

Congestion Externality and Autonomous Vehicles

Federico Boffa*(Free University of Bolzano)

Alessandro Fedele†(Free University of Bolzano)

Alberto Iozzi‡(University of Rome Tor Vergata)

VERY PRELIMINARY - EXTENDED ABSTRACT

September 9, 2018

Abstract

We analyze the market for autonomous vehicles (AVs). We show how different market structures and different rules for the available capacity utilization affect the internalization of congestion externality in a setting in which there is instant information on congestion and vehicles are part of fleets, therefore displaying a non-atomistic behavior. We also show how taxation should be implemented in such a setting in order to restore efficiency.

Keywords: autonomous vehicles; market structure; congestion externality

JEL Codes: R41; R11

1 Introduction

We analyze the market design for autonomous cars (ACs). ACs are self-driving vehicles, driven by a software that does not require human intervention. That of ACs is a relevant market for industrial organization, for at least three reasons. First, companies (Google, Uber, etc) are investing heavily in this market. Second, this may deeply affect the organization of the car transport market. Presumably, people will invest less in private cars, and will use companies (with a foreseeable convergence between taxi and car sharing services). Third, technology (GPS system) substantially improves information on traffic flows. This can potentially improve scheduling, and allow for coordinated forms of traffic centralization.

This change generates two important questions for IO. From a theoretical perspective, how do different market structures affect the extent of the internalization of congestion externalities? Second, from an applied perspective, how should an efficient and effective market design for autonomous cars look like?

*Email: federico.boffa@unibz.it

†Email: alessandro.fedele@unibz.it

‡Email: alberto.iozzi@uniroma2.it

2 Related Literature

There are two strands of literature relevant to our analysis. First, the traditional models of urban congestion, which treat travelers as atomistic. The standard framework is the bottleneck model (Vickrey, 1969, Arnott, de Palma and Lindsey, 1990), in which travelers have to choose when to leave. They face a tradeoff between length of the trip and early or late arrival. Efficiency in that context requires a dynamic toll. The toll takes the form of a traditional Pigouvian tax, equal to the marginal external cost (i.e., equal to the congestion externality) imposed on the other vehicles. Of course, then, before technology made traffic centralization envisageable, such a sophisticated toll was likely inapplicable. Several papers analyzed then the welfare properties of a variety of alternative coarse tolls, where efficiency was sacrificed in favor of applicability. To the best of our knowledge, there are only two papers that account for ACs in a bottleneck model. 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 ACs on road capacity.

Second, a recent, and somehow expanding literature on airport congestion. This literature points out the limited scope for congestion pricing at airports where carriers have market power, since carriers themselves internalize full (if monopolistic) or partial (if oligopolistic) congestion costs (Daniel, 1995, Brueckner, 2002). The result in Brueckner (2002) has been extended by Pels and Verhoef (2004) and Brueckner (2005) to the case of more realistic route structures, i.e., network structures with non-competing hub airports rather than simple route structures where a congested airport is connected to a single uncongested airport.¹ Their main finding is that the optimal congestion tolls are lower than what would be suggested by congestion costs alone. In particular, Brueckner (2005) show that the optimal tolls are decreasing in a carrier's airport flight share. The resulting reduced source of financing for airports can negatively affect their capacity investment and, in turn, future congestion; this issue is investigated by Zhang and Zhang (2006). Brueckner and Verhoef (2010) highlighted the following inconsistency in the above papers: large carriers do not take into account the impact of their own actions on the magnitude of congestion tolls. The authors address this problem by deriving manipulable toll rules, which are designed to achieve efficiency when agents fully anticipate the impact of their actions.

The empirical evidence on airport congestion is mixed. On one hand, Brueckner (2002), Mayer and Sinai (2003), and Rupp (2009) find that the most concentrated airports do not display the highest levels of delay; this result gives support to the idea that large airlines internalize the congestion they impose on themselves. On the other hand Daniel and Harback (2008) and Molnar (2013) find the opposite result due to strategic incentives. In particular, carriers can increase congestion externalities to deter entry. In particular, Molnar (2013) find that airlines schedule flights at their hub airports to create congestion at peak times and explain this evidence by arguing that the strategic benefits of entry deterrence outweigh the direct cost of congestion.

3 Results

We study the extent of congestion externalities - and hence of efficiency - in several market configurations, and we check how they differ vis-à-vis a welfare maximizer. We also study the welfare

¹Basso and Zhang (2007) explicitly consider competition between two airports and investigate its effect on capacities and congestion delays.

effects of allowing for multiple lanes with potentially different levels of congestion (hence, different levels of quality for the final users), under various licensing rules for the lanes, and, again, different market configurations. In particular, we analyze:

- Monopoly, with a single company owning (and managing) the AC traffic;
- duopoly, with two competing firms managing the AC traffic, with individual firms each managing a portion of the street, and each choosing its strategic variable (we have a version in which they choose quantity and one in which they choose price);
- duopoly, with two competing firms managing the AC traffic, with both firms using the same road, and each choosing its strategic variable (again, we have a version in which they choose quantity and one in which they choose price).
- We finally check the effects of a tax in such a framework.

Our preliminary **results** emphasize:

a) the determinants of overuse or underuse of the road in peak in the monopoly vis-à-vis the social planner;

b) the fact that, under a duopoly, congestion is internalized when individual firms each manage a portion of the street, and this produces results that are closer to social optimum with respect to the situation in which both firms share the same road (where there is excess use of the road in peak);

c) that, under some circumstances, duopoly with proprietary lanes is welfare superior to monopoly;

d) that an appropriate market design (in terms of lane property) with certain competitive structure may yield a welfare superior outcome with respect to a tax imposed on the monopolist.

The next **set of results** will look at taxes: their structure (and, of course, their level) under different market designs.

4 Modelling framework

Consider now a continuum of commuters using autonomous cars (ACs) to reach their workplace. Consumers are of different types, uniformly distributed in the $(0, 1)$ interval. Assume there is no congestion offpeak, while in peak there is congestion. Also, assume there are K lanes, with K being a finite number; each of these lanes, say lane k , has, in peak, its own level of congestion. A traveler of type θ has a utility of $B_O(\theta)$ if she travels offpeak, and of $B_P(\theta) - \theta t(n_k)$, where t , with $t' > 0$ and $t'' > 0$, is congestion cost, increasing and convex in the number of cars traveling in lane k , n_k . θ represents both a parameter of horizontal differentiation in consumers' willingness to travel onpeak and offpeak, and a parameter of preference for quality. Observe that our characterization of preferences implies that consumers that value traveling more are also more affected by congestion costs. An interpretation is that they travel for business and have a high opportunity cost of time. Assume single crossing condition, so that $B_O(\theta)$ and $B_P(\theta) - \theta t(n_k)$ cross only once. Conditions for this are: i) $B_O(0) > B_P(0)$; ii) $B_O(1) < B_P(0) - t(n_k)$; iii) a set of conditions on the curvature.

5 Sketch of the analysis

5.1 Social welfare

Social welfare function looks like:

$$W = \int_{\underline{\theta}}^{\theta_1} B_o(\theta) d\theta + \sum_{k=1}^K \int_{\theta_k}^{\theta_{k+1}} [B_p(\theta) - \theta t(n_k)] d\theta - c(1 - \underline{\theta}) - \sum_{k=1}^K n_k g(n_k) \quad (1)$$

If the number of lanes were physically unbounded or setting up new lanes would not cost additional money, optimal K would be infinity. In all other cases, K is finite; wlog, assume $K = 2$. Social welfare function then is

$$W = \int_{\underline{\theta}}^{\theta^*} B_o(\theta) d\theta + \int_{\theta^*}^{\theta^{**}} [B_p(\theta) - \theta t(\theta^{**} - \theta^*)] d\theta + \int_{\theta^{**}}^1 [B_p(\theta) - \theta t(1 - \theta^{**})] d\theta - c(1 - \underline{\theta}) - [\theta^{**} - \theta^*] g(\theta^{**} - \theta^*) - [1 - \theta^{**}] g(1 - \theta^{**}) \quad (2)$$

Let us first consider the social welfare function when the planner (for some reasons to be discussed) is **constrained** to have the same number of travellers in each lane. That amounts to imposing the additional constraint that $\frac{1+\theta^*}{2} = \theta^{**}$

FOCs with respect to θ^* are:

- θ^* :

$$\begin{aligned} \frac{\partial W}{\partial \theta^*} &= B_p(\theta_{SP}^*) - \theta_{SP}^* t\left(\frac{1 - \theta_{SP}^*}{2}\right) - B_o(\theta_{SP}^*) \\ &\quad - \frac{1 - (\theta_{SP}^*)^2}{4} t' \left(\frac{1 - \theta_{SP}^*}{2}\right) - \\ &\quad \left(g\left(\frac{1 - \theta_{SP}^*}{2}\right) + \frac{1 - \theta_{SP}^*}{2} g'\left(\frac{1 - \theta_{SP}^*}{2}\right)\right) = 0 \end{aligned} \quad (3)$$

We now move to an **unconstrained** social planner able to choose a different number of cars in the different lanes. We take the FOCs:

- θ^* :

$$\begin{aligned} \frac{\partial W}{\partial \theta^*} &= B_p(\theta_{PP}^*) - \theta_{PP}^* t(\theta_{PP}^{**} - \theta_{PP}^*) - B_o(\theta_{PP}^*) \\ &\quad - t'(\theta_{PP}^{**} - \theta_{PP}^*) \frac{(\theta_{PP}^{**})^2 - (\theta_{PP}^*)^2}{2} - \\ &\quad - (g(\theta_{PP}^{**} - \theta_{PP}^*) + [\theta_{PP}^{**} - \theta_{PP}^*] g'(\theta_{PP}^{**} - \theta_{PP}^*)) = 0 \end{aligned} \quad (4)$$

- θ^{**} :

$$\begin{aligned} \frac{\partial W}{\partial \theta^{**}} &= \theta_{PP}^{**} [t(\theta_{PP}^{**} - \theta_{PP}^*) - t(1 - \theta_{PP}^{**})] + \\ &\quad + t'(\theta_{PP}^{**} - \theta_{PP}^*) \frac{(\theta_{PP}^{**})^2 - (\theta_{PP}^*)^2}{2} - t'(1 - \theta_{PP}^{**}) \frac{1 - (\theta_{PP}^{**})^2}{2} - \\ &\quad - (g(1 - \theta_{PP}^{**}) + [1 - \theta_{PP}^{**}] g'(1 - \theta_{PP}^{**}) - g(\theta_{PP}^{**} - \theta_{PP}^*) - [\theta_{PP}^{**} - \theta_{PP}^*] g'(\theta_{PP}^{**} - \theta_{PP}^*)) = 0 \end{aligned} \quad (5)$$

5.2 Analysis

Result: $n_{PP}^{**} \neq n_{PP}^*$.

where the subscript PP denotes the unconstrained planner and n denotes the number of vehicles traveling in each of the two (* and **) lanes in peak.

Proof: By contradiction. Assume $n_{PP}^{**} = n_{PP}^*$. Then, $\theta_{PP}^{**} = \frac{1+\theta_{PP}^*}{2}$. Using this in (5) and simplifying, (5) becomes $-t_{PP}^* [\frac{1}{4}(1-\theta_{PP}^*)^2] = 0$, which is equal to 0 only when $t_{PP}^* = 0$ or $\theta_{PP}^* = 1$, a contradiction.

Result: $n_{PP}^{**} < n_{PP}^*$.

Proof: From the previous result, notice that, in (5), $LHS < RHS = 0$. Let dL and dR be the derivative of the LHS and RHS of (5) w.r.to θ_{PP}^{**} , respectively. The fact that $dL > 0$ and $dR < 0$ (shown below) proves the result.

$$\begin{aligned} dR &= \frac{\partial}{\partial \theta_{PP}^{**}} [g(1-\theta_{PP}^{**}) + [1-\theta_{PP}^{**}]g'(1-\theta_{PP}^{**}) - g_{PP}^{**} - \theta_{PP}^*] - [\theta_{PP}^{**} - \theta_{PP}^*]g'(\theta_{PP}^{**} - \theta_{PP}^*) \\ &= -g_{PP}^{**} - [1-\theta_{PP}^{**}]g'(1-\theta_{PP}^{**}) - [1-\theta_{PP}^{**}]g''(1-\theta_{PP}^{**}) \\ &\quad - g_{PP}^{**} - \theta_{PP}^* - g'(\theta_{PP}^{**} - \theta_{PP}^*) - [\theta_{PP}^{**} - \theta_{PP}^*]g''(\theta_{PP}^{**} - \theta_{PP}^*) < 0 \end{aligned}$$

$$\begin{aligned} dL &= \frac{\partial}{\partial \theta_{PP}^{**}} \left[\theta_{PP}^{**} [t(\theta_{PP}^{**} - \theta_{PP}^*) - t(1-\theta_{PP}^{**})] + t'(\theta_{PP}^{**} - \theta_{PP}^*) \frac{(\theta_{PP}^{**})^2 - (\theta_{PP}^*)^2}{2} - t'(1-\theta_{PP}^{**}) \frac{1 - (\theta_{PP}^{**})^2}{2} \right] \\ &= t(\theta_{PP}^{**} - \theta_{PP}^*) - t(1-\theta_{PP}^{**}) + \theta_{PP}^{**} [t_{PP}^{**} - \theta_{PP}^*] + t(1-\theta_{PP}^{**}) \\ &\quad + t''(\theta_{PP}^{**} - \theta_{PP}^*) \frac{(\theta_{PP}^{**})^2 - (\theta_{PP}^*)^2}{2} + t'(\theta_{PP}^{**} - \theta_{PP}^*) \theta_{PP}^{**} \\ &\quad + t''(1-\theta_{PP}^{**}) \frac{1 - (\theta_{PP}^{**})^2}{2} + t'(1-\theta_{PP}^{**}) \frac{1 - (\theta_{PP}^{**})^2}{2} \theta_{PP}^{**} > 0 \end{aligned}$$

We have shown that an unconstrained planner has more cars traveling in the peak lane designed for individuals with lower θ , and less cars traveling in the peak lane designed for individuals with higher θ , whose utility is more sensitive to congestion. A constrained social planner on which the condition $n^{**} = n^*$ is imposed increases n^{**} and reduces n^* with respect to an unconstrained social planner.

6 Monopoly

Notice first that a perfectly discriminating monopolist would split consumers equally across the k lanes, and he would generate no distortions with respect to first best.

Consider now the (more interesting, as it is more realistic) case of a non-discriminating monopolist.

Consider $k = 2$ lanes, and consider potentially two different levels of congestion in the 2 lanes. Denote:

- $\underline{\theta}$ the consumer indifferent between not travelling and travelling offpeak;
- θ^* the consumer indifferent between travelling offpeak and travelling in the high congestion (low quality) lane in peak;

- θ^{**} the consumer that is indifferent between travelling in the high congestion (low quality) lane and in the low congestion (high quality) lane in peak.

Let us look at the indifference conditions. The indifference condition that determines $\underline{\theta}$ is:

$$B_o(\underline{\theta}) - f_o = 0$$

The indifference condition that determines θ^* is:

$$B_o(\theta^*) - f_o = B_p(\theta^*) - \theta^* t(\theta^{**} - \theta^*) - f_l$$

The indifference condition that determines θ^{**} is:

$$\begin{aligned} B_p(\theta^{**}) - \theta^{**} t(\theta^{**} - \theta^*) - f_l &= B_p(\theta^{**}) - \theta^{**} t(1 - \theta^{**}) - f_h \\ f_h &= \theta^{**} (t(\theta^{**} - \theta^*) - t(1 - \theta^{**})) + f_l \end{aligned}$$

We therefore obtain:

$$f_o = B_o(\underline{\theta}) \tag{6}$$

$$f_l = B_p(\theta^*) - \theta^* t(\theta^{**} - \theta^*) - B_o(\theta^*) + B_o(\underline{\theta}) \tag{7}$$

$$f_h = \theta^{**} (t(\theta^{**} - \theta^*) - t(1 - \theta^{**})) + B_p(\theta^*) - \theta^* t(\theta^{**} - \theta^*) - B_o(\theta^*) + B_o(\underline{\theta}) \tag{8}$$

hence, $f_l < f_h$ requires $[1 - \theta^{**}] < [\theta^{**} - \theta^*]$.

A monopolist maximizes:

$$\pi = \int_{\underline{\theta}}^{\theta^*} f_o d\theta + \int_{\theta^*}^{\theta^{**}} f_l d\theta + \int_{\theta^{**}}^1 f_h d\theta - c[1 - \underline{\theta}] - [1 - \theta^{**}]g(1 - \theta^{**}) - [\theta^{**} - \theta^*]g(\theta^{**} - \theta^*)$$

After replacing, one gets:

$$\begin{aligned} \pi &= \int_{\underline{\theta}}^{\theta^*} B_o(\underline{\theta}) d\theta + \int_{\theta^*}^{\theta^{**}} (B_p(\theta^*) - \theta^* t(\theta^{**} - \theta^*) - B_o(\theta^*) + B_o(\underline{\theta})) d\theta \\ &+ \int_{\theta^{**}}^1 (\theta^{**} (t(\theta^{**} - \theta^*) - t(1 - \theta^{**})) + (B_p(\theta^*) - \theta^* t(\theta^{**} - \theta^*) - B_o(\theta^*) + B_o(\underline{\theta}))) d\theta \\ &- c[1 - \underline{\theta}] - [1 - \theta^{**}]g(1 - \theta^{**}) - [\theta^{**} - \theta^*]g(\theta^{**} - \theta^*) \end{aligned} \tag{9}$$

FOCs

- $\underline{\theta}$;

$$B_o(\underline{\theta}) - B'_o(\underline{\theta})[1 - \underline{\theta}] = c \tag{10}$$

less travel than socially optimal.

- θ^* :

$$\begin{aligned} &B_p(\theta_m^*) - B_o(\theta_m^*) \\ &- [B'_p(\theta_m^*) - B'_o(\theta_m^*)][1 - \theta_m^*] \\ &- \theta_m^* t(\theta_m^{**} - \theta_m^*) - \theta_m^* t'(\theta_m^{**} - \theta_m^*)[1 - \theta_m^*] \\ &+ t(\theta_m^{**} - \theta_m^*)[1 - \theta_m^*] + \theta_m^{**} t'(\theta_m^{**} - \theta_m^*)[1 - \theta_m^{**}] = \\ &+ g(\theta_m^{**} - \theta_m^*) + [\theta_m^{**} - \theta_m^*]g'(\theta_m^{**} - \theta_m^*) \end{aligned} \tag{11}$$

- θ^{**} :

$$\begin{aligned}
& \theta_m^* t' (\theta_m^{**} - \theta_m^*) [1 - \theta_m^*] \\
& - \theta_m^{**} t (1 - \theta_m^{**}) - \theta_m^{**} t' (1 - \theta_m^{**}) [1 - \theta_m^{**}] \\
& - t (\theta_m^{**} - \theta_m^*) [1 - \theta_m^{**}] - \theta_m^{**} t' (\theta_m^{**} - \theta_m^*) [1 - \theta_m^{**}] \\
& + t (1 - \theta_m^{**}) [1 - \theta_m^{**}] \\
& + \theta_m^{**} t (\theta_m^{**} - \theta_m^*) = \\
& g(1 - \theta_m^{**}) + [1 - \theta_m^{**}] g' (1 - \theta_m^{**}) - g(\theta_m^{**} - \theta_m^*) - [\theta_m^{**} - \theta_m^*] g' (\theta_m^{**} - \theta_m^*) \quad (12)
\end{aligned}$$

We can prove conditions under which monopoly yields overuse or underuse of the lines.

7 Next steps

We will then move to the:

- analysis of duopoly;
- introduction of a competitive fringe;
- analysis of the optimal tax in the different settings.

References

- [1] Arnott, de Palma, Lindsey. 1990. Economics of a Bottleneck. *Journal of Urban Economics*, 27, 111-130.
- [2] Basso, Zhang. 2007. Congestible facility rivalry in vertical structures. *Journal of Urban Economics* 61 218-237.
- [3] Brueckner, J. K. 2002. Airport congestion when carriers have market power. *American Economic Review* 92 (5) 1357-1375.
- [4] Brueckner, J. K. 2005. Internalization of airport congestion: A network analysis. *International Journal of Industrial Organization* 23(7-8) 599-614.
- [5] Brueckner, Verhoef. 2010. Manipulable congestion tolls. *Journal of Urban Economics* 67 315-321.
- [6] Daniel, J. I. 1995. Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues. *Econometrica* 63(2) 327-370.
- [7] Daniel, Harback. 2008. (When) Do hub airlines internalize their self-imposed congestion delays? *Journal of Urban Economics* 63(2), 583-612.
- [8] Lamotte, De Palma, Geroliminis. 2016. Sharing the road: the economics of autonomous vehicles *Elitisme Working Paper Series* 2016-002
- [9] Mayer, Sinai. 2003. Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays Are Evil. *American Economic Review*. 93(4) 1194-1215.

- [10] Molnar A. 2013 Congesting the commons: A test for strategic congestion externalities in the airline industry, mimeo.
- [11] Pels, Verhoef. 2004. The economics of airport congestion pricing. *Journal of Urban Economics* 55 (2) 257-277.
- [12] Rupp, N. G. 2009. Do carriers internalize congestion costs? Empirical evidence on the internalization question. *Journal of Urban Economics*. 65(1) 24-37.
- [13] van den Berg, Verhoef. 2016. Autonomous cars and dynamic bottleneck congestion: The effects on capacity, value of time and preference heterogeneity *Transportation Research Part B*. 94 43-60
- [14] Vickrey, W.S. 1969. Congestion theory and transport investment. *American Economic Review*. 59 (2) 251-260.
- [15] Zhang, Zhang. 2006. Airport capacity and congestion when carriers have market power. *Journal of Urban Economics* 60 (2) 229-247.