Congestion and Incentives in the Age of Driverless Cars

Federico Boffa
Free University of Bolzano and Collegio Carlo Alberto

Alessandro Fedele
Free University of Bolzano

Alberto Iozzi
Università di Roma "Tor Vergata" and SOAS, University of London

NBER Economics of Autonomous and Electric Vehicles Meeting
June 6-7, 2019

Motivation

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

- 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
- Larger incentives to use car sharing services will lead to **greater centralization of urban traffic** (mobility as a service)
  - traffic will likely be managed by few companies
  - companies will manage their cars with a fleet logic
- In spite of large debate around AVs, little is known on their impact through the changes in the organization of the mobility market that they will induce

Congestion externality

- With traditional vehicles, 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
    - ★ this induces overcongestion, above the socially optimal level (typically above zero anyways)
- When vehicles 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

The paper in a nutshell

- We analyze the welfare effect of the transition from a market with atomistic users to one managed by profit-maximizing companies running fleets of AVs in an environment without road charges
- We characterize optimal tax schemes during the transition and in the long-run, where the entire traffic will be managed by fleets
- In our setting both atomistic drivers and fleet-managing companies have access to softwares that perfectly predict traffic, to match the upcoming technological status
  - difference in incentives induced by centralization only (and by taxation schemes, when they are implemented)

Users and lanes

- An origin A and a destination B are connected by one road with **two** (segregated) lanes
- Continuum of users using vehicles to go from A to B
- Heterogeneous users $\theta \sim U[0,1]$
- 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 user chooses
  - whether or not to travel
  - if she travels, in which of the two lanes to do so

Travel options

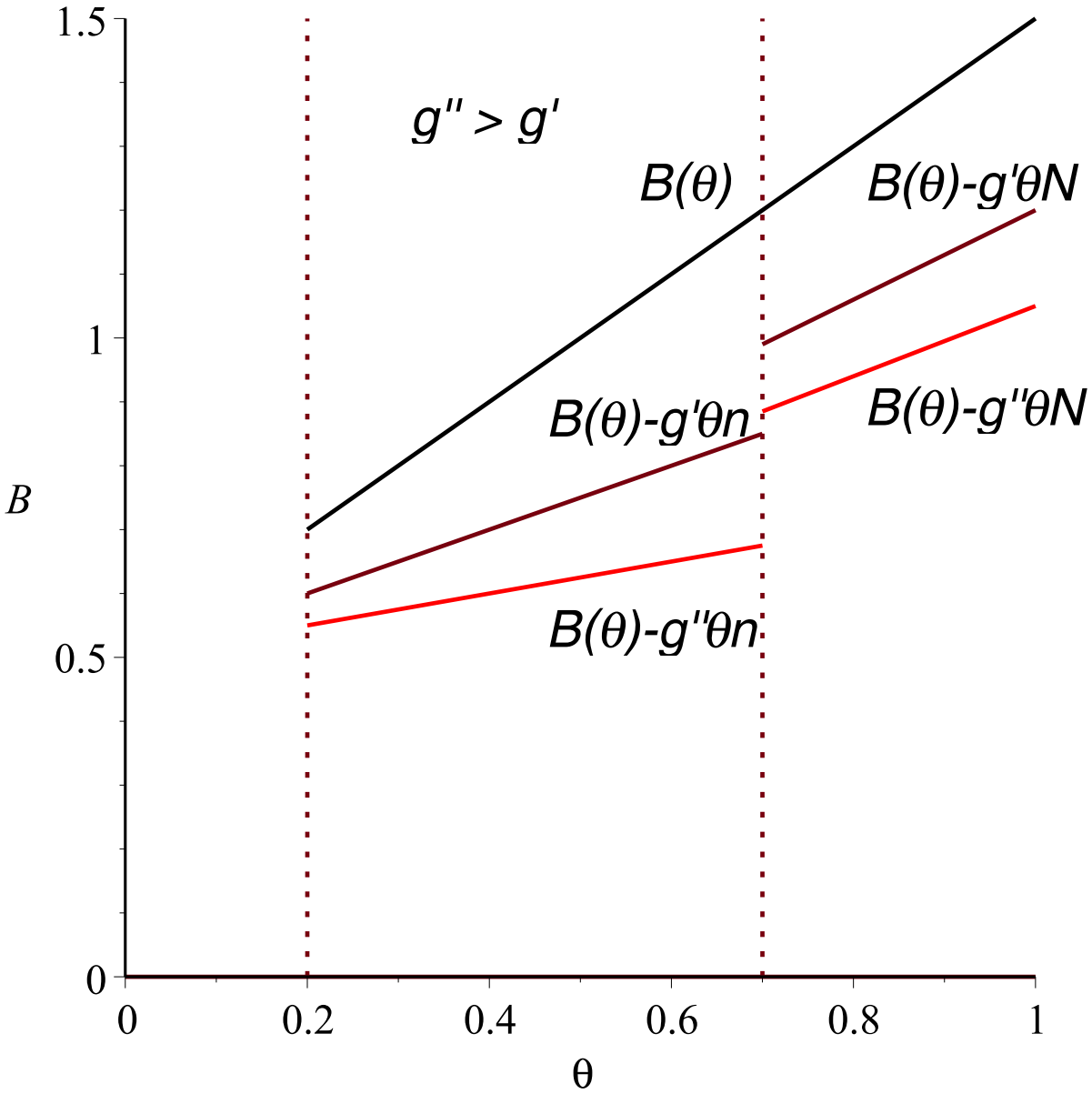
- Utility for a user with preference parameter θ when traveling in a lane with n other travelers

$$\mathcal{B}(\theta, n) = B(\theta) - \theta g n$$

with $\mathcal{B}(0) \geq 0$ and $\frac{\partial \mathcal{B}(\theta,n)}{\partial \theta} > 0$ (all figures using a linear specification for $B(\theta)$)

- θ is parameter of vertical differentiation, representing the value of time, or the disutility of congestion
- θ positively affects both the utility of traveling, $B(\theta)$, and the disutility from congestion, $\theta g n$
  - Consistent with evidence pointing to a positive relation between wage and value of time (see, e.g., Small, 2012)
- Low (high) congestion may be interpreted as high (low) quality

Utility



Preview of the main results

With no taxes

- at the **social optimum**
  - differentiation across lanes, but
  - some users may not travel
- when all users are **atomistic**, all (too many) users travel, and no (too little) differentiation across lanes
- when a **small company with a fleet logic** emerges, welfare increases only when the congestion problem is severe enough
- if the entire market is managed by a **single monopoly**
  - too much differentiation across lanes
  - (weakly) less users than under atomistic travel, but possibly too few

With taxes

-
  - first best with atomistic drivers requires conventional congestion charges
  - first best with a single company requires a very different tax/subsidy scheme. When congestion is severe enough, this scheme involves a net subsidy for the monopolist

Related literature - AVs

- Economics of carpooling: optimal tolling and optimal sharing of tolls across car poolers (Ostrovsky and Schwarz, 2018)
  - recognizes that carpooling is an essential feature associated to AVs
  - however, rules out heterogeneity in consumers' value of time
  - assumes that welfare maximization involves maximizing throughput – an assumption that works reasonably well for highways traffic, but less so for the urban traffic cases we are interested in
- Allocation of road infrastructures across conventional vehicles and AVs (Lamotte, De Palma and Geroliminis, 2016)
  - analysis of optimal tolling under coexistence of the two types of vehicles
  - analyze how road infrastructures should be allocated to conventional vehicles and AVs, and how tolls should be set when the two types of vehicles co-exist
  - AVs follow the prescription of a welfare-maximizing system operator on the departure time
  - however, travelers' choice of whether or not to follow the system operator's prescriptions left unmodeled
- Impact of AVs on road capacity (Van den Berg and Verhoef, 2016)

Related literature - urban congestion

- Bottleneck models (Vickrey, 1969; Arnott, de Palma and Lindsey, 1990)
  - congestion with atomistic travelers
  - no heterogeneity across travelers (at least in the baseline version)
  - all travelers have the same desired arrival time
  - all travelers in equilibrium have the same utility
  - one consumer arrives right on time, but face a lot of congestion
  - the others trade off less congestion with arriving early/late
  - a time-varying optimal toll achieves the desired level of congestion
- Value pricing and optimal differentiation across lanes (Hall, 2017)
  - fast lanes generate a Pareto improvement under hypercongestion (that is, a level of congestion that reduces throughput)

Related literature - airport congestion and more

- Airport congestion with non atomistic carriers with market power (Brueckner, 2002)
  - carriers internalize (fully if monopolistic, partially if oligopolistic) congestion costs
  - less scope for congestion charges
  - no consumers heterogeneity in their framework
- Quality level under multiproduct monopoly and duopoly and taxes (Mussa and Rosen, 1978; Cremer and Thisse, 1994; Lambertini and Mosca, 1999)

First best I

- Social planner maximizes:

$$\max_{\substack{n,N\in[0,1]\\ n+N\leq 1}} W = \int_{1-N-n}^{1-N} \left[B(\theta) - \theta g n\right] d\theta + \int_{1-N}^{1} \left[B(\theta) - \theta g N\right] d\theta$$

- Notation:
  - N : # of users traveling in the "low congestion/high quality" or "luxury" lane
  - n: # of users traveling in the "high congestion/low quality" or "popular" lane
- At the social optimum, a social planner differentiates the number of cars across the 2 lanes

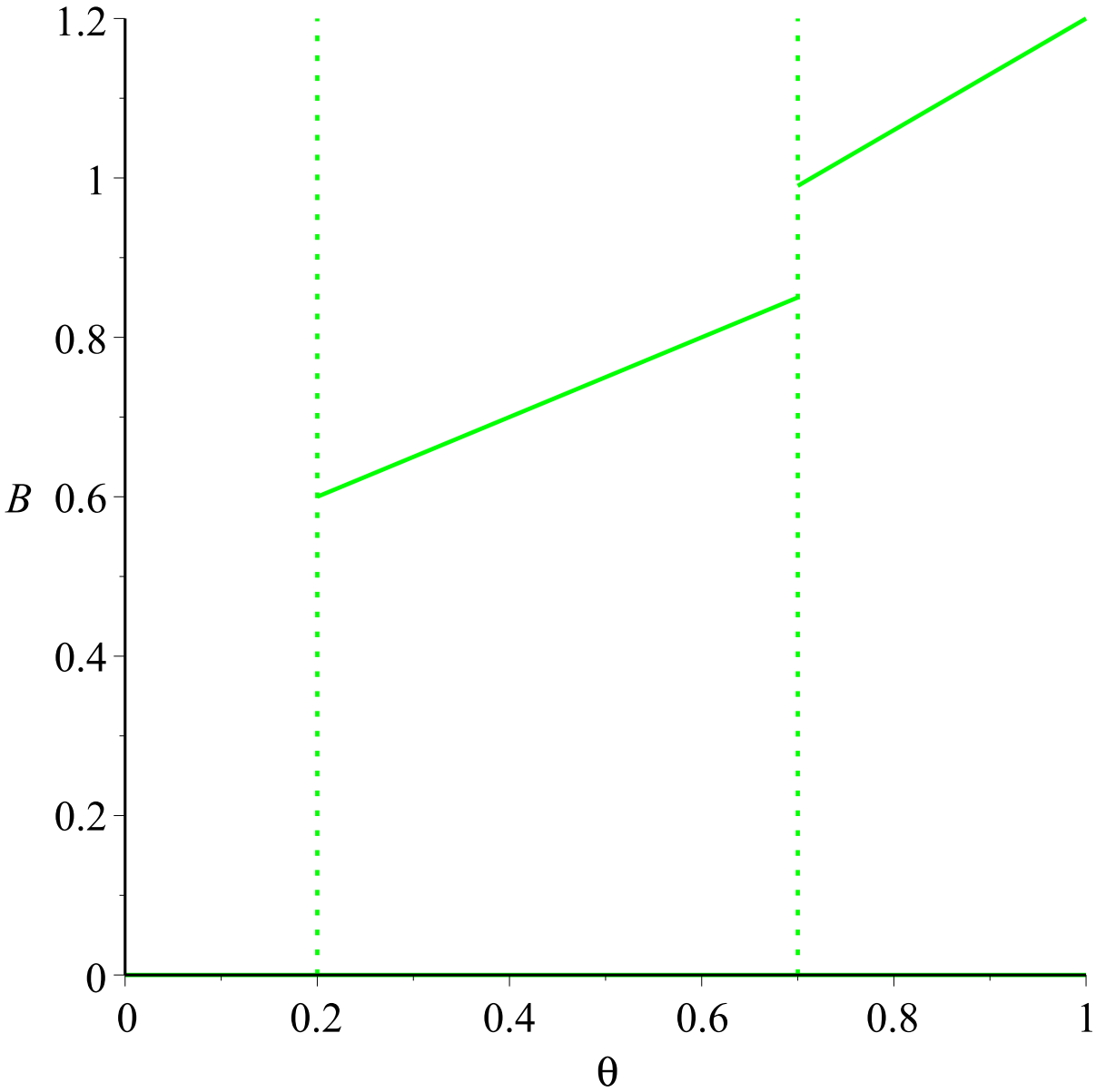
$$N_P < n_P$$

  - intuition: high θ have a higher disutility from traveling in a congested lane

First best II

- Users with **low** θ may not travel (when g low relative to $B(0)$)
- Users with **intermediate** θ travel in the low quality/high congestion lane
- Users with **high** θ travel in the high quality/low congestion lane
- Comparative statics: $\frac{\partial N_P}{\partial g}, \frac{\partial n_P}{\partial g} \leq 0$
  - Intuition: the larger the cost of congestion, the smaller the number of users in both lanes

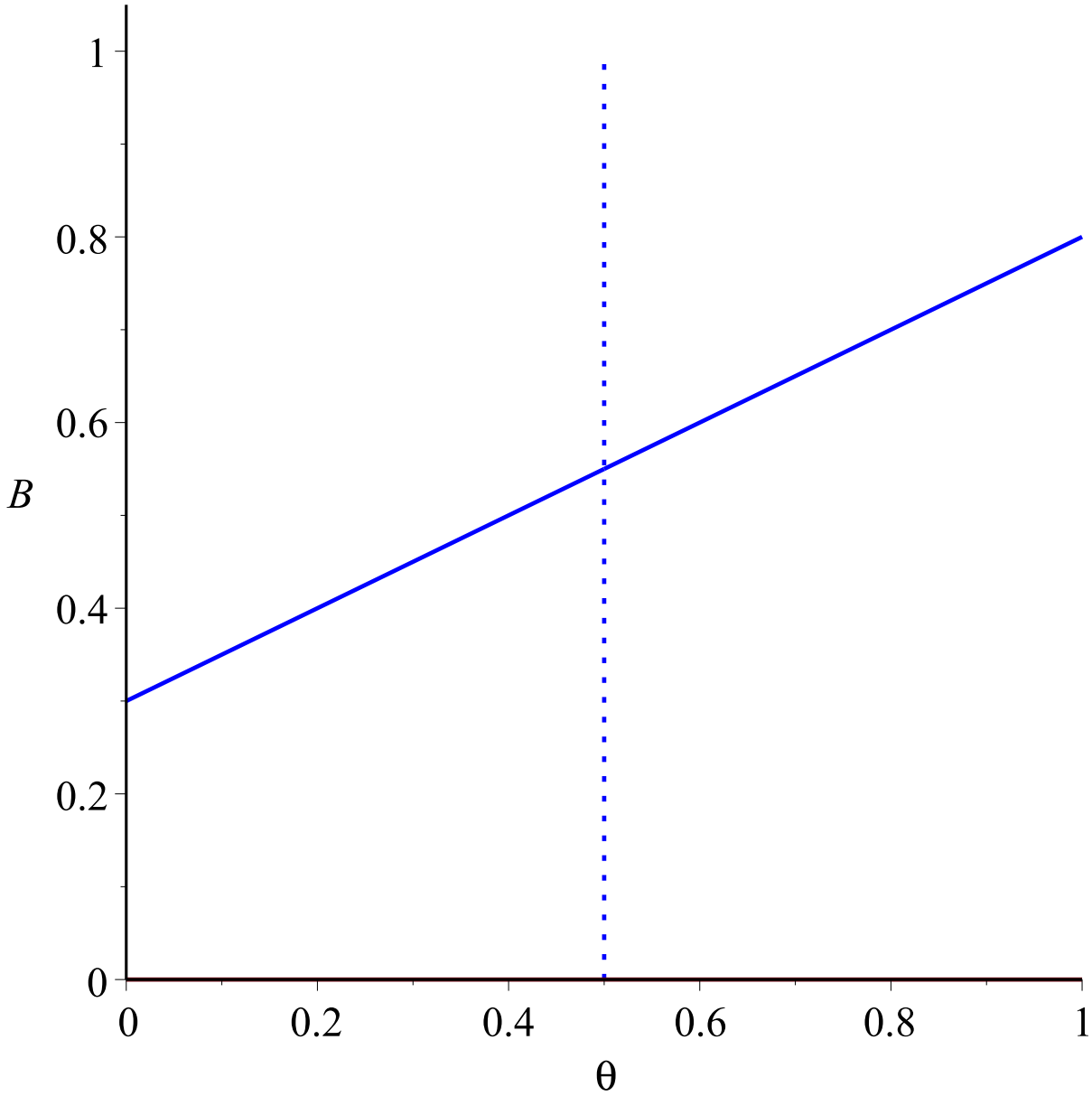
Utility: first best



Atomistic users only

- Market is populated with atomistic users only
- No charges in each lane (no fares, no taxes)
- Each user maximizes individual utility
- Features of the equilibrium
  - all users travel (as $B(0) \geq 0$ and $\frac{\partial \mathcal{B}(\theta)}{\partial \theta} > 0$)
  - they split equally in the two lanes, so that $n_A = N_A = \frac{1}{2}$
- Two types of distortion with respect to first best
  - there may be **excessive travel** (when the planner does not fully cover the entire market, that is, when congestion is a sufficiently severe problem to warrant traffic reduction at the optimum)
  - there is **no differentiation** across lanes
- No internalization of the congestion externality

Utility: atomistic travelers



The emergence of a monopolist

- A (large) share γ of users are atomistic and a (small) share $1-\gamma$ uses vehicles belonging to a company's fleet
- Company is effectively a monopolist on the $1-\gamma$ users
- No selection into ownership based on θ
  - For any finite partition of the θ space, there is an exogenously given proportion γ of *atomistic* commuters and a proportion $1-\gamma$ of *corporate* users
- Timeline of the game
  - monopolist sets two fares for corporate users: f for the popular lane, and F for the luxury lane. No price discrimination within lanes
  - simultaneously, all users, corporate and atomistic, choose if and where to travel

The emergence of a monopolist: equilibrium

- In equilibrium
  - mass of atomistic users large enough to bridge the congestion gap between corporate commuters across the two lanes
  - equally congested lanes (so $f = F$)
  - some corporate users may not travel
  - this happens when g small, i.e., $g < 2\left(B'\left(0\right) - B\left(0\right)\right)$
    - ★ stark difference with first best
- Monopolists' incentives to screen consumers out of the market not aligned with planner. Monopolist might restrict too much
  - When g is **low** relative to the **level** of willingness to pay, so congestion is relatively not severe, welfare is **reduced** by the presence of the monopolist
  - When g is **high** relative to the **level** of willingness to pay, so congestion is relatively more severe, welfare is **increased** by the presence of monopolist (as long as the monopolist does not restrict demand too much)

Monopoly only

- Timeline
  - monopolist sets two fares, f for the popular lane, and F for the luxury lane. No price discrimination
  - users choose if and where to travel
- Fares subject to standard IR and IC constraints under asymmetric info
- Monopolists' problem

$$\max_{f,F} fn + FN$$

$$\text{s.t. } f = b(1-n-N) - (1-n-N)gn \qquad \text{(IR-L)}$$

$$F = f + g(1-N)(n-N) \qquad \text{(IC-H)}$$

  - monopolist uses differentiation across lanes to extract value from high $\theta-$users
  - an increase in g decreases users' heterogeneity in willingness to pay

Equilibrium under monopoly

In equilibrium

- Monopolist **differentiates more** than social planner $\longrightarrow$ effect of IC-H
- **More or less people may travel** under monopoly vis-a-vis the social optimum
  - market is fully covered when g is large relative to difference between $B'(0)$ and $B(0)$
  - as g increases, the range of parameters for which full coverage occurs increases
    - ★ users' willingness to pay less heterogenous, so less incentives for monopolist to restrict output

Summarizing welfare analysis without taxes

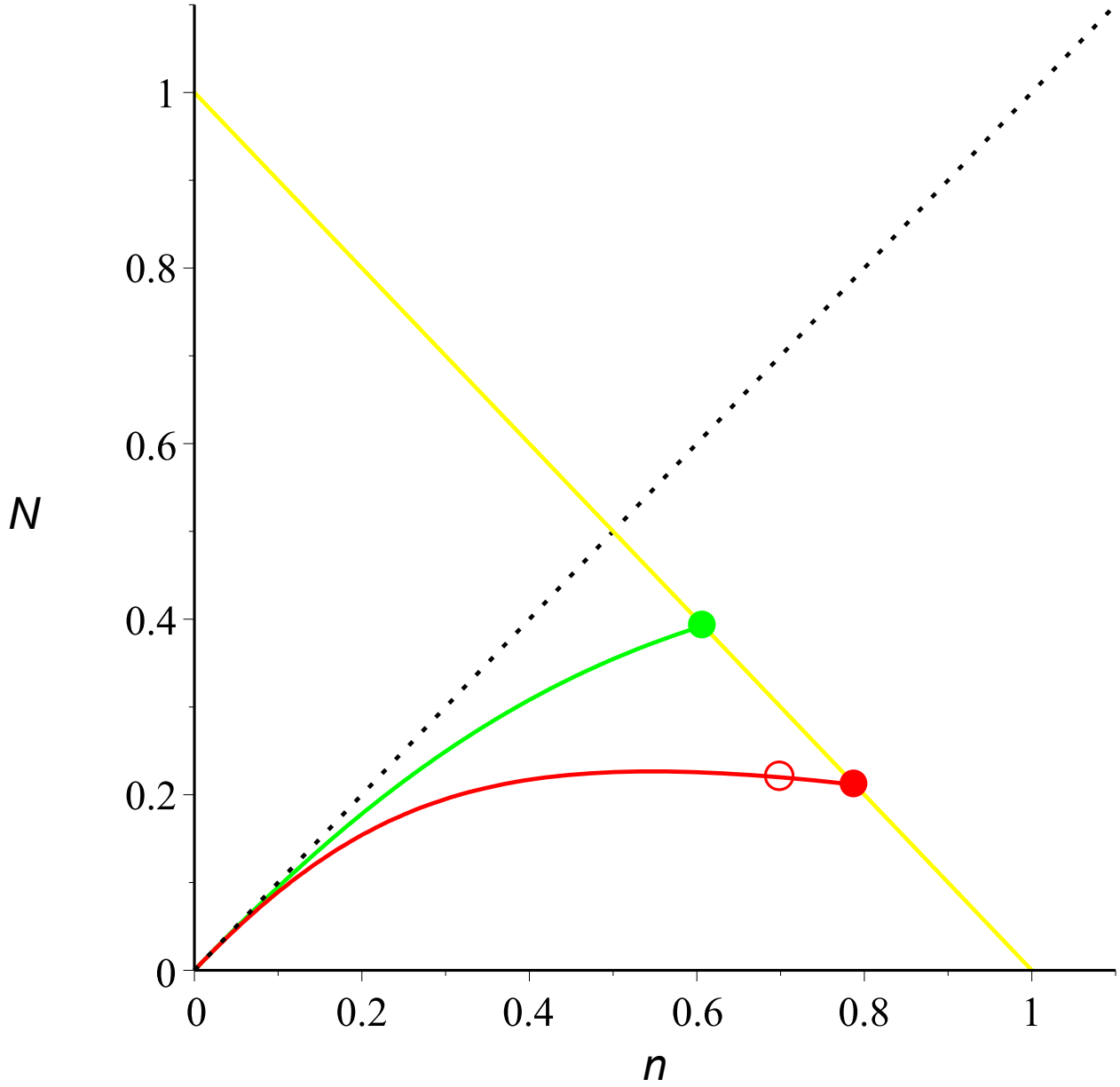
In the transition from atomistic to centralized travel, welfare changes due to

- change in differentiation across lanes:
  - moving from underdifferentiation with atomistic to overdifferentiation with monopoly
- change in the total number of vehicles on the road
  - total number of vehicles (weakly) reduced in the transition from atomistic to centralized

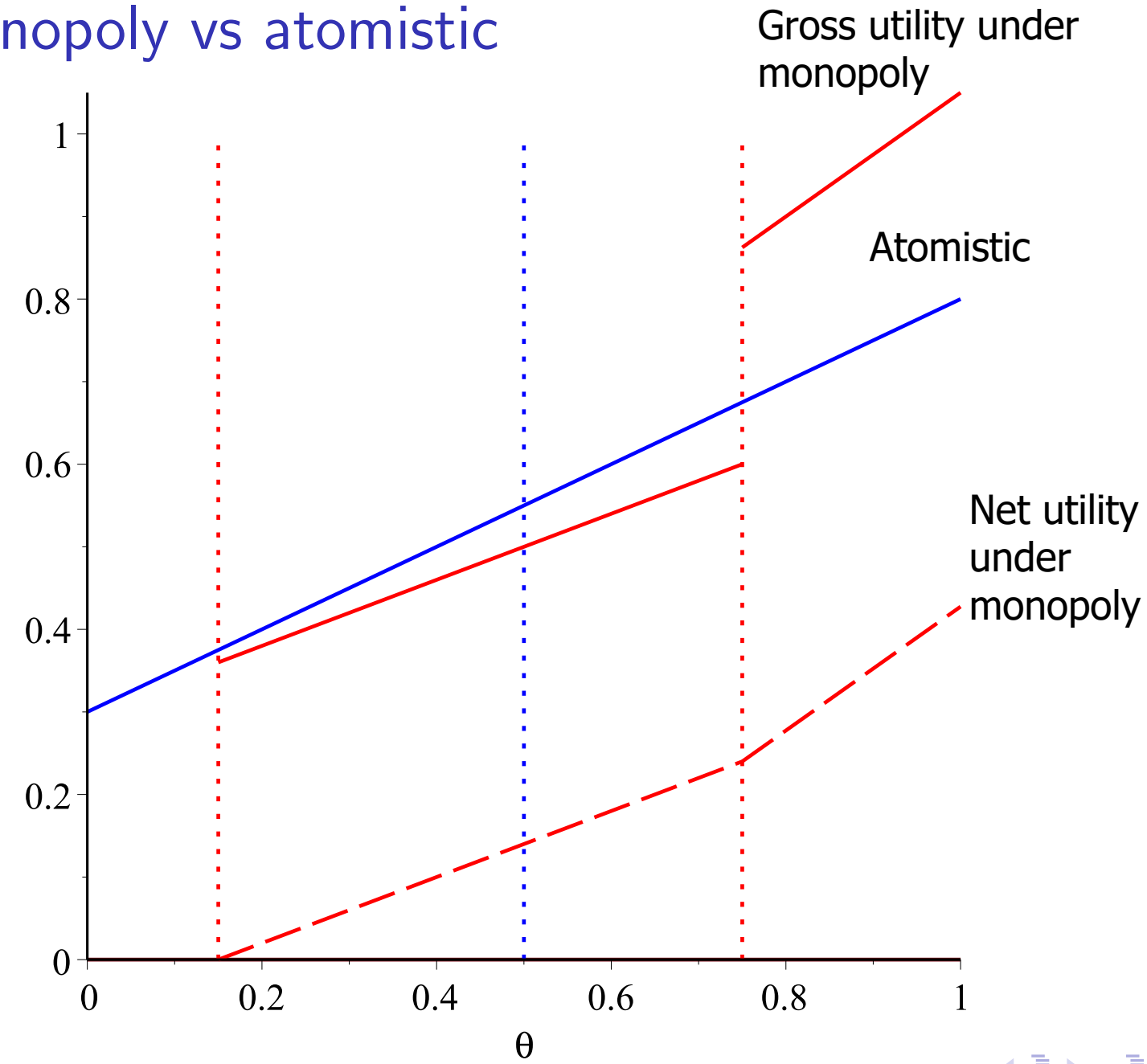
Welfare effects

- when everybody travels also under monopoly (so differentiation is the only change), total welfare turns out to be higher under atomistic users than under monopoly
- change in the number of vehicles dispatched in the transition towards monopoly has ambigous welfare implications
  - monopolist restrics usage when $B'(\theta)$ is large and g is small relative to $B(0)$
  - planner restricts usage when g is large relative to $B(0)$
  - monopolist might restrict too much (when $B'(\theta)$ is large and g is small relative to $B(0)$), and reduce welfare

Monopoly vs planner: comparison



Utility: monopoly vs atomistic



Taxes with atomistic users

- Government sets a per-vehicle tax equal to t in the popular lane, and equal to T in the luxury lane
- With atomistic users, taxes that restore first best are

$$\begin{cases} t_A & \leq B(0) \\ T_A & = t_A + \frac{g}{18}\left(5 - \sqrt{7}\right) \end{cases} \quad \text{if } g \leq 5.4179 \times B(0);$$

$$\begin{cases} t_A & = g\, n_P \left(1 - N_P - \frac{n_P}{2}\right) \\ T_A & = t_A + g\,(1 - N_P)\,(n_P - N_P) \end{cases} \quad \text{if } g \geq 5.4179 \times B(0).$$

- Standard congestion charge: each tax is equal to the congestion cost imposed on other users in the same lane as a result of the choice of the marginal user to travel in that lane
- Welfare-improving. But, in the absence of compensation, low θ's stand to lose: either they do not travel, or travel in a more congested lane

Taxes on monopoly

- Timeline
  - tax authority announces a per-vehicle tax equal to t in the popular lane, and equal to T in the luxury lane
  - monopolist sets F and f
  - corporate users make their travel decisions
- To restore first best

$$t_{MT} = gn - s_{MT}$$
$$T_{MT} = gN - s_{MT}$$

- Very different tax than that on atomistic users
  - gn and gN restore the optimal relation between n and N
    - ★ since $n > N$, tax is larger in the popular lane and discourage the monopolist to overcrowd it, thus reducing differentiation across lanes
    - ★ **not** a congestion charge, but a tax on quality (Cremer and Thisse, 1994)
  - s_{MT} is a subsidy to increase the monopolist's coverage of the market
- When congestion is sufficiently severe (g is large), subsidies exceed taxes
  - Absorbs funding from general taxation. Politically unappealing?
  - Possibly, to be compensated with an ex ante license

Taxes with an emerging monopoly

- Timeline
  - tax authority announces a per-vehicle tax scheme, possibly different between atomistic/corporate users and between users in the popular/luxury lane
  - monopolist sets F and f
  - corporate and atomistic users make their travel decisions
- Tax on atomistic users identical to the case with atomistic users only. Standard congestion charge
- Tax on corporate users is similar to the case of corporate users only, but with a difference. As the proportion of atomistic users increases
  - contribution of corporate users to congestion in each lane turns smaller
  - incentives for monopolist to allocate increasingly more corporate travelers to luxury
  - up to the point of underdifferentiation with respect to social planner, where the goal of the tax changes, and becomes to shift users to luxury
  - structure of the tax must be different when share of atomistic users changes

Conclusions

- Centralization associated to AVs affects congestion problems, with welfare and distributive effects
- With no taxes
  - when one moves from a world with atomistic travelers only to one with a small company with a fleet logic welfare may increase or decrease depending on how severe the congestion problem is in the first place
  - if congestion is severe, more likely that the introduction of a company is useful
  - if the entire market is managed by a single monopoly, too few travelers and too much differentiation across lanes
- With taxes
  - optimal tax on atomistic drivers is a congestion charge...
  - ...which is different from the optimal tax on consumers managed by a company (which, instead, is a quality tax)
  - if the entire market is managed by a single monopoly, optimal tax requires a subsidy for the monopoly when congestion is severe enough

Extensions

- Many competing companies
- Not only competition, but market design
  - 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 congestion)

THANKS!!