Proposal Title: Dynamic Transportation Markets in the Digital Economy: Matching Efficiency and the Value of Time

Investigators: Jakub Kastl (Princeton and NBER), Nick Buchholz (Princeton), Laura Doval (Caltech), Tobias Salz (MIT and NBER)

After going through a few rather uneventful decades the ground transportation sector is now experiencing a number of changes that will have important implications far beyond the sector itself. Changing agglomeration economies and the urban revival have drawn people back to city centers and, as a result, strained local transportation infrastructure. At the same time, technical innovations, such as GPS tracking and the wide adoption of smart-phones provide an opportunity to organize transportation in "real time". Those technical innovations have led to the emergence of new platform companies, such as Uber, Lyft, and Via, which are often referred to as transportation network companies (henceforth TNC's).

The economics literature has highlighted many of the positive implications of the services offered by TNC's. On the demand side, they have increased consumer surplus through a combination of new entry, the reduction of matching frictions, and flexible pricing (Buchholz (2018), Castillo *et al.* (2017), Frechette *et al.* (2019), Hall *et al.* (2017), Shapiro (2018)). On the supply side, they have created new opportunities for flexible work arrangements, thereby creating new earnings opportunities that are compatible with existing full time work arrangements (Chen *et al.* (2017)).

However, there are also negative side effects. Current estimates from NYC suggest that regular ride hail services achieve capacity utilization of only 60% since part of the road miles are spent on-route to the passenger or searching for the next passenger (Report of NYC's DOT (2019))[1]. This can lead to severe congestion (Mangrum and Molnar (2017), Agarwal *et al.* (2019)) and other types of externalities.

Both these positive and negative aspects of TNC's are part of an important and general trade-off in transportation markets between capacity utilization, on the one hand, and preference for flexibility and immediacy on the other hand. Public transit and personal transportation can be thought of as rather extreme points on this spectrum. Public transportation can achieve high capacity utilization but is typically inflexible with long waits in-between scheduled arrivals. Private transportation, on the other hand, provides a lot of flexibility but uses resources inefficiently. From a social point of view, high capacity utilization is beneficial because it reduces both environmental and congestion externalities.

The goal of this research project is to provide a framework for estimating both supply and demand in transportation markets as well as matching efficiencies. We use a unique and detailed data set from a European ride hail company. Given the exceptionally high quality data available to us we are able to generate unique insights into passengers' responsiveness to prices and waiting-times. The data also allows us to quantify the matching efficiencies of TNC platforms. In an almost completed project we use those insights to measure how the opportunity cost of time is distributed across a

[1] `https://www1.nyc.gov/assets/tlc/downloads/pdf/fhv_congestion_study_report.pdf`

city, which provides an alternative quantification to other measures of the heterogeneity of urban economic activity (for example, rents and real estate prices).

The more long run goal is to investigate the feasibility of a *new form of public transit* that harnesses the technical innovations of TNC's for public infrastructure. Such a new form of public transit would leverage all of the technological advancements that the industry currently employs and aim to achieve higher capacity utilization than personal transportation while, at the same time, providing more flexibility to riders than traditional public transit, which operates both on a fixed grid and fixed time schedule. Such a middle ground could be especially promising in cities that traditionally did not have much success in running a public transit system due to path dependence in institutions, or constraints in geography and infrastructure.

Data and Environment for current study

We have access to a unique dataset provided by a TNC called Liftago, which currently operates in the Czech Republic and Slovakia. Unlike regular taxis, which essentially clear on waiting times, and unlike Uber and Lyft, which essentially clear on prices, this company employs a decentralized market that can use both of these to reach an equilibrium. In particular, for any passenger requesting a ride, it asks several nearby available drivers who are required to be licensed cabs to submit bids and the passenger is free to choose among them. Due to the inherent randomness of this process (such as where the previous ride finished) different drivers will have different arrival times at which they can pick the passenger up. Since this is observed by passengers when making choices, this provides us with a very clean environment where passengers face a discrete choice from a set of alternatives with variation in price and waiting time (and few other observable characteristics such as car model, driver rating etc). We observe the universe of ride requests and rides performed over several years.

The Ingredients of the Framework: Demand, Supply, and Matching Efficiencies

Demand and Preferences for Flexibility. A better matching technology can lead to lower waiting times and, therefore, draw more passengers to a given TNC platform. A crucial economic parameter is, therefore, how passengers value time - and how this value of time varies across locations or time-of-day. Those preferences can be quantified in terms of demand elasticities with respect to waiting times. In Frechette *et al.* (2019) it is estimated as -1.2, however demand is not directly observed in that study, which means that the authors are relatively limited in the ability to study how this elasticity varies over time or across space. Platforms like Uber and Lyft employ "surge" pricing to equilibrate demand and supply, so that consumers do not get to directly express their preferences over prices and waiting times within the platform. Instead, since we observe a large panel of choices with useful variation in both waiting times and prices in the choice sets, we can employ a simple revealed preference estimator. We find large variation in time preferences across space, time of day, and, leveraging the panel structure of our data, also across consumers. For example, from 6AM to 9AM we estimate a waiting time elasticity between -0.74 to -0.66. In the evening, from 7PM to 11PM, demand becomes much less responsive to waiting times, with elasticities from -0.25 to -0.17. We further exploit our data to disaggregate these elasticities by latent demand

types (riders with high and low sensitivity to waiting) and trip types (trips arriving before or after important times of the day). The estimated demand model allows us to quantify how passengers would substitute between cabs on the platform and the outside option, permitting us to conduct a rich set of counterfactuals.

Supply: The market is essentially organized as an auction - drivers (bidders) are uncertain about the number of rivals they would be facing conditional on being asked to submit a bid, their locations and bids. Ultimately, for a given ride they face a decision on what, if any, bid to submit given the opportunity cost of leaving their current location. Since a candidate trip (in the event they would win the auction) would take some time to complete and would also take them to a different area, the opportunity cost includes an option value of staying put and obtaining a potentially more lucrative fare a bit later. This dynamic decision problem thus can be summarized by two value functions: $\mathcal{H}^t(\cdot)$ and $\mathcal{S}^t(\cdot)$. $\mathcal{H}^t(a_t, \hat{a}_{t+\tau}, \omega)$ is the value of a platform request for a trip from location a in time t to $\hat{a}$ taking τ periods and with $\omega \sim \mathcal{F}(.|a,t)$ being a draw of the off-platform per-period earnings opportunities, which could include a fare request outside of the platform that we study or simply the value of alternative usage of driver's time. $\mathcal{S}^t(a_t, \omega)$ is the continuation value of location a with ω at t. Let $\delta(a_t)$ denote the probability a driver is asked to bid on a ride. Then:

$$\mathcal{S}^t(a_t,\omega) = \delta(a_t)\cdot \mathbb{E}_{\hat{a},\tau}[\mathcal{H}^t(a_t,\hat{a}_{t+\tau},\omega)|a_t] + (1-\delta(a_t))\cdot\left[\omega + \mathbb{E}_{\omega}[\mathcal{S}^{t+1}|a_t]\right]$$

where

$$\mathcal{H}^t\left(a_t,\hat{a}_{t+\tau},\omega\right) = \max_b \left\{p(b)\cdot\left(b - f + \beta^\tau\cdot\mathbb{E}_{\hat{\omega}}\left[\mathcal{S}^{t+\tau}\left(\hat{a}_{t+\tau},\hat{\omega}\right)\right]\right) + \left(1-p(b)\right)\cdot\left(\omega + \beta\cdot\mathbb{E}_{\hat{\omega}}[\mathcal{S}^{t+1}(a_t,\hat{\omega})]\right)\right\}$$

where f is the fee collected by platform from drivers and $p(b)$ probability that passenger accepts bid b.

Let us assume for the moment that bids by drivers are chosen from a connected set (in practice, they are restricting to lie on a finite grid chosen by the driver corresponding to different fares per minute and mile in their meters). With continuous bids, we can follow the usual approach in empirical analysis of auction markets and estimate the opportunity cost that would rationalize the observed bids under the assumption the observed distribution of bids corresponds to what the drivers also believe they are bidding against. These opportunity costs can be expressed after collecting terms in $\mathcal{H}(\cdot)$ as:

$$v\left(a_t, a_{t+\tau}, \omega, t, \tau\right) \equiv -\left(\beta^\tau\cdot\mathbb{E}\left[\mathcal{S}^{t+\tau}|a_t,a_{t+\tau}\right] - \beta\mathbb{E}\left[\mathcal{S}^{t+1}|a_t\right] - \omega\right). \tag{1}$$

and hence, we can rewrite the bidder's problem as in a first price procurement auction:

$$\max_b p(b)\cdot\left(b - f - v\left(a_t, a_{t+\tau}, \omega, t, \tau\right)\right).$$

It follows that with our data $v(\cdot)$ can be estimated and from (1) one can recover the value functions by using the definitions of the value functions and appropriate projections. These estimates will then allow us to re-estimate drivers equilibrium choices as we conduct the counterfactuals of interest.

Matching Efficiencies: Since passengers have to be matched to the driver and other passengers in real time it is important to understand the properties of this matching process. A crucial quantity of interest are the returns to scale that can be achieved in the matching process when the market grows thicker. If the scale economies are large, sufficient market share might allow for very low waiting times and provide a high degree of flexibility to passengers. In Frechette *et al.* (2019) it is documented that the (random) matching process between passengers and taxis exhibits increasing returns to scale. A dispatch algorithm that assigns passengers and taxis optimally leads to much lower baseline waiting times but features lower scale economies. These results are based on simulations that take into account a stylized version of Manhattan's street grid.

In the current project we build on some of these results and investigate the properties of the matching technology of Liftago. One advantage of our data over the existing studies with taxis is that demand is directly observed in terms of requests that are made through the app. This means that both the inputs, vehicles and passengers, as well as outputs, matches, are directly observed. With this type of data we are able to non-parametrically estimate the matching function for pooled rides. We currently find that there are increasing but relatively modest returns to scale. This finding is in line with Frechette *et al.* (2019), which shows that a dispatch technology has much lower baseline waiting times than random matching but lower returns to scale.

Deliverables for this Proposal

The three ingredients (demand, supply, and the matching function) that we describe above provide a complete framework to study TNC's. In an almost completed study, we rely heavily on the demand side of this framework. In this project we (i) quantify the benefits from having a choice across transportation alternatives and (ii) micro-found the value of time. The latter essentially uses TNC data as a measurement device to explore heterogeneity of opportunity costs across different urban spaces. We see this as an interesting and novel way to complement standard measures of heterogeneity across urban spaces, such as real estate prices or time-use distributions. In addition, this exercise allows us to efficiently identify which origin-destination pairs in a city are relatively under-served by public transit infrastructure. These together show how private data sources might be used to improve public infrastructure. Without TNC data one might not be able to gather "price signals" across certain origin-destination pairs that are not currently served by public transit.

A second project, likely completed by early to mid next year, is using the full framework above to explore market design aspects of ride share platforms to increase their efficiency and better understand the economics of ride-share. The insights from this project can help illuminate regulatory aspects of ride share markets and, more broadly, the industry structure. In the next paragraph we provide a vision of how the insights from the current research might serve the longer run goal of a new form of public transit that capitalizes on the innovations brought about by TNC's.

Long Run Goal

For the long run goal of this project, we build on some of these results and investigate the properties of the matching technology of pooled rides. We are currently exploring another data

partnership with a company that predominantly offers pooled rides. This work would go beyond previous work in economics, which only analyzes the matching process between the driver and one passenger. With the new data one could non-parametrically estimate the matching function for pooled rides. Existing engineering studies (notably, Alonso-Mora *et al.* (2017)) suggest that an efficient real-time routing algorithm would be able to serve all Manhattan taxi trips with only 3,000 drivers (medallions) instead of the current 13,500. However, such simulations studies omit many practical issues and, at best, serve as a benchmark to which one can compare the empirical properties of a matching process for pooled rides. To begin with, this study assumes that demand is inelastic (in waiting time). An adequate analysis of such a market will therefore require, in addition to a realistic model of the matching function, a demand system that takes into account passengers' preferences for both prices but also waiting times. Our existing studies on the demand side, in particular how sensitive passengers are to waiting times, will therefore complement the analysis of the matching function of pooled rides.

Conflict of Interest Statement

Neither of us is in any way affiliated with the company that provided us with access to the data set and there were and are no strings attached to us using the data (other than a non-disclosure agreement) for academic research purposes.

References

Agarwal, S., Mani, D. and Telang, R. (2019). The impact of ride-hailing services on congestion: Evidence from indian cities. *Available at SSRN 3410623*.

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E. and Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, **114** (3), 462–467.

Buchholz, N. (2018). *Spatial equilibrium, search frictions and dynamic efficiency in the taxi industry*. Tech. rep., mimeo, Princeton University.

Castillo, J. C., Knoepfle, D. and Weyl, G. (2017). Surge pricing solves the wild goose chase. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, ACM, pp. 241–242.

Chen, M. K., Chevalier, J. A., Rossi, P. E. and Oehlsen, E. (2017). *The value of flexible work: Evidence from uber drivers*. Tech. rep., National Bureau of Economic Research.

Frechette, G. R., Lizzeri, A. and Salz, T. (2019). Frictions in a competitive, regulated market: Evidence from taxis. *American Economic Review*, **109** (8), 2954–92.

Hall, J. V., Horton, J. J. and Knoepfle, D. T. (2017). Labor market equilibration: Evidence from uber. *URL http://john-joseph-horton. com/papers/uber_price. pdf, working paper*.

Mangrum, D. and Molnar, A. (2017). The marginal congestion of a taxi in new york city. *Processed, Vamderbilt University*.

Shapiro, M. H. (2018). Density of demand and the benefit of uber.