

NBER WORKING PAPER SERIES

DO DIGITAL PLATFORMS REDUCE MORAL HAZARD?
THE CASE OF UBER AND TAXIS

Meng Liu
Erik Brynjolfsson
Jason Dowlatabadi

Working Paper 25015
<http://www.nber.org/papers/w25015>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2018

We thank Keith Chen, Dean Eckles, Andrey Fradkin, Xiang Hui, John Horton, and Erina Ytsma, as well as seminar participants at MIT, Uber, AEA 2018, Marketing Science 2018, SICS 2018, and 2018 NBER Summer Institute Industrial Organization Workshop and the NBER Digitization Workshop for their valuable comments and suggestions. The MIT Initiative on the Digital Economy provided generous research support, and Uber provided essential data. Dowlatabadi is a current employee at Uber. The views expressed here are those of the authors and do not necessarily reflect those of Uber Technologies, Inc or the National Bureau of Economic Research. All errors are ours.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Meng Liu, Erik Brynjolfsson, and Jason Dowlatabadi. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Do Digital Platforms Reduce Moral Hazard?
The Case of Uber and Taxis
Meng Liu, Erik Brynjolfsson, and Jason Dowlatabadi
NBER Working Paper No. 25015
September 2018
JEL No. D47,D8,D86,L15,L91,M52

ABSTRACT

Digital platforms like Uber can enhance market transparency and mitigate moral hazard via ratings of buyers and sellers, real-time monitoring, and low-cost complaint channels. We compare driver choices at Uber with taxis by matching trips so they are subject to the same optimal route. We also study drivers who switch from taxis to Uber. We find: (1) drivers in taxis detour about 7% on airport routes, with non-local passengers experiencing longer detours; (2) these detours lead to longer travel times; and (3) drivers on the Uber platform are more likely to detour on airport routes with high surge pricing.

Meng Liu
Washington University in St Louis
and MIT
mengliu@mit.edu

Jason Dowlatabadi
Uber Technologies
jasond@uber.com

Erik Brynjolfsson
MIT Sloan School of Management
100 Main Street, E62-414
Cambridge, MA 02142
and NBER
erikb@mit.edu

1 Introduction

Digital platforms are growing rapidly, and so are their economic effects. Examples include large platforms such as Uber for ride-hailing and Airbnb for accommodations, as well as a growing number of smaller platforms such as ClassPass for fitness studios and Rover for dog-walking. Digital platforms are often designed to mitigate information asymmetry problems through the use of new technologies and incentive systems, such as ratings of buyers and sellers, real-time monitoring, and low-cost complaint channels. For example, 73.5% of New York City (NYC) UberX trips are rated by passengers and Uber fare adjustments are made for 1 in every 170 trips. In contrast, NYC taxi complaints are much more difficult to lodge and occur only 1 in every 6,356 trips.

Do digital platforms significantly affect moral hazard or service quality, compared to traditional settings? The answer to this question is of crucial importance for a better understanding of the nature of online-offline competition and welfare in the digital economy. In this paper, we study this question by comparing a particularly successful and pervasive digital platform, Uber, with traditional taxis.

Specifically, we investigate driver detour, defined as the extra distance a driver adds to the fastest route. This is a measure of driver moral hazard in our context. In a hypothetical situation where a taxi driver and an Uber driver drive between the same two points at the same time, the difference in their routing decisions should reflect factors that affect the benefits and costs of detouring. To the extent that features such as shared GPS navigation, tech-aided monitoring, ratings, and digital feedback increase market transparency for passengers and therefore increase penalty of driver moral hazard, the Uber driver's routing is likely more efficient than that of the comparable taxi driver in situations with high moral hazard payoffs for both drivers.

A key challenge exists in identifying the effects of driver moral hazard — driver moral hazard is not directly observed. The inability to directly observe driver moral hazard is due to the lack of optimal routing benchmark at the time of the trip. For example, using a long-run average trip distance queried from routing engines such as Google Maps may underestimate the true real-time optimal route and overestimate the detour if there was a temporary road closure that required a longer route than indicated by the long-run average. Therefore, one needs to construct valid counterfactuals to infer opportunistic behavior by using detailed trip-level data of both taxis and

Uber. We overcome this challenge by leveraging public taxi trip records and proprietary UberX data in NYC and matching taxi and Uber trips at granular origin-destination-time level. As a result, the drivers of the matched trips are subject to the same real-time optimal routes, even if these are not directly observed.

These matched pairs of taxi and Uber trips then become our units of analysis, where we explore the variation in the within-match taxi-Uber routing difference across route types that represent different moral hazard incentives. However, the estimation approach (essentially a “difference-in-differences” approach) would face an identification challenge if route types were not randomly assigned to taxi and Uber drivers. We address this challenge by exploiting the institutional background of the taxi industry and the Uber platform. For taxis, there is no passenger selection of taxi drivers, as taxi drivers are *ex ante* homogeneous to passengers. Due to the strict taxi refusal law and sample over-representation of short trips in thick markets and airport trips, driver selection of passengers is at best limited. On Uber, rider assignment, performed by Uber’s algorithm, is practically random to individual drivers, and driver selection of passengers is deterred through multiple mechanisms. Therefore, the market itself approximates the experimental ideal. Nonetheless, we also exploit the within-driver variations in some specifications to further purge potential selection issues.

We find that drivers indeed respond opportunistically to changes in technology and incentives. When the fare is metered in trip distance, taxi drivers detour by an average distance of 7.4% on airport trips relative to Uber drivers. Taxi drivers detour even more when the rider is from outside the New York City area and on a metered airport trip. Uber drivers are not immune to moral hazard either: They are more likely to detour on metered airport trips during periods of high surge multipliers when they are paid more per mile.

These findings are consistent with our model of driver moral hazard, where the driver decides whether and how much to detour and the speed of travel, given the set of information, pricing schedule, and overall incentive devices at work. The key tension is a trade-off between the costs and benefits of the opportunistic behavior; the costs include expected monetary and reputation costs and expected forgone earnings lost due to the opportunistic behavior (for example, detouring usually prolongs travel time, reducing opportunities for additional trips), while the benefits are the extra pay earned on the trip. It then follows that drivers lack detour incentives when driving short

trips in thick markets (e.g., within-Manhattan trips), in light of low returns and high opportunity costs; and the detour incentive is greater on airport trips where the long distance rewards detour more. On the other hand, the driver's incentive to speed increases when the metered fare has lower compensation for driving time.

A careful analysis of alternatives reveals that drivers do not detour because they can save passengers time; in fact, detouring on average leads to longer trip durations. Furthermore, the efficiency loss due to detouring is large: Taxi drivers' detours result in 150 lost passenger hours per day in NYC, whereas the lost passenger time is 10 times smaller for Uber. Given that both taxi and Uber metered fares reward detours on airport trips, the reduced amount of unnecessary travel by Uber drivers in these situations suggests an important role for tech-aided market design in mitigating moral hazard.

However, when not detouring, taxis on average travel faster and finish trips earlier than Uber drivers.¹ This speeding incentive of taxi drivers is consistent with the taxi metered fare that rewards distance traveled more than driving time, relative to Uber's pricing schedule.

We strengthen our identification of driver moral hazard by exploring competing explanations. First, we show that the data are not compatible with the hypothesis that the difference in GPS adoption accounts for the observed taxi driver routing inefficiency. Second, we demonstrate that this is not mainly driven by drivers self-selecting into more profitable routes, as we observe no significant change in estimation results when Uber or taxi driver fixed effects are controlled for. Moreover, driver selection can also take place on the extensive margin (i.e., taxi and Uber may represent different distributions of driver types). Following 1,892 former taxi drivers who switched to Uber, we find that these drivers used to detour as taxi drivers, but they no longer do after joining Uber. This is strong evidence that drivers re-optimize under Uber's arrangement via behavioral change. Lastly, we rule out the possibility that taxi drivers route longer to avoid tolls for passengers, by observing a higher tendency of taxi drivers to take toll roads than Uber drivers.

The taxi industry offers a clean laboratory to study the relationship between technology and incentive design. First, the marketplace is highly competitive. Both taxi and Uber drivers offer a

¹Although passengers may prefer faster travel, driving faster than the ongoing road traffic creates safety issues. Also, drivers' weaving in and out of lanes contributes to traffic congestion (https://www.dmv.ca.gov/portal/dmv/detail/pubs/hdbk/idt_cong_rr_phones). Therefore, a full welfare account should include detours, time preference, safety, and externality.

homogeneous, well-defined service (namely, transporting a passenger from one point to another), take the price as given, and maximize earnings as individual entrepreneurs. Second, GPS coordinate data allow us to make precise and valid comparisons between taxis and Uber at the trip level. Such counterfactual groups can be difficult to form in other industries. As a result, evidence from this industry makes a strong and clear inference about moral hazard which can shed light on other industries and markets as well.

1.1 Literature and Contribution

Our paper is closely related to several strands of the literature. The first is on how technology, particularly information technology (IT), mitigates the agency problem in various settings (Tabarrok and Cowen (2015)). In the typical workplace, IT-enabled monitoring has been found to be productivity-enhancing through complementing performance pay (Aral et al. (2012), Bresnahan et al. (2002)), reducing employee shirking (Nagin et al. (2002)) or misconduct (Pierce et al. (2015)), and increasing standard process compliance (Staats et al. (2016)). In the context of trucking, Hubbard (2000) has found that on-board computers which facilitate monitoring of drivers increase productivity by improving both drivers' incentives and managers' resource allocation decisions. Duflo et al. (2012) have shown that incentive pay enabled by tech-aided monitoring can raise teachers' attendance rate and consequently student performance. Reimers et al. (2018) have found that insurance companies' monitoring technologies reduce driver moral hazard and fatal accidents. Sudhir and Talukdar (2015) have illustrated the role of IT in inducing business transparency by showing more corrupt businesses resist IT adoption more.

Besides the traditional settings, there are also studies on digital market designs that improve productivity by regulating agent incentives. Hui et al. (2016) have identified efficiency gains from eBay's buyer protection program as a result of reduced seller moral hazard and seller adverse selection. Klein et al. (2016) have shown that a change in eBay's policies that led to less biased buyer ratings of sellers also improved seller effort and quality without inducing sellers to exit the market. Gans et al. (2017) have evaluated the role of Twitter as a mechanism of consumer voice in disciplining firms for low quality. Liang et al. (2016) have found that IT-enabled monitoring mitigates moral hazard on an online labor platform.

While these aforementioned studies focus on technological improvements either within the

offline or online setting, we are among the first to provide a direct online-offline comparison to study the relationship between technology, agent incentives, and quality provision. As many sectors are being digitized, empirical studies of how incentives and quality provision differ between online and offline markets become crucial for a better understanding of the nature of online-offline competition.

The second strand of literature our paper is related to is the literature on digital disruption and online-offline competition (Bakos (1997), Brown and Goolsbee (2002), Brynjolfsson et al. (2003), Brynjolfsson and Smith (2000), Forman et al. (2009), among many others. See Goldfarb and Tucker (2017) for a review). In particular, this paper contributes to the studies of emerging tech-aided ride-hailing platforms. These platforms may reduce matching frictions between drivers and passengers (Buchholz (2015), Frechette et al. (2016)) with real-time technologies and dynamic pricing (Castillo et al. (2017), Hall et al. (2015)), as reflected in higher driver utilization (Cramer and Krueger (2016)), as well as quick adjustments to market equilibrium (Hall et al. (2017)). Specifically, efficiency induced by dynamic pricing critically depends on consumer preferences and the tradeoff between wait time and price (Lam and Liu (2017)), and driver labor supply that responds to wage fluctuations (Chen and Sheldon (2016)). Consumers benefit from ride-hailing platforms extensively, because of surge pricing (Cohen et al. (2016)), reduced drunk driving (Greenwood and Wattal (2017)), and improved service quality due to safer driving (Athey and Knoepfle (2018)). Drivers also benefit from these platforms due to flexible work arrangement (Chen et al. (2017), Hall and Krueger (2015)) and commission schemes that allow for driving without a lease (Angrist et al. (2017)). We find that these technological and organizational features have important implications on driver incentives and quality provision, and thus add an important layer in the analysis of efficiency.

Finally, our findings resonate with earlier empirical work on taxi driver opportunistic behavior, such as Balafoutas et al. (2013), Rajgopal and White (2015), Balafoutas et al. (2017), and Liu et al. (2017). We contribute to this literature by examining how moral hazard can be mitigated by tech-aided ride-hailing platforms.

2 NYC Taxis vs. Uber: Market Design and Pricing

2.1 Taxi and Uber Market Design

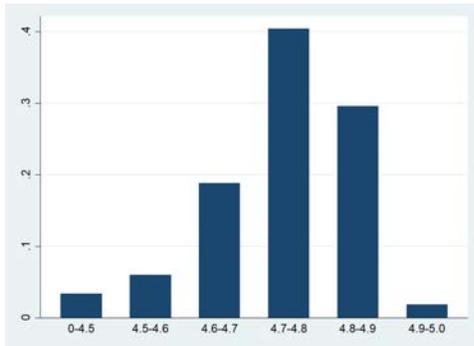
The market design for the Uber platform differs significantly from that of taxis. First, GPS navigation is widely adopted and used by Uber drivers, while taxi drivers mainly navigate without GPS. The Uber app is designed in a way that GPS navigation is integral to both driver and passenger: When the driver picks up a passenger and starts the trip, Uber's built-in GPS automatically initiates, or the app switches to the preferred GPS that the driver has set up (e.g., Google Maps and Waze). Therefore, driver routing on Uber is more transparent for passengers relative to taxis.

Second, Uber implements a set of institutional design choices that aim at aligning driver incentives and facilitate monitoring by the passengers. These are absent or costly with taxis. With the Uber app, passengers can readily monitor driver routing in real time; passengers can either monitor the route on their own smartphone app, or look at the driver's app, since the driver's phone is usually mounted in a way that it is visible to passengers. This way, passengers can easily tell whether or not the driver is taking the route that is given by the GPS.

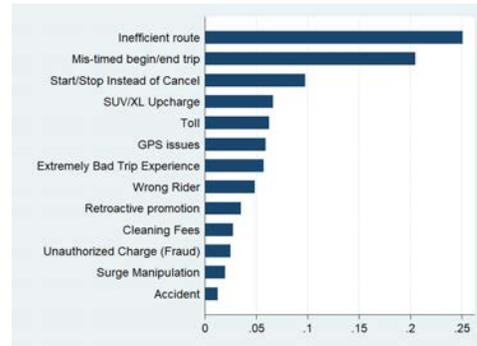
Uber uses a highly-visible rating system that is easy for users to update. After each ride, passengers are prompted to select a star rating, and therefore most passengers rate their drivers (73.5% for NYC UberX, January to June, 2016). Similar to other reputation systems that are subject to rating inflation (Filippas et al. (2018)), Uber driver ratings are highly concentrated with a mean of 4.74 out of 5 (see Figure 1a). Drivers with low ratings are constantly warned by Uber. Uber starts to consider deactivating a driver when the driver rating falls below a threshold (4.5 in NYC). Drivers appear to be very concerned about their ratings², and perhaps as a result, the actual deactivation risk is relatively low (about 3%).

²The qualitative study by Lee et al. (2015) states that "Drivers took their ratings seriously. High ratings such as 4.98 became a source of pride whereas a rating below 4.7 became a source of disappointment, frustration, and fear of losing their jobs."

Figure 1: Distance and Duration Ratios of Matched Taxi and Uber Trips



(a) NYC UberX Driver Rating Distribution



(b) Top UberX Fare Adjustment Reasons

Notes. Both plots are based on NYC UberX trip data, January to June, 2016. Figure (b) plots UberX fare adjustment reasons, conditional on a fare adjustment being made, for adjustments that account for 1% or more of the total.

In addition to monitoring and rating, verification and complaints can be made with little friction on Uber, thanks to electronic trip records. Passengers can revisit the historical trip summaries on their app to verify certain details. In the case of negative riding experiences, Uber passengers can easily file a complaint through the app, and Uber handles the conflict resolution by evaluating the trip records. By contrast, taxi passengers in these situations can either call the TLC hotline or visit the TLC website, but the process is usually long and may require legal procedures. In 2016, taxi complaints were 1 in every 6,356 trips, whereas Uber fare adjustments were 1 in every 170 trips. Figure 1b lists the main reasons of fare adjustments, with the number one reason being “inefficient route”.

2.2 Taxi and Uber Pricing

Pricing also differs between taxis and Uber. NYC taxi fares are set by the Taxi and Limousine Commission (TLC)³. Most routes are metered with a base fare of \$2.50 upon entry and \$0.50 for every $\frac{1}{5}$ miles traveled, plus taxes, fees, and tolls. A \$0.50 per-minute charge is applied in place of the per-mile charge when the traffic is slow (less than 12 miles per hour). The exception is that routes between Manhattan and JFK Airport are not metered; instead, a flat rate of \$52 applies. Some taxi drivers are medallion-owners who essentially run the business as an entrepreneur. Other drivers lease the medallions on a daily, weekly, or monthly basis, and they collect all the revenues minus gasoline and some vehicle maintenance costs. In both cases, drivers are residual claimants

³Refer to the official language on the pricing rule: http://www.nyc.gov/html/tlc/html/passenger/taxicab_rate.shtml

who are incentivized to maximize earnings.

Unlike the pricing of taxis, Uber’s pricing schedule is consistent in both fast and slow traffic. The UberX base fare includes a fixed component of \$2.55, \$0.35 per minute of travel, and \$1.75 per mile of travel, plus taxes, fees, and tolls. On top of the base fare, passengers also need to pay the surge multiplier in effect at the time of request. For a 2-mile, 10-minute trip with a surge multiplier of 2, UberX costs $2 \times (\$2.55 + \$0.35 \times 10 + \$1.75 \times 2) = \19.10 , plus taxes, fees, and tolls. Unlike taxis’ fixed fare on certain routes, all Uber routes in NYC are metered according to the same pricing formula. Uber drivers keep all trip earnings minus Uber’s commission, which usually runs between 20% and 30%. Uber drivers who operate using their own cars are responsible for all operation-related expenses, such as insurance, maintenance, and gasoline. Many NYC Uber drivers instead rent a vehicle from fleet owners due to heavy TLC requirements such as commercial insurance.

3 A Theoretical Framework of Driver Moral Hazard

In this section, we describe a theoretical framework of driver moral hazard that builds on [Liu et al. \(2017\)](#). A risk-neutral driver maximizes her payoff by choosing among alternative routes, where essentially the driver decides on the amount of detour (from the optimal passenger route) as well as driving speed. For a given route at a given time, let d_0 denote the distance of the optimal route given by a generic GPS that optimizes trip time. That is, any other route with a different distance than d_0 expects a longer travel time. Then the realized trip distance, d , is given by:

$$d = d_0(a + x + \epsilon), \tag{1}$$

where a represents the driver’s ability, e.g., driver’s knowledge of the streets and navigation skills, and $a \in [1, +\infty)$. Let x denote the amount of detour, where $x \in [0, +\infty)$. Let ϵ denote the random driver-route shock that affects routing efficiency, which is normally distributed with a mean 0⁴. For simplicity, let the realized travel time, denoted by t , be linear in trip distance:

$$t = \gamma d_0(a + x + \epsilon) + y, \tag{2}$$

⁴It is possible for the realized trip distance to be shorter than the GPS-suggested distance d_0 , when the random shock ϵ is sufficiently negative. This occurs, for example, when a road turn that is permitted during a certain time of the day shortens the route but is not captured by the GPS.

where γ measures how trip distance maps into trip time, and $\gamma \in (0, +\infty)$. Let y represent the extra travel time incurred when the driver drives at a different speed than the ongoing traffic: $y > 0$ when the driver drives relatively slow, and $y < 0$ when the driver drives relatively fast.

Let the metered fare be characterized by the base fare upon entry p_0 , the per-mile rate p_d , and the per-minute rate p_t ⁵. Let s denote the surge multiplier, where $s = 1$ for taxi trips, and $s \geq 1$ for Uber trips. Let q_e represent the probability of getting a subsequent passenger at the drop-off location and time if there was no detour. Let e_t denote the per-minute earning of the forgone trip.⁶ Therefore, $q_e e_t (\gamma d_0 x + y)$ measures the earnings from forgone service minutes. Then the driver chooses the amount of detour (x) and the speed of driving (equivalent to y) to maximize the following expected payoff function:

$$\begin{aligned} \text{Max}_{x,y} \text{ E} \{ & s[p_0 + p_d d_0(a + x + \epsilon) + p_t(\gamma d_0(a + x + \epsilon) + y)] \\ & - f(x; d_0, \Theta) - g(y; d_0, \Theta) - q_e e_t(\gamma d_0 x + y) \}, \end{aligned} \quad (3)$$

where f is the penalty cost of detour, which can be viewed as the probability of getting caught times the monetary cost and/or reputation cost. The cost may be in the form of fines (taxis), lost tips (taxis)⁷, low ratings (Uber), and refund to passengers (taxis and Uber). f is assumed twice differentiable in x , with a parameter set $\{d_0, \Theta\}$, and $f(x = 0) = 0$, $f_x > 0$, $f_{xx} > 0$. In addition, $s \in \Theta$, and $f_{xs} > 0$, meaning that the marginal detour penalty on Uber is greater when surge is greater⁸. Defined similarly as f , g is the expected monetary cost, reputation cost, or both associated with the travel speed: $g > 0$ for all y , i.e., both driving unnecessarily slow and unnecessarily fast relative to the traffic tend to be noticed and penalized by the passenger⁹; $g_y < 0$ for $y < 0$, $g_y > 0$ for $y > 0$, and $g_{yy} > 0$.

Taken together, the driver's problem in Equation 3 is to solve two trade-offs: One trade-off is between the monetary reward of detour and the opportunity cost of detour, where the opportunity

⁵Note that in NYC normal traffic, $p_t = 0$ for taxis.

⁶ e_t can be thought of as $\frac{p_0 + p_d D_e + p_t T_e}{T_e}$, where D_e and T_e are the expected length and duration of the forgone trip, respectively.

⁷By the end of our sample period, Uber had not implemented the tip feature in the application.

⁸This is because (1) Uber passengers are more incentivized to monitor driver routing when surge is high, and detour may result in worse ratings than when surge is not in effect; (2) fare adjustments reflect surge multipliers.

⁹While it is true in some cases passengers prefer fast driving, speeding and weaving in and out of lanes are among the leading factors of traffic accidents.

cost consists of the expected detour penalty and the forgone payoff; the other similar trade-off applies to driving speed. The first-order conditions are listed below,

$$f_x(x; d_0, \Theta) + q_e e_t \gamma d_0 - s d_0 (p_d + p_t \gamma) = 0, \quad (4)$$

$$g_y(y; d_0, \Theta) + q_e e_t - s p_t = 0. \quad (5)$$

Then, the comparative statics lead to the following testable implications (see the online appendix for proofs):

Hypothesis 1: *Drivers tend to detour more on longer routes than on shorter routes because longer distance increases detour payoffs unless the demand at the drop-off location is sufficiently high, marginal detour penalty increases significantly with trip distance, or both.*

Hypothesis 2: *Drivers detour more when the rider is a non-local passenger, and they detour less when the rider is a local passenger, as non-local passengers are less likely to notice the detour because they lack knowledge of local geography.*

Hypothesis 3: *Drivers detour more (respectively, less) during high surge prices if the increase in marginal detour payoff due to high surge dominates (respectively, is dominated by) the increase in marginal detour penalty due to high surge.*

Hypothesis 4: *Drivers detour less (respectively, more) when the demand at the drop-off location is higher (respectively, lower).*

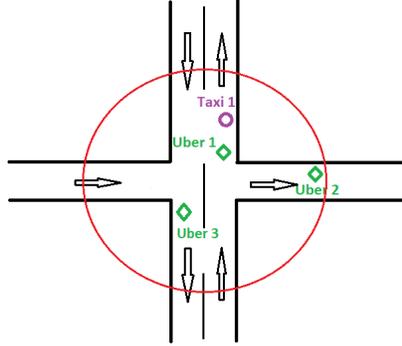
Hypothesis 5: *Everything else held constant, taxi drivers have greater incentives than Uber drivers to drive faster than other traffic on the road.*

4 Data and Sample Construction

4.1 Data

Our data combine NYC taxi trip records and Uber’s proprietary UberX trip records for two six-month periods: January to June, 2016, and July to December, 2013. Taxi trip records contain detailed information such as pick-up and drop-off time and GPS coordinates, trip distance and duration, and various fares and fees. The 2013 taxi data contain anonymized driver ID and medalion numbers, but the identifiers were subsequently removed by the TLC due to privacy concerns. UberX trip records contain similar information, plus extra information such as the surge multiplier,

Figure 2: Geographical Matching of Taxi and Uber Trips



anonymized driver ID, driver total number of trips on Uber, rider total number of trips on Uber, driver lifetime rating, and driver and rider rating for each trip.

4.2 Matching of Comparable Taxi and Uber Trips

This paper’s analyses build on a valid counterfactual construction of taxi and Uber, as will be made more clear in the empirical section. To that end, we conduct granular geographical matching of taxi and Uber trips such that the matched trips are subject to the same underlying optimal routing. In brief, we match an Uber trip and a taxi trip if they go from the same Point A to the same Point B, and begin at roughly the same time. In the remainder of this section, we detail the steps of the matching process.

Step 1: Because of the exceedingly high concentration of pick-ups and drop-offs around street intersections, we first define locations by dividing NYC into small Voronoi cells (see Figure A1) centered at street intersections, where each street intersection is approximately 100 meters from its closest neighboring intersections. Using Figure 2 as an illustration, this means that we initially match Taxi 1, Uber 1, Uber 2, and Uber 3 in the circled area.¹⁰

Step 2: We then restrict matched pick-ups to be on the same street,¹¹ because pick-ups on

¹⁰However, we show in Figure A2 that if we did nothing else, this matching criterion would yield a stark difference in the distribution of pick-ups between taxis and Uber; taxi pick-ups (purple) are more concentrated on major avenues and streets, whereas Uber pick-ups (green) are more from cross-town streets with slower traffic. A similar distribution applies to matched drop-offs as well. This may reflect the difference in drivers searching and matching with passengers. Specifically, taxi drivers mainly cruise on major avenues and streets to look for passengers, while Uber drivers more often pick up passengers at their doorsteps.

¹¹The accuracy of GPS coordinates can be adversely affected by tall buildings in an urban area. Indeed, there are more cases where taxi and Uber pick-up and drop-off GPS pinpoints fall on a building instead of on the street in Midtown Manhattan than in other areas with less concentration of tall buildings. In these cases, we assign the trip to

different streets are subject to different optimal routes even when they are going to the same destination. In Figure 2, this means that Taxi 1 will be matched with Uber 1 and Uber 3, but not with Uber 2.

Step 3: Following a similar logic as in Step 2, we further filter out matched pick-ups that follow different traffic directions of the same streets. Therefore, Taxi 1 and Uber 1 of Figure 2 remain in the matched sample.

We then apply the same filters (Steps 1-3) for drop-offs as well. For airport pick-ups and drop-offs, we match the trips based on the airport terminal.

Step 4: We further restrict matched trips to the ones that start within a short time window from each other so that they are subject to the same real-time traffic, road conditions, as well as other common factors. The time window for the main analysis is set at 15 minutes, and we apply other time windows (e.g., 5, 10, and 20 minutes) in the robustness checks.

Finally, we discover that in the raw TLC taxi trip records, there are two taxi meter vendors with about equal shares, where Vendor 1 reports trip distance to the first decimal place and Vendor 2 to the second decimal place. A casual check of dozens of randomly selected short trips in Manhattan from Vendor 1 against their Google Maps shortest distances makes us believe that this meter vendor may have rounded down the actual trip distance. Therefore, we drop Vendor 1 taxi trips from the sample. We discuss in detail the potential bias from rounding down in Section 6.4 and add Vendor 1 taxi trips back only as a robustness check.

When we use the 2016 taxi and Uber data, the matching process generates a sample of 95,357 pairs of matched trips that contains trips from 23,974 Uber drivers. The sample, summarized in Table 1, is over-represented by airport trips and short, within-Manhattan trips compared to the population of taxi trips and Uber trips, as a result of the matching criteria. *Non-local passenger* takes the value of 1 if the billing zip code of a given Uber passenger is outside of NYC, or if the billing zip code is missing, the passenger's city of most Uber trips is not NYC. The Uber surge multiplier is on average 1.11, and we subtract 1 from surge multipliers for easier interpretation of estimation results later.

The densities of Taxi-Uber distance ratios and duration ratios are separately plotted in Figure 3, across three route types: (1) routes between JFK and Manhattan where taxi fares are fixed; (2) non-

be on the street closest to its pinpoint.

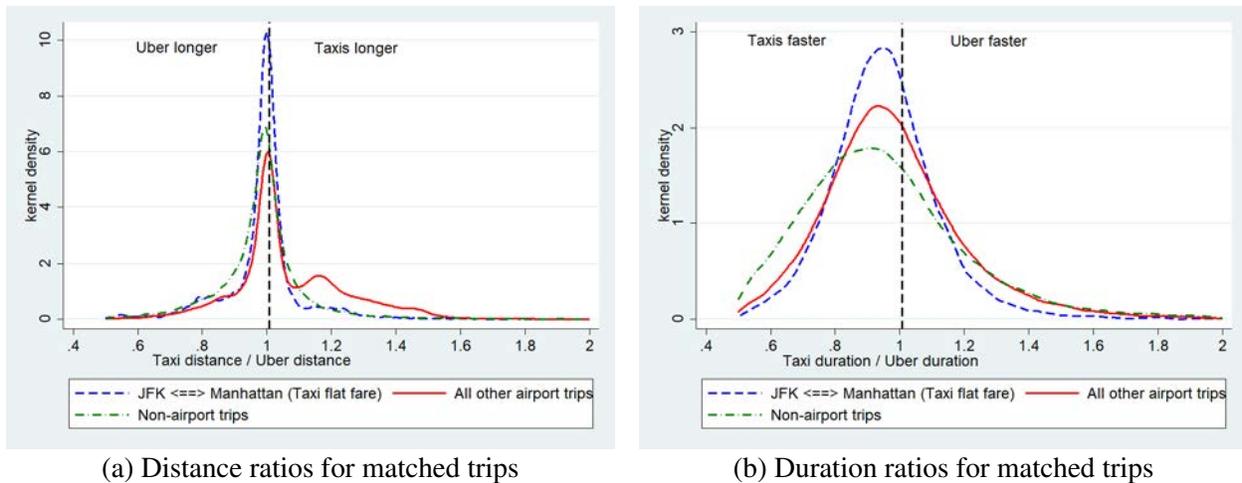
Table 1: Summary Statistics

Variable	Mean	Std. Dev.	10th	Median	90th
Taxi trip distance (miles)	8.69	5.32	1.12	9.67	16.55
Uber trip distance (miles)	8.43	5.34	1.14	9.15	16.66
Taxi distance / Uber distance	1.04	0.16	0.86	1.01	1.26
Taxi trip duration (minutes)	28.65	16.70	7.98	27	50.53
Uber trip duration (minutes)	30.03	17.30	8.72	28.32	52.77
Taxi duration / Uber duration	0.97	0.22	0.71	0.95	1.25
Airport	0.71	0.45	0	1	1
LaGuardia	0.60	0.49	0	1	1
JFK	0.10	0.30	0	0	1
Newark	0.02	0.12	0	0	0
Non-local passenger	0.51	0.50	0	1	1
Surge multiplier	0.11	0.27	0	0	0.50
Uber driver total trips	2489.19	2008.73	358	2018	5313
Uber driver lifetime rating	4.75	0.09	4.64	4.76	4.85
Uber rider total trips	115.34	170.12	5	55	293
N	95,357				
No. of Uber drivers	23,974				

Notes. The sample is from taxi and Uber trip records of NYC, January-June 2016. *Non-local passenger* takes the value of 1 if the billing zip code of a given passenger is outside of NYC, or if the missing billing zip code is missing, the passenger's city for most Uber trips is not NYC. *Surge multiplier* is the effective surge multiplier minus 1.

airport routes; and (3) all other airport routes where both taxi and Uber fares are metered, which include all LaGuardia trips, all Newark trips, and trips between JFK and NYC outer boroughs.

Figure 3: Distance and Duration Ratios of Matched Taxi and Uber Trips



The graphs reveal several interesting facts. Figure 3a shows that taxi and Uber trips are similar in trip distance for non-airport trips and JFK flat-fare trips, as illustrated by the high concentration

of the distance ratios around 1. However, taxi trips are significantly longer in distance than matched Uber trips for all other airport trips, as indicated by the second mode of the distribution around 1.15. Figure 3b shows that taxis overall arrive faster than Uber. These patterns are consistent with the implications of our theory and hypotheses: (1) compared to Uber drivers, taxi drivers route longer in routes that are anecdotally more lucrative (Hypothesis 1), and (2) taxi drivers in general drive at a greater average speed than Uber drivers (Hypothesis 5). In the next section, we turn to formal tests of driver moral hazard.

5 Empirical Analysis

5.1 Empirical Model

Let c denote taxis and u denote Uber. Equation (1) requires that for a given matched taxi-Uber pair,

$$\frac{d^c - d^u}{d_0} = a^c - a^u + x^{c*} - x^{u*} + \epsilon^c - \epsilon^u. \quad (6)$$

That is, the normalized difference in Uber and taxi routing is a function of driver skills, detours, and driver-trip random shocks. Given Hypotheses 1-4, the difference in detours is a function of various route characteristics that affect driver incentives.

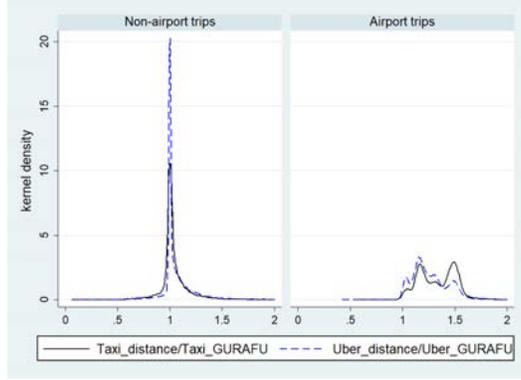
Particularly, both taxi and Uber drivers should have little incentive to detour on short trips in thick markets (e.g., trips start and end in Manhattan), because of the low marginal payoff of detouring on short trips (Hypothesis 1) and the high opportunity cost of detouring when finding another ride at drop-off is easy (Hypothesis 4). In fact, it is in drivers' best interest to take as many trips as possible to exploit the proportionately larger fixed component of the fare. On the other hand, both taxi and Uber drivers' detour incentive may increase in situations such as airport trips because long-distance trips reward detouring (Hypothesis 1).

However, the difference in detour cost of taxis and Uber, particularly the penalty costs due to different market designs, may create a difference in routing decisions when the detour payoff is high for both drivers (e.g., metered airport trips).

These assumptions are supported by the data. Figure 4 shows that both taxi and Uber trip distances closely concentrate on a measure of long-run average optimal routing given by GURAFU, Uber's internal routing engine, for non-airport trips, and not for airport trips. Therefore, we take

non-airport trips as the main control group and airport trips as one of the treatment groups in the estimation.

Figure 4: Taxi and Uber Driver Routing Weighted by Uber’s GURAFU



The theory also indicates that both taxi and Uber drivers may be more likely to detour when driving non-local passengers, due to information asymmetry (Hypothesis 2). In addition, Uber drivers may be more likely to detour in case of greater surge multipliers (Hypothesis 3).

Let r denote a given matched taxi-Uber pair. We specify and estimate the following empirical model on the sample that consists of metered trips by both taxis and Uber (that is, we exclude JFK taxi flat-fare routes):

$$\begin{aligned}
 \frac{d_r^c}{d_r^u} = & \alpha_0 + \alpha_1 \text{Airport}_r + \alpha_2 \text{Non_local}_r + \alpha_3 \text{Airport}_r \times \text{Non_local}_r \\
 & + \alpha_4 \text{Uber_surge_multiplier}_r + \alpha_5 \text{Airport}_r \times \text{Uber_surge_multiplier}_r \\
 & + \alpha_6 \log(\text{Uber_driver_total_trips}_r) + \alpha_7 \text{Uber_driver_rating}_r \\
 & + \alpha_8 \log(\text{Uber_rider_total_trips}_r) + \phi_{hw} + \eta_i + \epsilon_{ihwr}^c - \epsilon_{ihwr}^u.
 \end{aligned} \tag{7}$$

First, we include in the regression a set of route characteristics that affect the incentives and costs of detour, which include airport dummy, non-local dummy, airport times non-local, Uber surge multiplier, and airport times Uber surge multiplier. Second, we control for a set of Uber driver and Uber rider characteristics to further explain the variation in the Uber-taxi routing difference. Uber driver experience is measured by the driver’s total trips driven prior to the current trip, and it is expected to positively correlate with Uber driver routing efficiency if there is a learning-by-doing

effect. Uber driver routing efficiency is expected to positively correlate with Uber driver rating, as routing efficiency is an important metric in overall driver quality. We also account for Uber rider experience on the platform, measured by the total number of trips completed, as more experienced riders may make the trip more efficient by better communicating with the driver, choosing a more efficient pick-up or drop-off location, and so on. Let ϕ_{hw} represent fixed effects at the hour-of-week level. In some specifications, we include Uber driver fixed effects η_i to control for Uber driver-invariant unobservables.

5.2 Identification

Normality is assumed for the mean-zero driver-route random shocks ϵ_{ihwr}^u and ϵ_{ihwr}^c . For ordinary least squares estimation to produce unbiased coefficient estimates, the route characteristics in the empirical model need to be uncorrelated with the unobserved driver-route shock in the error term. This is supported by the institutional features of taxis and Uber.

For taxi drivers, the matching of passengers of certain destinations is close to randomly assigned, because (1) passengers do not select taxis as taxi cabs are *ex ante* homogeneous to passengers; (2) taxi refusal of passengers is heavily penalized by the TLC refusal law.¹² However, taxi drivers can indeed form expectations of passenger destinations and route profitability and develop their own geographical search strategies (Haggag et al. (2017), Zhang et al. (2016)), leading to a correlation between route characteristics and driver types. In this case, controlling for taxi-driver fixed effects is a good way to tease out the bias. In the absence of taxi driver IDs in 2016, we demonstrate in Section 6.2 that driver selection appears to be insignificant when the same estimation is run on 2013 data where taxi-driver fixed effects are controlled for.

Several features of the Uber platform limit the scope of endogeneity:

1. To Uber drivers, passenger assignment by the platform is virtually random by construction. Uber’s matching of drivers and passengers is mainly based on spatial proximity and dispatching efficiency, and it gives little weight to driver and rider characteristics in the matching.

2. Uber drivers have the option to cancel trip requests, but cancellation of rides is costly. Once

¹²Per the TLC refusal law, “It is against the law to refuse a person based on race, disability, or a destination in New York City. A taxicab driver is required to drive a passenger to any destination in the five boroughs.” Riders are encouraged to make a refusal complaint by calling 311. According to Haggag et al. (2017), “In 2009 the refusal punishment was \$200-\$350 for a first offense, \$350-\$500 and a possible 30-day license suspension for a second, and a mandatory license revocation for a third offense. The TLC received about 2,000 formal complaints per year in 2009 and 2010.”

assigned a rider, the driver cannot see the rider’s destination on the application until picking up that rider, which makes it difficult for drivers to “cherry pick” passengers before accepting a trip request. Moreover, frequent and suspicious ride cancellation is penalized on Uber, often in the form of warning, “time out”, or even deactivation. In addition, it is difficult for a driver to form expectations on the next rider’s profitability, making cancellation of the current ride risky and rare.

3. With volatile and unpredictable surge multipliers¹³, it is not in the driver’s best interest to chase the surge, at the cost of forgone earnings from trip requests missed or declined.¹⁴ In addition, the Uber app no longer shows the surge hot spots after driver’s accepting a ride, which prevents drivers from canceling the current non-surge ride in order to get a high-surge ride.

Taken together, these institutional details suggest that the correlation between route types and unobserved driver-route shocks is at best limited. To further reduce the potential bias, we control for Uber driver fixed effects in some specifications.

5.3 Regression Results

5.3.1 Baseline: Non-airport Routes and Metered Airport Routes

Table 2 shows the main regression results, which we performed on all trips that are both metered on Uber and taxis (that is, all routes except JFK taxi flat-fare routes).

In Specification 1, we find that the taxi-Uber distance ratio for an airport trip is on average 7.4% larger than for a non-airport trip, other things held constant. The effect of non-local passengers is positive and strong on airport trips, suggesting that taxi drivers tend to exploit the information advantage more than Uber drivers on airport routes taken by non-locals. However, it is important to note that “non-local” may be subject to measurement error, as we use the information of the Uber rider as a proxy for whether the taxi passenger is a local or not.¹⁵ Lastly, we find that Uber drivers tend to detour more relative to taxi drivers on airport trips with a surge price in effect. However, the effect of the surge multiplier is not found on non-airport trips. In addition, we find that Uber driver rating is positively correlated with the relative detour, suggesting that Uber drivers

¹³As shown in Lam and Liu (2017), Uber surge multipliers are extremely volatile and hard to predict by highly granular location-time fixed effects.

¹⁴Uber drivers commonly agree on this view, based on our conversations with Uber drivers in New York City and Boston, Massachusetts.

¹⁵We caution that the scope of measurement error of “non-local” should be smaller on airport routes than on non-airport routes, because it is more likely that the taxi rider and the Uber rider are either both locals or both non-locals when they head to the airport from the same specific place at the same time (e.g., a hotel).

Table 2: Taxi and Uber Driver Routing Difference

D.V. = Taxi dist. / Uber dist.	(1)	(2)	(3)
Airport	0.074*** (0.002)	0.074*** (0.003)	0.074*** (0.003)
Non_local	-0.002 (0.002)	-0.003 (0.002)	0.001 (0.003)
Airport \times Non_local	0.018*** (0.003)	0.019*** (0.003)	0.013*** (0.003)
Surge_multiplier	0.002 (0.003)	0.001 (0.003)	-0.000 (0.004)
Surge_multiplier \times Airport	-0.026*** (0.005)	-0.023*** (0.005)	-0.022*** (0.005)
Log (Uber_driver_total_trips)	0.000 (0.001)	-0.000 (0.001)	0.005** (0.002)
Uber_driver_rating	0.040*** (0.007)	0.039*** (0.007)	0.000 (.)
Log (Uber_rider_total_trips)	0.001 (0.000)	0.001 (0.000)	0.001 (0.001)
Hour of week FE	No	Yes	Yes
Uber driver FE	No	No	Yes
N	86,627	86,627	86,627
R^2	0.055	0.060	0.376

Notes. For all specifications, standard errors are cluster-robust at the hour-of-week level; *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

with better ratings are more efficient at routing, or drivers who take inefficient routes are more likely to get low ratings.

The estimated effects hardly change when we consider hour-of-week fixed effects in Specification (2). Further controlling for Uber driver fixed effects does not appear to affect the point estimates much, as shown in Specification (3). Interestingly, the effect of Uber driver total trips becomes strong, suggesting a routing improvement due to accumulating driving experience (Cook et al. (2018) document a similar learning-by-doing effect among Uber drivers).

5.3.2 Non-airport Routes and JFK Taxi Flat-fare Routes

It is important to note that the analysis so far identifies moral hazard in relative terms; the effect of 7.4% describes how taxi-driver routing in metered airport routes, referenced by Uber driver routing in these routes, compares with that in non-airport routes. Therefore, taxi drivers may in fact detour more in absolute terms (i.e., compared to GPS optimal routes) if matched Uber drivers also detour to some extent.

Table 3: Taxi and Uber Driver Routing Difference: JFK Taxi Flat-fare Routes and Non-airport Routes

D.V. = Taxi dist. / Uber dist.	(1)	(2)	(3)
JFK flat fare	0.003 (0.003)	0.002 (0.003)	-0.002 (0.004)
Non_local	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.003)
JFK flat fare \times Non_local	0.001 (0.003)	0.002 (0.004)	0.006 (0.005)
Surge_multiplier	0.002 (0.003)	0.000 (0.003)	-0.001 (0.004)
Surge_multiplier \times JFK flat fare	-0.007 (0.008)	-0.008 (0.008)	-0.003 (0.013)
Log (Uber_driver_total_trips)	0.002*** (0.001)	0.002*** (0.001)	0.000 (0.003)
Uber_driver_rating	0.029*** (0.008)	0.028*** (0.008)	0.000 (.)
Log (Uber_rider_total_trips)	0.001 (0.000)	0.001 (0.000)	0.000 (0.001)
Hour of week FE	No	Yes	Yes
Uber driver FE	No	No	Yes
N	36,587	36,587	36,587
R^2	0.001	0.007	0.520

Notes. For all specifications, standard errors are cluster-robust at the hour-of-week level; *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

Noting this, we leverage the JFK taxi flat-fare routes to identify the extent to which Uber drivers detour. These routes are an ideal scenario because taxi drivers do not have any detour incentive when the trip fare is flat, as drivers only incur costs for detouring without gaining any benefit. In particular, we estimate the same regressions as Equation 7 on the sample that consists of non-airport routes and routes between Manhattan and JFK.

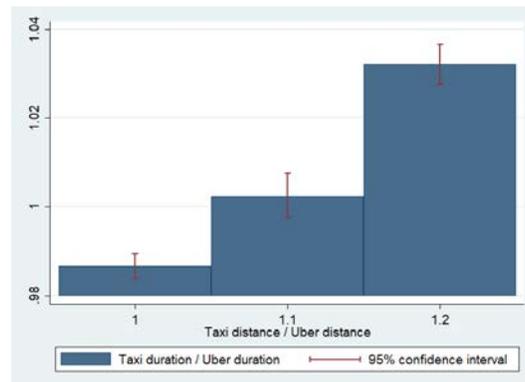
Shown in Table 3, the effect of JFK flat-fare dummy on the taxi-Uber distance ratio is not statistically different from 0, which suggests that Uber drivers do not detour on routes between JFK and Manhattan. Similarly, Uber drivers do not route less efficiently when the passenger is non-local or when the surge multiplier is high, for both JFK flat-fare routes and non-airport routes. To the extent that JFK flat-fare routes are significantly longer in distance than LaGuardia trips and thus warrant a higher detour benefit, no evidence of Uber driver detour on JFK routes implies no Uber driver detour on LaGuardia routes. Therefore, the effect of 7.4% in the baseline model appears to be an accurate estimate of the average taxi-driver detour.

The regression results in this section complement the baseline results. Combining these two sets of analyses yields our main empirical findings: (1) taxi drivers detour on airport trips that are metered, particularly for non-local passengers; (2) Uber drivers generally do not detour, except in cases of surged airport trips.

5.4 Distance vs. Time: Do Longer Routes Save Passengers Time?

Our analyses so far demonstrate that taxi drivers tend to add miles to the fastest routes (according to GPS), which would naturally add additional passenger travel time. However, it is possible that Uber drivers do not always follow the GPS. Or, taxi drivers in some cases possess superior routing knowledge than the GPS. In either of these above-mentioned cases, taxi drivers can indeed save passengers time by taking a longer route. We next show that the data are not compatible with this alternative hypothesis, but rather they lend more nuanced support to our main hypothesis of driver moral hazard.

Figure 5: Taxis on Average Arrive Later When on Longer Routes, Compared to Uber



First, we show in Figure 5 that when detouring (i.e., distance ratio > 1), taxi drivers on average finish trips *later* than matched Uber drivers. In addition, the relative travel time difference increases when taxi detour is longer.

Next, we investigate the way in which drivers detour via a case study and show that (1) taxi drivers appear to detour on airport routes by choosing a different bridge/tunnel than suggested by the optimal route; (2) detours indeed lead to longer travel times; (3) drivers with more route-specific experience seem to detour more.

The case study focuses on routes between LaGuardia Airport and an Midtown Manhattan area¹⁶

¹⁶This area consists of three NYC Neighborhood Tabulation Areas (NTAs): Midtown-Midtown South, Turtle Bay-

(see Figure 6), for two main reasons: First, this area has a high volume of taxi and Uber activities — approximately 37% of all LaGuardia taxi trips in 2016 either started or ended in this area; Second, the choice set of routes is relatively small and clear — the route in the middle via Queensboro Bridge is a toll-free route, usually the shortest and busiest among the three routes (we call this route S), whereas the other two routes have tolls of the same amount¹⁷, with the top route via FDR Drive generally longer in distance (we call this route L) than the bottom route via the Midtown Tunnel (we call this route M).

Figure 6: Alternative Routes Between Midtown Manhattan and LaGuardia

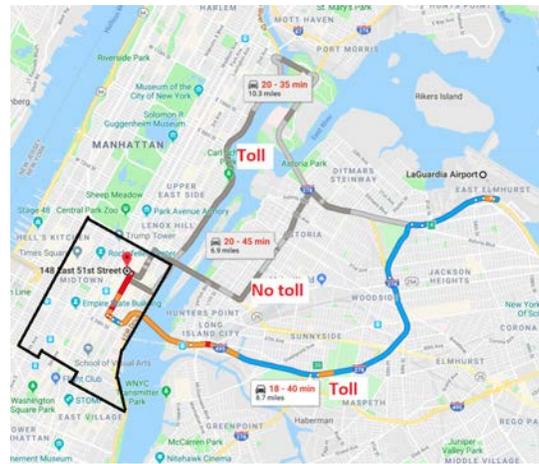
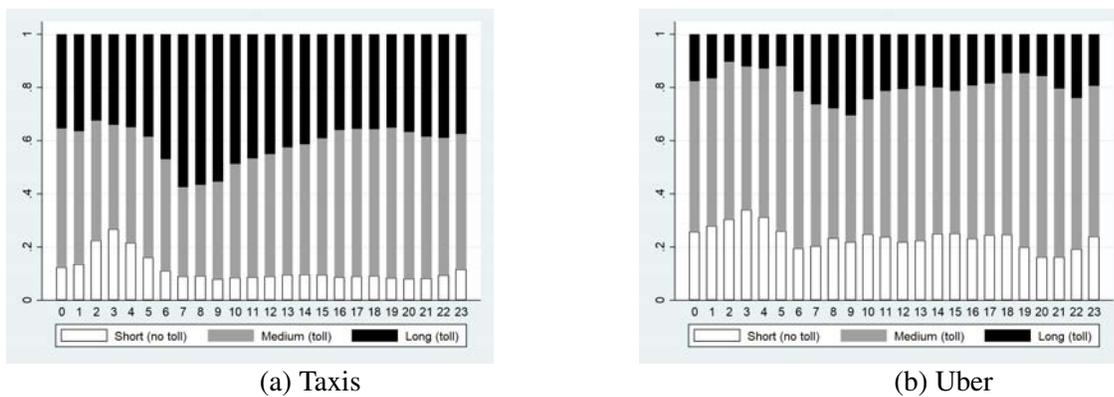


Figure 7: Taxi and Uber Driver Route Choices, Midtown to LaGuardia, by Hour



Taxi and Uber trip distance indeed exhibits a clear bimodal pattern for these two tolled routes

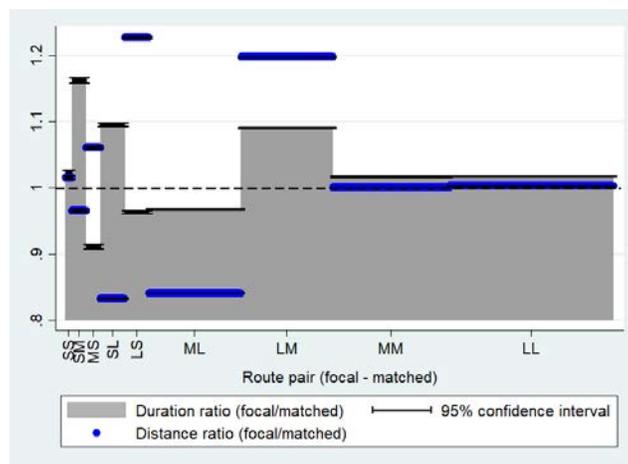
East Midtown, and Murray Hill-Kips Bay

¹⁷In 2016, the cash rate of the tolls was \$8 for taxi-like vehicles. Taxi and Uber drivers normally pay the discounted rate of \$5.54 by using an E-Z Pass.

(Figure A3), suggesting that the route choices are discrete. Therefore, using data on tolls as well as trip distance, we can identify, or at least proxy with good precision, which route drivers took.

Taxi drivers are overall more likely to take the long route than Uber drivers (Figure 7). Because Uber drivers generally drive slower than taxi drivers, we need to investigate whether taxi drivers save passengers time by choosing the long route, compared to other taxi drivers on similar routes. Therefore, we match taxi trips with other taxi trips, following the same approach as we did for taxi and Uber trips.¹⁸

Figure 8: Taxi Drivers on Average Cost Passenger Time on Long Routes



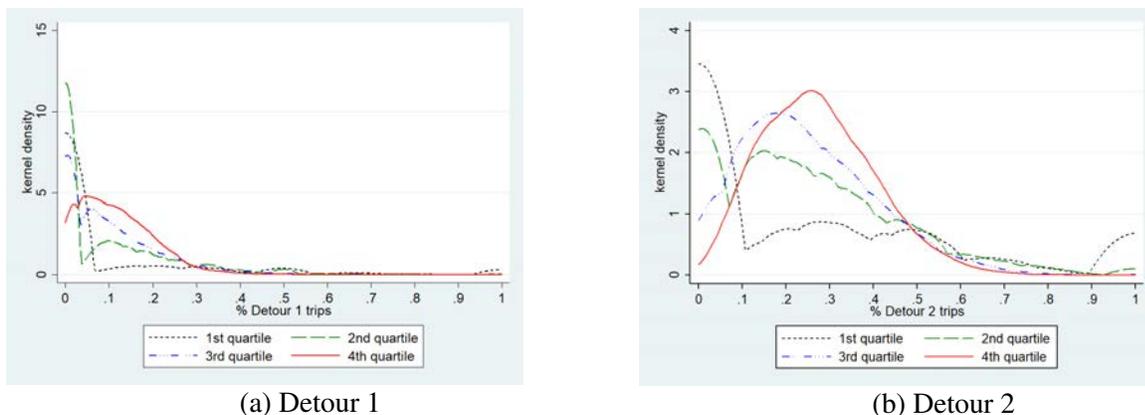
Notes. This figure is based on the matching of taxi drivers with other taxi drivers on similar routes between Midtown and LaGuardia in 2016. The shares of these nine route pairs are represented by their horizontal bars in the graph.

Figure 8 shows how focal taxi drivers’ travel time compares with that of the matched taxi drivers across situations when they take the same or different routes, where the occurrence of these nine situations is represented by the horizontal bars. For 52% of the time, the focal taxi driver and the matched taxi driver take the same route (SS, MM, and LL), and they spend about the same amount of time driving. When the focal driver takes S while the matched driver takes M or L (14% of the time), the matched driver indeed performs better in travel time (i.e., travel duration is shorter). However, in the case of ML and LM (34% of the time), drivers tend to waste passengers time instead of saving them time when choosing the long route. Given the shares of these situations, we conclude that taxi drivers on average cost passengers more time by choosing longer routes.

¹⁸Since we focus on airport routes here, we relax the matching process to be just Step 1 so that more trips can be matched without hurting precision much.

Finally, we provide evidence that the same-route experience is correlated with more detour. Using 2013 data, we split drivers into quartiles by their total number of trips between Midtown and LaGuardia. Figure 9a shows that drivers with more trips completed on this particular route (Midtown-LaGuardia) tend to detour more, as measured by Detour 1 — the incidence when the focal taxi driver takes a longer route than the matched taxi driver (i.e., LM and MS), *and* the focal taxi driver arrives later than the matched driver. Figure 9b shows a more pronounced pattern with the definition Detour 2, where the focal driver both logs more distance and more time. Therefore, this evidence further provides support for moral hazard, as it shows that inefficient routing generally comes from experienced drivers who knowingly take the long route rather than from less-experienced drivers.

Figure 9: Drivers with More Same-route Experience Tend to Detour More



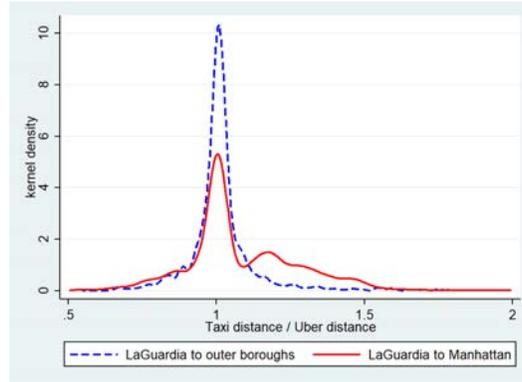
6 Competing Explanations and Robustness

We next examine three competing explanations and conclude that none of them is likely to have caused the observed patterns in the taxi-Uber routing difference.

6.1 A GPS Story?

First, the Uber-taxi routing difference can be driven by the difference in technology adoption, as Uber drivers generally use GPS navigation, while NYC taxi drivers mostly rely on their own judgment. For example, the effect of airport can also be rationalized if the benefits of GPS over taxi-driver knowledge are greater for these routes than for non-airport routes.

Figure 10: LaGuardia Trips to Manhattan vs. to Outer Boroughs



The regression results on the sample of JFK flat-fare routes directly speak against this possibility. JFK routes are among the longest routes of NYC, yet no significant difference in Uber and taxi routing is observed on these routes. However, it is likely that taxi drivers are more familiar with JFK routes, compared to other long routes that are infrequently traveled; as a result, JFK routes are not a conclusive counterargument. Noting this, we split LaGuardia pick-ups into trips to Manhattan and trips to outer boroughs. We show in Figure 10 that LGA-to-Manhattan routes which taxi drivers should be more familiar with, are exactly those where taxi drivers route *longer* in distance. Given the large share of LGA trips in our sample, the observed taxi-Uber routing difference cannot reasonably be interpreted as solely a GPS story.

6.2 Driver Selection?

Selection on the Intensive Margin As discussed in the identification, the unobserved driver types in the error term may correlate with route characteristics. It is revealing that we see no significant changes in the coefficient estimates when Uber-driver fixed effects are controlled for in the main analysis. However, we need to explore whether the effects remain when taxi-driver fixed effects are accounted for.

We repeat the baseline regression analysis on the 2013 data where taxi driver IDs are available. In the sample construction, we relax the matching criteria to only Step 1 and Step 4 with a time window of 30 minutes, because following the original matching procedure would lead to a sample size too small for identification, due to the small market share of Uber in 2013. The final sample consists of 16,989 matches on metered routes with 10,085 taxi drivers, 4,410 of which have more than one trip in the sample.

Table 4: Taxi and Uber Driver Routing Difference, 2013

D.V. = Taxi dist. / Uber dist.	(1)	(2)	(3)
Airport	0.105*** (0.007)	0.112*** (0.007)	0.101*** (0.011)
Non_local	-0.007** (0.003)	-0.010*** (0.003)	-0.009 (0.006)
Airport \times Non_local	0.019*** (0.007)	0.022*** (0.007)	0.023* (0.013)
Surge_multiplier	-0.001 (0.006)	-0.002 (0.004)	-0.006 (0.007)
Surge_multiplier \times Airport	-0.027** (0.011)	-0.021** (0.010)	-0.035 (0.024)
Log (Uber_driver_total_trips)	-0.002** (0.001)	-0.003** (0.001)	-0.003 (0.002)
Uber_driver_rating	0.056*** (0.020)	0.052** (0.021)	0.048 (0.039)
Log (Uber_rider_total_trips)	0.000 (0.001)	0.000 (0.001)	0.001 (0.002)
Hour of week FE	No	Yes	Yes
Taxi driver FE	No	No	Yes
N	16,989	16,989	16,989
R^2	0.082	0.102	0.644

Notes. The regression samples consist of matched pairs of taxi and Uber trips on metered airport routes and non-airport routes, using 2013 data. For all specifications, standard errors are cluster-robust at the hour-of-week level; *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

We find no evidence that driver selection into routes can explain our results. As shown in Table 4, the estimates are in general of the same sign and similar size as those of the 2016 analysis. Particularly, the effect of airport on distance ratio remains strong and large when taxi-driver fixed effects are controlled, and the effect of non-local on airport routes is borderline significant.

Selection on the Extensive Margin Drivers of different types may select differently into being taxi and Uber drivers. If this was the case, then the observed moral hazard would be an artifact of the driver type distributions of taxis and Uber. Not being able to directly observe driver types, we cannot rule out this possibility. However, we shed light on the extent of driver behavioral change by following 1,892 former taxi drivers who switched to Uber. Given that types are persistent, former taxi drivers who used to detour might be expected to continue their detour behavior on Uber in similar situations.

Table A1 shows the regression results of distance ratios on the sample of non-airport routes and metered airport routes for these 1,892 former taxi drivers and their matched Uber drivers in 2013.

The estimates confirm the moral hazard behavior of these former taxi drivers in 2013. Table A2 shows the regression results of distance ratios separately for the sample of metered airport routes and JFK flat-fare routes (both with non-airport routes as the benchmark), for these former-taxi, now-Uber drivers and their matched taxi drivers in 2016. The results show little evidence that these drivers continue to detour on profitable routes.

However, the behavior of the switchers may just reflect a common trend, for example, drivers may have become more honest over the years due to some unobserved factors. One good way to rule this out is by tracking taxi drivers who remained taxi drivers from 2013 to 2016; however, the exercise is not feasible due to missing taxi driver IDs in 2016. Nonetheless, we show in Figure A4 that 234 long-standing Uber drivers do not appear to have become more honest from 2013 to 2016, as suggested by their routing compared to taxi routing in the same time periods on JFK taxi flat-fare routes, where taxis are the “no detour” benchmark.

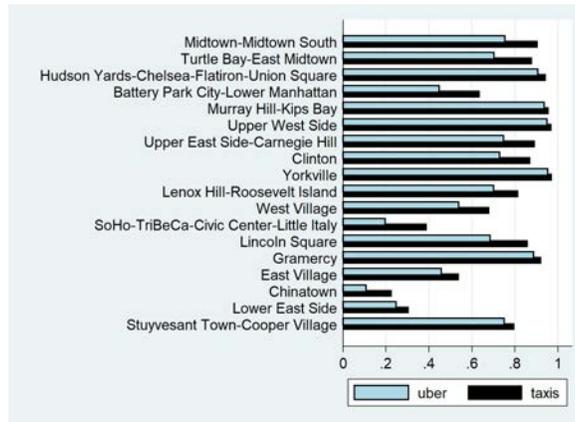
These analyses provide strong evidence that drivers adapt to changes in technology and market design via behavioral change, which is consistent with the driver moral hazard finding of this paper.

6.3 Do Taxi Drivers Route Longer to Avoid Tolls for Passengers?

There are cases when drivers can make passengers better off by choosing longer routes that minimize tolls (see Figure A5 for an example). However, a necessary condition for this to be true is that taxi drivers take toll-free routes more often than Uber drivers. Figure 11 shows precisely the opposite: Across major neighborhoods of Manhattan¹⁹, taxi drivers are *more* likely to take toll roads than Uber drivers do on their way to or from LaGuardia. Therefore, based on the data, we reject toll saving as the main explanation for taxi drivers’ detouring.

¹⁹These neighborhoods constitute Manhattan Core, where Manhattan Core is defined as roughly the part of Manhattan below the north edge of Central Park.

Figure 11: Taxi and Uber Shares of Tolled Trips between LaGuardia and Manhattan Neighborhoods



Notes. The figure plots the shares of toll roads, separately for taxis and Uber, between major Manhattan neighborhoods and LaGuardia, using 2016 data. These Manhattan neighborhoods are ordered by taxi trip shares.

6.4 Robustness Checks

We perform two sets of robustness checks to strengthen the identification of driver moral hazard. In the main analysis of driver detour, we constrained the matched taxi and Uber trips to be 15 minutes apart. In the first set of robustness checks, we perform the same analysis using alternative time windows, namely 5, 10, 20, and 30 minutes. As shown in Table A3, the estimated effects are stable and consistent across time window lengths. To the extent that trips within a narrower time range are more likely to be subject to the same real-time traffic, and thus better approximate the experimental ideal, significant effects of similar size even when using a time window as short as 5 minutes greatly enhance our identification.

Recall that our main sample contains only taxi trips reported by Vendor 2, because Vendor 1’s meter system appears to round down trip distance to the nearest first decimal place. Taxi trips may appear to be shorter because of the rounding, and the downward bias is greater in taxi-Uber distance ratios of shorter routes²⁰. This implies that the same regression analysis on Vendor 1 sample should yield an upward bias in the coefficient estimate of airport trips, instead of the opposite. In Table A4, we separately estimate the main regressions using Vendor 1 only, and then both Vendor 1 and

²⁰Consider a pair of matched Uber and taxi trips, where the Uber trip is 1 mile and the taxi trip is 0.95 miles, but the reported taxi trip length is rounded to 0.9 miles. Then the distance ratio would be 0.9 instead of 0.95, with a downward bias of -0.05. For a 9.95-mile taxi trip rounded to 9.9 miles with a matched 10-mile Uber trip, the downward bias is only -0.005 (0.99-0.995). Thus, the same amount of rounding error leads to proportionately greater downward bias on shorter routes.

Vendor 2, and compare these with the main regression results using only Vendor 2. We indeed find the upward bias, where the estimated effect of airport on distance ratio is 9.3% on Vendor 1 sample, 1.9% more than the effect on Vendor 2 sample. We find this upward bias, instead of a downward one, consistent with our main findings.

7 Discussion

7.1 Mechanisms

In this section, we discuss two mechanisms that account for the observed routing difference between taxi and Uber drivers.

The first mechanism is the set of technology-enabled incentive devices implemented by the Uber platform but not by taxis. These incentive devices include tech-aided monitoring and verification, tech-enabled rider rating of drivers, and tech-aided conflict resolution. Each of these essentially makes the cost function of moral hazard “steeper” for Uber than for taxis.

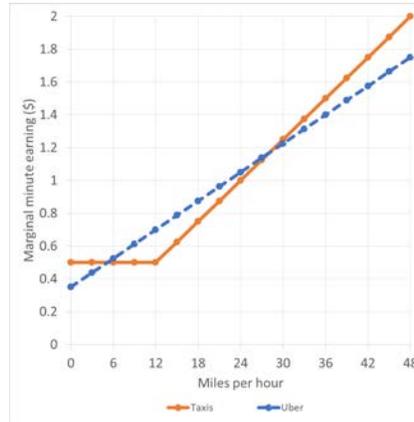
Our empirical results imply that these incentive devices (monitoring, rating, conflict resolution) enhance market transparency in most cases, as Uber drivers do not appear to detour on JFK airport routes, where the potential gain from detouring is large. Strikingly, Uber drivers do tend to detour on airport trips when surge pricing is high, suggesting that these incentive devices are not binding. Instead, as the agency model predicts, Uber drivers are willing to make a trade-off when a sufficiently high surge can offset the possible penalty.

One necessary condition for a working rating system to penalize strategic behavior is the negative correlation between passengers’ ratings to the drivers and driver routing inefficiency. This negative correlation is robustly present in the data, as shown in Table A5. One interesting nuance is that conditional on the Uber ride is shorter than the matched taxi ride, the shorter the Uber trip is, the less likely the Uber driver will get a high rating. One likely reason for this is that passengers dislike off-GPS routing and thus give low ratings even when Uber drivers found a shorter route. It seems plausible that passengers cannot easily assess whether a driver’s deviations from the prescribed GPS route are due to superior information used to shorten the route or an effort to extract a higher fare with a longer route. Therefore, tech-enhanced monitoring deters driver opportunistic behavior, yet on the other hand it may also create constraint for driver discretion.

Pricing is another important mechanism that predicts driver routing behavior. This is most

clearly reflected in taxi driver routing efficiency on flat-fare JFK routes, where taxi drivers optimize routes as residual claimants. When taxi fares are metered as a two-part tariff, taxi drivers tend to detour on airport routes mainly because the variable part of the fare can justify the detour. In addition, taxi drivers' detour payoff is greater than that of Uber drivers, as the taxi pricing formula puts a larger weight on distance than Uber pricing does.

Figure 12: Driver Marginal Minute Earning Across Travel Speeds



Furthermore, perhaps the pricing scheme is of first-order importance in explaining taxi driver travel speed. We document in the main analysis that taxi drivers, although detouring in some cases, in general drive at a faster speed than Uber drivers. This is expected because speeding is rewarded more for taxi drivers than for Uber drivers. In Figure 12, we compute driver marginal minute earnings at various traffic speeds, separately for taxis and Uber, following their pricing formulas. All the computations are net of the fixed component of the fare (i.e., \$2.5 for taxis and \$2.55 for Uber). An interesting divergence appears. For example, taxi-driver per-minute earning increases by 4 times from 12 miles per hour to 48 miles per hour (from \$0.5/min to \$2/min), while Uber-driver per-minute earning only increases by 2.5 times (from \$0.7/min to \$1.75/min). This difference stems from the weight given to trip distance in the pricing formula, where taxi distance is marginally more rewarding than that of Uber. One artifact of the taxi pricing schedule is that when traffic is flowing at about 12 miles per hour, a NYC taxi driver would earn a slightly higher fare by alternating between stopping and driving at 24 miles per hour.

7.2 Welfare Calculation

Detouring leads to welfare loss in the form of lost passenger time. Using the within-taxi matched sample, we estimate a 1.6% increase in travel time for the 7.4% increase in trip distance by taxi drivers on metered airport trips. For an average NYC metered airport trip that takes 32.7 minutes, the time waste is about 0.5 minutes. Given that there are 18,000 metered airport trips per day in NYC, a back-of-envelope calculation leads to an efficiency loss of 9,000 passenger minutes, or 150 passenger hours per day.

Passengers also pay more when drivers detour, although the payment is an income transfer to drivers rather than an efficiency loss. The detour on an average airport trip is about 0.7 miles, which translates into 12,600 added miles per day. This is equal to an extra passenger payment of \$37,800 per day (12,600 miles \times \$2.5 per mile \times (1+20%), assuming a 20% tip).

Uber drivers tend to detour on surge-priced airport trips (1,630 such trips per day). Given the estimated effect, we calculate an added travel time of 771 minutes, or 12.85 passenger hours per day. If we combine the added miles and minutes, the extra passenger payment is about \$1,200 per day. Therefore, the efficiency loss (in lost passenger hours) due to Uber driver detour is 10 times smaller than that of taxis. Passengers' extra payment is 30 times smaller for Uber than for taxis.

7.3 Mind vs. Machine

“GPS routes are slower in Manhattan, so I don’t use it.”

— Gurpreet Singh, NYC taxi driver, interviewed on July 9th, 2018.

The data show that taxi drivers are more efficient at routing than Uber drivers on short, non-airport routes. For short routes within Manhattan Core, taxi trip distance is on average 98.5% that of Uber, and the difference is statistically significant at the 1% level. This is evidence that human navigation can perform at least as well as the technology in dense markets.

One way that drivers can outperform GPS is by having more up-to-date information on the road networks and conditions, for example, temporary road closures, upcoming sporting events, or undocumented shortcuts. Another possibility is that experienced taxi drivers can suggest a better drop-off point than the exact address given by the passenger, based on the driver's extensive experience. For example, the driver might suggest dropping off the passenger on the opposite side of the street in order to avoid unnecessary travel. This is confirmed by an interview with Loai

Yousef ²¹, an NYC Uber and Lyft driver, who stated, “Sometimes the Uber GPS map has mistakes. Sometimes it makes the driver do a U-turn to arrive at the exact address even though it would be easy for rider to just cross the street. Taxis drivers often drop passengers off a short distance from exact address.”

NYC taxi driver expertise in routing should not be surprising, because as residual claimants, they are strongly motivated to learn the routes, optimize their routing, and take initiative when they can. In contrast, Uber drivers might not be as motivated to use their discretion, even in cases when they do possess better information than the GPS. The reason is that off-GPS routing might come across as suspicious behavior to the riders, which can result in bad ratings and complaints (as we witnessed in Table A5). Loai Yousef told us that “Uber passengers tend to want driver to go to the exact address even if it’s wasteful.” Therefore, the use of GPS, coupled with the monitoring and rating systems, can limit the incentives for human knowledge accumulation, as well as initiative and discretion.

8 Conclusion

In this paper, we study how a technology platform affects moral hazard and service quality. We provide evidence from the taxi and Uber setting in the form of driver routing choices from identical start and end points. By analyzing trip-level data from NYC, we find that taxi drivers tend to detour more relative to Uber drivers on airport routes and especially routes taken by non-local riders, while Uber drivers tend to detour more relative to taxi drivers on airport routes with greater surge pricing. In other words, as predicted by our theoretical model, the Uber technology platform and pricing scheme reduce driver moral hazard behavior where taxi moral hazard return is high, but at the same time create other margins of driver moral hazard. That said, we also find evidence that the incentives for creating and using driver routing expertise may be reduced by the Uber platform, relative to reliance on technologies such as GPS. These findings are consistent with the agency theory, while driver selection or the difference between taxis and Uber in navigation technologies is unlikely to have caused the observed patterns in driver routing.

We highlight the importance of incentive devices and pricing schemes as the underlying mechanisms for the observed taxi-Uber routing difference. As such, our findings have implications for

²¹Loai Yousef was interviewed on July 9th, 2018.

regulators and industry participants. For the TLC, our results provide support for the development and implementation of smart phone applications that handle taxi dispatching and matching with passengers, digital payment, and passenger monitoring. Also, it is important for the TLC to re-evaluate the current pricing scheme that rewards taxi cab speeding as well as the impacts of alternative pricing structures. For digital platforms such as Uber, our findings suggest an opportunity for machine-learning-based techniques to detect driver opportunistic behavior, which may further enhance market transparency and trust building.

There is growing body of research on the digital disruption and, in particular, the potential for digital platforms to mitigate moral hazard. We identify sizable efficiency gains due to reduced agency problems. Our study provides models and documents the relationship between technology and incentive design, creating a challenge and opportunity for new contractual relationships on digital platforms.

References

- Angrist, J. D., S. Caldwell, and J. V. Hall (2017). Uber vs. taxi: A drivers eye view. Technical report, National Bureau of Economic Research.
- Aral, S., E. Brynjolfsson, and L. Wu (2012). Three-way complementarities: Performance pay, human resource analytics, and information technology. *Management Science* 58(5), 913–931.
- Athey, Susan, J. C. C. and D. Knoepfle (2018). Service quality in the gig economy.
- Bakos, J. Y. (1997). Reducing buyer search costs: Implications for electronic marketplaces. *Management science* 43(12), 1676–1692.
- Balafoutas, L., A. Beck, R. Kerschbamer, and M. Sutter (2013). What drives taxi drivers? a field experiment on fraud in a market for credence goods. *Review of Economic Studies* 80(3), 876–891.
- Balafoutas, L., R. Kerschbamer, and M. Sutter (2017). Second-degree moral hazard in a real-world credence goods market. *The Economic Journal* 127(599), 1–18.
- Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics* 117(1), 339–376.
- Brown, J. R. and A. Goolsbee (2002). Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of political economy* 110(3), 481–507.
- Brynjolfsson, E., Y. Hu, and M. D. Smith (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science* 49(11), 1580–1596.
- Brynjolfsson, E. and M. D. Smith (2000). Frictionless commerce? a comparison of internet and conventional retailers. *Management science* 46(4), 563–585.
- Buchholz, N. (2015). Spatial equilibrium, search frictions and efficient regulation in the taxi industry. Technical report, University of Texas at Austin.

- Castillo, J. C., D. Knoepfle, and G. Weyl (2017). Surge pricing solves the wild goose chase. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pp. 241–242. ACM.
- Chen, M. K., J. A. Chevalier, P. E. Rossi, and E. Oehlsen (2017). The value of flexible work: Evidence from uber drivers. Technical report, National Bureau of Economic Research.
- Chen, M. K. and M. Sheldon (2016). Dynamic pricing in a labor market: Surge pricing and flexible work on the uber platform. In *EC*, pp. 455.
- Cohen, P., R. Hahn, J. Hall, S. Levitt, and R. Metcalfe (2016). Using big data to estimate consumer surplus: The case of uber. Technical report, National Bureau of Economic Research.
- Cook, C., R. Diamond, J. Hall, J. A. List, P. Oyer, et al. (2018). The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers. *Upubliceret paper. Tilgængelig på: <https://web.stanford.edu/~diamondr/UberPayGap.pdf>*. Besøgt, 26–04.
- Cramer, J. and A. B. Krueger (2016). Disruptive change in the taxi business: The case of uber. *American Economic Review* 106(5), 177–82.
- Duflo, E., R. Hanna, and S. P. Ryan (2012). Incentives work: Getting teachers to come to school. *American Economic Review* 102(4), 1241–78.
- Filippas, A., J. J. Horton, and J. Golden (2018). Reputation inflation. In *Proceedings of the 2018 ACM Conference on Economics and Computation*. ACM.
- Forman, C., A. Ghose, and A. Goldfarb (2009). Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management science* 55(1), 47–57.
- Frechette, G. R., A. Lizzeri, and T. Salz (2016). Frictions in a competitive, regulated market evidence from taxis.
- Gans, J. S., A. Goldfarb, and M. Lederman (2017). Exit, tweets and loyalty. Technical report, National Bureau of Economic Research.
- Goldfarb, A. and C. Tucker (2017). Digital economics. Technical report, National Bureau of Economic Research.
- Greenwood, B. N. and S. Wattal (2017). Show me the way to go home: An empirical investigation of ride-sharing and alcohol related motor vehicle fatalities. *MIS Quarterly* 41(1).
- Haggag, K., B. McManus, and G. Paci (2017). Learning by driving: Productivity improvements by new york city taxi drivers. *American Economic Journal: Applied Economics* 9(1), 70–95.
- Hall, J., C. Kendrick, and C. Nosko (2015). The effects of ubers surge pricing: A case study. *The University of Chicago Booth School of Business*.
- Hall, J. V., J. J. Horton, and D. T. Knoepfle (2017). Labor market equilibration: Evidence from uber. Technical report, Working Paper, 1–42.
- Hall, J. V. and A. B. Krueger (2015). An analysis of the labor market for ubers driver-partners in the united states. *ILR Review*, 0019793917717222.
- Hubbard, T. N. (2000). The demand for monitoring technologies: the case of trucking. *The Quarterly Journal of Economics* 115(2), 533–560.
- Hui, X., M. Saeedi, Z. Shen, and N. Sundaresan (2016). Reputation and regulations: evidence from ebay. *Management Science* 62(12), 3604–3616.
- Klein, T. J., C. Lambertz, and K. O. Stahl (2016). Market transparency, adverse selection, and moral hazard. *Journal of Political Economy* 124(6), 1677–1713.
- Lam, C. T. and M. Liu (2017). Demand and consumer surplus in the on-demand economy: the case of ride sharing.

- Lee, M. K., D. Kusbit, E. Metsky, and L. Dabbish (2015). Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 1603–1612. ACM.
- Liang, C., Y. Hong, and B. Gu (2016). Effects of it-enabled monitoring on labor contracting in online platforms: Evidence from a natural experiment.
- Liu, T., E. Vergara-Cobos, and Y. Zhou (2017). Pricing schemes and seller fraud: Evidence from new york city taxi rides.
- Nagin, D. S., J. B. Rebitzer, S. Sanders, and L. J. Taylor (2002). Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review* 92(4), 850–873.
- Pierce, L., D. C. Snow, and A. McAfee (2015). Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Science* 61(10), 2299–2319.
- Rajgopal, S. and R. White (2015). Cheating when in the hole: The case of new york city taxis.
- Reimers, I., B. R. Shiller, et al. (2018). Proprietary data, competition, and consumer effort: An application to telematics in auto insurance. Technical report.
- Staats, B. R., H. Dai, D. Hofmann, and K. L. Milkman (2016). Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in health-care. *Management Science* 63(5), 1563–1585.
- Sudhir, K. and D. Talukdar (2015). The peter pan syndrome in emerging markets: The productivity-transparency trade-off in it adoption. *Marketing Science* 34(4), 500–521.
- Tabarrok, A. and T. Cowen (2015). The end of asymmetric information. *Cato Unbound*.
- Zhang, Y., B. Li, and K. Ramayya (2016). Learning individual behavior using sensor data: The case of gps traces and taxi drivers.

Online Appendix

“Do Digital Platforms Reduce Moral Hazard? The Case of Taxis and Uber”

Meng Liu
(WUSTL and MIT)
Erik Brynjolfsson
(MIT and NBER)
Jason Dowlatabadi
(Uber Technologies)

A Theory

Recall that the first-order conditions are,

$$f_x(x; d_0, \Theta) + q_e e_t \gamma d_0 - s d_0 (p_d + p_t \gamma) = 0, \quad (8)$$

$$g_y(y; d_0, \Theta) + q_e e_t - s p_t = 0. \quad (9)$$

Then, the comparative statics lead to the following testable implications:

Hypothesis 1: *Drivers tend to detour more on longer routes than on shorter routes because longer distance increases detour payoffs unless the demand at the drop-off location is sufficiently high, marginal detour penalty increases significantly with trip distance, or both.*

Proof. $\frac{\partial x^*}{\partial d_0} = \frac{s(p_d + p_t \gamma) - q_e \gamma e_t - f_{x d_0}}{f_{xx}} \leq 0$, depending on the sign of $s(p_d + p_t \gamma) - q_e \gamma e_t - f_{x d_0}$.

(1) Clearly, when $f_{x d_0} \gg 0$, or marginal detour penalty increases significantly with trip distance, $\frac{\partial x^*}{\partial d_0} < 0$, or drivers detour less on longer routes than shorter routes.

(2) When $f_{x d_0} = 0$, $\frac{\partial x^*}{\partial d_0} > 0$ if and only if $q_e < \frac{s(p_d + p_t \gamma)}{\gamma e_t}$. In the case of taxis in normal NYC traffic, $s = 1$, $p_0 = 2.50$, $p_d = 2.50$, and $p_t = 0$. For simplicity, let $e_t = \frac{p_0 + p_d D_e + p_t T_e}{T_e}$, where D_e and T_e are the expected length and duration of the forgone trip, respectively. Then, $\gamma e_t = \gamma \times \frac{2.50 + 2.50 D_e}{\gamma D_e} = \frac{2.50 + 2.50 D_e}{D_e}$, by having $T_e = \gamma D_e$. Therefore, $\frac{\partial x^*}{\partial d_0} > 0$ requires q_e to be less than $\frac{D_e}{1 + D_e}$. For example, for an average NYC taxi trip of 3 miles, the threshold for q_e is 0.75. Therefore, drivers detour more on longer routes than shorter routes unless the drop-off demand q_e is sufficiently high. Similar conditions apply to Uber. ■

Hypothesis 2: *Drivers detour more when the rider is a non-local passenger, and they detour less when the rider is a local passenger, as non-local passengers are less likely to notice the detour because they lack knowledge of local geography.*

Proof. For a parameter $\theta \in \Theta$ that increases the marginal detour penalty ($f_{x\theta} > 0$), e.g., the passenger is local, $\frac{\partial x^*}{\partial \theta} = -\frac{f_{x\theta}}{f_{xx}} < 0$. ■

Hypothesis 3: *Drivers detour more (respectively, less) during high surge prices if the increase in marginal detour payoff due to high surge dominates (respectively, is dominated by) the increase in marginal detour penalty due to high surge.*

Proof. $\frac{\partial x^*}{\partial s} = \frac{d_0(p_d + p_t \gamma) - f_{xs}}{f_{xx}} \leq 0$, depending on the sign of $d_0(p_d + p_t \gamma) - f_{xs}$. ■

Hypothesis 4: *Drivers detour less (respectively, more) when the demand at the drop-off location is higher (respectively, lower).*

Proof. $\frac{\partial x^*}{\partial q_e} = -\frac{e_t \gamma d_0}{f_{xx}} < 0$. ■

Hypothesis 5: *Everything else held constant, taxi drivers have greater incentives than Uber drivers to drive faster than other traffic on the road.*

Proof. Let c denote taxis and u denote Uber. According to Equation 9, $g_y(y^*) = sp_t - q_e e_t$. Given $g_y < 0$ for $y < 0$, $g_y > 0$ for $y > 0$, and $p_t = 0$ for taxis, it then follows that $y^{c*} < 0$. The condition $p_t > 0$ for Uber makes speeding less profitable for Uber drivers than for taxi drivers, especially when the penalty cost of deviating from the road traffic is greater for Uber than for taxis, i.e., $g_{uy} < g_{cy}$ for $y < 0$ and $g_{uy} > g_{cy}$ for $y > 0$. Therefore, as long as other parameters (q_e and e_t) are similar between taxis and Uber, $y^{c*} < y^{u*}$. ■

B Figures

Figure A1: Dividing NYC into Voronoi Cells Centered at Street Intersections



Figure A2: Pick-up Locations of Uber and Taxi after Matching Step 1 (taxis in purple; Uber in green)

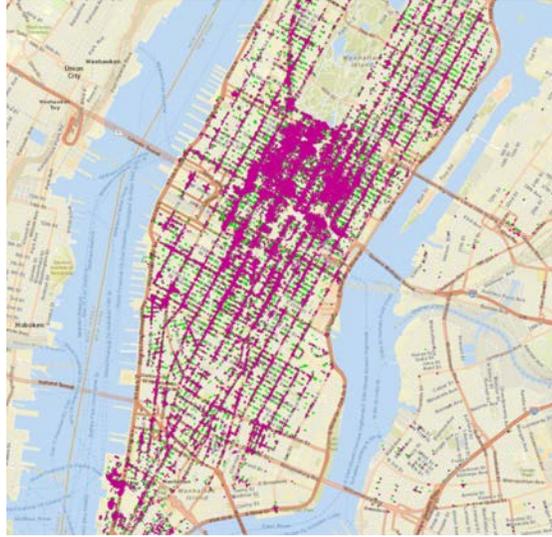
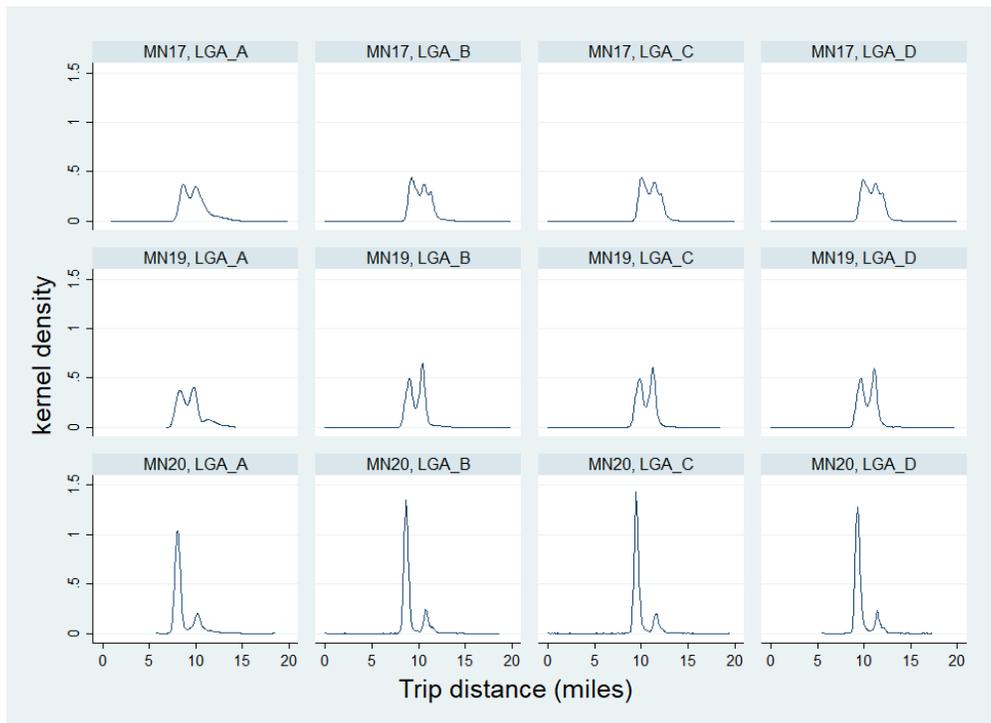
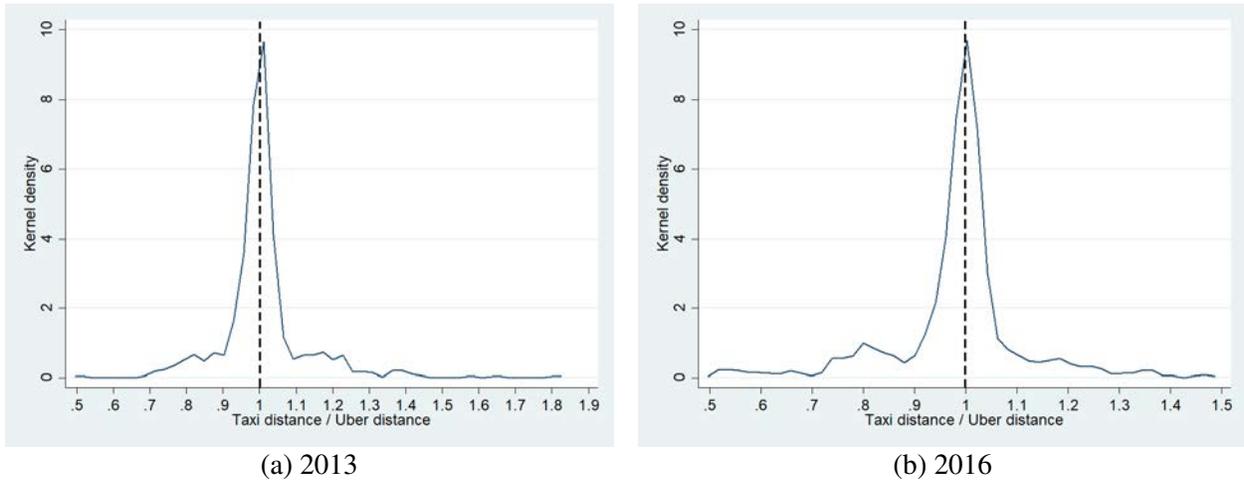


Figure A3: Taxi Trip Distance Exhibits a Bimodal Pattern for Tolloed Airport Routes



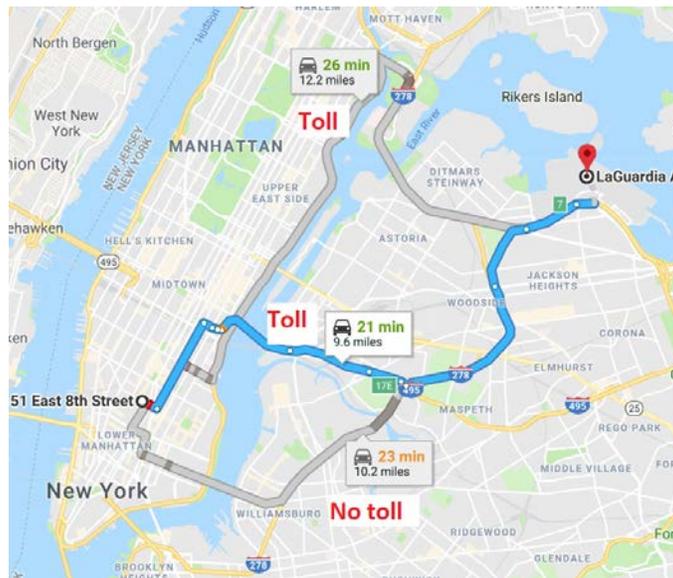
Notes. The plots are based on NYC yellow medallion taxi trips, January to June, 2016. Individual density plots demonstrate bimodal patterns of taxi distance on tolled trips for all routes between three NTAs in Midtown and four LaGuardia terminals. NTA code “MN17” denotes Midtown-Midtown South, “MN19” denotes Turtle Bay-East Midtown, and “MN20” denotes Murray Hill-Kips Bay. LaGuardia terminals are denoted by A, B, C, and D. Similar bimodal patterns are found for UberX trips.

Figure A4: Long-standing Uber Drivers Did Not Become More Honest with Time



Notes. These plots show that 234 drivers who remained Uber drivers from 2013 to 2016 do not appear to have become more honest in routing, as compared to taxi driver routing in the respective time periods, on JFK taxi flat-fare routes (where taxis are considered the “no detour” benchmark).

Figure A5: An Example Where the Longer, Toll-free Route Can Benefit Passengers



Notes. This is an example where the longer, toll-free route can benefit passengers. The toll-free route of 10.2 miles costs \$1.80 more than the shortest and fastest route of 9.6 miles ($0.6 \text{ miles} \times \$2.50/\text{mile} \times 1.20$, assuming a tip of 20%), according to the taxi pricing formula. Yet this route saves the passenger a toll of \$5.54. Thus, the longer, toll-free route is preferable by the passenger as long as the cost saving justifies the extra travel time (i.e., if passengers value their time less than $\$1.87$ per minute $(\$5.54 - \$1.80)/2$ minutes).

C Tables

Table A1: Former Taxi Drivers Compared to Uber Drivers in 2013

D.V. = Taxi dist. / Uber dist.	(1)	(2)	(3)
Airport	0.100*** (0.011)	0.108*** (0.013)	0.100*** (0.019)
Non_local	-0.009 (0.006)	-0.011* (0.006)	-0.013 (0.012)
Airport \times Non_local	0.026** (0.013)	0.025* (0.013)	0.026 (0.022)
Surge_multiplier	0.005 (0.008)	0.004 (0.008)	-0.001 (0.015)
Surge_multiplier \times Airport	0.002 (0.019)	0.006 (0.018)	-0.018 (0.030)
Log (Uber_driver_total_trips)	-0.004* (0.002)	-0.004* (0.002)	-0.003 (0.004)
Log (Uber_rider_total_trips)	0.001 (0.002)	0.001 (0.002)	0.001 (0.003)
Hour of week FE	No	Yes	Yes
Former taxi driver FE	No	No	Yes
N	3,952	3,952	3,952
R^2	0.081	0.127	0.594
No. of former taxi drivers	1,892	1,892	1,892
No. of Uber drivers in 2013	1,468	1,468	1,468

Notes. For all specifications, standard errors are cluster-robust at the hour-of-week level; *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

Table A2: Former Taxi Drivers Compared to Taxi Drivers in 2016

D.V. = Taxi dist. / Uber dist.	Airport is metered (1-3)			Airport is JFK flat fare (4-6)		
	(1)	(2)	(3)	(4)	(5)	(6)
Airport	0.047*** (0.004)	0.049*** (0.005)	0.049*** (0.006)	0.016 (0.012)	0.019 (0.014)	0.036 (0.023)
Non_Local	0.004 (0.005)	0.002 (0.006)	-0.002 (0.007)	0.009 (0.011)	0.009 (0.011)	0.010 (0.015)
Airport × Non_Local	0.003 (0.007)	0.004 (0.007)	0.004 (0.008)	-0.006 (0.015)	-0.005 (0.016)	-0.013 (0.026)
Surge_multiplier	-0.001 (0.009)	-0.000 (0.010)	0.001 (0.010)	0.016 (0.015)	0.018 (0.017)	0.004 (0.020)
Surge_multiplier × Airport	-0.019 (0.014)	-0.015 (0.015)	-0.017 (0.015)	-0.049 (0.032)	-0.040 (0.036)	-0.041 (0.054)
Log (Uber_driver_total_trips)	0.002 (0.002)	0.002 (0.002)	0.004 (0.007)	0.008* (0.004)	0.009** (0.005)	0.007 (0.018)
Uber_driver_rating	0.051** (0.024)	0.050** (0.024)	0.000 (.)	0.122*** (0.044)	0.150*** (0.047)	0.000 (.)
Log (Uber_rider_total_trips)	-0.001 (0.001)	-0.001 (0.001)	-0.002 (0.001)	-0.003 (0.002)	-0.004 (0.003)	-0.003 (0.004)
Hour of week FE	No	Yes	Yes	No	Yes	Yes
Uber driver FE	No	No	Yes	No	No	Yes
N	8,570	8,570	8,570	3,751	3,751	3,751
R ²	0.028	0.051	0.328	0.004	0.051	0.511
No. of Uber drivers in 2016 (former taxi drivers)	1,892	1,892	1,892	1,892	1,892	1,892

Notes. For all specifications, standard errors are cluster-robust at the hour-of-week level; *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

Table A3: Robustness: Various Time Windows

D.V. = Taxi dist. / Uber dist.	5 min.	10 min.	15 min.	20 min.	30 min.
Airport	0.074*** (0.005)	0.073*** (0.003)	0.074*** (0.003)	0.073*** (0.003)	0.072*** (0.002)
Non_local	-0.002 (0.005)	-0.003 (0.003)	0.001 (0.003)	-0.001 (0.002)	-0.003 (0.002)
Airport × Non_local	0.015** (0.006)	0.016*** (0.004)	0.013*** (0.003)	0.015*** (0.003)	0.016*** (0.002)
Surge_multiplier	0.009 (0.008)	0.002 (0.005)	-0.000 (0.004)	0.001 (0.003)	0.002 (0.003)
Surge_multiplier × Airport	-0.023** (0.011)	-0.024*** (0.007)	-0.022*** (0.005)	-0.020*** (0.005)	-0.024*** (0.004)
Log (Uber_driver_total_trips)	0.005 (0.004)	0.006*** (0.002)	0.005** (0.002)	0.005*** (0.002)	0.006*** (0.001)
Log (Uber_rider_total_trips)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.000 (0.000)	0.000 (0.000)
Hour of week FE	Yes	Yes	Yes	Yes	Yes
Uber driver FE	Yes	Yes	Yes	Yes	Yes
N	32,475	60,124	86,627	112,179	161,896
No. of Uber drivers	15,565	20,390	23,109	24,770	27,028
R^2	0.562	0.440	0.376	0.336	0.285

Notes. The regression samples consist of matched pairs of taxi and Uber trips on metered airport routes and non-airport routes, for various time windows. For all specifications, standard errors are cluster-robust at the hour-of-week level; *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

Table A4: Robustness: Taxi Meter Vendor 1

D.V. = Taxi dist. / Uber dist.	Vendor 1	Vendor 2 (Main sample)	Vendor 1 and 2
Airport	0.093*** (0.003)	0.074*** (0.003)	0.082*** (0.002)
Non_local	0.003 (0.003)	0.001 (0.003)	0.000 (0.002)
Airport × Non_local	0.007** (0.004)	0.013*** (0.003)	0.012*** (0.002)
Surge_multiplier	0.007 (0.005)	-0.000 (0.004)	0.001 (0.003)
Surge_multiplier × Airport	-0.032*** (0.007)	-0.022*** (0.005)	-0.025*** (0.004)
Log (Uber_driver_total_trips)	0.007*** (0.002)	0.005** (0.002)	0.005*** (0.001)
Log (Uber_rider_total_trips)	0.000 (0.001)	0.001 (0.001)	0.000 (0.000)
Hour of week FE	Yes	Yes	Yes
Uber driver FE	Yes	Yes	Yes
N	71,850	86,627	158,477
R ²	0.413	0.376	0.296

Notes. The regression samples consist of matched pairs of taxi and Uber trips on metered airport routes and non-airport routes, for Vendor 1 taxis, Vendor 2 taxis, and both taxi vendors, respectively. For all specifications, standard errors are cluster-robust at the hour-of-week level; *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

Table A5: Uber Driver Ratings by Passengers Are Correlated with Routing Efficiency

	<u>All JFK flat-fare routes</u>			<u>Uber dist.</u> <u>≥ Taxi dist.</u>	<u>Uber dist.</u> <u>< Taxi dist.</u>
	(1)	(2)	(3)	(4)	(5)
Uber_dist/Taxi_dist	-0.152** (0.062)	-0.150** (0.061)	-0.132 (0.082)	-0.283** (0.115)	0.518** (0.218)
Uber_dur/Taxi_dur	-0.253***	-0.204*** (0.046)	-0.206*** (0.053)	-0.269*** (0.077)	-0.123 (0.082)
Uber_driver_rating		0.862*** (0.096)	0.880*** (0.104)	0.904*** (0.149)	0.765*** (0.154)
Log (Uber_driver_total_trips)		0.013* (0.008)	0.015* (0.009)	0.021 (0.013)	0.011 (0.012)
Surge_multiplier		-0.111** (0.045)	-0.100* (0.053)	-0.121 (0.083)	-0.034 (0.079)
Non_local		0.019 (0.017)	0.021 (0.023)	0.039 (0.031)	0.007 (0.030)
Log (Uber_rider_total_trips)		0.028*** (0.005)	0.027*** (0.006)	0.035*** (0.009)	0.018** (0.008)
Hour of week FE	No	No	Yes	Yes	Yes
N	5,885	5,885	5,885	3,139	2,746
R ²	0.007	0.026	0.049	0.080	0.078

Notes. For Specifications (3), (4), and (5), standard errors are cluster-robust at the hour-of-week level; *** significant at the 1% level; ** significant at the 5% level. * significant at the 10% level.