

# Congestion and Incentives in the Age of Driverless Cars\*

Federico Boffa<sup>†</sup>      Alessandro Fedele<sup>‡</sup>      Alberto Iozzi<sup>§</sup>

May 16, 2019

## Abstract

GPS systems and Autonomous Vehicles (AVs) will likely open the way to forms of traffic coordination, or centralization. We analyze the welfare effects of moving from an environment with atomistic drivers to one in which few companies will manage the traffic. Differently than what happens with atomistic drivers, such companies or organizations will have an incentive to consider the congestion externality imposed by their vehicles on the other vehicles they dispatch. We analyze both a setting with no road taxes, to reflect their limited application and the popular opposition to them, as well as a setting with road taxes. We find that, without road taxes, the emergence of a small company supplying a small fraction of the travelers (while the others remain atomistic) increases (decreases) welfare if and only if the congestion problem was (was not) sufficiently severe in the first place. With road taxes, we find that, while congestion charges are optimal when all travelers are atomistic, the structure of the taxes differs markedly with a company that supplies a mass of customers. Restoring first best, in this case, may require subsidizing the company – something likely to be politically very unappealing.

**Keywords:** autonomous vehicles; market structure; congestion externality

**JEL Codes:** R41; R11

---

\*We are grateful to Giacomo Calzolari, Toomas Hinnosaar, Ignacio Monzon and Aleksey Tetenov for useful discussions and helpful comments. We thank seminar participants at Bolzano, Collegio Carlo Alberto, IMT Lucca, Porto, Siena, SOAS, Roma ‘Tor Vergata’ and session participants at AISRE, EARIE, ERSa, IAERE, SIE and SIET Conferences. All errors are ours.

<sup>†</sup>Free University of Bolzano and Collegio Carlo Alberto: [federico.boffa@unibz.it](mailto:federico.boffa@unibz.it).

<sup>‡</sup>Free University of Bolzano: [alessandro.fedele@unibz.it](mailto:alessandro.fedele@unibz.it).

<sup>§</sup>Università di Roma ‘Tor Vergata’ and SOAS University of London: [alberto.iozzi@uniroma2.it](mailto:alberto.iozzi@uniroma2.it).

# 1 Introduction

Technological advancement is rapidly changing the mobility industry. One of the most prominent innovation is the development of Autonomous Vehicles (AVs), that is, vehicles driven by a software that does not require human intervention. AVs open a host of relevant technological, legal and moral issues (Awad et al. 2018). But significant, and as of now underappreciated, impacts will come from the changes in the organization of the mobility market that they will entail, in particular by fostering a higher level of centralization. Centralization of transport decisions is crucial in the management of congestion, which has a negative effect of the mobility industry. The yearly congestion cost has been estimated to amount to more than one hundred billion dollars in the US, and to be steadily increasing over time (Schrang, Lomax and Eisele, 2011).

Congestion in the mobility industry not only derives from transport infrastructures being inadequate relative to demand, but is also the result of a standard externality. Indeed, in the current decentralized setting, drivers are atomistic, and they do not factor in their decisions the external effect in terms of congestion they impose on fellow travelers. Incentives will dramatically change with the diffusion of AVs. The increased benefits from car sharing will likely reduce consumers' investment into private cars (Fagnant and Kockelman, 2015), thus leading to the emergence of companies that will provide travel services. These companies will have an incentive to consider, in a fleet logic, the congestion cost imposed by their AVs on the other AVs they dispatch, thereby internalizing (at least in part) the externality. At the same, their presence will generate other distortions. This paper analyzes the welfare effects of such transition from the current travel market structure based on atomistic decisions to centralized traffic. We address the following two questions: how will the diffusion of AVs affect social welfare, in the different stages of the transition process? What are the optimal tax schemes to correct the distortions in each of them?

We consider a framework with travelers using AVs to travel over a road network, segmented into two separate parallel lanes. Both lanes are congested. In our framework, agents are assumed to be heterogeneous as to the utility they derive from the trip and to the disutility they derive from the congestion. In particular, we assume that the larger is the utility from the trip, the larger is the cost of congestion. This is consistent with evidence pointing to a positive relation between wage and value of time (see, for instance, Small, 2012). We first show that, in the welfare-maximizing first best benchmark, the two lanes have different levels of congestion, reflecting the heterogeneity in travelers' value of time. Furthermore, a commuter in the first best travels as long as her benefit from travelling exceeds the increase in aggregate congestion costs she imposes on fellow commuters. Finally, Thus, if the congestion cost is sufficiently large, efficiency requires to prevent some

low-value commuters from traveling.

We then study how the transition from a fully decentralized economy populated only by atomistic travelers to a centralized economy with one (or more) companies managing traffic affects welfare. We analyze the two polar cases of atomistic travelers only and of an individual company supplying the entire market, and we also consider the transition period, where some atomistic travelers coexist with others that are dispatched by a company. We run this analysis in two contexts, with and without road taxes.

Consider first the case of no road taxes. While economists advocate road pricing and congestion taxes, they are rarely implemented in practice, likely for political economy reasons (Oberholzer-Gee and Weck-Hannemann, 2002). In this case, we find that, at the beginning of the transition period, the emergence of a supplier that manages the vehicles with a fleet perspective, by internalizing the congestion externality, may reduce the aggregate amount of vehicles on the road. Whether this is desirable or not from the welfare standpoint depends on the severity of the congestion in the first place. If congestion is sufficiently severe that the social planner would efficiently exclude some low-value agents from traveling, the quantity reduction operated by the company may be efficient. If, to the contrary, congestion is not so severe in the first place, and the planner would dispatch all the travelers, the monopolist's screening is welfare-reducing. When the transition is complete, and all travelers are managed by the a monopolistic supplier, the congestion externality is fully internalized. However, two other distortions typical of monopoly emerge. With respect to the social optimum, there is overdifferentiation in the congestion levels across the two lanes, as the monopolist uses quality differentiation to extract profit through market segmentation. Also, the amount of vehicles dispatched differs from social optimum. Interestingly, output distortion can go in both directions, including an increase in the number of vehicles dispatched. This finding, at odds with the standard result that a monopolist sets total output below the socially optimal level, reflects the intuition that a monopolist underprovides quality relatively to the social optimum (Spence, 1979).

Our results for the monopoly situation bridge two streams of literature. On the one hand, they parallel those obtained in the analysis of quality levels with monopolistic carriers (see, for instance, the theory model by Basso, 2008 and the empirical counterparts estimating the relation between airport concentration and quality, see Mayer and Sinai, 2003, Rupp, 2009, Daniel and Harback, 2008, and Molnar, 2013). However, our analysis differs as it includes the dimension of multiplicity of lanes, crucial for the analysis of the welfare effects of AVs. On the other hand, Mussa and Rosen (1978) show that a monopolist providing multiple vertically differentiated goods has an incentive to underprovide quality of the low quality good with respect to social optimum. Our result that the centralized

monopolistic company has an incentive to overdifferentiate quality with respect to the social planner is reminiscent of this.

Second, we analyze the transition when a tax authority can impose taxes that restore social optimality. In the decentralized situation where all travelers are atomistic, a traditional congestion charge, i.e., a Pigouvian tax, equal to the marginal external cost imposed on the other vehicles, restores optimality. This mirrors the finding obtained in the bottleneck model (see Vickrey, 1969, and Arnott, de Palma and Lindsey, 1990)). Conversely, in the centralized situation in which all travelers are managed by the same company, the tax that restores social optimum differs markedly. Indeed, since the company internalizes the externality already, there is no scope for a congestion charge. This result aligns with those obtained in the literature on airports when carries have market power (Daniel, 1995, Brueckner, 2002, Pels and Verhoef, 2004, Brueckner 2005, Basso and Zhang, 2007, Silva and Verhoef, 2013). To the contrary, a tax to align the monopolist's incentives toward the welfare-maximizing quality choice restores the socially optimal allocation across lanes. In addition, a subsidy on all the cars dispatched by the monopolist restores the optimal number of vehicles. In particular, when the congestion problem is particularly severe, we show that the subsidy exceeds the tax, so that the monopoly receives a net subsidy. This may prove politically challenging, even more than traditional congestion pricing schemes, and may require some countermeasures by the tax authority that improve its political feasibility. An example of them could be the collection of license fees from the company, so as to balance the tax budget within the AV market. Finally, we show that when, during the transition, a group of atomistic travelers coexists with a group of travelers managed by a monopolistic supplier, the taxing schemes markedly differ across the two groups, and the scheme imposed to each of them closely resembles that charged to that group in isolation.

To the best of our knowledge, there are only two papers that relate market structure and congestion externality with reference to AVs. Lamotte, De Palma and Geroliminis (2016) investigate the commuters' choice between conventional and autonomous vehicles, while van den Berg and Verhoef (2016) focus on the impact of AVs on road capacity, studying the deployment of infrastructures resulting from the transition to the AVs framework.<sup>1</sup> Finally, our paper is close to Ostrovsky and Schwarz (2018), who investigate the interplay between autonomous transportation, carpooling, and road pricing to achieve socially efficient outcomes. Our analysis is complementary in that we disregard efficient carpooling with AVs, which is considered to be key to reduce congestion<sup>2</sup>, but study how congestion

---

<sup>1</sup>Two other papers analyze equilibria when drivers are non-atomistic (Silva et al. 2016, Lindsey, De Palma and Silva, 2019).

<sup>2</sup>For instance, Jerry Walters, a principal consultant at Fehr & Peers, a US transportation

can be mitigated under different market structures and different taxation systems vis-a-vis the socially efficient outcome.

The rest of the paper is organized as follows. Section 2 sketches the model. Section 3 illustrates the first best. Section 4 characterizes the equilibrium without taxes. Section 5 characterizes the equilibrium with road pricing. Section 6 concludes. The Appendix contains proofs of the results.

## 2 The model

**Travelers' utility and road network.** There is a unit mass of potential travelers (also referred to as agents), each with unit demand for a single trip. Trips occur along a road network composed of a single road connecting a single origin-destination pair. The road is divided into two ex-ante identical lanes. The lanes may, however, differ ex-post because of the possibly different mass of travelers in each of them. We denote by  $n$  the total mass of travelers in one lane and by  $N$  the total mass of travelers in the other lane, with  $0 \leq N \leq n$ . We assume that lanes are always small relative to demand, so that, at any  $n$  and  $N$ , lanes are congested and the utility from traveling is negatively affected by the mass of travelers in the same lane (see below). Given that travelers dislike congestion, we define the lane with less travelers as the *luxury* lane, and that with more travelers as the *popular* lane.

Travelers are heterogeneous with respect to an individual type  $\theta$ , which we assume to be uniformly distributed in the  $[0, 1]$  interval. A type- $\theta$  agent traveling in a lane with  $n \in \{n, N\}$  travelers derives the following utility from the trip:

$$\mathcal{B}(\theta, n) = B(\theta) - G(n, \theta). \quad (1)$$

The first term in (1) is the gross benefit from traveling,  $B(\theta)$ , which depends only on the individual type  $\theta$ . We let  $B(\theta)$  be increasing and weakly concave in  $\theta$ , so that  $B'(\theta) > 0$  and  $B''(\theta) \leq 0$ . The second term in (1) denotes the disutility from congestion. We assume  $G(\cdot)$  to be increasing in the mass of those who travel in the same lane, so that  $\frac{\partial G(n, \theta)}{\partial n} > 0$ . We also assume that  $\frac{\partial G(n, \theta)}{\partial \theta} > 0$ . This implies that the disutility of congestion is larger the larger the type  $\theta$ . This is consistent with evidence that points at the positive relation between wage and value of time (see, for instance, Small, 2012). We also assume that the marginal disutility of congestion weakly increases with the type  $\theta$ , so that  $\frac{\partial^2 G(n, \theta)}{\partial n \partial \theta} \geq 0$ . Accordingly,  $\theta$

---

consultancy, argues. "The key distinction is the number of people per vehicle. Without pretty radically increasing the number of people per vehicle, autonomous systems will increase total miles traveled." (source: <https://www.nytimes.com/2018/10/27/technology/driverless-cars-congestion.html>).

represents both a parameter of horizontal differentiation in agents' value of travel, and a parameter of disutility from congestion.

We make the following assumption on the travelers utility function:

**Assumption 1.** *i)  $B(0, n) = B(0) \geq 0$ ;*

*ii) for any  $\theta \in [0, 1]$ ,  $\frac{\partial B(\theta, n)}{\partial \theta} > 0$  for any  $n$ .*

Assumption 1.i) states that a type-0 agent's utility from traveling is non-negative. Assumption 1.ii) instead makes sure that the net utility from traveling is increasing in the type  $\theta$ , which, together with Assumption 1.i), implies that all travelers have a non-negative utility. The gross utility from traveling increases with  $\theta$  more rapidly than the increase in the disutility from congestion. This makes sure that the type  $\theta$  is sufficient for ordering travellers with respect to their net utility for any possible allocation of agents across the two lanes.

In the remainder of the analysis, we assume that the disutility of congestion increases linearly both in  $n$  and  $\theta$ , that is  $G(n, \theta) = \theta gn$ , with  $g > 0$ . Also, we sometimes use a linear specification for the gross benefit function  $B(\theta)$ :

**Assumption 2.**  $B(\theta) = b_0 + b\theta$ , with  $b_0 \geq 0$  and  $b > 0$ .

In this case, Assumption 1.ii) simply becomes  $b > g$ .

**Travelers' choice.** A traveler located at  $\theta$  chooses whether or not to travel, and, if she travels, in which lane she does it. The decision depends on the net benefit from traveling  $B(\theta)$ , on the potential fares or taxes she might pay when traveling in the popular or the luxury lane, denoted as  $\psi \geq 0$  and  $\Psi \geq 0$  respectively, and on the mass of travelers she anticipates to travel in the popular in the luxury lanes,  $n$  and  $N$  respectively. For any  $n$  and  $N$ , a traveler travels in the popular lane if and only if

$$B(\theta) - \theta gn - \psi \geq 0, \quad (2)$$

$$B(\theta) - \theta gn - \psi \geq B(\theta) - \theta gN - \Psi. \quad (3)$$

Similarly, she travels in the luxury lane if and only if

$$B(\theta) - \theta gN - \Psi \geq 0, \quad (4)$$

$$B(\theta) - \theta gN - \Psi \geq B(\theta) - \theta gn - \psi. \quad (5)$$

### 3 First best

In this Section, we consider an utilitarian welfare maximizing social planner, who is perfectly informed and can directly allocate travelers across the two lanes. We

characterize the first best allocation of travelers. In solving this problem, the planner chooses who should not travel (if any) and allocate the remaining mass of travelers to the popular lane (in a mass equal to  $n$ ) and to the luxury lane (in a mass equal to  $N$ ). Formally, the planner's problem is as follows

$$\begin{aligned} \max_{n,N} W &= \int_{1-n-N}^{1-N} [B(\theta) - \theta g n] d\theta + \int_{1-N}^1 [B(\theta) - \theta g N] d\theta \\ \text{s.t. } &0 \leq N \leq n \leq 1 \text{ and } n + N \leq 1. \end{aligned} \quad (6)$$

The utilitarian welfare  $W$  is given by the sum of two terms. The first term is the aggregate utility of the  $n$  travelers in the popular lane. The second term is that of  $N$  travelers in the luxury lane.

We denote by  $n_P$  and  $N_P$  the solutions to (6). When this problem has interior solutions, these are implicitly defined by the following system of FOCs:

$$B(1 - n_P - N_P) - 2gn_P \left(1 - N_P - \frac{3}{4}n_P\right) = 0; \quad (7)$$

$$B(1 - n_P - N_P) + g \left(n_P^2 + \frac{3}{2}N_P^2 - 2N_P\right) = 0. \quad (8)$$

From (7) and (8) we can derive the optimal choice of  $N_P$  as a function of  $n_P$ .<sup>3</sup> This relationship is given by

$$N_P(n_P) = \frac{1}{3} \left(2(1 + n_P) - \sqrt{7n_P^2 - 4n_P + 4}\right) \quad (9)$$

and is illustrated in Figure 1 by the green increasing line. This graph displays a crucial feature of the welfare-maximizing allocation of travelers. The green line always lies below the 45° dotted line, meaning that the social planner finds it optimal to differentiate between the two lanes. The popular lane contains strictly more travelers than the luxury lane,  $n_P > N_P$ . Travelers with a type  $\theta \in [1 - n_P - N_P, 1 - N_P]$ , less concerned about congestion, are allocated in the popular lane. Travelers with a type  $\theta \in [1 - N_P, 1]$ , more affected by congestion, are allocated in the luxury lane.

By implicit differentiation of the FOCs in (7) and (8), we find that  $\frac{\partial N_P}{\partial g} < 0$  and  $\frac{\partial n_P}{\partial g} < 0$ . This reflects the intuition that a larger disutility from congestion  $g$  yields a smaller social reward from a larger coverage of the market. To provide further comparative statics analysis, we rely on the linear gross benefit function in Assumption 2 and establish that  $\frac{\partial n_P}{\partial b}$ ,  $\frac{\partial N_P}{\partial b}$ ,  $\frac{\partial n_P}{\partial b_0}$  and  $\frac{\partial N_P}{\partial b_0}$  are all positive. A

---

<sup>3</sup>Explicit solutions to this system of equations, using Assumption 2, can be obtained but are very cumbersome and hard to interpret.

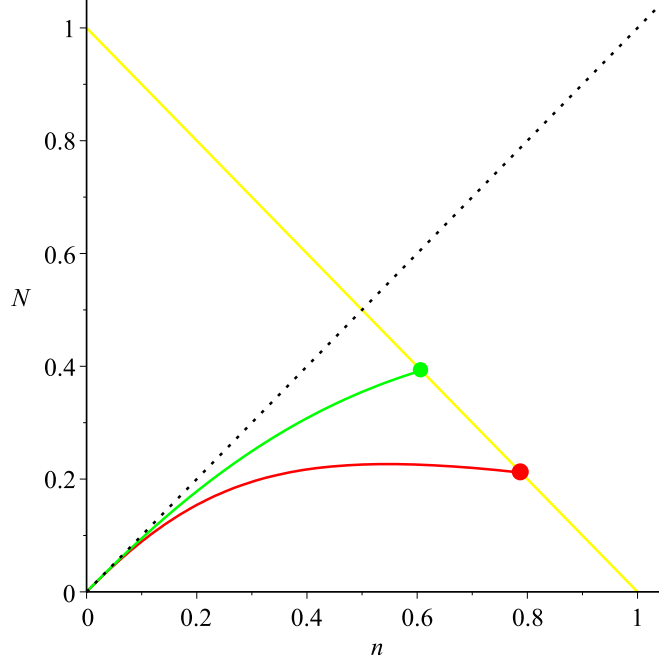


Figure 1: The implicit relationships  $N_P(n_P)$  (green line) and  $N_M(n_M)$  (red line)

larger benefit from traveling, both in its fixed component  $b_0$  and in the slope  $b$ , is associated to larger social reward in dispatching more cars.

When the planner finds it optimal to have all types traveling, the market is fully covered. This situation is illustrated in Figure 1 by the green solid circle located at the intersection between the green line (9) and the constraint  $n + N = 1$ . The next Proposition shows conditions under which the social planner finds it optimal to dispatch all travelers, and characterizes the first-best allocation in this circumstance.

**Proposition 1.** *Let  $\bar{n}_P$  and  $\bar{N}_P$  denote the solutions to problem (6) when there is full coverage, i.e., when  $\bar{n}_P + \bar{N}_P = 1$ . Then,*

$$\begin{aligned}\bar{n}_P &= \frac{1 + \sqrt{7}}{6} \cong 0.6076; \\ \bar{N}_P &= \frac{5 - \sqrt{7}}{6} \cong 0.3924.\end{aligned}\tag{10}$$



*These solutions occur if and only if*

$$g \leq \frac{2B(0)}{(1 - \bar{N}_P)^2} = \frac{72B(0)}{(1 + \sqrt{7})^2} \cong 5.4179 \times B(0). \quad (11)$$

Condition (11) illustrates that full coverage occurs when the congestion cost  $g$  is not that severe, and the benefit from traveling of the agent with the lowest  $\theta$  is sufficiently large relative to it. More precisely, full coverage arises when the benefit from traveling on the (popular) lane enjoyed by the agent with the lowest valuation,  $B(0)$ , is at least as high as the increase in aggregate congestion cost this traveler imposes on all agents in the popular lane. Notice that the implicit derivatives of  $n_P$  and  $N_P$  with respect to  $g$  and  $b_0 = B(0)$  show that the planner's choices smoothly converge toward the full coverage solution as the relation between  $b_0$  and  $g$  approaches that defined in condition (11).

## 4 Equilibrium analysis

We study under what conditions and to what extent deviations from optimality arise when travelers are atomistic and when they are served by a monopolistic supplier. We also study the case in which both types of agent coexist in the market.

We analyze a sequence of market structures that illustrate the expected stages in the transition towards a fully centralized markets with AVs. We first look at a market with atomistic travelers only, in line with the current predominant form of traffic organization in the real world. In this environment, agents do not take into account the impact of their traveling choices on others. We show that this results in overcongestion and in distortions in the travelers' allocation across lanes.

We then turn to coexistence of atomistic travelers with travelers managed by a monopolistic firm operating a fleet of AVs, to reflect the initial stages of diffusion of AVs. A monopolist cares about the effects of each vehicle on the utility of other travelers it dispatches. As a result, the congestion externality is, at least in part, internalized. We show that the welfare effects of this are, however, not clear-cut, and we characterize them in the different circumstances.

Finally, we look at the situation in which all travel services are provided by the monopolist, with no atomistic travelers on the road any longer. A monopolist fully internalizes the congestion externality, but other distortions, related to market power and to distorted incentives in terms of quality provision, emerge. We analyze the welfare implications of them.

This analysis is performed assuming no road taxes. Road pricing schemes, while often advocated by economists as effective tools to restore efficient outcomes, are

rarely implemented in practice, possibly because of their limited popular support. We will allow for road taxes in the next Section dedicated to them.

## 4.1 Atomistic travelers

In this Subsection, we study a fully decentralized economy populated by atomistic travelers only. Agents own their own car, and they pay no fee to travel. There is no organized market for transportation services.

When travelers are atomistic, it is actually immaterial whether they use an AV or a conventional vehicle. In both cases, agents make their individual decisions on if and how to travel only based on the utility they derive from the trip, net of the congestion cost, disregarding the costs they impose on other travelers. The vast use of navigation systems with significant computational ability, such as Google Driving Directions or Waze, currently provide travelers with increasingly accurate information on the traffic situation. This allows agents to optimize their decisions, given other travelers' choices. In this information context, it is meaningful to solve for the perfect-information pure strategy Nash equilibrium, which is characterized in the following Proposition:

**Proposition 2.** *Let  $n_A$  and  $N_A$  be the equilibrium mass of travelers in the two lanes when all travelers are atomistic. Then,*

$$n_A = N_A = \frac{1}{2}. \quad (12)$$

The Proposition illustrates that all atomistic travelers always choose to travel and they split equally across the two lanes, which feature the same level of congestion. Assumption 1.i) ensures that everyone travels. To get the intuition for the equal level of congestion across lanes, suppose by contradiction that the two lanes feature a different mass of travelers, so that, for instance,  $n_A > N_A$ . Then, any traveler, irrespective of her type  $\theta$ , would prefer to travel in the less congested lane because she enjoys higher utility.

Relative to the social planner case, atomistic travelers cause two distortions. First, there is too little differentiation across lanes. Travelers with the higher disutility of congestion are forced to travel in a lane whose level of congestion is too high relative to the first best. Second, when the congestion cost is high, there is excess travel. The first best prescribes that, under the conditions of high congestion cost (i.e., when  $g$  exceeds the right hand side of (11)), low- $\theta$  types do not travel. To the contrary, all atomistic travelers always travel, as they do not internalize the congestion cost imposed on fellow travelers.

## 4.2 Coexistence of atomistic travelers and a monopolistic supplier

In this Section, we look at the initial stage of diffusion of AVs. Two groups of travelers are coexisting. The proportion and the identity of agents in each group is exogenously given.<sup>4</sup> A proportion  $\gamma$  of travelers is atomistic. They own their car, and, similarly to the case of atomistic travelers in isolation analyzed above, they pay no fee to travel. A proportion  $1 - \gamma$  of travelers are non atomistic. They do not own a car and use the travel services of a firm operating as a monopolist. We refer to them as *corporate*. Corporate travelers have to pay a lane-specific fee to the monopolist in exchange for the service. The monopolist sets uniform fares within lanes: a fare  $f$  for the popular lane and a fare  $F$  for the luxury lane. Fares are used by the monopolist as the only instrument to direct travelers to the two lanes.

We assume that the distribution of corporate consumers is independent of the type  $\theta$ . At any interval  $[\theta, \theta + \epsilon]$ , with  $\theta \in [0, 1]$  and for any  $\epsilon > 0$  sufficiently small, there is a fraction  $\gamma \in (0, 1)$  of atomistic travelers and a fraction  $1 - \gamma$  of corporate travelers.<sup>5</sup>

We look at a sequential game, in which first the monopolist sets the fares  $f$  and  $F$ , and then, simultaneously, all agents, corporate and atomistic, make their travel decisions. All players are perfectly informed, and we solve for the subgame perfect Nash equilibrium of the game.

Let us now denote with  $m$  and  $M$  the mass of corporate travelers in the two lanes and, as before, with  $n$  and  $N$  the mass of total (corporate plus atomistic) agents traveling in the popular and luxury lane, respectively. The following Lemma illustrates the equilibrium allocation of travelers across the two lanes resulting from the last stage of the game.

**Lemma 1.** *Assume coexistence of atomistic and corporate agents. Then, for any pair of fares  $f$  and  $F$ , and for any  $m$  and  $M$  satisfying  $0 \leq M \leq m \leq 1 - \gamma$  and  $m + M + \gamma \leq 1$ , the mass of travelers in the two lanes is as follows:*

$$\left. \begin{array}{l} n = m \\ N = M + \gamma \end{array} \right\} \text{ if } \gamma < m - M \quad (13)$$

and

$$n = N = \frac{m + M + \gamma}{2} \text{ if } \gamma \geq m - M. \quad (14)$$

---

<sup>4</sup>Agents' allocation in the two groups depends on long-run decisions, involving, for instance, the choice to own a car or not, which are unmodeled here.

<sup>5</sup>Our results would change if we assumed that the distribution of corporate travelers is not independent of type. For an illustration of the choice between conventional and autonomous vehicles, see Lamotte, De Palma and Geroliminis (2016).

This Lemma illustrates that only two situations are possible. When  $n > N$ , all atomistic travelers travel in the luxury lane. This case, illustrated in (13), happens when the mass of atomistic travelers is relatively small (i.e.  $\gamma < m - M$ ). All atomistic travelers, given the choices made by the corporate travelers, choose to travel in the luxury lane. Because of their small mass, even after all atomistic agents travel in the luxury lane, the congestion level in the luxury lane is still lower than that in the popular lane. When instead  $n = N$ , atomistic travelers split themselves across the two lanes to equalize the total mass of travelers in each. This case, as given in (14), refers to a situation in which the mass of atomistic travelers is relatively large (i.e.  $\gamma \geq m - M$ ). Given corporate travelers' choices, if all atomistic travelers chose the luxury lane, its congestion level would exceed that of the popular lane. This clearly cannot be an equilibrium since any atomistic traveler in the luxury lane would benefit from switching to the popular lane, now less congested. Hence, travelers split between the two lanes so as to equalize the level of congestion across them.

We now analyze in details the case of a large fraction of travelers being atomistic (a sufficient condition for this is  $\gamma > \frac{1}{2}$ ), in which the mass of total (atomistic and corporate) travelers is identical in both lanes, as in (14).<sup>6</sup>

When the mass of atomistic travelers is large enough, equal congestions levels across lanes require equal fares in the two lanes, so that  $f = F$ . Hence, for all corporate travelers, the two incentive compatibility constraints in (3) and (5) are always trivially satisfied, and the individual rationality constraints (2) and (4) become identical and equal to

$$B(\theta) - \theta g \frac{m + M + \gamma}{2} - f \geq 0. \quad (15)$$

Since the monopolist charges the same fare to all corporate travelers, its profit is affected by their total mass only, and not by the way they split across lanes. Hence, denoting this total mass by  $\mathcal{M} \equiv m + M$ , and exploiting the fact that  $\mathcal{B}(\cdot)$  increases with  $\theta$ , the monopolist problem can then be written as

$$\begin{aligned} \max_f \quad & f\mathcal{M} \\ \text{s.t.} \quad & B(\theta_{MA}) - \theta_{MA}g \frac{\mathcal{M} + \gamma}{2} - f \geq 0 \\ & \mathcal{M} \leq 1 - \gamma \end{aligned} \quad (16)$$

where  $\theta_{MA}$  is the type of the marginal corporate traveler (where the subscript  $MA$  is a mnemonic for Monopoly & Atomistic). At the optimum, the IR constraint

---

<sup>6</sup>In this case, indeed, the mass of corporate travelers is less than  $\frac{1}{2}$  and so is the upper bound on  $m$ . This means that inequality  $\gamma \geq m - M$  is always strictly satisfied.

will always hold as equality for the marginal corporate traveler. Since corporate travelers are only a fraction  $1 - \gamma$  of the market and they are uniformly distributed along the unit line, the type  $\theta_{MA}$  of the marginal travelers is equal to  $\theta_{MA} = 1 - \frac{\mathcal{M}}{1-\gamma}$ . This allows us to rewrite the monopolist problem as

$$\begin{aligned} \max_{\mathcal{M}} \quad & \left( B \left( 1 - \frac{\mathcal{M}}{1-\gamma} \right) - \left( 1 - \frac{\mathcal{M}}{1-\gamma} \right) g \frac{\mathcal{M} + \gamma}{2} \right) \mathcal{M} \\ \text{s.t.} \quad & \mathcal{M} \leq 1 - \gamma \end{aligned} \quad (17)$$

We denote by  $\mathcal{M}_{MA}$  the solution to (17). When it is interior, this solution is implicitly defined by the following FOC:

$$\begin{aligned} \frac{1}{2(1-\gamma)} \times & \left( 3g\mathcal{M}_{MA}^2 - 2g\mathcal{M}_{MA}(1-2\gamma) - 2B' \left( 1 - \frac{\mathcal{M}_{MA}}{1-\gamma} \right) \mathcal{M}_{MA} + \right. \\ & \left. \left[ 2B \left( 1 - \frac{\mathcal{M}_{MA}}{1-\gamma} \right) - g\gamma \right] (1-\gamma) \right) = 0. \end{aligned} \quad (18)$$

In the special case of a linear benefit function, as in Assumption 2, we solve for the optimal total mass of corporate travelers:

$$\mathcal{M}_{MA} = \frac{g(1-2\gamma) + 2b - \sqrt{g^2\gamma^2 + (g^2 - 6gb_0)(1-\gamma) - 2gb\gamma - 2gb + 4b^2}}{3g} \quad (19)$$

Next, we establish conditions under which the market is fully covered, that is  $\mathcal{M}_{MA} + \gamma = 1$ .

**Proposition 3.** *When corporate travelers coexist with a large mass of atomistic travelers, full coverage (e.g.  $\mathcal{M}_{MA} + \gamma = 1$ ) occurs whenever*

$$g \geq 2 \times (B'(0) - B(0)). \quad (20)$$

Condition (20) shows a stark contrast with the full coverage conditions in first best in (11). The monopolist dispatches all vehicles when the congestion cost  $g$  is sufficiently high. To get the intuition for this possibly surprising result, consider that, as  $g$  gets larger, the aggregate demand curve turns more rigid, as the higher willingness to pay of higher  $\theta$  types is increasingly compensated by their higher congestion cost.<sup>7</sup> As a result, the monopolist increases market coverage. The other determinant of the demand curve faced by the monopolist is, perhaps more intuitively,  $B'$ . The monopolist's full coverage condition depends then also on  $B'(0)$ , with the same logic as above. This is also different than the planner, who considers, instead,  $B(0)$  only.

---

<sup>7</sup>This happens until  $g$  hits its upper bound  $B'(\theta)$ , determined by our Assumption 1.

**Welfare analysis.** The emergence of a supplier that manages the vehicles with a fleet perspective may have an impact on congestion, and, as a result, on welfare. This is because, by internalizing the congestion externality, it reduces the level of congestion. We show in this Section that this may be reflected into an improvement in welfare only in environments in which congestion is a sufficiently severe problem, so that the social planner would not cover the entire market, that is, it would exclude low- $\theta$  types. Under such circumstances, a monopolist finds it optimal to restrict supply, in line with the planner's choice, as long as the demand function is sufficiently steep. If, to the contrary, congestion is not such a severe problem, and the planner would cover the entire market, the monopolist may still find it optimal to restrict supply, but this would be welfare-reducing. Hence, while the intuition that the emergence of a share of traffic centralization (weakly) reduces congestion is valid, the welfare effects of this reduction are ambiguous.

To study the welfare effects of the initial steps of the process of traffic centralization through a monopoly, we first notice that, in the case of full coverage, social welfare is identical in the cases of atomistic travelers only and of coexistence between atomistic and corporate travelers. Everybody travels and the mass of travelers in the two lanes is identical (fares are just a transfer and do not play any role in the aggregate welfare evaluation). A welfare effect due to the emergence of a monopoly then comes up only when the monopoly excludes some low  $\theta$  types. To see this effect, we respectively denote as  $W_A$  and  $W_{MA}$  the expressions for total welfare in the cases of atomistic travelers only and of coexistence of atomistic and corporate travelers, and write

$$\begin{aligned}
W_A &= \int_0^1 \left( B(\theta) - \theta g \frac{1}{2} \right) d\theta; \\
W_{MA} &= (1 - \gamma) \int_{1-\frac{\mathcal{M}}{1-\gamma}}^1 \left( B(\theta) - \theta g \frac{\mathcal{M} + \gamma}{2} \right) d\theta + \\
&\quad + \gamma \int_0^1 \left( B(\theta) - \theta g \frac{\mathcal{M} + \gamma}{2} \right) d\theta.
\end{aligned} \tag{21}$$

In the case of atomistic agents only, all agents travel and suffer from the same level of congestion. Instead, in the case of an emerging monopoly, social welfare is the sum of welfare of corporate travelers (the first line in (21)) and of atomistic travelers (the second line in (21)). In both cases, the level of congestion is identical across all travelers, since each lane features a mass of  $\frac{\mathcal{M} + \gamma}{2}$  travelers. However, while the entire mass  $\gamma$  of atomistic travelers travels, only a fraction of corporate travelers do.

Take now the difference between these two expressions  $W_{MA} - W_A$  and denote it by  $\Delta W$ . Then:

$$\begin{aligned}
\Delta W = (1 - \gamma) \times & \quad (22) \\
& \left\{ \int_{1-\frac{\mathcal{M}}{1-\gamma}}^1 \left( B(\theta) - \theta g \frac{\mathcal{M} + \gamma}{2} \right) d\theta - \int_{1-\frac{\mathcal{M}}{1-\gamma}}^1 \left( B(\theta) - \theta g \frac{1}{2} \right) d\theta + \right. \\
& \left. - \int_0^{1-\frac{\mathcal{M}}{1-\gamma}} \left( B(\theta) - \theta g \frac{1}{2} \right) d\theta \right\} + \\
& + \gamma \left\{ \int_0^1 \left( B(\theta) - \theta g \frac{\mathcal{M} + \gamma}{2} \right) d\theta - \int_0^1 \left( B(\theta) - \theta g \frac{1}{2} \right) d\theta \right\}.
\end{aligned}$$

This expression clearly illustrates the pros and cons of the emergence of a monopoly that excludes some agents from traveling. The first 3 lines express the effects on corporate travelers. The second line illustrates the welfare gain for those still traveling, who are now facing a lower level of congestion, while the third line illustrates the welfare loss for those who are excluded from the market by the monopolist. Finally, the last line illustrates the positive effect on the welfare of atomistic travelers, who all still travel when a monopolist emerges, but who now face a lower level of congestion. The following Proposition illustrates, in the case of a linear benefit function as in Assumption 2, the conditions under which the emergence of a small monopolist fleet improves social welfare.

**Proposition 4.** *Assume Assumption 2 holds. Let  $\bar{\mathcal{M}} \equiv \frac{2b+g(1-2\gamma)-\sqrt{4b(b-g)+g^2-4g(4b_0-g)(1-\gamma)}}{2}$ . Then*

- *when  $g \geq 2(b - b_0)$ , so the monopolist fully covers the market, then  $W_{MA} = W_A$ ;*
- *when  $g < 2(b - b_0)$ , so the monopolist does not fully cover the market, then:*
  - *when  $g < 4b_0$ ,  $W_{MA} < W_A$  always holds;*
  - *when  $g > 4b_0$ , then  $W_{MA} < W_A$  if  $\mathcal{M}_{MA} < \bar{\mathcal{M}}$  and  $W_A < W_{MA}$  otherwise. In particular, for  $\gamma$  infinitesimally close to 1,  $W_A < W_{MA}$  if  $g > 2b_0 + \frac{2}{3}b$ .*

When  $g \geq 2(b - b_0)$ , the monopolist fully covers the market, so the allocation is identical to that with atomistic travelers only. When, instead,  $g < 2(b - b_0)$ , the monopolist reduces coverage. If the congestion cost is low relative to the benefit for the marginal type under full coverage ( $g < 4b_0$ ), excluding some agents is

welfare reducing<sup>8</sup>, so the monopolist's screening is inefficient no matter how large the monopoly is. If, instead, congestion is sufficiently high ( $g > 4b_0$ ), excluding some low types would be welfare enhancing. In this case, the emergent monopoly increases welfare provided that it does not exclude too many of them, which is ensured by the condition on  $\bar{\mathcal{M}}$ .

The emergence of a monopolist weakly reduces congestion vis-à-vis an environment with atomistic users only. However, this reduction has ambiguous welfare effects. When congestion is not such a severe problem to start with, the monopolist reduces welfare. When congestion becomes a problem, the monopolist may improve welfare, provided that its incentives to restrict coverage are not too large relative to the welfare-maximizing level.

### 4.3 Monopoly

In this Section, we investigate the case in which only corporate travelers exist. In setting the fares, the monopolist faces the problem of giving travelers the appropriate incentive to choose whether or not to travel and which lane to choose. Recalling that  $n$  and  $N$  denotes, as usual, the total mass of travelers in the popular and luxury lane, respectively, the monopolist problem is

$$\max_{f, F} fn + FN \quad (23)$$

$$s.t. \quad B(\theta'_M) - \theta'_M gn - f \geq 0, \quad (24)$$

$$B(\theta''_M) - \theta''_M gN - F \geq B(\theta'_M) - \theta'_M gn - f. \quad (25)$$

where  $\theta'_M$  and  $\theta''_M$  are the marginal travelers in the popular and luxury lane, respectively. Eqn. (24) is the individual rationality constraint for the marginal agent traveling in the popular lane, while equation (25) is the incentive compatibility constraint of the marginal agent traveling in the luxury lane. Because of Assumption 1, (25) implies that the individual rationality constraint for the marginal traveler in the luxury lane holds too. As to the incentive compatibility constraint for the marginal traveler in the popular lane, it holds whenever (24) holds as an equality. As it is standard, this constraint holds as an equality at the solution of the monopolist's problem. Also, (25) holds as an equality, which has to be true as long as there is a strictly positive amount of travelers in each lane. Using the two constraints (24) and (25) and the fact that  $\theta'_M = (1 - n - N)$  and  $\theta''_M = (1 - N)$ ,

---

<sup>8</sup>Note that we cannot use the social planner's full coverage conditions as our reference. When atomistic and corporate travelers coexist, the monopolist cannot differentiate congestion across lanes. As a result of this distortion, for any total mass of travelers in the two lanes, welfare is different (and lower) than with the social planner, where the two lanes have different levels of congestion.



fares are as follows

$$f = B(\theta'_M) - \theta'_M g n = B(1 - n - N) - [1 - n - N] g n, \quad (26)$$

$$F = f - g(\theta''_M)(1 - 2\theta''_M + \theta'_M) = f + g(1 - N)(n - N). \quad (27)$$

While the popular fare,  $f$ , is set to make the marginal traveler just willing to travel, the luxury fare,  $F$ , clearly illustrates the trade-off the monopolist faces in choosing the optimal degree of differentiation in the number of travelers in the two lanes. On the one hand, a large difference in the mass of travelers across the two lanes entails a larger extra-fee paid by the travelers in the luxury lane, hence a higher mark-up in the luxury lane. On the other hand, a large quality differentiation across the two lanes implies that the luxury lane, which yields the higher mark-up, is small. At the optimum, the monopolist strikes the balance between the two forces, by solving:

$$\begin{aligned} \max_{n, N} & (B(1 - N - n) - (1 - N - n) g n) n + \\ & (B(1 - N - n) - (1 - N - n) g n + g(1 - N)(n - N)) N \\ \text{s.t. } & 0 \leq N \leq n \leq 1 \text{ and } n + N \leq 1 \end{aligned} \quad (28)$$

At an interior solution, the solutions to problem (28), denoted by  $n_M$  and  $N_M$ , are implicitly defined by the following set of FOCs

$$\begin{aligned} B(1 - n_M - N_M) - B'(1 - n_M - N_M)(n_M + N_M) + \\ - g n_M(2 - 4N_M - 3n_M) = 0 \end{aligned} \quad (29)$$

$$\begin{aligned} B(1 - n_M - N_M) - B'(1 - n_M - N_M)(n_M + N_M) + \\ + g(3N_M^2 + 2n_M^2 - 2N_M) = 0 \end{aligned} \quad (30)$$

The optimal choice of  $N_M$  as a function of  $n_M$ , obtained by equating the equations (29) and (30) and solving it w.r.to  $N_M$  is given by

$$N_M(n_M) = \frac{1}{3} \left( 1 + 2n_M - \sqrt{7n_M^2 - 2n_M + 1} \right), \quad (31)$$

a relationship illustrated in Figure 1 by the red line.

Before discussing the features of the monopolist's choices, we characterize the equilibrium in case of full coverage:

**Proposition 5.** *Let  $\bar{n}_M$  and  $\bar{N}_M$  denote the solutions to problem (28) when there is full coverage. Then,*

$$\begin{aligned} \bar{n}_M &= \frac{3 + \sqrt{3}}{6} \cong 0.7887; \\ \bar{N}_M &= \frac{3 - \sqrt{3}}{6} \cong 0.2113. \end{aligned} \quad (32)$$

*These solutions occurs whenever*

$$g \geq \frac{6 (B'(0) - B(0))}{4 + \sqrt{3}} \cong 1.0467 \times (B'(0) - B(0)). \quad (33)$$

The optimal solutions to the monopolist's problem in (28) when there is full coverage is illustrated in Figure 1. This is given by the solid red circle at the intersection between the implicit relationship between  $n_M$  and  $N_M$  in (31) and the constraint  $n + N = 1$ .

Under some conditions, more travelers travel in a monopolistic market relative to the social optimum. Indeed, whenever  $0.8773 \times B'(0) < g < B'(0)$  (where the fact that  $g < B'(0)$  comes from Assumption 1), full coverage occurs under the monopoly but not under the social planner. This is clearly at odds with the standard result that a monopolist reduces total output. However, our result echoes the possibility that a monopolist underprovides quality relative to the social optimum (Spence, 1979). Intuitively, in our model, the monopolist's choices not only determine the mass of travelers but also the level of congestion and then quality in each lane. Hence, when  $B'(0)$  is sufficiently low relative to  $g$ , the monopolist has incentives to inefficiently dispatch low- $\theta$  travelers, which would be excluded by a social planner because of their external congestion cost.

We then look at the degree of differentiation across lanes, as compared to the socially optimal one. While the complexity of the explicit solutions makes it difficult to establish this comparison when solutions are interior, the result is clear cut in case of full coverage. When both a monopolist and a social planner choose to fully cover the market, the monopolist prefers a larger than optimal degree of quality (i.e. congestion) differentiation across lanes. Too few travelers travel in the luxury lane and too many travelers travel in the popular lane, as compared to the social optimum. This is the result of the IC constraint that, as it is clear from a simple inspection of the incentive compatible fare for the luxury lane in (27), allows to charge a higher luxury fare in the high quality market the larger the quality differential across lanes, providing an incentive to overdifferentiate quality with respect to the social optimum. This intuition squares with the results in Mussa and Rosen (1978).

A monopoly that manages all the travelers fully internalizes the congestion externality. At the same time, however, other distortions, related to market power and to distorted incentives in terms of quality provision, emerge. Such distortions have an unambiguous effect in terms of the level of differentiation of congestion across lanes: the monopolist overdifferentiates vis-à-vis the social planner. In moving from atomistic travelers to fully centralized travel, therefore, we transition from underdifferentiation to overdifferentiation. In the parameter space where the monopolist also fully covers the market, the only difference between the two market structures is the level of differentiation across lanes (under atomistic travelers the

market is always fully covered). In this case, welfare turns out to be larger with atomistic travelers than with travel managed by a monopolist, and the welfare dominance of atomistic travel over monopoly increases with the severity of the congestion cost  $g$ .

When, instead, the monopolist does not fully cover the market, the welfare effects of the transition turns ambiguous. In this case, besides overdifferentiating, the monopolist also has an incentive to restrict the number of vehicles vis-à-vis atomistic travel. This restriction can be excessive from the welfare standpoint, in particular when  $g$  is small relative to the slope of the benefit function.

## 5 Equilibrium analysis with taxes

The welfare effects of the transition from atomistic to centralized travel entailed by AVs are dramatically affected by the possibility to implement a taxation scheme. An appropriate taxation implements the first best. This Section characterizes the properties of that taxation scheme. We show that the structure of the optimal taxes with a centralized market differs sharply from that of a decentralized market with atomistic travelers. We then discuss the political feasibility of the first-best restoring tax schemes throughout the transition towards a centralized market. We focus on per-vehicle unit taxes and we denote with  $t$  and  $T$  the unit tax levied on AVs traveling in the popular and luxury lane, respectively.

Our main interest in the transition from a market with atomistic travelers only to one fully served by a centrally managed fleet of AVs. The market where atomistic and corporate travelers coexist is just an intermediate stage. In spite of this, for expositional reasons, we first present the two polar cases of atomistic and corporate travelers only and then illustrate the analysis of the case in which both types of travelers are traveling.

### 5.1 Taxation with atomistic travelers

In an environment with atomistic agents and externalities, it is well known that optimality is restored by Pigouvian taxes, which impose on each agent the non internalized social cost. In this Section, we characterize these charges in our framework, and we discuss their political feasibility as well as their distributional effects.

The timeline of the game we solve is the following. First, the tax authority announces a tax scheme. Second, travelers observe it and make their traveling decisions.

Road charges need to restore differentiation in congestion levels across lanes at its welfare-maximizing level, and, when full market coverage is not socially efficient possibly, to reduce total travel. Let  $t_{AT}$  and  $T_{AT}$  denote the unit taxes

levied on AVs traveling in the popular and luxury lane, respectively, that restore social optimum. These are described in the following:

**Proposition 6.** *Assume all travelers are atomistic. The pairs of taxes that replicate the social optimum are as follows*

$$\begin{cases} t_{AT} \leq B(0) \\ T_{AT} = t_{AT} + g \frac{5-\sqrt{7}}{18} \end{cases} \quad \text{if } g \leq \frac{72 B(0)}{(1 + \sqrt{7})^2} \cong 5.4179 \times B(0); \quad (34)$$

$$\begin{cases} t_{AT} = g n_P (1 - N_P - \frac{n_P}{2}) \\ T_{AT} = t_{AT} + g (1 - N_P) (n_P - N_P) \end{cases} \quad \text{if } g \geq \frac{72 B(0)}{(1 + \sqrt{7})^2} \cong 5.4179 \times B(0). \quad (35)$$

Taxes increase the relative price of the luxury relative to the popular lane ( $T_{AT} > t_{AT}$ ), thereby shifting some travelers accordingly.

In particular, when  $B(0)$  is small, misallocation of travelers across lanes is the only distortion to be solved, since the social planner as well covers the entire market. Hence, the optimal pair of taxes - as in (34) - should not restrict market coverage. A multiplicity of low enough  $t_{AT}$ , including  $t_{AT} = 0$ , delivers this. On the other hand, optimal differentiation across lanes is obtained through an appropriate difference between  $T_{AT} - t_{AT}$ , which ensures that the location of the traveler indifferent across the two lanes mimics that of the social planner. Notice that the indeterminacy of  $t_{AT}$  provides a flexible set of alternatives to the tax authority, ranging from solutions that minimize tax burden (when only those who travel in the luxury lanes are taxed), to others associated to a larger tax burden, paid by travelers in both lanes.

When instead  $B(0)$  is large, a social planner would screen out of the market agents with a low value for the travel, so both the total number of travelers and their allocation across lanes have to be corrected. The popular tax in (35) is then uniquely determined and ensures that the market coverage replicates the first best. The luxury tax still induces an optimal degree of differentiation across lanes.

Taxes as in (34) and (35) are congestion charges, which represent the external cost imposed by the marginal travelers on fellow riders through their traveling decisions. They align the incentives of the marginal travelers to social optimality.

The distributional consequences, hence the political feasibility, of such congestion charges depend on how the revenues from them is used (Small, 1992). Since congestion charges are welfare-improving, an appropriate redistribution scheme that fully compensates losers could be Pareto-improving. However, probably as a result of imperfect or unclear compensations, charges are usually not liked by citizens, and therefore rarely implemented (Oberholzer-Gee and Weck-Hannemann, 2002). Without compensation, low  $\theta$ 's stand to lose from the charge. If the market

remains fully covered after the charge is implemented, all the  $\theta$ 's then traveling in the popular lane face more congestion, and the resulting lower utility. If the charge excludes the lower portion of the  $\theta$ s, those excluded are worse-off as a result of the uncompensated charge.

## 5.2 Taxation in pure monopoly

We now look at the welfare maximizing taxation scheme to be imposed to a company that manages the entire AVs fleet. We show that it is remarkably different than the tax to be applied to atomistic travelers. Indeed, a monopolist perfectly internalizes congestion externality, so congestion charges are not appropriate. Alternative distortions, typical of a monopolist choosing quality levels, emerge. We show that the first best - restoring scheme turns out to be challenging, from a political economy perspective, when congestion is particularly severe ( $g$  is sufficiently high), so that efficiency requires restricting the total number of vehicles. In this case, the scheme actually involves a net transfer in favor of the company. To improve its acceptability, a policymaker could collect a license fee from the company to balance its budget within the AV market.

In our setting, the monopolist generates two distortions with respect to the social planner. First, it overdifferentiates traffic across the two lanes compared to the welfare optimal level, in order to better segment the market, *à la* Mussa and Rosen (1979). Second, it modifies the overall number of vehicles traveling.

The timeline of the game we solve is the following. First, the tax authority announces a tax scheme. Second, the monopolist observes it and sets the fees  $F$  and  $f$ . Third, travelers observe the fees and make their traveling decisions.

We consider a per-vehicle tax/subsidy, potentially differing by lane, imposed on the monopolist.<sup>9</sup> The maximization problem the monopolist faces is:

$$\begin{aligned} \max_{f, F} & (f - t(n, N))n + (F - T(n, N))N \\ \text{s.t.} & (24) - (25) \end{aligned} \tag{36}$$

where (24) and (25) are the same individual rationality and incentive compatibility constraints faced by the monopolist in the absence of taxes and already discussed in Section 4.3. Solving these two constraints w.r.to the monopolist's fares as in (26) and (27) and plugging them into the monopolist's problem in (36) allows to

---

<sup>9</sup>Imposing the tax on the monopolist or on the travelers produces identical effects.

rewrite it as follows

$$\begin{aligned} \max_{n,N} \quad & (B(1-N-n) - (1-N-n)gn - t(n, N))n + \\ & (B(1-N-n) - (1-N-n)gn + g(1-N)(n-N) - T(n, N))N \\ \text{s.t.} \quad & 0 \leq N \leq n \leq 1 \text{ and } n + N \leq 1. \end{aligned} \quad (37)$$

Let the solutions to this problem be denoted by  $n_{MT}$  and  $N_{MT}$ . When solutions are interior, these are implicitly defined by the following set of FOCs

$$\begin{aligned} & B(1 - n_{MT} - N_{MT}) - B'(1 - n_{MT} - N_{MT})(n_{MT} + N_{MT}) + \\ & - gn_{MT}(2 - 4N_{MT} - 3n_{MT}) - t(n_{MT}, N_{MT}) + \\ & - \frac{\partial t(n_{MT}, N_{MT})}{\partial n_{MT}}n_{MT} - \frac{\partial T(n_{MT}, N_{MT})}{\partial N_{MT}}N_{MT} = 0 \end{aligned} \quad (38)$$

$$\begin{aligned} & B(1 - n_{MT} - N_{MT}) - B'(1 - n_{MT} - N_{MT})(n_{MT} + N_{MT}) + \\ & + g(3N_{MT}^2 + 2n_{MT}^2 - 2N_{MT}) - T(n_{MT}, N_{MT}) + \\ & - \frac{\partial T(n_{MT}, N_{MT})}{\partial N_{MT}}N_{MT} - \frac{\partial t(n_{MT}, N_{MT})}{\partial N_{MT}}n_{MT} = 0 \end{aligned} \quad (39)$$

Let  $t_{MT}$  and  $T_{MT}$  denote the per-vehicle tax/subsidy in the popular and luxury lane respectively that restore the social optimum. We establish the following result:

**Proposition 7.** *Let*

$$s_{MT} \equiv \begin{cases} 0 & \text{if } g \leq \frac{36(B(0)-B'(0))}{(1+\sqrt{7})^2} \\ & \cong 2.7085 \times (B(0) - B'(0)); \\ B'(0) - B(0) + g \frac{4+\sqrt{7}}{18} & \text{if } \frac{36(B(0)-B'(0))}{(1+\sqrt{7})^2} \leq g \leq \frac{72B(0)}{(1+\sqrt{7})^2}; \\ B'(1 - n_P - N_P)(n_P + N_P) + \\ + g(2N_P - n_P^2 - \frac{3}{2}N_P^2) & \text{if } g \geq \frac{72B(0)}{(1+\sqrt{7})^2} \cong 5.4179 \times B(0). \end{cases} \quad (40)$$

*The social optimum is restored by a system of per-vehicle taxes of the following form*

$$\begin{aligned} t_{MT} &= gn - s_{MT}; \\ T_{MT} &= gN - s_{MT}. \end{aligned} \quad (41)$$

The Proposition illustrates that social optimality is restored by imposing on the monopolist a system of per-vehicle taxes/subsidies based on the mass of travelers, differentiated by lane.  $t_{MT}$  and  $T_{MT}$  consist of a tax component, which is larger

the larger the mass of travelers in that lane, and a subsidy component, exogenously determined by the tax authority, based on the socially optimal number of travelers and equal across the two lanes.

To understand the logic of this result, first notice that, by plugging the per-vehicle taxes given in (41) into the FOCs in (38) and (39) and then equalizing them to solve w.r.to  $N_{MT}$ , one obtains

$$N_{MT}(n_{MT}) = \frac{1}{3} \left( 2(1 + n_{MT}) - \sqrt{7n_{MT}^2 - 4n_{MT} + 4} \right). \quad (42)$$

This condition reproduces exactly the equivalent condition (9) for the case of the social planner. This implies that the tax components of  $t_{MT}$  and  $T_{MT}$  induce the monopolist to provide the same differentiation across lanes as the social planner. Indeed, (42) results only from the tax components of  $t_{MT}$  and  $T_{MT}$ , while it is unaffected by the subsidy component. The tax components of the tax provide an incentive to shift passengers from the popular to the luxury lane, as long as  $n > N$ , thereby softening the monopolist's incentive to overcrowd the popular lane.

The taxes restoring the welfare-maximizing allocation across lanes differs dramatically vis-à-vis the case of a market with atomistic travelers. Since the monopolist perfectly internalizes the congestion externality, the optimal taxes are not congestion charges. They instead align the monopolist's incentives to set the quality level of the two lanes to welfare optimality. The logic of this tax is similar, for instance, to that obtained by Mosca and Lambertini (1999), and by Cremer and Thisse (1994) in the context of a vertically differentiated oligopoly.

An additional instrument is however needed to eliminate the distortion caused by the monopolist in screening travelers out to the market in a socially inefficient manner. Without taxes, as discussed in Section 4.3, the total number of travelers may be below or above the social optimum, depending on the monopolist's incentives to overprovide or underprovide quality relative to social optimum. After introducing the tax components of  $t_{MT}$  and  $T_{MT}$ , however, the monopolist always reduces total travel below the efficient level. To ensure a socially efficient number of travelers, the monopolist must be granted a subsidy on the total number of cars traveling, as in (40).

The two components of the taxes serve very different purposes, and also have remarkably different features. The levels of the tax components only depend on the monopolist's choices, being dependent on  $n$  and  $N$ . In this respect, it is a very simple tax to set, since it does not require any specific knowledge by the tax authority, if not the value of  $g$ , which represents the magnitude of the effect of congestion on the travelers' benefits. On the other hand, the subsidy component requires a deeper knowledge of the market and a great computational ability by the tax authority, being based on a perfect knowledge of the the travelers' benefit function and of the solution to the first best problem.

Overall, in equilibrium, the subsidy component may exceed the tax component, so that the monopolist receives a net subsidy from the tax authority. This situation always occurs when congestion is severe enough ( $g$  sufficiently high) that efficiency commands not to fully cover the market. The AVs tax system therefore requires, under these circumstances, to absorb some funding from general taxation. Pels and Verhoef (2004) emphasize the political difficulties in implementing a negative tax scheme on airline companies. The same logic might well apply to the AVs market. The political feasibility of this tax scheme appears very dubious, even more so than the congestion charges. Indeed, congestion charges, with an appropriate redistribution scheme, can in principle improve each agent's welfare, while a negative tax would, on aggregate, harm citizens. A potential solution to improve political feasibility is to associate the scheme to an upfront fixed license that preserves the budget neutrality.

When, instead, congestion is not so severe a problem, and  $g$  is relatively low, the tax/subsidy scheme involves a net payment by the company, which improves acceptability.

The tax/subsidy scheme also has distributional consequences across different types of travelers, since it modifies the fares  $f$  and  $F$  charged in the popular and in the luxury lane, respectively.

Under full coverage,  $f$  remains unaltered, while travelers in the popular lane enjoy a lower congestion level. Travelers in the popular lane thus unambiguously benefit from the scheme.  $F$  instead decreases, as congestion in the luxury lane increases, with an unclear net utility effect for those travelers.

Under partial coverage, the mass of travelers increases or decreases depending on whether the mass of travelers without taxes is above or below social optimum. When the tax/subsidy is set to decrease the number of travelers (because the social planner dispatches more vehicles than the monopolist without taxes), the marginal  $\theta$  increases, and the congestion level in the popular lane decreases. As a result of both effects,  $f$  unambiguously rises. Travelers in the popular lane enjoy a lower congestion, so the sign of the net change in their utility is unclear. When, instead, the tax/subsidy is set to increase the amount of vehicles, the effects on  $f$  and  $F$  are a priori unclear.

### 5.3 Taxation with atomistic travelers and monopoly

In this Section, we characterize the set of taxes that restores the social optimum when a group of decentralized atomistic travelers coexist with a group of corporate travelers managed by a single company. We show that the taxing schemes markedly differ across the two groups. The taxing scheme imposed to each group closely resembles that charged to that group in isolation.



To find the first-best restoring set of taxes, we use the fact that travelers, atomistic and corporate, are exogenously allocated to their groups in fixed proportions, equal to  $\gamma$  and  $1 - \gamma$ , respectively. In solving for the optimal taxes for each group, we take the allocation across lanes of the travelers in the other group as given. A system of optimal taxes is then just composed of a system of taxes on corporate travelers that restores the social optimum for this group, *given* the system of optimal taxes for the atomistic travelers, and viceversa.

The timeline of the game we solve is the following. First, the tax authority announces a tax scheme  $T(n, N)$  and  $t(n, N)$ , possible differentiated across groups of travelers. Second, the monopolist observes it and sets the fees  $F$  and  $f$ . Third, corporate and atomistic travelers observe the fees, and simultaneously make their travel decisions.

**Taxes on atomistic travelers.** Denote the corporate travelers allocation that replicates, pro quota, the socially optimal allocation, by  $m_P \equiv (1 - \gamma)n_P$  and  $M_P \equiv (1 - \gamma)N_P$  for the popular and luxury lane, respectively. Let the system of taxes on atomistic travelers that, given the corporate traveler allocation, replicates the social optimum be denoted as  $t_{MAT}^A$  and  $T_{MAT}^A$  for the popular and luxury lane, respectively. Then,

**Proposition 8.** *Assume the market is populated by atomistic and corporate travelers. Let the corporate travelers be allocated as to replicate, pro quota, the first best, so that  $m = m_P$  and  $M = M_P$ . The pairs of taxes on atomistic travelers that replicate the social optimum are  $t_{MAT}^A = t_{AT}$  and  $T_{MAT}^A = T_{AT}$  as given in Proposition 6.*

The Proposition illustrates that, provided that corporate travelers are allocated optimally, a congestion tax identical to that derived in the case of atomistic travelers only is able to allocate them in a socially optimal manner. The equivalence stems from the incentives by atomistic travelers to neglect the external impact of their decisions. Given that corporate travelers are optimally allocated across lanes, the equilibrium level of congestion is the same in the two cases of atomistic only and of co-existence. Hence the tax that aligns the individual utility to welfare maximizing choice for the marginal atomistic agents (the  $\theta$  type indifferent between not traveling and traveling in the popular lane, and the  $\theta$  type indifferent between traveling in the popular and traveling in the luxury lane) is also the same.

**Taxes on corporate travelers.** Take now the allocation of atomistic travelers as given. Denote with  $a$  the mass of atomistic travellers allocated in the popular lane and with  $A$  the mass of those in the luxury lane. Denoting, as in Section 4.2, the mass of corporate travelers in the popular and luxury lane by  $m$  and  $M$ ,

respectively, the monopolist problem is

$$\max_{f,F} (f - t)m + (F - T)M \quad (43)$$

$$s.t. \quad B\left(\tilde{\theta}'_M\right) - \tilde{\theta}'_M g(a + m) - f \geq 0, \quad (44)$$

$$B\left(\tilde{\theta}''_M\right) - \tilde{\theta}''_M g(A + M) - F \geq B\left(\tilde{\theta}''_M\right) - \tilde{\theta}''_M g(a + m) - f. \quad (45)$$

where  $\tilde{\theta}'_M$  and  $\tilde{\theta}''_M$  are the marginal corporate travelers in the popular and luxury lane, respectively. Equation (44) is the individual rationality constraint for the marginal corporate consumer traveling in the popular lane, while equation (45) is the incentive compatibility constraint of the marginal consumer traveling in the luxury lane. Constraints (44) and (45) are the exact analog of the constraints faced by the monopolist without tax, i.e. (24) and (25), and the discussion provided there applies. There are however two differences relatively to the other constraints. First, the congestion cost faced by each traveler depends not only on the mass of corporate travelers in the same lane, but also on the mass of atomistic travelers. Hence, the congestion cost for a type- $\theta$ , say, in the popular lane is given by  $\theta g(a + m)$ . Second, the marginal corporate  $\theta$  is expressed in terms of corporate consumers only, that is only on the fraction  $1 - \gamma$  of all travellers. This implies that  $\tilde{\theta}'_M = 1 - \frac{m}{1-\gamma} - \frac{M}{1-\gamma}$  and  $\tilde{\theta}''_M = 1 - \frac{M}{1-\gamma}$ .

Using these considerations, fares are as follows

$$f = B\left(1 - \frac{m}{1-\gamma} - \frac{M}{1-\gamma}\right) - \left[1 - \frac{m}{1-\gamma} - \frac{M}{1-\gamma}\right] g(a + m), \quad (46)$$

$$F = f + g\left(1 - \frac{M}{1-\gamma}\right) (a + m - A - M). \quad (47)$$

This allows us to rewrite the monopolist's problem as follows

$$\max_{m,M} \left\{ B\left(1 - \frac{m}{1-\gamma} - \frac{M}{1-\gamma}\right) - \left[1 - \frac{m}{1-\gamma} - \frac{M}{1-\gamma}\right] g(a + m) - t + \right\} m + \quad (48)$$

$$\left\{ B\left(1 - \frac{m}{1-\gamma} - \frac{M}{1-\gamma}\right) - \left[1 - \frac{m}{1-\gamma} - \frac{M}{1-\gamma}\right] g(a + m) + \left[1 - \frac{M}{1-\gamma}\right] g[a + m - A - M] - T + \right\} M$$

$$s.t. \quad 0 \leq M \leq m \leq 1 \text{ and } m + M \leq 1 - a - A$$

Let the solutions to this problem be denoted by  $m_{MAT}$  and  $M_{MAT}$ . When solutions

are interior, these are implicitly defined by the following set of FOCs:

$$\left[ \frac{g(a + m_{MAT})}{1 - \gamma} - \frac{B' \left( 1 - \frac{m_{MAT}}{1 - \gamma} - \frac{M_{MAT}}{1 - \gamma} \right)}{1 - \gamma} - g \left( 1 - \frac{m_{MAT}}{1 - \gamma} - \frac{M_{MAT}}{1 - \gamma} \right) \right] \times \quad (49)$$

$$(m_{MAT} + M_{MAT}) + B \left( 1 - \frac{m_{MAT}}{1 - \gamma} - \frac{M_{MAT}}{1 - \gamma} \right) - g \left( 1 - \frac{m_{MAT}}{1 - \gamma} - \frac{M_{MAT}}{1 - \gamma} \right) \times \\ (a + m_{MAT}) - g(a + 2m_{MAT})x + g \left( 1 - \frac{M_{MAT}}{1 - \gamma} \right) M_{MAT} + s = 0;$$

$$\left[ \frac{g(a + m_{MAT})}{1 - \gamma} - \frac{B' \left( 1 - \frac{m_{MAT}}{1 - \gamma} - \frac{M_{MAT}}{1 - \gamma} \right)}{1 - \gamma} \right] (m_{MAT} + M_{MAT}) + \quad (50) \\ - g \left[ \frac{a + m_{MAT} - A - M_{MAT}}{1 - \gamma} + \left( 1 - \frac{M_{MAT}}{1 - \gamma} \right) + 1 \right] M_{MAT} + \\ B \left( 1 - \frac{m_{MAT}}{1 - \gamma} - \frac{M_{MAT}}{1 - \gamma} \right) - g \left( 1 - \frac{m_{MAT}}{1 - \gamma} - \frac{M_{MAT}}{1 - \gamma} \right) (a + m_{MAT}) + \\ + g \left( 1 - \frac{M_{MAT}}{1 - \gamma} \right) (a + m_{MAT} + A + M_{MAT}) - g(A + M_{MAT}) + s = 0.$$

Denote the atomistic travelers allocation which replicates, pro quota, the socially optimal allocation, by  $a_P \equiv \gamma n_P$  and  $A_P \equiv \gamma N_P$  for the popular and luxury lane, respectively. Let the system of taxes on corporate travelers that, given the atomistic travellers allocation, replicates the social optimum be denoted as  $t_{MAT}^M$  and  $T_{MAT}^M$  for the popular and luxury lane, respectively. Then,

**Proposition 9.** *Assume the market is populated by atomistic and corporate travelers and that atomistic travelers are allocated to replicate, pro quota, the socially optimal allocation, so that  $a = a_P$  and  $A = A_P$ . For any given  $\gamma \in [0, 1]$  and for any socially optimal allocation  $n_P$  and  $N_P$ , let:*

- $x_{MAT} \equiv x_{MAT}(\gamma, n_P, N_P)$ ,
- $s_{MAT} \equiv s_{MAT}(x_{MAT}; \gamma, n_P, N_P)$ ,

where the explicit forms of these functions are provided in the proof.

The social optimum is restored by a system of per-vehicle taxes on corporate travelers of the following form

$$t_{MAT}^M = g(m_{MAT} + a_P)x_{MAT} - s_{MAT}(x_{MAT}); \quad (51) \\ T_{MAT}^M = g(M_{MAT} + a_P) - s_{MAT}(x_{MAT}).$$

The Proposition illustrates that, provided that atomistic travelers are allocated optimally, the system of per-vehicle taxes/subsidies that restore the social optimum for corporate travelers shows some similarities to the case of a monopolist only. It features a tax component, different across lanes and based on the total mass of travelers in the same lane, and a subsidy component, exogenously given and identical across the two lanes.

There is, however, a clear difference between the two cases. The introduction of a fraction of atomistic travelers, already optimally allocated across lanes, modifies the monopolist's incentives in terms of differentiation. As  $\gamma$  grows larger, the contribution of corporate travelers to congestion in each lane turns smaller, and the monopolist allocates an increasingly high proportion of corporate travelers to the luxury lane. When  $\gamma$  is low, this effect is small, and the monopolist still overdifferentiates across lanes vis-à-vis the social planner. The optimal tax, then, is geared at shifting some travelers to the luxury lane. It is then lower in the luxury lane. When  $\gamma$  is high, instead, this effect prevails, and the monopolist underdifferentiates with respect to the planner. The optimal tax in this case induces some travelers to shift to the popular lane. It is then lower in the popular lane. The tax then must depend on  $\gamma$ , so as to restore the correct proportion of travelers across lanes for different sizes of atomistic travelers. To achieve this, we weight the tax on the popular lane by  $x_{MAT}$  function of  $\gamma$ . As  $\gamma$  increases from 0,  $x_{MAT}$  moves down from 1. For high enough values of  $\gamma$ ,  $x_{MAT}$  is small enough that the tax on the popular lane turns smaller than the tax on the luxury, thereby inducing travelers to shift to the popular lane.

## 6 Conclusions

AVs are expanding very rapidly. For instance, in 2016 General Motors bought Cruise, an autonomous vehicle start-up, for more than \$1bn, and Uber acquired Otto, a self-driving trucking company, for \$680 million. In 2017, Intel agreed to buy Mobileye, an Israeli car technology firm producing self-driving sensors and software, in a deal worth more than \$15bn.

AVs are triggering a large-scale debate. However, it mostly focuses on technological, ethical and legal issues (namely, liability). Our paper analyzes the incentives involved and the welfare effects of the transition from a market with atomistic drivers to a system of centralized traffic, managed by centralized companies.

We first analyze the transition process when there is no road pricing, reflecting the challenges in implementing this taxation scheme in the real world, due to the lack of popular support.

In this case, at the beginning of the transition process from decentralized to centralized traffic, when a small monopolistic supplier emerges, traffic (weakly)

reduces. The sign of the resulting welfare effect is ambiguous. It tends to be positive when congestion was severe in the first place, and negative otherwise. In the transition from decentralized to centralized traffic, we observe a shift from underdifferentiation to overdifferentiation of the congestion level across lanes. The monopolist uses differentiation of the congestion levels to extract profit through market segmentation. Overall, the analysis without taxes emphasizes the tradeoff between the benefits of monopolistic centralization, in terms of internalization of congestion externalities, and its costs, in terms of market power exploitation and quality distortions.

We then analyze first-best restoring taxes. Taxes are remarkably different in the atomistic and in the centralized monopolistic worlds. They address two very different types of distortion in the two cases - externality in the atomistic case, solved by a congestion charge, and market power and distorted quality level in the monopolistic case, solved by a tax/subsidy. We show that, when the congestion problem is particularly severe, restoring first best in monopoly requires a net subsidy to the monopolist, which may prove politically challenging. To improve its acceptability, a policymaker could collect a license fee from the company to balance its budget within the AV market. We finally show that, when travelers managed by a monopolist coexist with atomistic travelers, the optimal taxing schemes markedly differ across the two groups. The taxing scheme imposed to each group closely resembles that charged to that group in isolation.

## References

- [1] Arnott, R., de Palma, A., Lindsey, R. 1990. Economics of a bottleneck. *Journal of Urban Economics*, 27, 111-130.
- [2] Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., Rahwan I. 2018. The Moral Machine experiment. *Nature*. 563, 59-64.
- [3] Basso, L. 2008. Airport deregulation: Effects on prices and capacity. *International Journal of Industrial Organization*, 26, 1015-1031.
- [4] Basso, L., Zhang, A. 2007. Congestible facility rivalry in vertical structures. *Journal of Urban Economics* 61 218-237.
- [5] Brueckner, J. K. 2002. Airport congestion when carriers have market power. *American Economic Review* 92 (5) 1357-1375.
- [6] Brueckner, J. K. 2005. Internalization of airport congestion: A network analysis. *International Journal of Industrial Organization* 23(7-8) 599-614.

- [7] Brueckner, J. K, Verhoef, E. 2010. Manipulable congestion tolls. *Journal of Urban Economics* 67 315-321.
- [8] Cremer, H., Thisse, J. 1994. Commodity taxation in a differentiated oligopoly. *International Economic Review*, vol. 35, no. 3, 1994, pp. 613–633.
- [9] Daniel, J. I. 1995. Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues. *Econometrica* 63(2) 327-370.
- [10] Daniel, J. I., Harback, K. T. 2008. (When) Do hub airlines internalize their self-imposed congestion delays? *Journal of Urban Economics* 63(2), 583-612.
- [11] Lambertini, L., Mosca M. 1999. On the Regulation of a Vertically Differentiated Market. *Australian Economic Papers*, 38 (4), 354-366.
- [12] Lamotte, R., De Palma, A., Geroliminis, N. 2016. Sharing the road: the economics of autonomous vehicles. *Elitisme Working Paper Series* 2016-002.
- [13] Lindsey, R., De Palma, A., Silva, H. 2019. Equilibrium in the Bottleneck Model with Atomic and non Atomic Users. *Transportation Research Part B: Methodological*, 124, 82-107.
- [14] Mayer, C., Sinai, T. 2003. Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays Are Evil. *American Economic Review*. 93(4) 1194-1215.
- [15] Molnar, A. 2013 Congesting the commons: A test for strategic congestion externalities in the airline industry, mimeo.
- [16] Mussa, M., Rosen, S. 1978. Monopoly and Product Quality. *Journal of Economic Theory* 18, 301-317.
- [17] Oberholzer-Gee, F., Weck-Hannemann, H. 2002. Pricing road use: politico-economic and fairness considerations. *Transportation Research, Part D*, 7(2), 357-371.
- [18] Ostrovsky, M., Schwarz, M.. 2018. Carpooling and the Economics of Self-Driving Cars, mimeo.
- [19] Pels, E., Verhoef, E.. 2004. The economics of airport congestion pricing. *Journal of Urban Economics* 55 (2) 257-277.
- [20] Rupp, N. G. 2009. Do carriers internalize congestion costs? Empirical evidence on the internalization question. *Journal of Urban Economics*. 65(1) 24-37.

- [21] Schrank, D., Lomax, T., B. Eisele 2011. TTI's 2011 Urban Mobility Report. Texas Transportation Institute, The Texas A&M University System.
- [22] Silva, H, Verhoef, E. 2013. Optimal pricing of flights and passengers at congested airports and the efficiency of atomistic charges. *Journal of Public Economics*. 106, 1-13.
- [23] Silva H, Lindsey, R. de Palma, A, van den Berg, V., 2016. On the Existence and Uniqueness of Equilibrium in the Bottleneck Model with Atomic Users. *Transportation Science*, 51 (3), 863–881
- [24] Small, K. 1992. Using the revenues from congestion pricing. *Transportation*. 19, 359-381.
- [25] Small, K. 2012. Valuation of travel time. *Economics of Transportation*. 1, 2-14
- [26] Van Dender, K. 2005. Duopoly prices under congested access. *Journal of Regional Science*. 45 (2) 343–362.
- [27] Van den Berg, V., Verhoef, E. 2016. Autonomous cars and dynamic bottleneck congestion: The effects on capacity, value of time and preference heterogeneity. *Transportation Research Part B*. 94 43-60.
- [28] Vickrey, W.S. 1969. Congestion theory and transport investment. *American Economic Review*. 59 (2) 251-260.
- [29] Zhang, A., Zhang, Y. 2006. Airport capacity and congestion when carriers have market power. *Journal of Urban Economics* 60 (2) 229-247.

## Appendix

This Appendix contains the proofs of all Propositions of the paper.

**Proof of Proposition 1.** Results are obtained with the use of Kuhn-Tucker technique.

**Proof of Proposition 2.** First, we prove that  $n_A = N_A$ . By contradiction, assume that  $n_A > N_A$ : a traveler in the popular lane would find it optimal to switch to the luxury lane, where the cost of congestion is lower. An identical argument applies in the case of  $N_A > n_A$ . When, instead,  $n_A = N_A = \frac{1}{2}$ , no agent has an incentive to deviate, either by not traveling (since  $B(0) \geq 0$ ), or by traveling in another lane (since  $\frac{\partial G(n, \theta)}{\partial n} > 0$ ).

**Proof of Proposition 3.** Results are obtained with the use of Kuhn-Tucker technique.

**Proof of Proposition 4.** Let  $\Delta W \equiv W_{MA} - W_A$ . This is a cubic expression in  $\mathcal{M}$  whose three roots are  $\mathcal{M}_1 \equiv 1 - \gamma$ ,  $\mathcal{M}_2 \equiv \frac{2b+g(1-2\gamma)+\sqrt{4b(b-g)+g^2-4g(4b_0-g)(1-\gamma)}}{2}$  and  $\bar{\mathcal{M}} \equiv \frac{2b+g(1-2\gamma)-\sqrt{4b(b-g)+g^2-4g(4b_0-g)(1-\gamma)}}{2}$ .

It is easy to establish that  $\mathcal{M}_1 < \mathcal{M}_2$  and  $\bar{\mathcal{M}} < \mathcal{M}_2$ . Let  $\Delta \mathcal{M} \equiv \mathcal{M}_1 - \bar{\mathcal{M}}$ . Then  $\Delta \mathcal{M} = \frac{g-2b+\sqrt{4b(b-g)+g^2-4g(4b_0-g)(1-\gamma)}}{2g}$ . Next, notice that, iff  $b_0 > \frac{1}{4}g$ , we have that i)  $\Delta \mathcal{M} < 0$  and ii)  $\frac{\partial \Delta W}{\partial \mathcal{M}} \big|_{\mathcal{M}=1-\gamma} = \frac{1}{4}g - b_0 < 0$ .

Hence, two cases may occur:

- if  $b_0 > \frac{1}{4}g$ , then  $\mathcal{M}_1 < \bar{\mathcal{M}} < \mathcal{M}_2$  and  $\frac{\partial \Delta W}{\partial \mathcal{M}} \big|_{\mathcal{M}=1-\gamma} < 0$ .  $\Delta W$  is positive for  $\mathcal{M} < 1 - \gamma$  and, because of the ranking between the  $\mathcal{M}_i$ 's, for any admissible value of  $\mathcal{M}$ ;
- if  $b_0 < \frac{1}{4}g$ , then  $\bar{\mathcal{M}} < \mathcal{M}_1 < \mathcal{M}_2$  and  $\frac{\partial \Delta W}{\partial \mathcal{M}} \big|_{\mathcal{M}=1-\gamma} > 0$ .  $\Delta W$  is negative for any  $\mathcal{M}$  sufficiently close to  $1 - \gamma$  and, in particular, for any  $\mathcal{M} \in [\min\{\bar{\mathcal{M}}, \mathcal{M}_{MA}\}, \mathcal{M}_1]$ .

Next, notice that, as  $\gamma$  tends to 1, all  $\mathcal{M}_i$ 's and  $\mathcal{M}_{MA}$  tend to 0. Let  $L_{\mathcal{M}_1} \equiv \lim_{\gamma \rightarrow 1^-} \frac{\partial \mathcal{M}_1}{\partial \gamma}$  and  $L_{\mathcal{M}_{MA}} \equiv \lim_{\gamma \rightarrow 1^-} \frac{\partial \mathcal{M}_{MA}}{\partial \gamma}$ . It is easy to establish that  $L_{\mathcal{M}_1} < L_{\mathcal{M}_{MA}}$  iff  $b_0 < -\frac{1}{3}b + \frac{1}{2}g$ , which proves our result. ■

**Proof of Proposition 5.** Results are obtained with the use of Kuhn-Tucker technique.



**Proof of Proposition 6.** When  $B(0) \geq g \frac{\sqrt{7}+4}{36}$ , full coverage occurs both under the social planner and with atomistic travelers. Since the traveler located at  $\theta = 0$  gets utility  $B(0) - t_{AT}$  from traveling in the popular lane, because of Assumption 1 .i), a tax equal to  $t_{AT}$  means that full coverage occurs as in the social optimum. Given  $t_{AT}, T_{AT}$  as in (34) ensures that the marginal traveler in the luxury lane is as in the social optimum.

When  $B(0) \leq g \frac{\sqrt{7}+4}{36}$ , full coverage occurs under atomistic travelers only. Given  $t_{AT}, T_{AT}$  as in (35) ensures that the marginal traveler in the luxury lane is as in the social optimum. ■

**Proof of Proposition 7.** Equalize the FOCs (38) and (39) and substitute for  $t = t_{MT}$  and  $T = T_{MT}$ . Then, solve w.r.to  $N$  to get (42).

As to the subsidy  $s_{MT}$ , first notice that it enters linearly both FOCs (38) and (39), and then, when exogenously set, does not affect  $N_{MT}(n_{MT})$  in (42). In order to find the level of the subsidy that induces the monopolist to choose the social optimal quantity of travelers, focus first on interior solutions, i.e. no full coverage by the monopolist and the planner. Equalize the FOC w.r.to  $n$  in the monopolist problem in (37) - given in (38) - to the FOC w.r.to  $n$  in the social planner problem in (6) - given in (7) -, and solve with respect to  $s$ . This gives (40) when  $g \leq \frac{72 B(0)}{(1+\sqrt{7})^2}$ . An identical result is obtained when using the FOCs from the monopolist and planner problems w.r.to the mass of luxury travelers.

Focus now on the case of not interior solutions. Applying Kuhn-Tucker conditions, the monopolist subject to a system of tax/subsidy as in (41) fully covers the market (with a socially optimal degree of differentiation across lanes) when

$$g \leq \frac{36 (B(0) - B'(0) + s)}{(1 + \sqrt{7})^2}. \quad (52)$$

Notice that, when  $s = 0$ , (52) implies  $g \leq \frac{72 B(0)}{(1+\sqrt{7})^2}$ . That is, whenever the monopolist subject to a system of tax/subsidy as in (41) with  $s = 0$  covers the market, a social planner would equally do it, but not viceversa.

When  $s = 0$  and (52) holds, a positive subsidy is not needed to restore social optimality. When a positive subsidy is required for (52) to hold, the smallest subsidy is obtained by solving (52) w.r.to  $s$  when this holds as an equality. This gives  $s_{MT} = B'(0) - B(0) + g \frac{4+\sqrt{7}}{18}$ . This is optimal as long as full coverage is socially optimal, i.e.  $g \leq \frac{72 B(0)}{(1+\sqrt{7})^2}$ .

**Proof of Proposition 8.** Assume  $m = m_P \equiv (1 - \gamma)n_P$  and  $M = M_P \equiv (1 - \gamma)N_P$ . The proof of Proposition 6 applies whenever  $a = \gamma n_P$  and  $A = \gamma N_P$ .

**Proof of Proposition 9.** First, equalize the two FOCs (49) and (50). Then, substitute  $a = \gamma n_P$  and  $A = \gamma N_P$  and solve w.r.to  $M_{MAT}$ , and obtain

$$M_{MAT}(m_{MAT}) = \frac{1}{3}(1 - \gamma(1 - N_P + n_P)) + \frac{2}{3}n - \frac{1}{3} \left[ \gamma^2(N_P - n_P)^2 + (1 - \gamma)^2 \right. \quad (53)$$

$$\left. (1 - \gamma)(\gamma N_P + 2\gamma n_P - 2\gamma + 4m_{MAT} + 2) - 4\gamma m_{MAT}(N_P - n_P) + \right. \\ \left. + (1 - \gamma)(1 + \gamma(N_P - n_P) - 6m_{MAT}x - 3\gamma n_P x - \gamma - 2m_{MAT}) + 7m_{MAT}^2 \right] \frac{1}{2}$$

that gives the profit maximizing choice of  $M_{MAT}$  as a function of  $m_{MAT}$ . Recall that, at the social optimum, it must be that  $\frac{M}{m} = \frac{N_P}{n_P}$ . Solving this equality w.r.to  $M$  and substitute  $N_P = N_P(n_P)$  as in (9), it obtains  $M_{MAT}(m_{MAT}) = \frac{N_P(n_P)}{n_P} m_{MAT}$ . Given (9), for this equality to hold,  $x$  must be as follows

$$x_{MAT} \equiv \frac{1}{9} \frac{18n_P + \gamma \left( 2(n_P + 1)\sqrt{7n_P^2 - 4n_P + 4} - 10n_P^2 - 11n_P - 4 \right)}{n_P(2 - \gamma)}. \quad (54)$$

Using (54) in the maximization problem for the monopolist in (48), the implicit relationship between  $M_{MAT}$  and  $m_{MAT}$  becomes

$$M_{MAT}(m_{MAT}) = \frac{2}{3} + \frac{2}{3}p_L - \frac{1}{9}\gamma \left( 8 + 5p_L - \sqrt{7p_L^2 - 4p_L + 4} \right) + \quad (55) \\ - \frac{1}{3} \left[ 7p_L^2 - 4p_L + 4 - \frac{1}{9}\gamma \left( (48 - 20\gamma)(1 - p_L) + \right. \right. \\ \left. \left. + p_L^2(84 - 29\gamma) + 2(2 - p_L)(3 - 2\gamma)\sqrt{7p_L^2 - 4p_L + 4} \right) \right] \frac{1}{2}.$$

As to the subsidy  $s_{MAT}$ , first notice that it enters linearly both FOCs (49) and (50), and then, when exogenously set, does not affect (55).

In order to find the level of the subsidy that induces the monopolist to choose the social optimal mass of travelers, first plug  $x_{MAT}$  from (54) into the expression for profits of the monopolist in (48).

Next, focus on the case of interiors solutions both for the planner and for the monopolist. From the monopolist' problem in (48), take the FOC w.r.to  $m$  in (49) and substitute  $x = x_{MAT}$ . Denote the resulting expression as  $FOC_{MAT}^m(x_{MAT})$ . From the social planner problem in (6), denote the FOC w.r.to  $n$  - given in (7) - as  $FOC_P^n$ . Equalize these two expressions  $FOC_{MAT}^m(x_{MAT}) = FOC_P^n$  and, after substituting  $a = \gamma n_P$ ,  $A = \gamma N_P$ ,  $m = (1 - \gamma)n_P$  and  $M = (1 - \gamma)N_P$ , solve with

respect to  $s$ . This gives  $s_{MAT}$  as in the case of  $g \geq \frac{72B(0)}{(1+\sqrt{7})^2}$ . An identical result is obtained when using the FOCs from the monopolist and planner problems with respect to the mass of luxury travelers.

Focus now on the case of not interior solutions. Applying Kuhn-Tucker conditions, the monopolist subject to a system of tax/subsidy as in (51) fully covers the market (with a socially optimal degree of differentiation across lanes) when

$$g \leq \frac{36(B(0) - B'(0) + s)}{(1 + \sqrt{7})^2} - \gamma \frac{36((B(0) - B'(0) + s)(7 + 13\sqrt{7}))}{(1 + \sqrt{7})^2(11 + 5\sqrt{7} - g(7 + 13\sqrt{7}))}. \quad (56)$$

Notice that, when  $s = 0$ , (52) implies  $g \leq \frac{72B(0)}{(1+\sqrt{7})^2}$ . That is, whenever the monopolist subject to a scheme as in (51) with  $s = 0$  fully covers the market, a social planner would also do it, but not viceversa.

When  $s = 0$  and (56) holds, a positive subsidy is not needed to restore social optimality. When a positive subsidy is required for (56) to hold, the smallest subsidy is obtained by solving (56) w.r.to  $s$  when this holds as an equality, which gives  $s_{MAT} = B'(0) - B(0) + g \frac{11+5\sqrt{7}-\gamma(7+13\sqrt{7})}{18(1+\sqrt{7})}$ . This is optimal as long as full coverage is socially optimal, i.e.  $g \leq \frac{72B(0)}{(1+\sqrt{7})^2}$ .

Overall,  $s_{MAT}$  is as follows:

$$s_{MAT} \equiv \begin{cases} 0 & \text{if } g \leq g_1; \\ B'(0) - B(0) + g \frac{11+5\sqrt{7}-\gamma(7+13\sqrt{7})}{18(1+\sqrt{7})} & \text{if } g_1 \leq g \leq \frac{72B(0)}{(1+\sqrt{7})^2}; \\ B'(1 - n_P - N_P)(n_P + N_P) + \\ \quad + g(2N_P - n_P^2 - \frac{3}{2}N_P^2 + \\ \quad - \gamma N_P(2 - N_P)) & \text{if } g \geq \frac{72B(0)}{(1+\sqrt{7})^2}. \end{cases} \quad (57)$$

where  $g_1 \equiv \frac{36(B(0)-B'(0))}{(1+\sqrt{7})^2} - \gamma \frac{36(B(0)-B'(0))(7+13\sqrt{7})}{(1+\sqrt{7})^2(11+5\sqrt{7}-\gamma(7+13\sqrt{7}))}$ . ■