Mitigating Congestion and Environmental Impacts from Ride-Sharing Services: The Case of TNC Regulation in São Paulo, Brazil.

***Ciro Biderman*[·]**
***Fernando Haddad*[·]**

"[I]n no other major area are pricing practices so irrational, so out of date, and so conducive to waste as in urban transportation. Two aspects are particularly deficient: the absence of adequate peak-off differentials and the gross underpricing of some models relative to others"

William Vickrey (1963)

Abstract

Taxis have monopolized individual transport for commercial purposes for more than a century. This situation has changed drastically very recently with transportation network companies (TNC) such as Uber, Lyft and DiDi (among others). Differently from other cities, São Paulo opts to charge TNCs for using its roads. The fee is different according to the time of the day, the day of the week, the location in the city, etc. One of the rationales is that the infrastructure is public but the gains from using this public land are private. Furthermore, charging a price on individual transportation must be efficient since individual transportation generates negative externalities. In this paper we discuss the role of such a regulation to mitigate those negative externalities in an economy that is moving towards mobility as a service. Then we discuss the need for advancing the regulation towards pushing TNCs to open their data.

FALTA COMPLETAR AS REFERÊNCIAS E REVER A INTRODUÇÃO

Introduction

Taxis have monopolized individual transport for commercial purposes for more than a century. This situation has changed drastically very recently. Transportation network companies (TNC) such as Uber, Lyft and DiDi (among others operating at the local level) proposed a new product in the market. TNCs connect users to drivers and receive a fee for the service. Differently from other cities, São Paulo opts to charge TNCs for using its roads. The rationale is that the infrastructure is public but the gains from using this public land are private. Furthermore, charging a price on individual transportation might be efficient since individual transportation generates negative externalities. If the fee is able to reduce "overuse" of the road, it will improve welfare.

There is nothing new in charging a price for using public infrastructure. A classic example is charging for construction rights, the general rule in São Paulo. Real estate developers pay for the costs to serve their land or to payback large urban projects that

[·] Corresponding author. Centro de Estudo de Política e Economia do Setor Público (CEPESP/FGV).
[·] Insper

increase the value of their land. The similarity is that developers would make a private gain based on a public investment. Charging TNCs is different from charging developers though. First of all, the road system is a public good with congestion. In an empty road adding a vehicle does not affect the consumption of other users of the road. However, in a crowded road the cost of an extra car is exponentially high (for all other drivers). In other words, society costs changes depending on the use and the use changes over time and space. This is not true for real estate that has a constant use except when density changes what happens seldom. So, there is a need for charging just when density changes. The second difference is that real estate charges the wealth, i.e. the tax is applied over a stock that represents (the present value of) a future cash flow. Road use must be charged over its flow. The incidence (potentially) happens every minute.

In this paper we discuss the logic of charging for road use in general given that drivers do not internalize congestion costs. The following section discusses the congestion externality associated with driving. This is the main problem we are trying to solve when imposing a fee on the use of the roads. The third section discuss the main advances in urban transportation that happened in the last decade showing that the way we commute might change in the next decade but there is no guarantee that it will improve the welfare. The forth section discuss how to regulate the sharing economy that is leading to use capital as a service in general and transportation as service in particular. The main point is how to make a regulation that at the same time allows the new economy to grow and the society to profit from it. The fifth section reports how the regulation has evolved more recently and the sixth section presents the political economy behind the regulation of TNCs. The final section concludes.

Congestion Externality

Road congestion is a problem in almost any large city. Hours stuck on traffic jams represent a large loss in terms of welfare. The problem is that drivers do not internalize the costs they are imposing on all other drivers. Each extra driver slows down the road for all users. Although those costs are individually small, they add to a non-negligible amount. Some estimation would some up to 10% of the GDP. Although it is difficult to agree with such large estimation (sometimes the implicit counter-factual is zero commuting time) there is some consensus that the cost is pretty high.

To understand the source of externality from congestion, we start with a very simple model[1]. Imagine that there is just two ways to commute from your home to your job: you can use private transportation or public transit. Public transit is segregated so its speed does not depend on the amount of commuters using it[2]. Private transport on the other hand depends on the number of cars on the road. Up to some point (let us call it T_c – see Figure 1) there is no congestion on the road. An extra car does not reduce the speed of other cars. After that point an extra car does reduce the speed for everybody else using the road. To be more formal, let us define the cost of commuting in the road to be:

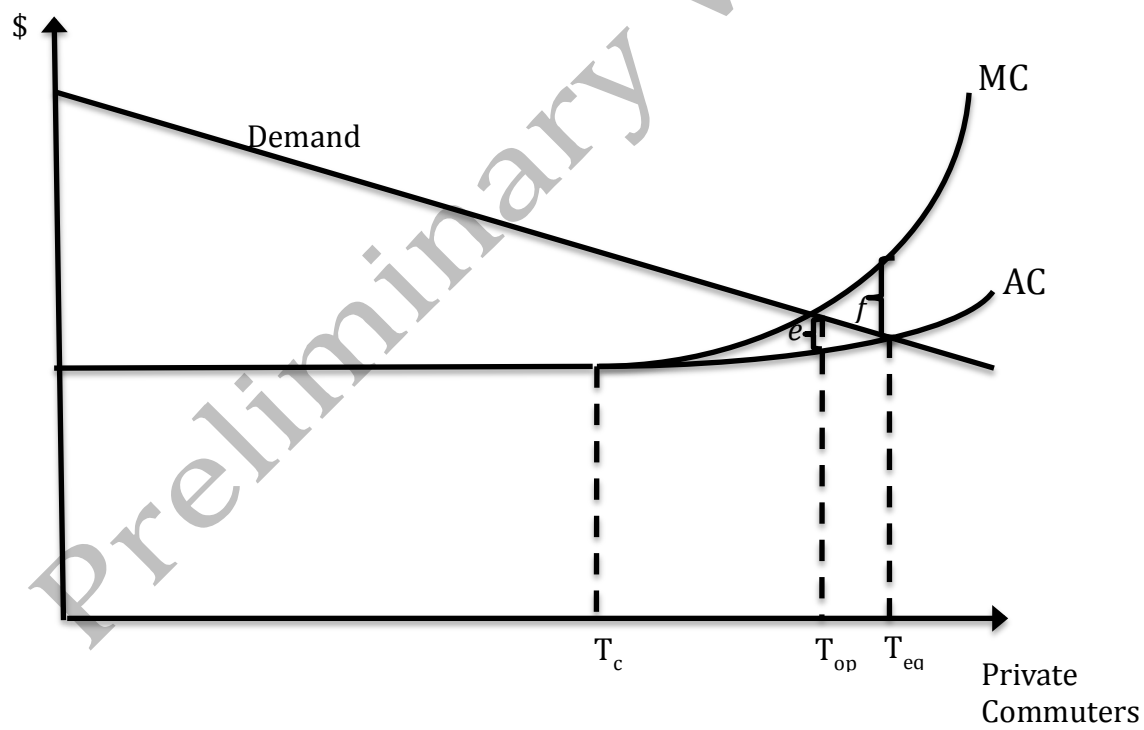
[1] This section follows very close Brueckner (2011); Small and Verhoef (2007).
[2] It may be a subway, a commuter rail or a BRT.

$c = m + wd/s$ (1)

Where c is total commuting cost, m is the monetary cost of commuting (fuel, parking, depreciation, etc.) w is the individual cost per unit of time, d is the distance traveled and s is the speed of the road. The cost of time might be the forgone wage per unit of time or some other measure[3]. Of course, d/s represents the time taken to travel distance d. If there are less than T_c cars on the road, this cost will not depend upon the number of commuters on the road (Figure 1). After this point, however, an extra car would slow down the road. So speed is a function of the number of cars on the road, i.e. $s = s(T)$. Given this fact, the cost of commuting is also a function of the number of cars in the roads after T_c. So, we can rewrite equation (1) as:

$c(T) = m + wd/s(T)$ (2)

Figure 1: Schematic Approach to Negative Externality of Private Commuting



Source: Brueckner (2011)

Consequently the total cost of commuting to the society is $Tc(T)$, i.e. we are aggregating the costs for all drivers. This aggregation assumes implicitly that all

[3] An "ideal" measure would be how much an individual would be willing to pay to reduce commuting time. A meta analysis of studies using this measure shows that the "revealed" cost of time is usually around half of the wage per hour.

drivers travel the same distance. Making a more realistic assumption would require integrating over all distances traveled making it more cumbersome with little gain in terms of interpretation. We can use the total cost to estimate the marginal cost of an extra driver on the road:

$c(T) + Tc'(T)$ (3)

Where $c'(T)$ is the derivative of the commuting time when a new car is added to the road. Notice that the average cost of commuting is simply $Tc(T)/T=c(T)$. Consequently the average cost of commuting after T_c is lower than the marginal cost. This is reflected in Figure 1 where the marginal cost (MC) is higher than the average cost (AC) after T_c. This is the point made before: the individual cost is the average cost but the social cost is the marginal cost. In other words, the "marginal" driver add $c'(T)$ for all drivers but from his standpoint the additional cost is just $c(T)$.

To find the market equilibrium we need to add the demand for transportation in the diagram. Evidently, the demand for one transport mode is always an indirect demand. The real demand is to commute from A to B. The mode decision always depends on the relative cost of all available modes. In this case we assume there is just two modes. Of course we allow commuters to have their own modal preference. So, someone for whom the cost of using public transit is very high will be in the highest part of the demand curve represented in Figure 1. If there is no congestion fee the market equilibrium will happen when the demand cross the average cost curve (T_{eq}). The reason is simple: for any other commuter the alternative mode is more "costly" than driving on the road.

The problem is that the market equilibrium is not socially optimal. The socially optimum equilibrium will happen when the marginal cost curve cross the demand. The reason is that an user valuing the use of the road less than the marginal cost (right after T_{eq}) will be willing to take the road since her alternate cost (for using public transit) is higher than her individual cost to use the road. However, the social cost is higher than the alternate cost for this commuter. Consequently society gain from getting rid from this driver: all drivers would be willing to pay her not to drive.

The best way to reduce the number of drivers on the road is charging for its use. The ideal "Pigou" taxation would make the average and marginal cost curves to be identical. If the number of commuters in Figure 1 is T_{eq}, the tax should be equal to f represented in the diagram. If the number of commuters is equal to T_{op} the tax should be equal to e. If there are fewer commuters than T_c the tax should be zero. An interesting result is that the tax to disincentive commuting at T_{eq} is f but once it is implemented it should eventually converge to e. Consequently a "honest" (local) government could begin with a larger tax and then reduce it.[4]

Moving to the real word, there is never a unique road connecting jobs to residency. Roads are congested around the Business Center. Furthermore, roads are not congested all the time. The main problem happens at peak hours. During "after hours" roads are typically not congested and there will be no reason to charge the road user

[4] An alternative would be charging e directly. In theory the final equilibrium should be in T_{op} as well.

just based on externalities arguments. So, an optimal taxation will change according to the time of the day, the day of the week and the location of the trip. It would also change prices according to the number of passengers in a car or the fuel used to move the car. We will call this "perfect" schedule of taxation (changing according to the time and space) as "Vickrey Pricing".

Singapore is probably the experience that is closest to "Vickrey Pricing". To implement this congestion tool they have to rely on technology. There are multiple sensors around the city attempting to correctly price congestion. Stockholm has a simple system while London is in the middle of the road using cameras to record plate numbers and charge the fee but it does not take into account the number of miles travelled. In the case of São Paulo it is potentially possible to implement Vickrey Pricing for TNCs in a very simple manner. Since TNC drivers must have a cell phone sending car's location to the platform, it is possible to the regulator or to the TNC itself to control the flow with no further ado. Furthermore it is possible to know how many passengers there are in the car, driver gender, fuel technology, etc. In other words it is possible to know all the details of each trip. Having access to platform data in real time would make it feasible to change the fee based on actual speed.

One of the main criticisms on this policy is that the theory would say that any car should be charged not just TNCs. Of course we will not dispute that a privately owned, single-occupant vehicles (SOVs) are the main responsible for the negative externality discussed in the previous section. However, it is not possible to enforce all drivers to have a cell phone reporting their location every minute. It is also debatable how much this would interfere with individual privacy. In other words, we are not claiming that SOVs should not be charged for the negative externalities they generate. We are just claiming that this is not the correct tool for charging them. This is unfortunate since the fee per mile based on driver's cell phone is a very simple, efficient and effective mechanism to implement such a fee.

The straightforward way to charge all cars, from the public policy perspective, is implementing a congestion tool. If you can do it like Singaporeans do, you will be very close to Vickrey Pricing. In this case you do not have to worry about over charging TNCs because the "social optimum" number of cars will be achieved. It is still reasonable to charge the commercial use of roads. The first point to call attention is that very few cities have successfully implemented a congestion tool. To the best of our knowledge, Singapore is the only city to implement it so precisely. Few other cities adopting such scheme usually have a second best solution charging a fix amount to enter the city center. In summary, although this is the best solution from the economic perspective it is politically very difficult to actually implement it for any SOV using city roads.

The second point is connected to the technology. To implement such refined pricing scheme not using driver's cell phone the city needs a system of sensors that are expensive to implement and to maintain. The revenue is higher than the sensors expense and the goal of the pricing scheme is to have a rational use of the car so the policy is still worth it. But it is always a good idea to save on public expenditure. This is the reason why Singapore is now moving from the sensors' monitoring system to cell phones' monitoring. The citizens are apparently accepting this but it is not clear if it would be acceptable anywhere in the world.

This situation raises one of the main issues regarding the new standards of mobility: the rights to information privacy. Although very many people in the world accept to give away his/her privacy to social media companies, the same folks are usually very disturbed by giving access to the government. Looking at the current status quo it is very difficult to believe that people would accept opening their location for the government to charge them even though the theory would say that they would be better off by the end of the day. The case of Singapore is different because they are already paying to drive; opening their data will just make it simpler and cheaper.

So, we can say that the only way people would accept to open their personal data is providing some benefits out of that. More than a billion people open their data to social media so they can participate in it. It is possible that people would be willing to open their data to participate in congestion pricing fee if there were congestion tools implemented in the first place. This is the reason why this scheme is not very likely to work unless you have implemented a congestion tool for all cars in the first place. Another alternative would be providing some other kind of return. For instance, cities with a plate control may allow drivers to use the road on the their "day off" as far as they use an app that would charge them on those days.

There are other ways to charge for the negative externalities related to (individual) driving though. One way is an excise tax on fuel. This is of course a second best option since the price for driving in non-congested roads will be the same as driving in congested roads. It is possible to slightly increase the efficiency of this pricing scheme changing the tax as one gets further from the business center but this scheme is complicated since it is necessary to avoid arbitrage.

Another way is dynamic pricing parking space that is often free in public streets. The proposal by Shoup (2005) although outdated in terms of technology (he is still working with the idea of parking meters) allows for a lot of flexibility in terms of road pricing. Since TNC drivers do not have to park, a dynamic pricing on parking is likely to work with individual drivers. If the parking fee is extended for companies that offer free parking (like in London, for instance) and an excise tax is charged to private parking in certain locations this is, in theory, a better second best than the fee on fuel[5]. The reason is that parking in downtown does not compete with parking in the periphery making arbitrage impossible.

All in all, mobility has been changing considerably in the last decade. There is a consensus that the way we commute in urban areas will be very different in two decades or less. One of the main issues with urban transportation, congestion, will not be solved by the new mobility because this is a social issue. The government is the one responsible to improve social gains. The only way congestion will be managed is if the government takes action. Actually there are many aspects of the changes in urban transportation that requires a change in governance. This paper is concerned with one of the changes presenting how it was faced by the City of São Paulo. Before

[5] A fee on fossil fuel might be a good environmental fee because it would reduce the consumption of this undesirable source of energy but it is certainly not the best way to solve the congestion problem.

continuing with details on how this regulation was implemented we will make a brief detour to discuss how the changes in mobility are interrelated.

The Future of Urban Transportation

There are four innovations going on that will very likely change drastically the way we commute in the next decade. The first one is the advance in batteries. There are already batteries ready to power a bus for 200 miles just charging it overnight. The second innovation is autonomous driving. Usually people are optimistic about the rate of innovation so the forecasts are often too generous and have to be delayed. For autonomous driving it is happening the opposite. The original schedules are being reviewed for a shorter term. The third innovation is on the information and communication technology (ICT). The increase in the capacity to store and process large databases and the advances in programming techniques summarized by the terms "big data" and "machine learning" opened the way to a new generation of applications. The forth innovation is in the business model. The sharing economy is changing the way people consume in many areas. For commuting the impact is already high.

The sharing economy is changing the way people commute somehow showing what transport specialists have knew for a long time. The very need is for commuting from A to B not to own a car. Although the car industry has been trying to convince users that owning a car is a an end in itself it is getting clear to more people everyday that it is a mean. Understanding that the mode is a mean (not an end) received the fancy name "Mobility as a Service" (MaaS) so people believe that this is a quite new idea. Of course the concept has been there at least since the 1950s but none has really implemented it so there are something new in the horizon indeed. Actually, without the advances in ICT it would not be possible to make this (transport economics) dream come true.

A shared vehicle is able to split capital costs in much more driven miles. A typical driver travels less than 20 miles per day. A sharing car can travel 200 miles in a working day. Since the fix cost is shared the total operational cost per mile is lower than using an individual car. If we consider that some trips might be done walking, biking or using public transit people is more and more realizing that, in most cases, it is considerably cheaper not to have a car and use a combination of modes especially in large cities. Even if you drive sometimes, it is likely that the car you demand changes according to the purpose of the trip. Going on vacations will probably require a larger car, maybe a FWD SUV. If you rent a car few times you can rent the ideal car for your trip and save money. Actually in many trips families and individuals are required to rent a car in the site and leave their car siting at home since it is too far away to do the whole trip by car.

Although it was probably cheaper to use a combination of modes instead of owning a vehicle before the advent of sharing commuting, this fact was not so clear. There are two reasons for that. The main reason is that Taxi's monopoly made it more expansive to use this specific mode. So, the economic advantage was lower due to high price. The second reason is the supply of vehicles. There was a lot of uncertainty

whether it would be possible to find a cab when needed and it was also inconvenient in many instances to wait on the street for a cab to arrive.

The combination of autonomous vehicles and the business model discussed above make it possible to have autonomous vehicles in a shorter term than previously forecasted. The main problem with autonomous vehicles is capital price. Notice that for TNCs running an autonomous vehicle will save on labor costs. So, as far as the flow associated with the increased cost of capital break even with the saving in avoiding paying for the driver, autonomous vehicles are economically sustainable for TNCs. This will happen much sooner than it would if individual drivers were supposed to buy the technology. An autonomous vehicle can be much more efficient in its job and may work 24/7. Car companies are already forecasting that sales will drop by 1/6 of current sales in two decades or less.

The combination of the advances in batteries technology combined with the sharing economy has a similar impact on the speed it will be implemented. Since the main increase in cost moving from a regular fueled car to a battery fueled car is in capital, TNCs and taxis will probably be able to change the fueling technology faster than individual users. The problem here is on the way alternate energy sources are developing and on the many alternatives there are in the market.

The fact is that there was not an energy revolution like there were in the first and second industrial revolution (steam power and combustion engine, respectively). The slow advance in alternate sources is at least partially a governance problem. There is not a game changer in physics or chemistry like there were in the two industrial revolutions. This is the reason why calling the recent advances as a third industrial revolution is very much a wishful thinking. Technology has advanced for sure but not in such a way to call it an industrial revolution. The same applies to the new business model that is the object of this paper. It is an advance but it is nothing like Ford's line assembling.

Given that, however, the government could be able to make alternate energies to develop faster than the current pace. The first mistake is the way tax incentives are being implemented. Carbon taxes may help the environment since they induce a lower use of "dirty" sources of energy. But they are very unlikely to buster a major innovation in technology (Acemoglu, 2012). The industry based on fossil energy can just keep their productivity above the productivity of the non-fossil energy "plus" subsidies/taxes. And they have been quite successful in doing so. A tax on fossil energy producers' profit would be much more efficient.

This paper is not directly concerned with this shortcoming in governance related to the advances (or lack of advances in this case) in urban transport. But there is one relevant issue very much related to this paper that could make a significant difference in emissions and congestion at the same time. If TNCs were able to increase car occupancy the impact on the environment and congestion would be immediate. Currently in most cities the rate of occupancy per vehicle is bellow 1.5 people per car. Doubling it would mean reducing the number of cars by half. This is much more than what is necessary to end congestion altogether and it will be reducing emissions by half. The problem is that this is not actually happening.

The pool option that is still not really feasible poses another question to the way we currently commute. Public transit and taxis are basically separate modes so the possible complementarity is not very much explored. Cities with good transit structure usually have a trunk system (often based on subway and/or commuter rail) and a local system using buses to feed the trunk. The experience of DiDi in China is and Via in New York showing that TNCs might be more efficient in supplying the feed needs than buses. This is probably true since one of the main problems with bus operations is the lack of demand in off-peak hours and in remote areas. Since the logic of traditional operation calls for a limit on waiting time, it is usual to see empty buses going around and people waiting too much to catch a bus. Working with mini buses and/or Vans TNCs may be able to do a better job at a lower cost on the feed side of the system.

Roads occupy around 25% of total land in urban areas and represent the bulk of public space in any city. If the main commuting mode will be taxis or the shared version of it the need for parking space would reduce considerably. Parking space represents something from 5 to 10% of total land available in urban areas. If walking, cycling and public transit is complement rather than substitute for taxis the number of motorized trips might diminish. Increasing the number of users will make pool options really feasible. Currently the pool option represents a small part of the rides but the "Bus on Demand" alternative is promising. If pooling trips become a reality, it would further reduce the need for land dedicated to cars. Increasing space for biking and walking will likely increase the demand for those modes reducing car trips once more. Electric vehicles make it easier to pedestrian and cyclists to share roads with cars. There is a possible improvement in mobility standards connected to the new mobility but once again it will depend on the government using the space opened by the new standards of mobility to improve the space for active modes.

All in all, the share of TNCs on total trips might increase in the next decade. So, although it currently represents a tiny part of total trips, no more than 1% in São Paulo (2017 figures), it is growing very fast. If we compare taxi's trips in 2007 with taxis plus TNC's trips in 2017 they grow by 424%; from 90.7 thousand trips in 2007 to 475.3 thousands in 2017. Taxi's trips grow 24.5% to 112.9 thousand trips well above total trips' growth in the period (8,7%). However, it is very impressive that in 2007 TNCs do not exist and in 2017 they totalized 362.4 thousand trips (0.9% of total trips)[6]. It will not be surprising if TNC trips grow larger than car trips. It means that we have to learn how to regulate and use the price system to make this sub-mode an asset rather than a liability for the urban transport system.

We do not expect TNCs to replace the trunk system though. The main corridors will probably need mass transit to work efficiently. The three innovations discussed in this section however have a relevant impact also for the trunk system especially in developing countries. Autonomous buses fueled by batteries in totally segregated lanes are almost perfect substitutes for subway. Many scholars have argued that BRTs are indeed a better solution for cities that do not have a significant subway network. Some scholars and practitioners have argued the opposite considering environmental (subways are fueled by electricity) and capacity issues. Both arguments on the superiority of the subway vis a vis the BRT are false giving the innovations discussed

[6] Metro (2019) "Resultados Preliminares da OD de 2017".

in this section. Actually, batteries powered at night are better for the environment than subways that use energy more intensely exactly when it is scarcer: during peak hours.[7] There is no reason to believe that BRTs with battery fueled autonomous buses will not be able to have subway's capacity with a similar environmental impact.

Regulating and pricing the sharing economy

What is called the "sharing economy" has been growing at a quite fast pace. In five years of operation, Uber was operating in 128 cities worldwide. After four years of operation, Airbnb was offering the same number of rooms for rent as Intercontinental Hotels Group, the largest hotel chain in the world that took 65 years to get there. Those companies are getting funds summing up to billions of dollars in the market. And the government has been watching this movement and either doing nothing or responding with the same tools it has been using for several decades.

We can find four types of regulations for TNCs around the world. The first is limiting TNC supply sometimes just banning the use. Another type of regulation is very rigid attempting to regulate very many details such as what should be offered in the vehicle (water, candies), the type of vehicle, etc. The relevant part of the rigid regulation is an attempt to fix the number of drivers allowed in the platform. There is also a very flexible regulation imposing almost no specific constraint on the service. Few cities created an excise tax. São Paulo adopted a different strategy charging a fee for TNCs per mile. São Paulo regulation intended to be in the middle position regulating relevant issues but allowing the market to grow as the demand increases. This strategy allows for using the pricing system to correct for negative externalities. It also allows the public sector to regulate the size of this market a relevant concern for many different reasons. In other words, the proposed price per mile is a regulatory price, not a revenue scheme. On the other hand it does have a non-negligible potential for increasing revenues in an efficient way. This is its merit but, at the same time, its risk.

Charging a fee per mile is at the same time relatively simple (for TNCs) and refined. It is a benefit fee and it can be used to regulate quantity. This is something that was difficult to understand. Most journalists and the secretary of transport himself when the fee was implemented had a hard time understanding the target system. The city defined a target that was settled in terms of "Taxi Equivalent" or, in other words, in miles. The city estimate how many miles per month five thousand taxis would travel and decided that this would be the target for the sum of miles run by all TNCs combined. Most observers believed that there would be an auction selling miles. This is not the way the system was regulated.

If TNCs travel more than the target there are two possible actions: increase the target if the city believes that the target was set too low or increase the price. A higher price might induce fewer trips. On 2017 the City decided to increase the target to 10

[7] Actually, for many countries including Brazil and most states in the US – notably California – increasing the use of energy during the day would rely on thermoelectric run by Diesel. On the other hand, given that batteries weight at least 25% of the weight of a vehicle, there is a need to waste energy transporting the batteries themselves.

thousand “Taxi Equivalent” miles but the price per mile has never changed since its inspection in 2016 despite de fact that Bus’ tariffs have increased 13% in the same period. The changes in the regulation following the initial proposal will be discussed later.

This solution is very different from the solutions in other cities that we are aware. In the US, 7 cities have implemented a fee and/or a tax rate on total fare (Kim and Puentes, 2018). The most usual is a fee per trip that varies from $0.24 in Seattle up to $2.75 in New York City. In the US just Philadelphia and Washington, DC currently have a tax levied on TNCs trips (1.4% and 6% on total fare, respectively). Mexico City has also imposed a tax rate on TNC trips. This is different from the system proposed by São Paulo City. The first difference is that if you charge a fee per trip you will be charging equally long and short trips. If we expect TNCs to substitute for first and last mile this is not very smart.

Furthermore the tax on trips do not allow for fine-tuning the policy. Two policy strategies are particularly important on this regard. In São Paulo the fee varies depending on the time of the day and the location of the trip. The fee is higher in the center where roads are indeed scarce and lower in the periphery where there is actually a lack of taxis’ supply. It is also more expansive in rush hours and close to zero in off hours when it is actually desirable to have a cheap option reducing the incentives for drink and drive. It is possible to include this scheme in the charge per trip regarding the time of the trip but rather complicated to make it for the location. To the best of our knowledge, none city has done it but São Paulo and cities that copied São Paulo’s regulation.

The rate solution will be similar to the fee per mile in regular conditions since total fare is a function of the miles driven. In rush hours it will increase given the surge price used by most TNCs. But there is no reason to believe that it will decrease in off hours. There is no reason to believe that it would be lower in the periphery as well. If there is a lack of drivers in the periphery compared to the demand it might be even higher. In theory it could be possible to change the rate depending on those factors but it is very unlikely that the council would approve such complicated scheme. In brief, the rate makes the city a “partner” of TNCs instead of their regulator.

The regulation created a committee (CMUV) including the secretaries of transport, finance, urban infrastructure and urban development and the president of the “São Paulo Business Company” (São Paulo Negócios in Portuguese) that was responsible for drafting the regulation. This committee has the power to change the fee, change the target, change the schedule of the fee and also create new schedules. It meets once a month to analyze TNC trip data and decide if it wants to change the regulation. This is a very flexible system that does not depend upon the council not even on a Mayor’s decree. This factor was not so publicized but it was crucial in the design of the policy.

Another aspect that was included in the regulation was discounts for the pooling option. The fee will be lower if there are two people sharing the trip than just one; lower for three people than two and so on. The fee will be close to zero for a trip shared by a party of four. The only city that we know that deal with pooled trips is New York City that reduces the fee from $2.75 to $0.75 per rider in a pooled trip. The policy behind it is not so clear to us since a trip with a party of three will pay more (in

total) than a trip with a party of two. A trip with a party of four will pay more than a trip with just one rider. Congestion concerns would recommend the opposite. In this case it would be simple to use the fee per trip and just charge a lower (total) amount that would be shared among riders.

The pool is actually very promising. SOVs are used to have a door-to-door service and are usually not willing very much to walk to a bus stop, wait for the bus, walk from the final stop to work or school… The TNC individual trip option is just substituting the driver with no impact on the way the person has the service: it will still be door-to-door. It would not affect congestion and emissions either if it does not substitute for a public transit trip. The pool option may affect emissions considerably at a relatively low cost for the user. Let us think about a two people trip. It would increase at most two stops comparing to an individual trip and reduce one trip for the city if those trips were originally by SOVs.

Why are pool trips not catching up? The problem is finding two (or more) trips that match so perfectly. Although TNCs trips have increased at a very fast place and now represent more than double of taxi trips in most cities it is still a small part of the trips. For instance in São Paulo City where Uber makes more trips than anywhere else in the world it is still 1/15 of the trips by car or public transit.

There is a variation on the P2P system that has been called “Bus on Demand” (BoD) and it is currently the best bet for transforming TNC’s individual trips into collective trips. The system is slightly different because it works just with groups; it is never an individual trip. Combining more than 5 people is key in BoD business model. TNCs operating with individual trips have been competing with public transit (although they refuse to admit) and consequently taking users away from the collective mode. This is not neutral in terms of congestion and emissions. It is actually adding trips to the city. Furthermore it is making public transit more costly. BoD models seem to be the middles of the road solution.

The way public transit is typically priced is charging less if you ride longer. This is evidently wrong from a pure efficiency standpoint but hard to change once it is established. In Latin American cities it may be justified on the base of equality since the poor live in the periphery. Short trips subsidize long trips. If there is an option that is close in price for short trips part of the users will change their commuting moving to e-hailing. Buses are losing users and consequently revenues with no reduction on costs. So, bus operators themselves have been trying to get back those costumers using BoD technology. BoD may attract other SOV trips as well and we may finally see a reduction in the total number of trips that was promised by TNCs but never delivered.

There are many different models for BoD but basically it offers a trip that is cheaper than e-hailing but more confortable and faster than traditional buses. It may pick you up closer to your origin and drop you off closer to your destination; makes less stops; and adjust the route to be more consistent with your needs. There are many current experiences but most of them are either contracted directly by the municipality or by the Bus operator itself. There are few experiences where the TNC have to survive directly from user’s fees.

All of those concerns bring to the point that to make the future of mobility better for the society we need to integrate modes including TNCs. If TNCs are taking people out of their cars to make the first or last mile of the trip, at least part of the trip will be done in a collective mode compared to previously when the whole trip used to be done by individual modes. If part of larger buses is moving to smaller buses through a BoD system but not increasing the number of vehicles (the larger bus was partially hidden), it is better for the society. A smaller bus has a smaller impact on congestion and emissions as far as the number of trips does not increase. It is possible that a small increase in the number of trips may be justifiable in two senses: increasing the financial sustainability of traditional bus operators (lowering its cost more than the revenue loss) and improving the quality of the service to users (increasing frequency).

Integration is currently very limited in most cities. Each mode and sub-mode has its own clearinghouse: TNCs, BoDs, Municipality, (Traditional) Bus Operators, Taxis (that might even not have a clearinghouse system), Scooters, Bicycles, etc. The only integration that is more usual is between the subway and the bus system. The challenge here is to create an integrated system that will foster innovation in payment systems and will use the income generated from floating the money to finance the transportation system itself. This step was not taken by any city, to the best of our knowledge, but this is the only way to really move towards Mobility as a Service.

The point here is that there is no reason to believe that the fee on TNCs is making it harder to make use of this complementarity between modes. Cities are not reducing the number of motorized trips because the systems are not physically or financially integrated regardless of the fee on TNCs. Integrating the system financially will imply in, for instance, returning part of the fee to the user when it transfer from an individual to a collective mode. It can also reduce the fare itself for each mode that will be splitting the total cost of the trip. The challenge to make the complementarity between hailing or e-hailing with public transit and active modes works is improving intermodal operation.

There is also one characteristic of the per mile fee: it is straightforward to use it for a kind of affirmative policy. In São Paulo case, one of the uses is to incentive women drivers. There are very few women driving taxis in São Paulo and this is usual around the world. There is much more women driving cars for TNCs. Given the demographics, it would be expected that women would represent half of the drivers. The idea behind this discount was that there is segregation in this job market that would vanish as soon as the number of woman drivers is significant. The other discounts were applied to non-fossil and hybrid vehicles and also for cars adapted for handicaps. Para-transit is a relevant issue in most large cities and increasing the number of adapted vehicles is an important step towards universal accessibility.

What's next?

On 2017 a new tenure started in São Paulo City and the regulation resisted. In part this is due to the fact that the revenues are not insignificant, around 50 million dollars per year. Although this is low for São Paulo City it is relevant from the investment budget. Furthermore, it might be considered illegal to reduce revenues with no good reason according to the Brazilian law of fiscal responsibility. Although the spirit of

the regulation was to use the fee as regulatory, rather than a revenue source, the revenue side of it makes it more resilient. The changes in the regulation were towards making it more rigid: regulating characteristics of the driver and the car.

CMUV resolution 16 is one example of this movement. It established two mandatory documents for TNC drivers that turn the regulation of TNCs closer to taxi regulation. The original idea was the very opposite: making taxi regulation more flexible. The documents were called CONDUAPP and CSVAPP that defined minimum requirements for drivers and for the insurance of the car, respectively. Resolution 16 defined, for instance, standards for driver's clothes. What calls attention to this part of the regulation is that the standards were defined just for man. This detail is saying a lot about who was behind the text.

Resolution 16 was a first step for Decree 58.595 issued by mayor Bruno Covas on January 4 2019. The decree attempted to control the number of drivers working for TNCs. The interpretation of the 10 thousand taxis equivalent was transformed into allowing just 10 thousand drivers in the platform. Asking drivers to have a CONDUAPP would allow the control of the number of drivers. The problem is that this is not the way TNCs work. Some drivers will work just on weekends or just after work. Some of them will be temporarily in the platform until they find another job. Consequently the platform is very dynamic with new drivers joining the platform every day and others leaving. This decree would be a killing decree for TNCs working in São Paulo. TNCs react to the decree that was considered not constitutional and in practice had no effect on the regulation that is still very similar to the original concept. Maybe the only difference is a sign in the windshield showing that the car is used as a TNC vehicle.

The regulation could have advanced in some areas but it did not. If something it went backward. The main point to advance would be pushing TNCs to share their data, something that the previous tenure failed to implement, and creating incentives to modal integration. The initial proposal was for TNCs to develop an API that would allow the city to access the data and control the place and time of each vehicle and consequently charge the correct amount per mile that could be adjusted in real time. Having the data would allow for more than monitoring the TNCs; it would be potentially an important tool for planning and monitoring the whole transport system.

Uber has never accepted to build the API claiming that it would affect privacy and reveal strategic information about the company. São Paulo has never proposed a way to push TNCs to open their data finding ways to protect privacy. This is a problem faced by many regulatory agencies in the world. Privacy is of course quite relevant. There is famous example in New York City when the Taxi and Limousine regulation agency open the data on driver's trips and a data scientist was able to find the routes taken by celebrities living in New York. Chicago and Los Angeles are probably the cities that have done more in terms of making this information available for the government and for the general public as well. Chicago idea was randomizing the exact position and time of the origin and destination of the trip within an area and time frame making it almost impossible to find out exactly where and when the trip started but allowing for making the analysis at a small geographical area and time period. Los Angeles is establishing a protocol to share data from any mode (bicycle, scooters, cars) that will protect privacy and allow opening the data for the general

public. After pioneering one of the most advanced regulation for TNCs, São Paulo is lacking behind not updating or improving the regulation.

Two months after the regulation was implemented, it was clear that the 5 thousand taxi equivalent target would be reached. On 2016 the city decided not to increase the target because Taxis were already complaining about the low demand they were facing. However, giving the political moment, it also decided not to increase the price per mile that was contradictory with the model logic. The problem for São Paulo in 2016 attempting to reduce TNCs activity using the pricing system was that the incumbent (Uber) had extremely large share of the market. It was clear that increasing prices would not have a large impact on their provision and consequently it would have a low impact on TNCs' activity.

Platforms have one characteristic that makes it almost a natural monopoly. This is called network externalities. A platform with a lot of drivers can supply the service fast attracting more clients and then more drivers since they will have more demand in this platform. There is a first mover advantage that is very relevant. Large consumers of inputs are usually not very sensitive to price. So, a linear increase in price would probably have a low impact on the main player and it could jeopardize smaller TNCs attempting to compete in the market. The decision was to implement a progressive tax, a scheme that is already used in many scarce inputs such as energy and water. Congested roads do qualify as a scarce input.

Uber had 95% or more of the trips when progressivity was implemented and react to the change in regulation. The judiciary suspended progressivity but São Paulo City government appeal and win reinstating the progressivity. In six months Uber share went down to 70%. More recently Uber appeal again and was able to suspend progressivity. The City Hall did not appeal and currently progressivity is suspended.

It is somehow schizophrenic the way City Hall currently deals with the regulation. On one side it is accepting taxis' complains and formally attempting to make it harder for TNCs to operate in São Paulo market. On the other hand it does not pressure TNCs to open their data, do not increase the price per mile and do not appeal to decisions taken by the court. The interpretation is that the City Hall is trying to avoid the political difficulties in regulating the service that is probably the reason why few cities have actually implemented a more rational regulation.

On one hand medallion owners are often politically powerful. Councilors do not see any gain from supporting TNCs and feel that they could lose votes supporting the regulation. It is well known that few councilors often have one or more medallions and were elected thanks to taxi drivers' and their families' votes. Those councilors are very active in protecting taxi drivers' rights, of course. Although we are talking about a few councilors the others do not see any advantage in fighting those colleagues. This is a typical "group of interest" case and there is nothing new about it. At least in Brazil this is likely the reason why many municipalities are not actually regulating the service.

There is something new however. Usually "group of interest" politics is explained by the fact that everybody else is hurt by so little that are not willing to spend time supporting the policy change. That is the usual rationale for import tariffs, for

instance. In this case, the equilibrium at the council seems to follow this rationale but this is not how citizens are reacting. It is clear that e-hailing is making a difference for users that could not use taxis before because of price or availability. Furthermore, there are much more e-haling drivers than taxi drivers. This group, their families and friends might be willing to support a consistent regulation of the system. This is probably the reason why many cities were just ignoring the existence of the service in their cities: there was no need to confront any group if you just ignore the phenomenon.

It was politically difficult to implement the new regulation in São Paulo. Some secretaries did not agree with the policy and were afraid of its political consequences. The fight in the council was tough with consequences for the relationship between the executive and the legislative branch. It is not a surprise why many cities just ignore the existence of TNCs: it is a political corner. If the city bans TNCs it has problems with part of the voters and if it regulates it will have problems with another part of the population. Curiously regulations that are not constitutional are attractive because you do not affect the service but apparently act for one part of the voters. That is probably one of the reasons why there are few rational regulations such as the one described in this article. If we analyze recent CMUV resolutions we can see that the City of São Paulo was adding regulation that was, in theory, protecting the taxis' group of interest but that were never actually implemented because they eventually lose in the court. This is the most likely political equilibrium but it is very bad for sound public policy.

It is the political economy, stupid

TNCs valuation is large for two main reasons. First they capture part of the rent generated by the original monopoly (Taxis). Second they solve information problems connecting clients to drivers. Matching demand and supply increase sales considerably. The evaluation system tells users that the driver is trustable diminishing the risk of adverse selection. The payment system reduces the risk for the driver that the client will not pay back (moral hazard). Any large or medium city had this service before but it was a very small part of the market supply. For instance, it was very usual to have private cars furnishing this kind of service in hotels for tourists that did not trust taxis. Actually, when a TNC enters a new market those are exactly the drivers they first approach.

When the information problem is solved the number of trips using this new service skyrocket because it is cheaper and more trustable than taxi services. The information gain is reflected in the extensive margin: a large increase in trips due to an increase in the number of drivers. It is possible that the drivers that were in the market before increased their trips as well but there were so few private drivers that the possible growth in the intensive margin is insignificant. There is not so much evidence on TNCs' trips but there is a perception that they are making much more trips than taxis (in São Paulo 3.2 times more[8]).

The rhetoric of the taxis' group of interest was very efficient before the existence of TNCs. In most cities the license to drive a taxi are concentrated in a small group, i.e.

[8] Metro (2019) "Resultados Preliminares da OD de 2017".

there are much more drivers than license owners. The usual rhetoric was arguing for an increase in taxi's tariffs to improve drivers' wages. However, since driving a taxi does not require relevant skills, any increase in tariffs was always capitalized in the license price. Medallion owners' have always been able to convince taxi drivers to riot for an increase in tariffs to increment their income but all it has done is incrementing the medallion price and consequently the taxy daily rent. Taxi tariffs might go down to compete with TNCs but the daily rent will go down accordingly. So, for most taxi drivers, net income will not be affected. As a matter of fact in many cities we can notice a significant reduction in medallion prices. In São Paulo taxis' tariffs are steady for 5 years and the number of trips using this model has increased faster than total trips.

One important point about the regulation is that the sharing economy cannot pay taxes based on the stock like the property tax. This would be totally incompatible with the idea of moving towards MaaS. On the other hand, the experiences with deregulating this market have been very bad. Jimmy Walker scandal and La Guardia medallions in 1930's New York is one example of this risk. With large unemployment rates it is easy to end up with too many cars on the street increasing congestion instead of reducing it. There is weak evidence that this is happening right now in some large cities. But this is exactly the idea of the regulation; if there are too many cars, just increase the price per mile.

Taxis work in a different way and there are reasons to believe that this business model should not disappear. São Paulo regulation was also willing to keep this modal competitive in the industry. Instead of paying a fee per mile, taxis pay for the license. In equilibrium license fees should be equal to the capitalization of the fee per mile[9] so, both groups are paying exactly the same except that taxis pay it in advance. The difference is that taxis implicitly commit to work constantly otherwise the fee per mile would be too high. The advantage to users of the service is that the driver will be usually more professional than a TNC driver. From the transport planning perspective there is also a potential gain in keeping the service alive. Since the government can control the number of licenses and this was the reason why license owners were accumulating rents before TNCs' entry, it is possible to implement transportation policies that would be more difficult to implement otherwise. For instance, not allowing cars in downtown areas would be criticize in terms of universal accessibility but somehow this would be compensated if taxis can still enter the regulated perimeter. It is not possible to allow TNC drivers to use those areas since you can be a TNC driver making very few trips per month.

Another relevant point is that there are two reasons for charging TNCs for using road space. We discuss carefully the externality side of it but there is another relevant issue. TNCs are using a public provided good (roads) to make profit out of it. Residents of the city have already paid for the roads when they pay property taxes. But this is different since TNCs are using the roads as input to generate profit. So, they must pay the cost of maintaining the roads. This is also the rationale for charging the license fee. The direct consequence for TNC regulation is that the negative externality pricing should be determined on the top of the maintenance cost.

[9] Of course, for this statement to be totally true, it would be necessary to have an auction for licenses.

The final point about the regulation of this new economy is the information generated in this process. The information is a quite powerful tool to improve the management of traffic in this case. Pushing TNCs to share their information with the public sector will give a new set of tools that would be very costly to obtain otherwise. For instance, the most recent fad is scooters' sharing. Once again cities are struggling to regulate the service and some are already prohibiting the new service. As we discussed before, prohibition is in general a very weak public policy. Scooters are a case in which real-time information is most needed because the problems related to this mode of transportation is the way users park the vehicle and safety. To efficiently regulate parking and safety it is necessary to have real time information on parking and speed. In other words, municipalities must regulate the use of their roads and they can make it with TNCs charging for the use of the road and accessing their information always respecting users' privacy.

It is clear that TNCs' business model will advance both in the proportion of trips as in scope. A general regulation of the services would include the concept that any new provider should open their data to the regulatory agency and be subject to paying a fee for using City's roads. The fee might be even "negative" (a subsidy). Using TNC information cities would be able to regulate the service that is getting more complex. Currently the best way to do it is using an API but cities must be prepared to change the regulation as soon as a new technology is developed. What is important is having access to the information in real time using the best technology available. This is currently one of the main issue in TNC regulation: who should have access to which information; how to protect privacy; how to open information to the general public to foster further innovation? Another main issue is how to integrate this myriad of transportation services that are now operating in any large city in the world. We will see many advances in this front in the next years but São Paulo is not anymore in the frontier of this movement.

Conclusion

The new business model using the sharing economy is among the most dynamic sectors in the economy in the last decade. It is changing the way we think about mobility especially when we paired it with the advances in battery technology and autonomous vehicles. It might represent a significant share of the trips in the future and it may help changing drastically the modal share in large cities. The current pattern is probably not sustainable.

Governments need to regulate this new system. It is not clear why the government usually charges property tax but it is not charging appropriately the use of roads, the largest proportion of public land in any city in the world. Given the way the new economy operates there is a very good opportunity to make it a perfect tax that increases efficiency instead of creating a burden in the economy, as it is usual for the tax system. It is possible to do a perfect pricing at a very low cost.

São Paulo proposal deal with this problem and it was successful implemented. Local governments around the world have to start thinking creatively on how to regulate and tax these new entrants in the market without affecting their operation and growth. The

opportunity is siting on the table and if we do not catch it we will loose an amazing opportunity to increase revenues and efficiency at the same time.

There are two other issues that were not resolved by São Paulo regulation that are likely to be the main issues in mobility in the next decade: how to increase interoperability among different modes of transportation and how to open mobility data. Cities are currently experimenting with these issues and Chicago and Los Angeles are probably the more advanced examples in data accessibility. Finding a way to guarantee privacy and at the same time open the data to foment new business and improve monitoring and management capabilities from the government perspective is a must as well as integrating the new modes with public transit.

References

Acemoglu D., P. Aghion, L. Bursztyn e D. Hemous (2012) "The Environment and Directed Technical Change". *American Economic Review*, 102(1): 131–166.

Brueckner, J.K. (2011) *Lectures on Urban Economics*. MIT Press.

Kim, So Jung and R. Puentes (2018) "Taxing New Mobility Services: What's Right? What's Next?". *Eno Brief*. Eno Center for Transportation.

Metro (2019) "Resultados Preliminares da OD de 2017".

Shoup , D. (2005) *The High Cost of Free Parking*. Planners Press.

Small , K. A. , and E. T. Verhoef (2007) *The Economics of Urban Transportation*. Routledge.

Vickrey, W. S. (1963) "Pricing In Urban And Suburban Transport" *American Economic Review*, 53(2): 452-465.