Congestion and Incentives in the Age of Driverless Cars

Federico Boffa\textsuperscript{1}  \hspace{1cm} Alessandro Fedele\textsuperscript{2}  \hspace{1cm} Alberto Iozzi\textsuperscript{3}

\textsuperscript{1,2} Free University of Bolzano  \\
\textsuperscript{3} Università di Roma Tor Vergata

AIEE, December 15, 2022
Motivation

Automation and artificial intelligence are rapidly changing the structure of the automobiles market and the organization of traffic

- Autonomous vehicles (AVs)
  - vehicles driven by a software that does not require human intervention (5 levels of automation; Level-4 AVs are just behind the corner)
- AVs entail relevant technological, legal, and moral issues (Awad et al. 2018)
- But significant, and often underappreciated, economic aspects as well.
- Widespread use of software for traffic management allows for greater traffic coordination, due to
  - reliable real-time information on traffic flows
  - opportunity for cars to make/change their travel plans contingent on other vehicles behavior
Motivation

Boffa, Fedele & Iozzi (AIEE 2022)

Congestion & Incentives in the Age of Driverless Cars
Centralization of urban traffic

- Robotaxis, i.e., taxis operated by ride service companies through AVs
  - Partnership Waymo-Lyft to offer robotaxi services in selected urban areas in the US
  - General Motor’s subsidiary Cruise expected to roll out its robotaxis fleet soon
  - Tesla will soon stop selling cars to private owners, and use them for - comparatively more profitable - robotaxis fleet instead

- Investment into private cars is bound to shrink (Fagnant and Kockelman, 2015)

- Fleets are cheaper (labor cost saving thanks to AVs) and with higher utilization rate

- Transition from private vehicles to fleets of AVs - centralization of urban traffic
Congestion externality

- Congestion big issue (congestion costs > 100 bln $ per year in US)
- Likely to remain important (even more so) when AVs will be in place
  - more rational use of infrastructure → higher road capacity → less congestion
  - travelers may spend more productively their time on vehicles → less congestion costs
  - induced demand for travel → more congestion
- When private (traditional or autonomous) vehicles are not organized in fleets, congestion externalities emerge: a driver driving on a congested road is contributing to increasing the congestion
  - this not only affects the driver, but it affects the other drivers as well
  - drivers are atomistic, and, unless they are taxed appropriately, they do not consider the extra cost that their decisions to drive in congested hours/places imposes on their fellow drivers
- When AVs are organized in fleets and centrally managed, they do not behave atomistically
  - each company managing a set of AVs has an incentive to consider the impact of congestion costs on its profit
  - depending on market structures, there may be incentives to (at least partially) internalize congestion externalities
What we do

- We analyze and the welfare effects of the transition from a decentralized market with atomistic travelers to a centralized market with a single company managing a fleet of vehicles, including the analysis of the transition period where AVs managed by a company will coexist with privately owned AVs.
  - It may seem intuitive that, as a result of the internalization of congestion under centralization, welfare increases. We show that this intuition is incomplete.
- We characterize optimal tax scheme throughout the centralization process.
  - We show that the optimal tax under decentralization—a congestion charge—differs markedly from the optimal tax under centralization—a tax to correct quality distortions.
Heterogeneity

- Heterogeneity in disutility from congestion costs is substantial (Small, 2012) and reflects heterogeneity in individuals’ value of time, as well as in value of reliability
  - Difference in value of time between the 75th and the 25th percentile is about half of the median value for the value of time (with the median being about 21 $ per hour, and the interquartile range being about 10 $ per hour)
  - Value of reliability varies between the third and the first quartile by more than 100% of the median value (with a median of approximately 12 $ per hour, and the interquartile range about 13 $ per hour)

- With heterogenous travelers, reduction in aggregate congestion costs (as well as welfare maximization) requires to act also on the efficient sorting of travelers
  - Consider a highway with two lanes. With heterogeneous disutility from congestion, efficiency may require to differentiate the speed across the two lanes, with a faster lane for fewer travelers that dislike congestion a lot, and a slower lane for the others
Sorting

- Efficient sorting is economically very relevant
- A little less than the average hourly disutility from congestion can be mitigated by efficiently sorting travelers, without any increase in infrastructure capacity or travelers rationing
  - AVs and other recent technological advancements are dramatically decreasing the cost of the sorting technology, thereby making it a viable and important alternative to manage congestion
- Our stylized framework will allow for both margins to mitigate congestion:
  - rationing
  - sorting
- Welfare effects of the transition depends crucially on what margin is dominant
Travelers and lanes

- An origin A and a destination B are connected by one road with *two segregated and congested* lanes
- Continuum of heterogeneous travelers, denoted by parameter $\theta \sim U[0, 1]$, using AVs to go from A to B
- Lanes ex ante identical and ex post potentially different in terms of congestion and in the price and/or taxes that are charged in each of them
- A traveler chooses
  - whether or not to travel
  - if she travels, in which of the two lanes to do so
- An outcome is an assignment of each type $\theta$ to a lane, or to not traveling
Travel options

- Net utility for a traveler with preference parameter $\theta$ when travelling in a lane with $n$ other travelers
  \[ B(\theta) - \theta gn \]

- $\theta$ affects the utility of traveling, $B(\theta) \geq 0$, with $B' > 0$, and the idiosyncratic component $\theta$ of the congestion disutility $\theta gn$
  - Consistent with evidence pointing to a positive relation between wage and value of time (see, e.g., Small, 2012)

- $g$, common component of the congestion disutility $\theta gn$, sufficiently low so that $\frac{\partial(B(\theta) - \theta gn)}{\partial \theta} > 0$ for any $\theta$ and $n$: utility is increasing in $\theta$
Utility
First best

- Social planner maximizes:

\[
W = \max_{s \geq 0, f \geq 0, s + f \leq 1} \int_{1-s-f}^{1-f} \left[ B(\theta) - \theta gs \right] d\theta + \int_{1-f}^{1} \left[ B(\theta) - \theta gf \right] d\theta
\]

- Notation:
  - \( s \): proportion of travelers traveling in slow lane
  - \( f \): proportion of travelers traveling in fast lane

- At the social optimum, there is differentiation across the 2 lanes

\[ s_{FB} > f_{FB} \]

- Travelers with very low \( \theta \) may not travel (when \( g \) high relative to \( B(0) \))
- Travelers with low \( \theta \) travel in the low quality/high congestion lane: slow lane
- Travelers with high \( \theta \in [1 - f_{FB}, 1] \) travel in the high quality/low congestion lane: fast lane
Sorting

- We first focus on the sorting problem, and analyze welfare in the transition between decentralized and centralized travel assuming the market is fully covered in all circumstances.
- This turns out to happen, given our parameters, when
  \[
  \frac{B'(0)}{B(0)} \leq 1
  \]
- \(g\) sufficiently low relative to \(B(0)\) so that the planner wants to dispatch all travelers.
- the value of traveling sufficiently flat so that a firm does not want to screen some consumers out of the market.
Fully decentralized market

- Market is populated with atomistic travelers only (= private cars)
- No internalization of congestion externality
- No charges in each lane (no fares, no taxes)
- Each traveler maximizes individual utility
- Features of the equilibrium
  - Atomistic travelers split equally in the two lanes (no matter their type), so that
    \[ s_A = f_A = \frac{1}{2} \]
- Distortion with respect to first best:
  - There is no differentiation across lanes (or underdifferentiation): fast lane is too slow, slow lane is too fast
Utility in decentralized market vs first best
Centralized market

- All travelers are corporate, $\mu = 1$
- They all use a single firm that manages a fleet
- Firm’s problem

$$\max_{s,f} ps + Pf$$

s.t. $p = B(0)$ (participation of type-0 traveler)

$$P = p + (1 - f) g(s - f)$$ (incentive compatibility of marginal traveler)

- firm uses fares to direct travelers across lanes
- it uses differentiation across lanes to extract value from high $\theta$-travelers

In equilibrium

- Firm differentiates more than social planner (overdifferentiation): fast lane is too fast ($f_M < f_{FB}$), slow lane is too slow ($s_M > s_{FB}$)
- Intuition: monopolist distorts quality
Utility
Welfare effects

- While, as the level of market centralization $\mu$ increases, the firm is increasingly able to internalize the congestion externality, this does not reflect into an increase in welfare. To the contrary, welfare goes down.

- Two distortions. One involves the mass of travelers traveling and the resulting level of differentiation across lanes:
  - Overdifferentiation with centralization more pronounced than underdifferentiation with decentralization

- A second one involves the identity of travelers, during the transition period

- With high $\mu$, all the atomistic travelers, including low $\theta$, travel in the fast lane. This misallocation reduces welfare severely.

- Distributional effects:
  - Centralization makes corporate travelers with high $\theta$, as well as travelers that remain atomistic, better off. Low $\theta$ corporate travels stand to lose from centralization.
Rationing

- Social planner rations when congestion cost $g$ is high relative to value for traveling for marginal type $B(0)$
- Firm rations when $B'(0)$ is large relative to $B(0)$
  - standard monopoly tradeoff
  - firm can screen too many or too few travelers out of the market with respect to social planner
- With fully decentralized market, all consumer always travel
- Effect of centralization ambiguous
  - scope for welfare improving centralization when $g$ is high relative to $B(0)$ - so that planner would ration - and $B'(0)$ large relative to $B(0)$ - so that firm would ration.
  - but parameters must ensure that firm is not rationing too much.
Welfare with rationing

- Welfare may improve with centralization when rationing occurs.
- This requires two circumstances:
  - Rationing under centralization occurring with a high enough $g$ (so that the planner would also ration)
  - Parameter values ensure not too much restriction ($B'(0)$ not too high)
Taxes with atomistic travelers

- We design first-best restoring taxes in the polar cases $\mu = 0$ (decentralized travel) and $\mu = 1$ (centralized travel, only corporate travelers).

- Government sets a per-vehicle tax equal to $t$ in the slow lane, and equal to $T$ in the fast lane:

$$t_{AT} \leq B(0)$$

$$T_{AT} = t_A + \frac{g}{18} \left(5 - \sqrt{7}\right)$$

- With atomistic travelers only, taxes that restore first best are congestion charges: each traveler internalizes the congestion externality she imposes on fellow travelers.

- Taxes are welfare-improving, but, in the absence of compensation, low $\theta$’s stand to lose because they travel in a more congested lane.
Taxes on monopoly

- **Timing**
  - first, tax authority announces a per-vehicle tax equal to \( t \) in the slow lane, and equal to \( T \) in the fast lane
  - second, firm sets fares \( p \) and \( P \)
  - third, travelers make their travel decisions

- **First-best-restoring tax scheme:**
  
  \[
  t = g_s \\
  T = g_f
  \]

- **Very different tax than that on atomistic travelers**
  - \( g_s \) and \( g_f \) restore first-best differentiation (i.e., first-best relation between \( s \) and \( f \))
  - since \( s > f \), tax is larger in the slow lane and discourage the monopolist to overcrowd it, thus reducing differentiation across lanes
  - a congestion charge, but a tax on quality (Cremer and Thisse, 1994)
Taxes with rationing

- First best restoring taxes under full centralization when rationing occurs in the equilibrium without taxes:

\[ t = gs - d \]
\[ T = gf - d \]

- where \( d \) is a subsidy.

- When congestion is sufficiently severe (\( g \) is large), subsidies exceed taxes.

- They would absorb funding from general taxation. Politically unappealing?
  - possibly, to be compensated with an ex ante license.
Conclusions

- We analyze welfare effect of transition from decentralized to centralized traveling under heterogenous travelers
- We show that they differ depending on whether the main issue is sorting or rationing
- In spite of a better ability to internalize congestion externality, centralization fails to be welfare-improving when the issue is rationing
- When the main problem is rationing (because the congestion costs are severe enough with respect to utility from traveling), centralization may be welfare-improving.
- Taxes that restore first best are starkly different in the decentralized and in the centralized world
- With full centralization, restoring first best may require a subsidy
Extensions

- Oligopoly competition
- Not only competition, but market design, for instance...
  - ...exclusive lanes
- Endogenous choice of owning the car
  - fares and, more in general, transport menus need to be incentive compatible across transport modes
- Acceptability
  - are we ready to surrender our individual decision making for a public good (such as the reduction of aggregate congestion disutility)

THANKS
federico.boffa@unibz.it