosgerau does not indicate how the time-dependent capacity management would be achieved in practice, metering seems to be the most natural interpretation for road applications. Fosgerau then demonstrates that such a scheme is always Pareto-improving and that it “can achieve exactly the same travel time and queuing outcomes as a [socially optimal two-level] coarse toll”.

A very similar outcome can be obtained by replacing metering with third-degree price discrimi-

<sup>1</sup>This was referred to as “congestion externality redistribution” by Shen and Zhang (2010), but we find this expression somewhat misleading because it suggests that the sum of externalities is conserved.

<sup>2</sup>With a given demand and modal split, static reserved lanes cannot be Pareto-improving because the congested period for non-priority users lasts unavoidably longer than without HOV lanes (Konishi and Mun 2010).

nation. Daganzo and Garcia (2000) for instance imposes a fine toll only on non-priority vehicles and only during a part of the peak period, such that priority vehicles can all pass without competing with non-priority ones. This also leads to a Pareto-improvement and a reduction of the congestion cost of up to 25 % with homogeneous  $\alpha - \beta - \gamma$  schedule preferences (see Section 3.1 for a definition). This toll is suboptimal in that it does not remove all queuing, but it has the advantage of requiring fewer toll payments, and might be socially more acceptable.<sup>3</sup>

Knockaert, Verhoef, and Rouwendal (2016) considers a similar set-up but imposes a toll also on the priority vehicles, thereby generating even larger cost reductions, but with a significantly larger total toll revenue. These toll-based policies do not require a physical separation of priority and non-priority users, but they require toll booths and would unavoidably raise equity issues if they were implemented, as they would not be as dissuasive for high- and low-income travelers.

All these papers assume  $\alpha - \beta - \gamma$  preferences, except for Fosgerau (2011) that allows for a more general specification.<sup>4</sup> They also all assume a homogeneous population, except for Daganzo and Garcia (2000) which includes an extension with some heterogeneity in desired arrival time, but only within some bounds and with  $\alpha - \beta - \gamma$  preferences, such that it has almost no effect. Overall, it appears that the literature has largely focused on the ideal case that best illustrates the benefits of priority schemes, without investigating whether negative effects could arise under more realistic assumptions, and in particular heterogeneity of preferences.

To conclude this literature review, we shall also mention the rich literature dedicated to departure time choice with static reserved lanes - see e.g. Qian and Zhang (2011) and Zhong et al. (2020). Static reserved lanes are admittedly simpler to implement and to study (because capacity is independent of time), but they are also less efficient, as permanently granting a more than proportional share of the capacity to some users unavoidably implies some capacity losses (Konishi and Mun 2010, Chen, Varaiya, and Kwon 2005).

### 3 Assumptions

Vickrey’s bottleneck problem considers departure time choice in the simplest possible setting: some population of selfish and non-cooperating users approximated by a continuum, wishes to go from  $A$  to  $B$ , with some schedule preferences (see Section 3.1). The originality of our approach lies in the congestion mechanism introduced in Section 3.2. The most important notations (those used across different sections) are listed in Table 1.

#### 3.1 Schedule Preferences

This paper considers the following general specifications of the schedule preferences similar to those used by Lindsey (2004) to establish existence and uniqueness of equilibrium.

**Assumption 1.** *The population is a set  $\mathcal{J} = \{1, \dots, J\}$  of homogeneous groups of users. Group  $j \in \mathcal{J}$  has a size  $N_j > 0$ , a desired arrival time  $t_j^*$  and a trip cost function  $c_j(T, t) = \alpha T + \text{SP}_j(t)$ , where  $\text{SP}_j$  denotes the schedule penalty of group  $j$ ,  $T$  denotes the extra travel time due to congestion and  $\alpha > 0$  is the value of time. The schedule penalty  $\text{SP}_j$  is continuous, non-negative, reaches 0 for  $t = t_j^*$  and is such that for all feasible  $t$ ,*

$$\lim_{\Delta t \rightarrow 0^+} \left( \frac{\text{SP}_j(t + \Delta t) - \text{SP}_j(t)}{\Delta t} \right) > -\alpha, \forall j \in \mathcal{J}. \quad (1)$$

The schedule penalty functions  $(\text{SP}_j)_{j \in \mathcal{J}}$  capture the inconvenience associated to arriving at a time different from the preferred one. They do not need to be differentiable. The  $\alpha - \beta - \gamma$  preferences of Arnott, de Palma, and Lindsey (1993) represent a widely-used example:

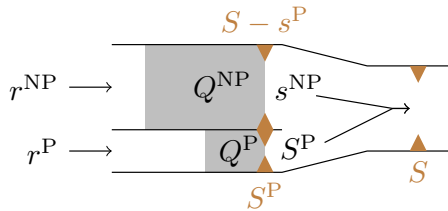
$$\text{SP}(t) = \begin{cases} \beta(t^* - t), & \text{if } t < t^* \\ \gamma(t - t^*), & \text{otherwise.} \end{cases} \quad (2)$$

<sup>3</sup>Theoretically, the collected revenue could even be brought arbitrarily close to zero when users differ only in terms of their desired arrival times (Daganzo and Garcia 2000).

<sup>4</sup>Specifically, Fosgerau (2011) assumes that utility is a separable function of departure time and arrival time, strictly concave, increasing with the departure time and decreasing with the arrival time.

Variables	Unit	Description
$\alpha$	\$/h	unit cost of travel time
$\beta$	\$/h	unit cost of arriving early
$\gamma$	\$/h	unit cost of arriving late
$\delta$	\$/h	$\triangleq \frac{\beta\gamma}{\beta+\gamma}$
$t^*$	h	desired arrival time
$T$	h	queueing time
$\mathcal{T}$	-	$\triangleq (t_0, t_e)$ , set of allowed departure times
$\mathcal{J}$	-	set of user groups
$N$	users	total number of users
$\tilde{N}$	veh	total number of vehicles
$N_j$	users	number of users in group $j \in \mathcal{J}$
$\tilde{N}_j$	veh	number of vehicles in group $j \in \mathcal{J}$
$S$	veh/h	capacity of the entire bottleneck
$\tau$	h	demand-to-capacity ratio
$S^P$	veh/h	capacity preemptable by priority users
$g$	users/veh	vehicle occupancy of priority vehicles
$\bar{g}$	users/veh	average vehicle occupancy
$p$	-	proportion of priority users
$q$	-	proportion of priority vehicles
$\bar{q}$	-	$\triangleq S^P/S$ , maximal proportion of priority vehicles
$\bar{p}$	-	$\triangleq gS^P/(S - S^P + gS^P)$ , maximal proportion of priority users
SP	\$	schedule penalty function
$\mathcal{C}$	-	image of $\mathcal{T}$ under SP
$C$	\$	reduced form cost function
$C^P, C^{NP}$	\$	cost for priority (resp. non-priority) users
TC	\$	total cost
$TC^{\text{ref}}$	\$	total cost without a priority scheme
$\% \Delta TC$	-	$\triangleq (TC^{\text{ref}} - TC)/TC^{\text{ref}}$ , relative total cost savings
$q^o$	-	proportion of priority vehicles maximizing $\% \Delta TC$
$p^*, q^*$	-	equilibrium proportion of priority users (resp. vehicles)
$L$	h	length of the interval over which $t^*$ is distributed
$\theta$	\$	carpooling inconvenience
$F_\theta$	-	cumulative distribution function of $\theta$
$\bar{\theta}, \theta^-, \theta^+$	\$	average, minimum and maximum carpooling inconvenience
$A$	\$	congestion advantage derived from being prioritized
$w$	$h^{-1}$	parameter of the function SP defined in the note below Fig. 7

Table 1: Notations



**Figure 1:** Schematic view of the metering-based priority scheme with parallel queues

The coefficients  $\beta$  and  $\gamma$  account for the costs of earliness and lateness respectively. With these preferences, the condition in Eq. (1) reduces to  $\beta < \alpha$ . In plain words, it ensures that users always prefer being at their destination to being on the road.

Note that while Lindsey (2004) allowed the value of time  $\alpha$  to be group-specific, we can assume without any loss of generality that it is homogeneous in the absence of pricing.

## 3.2 Bottleneck Dynamics with Metering-Based Priority

### 3.2.1 Mathematical Description

The supply side in Vickrey’s bottleneck problem is highly stylised: it considers a single bottleneck of constant capacity  $S$ , where all queueing occurs in a FIFO manner, without explicit spatial propagation. The variant considered in this paper is similar to a highway merge with ramp metering. Vehicles waiting upstream of the bottleneck are separated into two queues, denoted NP and P (see Fig. 1). Queue NP is open to all vehicles, while queue P is reserved for priority vehicles. Queue P has direct (i.e. un-metered) access to a subpart of the bottleneck with constant capacity  $S^P \leq S$ . If we denote  $r^P(t)$  and  $s^P(t)$  the flows of priority vehicles upstream and downstream of the bottleneck and  $Q^P(t)$  the queue length, we have the classic formulation of bottleneck dynamics (de Palma et al. 1983):

$$\left( \frac{dQ^P}{dt}(t), S^P(t) \right) = \begin{cases} (r^P(t) - S^P, S^P), & \text{if } Q^P(t) > 0 \text{ or } r^P(t) > S^P \\ (0, r^P(t)), & \text{otherwise.} \end{cases} \quad (3)$$

For NP users however, metering leads to a time-dependent capacity, which depends on the flow on the non-metered approach. If we define  $r^{NP}(t)$  and  $s^{NP}(t)$  as the flows of non-priority vehicles upstream and downstream of the bottleneck and  $Q^{NP}(t)$  the number of non-priority vehicles queueing at time  $t$ , we have:

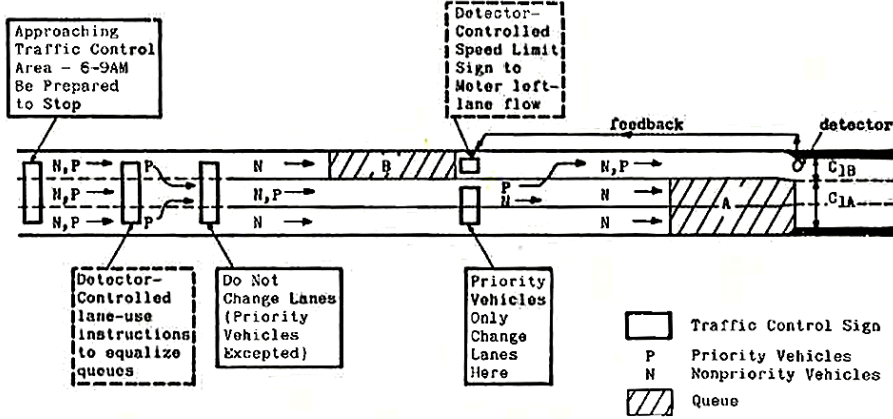
$$\left( \frac{dQ^{NP}}{dt}(t), s^{NP}(t) \right) = \begin{cases} (r^{NP}(t) - (S - S^P(t)), S - S^P(t)) & \text{if } Q^{NP}(t) > 0 \text{ or } r^{NP}(t) > (S - S^P(t)) \\ (0, r^{NP}(t)), & \text{otherwise.} \end{cases} \quad (4)$$

We refer to this scheme as *metering-based priority* (MBP) and to  $S^P$  as *preemptable capacity*, to distinguish it from conventional static schemes in which the capacity  $S^P$  is *reserved* at all times. Unlike a static reserved lane, a preemptable capacity cannot be underutilized when non-priority users are queueing.

Note also that the preemptable capacity might be relatively small, while the metered queue should have access to the full bottleneck capacity when there is no priority user. In that sense, our scheme is closer to “mainline-metering” than to traditional ramp-metering.

### 3.2.2 Practical View

The schematic view provided in Fig. 1 may seem to suggest that MBP requires the road upstream of the bottleneck to have a capacity of at least  $S + S^P$ . This is actually not necessary, as queues may be stored at different mileposts. Small (1983) provides an example of such scheme for highways. We reproduce its illustration in Fig. 2, but note that the notations used (explained in the figure’s legend)



**Figure 2:** A practical view of metering-based priority scheme with a fixed number of lanes (Small 1983)

differ from ours. In short, the scheme requires non-priority vehicles to stay on their lane, while allowing priority vehicles to slalom around the queues. It is similar to the concept of pre-signals ahead of signalized intersections (Wu and Hounsell 1998, Guler and Menendez 2014).

## 4 Random Priority

This section focuses on Random, Metering-Based Priority (R-MBP), a scheme in which the priority status is assigned to a proportion  $q$  of randomly selected vehicles. The priority status is allocated independently of the schedule preferences and of the vehicle occupancy, so that the metering scheme does not influence vehicle occupancy (no modal shift). These assumptions ensure that (i) the proportion of prioritized users  $p$  equals  $q$  and that (ii) the total passenger delay is equal to the average occupancy  $\bar{g}$  times the total vehicle delay (no ordering effect).

Let  $N = \sum_{j \in \mathcal{J}} N_j$  be the total population size. Because the bottleneck capacity is expressed in vehicles per hour, it is convenient to define  $\tilde{N} = N/\bar{g}$ , the total number of vehicles. For priority to be effective, the demand-to-capacity ratio of priority vehicles should be smaller than the one without metering, i.e.  $q\tilde{N}/S^P \leq \tilde{N}/S$ . Put differently, the prioritized proportion  $q$  should be kept within the interval  $[0, \bar{q}]$ , where  $\bar{q} = S^P/S$ .

### 4.1 Equilibrium: Definition, Existence and Uniqueness

In the present context, an equilibrium is a situation such that no individual user can reduce her travel cost by unilaterally changing departure time. Because of its deterministic demand model, the type of equilibrium considered here is commonly referred to as Deterministic Dynamic User Equilibrium (DDUE). The existence and uniqueness of DDUE were shown by Smith (1984) and Daganzo (1985) for a bottleneck of constant capacity and a continuum of users differing only in their desired arrival times  $t^*$ , with convex schedule penalty functions.

As explained by Lago and Daganzo (2007), these results can be extended to the case with a time-dependent capacity by rescaling the time axis. The case with metering-based priority then follows by considering the two groups of users sequentially. The priority users are not influenced by the non-priority ones and are faced with a bottleneck of constant capacity  $S^P$ , so a unique equilibrium exists for them. The departure-rate of priority users can then be considered as given, which means that the non-priority users can also be considered in isolation, but with a time-varying capacity.

The same approach is adopted here, but starting from the existence and uniqueness results of Lindsey (2004) and Akamatsu et al. (2018) instead, which allow for more general heterogeneity and non-convex schedule penalty functions. This leads to the following key result (the derivations are detailed in Appendix A).

**Assumption 2.** *The bottleneck is accessible with constant capacity  $S > 0$  during a time interval  $\mathcal{T} = (t_0, t_e)$ , such that  $(t_e - t_0)S > \tilde{N}$  and  $t_j^* \in \mathcal{T}, \forall j \in \mathcal{J}$ .*

**Proposition 1.** *Let Assumptions 1 and 2 hold. Consider a R-MBP scheme with  $S^P \in (0, S)$  and  $q \in (0, \bar{q})$ . There exists at least one deterministic departure-time user equilibrium. It may not be unique, but the travel costs of each group are uniquely defined.*

The departure-time user equilibrium is not necessarily unique in the sense that during uncongested periods, several departure profiles may correspond to the same equilibrium costs. For the purpose of this paper however, it suffices to know that the equilibrium costs are unique.

## 4.2 Benchmark: Identical Schedule Preferences

This section considers the effect of R-MBP on a population having identical schedule preferences. This case is already treated in the literature, but we revisit it in a way that prepares for cases with heterogeneous users.

### 4.2.1 Equilibrium Description

With a homogeneous population and a constant capacity bottleneck, the unique individual cost only depends on the demand-to-capacity ratio  $\tau = \tilde{N}/S$ , via a function  $C$ , known as the reduced form cost function (Arnott, de Palma, and Lindsey 1993). Given Assumption 1, this function can be expressed as:

$$C : \mathcal{R}_+ \rightarrow \mathcal{R}$$

$$\tau \mapsto C(\tau) = \sup \left\{ c \in \mathcal{C} \mid \int_{\mathcal{T}} [\text{SP}(t) < c] dt \leq \tau \right\}, \quad (5)$$

where  $\mathcal{C}$  denotes the image of  $\mathcal{T}$  under SP and where  $[P]$  is equal to 1 if P is true and 0 otherwise (Iverson bracket). Eq. (5) can be seen as a side-result of a more general expression obtained with time-dependent capacity (Lemma 3, Appendix B.1). With the  $\alpha - \beta - \gamma$  preferences, it reduces to  $C(\tau) = \delta\tau$ , with  $\delta = \frac{\beta\gamma}{\beta+\gamma}$ . We extend the scope of this concept and apply it to describe situations with metering-based priority.

**Proposition 2.** *Consider a homogeneous population satisfying Assumption 1 and a bottleneck satisfying Assumption 2. Under DDUE with metering-based priority,  $C^P = C(q\tilde{N}/S^P)$  and  $C^{\text{NP}} = C(\tilde{N}/S)$ . The DDUE with metering-based priority represents a Pareto improvement compared to the DDUE with no priority.*

The proof is provided in Appendix B.1.

The effect of R-MBP on priority users is relatively intuitive: by providing them with a more than proportional share of the capacity ( $q \leq S^P/S$ ), it reduces their demand-to-capacity ratio, and therefore their cost. A more surprising result is that the cost of non-priority users remains the same with and without metering. This is because with homogeneous users, the equilibrium individual cost is equal to the largest schedule penalty experienced among all users (see Eq. (5)). Since the total population size does not change and no capacity is wasted, this value does not change when implementing metering-based priority. As this could offer opportunities for practical implementations, one of the primary objectives of this paper is to determine whether this result still holds under more realistic conditions.

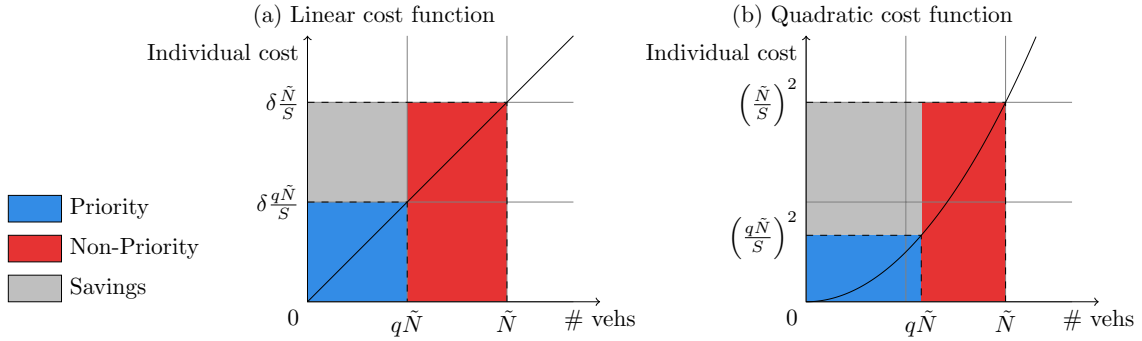
By contrast, the individual costs with a static reserved lane would be  $C^P = C(q\tilde{N}/S^P)$  (as with R-MBP) and  $C^{\text{NP}} = C((1-q)\tilde{N}/(S - S^P))$ . Since  $q < \bar{q} \Leftrightarrow (1-q)/(S - S^P) > 1/S$ , such schemes cannot make priority users better-off without making non-priority users worse off.

### 4.2.2 Optimal Priority Scheme

This section considers the effect of priority on the total cost at equilibrium, TC. Letting  $\text{TC}^{\text{ref}}$  denote the total cost without metering, we define the relative total cost savings as follows:

$$\% \Delta \text{TC} \triangleq \frac{\text{TC}^{\text{ref}} - \text{TC}}{\text{TC}^{\text{ref}}}. \quad (6)$$





**Figure 3:** Linear and quadratic reduced cost functions, and their optimal priority schemes

With homogeneous schedule preferences and R-MBP, the total cost  $TC$  depends on the preemptible capacity  $S^P$  and on the proportion  $q$  of priority vehicles via:

$$TC(S^P, q) = qNC^P + (1 - q)NC^{NP}. \quad (7)$$

As  $C^{NP} = C(\tilde{N}/S)$  and  $TC^{\text{ref}} = NC(\tilde{N}/S)$ , the relative total cost savings reduces to

$$\% \Delta TC(S^P, q) = \frac{NC(\tilde{N}/S) - qNC^P - (1 - q)NC(\tilde{N}/S)}{NC(\tilde{N}/S)} = q \left( 1 - \frac{C^P}{C(\tilde{N}/S)} \right). \quad (8)$$

Since  $C^P$  decreases with  $S^P$  while  $C^{NP}$  remains constant (Proposition 2), the model suggests that total cost  $TC$  decreases with  $S^P$ , such that it is socially optimal to let priority users preempt the entire bottleneck capacity ( $S^P = S$ ). In practice, however, this would raise several issues (unplanned trips, uncertainty in bottleneck access time, etc.) likely to penalize non-priority vehicles beyond acceptable levels. Thus, the model seems to be more appropriate for situations with a relatively small preemptible capacity  $S^P$ . This will naturally be the case when prioritizing only High-Occupancy Vehicles (HOV), see Section 5.

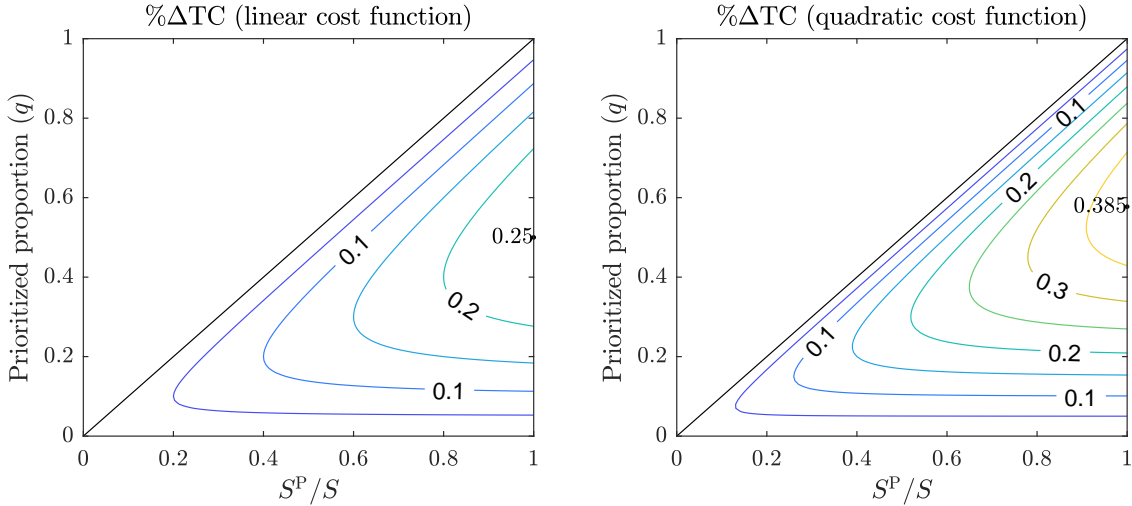
Let us now turn to the influence of the proportion of priority vehicles,  $q$ . With a linear cost function  $C(\tilde{N}/S) = \delta\tilde{N}/S$ , the total cost  $TC(S^P, q) = \delta N\tilde{N} [q^2/S^P + (1 - q)/S]$  is a second order polynomial, which is minimized for  $q = S^P/(2S)$ . In this case, priority vehicles experience a cost half as large as the others, such that  $\% \Delta TC = S^P/(4S)$ . If we set  $S^P = S$ , we recover the situation studied by Daganzo and Garcia (2000) and illustrated in Fig. 3a. The blue and red rectangles represent the contributions of priority and non-priority vehicles to the social cost. The sum of the two rectangles' areas (i.e. the total cost) is minimized for  $q = 0.5$ , leading to a relative total cost savings (the gray area) of 25%.

The comparison with Fig. 3b suggests that metering-based priority schemes can provide even larger benefits when users have a convex cost function. This is formalized in the following proposition (proven in Appendix C).

**Proposition 3.** *Let assumption 2 hold. Consider a population having homogeneous schedule preferences satisfying Assumption 1, with a continuously differentiable, strictly increasing and convex reduced form cost function. There exists a unique proportion  $q^o$  maximizing the relative total cost savings  $\% \Delta TC$ . It satisfies  $q^o > \bar{q}/2$  and  $\% \Delta TC(q^o) \geq \bar{q}/4$ .*

The reduced form cost function is convex when the schedule penalty function is convex and satisfies Assumption 1, but a convex schedule penalty function is not required for the convexity of the cost function (see Appendix B.2).

Fosgerau (2011) includes a similar result for a different class of utility functions (strictly concave in departure time and arrival time, increasing with arrival time and decreasing with departure time). Specifically, Fosgerau (2011) showed that with such utility functions, priority schemes achieve at least half the cost savings of the ideal fine toll. Since the ideal fine toll reduces the total cost by at least 50% with these utility functions and since our utility functions linearly decrease with the time of departure from the origin, Fosgerau's specification admits ours as a limiting case.



**Figure 4:** Relative total cost savings with homogeneous users

*Note:* Analytical forms: (a)  $\% \Delta \text{TC}(S^P, q) = q \left(1 - \frac{qS}{S^P}\right)$ , (b)  $\% \Delta \text{TC}(S^P, q) = q \left(1 - \left(\frac{qS}{S^P}\right)^2\right)$ .

Fig. 4 shows the relative total cost savings  $\% \Delta \text{TC}(S^P, q)$  obtained with linear and quadratic<sup>5</sup> cost functions for all the range of possible  $S^P/S$  and  $q$ . The maximum total cost savings are  $\% \Delta \text{TC}(S, 0.5) = 25\%$  and  $\% \Delta \text{TC}\left(S, \sqrt{1/3}\right) \simeq 38.5\%$ , respectively.

### 4.3 R-MBP with Heterogeneous Schedule Preferences

This section investigates whether the promising properties of R-MBP established in the literature (and summarized in Section 4.2) still hold under heterogenous schedule preferences. We distinguish two types of heterogeneity that we treat first separately to maintain tractability (see Section 4.3.1 and 4.3.2), and then combined in a numerical example (see Section 4.4).

#### 4.3.1 Users with Different Flexibility

Consider the following variation of Assumption 1, which constrains all groups to have proportional schedule penalty functions.

**Assumption 1'.** *The population is a set  $\mathcal{J} = \{1, \dots, J\}$  of homogeneous groups of users. Group  $j \in \mathcal{J}$  has a size  $N_j > 0$ . All groups have the same desired arrival time  $t^*$  and a trip cost function  $c_j(T, t) = \alpha T + k_j \text{SP}(t)$ , where  $\text{SP}$  denotes the common base schedule penalty,  $k_j$  denotes group  $j$ 's inflexibility,  $T$  denotes the extra travel time due to congestion and  $\alpha > 0$ . The groups are numbered by increasing inflexibility, such that  $0 < k_1 < k_2 < \dots < k_J$ . The schedule penalty  $\text{SP}$  is continuous, non-negative, reaches 0 for  $t = t^*$  and is such that for all feasible  $t$ ,*

$$\lim_{\Delta t \rightarrow 0^+} \left( k_j \frac{\text{SP}(t + \Delta t) - \text{SP}(t)}{\Delta t} \right) > -\alpha, \forall j \in \mathcal{J}.$$

The case without priority has already been studied multiple times with  $\alpha - \beta - \gamma$  preferences (see e.g. Arnott, de Palma, and Lindsey (1988), van den Berg and Verhoef (2011)). By applying the same ideas with our more generic schedule penalty functions, one can show that the individual cost of group  $j$  is

$$C_j^{\text{ref}} = k_j C(\tau_j^{\text{ref}}) + \sum_{i=1}^{j-1} k_i (C(\tau_i^{\text{ref}}) - C(\tau_{i+1}^{\text{ref}})),$$

where  $\tau_j^{\text{ref}} = (\sum_{i=j}^J \tilde{N}_i)/S$  is the minimum duration required to serve all the groups that are at least as inflexible as group  $j$ , and  $C$  is the reduced form cost function of a homogeneous population with schedule penalty  $\text{SP}$  and flexibility  $k = 1$ .

<sup>5</sup>It is easy to show that a quadratic cost function emerges when the marginal utility rates at the origin and at the destination are both linear in the time interval considered.

The first component ( $k_j C(\tau_j^{\text{ref}})$ ) results from the competition among vehicles that are at least as inflexible as group  $j$  (note that their exact level of inflexibility does not matter), while the second is the sum of delays imposed by all the groups strictly more flexible than  $j$ . Defining for convenience a fictive group with  $k_0 = 0$ , the individual costs can be rewritten as:

$$C_j^{\text{ref}} = \sum_{i=1}^j (k_i - k_{i-1}) C(\tau_i^{\text{ref}}), \quad \forall j \in \mathcal{J}. \quad (9)$$

With R-MBP, the individual costs  $(C_j^{\text{P}})_{j \in \mathcal{J}}$  and  $(C_j^{\text{NP}})_{j \in \mathcal{J}}$  are given by expressions similar to Eq. (9), but with different periods  $(\tau_j^{\text{P}})_{j=1 \dots J}$  and  $(\tau_j^{\text{NP}})_{j=1 \dots J}$ . For priority vehicles, it suffices to scale down both the population size and the capacity, so  $\tau_j^{\text{P}} = q(\sum_{i=j}^J \tilde{N}_i)/S^{\text{P}}$ . For non-priority vehicles, two cases must be distinguished:

$$\tau_j^{\text{NP}} = \begin{cases} \frac{(1-q) \sum_{i=j}^J \tilde{N}_i}{S - S^{\text{P}}}, & \text{if } \frac{(1-q) \sum_{i=j}^J \tilde{N}_i}{S - S^{\text{P}}} \leq \frac{q\tilde{N}}{S^{\text{P}}} \\ \frac{q\tilde{N} + (1-q) \sum_{i=j}^J \tilde{N}_i}{S}, & \text{otherwise} \end{cases} \quad \forall j \in \mathcal{J}. \quad (10)$$

The first expression is used when all non-priority vehicles of group  $j$  and those that are less flexible travel simultaneously with priority vehicles (they then see a constant capacity  $S - S^{\text{P}}$ ). When this is not the case, the duration required to serve the non-priority vehicles that are at least as inflexible as group  $j$  (with the capacity left unused by priority vehicles) is equal to the duration required to serve these vehicles *plus all priority ones, with the full capacity  $S$* . The second expression in Eq. (10) is the direct mathematical translation of this definition.

Before comparing the effects of R-MBP for the heterogeneous and homogeneous cases, let us define some additional notations. Let  $\mathbf{K} \triangleq (k_j)_{j \in \mathcal{J}}$  and  $\mathbf{N} \triangleq (N_j)_{j \in \mathcal{J}}$ . For all  $j \in \mathcal{J}$ , let  $\%C_j^{\text{P}} \triangleq C_j^{\text{P}}/C_j^{\text{ref}}$  and  $\%C_j^{\text{NP}} \triangleq C_j^{\text{NP}}/C_j^{\text{ref}}$  denote the relative individual costs of group  $j$ 's priority and non-priority vehicles, and let  $\%\Delta\text{TC}(S^{\text{P}}, q, \mathbf{K}, \mathbf{N})$  denote the relative total cost savings. For comparison purposes, we also define their counterparts  $\%C^{\text{P}} \triangleq C^{\text{P}}/C^{\text{ref}}$ ,  $\%C^{\text{NP}} \triangleq C^{\text{NP}}/C^{\text{ref}}$  and  $\%\Delta\text{TC}(S^{\text{P}}, q)$  for the case where all  $N$  users have the same schedule penalty function  $\text{SP}$ , with the same  $k$ .<sup>6</sup>

**Proposition 4.** *Let assumptions 1' and 2 hold. Consider a R-MBP scheme with  $S^{\text{P}} \in (0, S)$  and  $q \in (0, \bar{q})$ .*

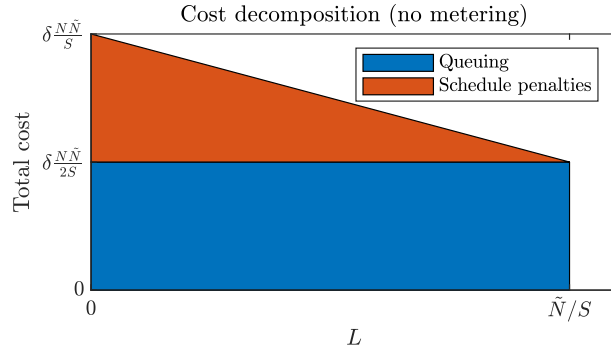
- a)  $C_1^{\text{NP}} = C_1^{\text{ref}}$ .
- b) *If the reduced form cost function  $C$  is strictly increasing, then  $C_j^{\text{NP}} > C_j^{\text{ref}}$  for all  $j = 2, \dots, J$ .*
- c) *If, in addition,  $x \mapsto \frac{C(\nu x)}{C(x)}$  is non-increasing for all  $\nu \in (0, 1)$  and  $x \geq 0$ , then*
  - i)  $\%C_j^{\text{P}} \geq \%C^{\text{P}}$  and  $\%C_j^{\text{NP}} \geq \%C^{\text{NP}}$  for all  $j \in \mathcal{J}$ ;
  - ii)  $\%\Delta\text{TC}(S^{\text{P}}, q, \mathbf{K}, \mathbf{N}) \leq \%\Delta\text{TC}(S^{\text{P}}, q)$ .

The proof is relegated to Appendix C.

Proposition 4.b contrasts with the homogeneous case, because the R-MBP scheme does not Pareto-dominate the laissez-faire strategy. Under some additional restrictions,<sup>7</sup> Proposition 4.c goes further and establishes that all individual relative cost savings are smaller than in an analogue situation with a homogeneous population. As a result, the relative total cost savings  $\%\Delta\text{TC}$  is also smaller. In light of these results, the homogeneous situation appears as a best case for the R-MBP policy, which seems exceedingly optimistic.

<sup>6</sup>With homogeneous users, the relative individual costs and relative total savings do not depend on  $k$ .

<sup>7</sup>Examples of functions  $C$  such that  $x \mapsto \frac{C(\nu x)}{C(x)}$  is non-increasing include  $C(x) = ax^b$  with  $a, b > 0$  (incl. the  $\alpha - \beta - \gamma$  preferences) and logarithmically convex functions.



**Figure 5:** Cost decomposition without metering, with  $\alpha - \beta - \gamma$  preferences and uniformly distributed  $t^*$

Note:  $\alpha = 1$ ,  $\beta = 0.3$ ,  $\gamma = 1.2$ .

### 4.3.2 Users with Different Desired Arrival Times

This section shows that R-MBP is also unlikely to be Pareto-improving with heterogeneity in desired arrival times, but that in this case, its benefits may exceed those estimated with homogeneous users.

We start by analyzing the cost savings of R-MBP for a population of users having homogeneous  $\alpha - \beta - \gamma$  coefficients, but where  $t^*$  is uniformly distributed on an interval  $[t_1^*, t_2^*]$  such that  $(t_2^* - t_1^*) \triangleq L \in [0, \tilde{N}/S)$ . Note that this set-up is already partially covered by Daganzo and Garcia (2000), which considers a more general sigmoid cumulative distribution of desired arrivals, but only within some bounds, such that even priority users always queue. Our treatment of the case with  $\alpha - \beta - \gamma$  preferences extends Daganzo and Garcia (2000) not only in that it addresses the case where priority vehicles do not have to queue, but it also provides a quantitative analysis of the cost savings.

The case with continuously distributed desired arrival times is not covered by Assumption 1 because of the infinite number of group, but it is well covered by the literature (starting with Vickrey (1969)). The social cost decomposition at equilibrium is displayed in Fig. 5 as a function of  $L$ . The total queuing time is constant for  $L \in [0, \tilde{N}/S)$ , but the total schedule penalty decreases linearly and converges to 0 as  $L$  tends towards  $\tilde{N}/S$ . The case  $L = \tilde{N}/S$  is degenerate<sup>8</sup>, and the cases  $L > \tilde{N}/S$  exhibit no congestion at all, because the “demand density”,  $\tilde{N}/L$ , is everywhere smaller than the capacity  $S$ .

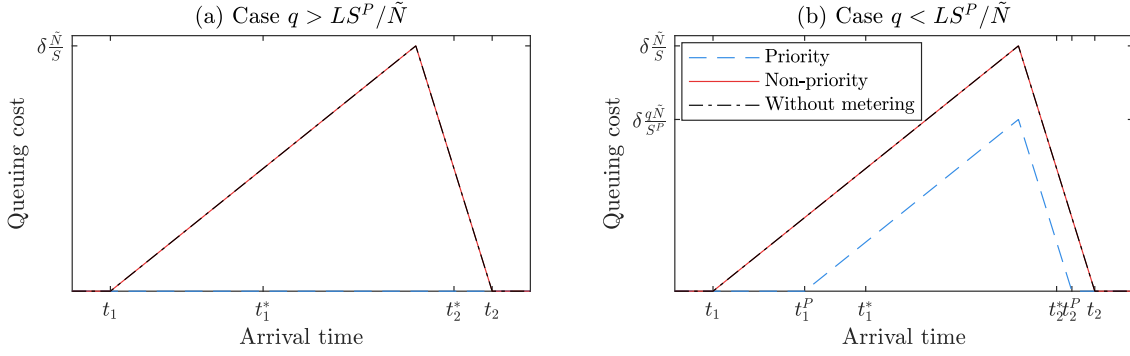
Let  $\% \Delta \text{TC}(S^P, q, L)$  and  $\% \Delta \text{TC}(S^P, q)$  denote the relative total cost savings achieved with heterogeneous and homogeneous  $t^*$  respectively. As shown hereafter, the R-MBP strategy leverages the particular properties of the equilibrium with heterogeneous  $t^*$  to alleviate congestion even more significantly than with homogeneous users.

**Proposition 5.** *Consider a bottleneck of constant capacity  $S$  and a population of size  $N$  having the same  $\alpha - \beta - \gamma$  preferences (with  $\beta < \alpha$ ) but with  $t^*$  uniformly distributed on  $[\bar{t}^* - \frac{L}{2}, \bar{t}^* + \frac{L}{2}]$ ,  $L \in (0, \tilde{N}/S)$ . Assume the bottleneck is accessible during a sufficiently large time interval. Applying R-MBP leads to  $\% \Delta \text{TC}(S^P, q, L) \geq \% \Delta \text{TC}(S^P, q)$  and  $\sup_{q,L} \% \Delta \text{TC}(S^P, q, L) = 100\%$ .*

*Proof.* Proof. Two cases should be distinguished, depending on whether the priority vehicles experience congestion. Since their demand density is  $q\tilde{N}/L$  and their capacity is  $S^P$ , this depends on whether  $q$  is smaller or larger than  $LS^P/\tilde{N}$ .

**Case with No Congestion for Priority Vehicles ( $q \leq LS^P/\tilde{N}$ ).** Let us consider separately the sum of schedule penalties and the sum of queuing costs. The sum of schedule penalties over all users (both priority and non-priority) is the same with and without R-MBP. Indeed, the arrival order at destination with R-MBP can be obtained by reallocating the arrival times of early (resp. late) or on-time vehicles without R-MBP, in such a way that all these vehicles remain early (resp. late) or on time. With homogeneous  $\beta$  and  $\gamma$  coefficients, such modifications leave the sum of schedule penalties unchanged.

<sup>8</sup>There can be two equilibria, one with delays (the limit case as  $L \rightarrow \tilde{N}/S$  from below) and one without (the limit case as  $L \rightarrow \tilde{N}/S$  from above).



**Figure 6:** Equilibrium queuing times with  $\alpha - \beta - \gamma$  preferences and uniformly distributed  $t^*$   
Note:  $\alpha = 1$ ,  $\beta = 0.3$ ,  $\gamma = 1.2$ .

Let us now consider the queuing time profiles, illustrated in Fig. 6a. For non-priority vehicles, the profile is the same as without metering. Thus, the only difference in terms of total cost between R-MBP and the no-priority case is that  $qN$  vehicles uniformly distributed between  $t_1^*$  and  $t_2^*$  experience no delay at all. Without metering, these vehicles would have experienced queuing time uniformly distributed in the interval  $[\delta(\frac{\tilde{N}}{S} - L), \delta\frac{\tilde{N}}{S}]$ . Thus, the total queuing cost saved is  $qN\delta(\frac{\tilde{N}}{S} - \frac{L}{2})$ . Since the overall congestion cost without metering is  $\delta N(\frac{\tilde{N}}{S} - \frac{L}{2})$ ,  $\% \Delta \text{TC}(S^P, q, L) = q$ . It is larger than with homogeneous users ( $\% \Delta \text{TC}(S^P, q) = q(1 - \frac{qS}{S^P})$ ) and its limit as  $q$  tends towards  $LS^P/\tilde{N}$  and  $L$  tends towards  $\tilde{N}/S$  is 100 %.

**Case with Congestion for All Vehicles ( $q \geq LS^P/\tilde{N}$ ).** As in the previous case, the sum of schedule penalties remains unchanged. Here however, priority vehicles do experience queuing at equilibrium. An example of such scenario is shown in Fig. 6b. The  $qN$  priority users traveling between  $t_1^P$  and  $t_2^P$  benefit from a reduction of  $\delta(\frac{\tilde{N}}{S} - \frac{q\tilde{N}}{S^P})$  in their individual queuing cost. Thus, the overall social cost savings is

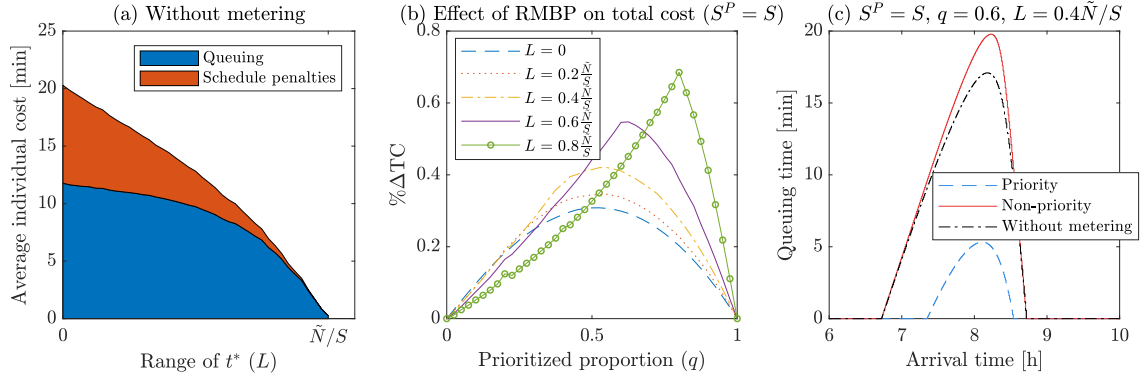
$$\% \Delta \text{TC}(S^P, q, L) = \frac{q(\frac{\tilde{N}}{S} - \frac{q\tilde{N}}{S^P})}{\frac{\tilde{N}}{S} - \frac{L}{2}} = \frac{q(1 - \frac{qS}{S^P})}{1 - \frac{LS}{2\tilde{N}}}.$$

Since  $L < \tilde{N}/S$ , the denominator  $(1 - \frac{LS}{2\tilde{N}}) \in (\frac{1}{2}, 1)$ , so  $\% \Delta \text{TC}(S^P, q, L) > q(1 - \frac{qS}{S^P}) = \% \Delta \text{TC}(S^P, q)$ .

The degenerate cases  $q = LS^P/\tilde{N}$  admit two equilibria (one with congestion, and one without), but both equilibria satisfy the results stated in the proposition.  $\square$

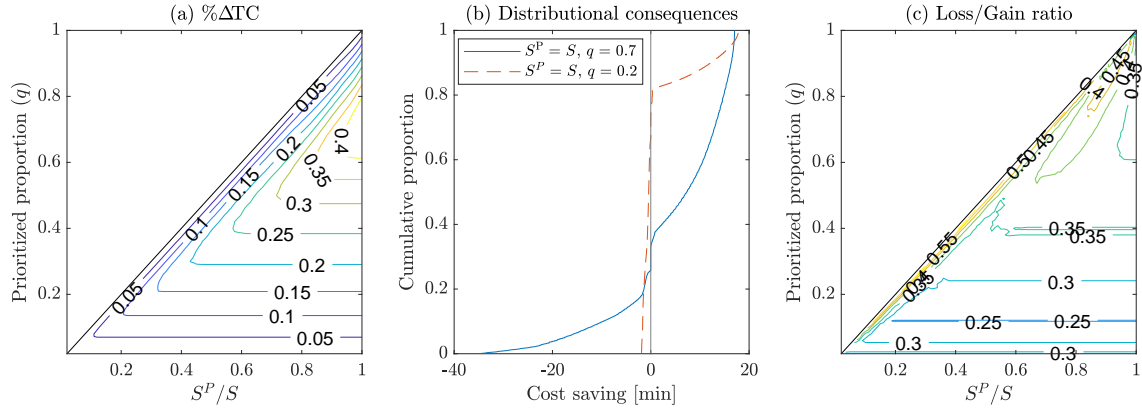
The R-MBP strategy Pareto-dominates the *laissez-faire* under the assumed conditions. This result depends however heavily on the  $\alpha - \beta - \gamma$  preferences postulated in Proposition 5. Recall the first-order equilibrium condition: if a user with a differentiable schedule penalty function SP arrives at a time  $t$ , then  $\alpha T'(t) = -SP'(t)$ . If  $SP'(t)$  increased smoothly from  $-\beta$  to  $+\gamma$ , users would thus impose larger delay variations  $|T'|$  when arriving further from their  $t^*$ .

To illustrate this, consider the situation in Fig. 7, which is similar to the one of Fig. 6, but with a convex approximation of the  $\alpha - \beta - \gamma$  preferences. Because the equilibrium is more complicated to derive analytically, we rely instead on a discrete time approximation of equilibrium (see the algorithm described in Appendix E). The total cost without metering is represented as a function of  $L$  in Fig. 7a. For convenience, we choose a value of time  $\alpha = 1$ , such that the costs can be expressed in terms of equivalent waiting time. As  $L$  tends towards  $\tilde{N}/S$ , users arrive closer to their desired time  $t^*$ , and thus experience not only smaller schedule penalties, but also smaller delays, because  $|T'|$  takes smaller values. This also explains why here, unlike with the  $\alpha - \beta - \gamma$  preferences, the total cost decreases towards 0 as  $L$  tends towards  $\tilde{N}/S$ . This situation is less favorable for the R-MBP strategy, because modifying the sequence of vehicles (which naturally sort by increasing  $t^*$ ) leads to larger delays, as illustrated in Fig. 7c. Yet, heterogeneity in  $t^*$  still allows for maximum relative total cost savings  $\% \Delta \text{TC}$  that are larger than with homogeneous users ( $L = 0$ ) - see Fig. 7b.



**Figure 7:** Equilibrium queuing times and cost decomposition with a smooth approximation of the  $\alpha - \beta - \gamma$  preferences and uniformly distributed  $t^*$

*Note:* Results obtained with  $\tilde{N}/S = 2$  h, discrete arrival times spanning over a 4-hour period centered around the mean  $t^*$  (time step: 0.01 h). Schedule preferences:  $SP(t) = \int_{t^*}^t \frac{\delta}{\pi} \tan^{-1}(w(s - \tilde{t})) ds$ , with  $\tilde{t} = t^* + \tan\left(\frac{\pi(\gamma - \beta)}{2(\gamma + \beta)}\right) \frac{1}{w}$ ,  $w = 10 \text{ h}^{-1}$ ,  $\alpha = 1$ ,  $\beta = 0.3$  and  $\gamma = 1.2$ .



**Figure 8:** Case-study with R-MBP

*Note:*  $\tilde{N}/S = 2.5$  h. Time discretization: every 1 min, from 5 AM to 10 AM.

#### 4.4 Numerical Example

We now combine heterogeneity in flexibility and in desired arrival time. Since the R-MBP scheme is not expected to be Pareto-improving, this section compares the gains and losses incurred by all individuals. All users have the same generic smooth and convex schedule penalty function introduced in Fig. 7 (with  $w = 10 \text{ h}^{-1}$ ), but their flexibility levels are uniformly distributed between three values ( $(\alpha, \beta, \gamma) \in \{(1, 0.3, 1.2), (1, 0.6, 2.4), (1, 0.9, 3.6)\}$ ), while their desired arrival times  $t^*$  are normally distributed (centered around  $t = 8$  h, with  $\sigma = 0.25\tilde{N}/S$ ). Flexibility and desired arrival time are independent.

The three graphs in Fig. 8 summarize the global and distributional consequences of the R-MBP scheme. From a global perspective, the R-MBP scheme reduces the total cost for all  $S^P > 0$  and  $q \in (0, \bar{q})$ . When  $S^P$  is large and the proportion of prioritized vehicles is near  $0.7S^P/S$ , savings represent up to 45.6% of the total cost (value obtained for  $S^P = S$  and  $q = 0.7$ ), as illustrated in Fig. 8a. This is slightly larger than the results obtained with homogeneous users (cf Fig. 4).

At the individual level however, the consequences are more contrasted. Fig. 8b shows the cumulative distribution of individual changes in congestion cost (in units of equivalent waiting time), for the set-up that maximizes the total cost savings ( $S^P = S$  and  $q = 0.7$ ) and for another set-up with a smaller prioritized proportion ( $q = 0.2$ ). In the socially optimal set-up, the maximum loss is equivalent to about 34.5 min of queuing time, the maximum gain to 17 min, and the average cost savings to 5 min. Although only a minority of the population (about 25%) is worse-off, the very large cost increase imposed on a few users would most likely undermine the acceptability of

this socially optimal R-MBP scheme.

The set-up  $q = 0.2$  has very different distributional consequences. The proportion of users made worse-off by the R-MBP scheme is much larger (67 %), but the largest loss is now only 1.88 min, for an even larger maximum gain (17.9 min) and an average gain of 1.56 min. Yet, the overall ratio of total losses over total gains is still about 29 %, while it reaches 32 % with  $q = 0.7$ , and takes similar values for a wide range of other set-ups ( $S^P, q$ ), as shown by Fig. 8c. Thus, there is a trade-off between large cost savings and a broadly distributed total loss, but the overall loss-over-gain ratio remains approximately the same. It contrasts rather sharply with the Pareto-improvement result obtained in the literature and in Section 4.2 for homogeneous users, and strongly impairs the acceptability of R-MBP.

## 5 Targeted Priority

This section considers schemes where the priority status is not randomly allocated, but given to carefully chosen categories of users. In the context of road traffic, these could be vehicles with several occupants (known as High-Occupancy Vehicles, or HOV) and/or vehicles with a small capacity usage. In the context of crowded train stations, they could be passengers with a physical condition that make waiting particularly inconvenient (pregnant women, elderly people), or again, passengers with a small capacity usage (e.g. those without luggage). Depending on the chosen category, the compartmentalization effect is reinforced by either only an ordering effect (e.g. for pregnant women or elderly people), or by both an ordering effect and a mode choice effect. To simplify the exposition, we depict targeted priority schemes with a concrete example, HOV bypasses, and refer to the corresponding scheme as HOV-MBP.

### 5.1 Two simplifying assumptions

This section introduces two assumptions that simplify the integration of carpools in the departure time choice problem.

**Assumption 3.** *Carpools all have  $g > 1$  occupants, while others have only one.*

**Assumption 4.** *Carpoolers travel only with members of their own group  $j \in \mathcal{J}$ .*

Assumption 3 implies that the proportion  $p$  of prioritized users now differs from the proportion  $q$  of prioritized vehicles. With these notations,  $\bar{g} = qg + 1 - q$ ,  $p = qg/\bar{g}$  and  $q = p/(p + g - pg)$ . As in Section 4, the proportion of priority vehicles should be kept within the interval  $[0, \bar{q} = S^P/S]$  to ensure that they experience less congestion than without metering. In terms of passengers, this translates into a similar condition  $p \in [0, \bar{p}]$ , with

$$\bar{p} = gS^P / (S - S^P + gS^P).$$

Assumption 4 ensures that carpoolers choose their departure times as if they were traveling on their own, as also assumed by Yu, van den Berg, and Verhoef (2019). Although the bargaining process that would normally lead to the choice of departure time among users with heterogeneous preferences certainly has important consequences (see e.g. (Picard, Dantan, and de Palma 2018) in the context of mode choice within couples), we leave it as a future research direction for a paper that would focus on carpooling itself (and not on priority schemes).

### 5.2 Predetermined Carpooling Decision

This section considers the proportion of carpoolers as given and compares the HOV-MBP scheme with a R-MBP scheme that would prioritize exactly the same number of vehicles. The difference between the two schemes illustrates the ordering effect.

To simplify the analysis, we rely on the following assumption:

**Assumption 5.** *Every group  $j \in \mathcal{J}$  contains the same proportion  $p_0$  of carpoolers.*

Assumption 5 ensures that the equilibrium queuing profile with HOV-MBP is the same as with the R-MBP scheme that prioritizes a proportion of vehicles equal to  $q_0 = p_0/(p_0 + g - p_0g)$ , the proportion of HOV. There is the same number of priority vehicles with both schemes, but the priority vehicles have more occupants on average with HOV-MBP, so that the proportion of users made worse-off by HOV-MBP is smaller than with the R-MBP scheme. The same holds for the total cost, as formalized hereafter (and proven in Appendix C).

**Proposition 6.** *Let assumptions 1, 2, 3, 4 and 5 hold. Consider the HOV-MBP scheme and the associated R-MBP scheme that prioritizes the same proportion of vehicles  $q_0$ . The relative total cost savings they achieve satisfy  $\% \Delta \text{TC}^{\text{R-MBP}} \leq \% \Delta \text{TC}^{\text{HOV-MBP}}$ .*

*In the special case where users have homogeneous schedule preferences,*

$$\frac{\% \Delta \text{TC}^{\text{R-MBP}}}{\% \Delta \text{TC}^{\text{HOV-MBP}}} = \frac{\bar{g}}{g} = \frac{1}{g} + q_0 \left(1 - \frac{1}{g}\right). \quad (11)$$

*In the special case where flexibility is the only source of heterogeneity (Assumption 1'),*

$$\frac{\% \Delta \text{TC}^{\text{R-MBP}}}{\% \Delta \text{TC}^{\text{HOV-MBP}}} \leq \frac{\bar{g}}{g} = \frac{1}{g} + q_0 \left(1 - \frac{1}{g}\right). \quad (12)$$

Note that the ratio  $\bar{g}/g$  linearly increases from  $1/g$  to 1 as  $q_0$  varies from 0 to 1, i.e. as the proportion of carpools increases. In other words, the ordering effect is particularly strong when a small proportion of vehicles is prioritized.

The main drawback of HOV-MBP is that the proportion of prioritized users can only be controlled via the choice of the occupancy threshold for priority users. This limitation might be problematic for pure welfare-maximizing strategies (which typically require large prioritized proportions, as seen in Section 4), but acceptability also matters and Section 5.3 shows that the HOV-MBP scheme has significant advantages in this regard.

## 5.3 Endogenous Carpooling Decision

### 5.3.1 Carpooling inconvenience

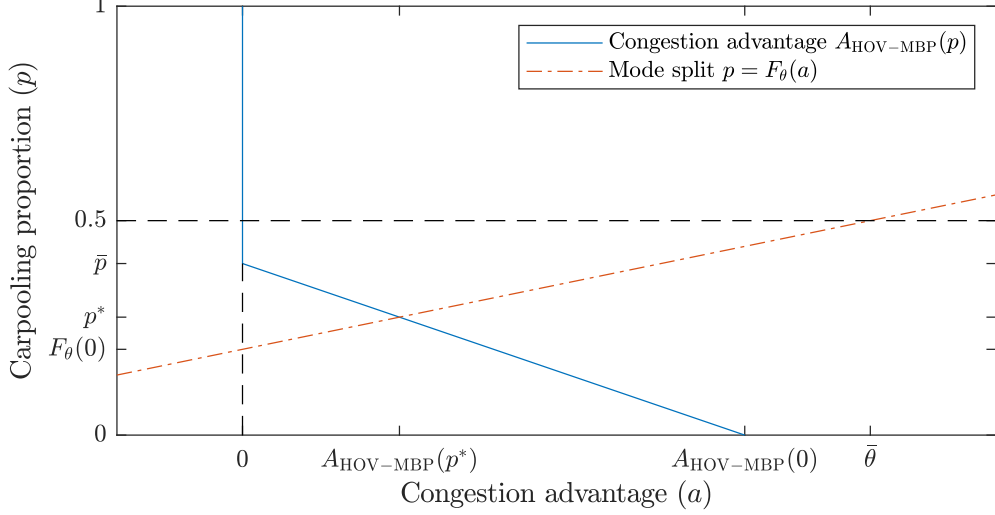
Let us now consider the case where the population size  $N$  is given, but the number of vehicles  $\tilde{N}$  may change as users choose whether to carpool or not. As explained in Appendix D, many complex carpooling models have been proposed in the literature but as our focus here is rather on the interaction between carpooling, priority and departure time choice, we prefer a parsimonious model. Users take this second decision by weighing the difference in congestion cost against their own preferences for the two modes. Following Konishi and Mun (2010), we model this with an individual specific inconvenience  $\theta \in \mathbb{R}$  incurred when carpooling. This cost can be positive or negative, and accounts for inconveniences such as the need to detour, the privacy loss, or the extra organizational load, as well as for carpooling advantages, such as the opportunity to socialize or the sharing of fixed travel costs.

**Assumption 6.** *The carpooling inconvenience  $\theta$  is independent of the schedule preferences. The support of its probability density function  $f_\theta$  is an interval  $[\theta^-, \theta^+] \subset \mathbb{R}$ , with  $\theta^- < 0$ , and  $\theta^+ > 0$  large enough to ensure that some users never carpool.*

Letting  $F_\theta$  denote the cumulative distribution function of the carpooling inconvenience, Assumption 6 ensures that there exists a proportion  $F_\theta(0) > 0$  of “natural carpools”, who carpool even without the HOV-MBP scheme.<sup>9</sup> To improve readability, we also denote  $\theta(p)$  the inverse of the restriction of  $F_\theta$  to  $[\theta^-, \theta^+]$ . In other words, a proportion  $p$  of users has an inconvenience cost  $\theta \leq \theta(p)$ .

<sup>9</sup>This is different from Konishi and Mun (2010), where  $\theta$  was assumed to be always positive, such that no user would carpool in the absence of HOV lanes.





**Figure 9:** Equilibrium with endogenous proportion of efficient users

*Note:* With  $\alpha - \beta - \gamma$  preferences, a uniform distribution of the carpooling inconvenience  $\theta$  of mean  $\bar{\theta}$  and with  $S^P = S/3$  and  $g = 2$ .

### 5.3.2 Characterization of the user equilibrium

When users have homogeneous schedule preferences, the equilibrium proportion of carpoolers must satisfy  $p = F_\theta(a)$ , where  $a$  denotes the congestion cost savings derived from being prioritized. Yet, with HOV-MBP and a given preemptable capacity  $S^P$ , this advantage  $a$  depends on the proportion  $p$  of carpoolers via:

$$a = A_{\text{HOV-MBP}}(p) = \begin{cases} C\left(\frac{\tilde{N}(p)}{S}\right) - C\left(\frac{pN}{gS^P}\right), & \text{if } p \in [0, \bar{p}], \\ 0, & \text{if } p \in [\bar{p}, 1], \end{cases}$$

where  $\tilde{N}(p) = N/\bar{g} = N(p/g + 1 - p)$ . Thus, solving for the equilibrium reduces to solving the fixed point problem  $p = F_\theta(A_{\text{HOV-MBP}}(p))$ . This problem admits a unique solution, as illustrated in Fig. 9 and formalized hereafter.

**Lemma 1.** *Let assumptions 1, 2, 3, 4, and 6 hold. Assume that the population has homogeneous schedule preferences and that carpoolers benefit from a HOV-MBP scheme with  $S^P \in (0, S)$ . There exists a unique equilibrium proportion of carpoolers  $p^*(S^P)$ . It is equal to  $F_\theta(0)$  if  $\bar{p} \leq F_\theta(0)$ , and it belongs to  $(F_\theta(0), \bar{p})$  otherwise.*

Lemma 1 simply states that if the capacity allocated to priority users is large enough (such that  $F_\theta(0) < \bar{p}$ ), the critical user that is indifferent between carpooling and driving alone has a positive carpooling inconvenience.

It is difficult to provide a more precise characterization of the equilibrium in the general case, but some interesting results can be derived by assuming  $\alpha - \beta - \gamma$  preferences and  $S^P = S$ . Recall from Section 4.2.1 that in this case, the reduced form cost function is simply  $C(\tau) = \delta\tau$ . The total cost is then:

$$\begin{aligned} \text{TC}(p) &= pNC\left(\frac{pN}{gS}\right) + (1-p)NC\left(\frac{pN}{gS} + \frac{(1-p)N}{S}\right) + N \int_0^p \theta(u) du \\ &= \frac{\delta N^2}{S} \left( (1-p)^2 + \frac{p}{g} \right) + N \int_0^p \theta(u) du. \end{aligned} \quad (13)$$

This is a sum of convex functions including one strictly convex component,  $(1-p)^2$ , so TC is strictly convex as well and admits a unique minimum.

In order to position the user equilibrium with respect to the proportion  $p$  minimizing the total cost, let us examine the sign of  $\frac{dTC}{dp}$  at the user equilibrium. By differentiating Eq. (13):

$$\begin{aligned}\frac{dTC}{dp}(p^*) &= \frac{\delta N^2}{S} \left( -2(1-p^*) + \frac{1}{g} \right) + N\theta(p^*) \\ &= \frac{\delta N^2}{S} \left( p^* - \left( 1 - \frac{1}{g} \right) \right),\end{aligned}\tag{14}$$

where we rely on the equation  $\theta(p^*) = \delta(1-p^*)N/S$  that derives from the User Equilibrium condition. Eq. (14) implies that the proportion of *users* that carpool at equilibrium is optimal iff  $p^* = 1-g^{-1}$ , which translates into a proportion of *vehicles*  $q^* = p^*/(p^* + (1-p^*)g) = (g-1)/(2g-1)$ .

Let us now analyze a few special cases. The limit case  $g = 1$  corresponds to a situation where “carpoolers” actually have the same average occupancy as other vehicles. In this case, the critical value of  $p^*$  is 0. This implies that these not-so-efficient “carpoolers”, if they exist, are always too numerous at equilibrium. This is because the users that switch to carpooling impose a negative externality on those that were already carpooling, without influencing the cost of non-prioritized users. With  $g = 2$ , the critical value of  $p^*$  is  $p^* = 0.5$ , which corresponds to  $q^* = 1/3$ . Thus, if we assume that all carpools have two occupants and that users have homogeneous  $\alpha - \beta - \gamma$  preferences, we know that as long as carpools represent less than one third of all vehicles at equilibrium, the total cost could be reduced by further encouraging carpooling. If the proportion of carpools is larger however, too many vehicles are prioritized at equilibrium. This implies that the loss associated to a poor compartmentalization outweighs those obtained from the reduced number of vehicles. As  $g$  further increases, the critical value of  $p^*$  tends towards 1 and that of  $q^*$  towards 0.5. Thus, regardless of the carpool occupancy, it is never optimal to have more than 50 % of priority vehicles at a bottleneck with the HOV-MBP policy. In cases where this threshold would be exceeded, a more stringent criteria for priority might be worth considering (e.g. a larger minimum vehicle occupancy).

### 5.3.3 Comparison with static HOV lanes

Let  $p_{\text{HOV-MBP}}(S^P)$  denote the user equilibrium associated to the HOV-MBP policy. A similar fixed point problem arises when carpools are provided with a *static* HOV lane of capacity  $S^P$ . In that case, the congestion advantage depends the carpooling proportion via

$$a = A_{\text{static}}(p) = \begin{cases} C\left(\frac{N(1-p)}{S-S^P}\right) - C\left(\frac{pN}{gS^P}\right), & \text{if } p \in [0, \bar{p}], \\ 0, & \text{if } p \in [\bar{p}, 1]. \end{cases}$$

This problem also admits a unique solution, as stated hereafter.

**Lemma 2.** *Let assumptions 1, 2, 3, 4, and 6 hold. Assume that the population has homogeneous schedule preferences and that carpools benefit from a static reserved lane, of capacity  $S^P \in (0, S)$ . There exists a unique equilibrium proportion of carpools  $p_{\text{static}}(S^P)$ . It is equal to  $F_\theta(0)$  if  $\bar{p} \leq F_\theta(0)$ , and it belongs to  $(F_\theta(0), \bar{p})$  otherwise.*

Let us now compare the individual costs under these two schemes. Let  $C_{\text{static}}^P(S^P, p)$ ,  $C_{\text{HOV-MBP}}^P(S^P, p)$ ,  $C_{\text{static}}^{\text{NP}}(S^P, p)$  and  $C_{\text{HOV-MBP}}^{\text{NP}}(S^P, p)$  denote the costs of priority and non-priority users with static HOV lanes and HOV-MBP, when a proportion  $p$  of the population is carpooling.

**Proposition 7.** *Let assumptions 1, 2, 3, 4, and 6 hold, and assume that the population has homogeneous schedule preferences. For all  $S^P \in (0, S)$ ,*

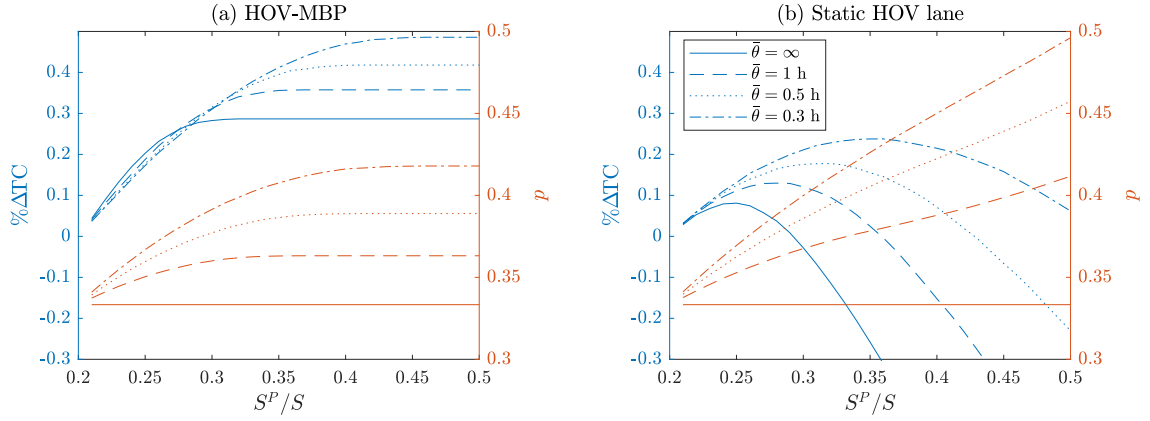
$$p_{\text{static}}(S^P) \geq p_{\text{HOV-MBP}}(S^P),\tag{15}$$

$$C_{\text{static}}^P(S^P, p_{\text{static}}(S^P)) \geq C_{\text{HOV-MBP}}^P(S^P, p_{\text{HOV-MBP}}(S^P)),\tag{16}$$

$$C_{\text{static}}^{\text{NP}}(S^P, p_{\text{static}}(S^P)) \geq C_{\text{HOV-MBP}}^{\text{NP}}(S^P, p_{\text{HOV-MBP}}(S^P)).\tag{17}$$

The proofs of Lemmas 1, 2 and Proposition 7 are relegated to Appendix C.

Proposition 7 confirms the superiority of the HOV-MBP scheme over its static counterpart and in the process, highlights why HOV priority schemes should not be evaluated solely based on the



**Figure 10:** Relative total cost savings and proportion of carpoolers with the HOV-MBP scheme and static HOV lanes

*Note:* Results obtained with  $N$  such that  $\tilde{N}/S = 2.5$  h when the only carpoolers are the natural ones, discrete times spanning over a 5-hour period centered around the mean  $t^*$  (time step: 1 min).

modal split they generate, but also on the overall efficiency. Here, a larger proportion shifts to carpooling under the static scheme than under HOV-MBP, but all users are weakly better-off under HOV-MBP than under a static scheme.

## 5.4 Numerical Example

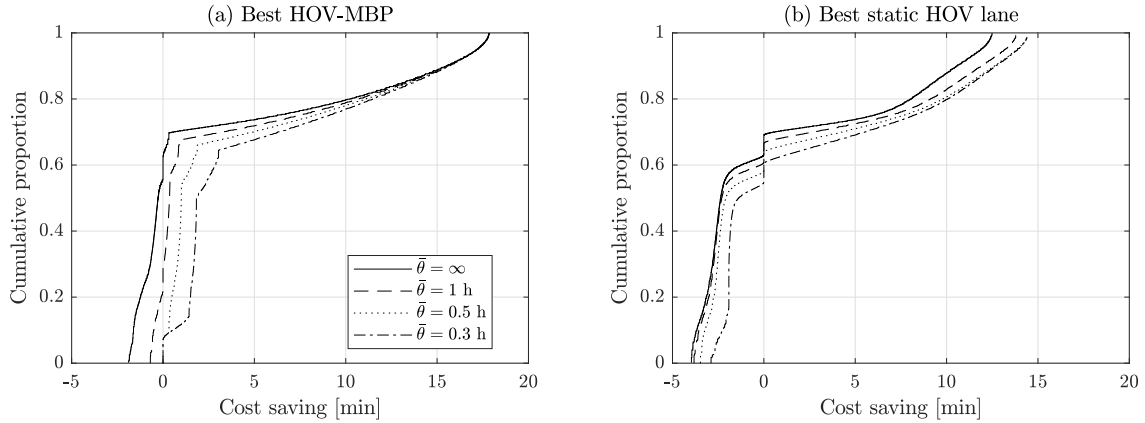
This case study replicates the same simulations as in Section 4.4, but with the HOV-MBP scheme and a conventional static reserved lane. We analyze the consequences at the global and individual levels, with both exogenous and endogenous mixes of vehicle occupancies.

Note that the previous analytical results establishing existence and uniqueness of equilibrium costs under metering-based priority do not apply with both an endogenous mix and heterogeneous schedule preferences. We believe that unique equilibrium costs still exist but we have not been able to demonstrate it under these general assumptions. Yet, since all the simulations we ran numerically converged (using the algorithms described in Appendix E), we leave aside this technical difficulty and still present numerical results.

The carpool occupancy  $g$  is set to 2 and the proportion of natural carpoolers  $F_{\theta}(0)$  to  $1/3$  in the simulations hereafter. This corresponds to a proportion  $q_0 = 20\%$  of vehicles, in line with real world observations for work trips (Federal Highway Administration 2017). More people may switch to carpooling if the priority scheme offers a sufficiently large advantage. In general, the total modal shift is determined by the slope  $f_{\theta}(0) = F'_{\theta}(0)$ , and by the higher order derivatives of  $F_{\theta}$  at 0. We discard here the higher-order effects by assuming a uniform distribution for  $\theta$ . Given the constraint  $F_{\theta}(0) = 1/3$ , we only need one additional parameter to fully determine the distribution. We choose  $\bar{\theta}$ , the mean carpooling inconvenience. The slope is then  $f_{\theta}(0) = (F(\bar{\theta}) - F(0))/(\bar{\theta} - 0) = (1/2 - 1/3)/\bar{\theta}$ . There are very few empirical estimates of the average carpooling inconvenience, but Small, Winston, and Yan (2006) suggests that a realistic value is likely to be  $\bar{\theta} = 0.5$  h. In other words, half of users would carpool if we could guarantee a travel time reduction of 30 min. Yet, the simulations hereafter are replicated for various values of  $\bar{\theta}$ , ranging from 0.3 h (18 min) to infinity<sup>10</sup> to account for the uncertainty about the exact value of  $\bar{\theta}$ .

Figure 10 shows the relative total cost savings obtained under the HOV-MBP scheme and with static HOV lanes. Note that the total cost is now a combination of queueing delays, schedule delays, and carpooling inconveniences. With HOV-MBP, the relative total cost savings increases with  $S^P$  until it reaches a threshold. The value of  $\% \Delta \text{TC}$  at this threshold and the smallest  $S^P$  for which it is first reached depend on the mean carpooling inconvenience  $\bar{\theta}$ . In this specific example, an exogenous mix ( $\bar{\theta} = \infty$ ) leads to  $\% \Delta \text{TC} \simeq 29\%$  for  $S^P/S \geq 32\%$ , while the more realistic estimate ( $\bar{\theta} = 0.5$  h) leads to  $\% \Delta \text{TC} \simeq 42\%$  for  $S^P/S \geq 41\%$ . In comparison, the maximum cost savings with the R-MBP scheme were 22% for  $S^P/S = 32\%$  and 27% for  $S^P/S = 41\%$  (see Fig 10a).

<sup>10</sup>The case  $\bar{\theta} = \infty$  is characterized by  $F_{\theta}(\theta) = F_{\theta}(0)$  for all  $\theta > 0$ , i.e. an exogenous mix.



**Figure 11:** Distribution of individual cost savings

*Note:* Results shown for the socially optimal HOV-MBP schemes and static HOV lanes from Fig. 10.

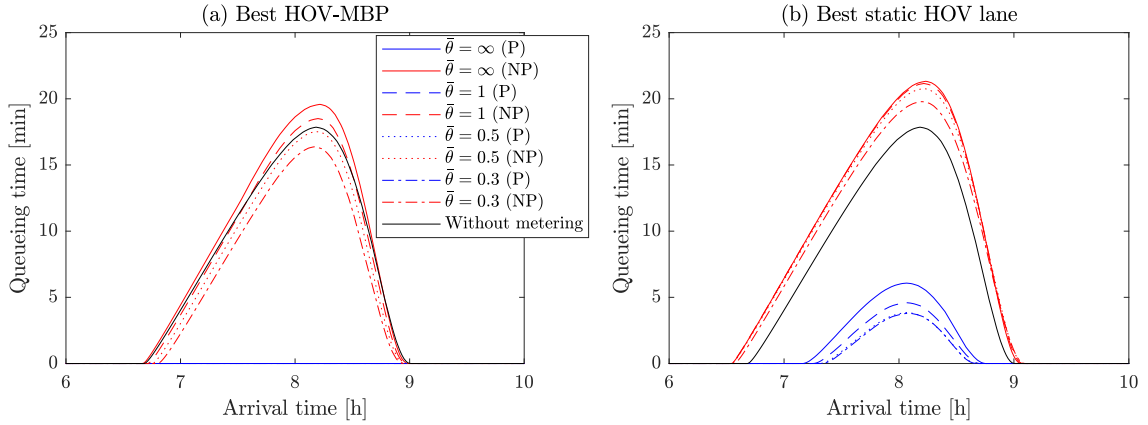
With static HOV lanes, the maximum relative total cost savings are much smaller, and are only reached for a specific capacity split  $S^P/S$ . This is problematic as static capacity splits can typically only be chosen among a small set of fractions ( $\{1/m, 2/m, \dots, (m-1)/m\}$ , where  $m$  is the number of lanes). With  $\bar{\theta} = 0.5$  h,  $\% \Delta \text{TC}$  reaches a maximum of 18% for  $(S^P/S) \simeq 32\%$ . This is not even half the maximum cost savings achieved with HOV-MBP for the same  $\bar{\theta}$ . Static HOV lanes perform particularly poorly when the average carpooling inconvenience  $\bar{\theta}$  is large. With an exogenous mode split ( $\bar{\theta} = \infty$ ), static HOV lanes only reduce the total cost for  $(S^P/S) \in [20\%, 29\%]$ , and increase it outside of this interval.

Let us now analyze the distributional consequences of HOV-MBP and static reserved lanes, with the socially optimal preemptable/reserved capacities  $S^P$  found in Fig. 10. Our analysis relies on two figures. Fig. 11 shows the cumulative density functions of cost savings (including individual carpooling inconveniences) for both schemes, while Fig. 12 shows the associated queueing times, thereby facilitating the interpretation. Although user type is not displayed here to save space, some information is relatively easy to infer from these two figures. First, note in Fig. 11 that there is always a positive measure of users with exactly zero cost saving. These have a desired arrival time out of the peak period both with and without the priority scheme (either HOV-MBP or static lanes), and are thus indifferent. Then, one can clearly distinguish two other groups in Figs. 11a and 11b, separated by a kink in the cumulative proportion (occurring between 60% and 70%). The group of users above the kink (who thus represent about 30-40% of the population) are all better-off and correspond to carpoolers while the group below (except those that are indifferent) are all solo-drivers. One can verify that these proportions are in agreement with Fig. 10.

When the lane reservation is static, solo-drivers are all either worse-off or indifferent, depending on whether they travel during the peak period or not. This is consistent with the corresponding queueing times (NP) in Fig. 12b, which are systematically greater than (or equal to) those without metering. A few solo-drivers are better-off with HOV-MBP and  $\bar{\theta} = \infty$ , because the boundaries of the congested period on non-priority lanes are slightly shifted towards the early morning (hardly visible in Fig. 12a). With HOV-MBP and  $\bar{\theta} < \infty$  however, the modal shift reduces the total length of the peak period for non-priority users, so that many (if not all) solo-drivers are better-off. While the ordering effect already reduces drastically the loss/gain ratio (from 29% with R-MBP to 14.3% with HOV-MBP and  $\bar{\theta} = \infty$ ), the modal shift can reduce this ratio even further: it equals 2.4% for  $\bar{\theta} = 1$  h, and reaches 0% for  $\bar{\theta} = 0.5$  h and 0.3 h. In these last cases, the addition of the ordering and modal shift effects to the original compartmentalization effect is sufficient to restore a Pareto-improvement (all users have non-negative cost savings).

In comparison, the loss/gain ratio is much larger with static HOV lanes. With the socially optimal reserved capacity, it is equal to 66.8% for  $\bar{\theta} = \infty$ , 54.4% for  $\bar{\theta} = 1$  h, 43.9% for  $\bar{\theta} = 0.5$  h and 29.3% for  $\bar{\theta} = 0.3$  h. These ratios are of the same magnitude as those obtained with the R-MBP scheme, or even larger.

Note that the modal shift and absolute cost savings strongly depend on the initial congestion cost. Here, the maximum congestion cost without metering-based priority is equivalent to approximately



**Figure 12:** Queueing times

*Note:* Results shown for the socially optimal HOV-MBP schemes and static HOV lanes from Fig. 10.

18 min (0.3 h) of waiting time. This is a rather long recurrent delay for a single bottleneck, but it is very reasonable if we consider all the congestion delays experienced over a trip during the peak period.

## 6 Discussion

This paper examined the potential of metering-based priority to alleviate congestion in a socially acceptable way. We have shown that the homogeneity hypothesis previously made in the literature concealed the adverse effects of random priority schemes. With heterogeneous preferences, random priority leaves many non-prioritized users worse off. The relative cost savings achieved by random priority schemes are also affected by heterogeneity: they tend to decrease with heterogeneity in flexibility, but may increase with heterogeneity in desired arrival times.

The drawbacks of priority schemes can be alleviated by prioritizing targeted users. Depending on the chosen criterion, targeted priority may induce an additional ordering effect and/or a modal shift. Both contribute to improving the acceptability of priority schemes: the ordering effect amplifies the gains and reduces the losses, while the modal shift alleviates congestion for non-priority users. On the other hand, targeted priority imposes a constraint on the proportion of the population that can be prioritized, thereby restricting the total cost savings. We believe however that this is a mild restriction compared to the gains it provides in terms of distributional consequences.

The range of situations in which MBP schemes could be implemented is quite large, but additional network aspects need to be considered. It is known since Rogers (1985) that implementing metering by-passes only at a few on-ramps of a highway leads solo-drivers to re-route towards other on-ramps. The same phenomenon would most likely arise in cities, if priority schemes were to be implemented at randomly chosen intersections. This is another argument in favor of combining MBP with other metering strategies such as perimeter control (Ramezani, Haddad, and Geroliminis 2015), which inherently control all the access points to a zone.

Another topic deserving further research is the identification of selection criteria that are not only efficient and socially acceptable, but also easy to enforce. Vehicle occupancy is certainly a strong candidate, but one may also consider discriminating vehicles based on their destination or their trip length (as in Daganzo and Lehe 2015). Alternatively, transport authorities may choose to prioritize low-emission modes, to combine operational benefits with environmental ones.

Depending on the chosen criterion, there might also be issues related to demand elasticity: persons that do not utilize this bottleneck in the laissez-faire regime might start to do so when they are prioritized. Although this is analytically difficult to quantify (see Fosgerau (2011) for some first derivations), this effect might threaten the Pareto improvement. The consequences of demand elasticity on the social cost are more ambiguous, in particular for cases like HOV-MBP. The laissez-faire regime has the advantage of generating a self-selection, where those that travel are also those who are ready to queue the most. This self-selection is however independent of the individual contribution to congestion, and therefore suboptimal.

If metering-based priority is actually based on vehicle occupancy, the matching process would deserve to be refined to model the trade-off between the similarity of schedule preferences, affinity between carpoolers and spatial considerations. This then points towards two other research directions, as (i) the aggregation of carpoolers with different desired arrival times may induce a concentration close to the peak (as in Picard, Dantan, and de Palma 2018) and (ii) detours caused by carpooling may also contribute to congestion.

## Acknowledgments

The authors would like to thank Takamasa Iryo for useful explanations regarding results included in Iryo and Yoshii (2007) and Akamatsu et al. (2018), about existence and uniqueness of equilibrium. André de Palma benefits from Délégation CNRS (F), 2019-2020. Nikolas Geroliminis and Raphaël Lamotte acknowledge the financial support provided by ERC starting grant METAFERW.

## A Existence of Equilibrium and Uniqueness of Equilibrium Costs with a Bottleneck of Time-Varying Capacity

This appendix explains how existence and uniqueness results established for a bottleneck of constant capacity can be transposed to the case with a time-varying capacity. Specifically, we consider the case where Assumption 2 is replaced by the following reformulation.

**Assumption 7.** *The bottleneck is accessible with a piece-wise continuous capacity  $S(t) > 0$  during a time interval  $\mathcal{T} = (t_0, t_e)$ , such that  $\int_{\mathcal{T}} S(t) dt > \bar{N}$  and  $t_j^* \in \mathcal{T} \forall i = 1, \dots, J$ .*

### A.1 Existence of Equilibrium

Two approaches compatible with our assumptions have been proposed in the literature: one by Lindsey (2004), and another by Iryo and Yoshii (2007) and Akamatsu et al. (2018). We choose to transpose the second, because it also provides the intuition behind the algorithms presented in Appendix E. This section refers hereafter more specifically to the work of Akamatsu et al. (2018) because it allows for a continuum of arrival times, unlike Iryo and Yoshii (2007).

The method proposed by Akamatsu et al. (2018) consists in identifying a set of arrival rate functions  $(s_j)_{j \in \mathcal{J}}$  (at destination) and a queueing time function  $T$  that satisfy the optimal choice condition (A.1), the queueing condition (A.2) and that together, define departure rates (from the origin) that are compatible with the FIFO condition and do not entail mass departures. With a constant capacity  $S$ , the optimal choice and queueing conditions are the following:

$$\begin{cases} C_j = \alpha T(t) + SP_j(t) & \text{if } s_j(t) > 0 \\ C_j \leq \alpha T(t) + SP_j(t) & \text{if } s_j(t) = 0 \end{cases} \quad \forall j \in \mathcal{J}, t \in \mathcal{T} \quad (\text{A.1})$$

$$\begin{cases} \sum_{j=1}^J s_j(t) = S & \text{if } T(t) > 0 \\ \sum_{j=1}^J s_j(t) \leq S & \text{if } T(t) = 0 \end{cases} \quad \forall t \in \mathcal{T}. \quad (\text{A.2})$$

We follow the same approach here and leverage key results of Akamatsu et al. (2018) by transposing the problem in a different time coordinate system.

**Step 1.** Starting from a problem with time-dependent capacity (as in Assumption 7), we define the bijection  $\tilde{t} = \Phi(t)$ , where:

$$\begin{aligned} \Phi: \mathcal{R} &\rightarrow R \\ t &\mapsto \int_0^t S(\tau) d\tau. \end{aligned}$$

The transformation  $\Phi$  is such that the maximum number of passages at the bottleneck over any interval  $[\tilde{t}_1, \tilde{t}_2] \subset \Phi(\mathcal{T})$  is  $\tilde{t}_2 - \tilde{t}_1$ , i.e. the bottleneck has capacity 1 in the time system  $\tilde{t}$ . The cost functions of the different groups keep the same form, except that the schedule penalty is now

$\tilde{\text{SP}}_j(\tilde{t}) = \text{SP}_j(\Phi^{-1}(\tilde{t}))$ . These functions inherit the continuity from the functions  $\text{SP}_j$  and they are minimized for  $\tilde{t}_j^* = \Phi(t_j^*)$ . Only the condition on the rate of change (1) may not be inherited, but this is addressed in step 3.

**Step 2.** After transposing the problem in the  $\tilde{t}$  time system, one can directly apply the part of the demonstration in Akamatsu et al. (2018) that shows the existence of functions  $(\tilde{s}_j)_{j \in \mathcal{J}}$  and  $\tilde{T}$  satisfying the conditions (A.1) and (A.2) (but with  $\tilde{t}$ ,  $\Phi(\mathcal{T})$  and  $\tilde{\text{SP}}_j$  instead of  $t$ ,  $\mathcal{T}$  and  $\text{SP}_j$ ). This is done by relating the departure time choice problem to the following optimal transport problem:

$$[2\text{D-LP}(\tilde{s})] \underset{x \geq 0}{\text{minimize}} \quad Z(x) = \sum_{j=1}^J \int_{\Phi(\mathcal{T})} \frac{\tilde{\text{SP}}_j(\tilde{t})}{\alpha} \tilde{s}_j(\tilde{t}) \, d\tilde{t} \quad (\text{A.3a})$$

$$\text{subject to} \quad \sum_{j=1}^J \tilde{s}_j(\tilde{t}) \leq 1, \quad \forall \tilde{t} \in \Phi(\mathcal{T}), \quad (\text{A.3b})$$

$$\int_{\Phi(\mathcal{T})} \tilde{s}_j(\tilde{t}) \, d\tilde{t} = N_j, \quad \forall j \in \mathcal{J}. \quad (\text{A.3c})$$

The optimal transport literature has established that provided that the functions  $\tilde{\text{SP}}_j$  are lower-semicontinuous, optimal solutions exist (see e.g. Theorem 4.1 in Villani (2008), or Galichon (2016)). Since continuity is a stronger property, optimal solutions exist here as well. Then, Akamatsu et al. (2018) shows that if we take for  $T(t)$  and  $C_j$  the Lagrange multipliers of Eqs. (A.3b) and (A.3c), the solutions of  $[2\text{D-LP}(\tilde{s})]$  necessarily satisfy the optimal choice condition (A.1) and the queueing condition (A.2) (again, with the variables and functions of the  $\tilde{t}$  time system). If we denote  $\tilde{s}_j$ ,  $\tilde{C}_j$  and  $\tilde{T}$  the solutions with the constant capacity bottleneck in the  $\tilde{t}$  time system, it is then straightforward to show that the functions defined in the original time system for all  $t \in \mathcal{T}$  and  $j \in \mathcal{J}$  by  $s_j(t) = S(t)\tilde{s}_j(\Phi^{-1}(t))$  and  $T(t) = \tilde{T}(\Phi^{-1}(t))$  also satisfy the optimal choice condition (A.1) with  $C_j = \tilde{C}_j$  and the queueing condition (A.2), but with time-dependent capacity  $S(t)$ .

**Step 3.** The last step is to verify that the resulting departure rates are consistent with the First-In, First-Out condition and do not entail mass departures. Consider some group  $j \in \{1, \dots, J\}$  and two times  $t_1, t_2 \in \mathcal{T}$  such that  $t_1 < t_2$ ,  $s_j(t_1) > 0$  and  $s_j(t_2) > 0$ . The optimal choice condition (A.1) imposes that  $C_j = \alpha T(t_1) + \text{SP}_j(t_1) = \alpha T(t_2) + \text{SP}_j(t_2)$ . This can be rewritten

$$\frac{T(t_2) - T(t_1)}{t_2 - t_1} = - \frac{\text{SP}_j(t_2) - \text{SP}_j(t_1)}{\alpha(t_2 - t_1)},$$

such that condition (1) of Assumption 1 ensures that  $(T(t_2) - T(t_1))/(t_2 - t_1) > 1$ . Thus, the departure time of users passing at  $t_2$  (i.e.  $t_2 - T(t_2)$ ) is strictly smaller than the departure time of users passing at  $t_1$  (i.e.  $t_1 - T(t_1)$ ), which concludes.

## A.2 Uniqueness of Equilibrium Costs

Lindsey (2004) already proved that equilibrium costs have to be unique with a bottleneck of constant capacity. We revisit his proof hereafter with a time-dependent capacity.

The proof relies heavily on the concept of isocost curves. An isocost queueing time of a group  $i$  is a function  $t \mapsto T_j(t, C_j) \triangleq \alpha^{-1}(C_j - \text{SP}_j(t))$ . It represents the time that users of group  $j$  passing at time  $t$  should queue to experience the cost  $C_i$ . Letting  $\mathbf{C} = (C_j)_{j \in \mathcal{J}}$ , the interactions between different groups can then be captured with the following additional concept.

**Definition 1.** For any  $\mathcal{I} \subseteq \mathcal{J}$ , let  $t \mapsto T_{\mathcal{I}}(t, \mathbf{C}) \triangleq \max(0, \max\{T_j(t, C_j), j \in \mathcal{I}\})$  be the upper non-negative envelope of all the isocost curves belonging to groups of  $\mathcal{I}$ .

At equilibrium, the bottleneck dynamics impose that:

1. the bottleneck should be used at capacity at all times such that  $T_{\mathcal{J}}(t, \mathbf{C}) > 0$ ,
2. only the groups  $j$  such that  $T_j(t, C_j) = T_{\mathcal{J}}(t, \mathbf{C})$  can pass at time  $t \in \mathcal{T}$ .

Together, these principles imply that for any  $\mathcal{I} \subset \mathcal{J}$ ,

$$\int_{\mathcal{T}} [T_{\mathcal{I}}(t, \mathbf{C}) > T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C})] S(t) dt \leq \sum_{j \in \mathcal{I}} N_j \leq \int_{\mathcal{T}} [T_{\mathcal{I}}(t, \mathbf{C}) \geq T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C})] S(t) dt. \quad (\text{A.4})$$

We then proceed by contradiction and assume that there exist two distinct equilibrium cost vectors  $\mathbf{C}$  and  $\mathbf{C}'$ . Assume without loss of generality that there exists  $j \in \mathcal{J}$  such that  $C_j > C'_j$  and let  $\Delta = \max\{\alpha^{-1}(C_j - C'_j), j \in \mathcal{J}\}$ .

The first step of the proof consists in showing that all groups actually have the same normalized increase in cost, i.e.  $\Delta = \alpha^{-1}(C_j - C'_j)$  for all  $j \in \mathcal{J}$ . By contradiction, assume that  $\Delta = \alpha^{-1}(C_j - C'_j)$  only for  $j \in \mathcal{I} \subsetneq \mathcal{J}$ , and denote  $\Delta' = \max\{\alpha^{-1}(C_j - C'_j), j \in \mathcal{J} \setminus \mathcal{I}\}$ . By applying the left part of inequality (A.4) to  $\mathbf{C}$  and the right part to  $\mathbf{C}'$ , we obtain that

$$\int_{\mathcal{T}} [T_{\mathcal{I}}(t, \mathbf{C}) > T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C})] S(t) dt \leq \sum_{j \in \mathcal{I}} N_j \leq \int_{\mathcal{T}} [T_{\mathcal{I}}(t, \mathbf{C}') \geq T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C}')] S(t) dt. \quad (\text{A.5})$$

However,

$$\begin{aligned} \int_{\mathcal{T}} [T_{\mathcal{I}}(t, \mathbf{C}') \geq T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C}')] S(t) dt &\leq \int_{\mathcal{T}} [T_{\mathcal{I}}(t, \mathbf{C}' + \Delta') \geq T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C}' + \Delta')] S(t) dt \\ &\leq \int_{\mathcal{T}} [T_{\mathcal{I}}(t, \mathbf{C}' + \Delta') \geq T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C})] S(t) dt \\ &< \int_{\mathcal{T}} [T_{\mathcal{I}}(t, \mathbf{C}) \geq T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C})] S(t) dt. \end{aligned} \quad (\text{A.6})$$

The last inequality is strict because  $C'_j + \Delta' < C_j$  for all  $j \in \mathcal{I}$ . Since the continuity of the isocost curves is inherited by their envelopes, the set of times  $t \in \mathcal{T}$  such that  $T_{\mathcal{J} \setminus \mathcal{I}}(t, \mathbf{C}) \in (T_{\mathcal{I}}(t, \mathbf{C}' + \Delta'), T_{\mathcal{I}}(t, \mathbf{C}))$  has a positive measure. Thus, Eq. (A.6) contradicts Eq. (A.5).

A similar reasoning shows that the total population cannot keep the same size when all groups have a normalized cost increase  $\Delta$ . Indeed,  $\int_{\mathcal{T}} [T_{\mathcal{J}}(t, \mathbf{C}') \geq 0] S(t) dt < \int_{\mathcal{T}} [T_{\mathcal{J}}(t, \mathbf{C}' + \Delta) > 0] S(t) dt = \int_{\mathcal{T}} [T_{\mathcal{J}}(t, \mathbf{C}) > 0] S(t) dt$ , but applying Eq. (A.4) to all groups together implies that  $\int_{\mathcal{T}} [T_{\mathcal{J}}(t, \mathbf{C}) > 0] S(t) dt \leq \sum_{j \in \mathcal{I}} N_j \leq \int_{\mathcal{T}} [T_{\mathcal{J}}(t, \mathbf{C}') \geq 0] S(t) dt$ . Again, there is a contradiction.

## B Individual Costs and Reduced Form Cost Function

### B.1 Proof of Proposition 2

This section proves the expression of the individual costs stated in Proposition 2, in which metering-based priority is applied to a population with homogeneous schedule preferences. We rely for this on the following lemma, which provides an expression of the individual costs for a single population (no priority scheme), with a possibly time-dependent capacity.

**Lemma 3.** *Consider a problem satisfying Assumptions 1 and 7, but with a homogeneous population ( $J = 1$ ). In deterministic departure-time user equilibrium, the individual cost is*

$$\tilde{C}(\tilde{N}, S) = \sup \left\{ c \in \mathcal{C} \mid \int_{\mathcal{T}} [\text{SP}(t) < c] S(t) dt \leq \tilde{N} \right\}, \quad (\text{B.1})$$

where  $\mathcal{C}$  denotes the image of  $\mathcal{T}$  under SP.

*Proof.* Proof. Let  $s(t)$  denote the flow passing the bottleneck at time  $t$ . Equilibrium requires that no user can be better-off by changing departure time. Thus, all times  $t$  such that  $\text{SP}(t) < \tilde{C}(\tilde{N}, S)$  must satisfy  $T(t) > 0$ , and therefore  $s(t) = S(t)$ . By integrating over time and using the Iverson bracket notation, we have that  $\tilde{C}(\tilde{N}, S)$  must be part of the set

$$Z_1 = \left\{ c \in \mathbb{R} \mid \int_{\mathcal{T}} [\text{SP}(t) < c] S(t) dt \leq \tilde{N} \right\}.$$



Similarly, since the queueing time  $T$  cannot be strictly negative, all times such that  $s(t) > 0$  must satisfy  $\text{SP}(t) \leq \tilde{C}(\tilde{N}, S)$ . By integrating over time,  $\tilde{C}(\tilde{N}, S)$  must also be part of the set

$$Z_2 = \left\{ c \in \mathbb{R} \mid \int_{\mathcal{T}} [\text{SP}(t) \leq c] S(t) dt \geq \tilde{N} \right\}.$$

Besides, by assumption 7, we can take  $t \in \mathcal{T}$  such that  $s(t) < S(t)$ , and therefore  $T(t) = 0$  and  $\text{SP}(t) \geq \tilde{C}(\tilde{N}, S)$ . Since  $N > 0$ , we can find  $t'$  such that  $s(t') > 0$  and therefore  $\text{SP}(t') \leq \tilde{C}(\tilde{N}, S)$ . Since  $\text{SP}$  is continuous over the closed interval defined by  $t$  and  $t'$ , we obtain that  $\tilde{C}(\tilde{N}, S) \in \mathcal{C}$ .

Let us now assume that we have  $c_1 \in Z_1 \cap \mathcal{C}$ ,  $c_2 \in \mathcal{C}$  and  $c_2 < c_1$ . Clearly,

$$\int_{\mathcal{T}} [\text{SP}(t) \leq c_2] S(t) dt = \int_{\mathcal{T}} [\text{SP}(t) < c_1] S(t) dt - \int_{\mathcal{T}} [\text{SP}(t) \in (c_2, c_1)] S(t) dt.$$

Since  $c_1 \in Z_1$ ,  $\int_{\mathcal{T}} [\text{SP}(t) < c_1] S(t) dt \leq \tilde{N}$ , and since both  $c_1$  and  $c_2$  belong to  $\mathcal{C}$  and  $D$  is continuous,  $\int_{\mathcal{T}} [\text{SP}(t) \in (c_2, c_1)] S(t) dt > 0$ . Thus,  $\int_{\mathcal{T}} [\text{SP}(t) \leq c_2] S(t) dt < \tilde{N}$ , i.e.  $c_2 \notin Z_2$ . The contrapositive is that all elements of  $Z_2 \cap \mathcal{C}$  are greater than all elements of  $Z_1 \cap \mathcal{C}$ . Since  $\tilde{C}(\tilde{N}, S)$  belongs to both, it is necessarily the supremum of  $Z_1 \cap \mathcal{C}$ .  $\square$

In the particular case where the capacity  $S(t)$  is constant, the individual cost  $\tilde{C}(\tilde{N}, S)$  can be rewritten as a function mapping a demand-to-capacity ratio  $\tau = \tilde{N}/S$  to a cost, as in Eq. (5). We now have all the tools required to prove Proposition 2.

Since priority users do not compete with non-priority ones and experience a constant capacity  $S^{\text{P}}$ , their individual cost is simply  $C^{\text{P}} = C(p\tilde{N}/S^{\text{P}})$ . Non-priority users then compete among themselves for the remaining time-dependent capacity,  $S - S^{\text{P}}(t)$ . Using Lemma 3, their individual cost is

$$\begin{aligned} C^{\text{NP}} &= \tilde{C}((1-p)N, S - S^{\text{P}}) \\ &= \sup \left\{ c \in \mathcal{C} \mid \int_{\mathcal{T}} [\text{SP}(t) < c] S dt - \int_{\mathcal{T}} [\text{SP}(t) < c] S^{\text{P}}(t) dt \leq (1-p)\tilde{N} \right\}. \end{aligned}$$

Since  $S^{\text{P}}(t) > 0 \Rightarrow \text{SP}(t) \leq C^{\text{P}}$  and  $C^{\text{P}} \leq C^{\text{NP}}$ ,  $\int_{\mathcal{T}} [\text{SP}(t) < C^{\text{NP}}] S^{\text{P}}(t) dt = \int_{\mathcal{T}} S^{\text{P}}(t) dt = p\tilde{N}$ . Thus,

$$C^{\text{NP}} = \sup \left\{ c \in \mathcal{C} \mid \int_{\mathcal{T}} [\text{SP}(t) < c] S dt \leq \tilde{N} \right\} = C(\tilde{N}/S).$$

## B.2 Conditions for a Convex Reduced Form Cost Function

The general form of  $C(\tau)$  provided by Eq. (5) implies that  $C$  is non-decreasing, but it is not very intuitive. If we further assume that  $\text{SP}$  is strictly decreasing for  $t < t^*$  and strictly increasing for  $t > t^*$ , Eq. (5) reduces to

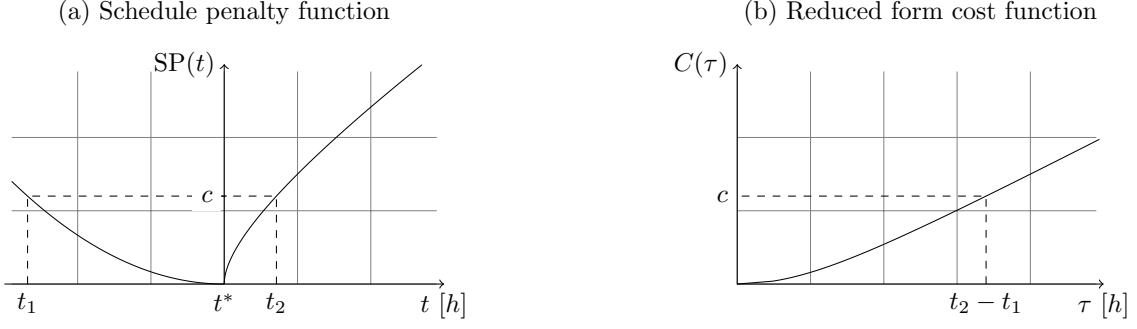
$$\int_{\mathcal{T}} [\text{SP}(t) \leq C(\tilde{N}/S)] dt = \tilde{N}/S.$$

With the  $\alpha - \beta - \gamma$  preferences introduced in Section 3.1, the solution to this equation is simply  $C(\tau) = \delta\tau$ , where  $\delta = \beta\gamma/(\beta + \gamma)$ .

Let us now examine the conditions under which  $C$  is convex (or concave). To obtain some intuition, we focus on a situation like the one illustrated in Fig. 13a, where  $\text{SP}$  is continuously differentiable on  $(t_0, t^*) \cup (t^*, t_e)$ , with  $\text{SP}'(t) < 0$  for early arrivals ( $t < t^*$ ) and  $\text{SP}'(t) > 0$  for late arrivals ( $t > t^*$ ).

In such a situation, the congested period at equilibrium always consists of a single interval  $(t_1, t_2)$ . If the demand is small enough for both  $t_1$  and  $t_2$  to be different from  $t_0$  and  $t_e$ , these bounds can be defined as functions of the equilibrium cost  $c$  as the only times in  $(t_0, t^*)$  and  $(t^*, t_e)$  satisfying  $\text{SP}(t_1(c)) = \text{SP}(t_2(c)) = c$ . Differentiating this with respect to  $c$  leads to  $t_1'(c) = (\text{SP}'(t_1))^{-1}$  and  $t_2'(c) = (\text{SP}'(t_2))^{-1}$ . Note also that  $C^{-1}(c) = \tau = t_2(c) - t_1(c)$ . Combining these results leads to:

$$\begin{aligned} C'(\tau) &= [C^{-1'}(C(\tau))]^{-1} \\ &= [t_2'(C(\tau)) - t_1'(C(\tau))]^{-1} \\ &= \left[ (\text{SP}'(t_2(C(\tau))))^{-1} - (\text{SP}'(t_1(C(\tau))))^{-1} \right]^{-1}. \end{aligned}$$



**Figure 13:** A non-convex schedule penalty and its convex reduced cost function

If SP is convex (resp. concave) on both  $(t_0, t^*)$  and  $(t^*, t_e)$ , the functions  $\tau \mapsto \text{SP}'(t_2(C(\tau)))$  and  $\tau \mapsto -\text{SP}'(t_1(C(\tau)))$  are both positive and increasing (decreasing), so  $C'$  is also increasing (decreasing) and  $C$  is therefore convex (concave). Yet, if one term is convex and the other concave (as in Fig. 13a), it is still possible that the derivative of one always dominates the derivative of the other, so that the overall function  $C$  is convex (as in Fig. 13b) or concave.

## C Other Proofs

*Proof.* Proof of Proposition 3. Differentiating Eq. (7) with respect to  $q$  leads to

$$\frac{\partial \text{TC}}{\partial q}(S^{\text{P}}, q) = NC \left( \frac{q\tilde{N}}{S^{\text{P}}} \right) + qNC' \left( \frac{q\tilde{N}}{S^{\text{P}}} \right) - NC \left( \frac{\tilde{N}}{S} \right).$$

For all  $q \in (0, \bar{q}]$ , this is of the same sign as

$$\frac{1}{q} \frac{\partial \text{TC}}{\partial q}(S^{\text{P}}, q) = \frac{C \left( \frac{q\tilde{N}}{S^{\text{P}}} \right) - C \left( \frac{\tilde{N}}{S} \right)}{q} + C' \left( \frac{q\tilde{N}}{S^{\text{P}}} \right).$$

The convexity of  $C$  implies that both  $\left( C(q\tilde{N}/S^{\text{P}}) - C(\tilde{N}/S) \right) q^{-1}$  and  $C'(q\tilde{N}/S^{\text{P}})$  increase with  $q$ , and so does their sum. Besides,  $\frac{\partial \text{TC}}{\partial q}(S^{\text{P}}, 0) = N(C(0) - C(\tilde{N}/S)) < 0$  and  $\frac{\partial \text{TC}}{\partial q}(S^{\text{P}}, \bar{q}) = \bar{q}NC'(\tilde{N}/S) > 0$ . Thus, for any given  $S^{\text{P}}$ , there exists a unique  $q^o \in (0, \bar{q})$  minimizing  $q \mapsto \text{TC}(S^{\text{P}}, q)$  and  $\text{TC}(S^{\text{P}}, q)$  decreases with  $q$  on  $q \in [0, q^o]$  and increases with  $q$  on  $q \in [q^o, \bar{q}]$ .

Let us now focus on the case  $q = \bar{q}/2$ . Recall that  $\bar{q}/S^{\text{P}} = 1/S$ . Thus,  $\left( C(\bar{q}\tilde{N}/(2S^{\text{P}})) - C(\tilde{N}/S) \right) (\bar{q}/2)^{-1}$  represents the negative of the average slope of the function  $q \mapsto C(q\tilde{N}/S^{\text{P}})$  between the points  $\bar{q}/2$  and  $\bar{q}$ . The convexity of  $C$  imposes that  $\left( C(\bar{q}\tilde{N}/(2S^{\text{P}})) - C(\tilde{N}/S) \right) (\bar{q}/2)^{-1} \geq C'(\tilde{N}/(2S))$ . Thus,  $\frac{\partial \text{TC}}{\partial q}(S^{\text{P}}, \bar{q}/2) \leq 0$ , and therefore  $(\bar{q}/2) \in (0, q^o]$ .

Finally, the convexity of  $C$  also imposes that  $C(\tilde{N}/(2S)) \leq (C(0) + C(\tilde{N}/S))/2 = C(\tilde{N}/S)/2$ , which means that  $\text{TC}(S^{\text{P}}, q^o) \leq \text{TC}(S^{\text{P}}, \bar{q}/2) \leq \bar{q}C(\tilde{N}/S)/4 + (1 - \bar{q}/2)C(\tilde{N}/S) = (1 - \bar{q}/4)C(\tilde{N}/S)$ . Thus, the maximum total cost savings is at least  $\bar{q}/4$ .  $\square$

*Proof.* Proof of Proposition 4. **Part a)** Note that with  $j = 1$ , the condition in Eq. (10) reduces to  $\frac{(1-q)\tilde{N}}{S - S^{\text{P}}} \leq \frac{q\tilde{N}}{S^{\text{P}}}$ , which is incompatible with  $q < \bar{q}$ . Thus,  $\tau_1^{\text{NP}} = \frac{q\tilde{N} + (1-q)\tilde{N}}{S} = \frac{\tilde{N}}{S}$  and  $C_1^{\text{NP}} = C_1^{\text{ref}}$ .

**Part b)** Let  $j \in \{2, \dots, J\}$ .  $q \in (0, \bar{q})$  implies that  $(1-q)/(S - S^{\text{P}}) > S$ . Thus, if

$$\frac{(1-q) \sum_{i=j}^J \tilde{N}_i}{S - S^{\text{P}}} \leq \frac{q\tilde{N}}{S^{\text{P}}},$$

then  $\tau_j^{\text{NP}} > (\sum_{i=j}^J \tilde{N}_i)/S = \tau_j$ . Otherwise,

$$\tau_j^{\text{NP}} = \frac{q\tilde{N} + (1-q) \sum_{i=j}^j \tilde{N}_i}{S} = \frac{q \sum_{i=1}^{j-1} \tilde{N}_i + \sum_{i=j}^j \tilde{N}_i}{S} > \frac{\sum_{i=j}^j \tilde{N}_i}{S} = \tau_j.$$

Since  $C$  is assumed strictly increasing,  $C_j^{\text{NP}} = \sum_{i=1}^j (k_i - k_{i-1})C(\tau_i^{\text{NP}}) > \sum_{i=1}^j (k_i - k_{i-1})C(\tau_i) = C_j^{\text{ref}}$ ,  $\forall j > 1$ .

**Part c)** The result is straightforward for priority and non-priority users of group  $j = 1$  (because  $\%C_1^{\text{P}} = \%C^{\text{P}}$  and  $\%C_1^{\text{NP}} = \%C^{\text{NP}}$ ) and it is a direct consequence of part a) for non-priority users of the groups  $j = 2, \dots, J$  (because  $\%C_j^{\text{NP}} > 1$  and  $\%C^{\text{NP}} = 1$ ). Only the priority users of the groups  $j = 2, \dots, J$  require some special care. There,

$$\%C_j^{\text{P}} = \frac{\sum_{i=1}^j (k_i - k_{i-1})C\left(q\left(\sum_{l=i}^J \tilde{N}_l\right)/S^{\text{P}}\right)}{\sum_{i=1}^j (k_i - k_{i-1})C\left(\left(\sum_{l=i}^J \tilde{N}_l\right)/S\right)}.$$

Since  $C\left(\left(\sum_{l=i}^J \tilde{N}_l\right)/S\right) > 0$  for all  $i = 1, \dots, j$ , this can be seen as the weighted average of the  $j$  ratios  $\left(C\left(q\left(\sum_{l=i}^J \tilde{N}_l\right)/S^{\text{P}}\right)\right) / \left(C\left(\left(\sum_{l=i}^J \tilde{N}_l\right)/S\right)\right)^{-1}$ . Since  $(qS/S^{\text{P}}) \in (0, 1)$ , the function  $x \mapsto C(qSx/S^{\text{P}})/C(x)$  is non-increasing for all  $x \geq 0$  and in particular:

$$\frac{C\left(q\left(\sum_{l=i}^J \tilde{N}_l\right)/S^{\text{P}}\right)}{C\left(\left(\sum_{l=i}^J \tilde{N}_l\right)/S\right)} \geq \frac{C(q\tilde{N}/S^{\text{P}})}{C(\tilde{N}/S)} = \%C^{\text{P}}.$$

The relative cost  $\%C_j^{\text{P}}$  is thus the weighted average of terms that are all greater than or equal to  $\%C^{\text{P}}$ .  $\square$

*Proof.* Proof of Proposition 6. The number of priority *vehicles* in each group  $j \in \mathcal{J}$  is the same with both schemes. As a consequence, the individual gains  $\Delta C_j^{\text{P}} = C_j^{\text{ref}} - C_j^{\text{P}}$  and  $\Delta C_j^{\text{NP}} = C_j^{\text{ref}} - C_j^{\text{NP}}$  experienced by priority and non-priority vehicles are also the same. Yet, the proportion of priority *users* is equal to  $qg/\bar{g}$  with HOV-MBP, but only to  $q$  with R-MBP. Since  $\Delta C_j^{\text{P}} \geq \Delta C_j^{\text{NP}}$ , we have indeed

$$\sum_{j \in \mathcal{J}} N_j (q\Delta C_j^{\text{P}} + (1-q)\Delta C_j^{\text{NP}}) \leq \sum_{j \in \mathcal{J}} N_j (q\frac{g}{\bar{g}}\Delta C_j^{\text{P}} + (1-q\frac{g}{\bar{g}})\Delta C_j^{\text{NP}}),$$

i.e., after division of both sides by  $\text{TC}^{\text{ref}}$ ,  $\% \Delta \text{TC}^{\text{R-MBP}} \leq \% \Delta \text{TC}^{\text{HOV-MBP}}$ .

In the case with homogeneous users, the total costs savings reduce to  $Nq\Delta C^{\text{P}}$  and  $Nq\frac{g}{\bar{g}}\Delta C_j^{\text{P}}$  respectively, and Eq. (11) follows directly.

In the case where flexibility is the only source of heterogeneity, all non-priority users have a cost at least as large as without the priority scheme ( $\Delta C_j^{\text{NP}} \leq 0$  for all  $j \in \mathcal{J}$ ). Thus,

$$\begin{aligned} \% \Delta \text{TC}^{\text{R-MBP}} &= \frac{\sum_{j \in \mathcal{J}} N_j (q\Delta C_j^{\text{P}} + (1-q)\Delta C_j^{\text{NP}})}{\text{TC}^{\text{ref}}} \\ &= \frac{\bar{g}}{g} \frac{\sum_{j \in \mathcal{J}} N_j (q\frac{g}{\bar{g}}\Delta C_j^{\text{P}} + \frac{g}{\bar{g}}(1-q)\Delta C_j^{\text{NP}})}{\text{TC}^{\text{ref}}} \\ &\leq \frac{\bar{g}}{g} \frac{\sum_{j \in \mathcal{J}} N_j (q\frac{g}{\bar{g}}\Delta C_j^{\text{P}} + (1-q\frac{g}{\bar{g}})\Delta C_j^{\text{NP}})}{\text{TC}^{\text{ref}}} \\ &= \frac{\bar{g}}{g} \% \Delta \text{TC}^{\text{HOV-MBP}}. \end{aligned}$$

$\square$

*Proof.* Proof of Lemma 1. Assumption 6 ensures that  $F_\theta$  is continuous and strictly increasing on  $[\theta^-, \theta^+]$ . On one hand,  $p \rightarrow \theta(p)$  is continuous and strictly increasing on  $[0, 1]$ , from  $\theta(0) = \theta^- \leq 0$  to  $\theta(1) = \theta^+ > 0$ . On the other, the function  $p \mapsto A_{\text{HOV-MBP}}(p) = C_{\text{HOV-MBP}}^{\text{NP}}(S^{\text{P}}, p) - C_{\text{HOV-MBP}}^{\text{P}}(S^{\text{P}}, p)$  is continuous and decreases on  $[0, \bar{p}]$  from  $A_{\text{HOV-MBP}}(0) = C(N/S) - C(0) \geq 0$  to  $A_{\text{HOV-MBP}}(\bar{p}) = 0$ , and then remains equal to 0 on  $[\bar{p}, 1]$ . Thus, the function  $\theta(p)$  and  $A_{\text{HOV-MBP}}(p)$  have a unique intersection  $p^* \in [0, 1]$ . If  $\bar{p} \leq F_\theta(0)$ , they intersect for  $p = F_\theta(0)$ . Else, they intersect for  $p \in (F_\theta(0), \bar{p})$ .  $\square$

*Proof.* Proof of Lemma 2. The same reasoning applies as for the proof of Lemma 1, after noticing that  $p \mapsto A_{\text{static}}(p) = C_{\text{static}}^{\text{NP}}(S^{\text{P}}, p) - C_{\text{static}}^{\text{P}}(S^{\text{P}}, p)$  is continuous and decreases on  $[0, 1]$ , from  $A_{\text{static}}(0) = C(N/S^{\text{P}}) - C(0) \geq 0$  to  $A_{\text{static}}(1) = 0$ .  $\square$

*Proof.* Proof of Proposition 7. Let  $S^{\text{P}} \in (0, S)$ . If  $\bar{p} \leq F_{\theta}(0)$ , then  $p_{\text{static}} = p_{\text{HOV-MBP}} = F_{\theta}(0)$ , and priority users have no advantage, so the desired results hold.

Let us now assume that  $\bar{p} > F_{\theta}(0)$ . As in Section 4.2.1, the static and dynamic schemes lead to the same costs for priority users, i.e.

$$C_{\text{HOV-MBP}}^{\text{P}}(S^{\text{P}}, p) = C_{\text{static}}^{\text{P}}(S^{\text{P}}, p), \forall p \in [0, \bar{p}], \quad (\text{C.1})$$

but the static scheme leads to a capacity under-utilization, and therefore  $C_{\text{HOV-MBP}}^{\text{NP}}(S^{\text{P}}, p) \leq C_{\text{static}}^{\text{NP}}(S^{\text{P}}, p)$ , for all  $p \in [0, \bar{p}]$ . This establishes that

$$A_{\text{HOV-MBP}}(p) \leq A_{\text{static}}(p), \forall p \in [0, \bar{p}]. \quad (\text{C.2})$$

This applies in particular to  $p_{\text{HOV-MBP}}$ , so  $A_{\text{static}}(p_{\text{HOV-MBP}}) \geq A_{\text{HOV-MBP}}(p_{\text{HOV-MBP}}) = \theta(p_{\text{HOV-MBP}})$ . As a consequence,  $A_{\text{static}}(p)$  intersects  $\theta(p)$  for  $p \geq p_{\text{HOV-MBP}}$ , which establishes Eq. (15).

Marshalling Eq. (15), Eq. (C.1) and the fact that  $C_{\text{HOV-MBP}}^{\text{P}}$  is non-decreasing with respect to  $p$  leads to Eq. (16). Then, since  $A_{\text{HOV-MBP}}$  is also non-decreasing with respect to  $p$ ,  $A_{\text{HOV-MBP}}(p_{\text{HOV-MBP}}) \leq A_{\text{HOV-MBP}}(p_{\text{static}}) \leq A_{\text{static}}(p_{\text{static}})$ . Combining this with Eq. (16) leads to Eq. (17).  $\square$

## D More complex carpooling models

Section 5.3 considers a very simple carpooling model, in which the carpooling inconvenience is considered an individual-specific constant. This appendix reviews some alternative assumptions proposed in the literature.

Many works distinguish carpool drivers from carpool passengers: drivers cover the vehicle-related costs and incur an extra detouring cost to pick-up and drop-off passengers, while passengers incur an access cost and/or a waiting time. As the inconvenience is larger for drivers, passengers are often assumed to compensate the drivers financially, such that the generalized costs for carpool drivers and passengers are equal at equilibrium (see e.g. Ma and Zhang 2017, Liu and Li 2017). The distinction between carpool drivers and carpool passengers is then only relevant in terms of cost components, not in terms of generalized cost.

One common assumption is that the value of time (VoT) for a given person may depend on whether this person drives alone, or carpools. Two radically different assumptions have been made in this regard. On one hand, Liu and Li (2017) assumes that carpooling generates an inconvenience that is the sum of a component proportional to the distance and of another component proportional to the time spent together with the other carpoolers, with different proportionality coefficients for carpool drivers and carpool passengers. As Liu and Li (2017) does not model detours and assumes a classic bottleneck model, this is equivalent to assuming that carpoolers have a constant carpooling inconvenience, plus a larger value of time. On the other hand, Ma and Zhang (2017) and Zhong et al. (2020) assume that carpool drivers keep the same VoT while carpool passengers have a lower one, because they do not need to drive. The fact that the VoT of carpool passengers is larger than that of solo drivers in Liu and Li (2017), but smaller in Ma and Zhang (2017) and Zhong et al. (2020) highlights the lack of consensus and the need for more empirical observations. A Stated-Preferences (SP) experiment recently conducted in France (Monchambert 2020) partially supports the assumption of Liu and Li (2017): drivers were found to have a VoT that is 13 % larger when they drive a carpool than when they drive alone.

Besides the influence of carpool on the VoT, various models have also been proposed for the costs associated with passenger pick-up and/or drop-off. Ma and Zhang (2017) and Zhong et al. (2020) consider a modeling framework somehow different from ours: they focus on expected carpooling costs and on the expected number of carpool passengers, which is not necessarily an integer. The cost of detours then varies with the proportion of the population that carpools, but this can be considered as an effect of the use of expected values: the average carpooling detour increases because drivers are more likely to pick-up carpool passengers, not because these passengers are further away

(the length of the detour per additional passenger is assumed to be constant). The only study that we are aware of that relaxes this assumption is Wang, Yang, and Zhu (2018). The authors consider a mode choice problem including carpooling as an alternative and allow the length of detours to vary with the number of carpool drivers and carpool passengers on the market. Specifically, they prove the existence of a mode choice equilibrium by only requiring that the distance between a pair of matched drivers and passengers satisfies two constraints: (i) it should remain between a lower bound and an upper bound and (ii) its partial derivatives with respect to the number of carpool drivers and the number of carpool passengers should be negative. In the example they consider, this distance is assumed to be inversely proportional to the number of matched carpooling pairs. It might be possible to test this assumption analytically and/or provide an alternative functional form. Yet, it seems that the research community has mostly focused so far on how to implement such sharing services (various matching algorithms have been proposed - see e.g. Agatz et al. (2012)), rather than on evaluating how their performance varies with the total demand.

## E Simulations

Except for a few specific cases (e.g. those from the sections 4.2 and 4.3), the DDUE cannot be characterized exactly with both a continuum of users and continuous time. The alternative presented here is to discretize time and to find the equilibrium by solving a finite-dimensional version of the problem [2D-LP( $\tilde{s}$ )] from Appendix A. The core algorithm is based on the approach proposed by Iryo and Yoshii (2007), but it has been adapted to allow for a time-dependent capacity. It is called by other programs to estimate the DDUE in various situations (without any priority scheme, with R-MBP, with HOV-MBP and with static HOV lanes).

### E.1 Core Algorithm: One Population, One Bottleneck

The core algorithm takes as inputs:

- a vector  $\mathbf{t}$  of length  $n_t$  of discrete times,
- a vector  $\mathbf{S}$  of length  $n_t$  with the capacity (in vehicles) of the bottleneck at each discrete time (this requires multiplying the bottleneck capacity (in veh/h) by the width of each time slot),
- a vector  $\tilde{\mathbf{N}} = (\tilde{N}_j)_{j \in \mathcal{J}}$  indicating the number of vehicles of each group,
- $J$  normalized schedule penalty functions  $\text{SP}_j/\alpha$ .

It then relies on a commercial solver (Matlab's dual-simplex) to find an optimal solution to the following linear program:

$$\begin{aligned} & \underset{x_{i,j} \geq 0}{\text{minimize}} && \sum_{j \in \mathcal{J}} \sum_{i=1}^{n_t} \frac{\text{SP}_j(t_i)}{\alpha} x_{i,j} \end{aligned} \tag{E.1a}$$

$$\text{subject to} \quad \sum_{j \in \mathcal{J}} x_{i,j} \leq S_i, \quad \forall i = 1, \dots, n_t, \tag{E.1b}$$

$$\sum_{i=1}^{n_t} x_{i,j} = \tilde{N}_j, \quad \forall j \in \mathcal{J}. \tag{E.1c}$$

The algorithm returns as outputs the matrix  $x_{i,j}$  of bottleneck departures, the Lagrange multipliers of the demand constraints (E.1c), which correspond to the normalized costs of all groups and if necessary, the Lagrange multipliers associated to the capacity constraints (E.1b), which correspond to the queueing times.

### E.2 Metering-Based Priority and Static HOV Lanes - Exogenous Mix

Given some partition of the population between priority and non-priority and some capacity  $S^P$ , the cases with R-MBP, HOV-MBP and static HOV lanes can be solved by running the core algorithm separately for priority and non-priority vehicles.

With static HOV lanes, the two sub-equilibria can be computed either sequentially or in parallel, because each population has access to pre-determined capacity. If the difference between any two consecutive values of  $\mathbf{t}$  is  $\delta t$ , the vector of capacities is defined for priority vehicles by  $S_i = S^P \delta t, \forall i = 1, \dots, n_t$  and for non-priority vehicles by  $S_i = (S - S^P) \delta t, \forall i = 1, \dots, n_t$ .

With R-MBP and HOV-MBP, the core algorithm must be run first for priority vehicles with  $S_i = S^P \delta t, \forall i = 1, \dots, n_t$ . Once the results are known, the core algorithm can be run for non-priority ones, with  $S_i = S \delta t - \sum_{j \in \mathcal{J}} x_{i,j}^P$ , where  $(x_{i,j}^P)_{i=1, \dots, n_t, j \in \mathcal{J}}$  are the priority arrivals obtained earlier.

### E.3 Metering-Based Priority and Static HOV Lanes - Endogenous Mix

When the partition of the population between priority and NP is endogenous, a variant of the Method of Successive Averages (MSA) is utilized (Algorithm 1). Every iteration starts with a

---

#### Algorithm 1 Endogenous mix (MSA)

---

```

P =  $F_\theta(0) \times \mathbf{1}_{1 \times \mathcal{J}}$  ▷ Initial proportion of carpoolers
for iteration = 1 ... maxIterations do
   $\tilde{\mathbf{N}}^P = \mathbf{N} \circ \mathbf{P} / g$  ▷  $(\mathbf{N} \circ \mathbf{P})_i = N_i P_i$  (Hadamard product)
   $\tilde{\mathbf{N}}^{NP} = \mathbf{N} \circ (\mathbf{1} - \mathbf{P})$ 
   $[\mathbf{C}^P, \mathbf{C}^{NP}] = \text{computeEquilibrium}(\mathbf{t}, (SP_j)_{j \in \mathcal{J}}, \tilde{\mathbf{N}}^P, \tilde{\mathbf{N}}^{NP}, S^P, S)$ 
  targetP =  $F_\theta(\mathbf{C}^P - \mathbf{C}^{NP})$ 
   $\mathbf{P} = w_{iteration} \mathbf{targetP} + (1 - w_{iteration}) \mathbf{P}$ 
end for

```

---

vector  $\mathbf{P}$  indicating the proportion of carpoolers in each group  $j \in \mathcal{J}$  (at the beginning of the first iteration, the only carpoolers are the natural ones). After translating these proportions into numbers of priority and non-priority vehicles, the departure-time equilibria are determined using the function `computeEquilibrium`, using one of the procedures described in Section E.2. The resulting departure-time equilibria are used to estimate the congestion advantage that each group enjoys when carpooling, so that a target vector **targetP** can be computed. To improve stability, the vector  $\mathbf{P}$  utilized in the next iteration is a weighted average of this target and of the previous value. After sufficiently many iterations, the procedure should converge towards a simultaneous mode and departure time equilibrium. Here, after some trial and error, we used 15 iterations and a weight function  $w_{iteration} = 1/\sqrt{iteration}$ .

## References

- Agatz N, Erera A, Savelsbergh M, Wang X, 2012 *Optimization for dynamic ride-sharing: A review*. *European Journal of Operational Research* 223(2):295–303.
- Akamatsu T, Wada K, Iryo T, Hayashi S, 2018 *Departure time choice equilibrium and optimal transport problems*. Working paper, URL <https://mpra.ub.uni-muenchen.de/90361/>, accessed 2019-04-16.
- Arnott R, de Palma A, Lindsey R, 1988 *Schedule delay and departure time decisions with heterogeneous commuters*. *Transportation Research Record* 1197:56–67.
- Arnott R, de Palma A, Lindsey R, 1993 *A structural model of peak-period congestion: A traffic bottleneck with elastic demand*. *The American Economic Review* 83(1):161–179.
- Chen C, Varaiya P, Kwon J, 2005 *An empirical assessment of traffic operations*. *Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, 105–124.
- Daganzo CF, 1985 *The uniqueness of a time-dependent equilibrium distribution of arrivals at a single bottleneck*. *Transportation science* 19(1):29–37.
- Daganzo CF, Garcia RC, 2000 *A Pareto improving strategy for the time-dependent morning commute problem*. *Transportation Science* 34(3):303–311.
- Daganzo CF, Lehe LJ, 2015 *Distance-dependent congestion pricing for downtown zones*. *Transportation Research Part B: Methodological* 75(0):89–99.
- de Palma A, Ben-Akiva M, Lefevre C, Litinas N, 1983 *Stochastic equilibrium model of peak period traffic congestion*. *Transportation Science* 17(4):430–453.
- de Palma A, Proost S, Seshadri R, Ben-Akiva M, 2018 *Congestion tolling - dollars versus tokens: A comparative analysis*. *Transportation Research Part B: Methodological* 108:261–280.

- Federal Highway Administration, 2017 National Household Travel Survey (NHTS). <http://nhts.ornl.gov> [Accessed: August 2019].
- Fosgerau M, 2011 *How a fast lane may replace a congestion toll*. *Transportation Research Part B: Methodological* 45(6):845–851.
- Galichon A, 2016 *Optimal Transport Methods in Economics* (Princeton University Press).
- Geroliminis N, Srivastava A, Michalopoulos P, 2011 *A dynamic-zone-based coordinated ramp-metering algorithm with queue constraints for minnesota’s freeways*. *IEEE Transactions on Intelligent Transportation Systems* 12(4):1576–1586.
- Grant-Muller S, Xu M, 2014 *The role of tradable credit schemes in road traffic congestion management*. *Transport Reviews* 34(2):128–149.
- Guler SI, Menendez M, 2014 *Analytical formulation and empirical evaluation of pre-signals for bus priority*. *Transportation Research Part B: Methodological* 64:41–53.
- Iryo T, Yoshii T, 2007 *Equivalent optimization problem for finding equilibrium in the bottleneck model with departure time choices*. *4th IMA International Conference on Mathematics in Transport*, 231–244.
- Knockaert J, Verhoef ET, Rouwendal J, 2016 *Bottleneck congestion: Differentiating the coarse charge*. *Transportation Research Part B: Methodological* 83:59–73.
- Konishi H, Mun S, 2010 *Carpooling and congestion pricing: Hov and hot lanes*. *Regional Science and Urban Economics* 40(4):173 – 186.
- Lago A, Daganzo CF, 2007 *Spillovers, merging traffic and the morning commute*. *Transportation Research Part B: Methodological* 41(6):670–683.
- Lamotte R, de Palma A, Geroliminis N, 2017 *On the use of reservation-based autonomous vehicles for demand management*. *Transportation Research Part B: Methodological* 99:205–227.
- Lindsey R, 2004 *Existence, uniqueness, and trip cost function properties of user equilibrium in the bottleneck model with multiple user classes*. *Transportation Science* 38(3):293–314.
- Liu Y, Li Y, 2017 *Pricing scheme design of ridesharing program in morning commute problem*. *Transportation Research Part C: Emerging Technologies* 79:156–177.
- Ma R, Zhang H, 2017 *The morning commute problem with ridesharing and dynamic parking charges*. *Transportation Research Part B: Methodological* 106:345–374.
- Monchambert G, 2020 *Why do (or don’t) people carpool for long distance trips? a discrete choice experiment in France*. *Transportation Research Part A: Policy and Practice* 132:911–931.
- Papamichail I, Kotsialos A, Margonis I, Papageorgiou M, 2010 *Coordinated ramp metering for freeway networks – a model-predictive hierarchical control approach*. *Transportation Research Part C: Emerging Technologies* 18(3):311–331, 11th IFAC Symposium: The Role of Control.
- Picard N, Dantan S, de Palma A, 2018 *Mobility decisions within couples*. *Theory and Decision* 84(2):149–180.
- Qian Z, Zhang M, 2011 *Modeling multi-modal morning commute in a one-to-one corridor network*. *Transportation Research Part C: Emerging Technologies* 19(2):254–269.
- Ramezani M, Haddad J, Geroliminis N, 2015 *Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control*. *Transportation Research Part B: Methodological* 74:1–19.
- Rogers CA, 1985 *Effects of ramp metering with hov bypass lanes on vehicle occupancy*. *Transportation Research Record* (1021):10–15.
- Shen W, Zhang H, 2010 *Pareto-improving ramp metering strategies for reducing congestion in the morning commute*. *Transportation Research Part A: Policy and Practice* 44(9):676–696.
- Small KA, 1983 *Bus priority and congestion pricing on urban expressways*. *Research in Transportation Economics* 1.
- Small KA, Winston C, Yan J, 2006 *Differentiated road pricing, express lanes, and carpools: Exploiting heterogeneous preferences in policy design*. *Brookings-Wharton Papers on Urban Affairs* 53–96.
- Smith MJ, 1984 *The existence of a time-dependent equilibrium distribution of arrivals at a single bottleneck*. *Transportation Science* 18(4):385–394.
- van den Berg V, Verhoef ET, 2011 *Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay*. *Journal of Public Economics* 95(7-8):983–992.
- Vickrey WS, 1969 *Congestion Theory and Transport Investment*. *American Economic Review* 59(2):251–60.
- Villani C, 2008 *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften* (Springer Science & Business Media).

- Wang X, Yang H, Zhu D, 2018 *Driver-rider cost-sharing strategies and equilibria in a ridesharing program*. *Transportation Science* 52(4):868–881.
- Wu J, Hounsell N, 1998 *Bus priority using pre-signals*. *Transportation Research Part A: Policy and Practice* 32(8):563–583.
- Xu Xy, Liu J, Li Hy, Hu JQ, 2014 *Analysis of subway station capacity with the use of queueing theory*. *Transportation Research Part C: Emerging Technologies* 38:28–43.
- Yu X, van den Berg VA, Verhoef ET, 2019 *Carpooling with heterogeneous users in the bottleneck model*. *Transportation Research Part B: Methodological* 127:178–200.
- Zhong L, Zhang K, (Marco) Nie Y, Xu J, 2020 *Dynamic carpool in morning commute: Role of high-occupancy-vehicle (HOV) and high-occupancy-toll (HOT) lanes*. *Transportation Research Part B: Methodological* 135:98–119.