Congestion and Incentives in the Age of Driverless Cars

Federico Boffa* Alessandro Fedele† Alberto Iozzi‡

January 21, 2019

PRELIMINARY AND INCOMPLETE
PLEASE DO NOT CITE WITHOUT PERMISSION

Abstract

GPS systems and Autonomous Vehicles (AVs) will likely open the way to forms of traffic coordination, or centralization. We analyze the 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 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

*Free University of Bolzano and Collegio Carlo Alberto: federico.boffa@unibz.it.
†Free University of Bolzano: alessandro.fedele@unibz.it.
‡Università di Roma 'Tor Vergata' and SOAS University of London: alberto.iozzi@uniroma2.it.

1 Introduction

Technological advancement is rapidly changing the structure of the automobiles market and the organization of traffic. GPS systems and Autonomous Vehicles (AVs), that is, vehicles driven by a software that does not require human intervention, will open the way to forms of traffic coordination, or centralization.

This will entail a transition. In the current decentralized setting, drivers are essentially atomistic, and do not factor in their decisions the external effect they impose on fellow travelers. In the near future, instead, this internalization will (at least in part) occur, as a result of the presence of companies or organizations that will have an incentive to consider, in a fleet logic, the congestion externality imposed by their vehicles on the other vehicles they dispatch. The emergence of such companies may be favored by the reduced consumers' investment into private cars, following the incrased benefits from car sharing (Fagnant and Kockelman, 2015). The paper analyzes the welfare effects of such transition, from atomistic to centralized driving.

We consider a framework with a continuum of commuters use AVs to go from A to B. There is a road connecting point A to point B and this road is segmented into two separate parallel lanes. Both lanes are congested. Congestion will still be an issue in the presence of AVs, notwithstanding the beneficial effects on congestion of the improved coordination with AVs. There is in fact some consensus that AVs, by reducing the opportunity cost of spending time in traffic through the automation of driving, may increase the demand for mobility, as well as induce more travellers to use AVs rather than public transportation during already congested times. In the end, an increase in overall congestion with AVs is not unforeseeable.[1]

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. A commuter in the first best travels as long as the benefit from travelling she enjoys exceeds the aggregate congestion cost that she imposes. If the congestion cost is sufficiently large, efficiency requires to exclude some low-value travelers 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 do this in two contexts.

[1] Several practitioners believe this will be the case. See, e.g.: https://www.nytimes.com/2018/10/27/technology/driverless-cars-congestion.html; https://www.ptua.org.au/myths/robotcar/; https://electrek.co/2017/05/01/tesla-network-elon-musk-autonomous-ride-sharing-vision/

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-Hannemannn, 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 travelers 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, too few vehicles are dispatched while, at the same time, there is overdifferentiation in the congestion levels across the two lanes. 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). 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 a political economy perspective, such a subsidy may be particulary unpopular, even more than a traditional road pricing scheme.

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.

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,[2] but study how congestion 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.

2 The model

Utility and road network

There is a unit mass of heterogenous travelers, all with unit demand for a single trip from point A to B.

There is an individual road connecting A and B. This road is composed of two lanes that are ex-ante identical. However, the roads may potentially differ ex post in terms of the mass of travelers traveling in each, as a result of travelers' choices. When this is the case, we denote by n the lane where a larger mass of travelers travel, and by N the lane where the smallest mass of travelers travel. When the same mass of travelers travel on each lane, we have $n = N$. We assume that, at any $n, N > 0$, the roads are congested. Given that travelers dislike congestion, we define the road with less (N) travelers as the luxury lane, and the road with more (n) travelers are the popular lane.

[2]For instance, Jerry Walters, a principal consultant at Fehr & Peers, a US transportation 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).

Each traveler has an individual taste parameter θ, which we assume to be uniformly distributed in the $[0,1]$ interval.

An individual with a taste parameter θ travelling in a lane with a mass of fellow travellers equal to m derives an utility from the trip equal to $\mathcal{B}(\theta, m) = B(\theta) - \theta\alpha(m)$. Individual utility for consumer θ is composed of the gross benefit from traveling $B(\theta)$, which depends only on the individual type, and on the cost of congestion $-\theta\alpha(m)$. We assume $B(\theta)$ is increasing in θ, i.e. $B' > 0$, and $\alpha(m)$ is increasing in the mass of travellers in the same lane, m. We assume the relationship to be linear in m, so that $\alpha(m) = a\,m$, with $a > 0$. The disutility consumers draw from congestion differs across travelers. In particular, we assume that, for any m, the disutility is increasing in θ, so that commuters who value traveling more are also more affected by congestion costs. This is consistent with the evidence that points at the positive relation between wage and value of time (see, for instance, Small, 2012). Accordingly, θ represents both a parameter of horizontal differentiation in commuters' value of travel, and a parameter of disutility of congestion.

Throughout the rest of the paper, we make the following assumption:

Assumption 1. *For any $\theta \in [0,1]$:*

i) $\mathcal{B}(0, m) = B(0) > 0$*;*

ii) $\mathcal{B}(\theta, m) = B(\theta) - \theta a m$ *is increasing in* θ*.*

Assumption 1 ensures that all travelers have a positive individual net utility from traveling absent any price for traveling or tax. Part ii) of 1 imposes an additional regularity condition that simplifies the characterization of equilibria when travelers (or a fraction of them) are managed by a company.

We will sometimes use a linear demand function of the form:

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

Travelers' decision

When making the travel decision, a traveler located at θ chooses if and where (i.e. in which lane) to travel so as to satisfy the constraint that she derives a positive utility from travelling and that she choses the lane delivering the highest utility, net of fares or taxes possibly imposed on that lane. In other words, by denoting by $\tilde{n}$ and $\tilde{N}$ the anticipated total number (determined by the choice of all the other consumers) of travellers in the popular and luxury lane, respectively, and by $\phi \geq 0$ and $\Phi \geq 0$ the fares and/or taxes she pays when she travels in the popular and the luxury lane respectively, a traveler θ chooses to travel in the popular lane if

$$B(\theta) - \theta a\tilde{n} - \phi \geq 0, \tag{1}$$

$$B(\theta) - \theta a\tilde{n} - \phi \geq B(\theta) - \theta a\tilde{N} - \Phi. \tag{2}$$

Similarly, traveler θ chooses to travel in the luxury lane if

$$B(\theta) - \theta a\tilde{N} - \Phi \geq 0, \tag{3}$$

$$B(\theta) - \theta a\tilde{N} - \Phi \geq B(\theta) - \theta a\tilde{n} - \phi. \tag{4}$$

Because of Assumption 1, (4) implies that the IR constraint for the marginal traveler, (3), in the luxury lane holds too.

3 Social planner

In this Section we solve for the first best under our constraint of travelers being allocated across two lanes only. This amounts to characterizing the utilitarian welfare-maximizing allocation of consumers.

The perfectly foresighted social planner who can direct travelers in the two lanes solves the problem of finding which (if any) consumers should not travel, and then optimally allocating a mass of travelers n_P to the popular lane, and a mass of travelers N_P to the luxury lane.

$$\max_{\substack{n,N\in[0,1]\\ n+N\leq 1}} W = \int_{1-n-N}^{1-N} (B(\theta) - \theta\, a\, n)\, d\theta + \int_{1-N}^{1} (B(\theta) - \theta\, a\, N)\, d\theta. \tag{5}$$

When it is interior, it is implicitly defined by the following set of FOCs

$$B(1 - n_P - N_P) - 2an_P \left(1 - N_P - \frac{3}{4}n_P\right) = 0 \tag{6}$$

$$B(1 - n_P - N_P) + a\left(n_P^2 + \frac{3}{2}N_P^2 - 2N_P\right) = 0 \tag{7}$$

FOCs enable us to derive the optimal choice of n_P as a function of N_P[3]. This is given by

$$N_P(n_P) = \frac{1}{3}\left(2\,(1 + n_P) - \sqrt{7n_P^2 - 4n_P + 4}\right). \tag{8}$$

This relationship is illustrated in Figure 3 by the green increasing line.

We first ask under what conditions the planner finds it optimal to dispatch all travelers ($n_P + N_P = 1$), thereby covering the entire market, and how she allocates them in this circumstance.

[3]Even in the simple case of a linear benefit function $b(\theta)$, the explicit solutions of this system of equations are quite cumbersome and are omitted here.

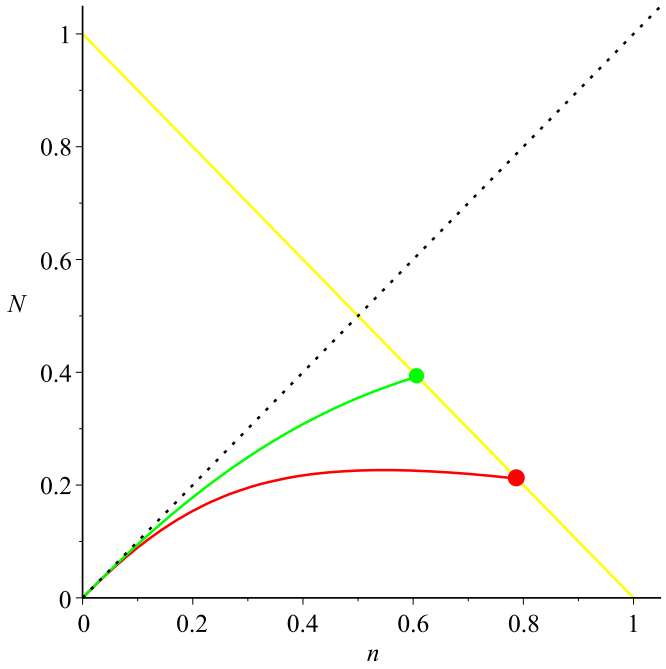


Figure 1: Implicit solutions (green line for the social planner, red line for the monopolist

Proposition 1. *Let $\bar{n}_P$ and $\bar{N}_P$ denote the solutions to problem (5) when there is full coverage. Then,*

$$\bar{n}_P = \frac{1}{6}\left(1+\sqrt{7}\right) \cong 0.6076; \tag{9}$$
$$\bar{N}_P = \frac{1}{6}\left(5-\sqrt{7}\right) \cong 0.3924.$$

These solutions occurs whenever

$$B(0) \geq a\,\frac{\left(1-\bar{N}_P\right)^2}{2} \cong a \times 0.1846. \tag{10}$$

The social planner covers the entire market when the benefit of serving the consumers with the lowest valuation of the trip is large enough and when the congestion cost of adding a traveller a is small enough. In particular, full coverage requires that the benefit from traveling of the traveler with the lowest valuation, $B\left(0\right)$, exceeds the aggregate congestion cost that this additional consumer located at $\theta = 0$ imposes on all the other travelers traveling in the popular lane. In this circumstance, it is worth noticing that the allocation of travelers across the lanes under full coverage is independent of the specific nature of the benefit function

$B(.)$. Indeed, all travelers enjoy the benefit $B(.)$, so welfare maximization simply requires the minimization of the aggregate congestion externality, which, in turn, does not depend on $B(.)$.

Along the green increasing line in 3, it is also possible to find the optimal solution to the problem (5) under full coverage. This is illustrated by the green solid circle located at the intersection of the implicit equilibrium relationship between n_P and N_P (i.e. $N_P(n_P)$ in (8)) and the constraint $n_P + N_P = 1$.

While we cannot fully characterize the welfare-maximizing allocation of travelers outside full coverage, implicit differentiation of the FOCs in (6) and (7) yields $\frac{\partial N_P}{\partial a}$, $\frac{\partial n_P}{\partial a} < 0$ for parameter values such that $n_P + N_P < 1$.This reflects the intuition that the larger is a traveller's disutility from congestion, the smaller is the social reward of a larger coverage of the market.

In the case of a linear benefit function (as in Assumption 2), we also 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. The larger is a traveller's benefit from travelling, both in its fixed component and in the component that is a function of the preference parameter θ, the larger is the social reward in covering a larger fraction of the market. These comparative statics exercises show that the planner's choice smoothly converges toward the full coverage solution as the relation between a and b_0 approaches that defined in condition (10).

4 Analysis of the model without 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-Hannemannn, 2002). Thus, understanding the welfare effects when neither road pricing nor other forms of taxes contingent on the mass of vehicles traveling can be imposed is a crucial starting point. This Section is devoted to this problem.

4.1 Atomistic travellers

We first study a fully decentralized economy populated by atomistic travelers only. This form of traffic organization prevails before the transition towards the centralization brought about by AVs start. There is no organized market for transportation services and all travelers make their individual decision on if and how to travel based on their individual preferences. We abstract from coordination issues, and we assume that technology allows all travelers, even when atomistic, to make their individually optimal choice given all the other travelers' decisions.

The next Proposition characterizes the equilibrium allocation of atomistic travelers.

Proposition 2. *Let n_A and N_A be the equilibrium mass of travellers in the popular and luxury lane, respectively. Then, it is always the case that $n_A = N_A = \frac{1}{2}$.*

Proof. First, note that Assumption 1.i implies that all travelers decide to travel when atomistic. Suppose n and $N < n$ AVs are already traveling in the two lanes. In this case, an additional traveler prefers to travel in the less congested lane since $b(\theta) - \theta aN > b(\theta) - \theta an$ for any given θ. This holds for any traveler as long as $N < n$ and implies that the equilibrium allocation is such that $N = n$. □

Proposition 2 simply says that, whenever all consumers are atomistic, everyone travels and the level of congestion in the two lanes is identical. Relatively to the social planner case, atomistic travelers cause two possible distortions. First, there is too little differentiation across lanes. Travelers with the higher disutility of congestion are forced to travel in a luxury lane which experiences the same level of congestion as the popular lane. Second, too many travelers may travel. This occurs when the social planner excludes some consumers from traveling. In choosing whether or not a consumer travels, the social planner internalizes the congestion cost imposed on other travelers, while atomistic travelers do not.

4.2 Coexistence of atomistic travellers and a monopolistic supplier

In this Section, we look at the transition period in the introduction of AVs. We assume that only a fraction of travelers travels independently. The remaining travelers do not own a car, and use the travel services of a firm operating as a monopolist, paying a fare accordingly. We refer to the atomistic travelers as *independent* and, by contrast, to those travelers using the services offered by the monopoly as *corporate* travelers. Independent travelers may travel without paying the fares imposed by the firms operating in the market and choose the lane they prefer. We let *independent* travelers choose if and in which lane to travel once *corporate* travelers have made their choices. In equilibrium, *corporate* travellers correctly anticipate the number of travellers in each lane, which clearly depends on the choices made by the *independent* consumers.

More formally, we assume that at any interval $[\theta, \theta + \epsilon]$, with $\theta \in [0, 1]$ and for any ϵ arbitrarily small, there is a fraction γ of *independent* consumers and a fraction $1 - \gamma$ of *corporate* consumers. We denote with n and N the mass of corporate travelers in the popular and luxury lane, respectively, and with n^* and N^* the mass of total (corporate and independent) consumers traveling in the popular and luxury lane, respectively.

Crucially, we assume that consumers do not select into car ownership based on their taste parameter θ (see Lamotte, De Palma and Geroliminis, 2016, for an illustration of the choice between conventional and autonomous vehicles).

We first establish the way in which corporate and independent travelers allocate themselves across the two lanes. The following Lemma directly follows from incentive compatibility

Lemma 1. *A necessary condition for independent travelers to travel in both lanes is that $n^* = N^*$. Instead, when $N^* < n^*$, all independent travelers travel in the luxury lane.*

Proof. An independent traveler located at θ travels in the luxury lane if and only if $b(\theta) - \theta aN \geq 0$ and $b(\theta) - \theta aN \geq b(\theta) - \theta an$, where the latter implies $N \leq n$. □

Two situations may potentially arise. Independent travelers may all travel in the luxury lane, or, alternatively, they may split themselves across the two lanes to equalize the total number of travelers in each lane.[4] Which of the two situations occurs depends on the relationship between the size of corporate and independent travelers. In particular,

$$\left.\begin{array}{c} n^* = n \\ N^* = N + \gamma \end{array}\right\} \textit{if } \gamma < n - N \tag{11}$$

and

$$n^* = N^* = \frac{n + N + \gamma}{2} \textit{ if } \gamma \geq n - N \tag{12}$$

Both cases are pretty simple to explain. The first case refers to a situation in which the mass of independent travelers is relatively small (i.e., $\gamma < n - N$). All independent travelers, after having observed the choices made by the corporate travelers, choose to travel in the luxury lane. Because of their small number, this lane maintains a congestion level lower than the popular lane. The second case refers instead to a situation in which the mass of independent travelers is relatively large (i.e., $n - N < \gamma$). Once the choice of the corporate travelers has been made, if all independent travelers choose the luxury lane, the congestion level of the luxury lane would exceed that of the popular lane. This, as explained above, cannot be an equilibrium. Hence, travelers split between the two lanes so as to equalize the level of congestion across them.

The presence of the independent travelers impacts the problem faced by their corporate counterpart. We assume that, when making their choices, they correctly anticipate the behavior of independent travelers, so that, in equilibrium, their expectations are fulfilled. The existence of the independent travelers matters for corporate travelers, as it contributes to determining the congestion level of each lane.

[4]Notice that the occurrence that all independent travelers travel in the popular lane cannot be an equilibrium, since there would always be an incentive for an independent traveller to move to the luxury lane and enjoy a lower level of congestion.

We first analyze the case of the fraction of independent travelers being large.

Large fringe. When the mass of atomistic travelers is large enough, the mass of total (independent and corporate) travelers is identical in both lanes, as determined by (12). This situation occurs whenever the fringe is larger than $\frac{1}{2}$. An equal number of travelers across the two lanes implies that the company can only charge the same fare in the two lanes. In turn, this implies that the IR constraints in (2) and (4) trivially holds and that the fare(s) is determined by the IR constraints (1) and (3) only. Denoting with f and F the fares charged to the corporate travelers in popular and the luxury lanes respectively, this implies

$$\begin{aligned} f &= F \\ &= B\left(1-N^*-n^*\right)-\left[1-N^*-n^*\right]an^* \\ &= B\left(1-N^*-n^*\right)-\left[1-N^*-n^*\right]aN^* \end{aligned} \tag{13}$$

Re-expressing these fares in terms of the monopolist choice variables n and N, we have

$$\begin{aligned} f &= F \\ &= B\left(1-\frac{N}{1-\gamma}-\frac{n}{1-\gamma}\right)-\left[1-\frac{N}{1-\gamma}-\frac{n}{1-\gamma}\right]a\frac{n}{1-\gamma} \\ &= B\left(1-\frac{N}{1-\gamma}-\frac{n}{1-\gamma}\right)-\left[1-\frac{N}{1-\gamma}-\frac{n}{1-\gamma}\right]a\frac{N}{1-\gamma}. \end{aligned} \tag{14}$$

Since $f=F$, it is clear that the monopolist's profits are affected by the total mass of corporate travelers only. Hence, denoting by $\breve{n}$ the total mass of corporate travelers (i.e. $\breve{n}=n+N$), the monopolist problem can then be written as

$$\max_{\breve{n}} \; f\breve{n} \tag{15}$$

At an interior solution, the solution to this problem, denoted by $\breve{n}_{LF}$, is implicitly defined by the following FOC

$$3a\breve{n}_{LF}^2-2a\breve{n}_{LF}\left[1-2\gamma\right]-2b'\left(1-\frac{\breve{n}}{1-\gamma}\right)\breve{n}_{LF}-\left[2b\left(1-\frac{\breve{n}_{LF}}{1-\gamma}\right)+a\gamma\right]\left[1-\gamma\right]=0 \tag{16}$$

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 is 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 prevent consumers with

a low value of travel from travelling. 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 it is true that the presence of a company (weakly) reduces congestion, the welfare effects of this reduction are ambiguous.

To formally support this argument, we first establish the condition for full coverage when the proportion of corporate travelers is small:

Lemma 2. *For any admissible shape of the benefit function $B\theta$), $\breve{n}_{LF} + \gamma = 1$ occurs whenever*

$$B(0) \geq B'(0) - \frac{1}{2}a. \tag{17}$$

If we compare this condition to the full coverage condition of the planner, we see that the larger is $b(0)$, the larger is the set of parameters for which full coverage occurs both under the planner and under this circumstance. To the contrary, when $b'(0)$ is small (large), full coverage becomes more (less) likely under this structure of supply, but nothing changes, instead, under the social planner. Finally, when a is large, full coverage is less likely under the social planner (which internalizes the congestion costs), but more likely under the currently analyzed structure of supply (because it makes the benefit function flatter). We therefore may have three cases. A first one in which the market is fully covered both under the planner and under the current supply structure. In this case, welfare remains unaltered with respect to the situation with atomistic travelers only. A second one in which both the planner and the small company with fringe do not fully cover the market. Here, the company's incentives to reduce the total mass of travelers are in line with those of the planner, hence welfare increases with respect to the case of pure atomistic travelers. Finally, when the benefit function is particularly steep, that is $B'(0)$ is large or a is small, the monopolist has an incentive to exclude some travelers, while the planner, as long as $B(0)$ is sufficiently large, does not. In this case, adding a small company to the atomistic travelers reduces welfare. We establish these results formally in the next Proposition.

Proposition 3. *Relatively to the case of atomistic travellers only, the welfare effects of a small company are as follows:*

i.) when full coverage occurs under the small company with fringe (i.e. $B(0) \geq B'(0) - \frac{a}{2}$), the existence of a small company does not alter social welfare;

ii.) when full coverage occurs with a social planner (i.e. $B(0) \geq a\,0.18$) but not under the small company with fringe (i.e. $B(0) < B'(0) - \frac{a}{2}$), the existence of a small company reduces social welfare since it reduces coverage;

iii.) when full coverage does not occur neither with a social planner nor under the small company with fringe ($B(0) < \min\{a\,0.18, B'(0) - \frac{a}{2}\}$), the existence of a small company increases social welfare since it reduces coverage.

Observe that, as the full coverage condition for the small company with fringe, does not depend on the size of the fringe (as long as γ is above 1/2), the result under i. remains true for all values of $\gamma > \frac{1}{2}$.

Small fringe. With a small fringe, the mass of travelers in each lanes are different. Hence, the monopolist can charge different fares, which however need to satisfy the incentive compatibility and individual rationality constrains in (1)-(4). constraints. Denoting with f and F such fares, and reexpressing the total number of travellers in each lane in terms of the corporate travellers in the popular and luxury lanes (i.e. n and N, respectively), the relevant constraints become

$$B\left(1-\frac{n}{1-\gamma}-\frac{N}{1-\gamma}\right)-\left[1-\frac{n}{1-\gamma}-\frac{N}{1-\gamma}\right]an-f\geq 0, \tag{18}$$

$$B\left(1-\frac{N}{1-\gamma}\right)-\left[1-\frac{N}{1-\gamma}\right]aN-F\geq B\left(1-\frac{N}{1-\gamma}\right)-\left[1-\frac{N}{1-\gamma}\right]an-f. \tag{19}$$

When these hold as an equality, we could solve them for the corporate fares

$$f = B\left(1-\frac{n}{1-\gamma}-\frac{N}{1-\gamma}\right)-\left[1-\frac{n}{1-\gamma}-\frac{N}{1-\gamma}\right]an, \tag{20}$$

$$F = f + a\left(1-\frac{N}{1-\gamma}\right)(n-N). \tag{21}$$

The problem faced by the monopolist can then be written as follows:

$$\max_{n,N}\ f\,n + F\,N \tag{22}$$

Since demand is now given as in (11), this problem may be rewritten as

$$\begin{aligned}\max_{n,N}\ &\left[B\left(1-\frac{n}{1-\gamma}-\frac{N}{1-\gamma}\right)-\left[1-\frac{n}{1-\gamma}-\frac{N}{1-\gamma}\right]an\right][n+N]+ \\ &+\left[1-\frac{N}{1-\gamma}\right]a\,[n-N-\gamma]\,N\end{aligned} \tag{23}$$

At an interior solution, the solutions to this problem, denoted by n_{SF} and N_{SF}, are implicitly defined by the following set of FOCs

$$\frac{1}{1-\gamma}\Bigg[B\left(1-\frac{n_{SF}}{1-\gamma}-\frac{N_{SF}}{1-\gamma}\right)[1-\gamma]-B'\left(1-\frac{n_{SF}}{1-\gamma}-\frac{N_{SF}}{1-\gamma}\right)[n_{SF}+N_{SF}]+ \\ -an_{SF}\left[2[1-\gamma]-4N_{SF}-3n_{SF}\right]\Bigg]=0 \tag{24}$$

$$\frac{1}{1-\gamma}\Bigg[B\left(1-\frac{n_{SF}}{1-\gamma}-\frac{N_{SF}}{1-\gamma}\right)[1-\gamma]-B'\left(1-\frac{n_{SF}}{1-\gamma}-\frac{N_{SF}}{1-\gamma}\right)[n_{SF}+N_{SF}]+ \\ -a\left[\gamma\left[1-\gamma\right]-4\gamma N_{SF}-3N_{SF}^2-2n_{SF}^2+2N_{SF}\right]\Bigg]=0 \tag{25}$$

The optimal choice of N_{SF} as a function of n_{SF}, obtained by equating the equations (24) and (25) and solving it w.r.to N_{SF}, is given by

$$N_{SF}(n_{SF})=\frac{1}{3}\left(1+2\left(n_{SF}-g\right)-\sqrt{7n_{SF}^2-2n_{SF}+1-\gamma(1-\gamma)-2\gamma n_{SF}}\right) \tag{26}$$

In the case of full coverage, we have that

Lemma 3. *Let $\bar{n}_{SF}$ and $\bar{N}_{SF}$ denote the solutions to problem (23) when there is full coverage. Then,*

$$\bar{n}_{SF}=\frac{1}{2}\left(1-\frac{2}{3}\gamma+\frac{1}{3}\sqrt{3+4\gamma^2-6\gamma}\right); \tag{27}$$
$$\bar{N}_{SF}=\frac{1}{2}\left(1-\frac{4}{3}\gamma-\frac{1}{3}\sqrt{3+4\gamma^2-6\gamma}\right).$$

For any admissible shape of the benefit function $B(\theta)$, these solutions occurs whenever

$$B(0)\geq B'(0)-\frac{1}{1-\gamma}\left(a\left(1-N_{SF}^2\right)-a\gamma\left(2-\gamma\right)\right). \tag{28}$$

4.3 Monopoly

Consider now a steady-state situation, in which all consumers are supplied by companies that manage AVs traffic. Suppose this company is monopolistic. While the monopolist can charge different prices for the two lanes, it is not allowed to price-discriminate across different consumers within the same lane. Furthermore,

differently than the social planner, she has to rely on prices only to direct travellers to the two lanes. This implies that fares in the two lanes have to be set in an incentive compatible way so as to induce the optimal allocation of travellers in the luxury and in the popular lane. Denoting by f and F the fares charged to travellers in the popular and luxury lanes respectively, fares must satisfy

$$B\left(1-n-N\right)-\left[1-n-N\right]an-f\geq 0, \tag{29}$$

$$B\left(1-N\right)-\left[1-N\right]aN-F\geq B\left(1-N\right)-\left[1-N\right]an-f. \tag{30}$$

Eqn. (29) is the individual rationality constraint for the marginal consumer traveling in the popular lane. Eqn. (30) is the incentive compatibility constraint of the marginal consumer traveling in the luxury lane. Because of Assumption 1, (30) implies that the IR constraint for the marginal traveler in the luxury lane holds too. Eqn. (30) also tells us that the fare in the popular lane f has to be sufficiently lower than the fare in the luxury lane F, so as to balance the lower quality (due to higher congestion) of the popular lane with a lower price, thereby separating the travelers. Those with a sufficiently low value for quality (i.e., a sufficiently low disutility from congestion) choose the popular lane, whereas those with a sufficiently high value for quality (i.e., a sufficiently high disutility from congestion) choose the luxury lane.

The monopolist problem can then be written as follows

$$\begin{aligned}&\max_{f,F}\ fn+FN \\ &s.t.\ \ (29)-(30)\end{aligned} \tag{31}$$

As it is standard, (29) holds as an equality at the solution of the monopolist's problem. Also, (30) holds as an equality, which has to be true as long as there is a strictly positive amount of travelers in each lane. Hence, the two constraints (29) and (30) may be rewritten as

$$f=b\left(1-n-N\right)-\left[1-n-N\right]an, \tag{32}$$

$$F=f+a\left(1-N\right)\left(n-N\right). \tag{33}$$

While the popular fare, f, is set to make the marginal traveller 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 travellers in the two lanes. On the one hand, a large difference in the mass of travellers across the two lanes, which can be interpreted as a large difference in quality across the two lanes, entails a larger extra-fee paid by the travelers in the luxury lane, hence a higher mark-up (i.e., a higher profit per traveller) in the luxury lane. On the other hand, since quality is inversely related to congestion, and hence to quantity,

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:

$$\max_{\substack{n,N\in[0,1]\\ n+N\leq 1}} \left(B\left(1-N-n\right)-\left(1-N-n\right)an\right)n+ \tag{34}$$
$$\left(B\left(1-N-n\right)-\left(1-N-n\right)an+a\left(1-N\right)\left(n-N\right)\right)N$$

At an interior solution, the solutions to problem (34), denoted by n_M and N_M, are implicitly defined by the following set of FOCs

$$\begin{aligned} B\left(1-n_M-N_M\right)-B'\left(1-n_M-N_M\right)\left(n_M+N_M\right)+\\ -an_M\left(2-4N_M-3n_M\right)=0 \end{aligned} \tag{35}$$

$$\begin{aligned} B\left(1-n_M-N_M\right)-B'\left(1-n_M-N_M\right)\left(n_M+N_M\right)+\\ +a\left(3N_M^2+2n_M^2-2N_M\right)=0 \end{aligned} \tag{36}$$

The optimal choice of N as a function of n, obtained by equating the equations (35) and (36) and solving it w.r.to N_M is given by

$$N_M(n_M)=\frac{1}{3}\left(1+2n_M-\sqrt{7n_M^2-2n_M+1}\right) \tag{37}$$

This non monotone relationship is illustrated in Figure 3 by the red line. One can grasp the intuition behind this non monotonicity result by noting that the benefit from differentiating the quality level (i.e., the congestion level) across the two lanes increases with n. When n is above 0.5469, the benefits from differentiation turn so large that an increase in n is associated to a decrease in N. The optimal solution to the monopolist's problem in (34) when there is full coverage is also illustrated in the Figure 3. This is given by the solid red circle at the intersection between the implicit relationship between n_M and N_M in (37) and the constraint $n_M+N_M=1$. In the case of full coverage, we have that

Lemma 4. *Let $\bar{n}_M$ and $\bar{N}_M$ denote the solutions to problem (34) when there is full coverage. Then,*

$$\begin{aligned} \bar{n}_M&=\frac{1}{2}\left(1+\frac{1}{3}\sqrt{3}\right)\cong 0.7887; \\ \bar{N}_M&=\frac{1}{2}\left(1-\frac{1}{3}\sqrt{3}\right)\cong 0.2113. \end{aligned} \tag{38}$$

For any admissible shape of the benefit function $B(\theta)$, these solutions occurs whenever

$$B(0)\geq B'(0)-a\left(1-\bar{N}_M^2\right)\cong B'(0)-a\times 0.9553. \tag{39}$$

The optimal choices of a monopolistic firm when it chooses to fully cover the market, as in the case of the social planner, are independent of the nature of the function $B(.)$, as long as it satisfies Assumption 1. More importantly, a comparison between these choices and those of the social planner reveals that the monopolist, when she chooses to fully cover the market, prefers a larger than optimal degree of quality (i.e. congestion) differentiation across lanes relatively to the social planner; in other words, less travellers travel in the luxury lane and more travellers 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 (33), allows to charge a higher luxury fare in the high quality market (with the higher mark up) the larger is 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).

The Lemma 4 also provides the condition ensuring that the monopolist prefers to cover the entire market.

Comparing 39 and 10, one finds that, for all the admissible parameter values, the social planner always fully covers the market when the monopolist does so, while the reverse does not hold true. This mirrors the standard intuition that a monopolist screens some consumers out of the market.

When the market is not fully covered, we can perform the same comparative statics exercises as in the case of the social planner's problem. By implicitly differentiating (35) and (36), we find that $\frac{\partial n_M}{\partial a}$ may be positive, in stark contrast with the social planner. This happens in particular when n_M is sufficiently large (i.e., $n \gtrsim 0.3812$), implying that it happens also in the neighborhood of full coverage. The positive effect of a on n_P derives from the combination of three effects, operating in different directions. Two effects contribute to a positive relation between a and n_M. First, when a when a gets larger, the aggregate demand curve becomes more rigid, as the higher θ consumers's WTP is increasingly compensated by their higher congestion cost (until the upper bound of a, determined by our Assumption 1. As a result, the monopolist reacts *à la* Ramsey, that is increasing the quantity supplied. An additional effect moving in the same direction of a positive correlation between a and n_M goes through the IC constraint; as it is clear from (33), the benefit on the fare F from differentiating the amount of travelers across the two lanes positively depends on a. The starker incentive to differentiate leads the monopolist to increase n_M, while at the same time reducing N_M, as a increases. Finally, a third effect pushes in the opposite direction. As the total number of travelers shrinks, f, the utility of the marginal consumer, is increasingly negatively affected by the level of a. As a result, this effect reduces the number of travelers as a increases. The latter effect prevails when n_M is sufficiently small, while the first two effects prevail when n_M is sufficiently large. The sign of $\frac{\partial N_M}{\partial a}$ may be

interpreted in the light of the non-monotone relation between N_M and n_M, due to the higher incentives to differentiate quality across lanes as n_M increases.

Turning to the analysis of $\frac{\partial n_M}{\partial b_0}$, we find that, as intuitive, an increase in b_0, which shifts each consumers' WTP upwards, also increases the number of travelers in the popular lane. The effect on the luxury lane is ambiguous, due to the non-monotone relation between n_M and N_M. Finally, $\frac{\partial n_M}{\partial b}$ is ambiguous, and driven by two countervailing effects. First, as b gets larger, the WTP becomes less similar across different θ, so the demand turns more elastic. This, ceteris paribus, decreases n_M. On the other hand, a larger b, as long as the monopolist does not fully cover the market, increases the marginal consumers' WTP, thereby increasing n_M. This effect gets larger as n_M turns lower. For low values of n_M, the latter effect prevails, and $\frac{\partial n_M}{\partial b} > 0$. For large values of n_M, the former effect prevails and $\frac{\partial n_M}{\partial b} < 0$. As usual, the shape of $\frac{\partial N_M}{\partial b}$ is also determined by the non-monotonic nature of $N_M(n_M)$.

4.4 Welfare effects without taxation

When not subject to a tax, private AVs fail to internalize congestion externality because they behave atomistically. This leads to overcongestion compared to the first best benchmark in that too many AVs travel, on the one hand, and to underdifferentiation across lanes, in that the two lanes contain the same amount of travellers. Instead, fleets are non-atomistic, hence they internalize congestion externality. This can lead to market undercoverage, on the one hand, and to overdifferentiation across lanes, on the other hand. At the beginning of the transition period, when a small fraction of consumers are managed by a company, while the majority are atomistic, welfare may increase as long as the congestion problem is sufficiently severe that the social planner would exclude consumers with a low travel benefit from traveling. At the end of the transition period, with a single company supplying to all the travelers, excess differentiation as well as a reduction in the mass of travelers emerge.

5 Taxation

We now move to the case of a tax authority being able to implement a road pricing policy. This Section characterizes the taxes that implement the first best allocation and discusses their properties.

5.1 Taxation with atomistic travellers

To characterize taxes that replicates the social optimum under purely decentralized atomistic travel, we first denote with t and T the unit tax levied on AVs traveling in the popular and luxury lane, respectively. We then observe that the equilibrium allocation described in eqns. (35) and (36) differs from the social optimum in two respects: (i) underdifferentation of congestion levels across lanes; (ii) full coverage always arising with atomistic travellers, while the market may not be fully covered in the first best best.

The set of unit taxes that restores social optimum is as follows:

Proposition 4. *When all travelers are atomistic, the pair(s) of taxes that replicate the social optimum is as follows*

$$\begin{cases} t & \leq b\left(0\right) \\ T & = t + \left(1 - \bar{N}_P\right) a \left(1 - 2\bar{N}_P\right) \end{cases} \quad \text{if } b\left(0\right) \geq \frac{\sqrt{7}+4}{36} a \tag{40}$$

$$\begin{cases} t & = an\left(1 - N - \frac{n}{2}\right) \\ T & = t + \left(1 - N_P\right) \times a \times \left(n_P - N_P\right) \end{cases} \quad \text{if } b\left(0\right) < \frac{\sqrt{7}+4}{36} a \tag{41}$$

Proof. When $b\left(0\right) \geq \frac{\sqrt{7}+4}{36} a$, full coverage occurs both under the social planner and with atomistic travelers. Note that the traveler located at $\theta = 0$ gets utiliy $b\left(0\right) - t$ from traveling in the popular lane. Given Assumption 1.*i)*, this means that full coverage occurs as in the social optimum. By contrast, if $t > b\left(0\right)$ some low-θ travellers do not travel here unlike at the social optimum. Given $t \leq b\left(0\right)$, T as in (40) ensure that the FOCs of the two problem are identical.

When $b\left(0\right) \leq \frac{\sqrt{7}+4}{36} a$, full coverage occurs under atomistic travelers only. The first equation in 41 ensures that the lowest-θ traveller is as in the social optimum. The second equation ensures that the allocation of AVs across the two lanes is as the FOC of the social planner problem when $n_P + N_P < 1$. □

Note that 40 allows for a large range of taxes, which include also 41.

When $b(0) \geq \frac{\sqrt{7}+4}{36} a$, full coverage occurs also under the social planner and the only distortion introduced by atomistic travelers is in the degree of differentiation across lanes. The optimal pair of taxes thus simply needs to make sure that, in the first place, all travelers travel; a sufficiently low t is enough for this, thus giving rise to the multiplicity of solutions to the problem. On the other hand, the luxury tax simply has to make sure that the IC constraint becomes binding for the traveler that gives rise to the socially optimal degree of differentiation. Notice that the indeterminacy of t allows both for solutions that maximize tax revenues, taxing both types of travelers, and for solutions, which, instead, reduce to zero the tax burden on those who travel in the popular lane, and which impose a positive tax on luxury travelers only.

When instead $b(0) < \frac{\sqrt{7}+4}{36}a$, the social planner excludes those with a low value for the travel from traveling. The pair of taxes ensures that this exclusion is mirrored and that the degree of differentiation across lanes is optimal.

Not surprisingly, with atomistic travelers, this result is obtained by a pair of congestion charges, equal to the external cost imposed by the marginal travelers by her decision to travel. When subject to the congestion charges, the marginal traveler considers the full social cost of her decision to travel when she takes it.

5.2 Taxation in pure monopoly

When traffic is centralized and travelers are supplied by a monopolist, the form of the taxes that restore social optimum dramatically changes. We investigate whether appropriately designed taxes on producers can restore the social optimum or, at least, mitigate inefficiency. In our setting, the monopolist suffers from 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. Second, it modifies the overall number of vehicles travelling. We look at the case of a lane-specific unit tax imposed on the monopolist. The monopolist perfectly internalizes the congestion externality, so the optimal tax is not a congestion charge. We consider, instead, a specific (as opposed to an ad valorem) tax/subsidy, potentially differing by lane (that is, as long as we are willing to use the quality interpretation, potentially differing by quality level). Denote with $t(n, N)$ and $T(n, N)$ the tax/subsidy to be imposed on traffic in the popular and luxury lane, respectively. The maximization problem the monopolist faces can then be stated as

$$\begin{aligned} &\max_{f,F} \ (f - t(n,N))\, n + (F - T(n,N))\, N \\ &s.t. \ \ (29) - (30) \end{aligned} \tag{42}$$

Let the solutions to this problem be denoted by n_{MT} and N_{MT}. Using (32) and (33) and assuming interior solutions, these are implicitly defined by the following set of FOCs

$$\begin{aligned} &B(1 - n_{MT} - N_{MT}) - B'(1 - n_{MT} - N_{MT})(n_{MT} + N_{MT}) + \\ &- a n_{MT}(2 - 4N_{MT} - 3n_{MT}) - t(n_{MT}, N_{MT}) + \\ &- \frac{\partial t(n_{MT}, N_{MT})}{\partial n_{MT}} n_{MT} - \frac{\partial T(n_{MT}, N_{MT})}{\partial n_{MT}} N_{MT} = 0 \end{aligned} \tag{43}$$

$$
\begin{aligned}
&B\left(1-n_{MT}-N_{MT}\right)-B'\left(1-n_{MT}-N_{MT}\right)\left(n_{MT}+N_{MT}\right)+ \\
&+a\left(3N_{MT}^2+2n_{MT}^2-2N_{MT}\right)-T\left(n_{MT},N_{MT}\right)+ \\
&-\frac{\partial T\left(n_{MT},N_{MT}\right)}{\partial N_{MT}}N_{MT}-\frac{\partial t\left(n_{MT},N_{MT}\right)}{\partial N_{MT}}n_{MT}=0
\end{aligned} \tag{44}
$$

We establish the following result:

Proposition 5. *Let*

$$
s_{MT}\equiv b'(1-n_P-N_P)(n_P+N_P)+2an_P\left(1-N_P-\frac{3}{4}n_P\right). \tag{45}
$$

A system of unit taxes of the following form

$$
\begin{aligned}
t(n,N)&=an-s_{MT} \\
T(n,N)&=aN-s_{MT}
\end{aligned} \tag{46}
$$

restores the social optimum under monopoly.

Proof. TO BE WRITTEN □

The Proposition illustrates that socially optimality is restored by imposing on the monopolist a system of unit taxes on each traveller, differentiated by the lane. Each tax consists of a proper tax, whose value increases with the total number of travellers in that lane, and a subsidy, exogenously determined and equal across the two lanes.

To understand the logic of this result, first notice that, by plugging the unit taxes given in (46), solving for the FOCs and then equalizing them to solve w.r.to N_{MT}, one obtains

$$
N_{MT}(n_{MT})=\frac{1}{3}\left(2\left(1+n_{MT}\right)-\sqrt{7n_{MT}^2-4n_{MT}+4}\right) \tag{47}
$$

which expresses the optimal N_{MT} as a function of n_{MT}. This condition reproduces exactly the equivalent condition (8) for the case of the social planner; this ensures that the taxes induce the monopolist to a socially optimal degree of differentiation across lanes. Why is this the case? The components an and aN of the tax provide an incentive to shift passengers from the popular to the luxury lane, as long as $n\neq N$. The marginal increase in profit for the monopolist from shifting a travelers across the two lanes is $a\left(n-N\right)$, which counteracts the incentives faced by the monopolist provided by his IC constraint. By eliminating the monopolist's incentive to overcrowd the popular lane in order to be able to apply an extracharge to those who travel in the luxury lane, the taxes restore the socially optimal degree

of differentiation. This tax could be regarded as an example of tax on quality, similar, for example, to Cremer and Thisse (1994).

An additional instruments is however needed to eliminate the distortion caused by the monopolist in screening travellers out to the market in a socially inefficient manner. Indeed, this inefficiency, already present in the case of the absence of taxes, it is only exacerbated by the fact that taxes impose an extra cost per traveller on the monopolist. To ensure a socially efficient number of travellers, the monopolist must be granted a subsidy on the total number of cars traveling, as in (45). Since the subsidy depends linearly on n and N, it does not affect the first order condition and thus the equilibrium differentiation for the monopolist, which is restored to optimality by the components an and aN.

5.3 Taxation in monopoly with atomistic travellers

TO BE WRITTEN

5.4 Welfare effects with taxation

When road pricing is possible, the welfare-maximizing allocation is restored through a conventional congestion charge when drivers are fully atomistic. When, instead, drivers are managed by a monopolistic company, the taxing scheme differs markedly. Reaching first best requires indeed a combination of a quality tax that restores the optimal allocation of travelers across lanes, and a subsidy to countervail the monopolist's incentive to screen some consumers out of the market. The issue of the extent to which a subsidy to a monopolist is popularly acceptbale warrants further discussion.

6 Conclusions

AVs are expanding very rapidly, as important companies are investing heavily in this market. 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 attempts to analyze the incentives involved and the welfare effects of the transition to a system of centralized traffic, managed by centralized companies.

We highlight the following tradeoff between decentralized versus (partially) centralized vehicle ownership. When not subject to a tax, private AVs fail to internalize congestion externality because they behave atomistically. This leads to overcongestion compared to the first best benchmark in that too many AVs travel, on the one hand, and to underdifferentiation across lanes, in that the two lanes contain the same amount of travellers. Instead, fleets are non-atomistic, hence they internalize congestion externality. This can lead to market undercoverage, on the one hand, and to overdifferentiation across lanes, on the other hand. At the beginning of the transition period, when a small fraction of consumers are managed by a company, while the majority are atomistic, welfare may increase as long as the congestion problem is sufficiently severe that the social planner would exclude consumers with a low travel benefit from traveling. In the absence of taxation, the benefits are unclear.

When road pricing is possible, the welfare-maximizing allocation is restored through a conventional congestion charge when drivers are fully atomistic. When, instead, drivers are managed by a monopolistic company, reaching first best requires a combination of a quality tax that restores the optimal allocation of travelers across lanes, and a subsidy to countervail the monopolist's incentive to screen some consumers out of the market.

References

[1] Arnott, R., de Palma, A., Lindsey, R. 1990. Economics of a bottleneck. Journal of Urban Economics, 27, 111-130.

[2] Basso, L. 2008. Airport deregulation: Effects on prices and capacity. International Journal of Industrial Organization, 26, 1015-1031.

[3] Basso, L., Zhang, A. 2007. Congestible facility rivalry in vertical structures. Journal of Urban Economics 61 218-237.

[4] Brueckner, J. K. 2002. Airport congestion when carriers have market power. American Economic Review 92 (5) 1357-1375.

[5] Brueckner, J. K. 2005. Internalization of airport congestion: A network analysis. International Journal of Industrial Organization 23(7-8) 599-614.

[6] Brueckner, J. K, Verhoef, E. 2010. Manipulable congestion tolls. Journal of Urban Economics 67 315-321.

[7] Cremer, H., Thisse, J. 1994. Commodity taxation in a differentiated oligopoly. International Economic Review, vol. 35, no. 3, 1994, pp. 613–633.

[8] Daniel, J. I. 1995. Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues. Econometrica 63(2) 327-370.

[9] 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.

[10] Lamotte, R., De Palma, A., Geroliminis, N. 2016. Sharing the road: the economics of autonomous vehicles. Elitisme Working Paper Series 2016-002.

[11] 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.

[12] Molnar, A. 2013 Congesting the commons: A test for strategic congestion externalities in the airline industry, mimeo.

[13] Mussa, M., Rosen, S. 1978. Monopoly and Product Quality. Journal of Economic Theory 18, 301-317.

[14] Oberholzer-Gee, F., Weck-Hannemann, H. 2002. Pricing road use: politico-economic and fairness considerations. Transportation Research, Part D, 7(2), 357-371.

[15] Ostrovsky, M., Schwarz, M.. 2018. Carpooling and the Economics of Self-Driving Cars, mimeo.

[16] Pels, E., Verhoef, E.. 2004. The economics of airport congestion pricing. Journal of Urban Economics 55 (2) 257-277.

[17] Rupp, N. G. 2009. Do carriers internalize congestion costs? Empirical evidence on the internalization question. Journal of Urban Economics. 65(1) 24-37.

[18] Small, K. 2012. Valuation of travel time. Economics of Transportation. 1, 2-14

[19] Van Dender, K. 2005. Duopoly prices under congested access. Journal of Regional Science. 45 (2) 343–362

[20] 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

[21] Vickrey, W.S. 1969. Congestion theory and transport investment. American Economic Review. 59 (2) 251-260.

[22] Zhang, A., Zhang, Y. 2006. Airport capacity and congestion when carriers have market power. Journal of Urban Economics 60 (2) 229-247.