DECLINE OF RIDE-SPLITTING: A CASE STUDY OF NEW YORK CITY

Ayush Pandey, Corresponding Author
ayushp3@illinois.edu
Department of Civil and Environmental Engineering
University of Illinois Urbana-Champaign

Hao Liu, Ph.D.
hao.liu@jsums.edu
Department of Civil & Environmental Engineering and Industrial Systems & Technology
Jackson State University

Xin Dong
xjd5036@psu.edu
Department of Civil and Environmental Engineering
The Pennsylvania State University

Lewis J. Lehe, Ph.D.
lehe@illinois.edu
Department of Civil and Environmental Engineering
University of Illinois Urbana-Champaign

Vikash V. Gayah, Ph.D.
gayah@engr.psu.edu
Department of Civil and Environmental Engineering
The Pennsylvania State University

Word Count: 5907 words + 6 table(s) × 250 = 7407 words

Submission Date: September 19, 2024

ABSTRACT

Ride-splitting is a service offered by Transportation Network Companies that pairs riders who choose to share a portion of their ride-hailing trip with a stranger, in exchange for a discounted fare. This paper studies the decline of ride-splitting using data from New York City. We document the events concerning ride-splitting and present plausible explanations for them. The decline of ride-splitting pre-COVID can possibly be explained by Uber's withdrawal of subsidies for UberPool; no other TNC reduces their discounts during this period. Post-COVID, we provide a plausible reason for why service has settled at a much lower level and analyze the effects of the same. We hypothesize that ride-splitting is stuck at a lower-sharing equilibrium. There are multiple equilibria because ride-splitting exhibits strong scale economies. Finally, we establish the presence of scale economies in ride-splitting using an Instrumental Variables approach.

Keywords: ride-pooling, ride-splitting, scale economies, instrumental variables, ridesharing

INTRODUCTION

Ride-splitting, also called ride-pooling or shared ride-hailing[1], is a service where passengers can choose to share their ridehailing trips with strangers to make their trips more affordable (2). Passengers use an app to input their origin, destination and willingness to share a ride and the Transportation Network Company (TNC), like Uber or Lyft, tries to match them with other riders traveling along the same route; the TNC may or may not successfully match a ride-splitting trip. Ride-splitting has garnered a lot of interest since UberPool and Lyft Line were launched in August 2014. UberPool served over 100 million trips by 2016 and accounted for more than half of Uber's trips in many cities; Lyft Line accounted for 30% of all Lyft rides with the proportion being closer to 50% in cities like New York and San Francisco (3). However, the share of ride-splitting trips has declined to single digit percentages today. Figure 1 shows the number of ridehailing trips requested in New York City from February 2019 to January 2024. The share of ride-splitting trips decreases from about 30% to 3% during this period. In this paper, we chronicle the significant events from 2019 to 2023 that have impacted ride-splitting. Additionally, we propose a plausible hypothesis for its failure to rebound post-COVID.

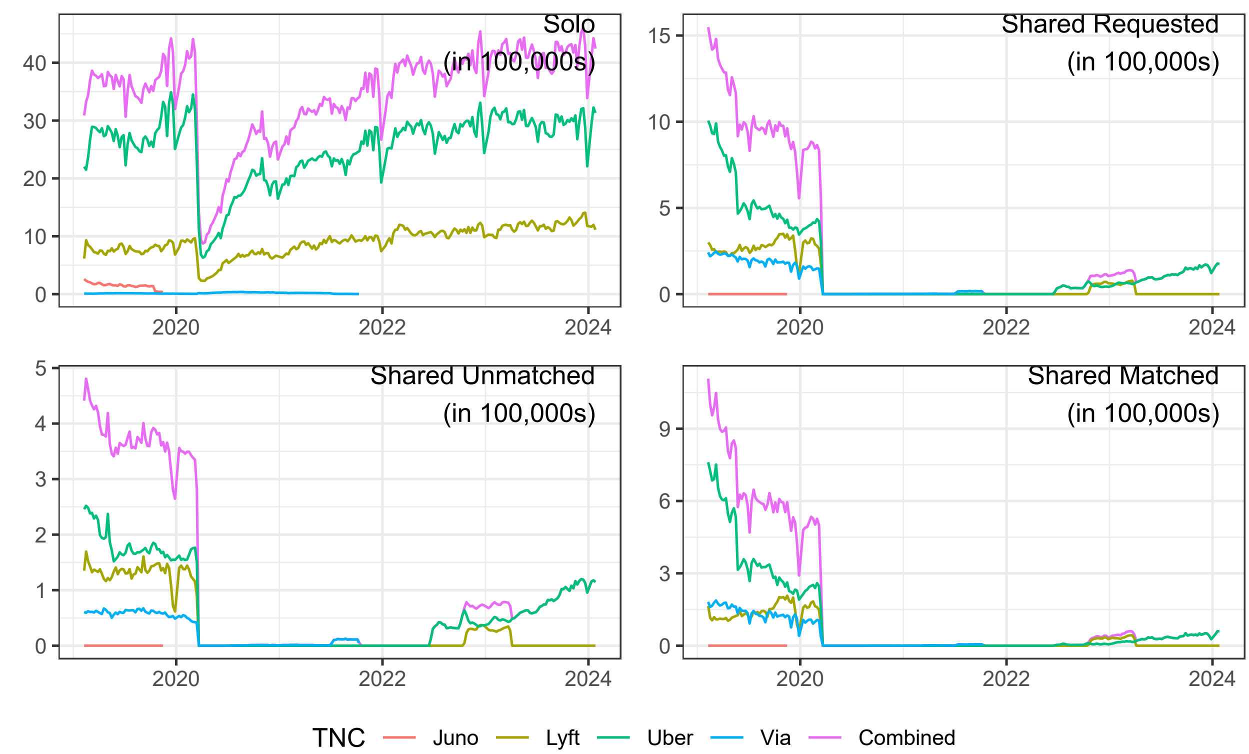


FIGURE 1: Number of trips taken each week in NYC from February 2019 to January 2024

In this paper we classify different types of trips as follows: Trips that do not want to be pooled are called *solo* trips. Trips that request ride-splitting are called *shared requested* trips. Among those, the trips that successfully get matched with another passenger are called *shared matched* and ones that do not get matched are called *shared unmatched*.

[1]See Shaheen et al. (1) for taxonomy.

A rider almost certainly suffers detours, to pick-up and drop-off other passengers, if their trip is successfully matched. TNCs offer discounts on ride-splitting to make up these for these detours and attract passengers (4), and at least a part of these discounts is not contingent on being matched. There are more passengers in the car in a shared matched trip and the fare for each passenger may be even lower. Ride-splitting puts more people in the same vehicle which can lower the number of vehicles required to serve all trips (1). State and local governments recognize that ride-splitting might lead to less congestion, fuel use and need for parking. It is a more sustainable option compared to traveling alone. To promote ride-splitting, cities offer discounts to passengers on taxes and surcharges levied on all ride-hailing trips. Chicago, IL; New York City, NY; San Francisco, CA; and New Jersey, Georgia, and some other cities and states levy a lower tax on shared ride-hailing trips than on solo ones (5).

Another reason to subsidize ride-splitting is that many professionals (as well as academics) believe that ride-splitting will become more efficient at scale. i.e., ride-splitting becomes more efficient as more people request a shared ride-hailing trip. As more people request to share their rides, the TNC will have a large pool to match people from which should allow them to (i) successfully match more trips Kidd (6), and (ii) find better matches with smaller detours. At scale, ride-splitting is expected to have small detours and low travel costs for passengers; higher occupancy and smaller fleet sizes for TNCs; and lower congestion, smaller emissions, and lower demand for parking for cities.

Castillo and Mathur (7) show that matching between drivers and riders gets more efficient as more drivers and riders participate in the system. Lehe et al. (8) show evidence of scale economies in carpooling using data from SCOOP (a carpooling platform) and Ke et al. (9) uses simulations to show that the match rates increase and detours decrease when more people demand pooled rides. Using Chicago's 2019 TNC data, Liu et al. (4) show that detours decrease and match rates increase as more people authorize to share rides with others. However, establishing scale economies in matching is not as simple because of the endogeneity between demand and efficiency measures. Empirical evidence (4, 8) shows that detours decrease as more people request to share rides; but it does not establish if it happens because detours get shorter with more requests, or do people request more when detours are shorter? Probably both are true simultaneously. Since the efficiency and demand for ride-splitting are endogenous, ordinary least square regressions cannot estimate the causal effect of number of requests on the efficiency of the system. This paper utilizes an instrumental variables based approach to show that ride-splitting exhibits scale economies.

We hypothesize that the presence of scale economies is a *plausible* explanation for ride-splitting's failure to rebound post-COVID. Transportation systems with scale economies *can* have multiple equilibria (10, 11) which may help explain why different cities can sometimes end up in different equilibria (12). We propose the following explanation: Ride-splitting was already in decline in 2019 (Figure 1) but the system was at a high-sharing equilibrium. The proportion of ride-splitting was high (between 25-30%) and more trips were getting matched (over 60%). The COVID19 pandemic forced TNCs to shut-down all ride-splitting offerings. Demand recovered slowly after COVID subsided and agencies resumed their ride-splitting options. But low initial demand meant that matching was not very efficient, passengers had to face with higher detours, and lower discounts; and TNCs lost more money due to low match rates. TNCs responded to loses by shutting down ride-splitting entirely or by greatly reducing the discounts offered to shared trips, leaving us with a low-sharing equilibrium. To be precise about our claim, the existence of multiple equilibria with vastly different demand and match rates does not necessarily mean that the system exhibits scale economies; but scale economies can be a reason for the existence of multiple equilibria. We can *possibly* shift from the low-sharing to the high-sharing equilibrium if ride-splitting is highly subsidized for a long period, but it seems unlike that TNCs will choose to do that.

The next subsection presents a timeline of ride-splitting in the United States. We then describe the data used for this study. Next, we discuss the reasons behind the decline of ride-splitting in 2019, and TNCs response to the same issues post COVID. The following Section discusses the effects of Lyft's exit from ride-splitting market in detail. The next section established the existence of scale economies ride-splitting using an Instrumental Variables based analysis. Final section 6 concludes.

Timeline

Table 1 gives an overview of the key events concerning ride-splitting in the United States starting June 2014, when Hitch became the first company to offer ride-splitting service in San Francisco, CA (13). Just a few months later, Uber announced it's ride-splitting option, *UberPool*, on August 5, 2014 (14). The service was initially launched as a beta test available to a few customers and was expanded to a public beta in the San Francisco Bay Area on August 15, 2014. Following Uber's announcement, Lyft launched *Lyft Line* to all customers in San Francisco on August 6, 2014 (15), and acquired other companies like Hitch and Rover to jump start their ride-splitting efforts (16); by March 2015, Lyft Line accounted for majority of Lyft's business in San Francisco. UberPool had served more than 100 million trips by March 2016 and more than half the trips taken in an Uber in many cities requested UberPool(17). Uber's CEO at the time, claimed that UberPool was profitable in many of the 29 cities it operated in (17). During the same time, Lyft Line was serving about 30% of all Lyft rides and the proportion was over 50% in San Francisco and New York. Lyft re-branded Lyft Line to Lyft Shared in June 2018 (18).

Uber and Lyft were the first large companies to offer ride-splitting in the US. Looking to expand their pooling business, Uber launched Uber *ExpressPool* in February 2018 (19) asking passengers to walk a short distance in exchange for larger discounts. Lyft announced Lyft *Shared Saver*, a service similar to Uber ExpressPool, on February 20, 2019 (20). To match passengers along routes with high demand, Lyft introduced a (now defunct) shuttle option in 2017 which plied on fixed routes during commute hours and charged a fixed fare (21). Uber announced a similar service, Uber *Shuttle*, in May 2024, (22, 23).

All ride-splitting services were suspended due to the COVID19 pandemic. UberPool, Uber ExpressPool, LyftLine, and other services were shutdown in March 2020. Uber brought back its ride-splitting option in November 2021, in Miami, with a new name: UberX Share (24). Lyft brought back Lyft Shared in July 2021 for Chicago, Philadelphia and Denver (25) and expanded it to Miami, Atlanta, Las Vegas, San Francisco, San Jose, Los Angeles, Nashville, Washington D.C., Boston, Portland, and San Diego., in May 2022 (26). Lyft shuttered their Lyft Shared in March 2023 (27). Uber still offers UberX Share.

Many apps came out during these years trying to compete with Uber and Lyft, most of which have since shut down or pivoted to other business models. A New York Times article (28) sums them up in this paragraph: "[the app] Gett came out against surge pricing ... Curb lets riders hail taxis either by hand or via its app, and then pay for them with the app. Lyft is perceived as a less corporate option ... And Juno's approach is to promise drivers a supportive corporate culture and a larger cut of its business." Via used to offer ride-splitting in NYC in 2019, but now Via's

TABLE 1: Timeline of ride-splitting in the United States.

June 2014	Hitch starts pairing up riders with similar origins and destimations.
August 5, 2014	Uber launches UberPool as a best test.
August 6, 2014	Lyft launches Lyft Line to all customers in San Francisco, CA.
February 2018	UberPool offered in 36 cities worldwide, 16 of them in the US.
February 22, 2018	Uber ExpressPool launched in 8 cities in the US.
June 2018	Lyft Line re-branded as Lyft Shared.
February 20, 2019	Lyft launches Lyft Shared Saver.
March 18, 2020	Ridesplitting services suspended due to COVID.
July 2021	Lyft brings back its ridesplitting option, Lyft Shared, in Chicago.
November 2021	Uber brings back UberPool post COVID; rebranded as UberX Share.
June 20, 2022	UberX Share introduced to New York City, NY.
June 2022	UberX Share available in 9 US cities.
October 19, 2022	Lyft Shared comes back to NYC.
March 23, 2023	Uber raises price per mile for UberX Share.
April 1, 2023	Lyft ceases offering Lyft Shared.
May 18, 2023	Uber brings back prices to previous levels.
May 15, 2024	Uber announces its shuttle service, Uber Shuttle.
July 2024	UberX Share available in 34 metropolitan regions in the US, and 60 cities globally.

TABLE 2: Regression results for average discount in New York City, 2019

	Uber	Lyft	Via	Combined
(Intercept)	8.184***	7.795***	11.632***	8.422***
week-of-year	-0.036**	-0.003	-0.015	-0.044***
Num.Obs.	47	47	47	47
R2	0.174	0.011	0.016	0.271

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

business relies on "working directly with transit agencies, rather than building up a de facto taxi service" (29). Today, Uber is the only major ride-hailing company offering ride-splitting services in 32 metropolitan regions in the US, and 60 cities globally (30).

Data

We use the TNC data published by the New York City's Taxi and Limousine Commission (TLC) from February 2019 to January 2024. Most of the analysis focuses on 2019 and 2023, using them as proxies for the pre- and post-COVID years. The TLC dataset starts from February 2019 and contains the records of (almost) all TNC trips in New York City (NYC), with their fare; distance; duration; origin and destination taxi zones; if the trip was authorized to be shared; if the trip was successfully matched with another trip; and the TNC that trip was requested on. NYC had 4 TNCs (Juno, Via, Uber, and Lyft) in 2019 and only two remained (Uber and Lyft) in 2023.

We calculate weekly aggregates for all of NYC. Statistics calculate by aggregating data from all TNCs is labeled as "Combined". For example, the total number of trips in all of NYC in Figure 1, calculated by taking the sum of all trips in different TNCs, is labeled as "Combined". Trips shorter than 2 minutes or 0.1 miles were filtered out. We calculate the number of *solo*, *shared requested*, *shared unmatched* and *shared matched* trips. The match rate is ratio of number of shared matched with the number of shared requested. We also calculate other aggregate statistics, like average fare, trip distance, and trip duration, for each type of trip for each week. All analysis, except in Section 5, uses weekly aggregates for all of NYC to get a city-wide big-picture perspective of ride-splitting in NYC.

PRE-COVID DECLINE

It might be tempting to ascribe the downfall of ride-splitting solely to persistent effects from COVID. But, ride-splitting was already in decline in 2019. The number of requests for shared trips fell from 1.55 million during the week of February 4-10, 2019, to about a million during the week of September 9-15, 2019, and then to 0.55 million during the holidays in December 23-29, 2019 (Figure 2). In contrast the number of requests for solo trips did no change much during 2019. 3.09 million solo trips during February 4-10, 2019; 3.63 million during September 9-15, 2019; and 3.2 million during in December 23-29, 2019.

The fall in number of shared trips requested is almost entirely due to the fall in demand for UberPool, which made up about two-thirds of all shared requested trips in February 2019; the demand for Lyft Share and Via is largely constant throughout 2019. The decrease in demand for UberPool might be a direct effect of Uber's financial troubles. UberPool offered discounts as high as 50% to attract customers, losing up to a million dollars a week in San Francisco alone (31). In

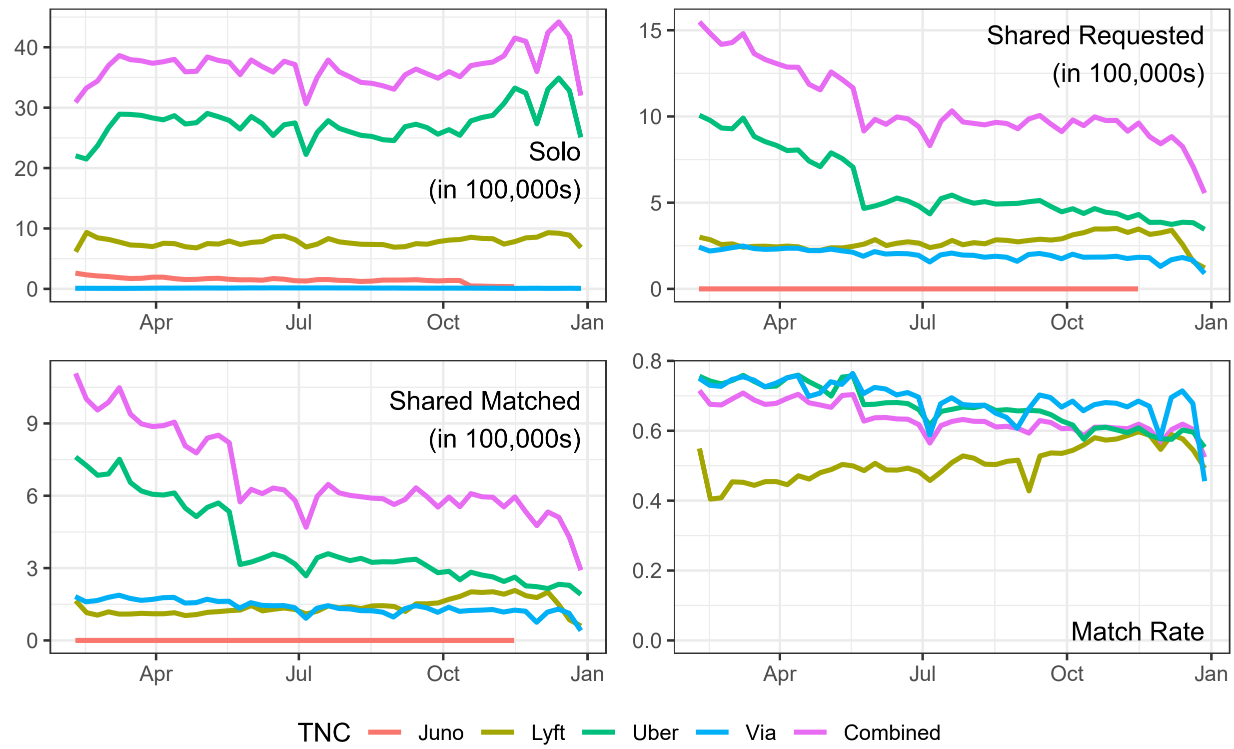


FIGURE 2: Weekly aggregates in NYC in 2019

2019, Uber was already in the process of reducing the discounts offered to shared rides. Figure 3 shows that the difference in average fare between a solo trip and a shared requested trip keeps decreases with time[2]. We regress the fare discount over week-of-year for all TNCs; Table 2 shows the results. The average discount offered by Lyft and Via did not have any statistically significant variation over the course of the year, while the discount offered by Uber decreased. Since Uber made up a large proportion of the combined ride-splitting market, the effect is statistically significant even when data from all TNCs was combined. Abkarian et al. (33) also notices a decreasing trend in the demand for shared rides in Chicago in 2019, suggesting that this withdrawal of subsidies is not unique to NYC.

The financial burden imposed on the TNCs by their ride-splitting options persisted after the COVID19 pandemic subsided. Via exited the ride-splitting market after a short comeback in 2021. Lyft stared re-introducing its ride-splitting option, Lyft Shared, in June 2021. The service was re-introduced in New York City on October 19, 2022 and Lyft exited the ride-splitting market on April 1, 2023; next section discusses the effect of this exit. Uber re-branded their ride-splitting service as UberX Share. When using UberPool, passengers were guaranteed a discount whenever they requested a shared ride, irrespective of the trip being successfully matched with other rider or not. With UberX Share "riders will receive an upfront discount if they choose UberX Share, and get up to 20% off the total fare, if matched with a co-rider along the way (34)," i.e., the riders get

[2]We call this value *fare discount*. There can be different ways calculating the fare discounts (Liu et al. (32) compares the fares for the same Origin-Destination pairs to get the fare ratio) but we compare the city wide averages to observe trends at the city level.

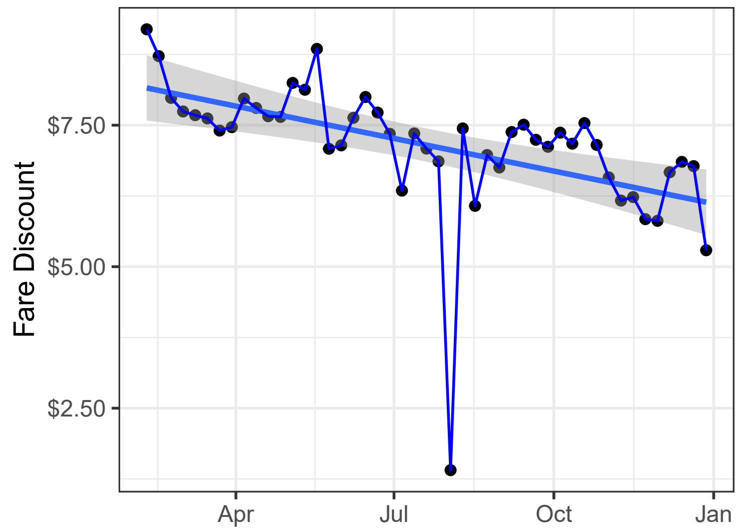


FIGURE 3: Average discount offered to shared rides in NYC in 2019

a smaller discount upfront for requesting a shared ride but only get the larger discount if the ride is successfully matched with another passenger. Uber was loosing a lot of money before COVID due to unmatched trips which probably forced them to try to match as many trips as possible, even at the cost of large detours. This is reflected in Uber's high match rate of over 60% in 2019. Post COVID, the loss on each unmatched trip would not be too high, meaning Uber did not worry too much about their match rate which fell to about 30% in 2023 (Figure 5).

TABLE 3: Top 10 OD pairs with highest shared rides requested in 2023

Origin	Destination	Total Trips	Shared Requested	Shared Matched	Match Rate
South Ozone Park	JFK Airport	243528	65879	1014	0.02
Springfield Gardens North	JFK Airport	58462	19862	411	0.02
Baisley Park	JFK Airport	76601	16709	522	0.03
East New York	East New York	487516	16321	1154	0.07
East Elmhurst	LaGuardia Airport	45773	14931	13	0.00
Canarsie	Canarsie	300470	11525	593	0.05
JFK Airport	JFK Airport	96894	11283	1136	0.10
Crown Heights North	Crown Heights North	235864	8405	873	0.10
Jackson Heights	LaGuardia Airport	50664	6366	130	0.02
Springfield Gardens South	JFK Airport	36685	6226	400	0.06

Another issue with ride-splitting is that passengers might strategically request shared trips — passengers request a shared trip when they do not expect to get matched which allows them to get a discount without suffering additional detours. This means that in certain cases we may observe high demand for shared trips but very low match rates. All of the top ten origin-destination (OD) pairs, by demand for shared trips, in NYC have a match rate of 10% or less; seven out of ten of these OD pairs end at an Airport. The highest number of requests for shared trips is from

South Ozone Park to the JFK Airport, over 3 times the the number of requests for OD pair with second highest trips. This high demand is probably because South Ozone Park is next to the JFK Airport and has many hotels and commercial parking lots. Passengers traveling to the airport might be requesting a shared ride knowing that their probability of getting matched with another rider is low. Such behavior makes the analysis of these ride-splitting systems even harder. Observing the effects of improvements in the matching algorithm would be much harder if people request shared trips only when they expect not to get matched. If people are indeed strategically requesting trips, a more efficient ride-splitting system with higher match rate might push people away from requesting shared rides. This will make it much harder to observe scale economies in ride-splitting.

EFFECT OF LYFT'S EXIT

In the presence of scale economies, it is natural to expect that ride-splitting will be more efficient if only one TNC offers the service. This is probably why Uber and Lyft invested heavily in ride-splitting, as, eventually, the winner would take the whole market. Uber did win the market and Lyft stopped offering their ride-splitting option, Lyft Shared, from April 1, 2023.

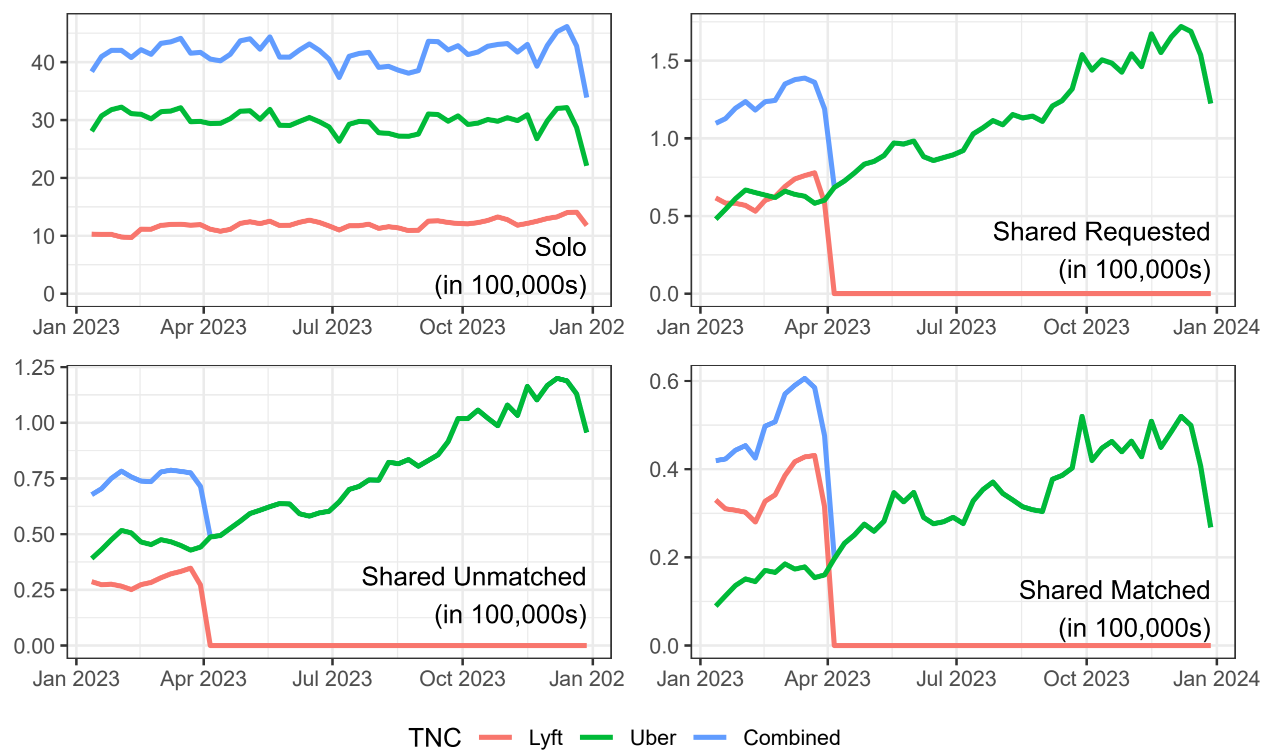


FIGURE 4: Weekly trips in NYC in 2023

Davalos (35) article from May 11, 2023 announces the cessation of Lyft Shared. Figure 4 shows that the number of Lyft Shared rides requested in NYC drop to zero starting April 1, 2023. The number of shared trips requested, and matched, on Uber kept increasing throughout 2023, while the number of solo trips was largely constant (Figure 4). The total number of shared trips requested in December 2023 was higher than the number of requests in March 2023 (with Lyft and Uber combined), indicating that the demand for ride-splitting was increasing and people

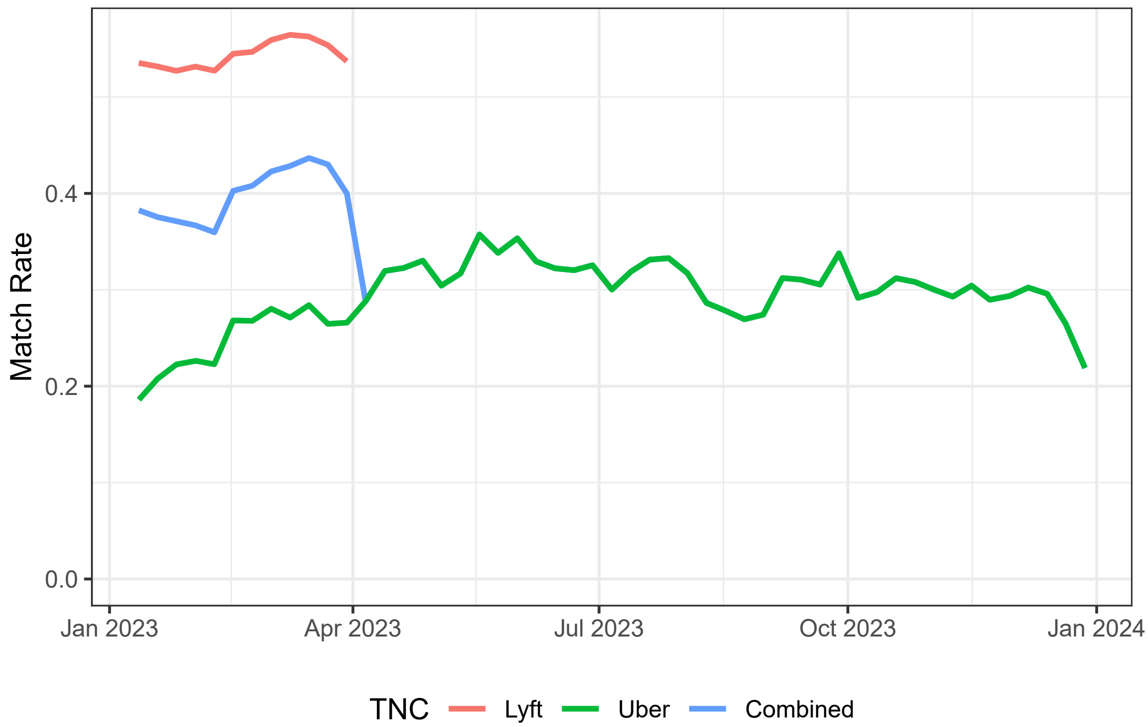


FIGURE 5: Match rate in NYC in 2023

slowly shifted to UberX Share from Lyft Shared; Lyft Shared was discontinued despite increasing demand.

The number of shared trips requested on Uber doubled from about 63,000 per week during the four weeks between March 5th and April 1st, to about 150,000 per week in October, a 137% increase. If there are city-wide scale economies in ride-splitting, then the match rate should increase with the doubling of demand. But Figure 5 shows that the match rate did not change much after Lyft's exit; going from around 27% in March to around 30% in October. This suggests that the UberX Share's matching efficiency did not change much despite the demand more than doubling.

In contrast to the hypothesis, the overall match rate actually decreased after Lyft's exit despite an increase in total demand that was served by only one TNC. Lyft's match rate of 53% was much higher than Uber's, suggesting that Lyft was much better at matching trips. A plausible explanation for Lyft's higher match rate is the following: Trips requested to be shared on Lyft are cheaper than the ones on UberX Share, per mile and per minute; Lyft Shared costs about $3 a mile while UberX Share is closer to $3.5 on average. Price sensitive passengers, who are the ones taking pooled trips, with longer trips might prefer using Lyft Shared over UberX Share. Hence, the average trip length of a shared requested trip on Lyft is longer, both in terms of time and distance, than the ones requested on Uber (6). Furthermore, the average trip duration and distances for the matched trips are comparable for both Uber and Lyft, suggesting that trips that get matched have similar characteristics, no matter the TNC. Since, Lyft can attract more passengers with longer trips, they have more opportunities to match them, and consequently, have a higher match rate. At the risk of being speculative, we might presume that Lyft was keeping their prices lower to compete with Uber and was losing a lot more money on pooled trips. This might have been the

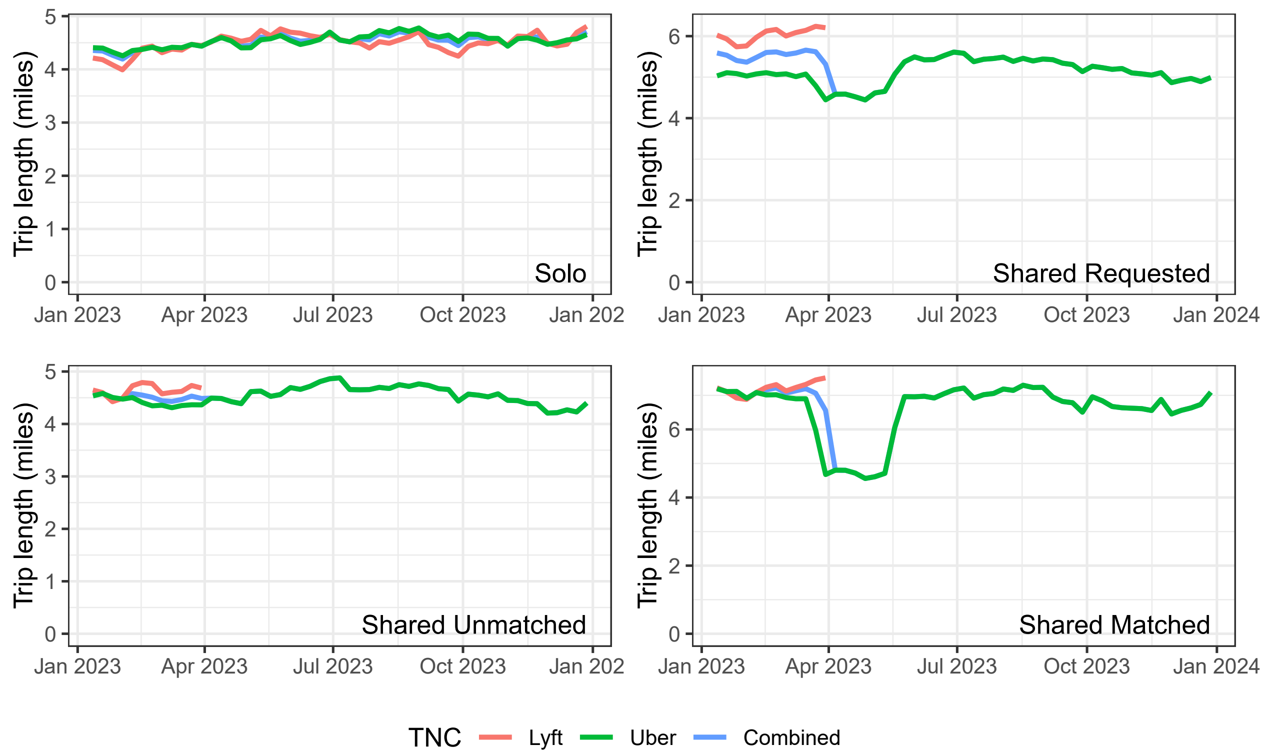


FIGURE 6: Average trip lengths in NYC in 2023

reason behind their decision to discontinue Lyft Shared. The average match rate, for all trips in NYC, declined from about 40% to about 30% after Lyft's exit.

We also observe a sharp fall in the average trip length of the trips Uber successfully matched between March 23, 2023 and May 18, 2023 in Figure 6. The trip length for shared requested trips also has a similar but smaller dip. There is no change in the length of solo or unmatched trips. We investigated the average travel time, fare, fare per mile, and fare per minute for all types of trips, i.e., solo, shared requested, unmatched and matched. Average travel time, fare, fare per minute do not show any sudden changes for any type of trip[3]. However, fare per mile for shared matched trips jumps abruptly during the same time period. Figure 7 plots the average travel time, fare, fare per minute and fare per mile for matched trips. Uber raised their prices, possibly, in anticipation of Lyft leaving the ride-splitting market. A week before Lyft's exit, the average fare per mile for successfully matched trips on UberX Share jumped from \$3/mile to \$4.5/mile. Remember that one of the key difference between UberPool and UberX Share is that a part of discount in UberX Share in contingent on being matched with another rider. What we observe here is, perhaps, Uber decreasing that conditional discount, or conversely increasing the fare per mile for the matched trips. The fare hike results in a decrease in the average trip length of pooled trips requested probably because some riders with long trips, that regularly got matched, stopped using UberX Share as the the higher fare was not worth the hassle of sharing their ride with a stranger. However, this price change too did not result in a notable difference in the match rate.

This indicates that the departure of Lyft did not increase the match rate for UberX Share by

[3]We have omitted these plots to avoid clutter.

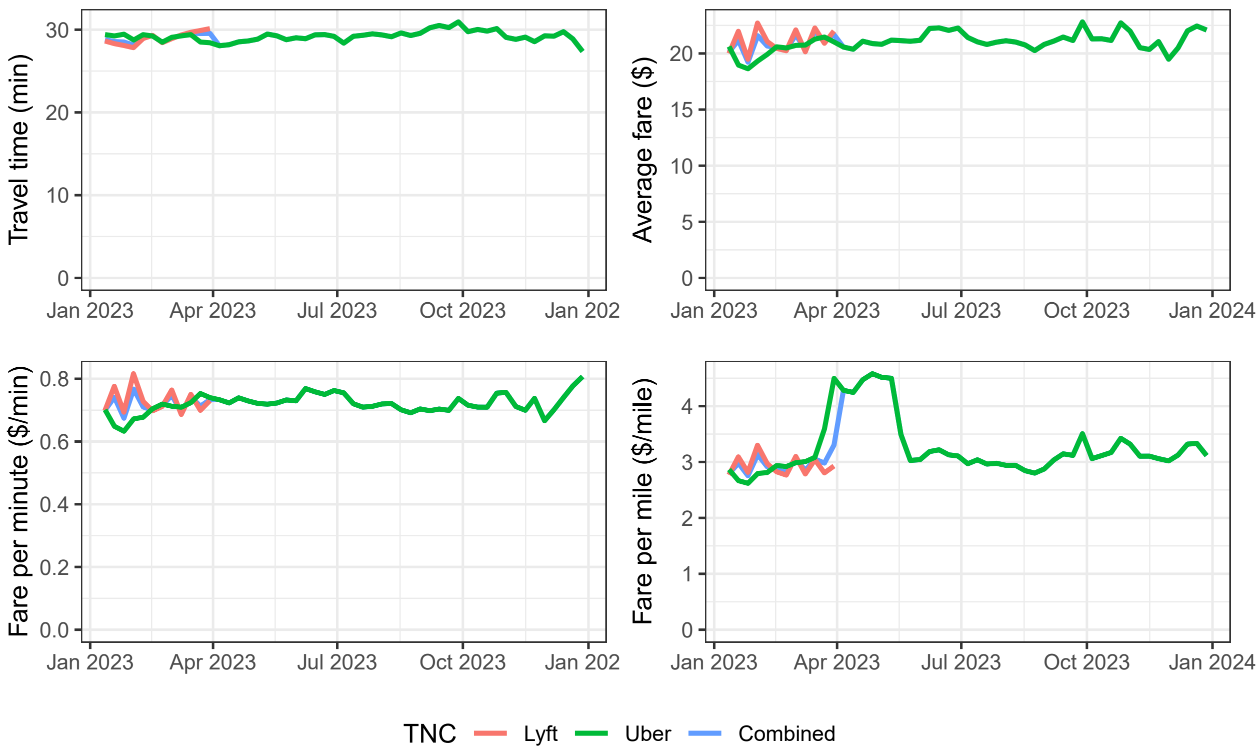


FIGURE 7: Travel time, average fare, average fare per minute and average fare per mile for matched trips in 2023

a lot, even when shared requested trips more than doubled. The presence or absence of competition did not affect how efficient the matching process was. However, the absence of competition did allow Uber to perhaps abuse their market power and raise their prices for a pricing experiment. The pricing change too did not affect the match rate much. A plausible reason why we do not observe large changes in match rates could be that the increase in demand for UberX Share, or the price change, was not large enough to shift us from a low-sharing to a high-sharing equilibrium. Even though the number of requests on UberX Share more than doubled between March and October, 2023; the number in October 2023 (150,000 per week) is less than half the number of requests in October, 2019 (440,000 per week). Maybe the decline in 2019 is a local change in the high-sharing equilibrium due to withdrawal of subsidies. And the effects of Lyft's exit are local changes in the low-sharing equilibrium. Presence of scale economies *can* give rise to multiple equilibria in systems. Hence, the next section tries to empirically identify scale economies in ride-splitting.

SCALE ECONOMIES

Since ride-splitting relies on finding similar trips to match together,it is natural to expect that the system will become more efficient as the number of people requesting to share rides increases, i.e., the system is expected to have scale economies. Lehe et al. (8) and Liu et al. (4) find evidence indicating the existence of scale economies, but empirically proving its existence is much harder. The quality of matching and the number of people requesting to share rides are endogenous. The presence of scale economies means that we expect the quality of matching to improve as more people request to share rides. However, more people might also request shared rides as the quality of matching improves. Hence, a positive correlation between quality and number of requests does not necessarily mean the system exhibits scale economies.

Assuming the quality of matching (η) and the number of people requesting to share rides (D^s) are linearly related, we can model the dependence of η on D^s as:

$$\eta = \gamma_0 + \gamma_1 D^s + \mu, \tag{1}$$

and the dependence of the number of people requesting to share rides on the quality of matching can be modeled as:

$$D^s = \alpha_0 + \alpha_1 \eta + \varepsilon. \tag{2}$$

Equations (1) and (2) have a *causal interpretation* and are analogous to simultaneous supply and demand relations from Wooldridge (36, Sec 16.3) and Romer (37)[4]. The supply curve (1) says that a change in D^s *causes* a change in η, and, similarly, the demand curve (1) says that a change in η *causes* a change in D^s. Claiming that the system has scale economies is equivalent to claiming that $\gamma_1 > 0$, i.e., η with D^s. However, we cannot estimate equations (1) and (2) as they have four unknowns $\gamma_0, \alpha_0, \gamma_1, \alpha_1$, and only two exogenous variables, η, D^s. The system does not have sufficient variation to identify any coefficients in either of the equations. We can write the reduced

[4]See Chapter 16 of Wooldridge (36) for a through analysis of systems with simultaneous equations.

form of (1) and (2) as

$$\gamma_0 + \gamma_1 D^s + \mu = \frac{D^s}{\alpha_1} - \frac{\alpha_0}{\alpha_1} - \frac{\varepsilon}{\alpha_1}$$
$$\implies D^s = \frac{\gamma_0\alpha_1 + \alpha_0}{1-\alpha_1\gamma_1} + \frac{\mu\alpha_1 + \varepsilon}{1-\alpha_1\gamma_1} \equiv \pi_{D^s 0} + \varepsilon_{D^s}. \quad (3)$$

Similarly,

$$\eta = \frac{\alpha_0\gamma_1 + \gamma_0}{1-\alpha_1\gamma_1} + \frac{\varepsilon\gamma_1 + \mu}{1-\alpha_1\gamma_1} \equiv \pi_{\eta 0} + \varepsilon_{\eta 0}. \quad (4)$$

Eqns. (3) and (4) do not have an exogenous variable on the right hand side and hence none of the coefficients from (1) and (2) can be estimated.

To estimate the coefficients in (1), we need some exogenous variable z_d that affects D^s but not η. We can think of z_d as some variable that introduces an exogenous variation to the demand curve which 'shifts' the demand curve without having any effect on the supply curve, allowing us to estimate the supply curve. If our goal was estimating the demand curve, we would need a similar exogenous variation in the supply side of our system. Consider some variable z_d that only affects the demand for shared trips but not the quality of matching. Our system of equations then becomes

$$\eta = \gamma_0 + \gamma_1 D^s + \mu, \quad (5)$$
$$D^s = \alpha_0 + \alpha_1\eta + \alpha_2 z_d + \varepsilon. \quad (6)$$

Solving for the reduced form, we get

$$\eta = \frac{\gamma_0 + \gamma_1\alpha_0}{1-\gamma_1\alpha_1} + \frac{\gamma_1\alpha_2}{1-\gamma_1\alpha_1} z_d + \frac{\mu + \gamma_1\varepsilon}{1-\gamma_1\alpha_1} \quad (7)$$
$$\equiv \pi_{\eta 0} + \pi_{\eta z_d} z_d + \varepsilon_\eta, \quad (8)$$

and

$$D^s = \frac{\alpha_0 + \alpha_1\gamma_0}{1-\gamma_1\alpha_1} + \frac{\alpha_2}{1-\gamma_1\alpha_1} z_d + \frac{\varepsilon + \alpha_1\mu}{1-\gamma_1\alpha_1} \quad (9)$$
$$\equiv \pi_{D^s 0} + \pi_{D^s z_d} z_d + \varepsilon_D^s. \quad (10)$$

Since z_d is exogenous to both η and D^s, we can estimate the coefficients of (8) and (10) using linear regression. This allows us to then estimate the effect of demand on quality of matching using

$$\gamma_1 = \frac{\pi_{\eta z_d}}{\pi_{D^s z_d}}. \quad (11)$$

$\frac{\pi_{\eta z_d}}{\pi_{D^s z_d}}$ is the IV estimate of γ_1, using z_d as an instrument for D^s. Hence, we can estimate the causal effect of the number of addition shared trips requested on the quality of matching.

Quality of matching

Quality of matching is a measure of how well the system is able to match trips. We cannot have a direct measure of this unless we know what objective function the TNCs are seeking to optimize, which is proprietary information. However, we can safely guess that an improvement in matching would result in an increase in match rates and a decrease in detour distances for successfully matched riders. We can use either of these measures as a proxy for η. We use the match rate in this study for the following reasons.

1. While the TNCs might want to limit the detour distances for the sake of the riders, they might also opt to set a limit on the detour distance, as a service guarantee, and not optimize further (or do so in a limited fashion). This is because the match rate affects the bottom line of the TNCs, while detour distance does not. Anyone who requests a pooled trip is offered a discounted price *regardless of the trip actually getting matched*. This means that the TNCs subsidize the pooled trips that do not get matched. Increasing their match rate is essential for the TNCs to make ride-splitting business sustainable. TNCs might choose to optimize for the match rate while limiting the detour distance to some maximum value. Hence, any increase in demand is likely to improve the match rate more than the detour distance.
2. The number of trips successfully matched in United States is still low which makes it hard to estimate the detour distances. For example, (4) estimates detour distances between OD pairs by calculating the average travel distance between the OD pair for solo trips and for successfully matched trips, and then taking their difference to get the detour for the OD pair. However, they ignore the randomness in trip lengths and only use the means. Since there are very few matched trips, the standard deviation for matched trip lengths are large, making the detour distances unusable.

Figure 8 plots the normalized detour, the ratio of detour and the distance between an OD pair, against the number of shared trips requested. The detour is calculated as the difference between the average trip length for solo trips and the average trip length for matched trips. We also calculate upper/lower bounds of the detour by adding/subtracting the standard deviation of the trip lengths of matched trips to the detour distance. We can observe that the detour distance has a lot of noise. The standard deviations are higher when few trips are being requested and matched.

Proposed Instrument

We need an instrument that affects the number of people requesting to share rides but does not affect the quality of matching. We propose to use stadium events. People naturally gather within specific time windows at stadiums for various events such as concerts or professional sports. As there are more arriving at and leaving the venue when an event is happening, more rides get requested during events and proportionally more people request to share rides. This increase in the number of requests is likely to be irrespective of the quality of matching. Hence, we can use if an event is happening at a stadium as an instrument for the number of people requesting to share rides and use the IV estimate to measure the causal effect of the number of people requesting to share rides on the match rate.

Data

We focus this analysis on the four Taxi Zones around Barclays center in Brooklyn (Figure 9), as the Barclays center is a major events venue in NYC and the date and start time of historical events

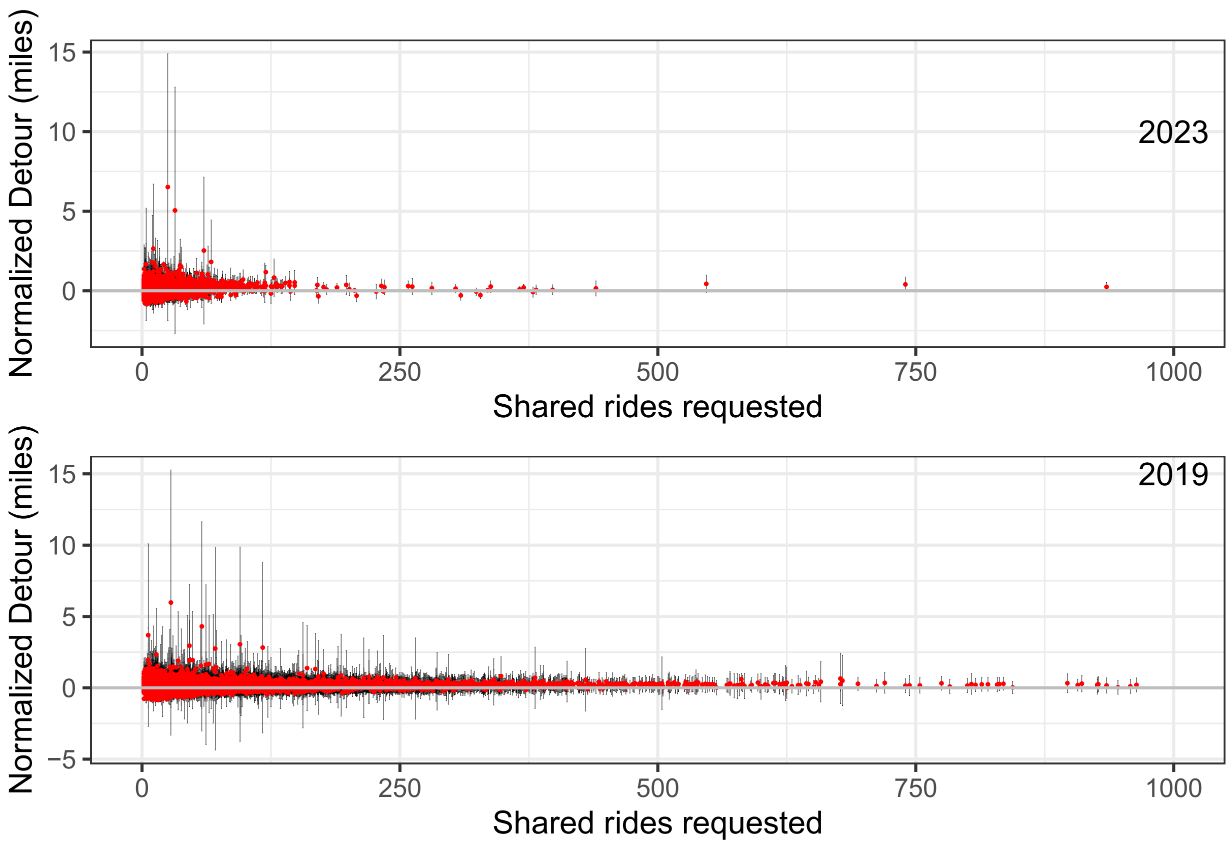


FIGURE 8: Normalized Detours between OD pairs in NYC

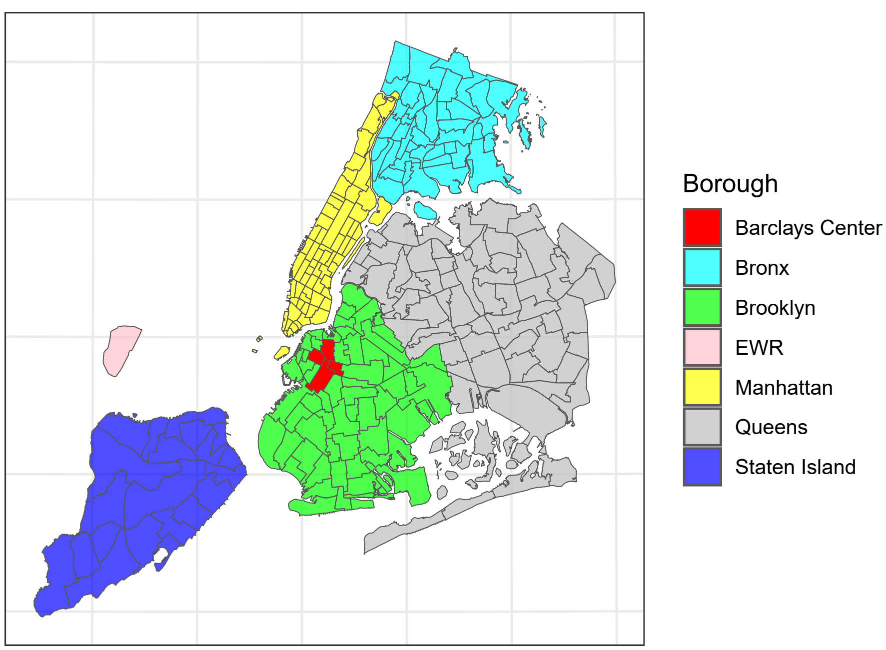


FIGURE 9: Taxi zones in New York City as given by the TLC data. Newark Airport (EWR) is not a borough of New York City but is included as such by the TLC.

is available from their website [5]. We use the TNC data from NYC TLC for 2019 and 2023. We use 2019 and 2023 as proxies for pre- and post-COVID data. We calculate the number of shared trips requested, starting or ending around the four taxi zones in Barclays center, and the number of trips successfully matched for *each hour* in 2019 and 2023.

The data for the events at Barclays center is collected manually from their website. This data includes the date and the start time of the event. We do not have any information about when these events end. Figure 10 shows that the number of shared trips requested to be dropped off at Barclays center increases in the hour before an event starts, and the number of shared trips requested to be picked up from Barclays center increases 2-4 hours after an event starts; we say that the effect of an event is *active* during these hours.

Since there can be a lot of variation in end times based on the type of event happening at the venue, we collect data about which of these events are basketball games. Basketball games last for a nearly fixed duration of 2.5 hours and we can use this information to estimate the end times for all events that were basketball games.

We have 8 total datasets. Two each for pickups and dropoffs in 2019 and 2023, and 2 each of these four for both instruments, (i) one constructed with all events, and (ii) the other with just basketball games. Each observation in a dataset is an hour of the year.

Regressions

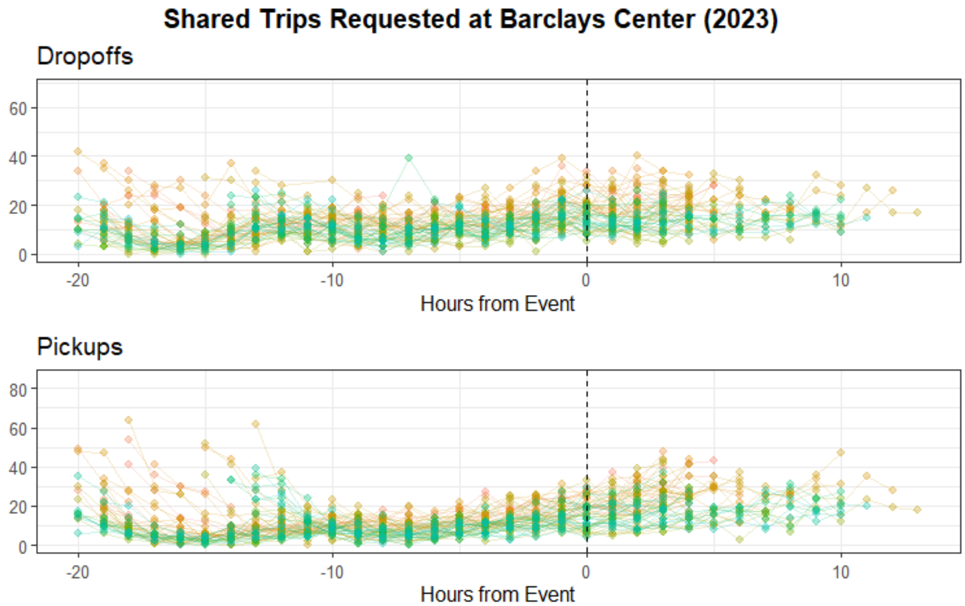
We run two sets of regressions on the data for 2019 and 2023. All regressions include a fixed effect for hour-of-day. We set $z_d = 1$ when the effect of an event is active to compare against times when events are not active.

1. The first set of regressions uses all the events as instruments. For drop-offs, we set the instrument $z_d = 1$ from two hours before the start to the start of an event, and $z_d = 0$ otherwise. For pick-ups, we set $z_d = 1$ for two to four hours after the start of an event, and zero otherwise. The wider window accounts for the fact that arrivals and departures might be more dispersed for large concerts.
2. The second set of regressions uses only basketball games as instruments. Days with any other events are removed from the dataset. Months with no basketball games at all are also removed. We set $z_d = 1$ for 1 hour before the start of the game for drop-offs. All games last for about 2.5 hours, so we set $z_d = 1$ for pick-ups from 2 to 3 hours after the start time of a game.

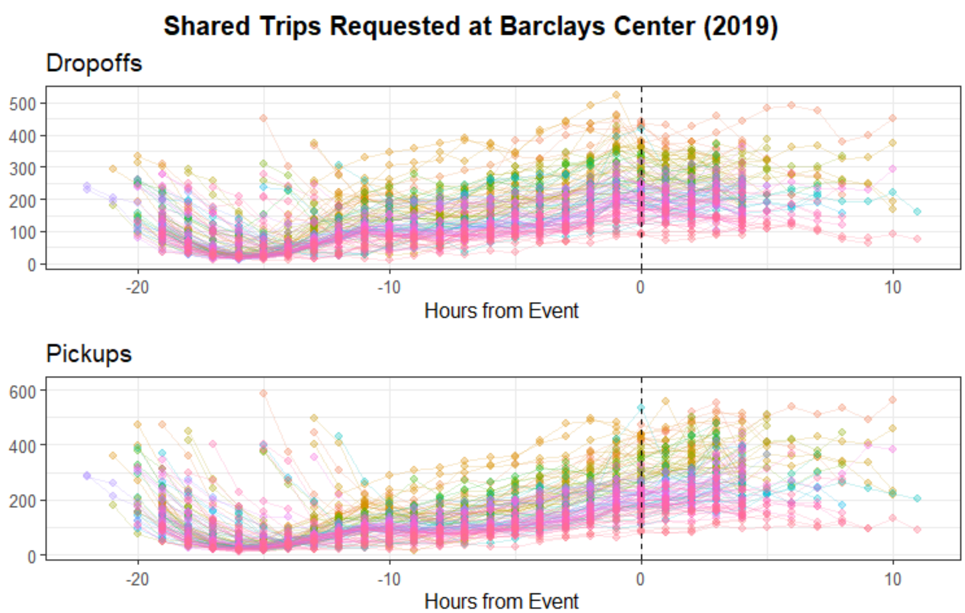
Table 4 shows the results for the first stage regression, as defined in (10). Considering all events, an "active" event is a strong instrument for drop-offs in 2023 and 2019, and for pickups in 2019; the event does not have a statistically significant effect for pickups in 2023. When we look at only basketball games, we notice that the estimates for drop-offs and pickups in 2019 are approximately equal to the estimates when that dataset considers all events. However, the statistical significance of the estimate for drop-offs in 2023 in slightly reduced; the effect on pickups in 2023 is still statistically insignificant. The F-statistic is greater than the critical value of 10 for all regressions.

Table 5 shows the results for the reduced form regressions, i.e., the effect of an "active" event on the match rate as given in (8). Active events have a positive effect on match rate. However, the effect is statistically insignificant for 2023 if we only consider basketball games. Table 6 shows

[5] https://www.barclayscenter.com/events-tickets/event-calendar



(a) 2023



(b) 2019

FIGURE 10: This figure shows the number of shared trips requested to be picked up and dropped off in the four taxi zones around the Barclays Center. The plots show a small increase in dropoffs before an event starts and an increase in pickups a couple hours after the event ends.

TABLE 4: First stage regressions

	Dropoff 2019	Pickup 2019	Dropoff 2023	Pickup 2023
All events				
active	27.622***	39.255***	1.053*	0.324
	(3.558)	(4.757)	(0.491)	(0.659)
Num.Obs.	8015	8015	7271	7272
R2	0.484	0.548	0.281	0.445
F	312.032	403.535	118.117	241.831
Basketball games				
active	28.731**	45.387***	2.241+	-0.653
	(10.694)	(11.857)	(1.337)	(1.414)
Num.Obs.	1560	1560	1690	1690
R2	0.142	0.313	0.147	0.373
F	19.721	54.101	22.214	76.743

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 5: Reduced form regressions

	Dropoff 2019	Pickup 2019	Dropoff 2023	Pickup 2023
All events				
active	0.026***	0.028***	0.079***	0.063***
	(0.005)	(0.006)	(0.015)	(0.018)
Num.Obs.	8015	8015	7239	7242
R2	0.194	0.389	0.085	0.139
F	80.082	212.070	27.753	48.606
Basketball games				
active	0.021*	0.029**	0.019	0.008
	(0.009)	(0.010)	(0.031)	(0.029)
Num.Obs.	1560	1560	1690	1690
R2	0.105	0.347	0.144	0.199
F	14.018	63.277	21.600	31.993

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 6: IV regressions

	Dropoff 2019	Pickup 2019	Dropoff 2023	Pickup 2023
All events				
shared requested	0.001***	0.001***	0.075*	0.194
	(0.000)	(0.000)	(0.035)	(0.387)
Num.Obs.	8015	8015	7239	7242
Wu-Hausman (p-val)	0.21	0.79	9.9e-7	4.61e-3
Basketball games				
shared requested	0.001**	0.001***	0.009	-0.012
	(0.000)	(0.000)	(0.014)	(0.055)
Num.Obs.	1560	1560	1690	1690
Wu-Hausman (p-val)	0.54	0.57	0.69	0.73

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the results for the *Instrumental variable* (IV) regression conducted using the ivreg package in R[6]. The estimated coefficient for the IV regression can also be obtained by dividing the estimate from the reduced form regression by the estimate of the first stage regression.

The IV estimate is positive, meaning the an increase in the number of shared trips requested *causes* the match rate to increase, whenever the relationship is statistically significant. We also note that the magnitude of scale economies have changed significantly between 2019 and 2023. The most probable reason could be that the the relationship between demand for shared trips and quality of matching is not linear. We assume a linear relationship D^s and η because the goal of this study was to causally establish the existence of scale economies; a robustness check with log-log regression also confirmed the presence of scale economies.

The IV estimate for some regressions are not statistically significant because binary z_d does not introduce enough variation in D^s to estimate the supply curve. As a robustness check, we run another regression on the pickup data from 2023 where we use the number of people entering the Atlantic-Avenue Barclays Center subway station[7] each hour as the instrument; this regression also includes fixed effect for hour-of-day. The IV estimate for γ_1 in (1) is 0.0215 (p-value 0.002) showing that the system exhibits scale economies. The regression also rejects the Null hypothesis for the Weak Instrument and Wu-Hausman test.

This analysis shows that ride-splitting does exhibit scale economies. However, the significance of the estimate also depends on the choice of the instrument. Future work can explore how different sources of exogenous variation affect the estimation; ideally all instruments should result in similar estimates.

[6] `https://cran.r-project.org/web/packages/ivreg/vignettes/ivreg.html`

[7] Available since February 2022 at `https://data.ny.gov/Transportation/MTA-Subway-Hourly-Ridership-Beginning-February-202/wujg-7c2s/data`

CONCLUSION

Ride-splitting has the potential to make urban transportation more sustainable while reducing travel costs for passengers. Despite heavy investments by TNCs and tax incentives by governments, only a small proportion or ride-hailing trips request to share their rides. An even smaller number of trips actually get matched.

In this paper we provide plausible explanations for different challenges to ride-splitting, pre- and post-COVID. Possibly, the pre-COVID decline in ride-splitting was largely a result of Uber's withdrawal of subsidies. We also hypothesize that ride-splitting's lack of rebound post-COVID can possibly be a result of the system being stuck in a low-sharing equilibrium. The presence of scale economies enables the possibility of multiple equilibria. Finally, we use an Instrumental Variables based approach to empirically establish the existence of scale economies using different regressions.

ACKNOWLEDGMENTS

This research was supported by National Science Foundation Grants CMMI-2052337, CMMI-2052512 and CMMI 23-39943 CAR.

REFERENCES

1. Shaheen, S., A. Cohen, I. Zohdy, and United States. Federal Highway Administration, *Shared Mobility: Current Practices and Guiding Principles*, 2016.
2. Shaheen, S. and A. Cohen, Shared Ride Services in North America: Definitions, Impacts, and the Future of Pooling. *Transport Reviews*, Vol. 39, No. 4, 2019, pp. 427–442.
3. Manjoo, F., Car-Pooling Helps Uber Go the Extra Mile. *The New York Times*, 2016.
4. Liu, H., S. Devunuri, L. Lehe, and V. V. Gayah, Scale effects in ridesplitting: A case study of the City of Chicago. *Transportation Research Part A: Policy and Practice*, Vol. 173, 2023, p. 103690.
5. Lehe, L., S. Devunuri, J. Rondan, and A. Pandey, Taxation of Ride-hailing. *FHWA-ICT-21-029*, 2021.
6. Kidd, C., *Matchmaking in Lyft Line*, 2018, accessed: 2024-07-28.
7. Castillo, J. C. and S. Mathur, Matching and Network Effects in Ride-Hailing. *AEA Papers and Proceedings*, Vol. 113, 2023, pp. 244–247.
8. Lehe, L., V. V. Gayah, and A. Pandey, Increasing returns to scale in carpool matching: Evidence from Scoop. *Transport findings*, 2021.
9. Ke, J., Z. Zheng, H. Yang, and J. Ye, Data-driven analysis on matching probability, routing distance and detour distance in ride-pooling services. *Transportation Research Part C: Emerging Technologies*, Vol. 124, 2021, p. 102922.
10. Kitamura, R., S. Nakayama, and T. Yamamoto, Self-Reinforcing Motorization: Can Travel Demand Management Take Us out of the Social Trap? *Transport Policy*, Vol. 6, No. 3, 1999, pp. 135–145.
11. Ying, J. Q. and H. Yang, Sensitivity Analysis of Stochastic User Equilibrium Flows in a Bi-Modal Network with Application to Optimal Pricing. *Transportation Research Part B: Methodological*, Vol. 39, No. 9, 2005, pp. 769–795.
12. David, Q. and R. Foucart, Modal Choice and Optimal Congestion. *Regional Science and Urban Economics*, Vol. 48, 2014, pp. 12–20.
13. Lawler, R., New Ride-Sharing App Hitch Tries To Pair You With A Driver... And Other Passengers | TechCrunch, 2014.
14. Russell, K., UberPool Lets You Split Uber Fares With Other Passengers Along The Same Route. *TechCrunch*, 2014.
15. Lawler, R., *With Lyft Line, Passengers Can Split Fares For Shared Rides*, 2014.
16. Lawler, R., *Lyft Acquires Shared Ride Startup Hitch To Bolster Its Lyft Line Service*, 2014.
17. Manjoo, F., Car-Pooling Helps Uber Go the Extra Mile. *The New York Times*, 2016.
18. Marshall, A., Lyft Redesigns Its App—and Strategy—for the Age of Sharing. *Wired*, 2018.
19. Stock, b., *Uber Newsroom*, 2018.
20. Westhagen, A., *Introducing Shared Saver, Our Most Affordable Ride | Lyft*, 2019.
21. Hawkins, A. J., Lyft Shuttle Mimics Mass Transit with Fixed Routes and Fares. *The Verge*, 2017.
22. Duffy, C., Uber Will Soon Let Riders Book a Shuttle to the Airport, 2024.
23. Uber, *Uber Shuttle FAQ | Riders | Uber Help*, 2024.
24. Peters, J., Uber Reintroduces Shared Rides with a New Name. *The Verge*, 2021.
25. Bellon, T., Lyft to Resume Shared Rides in U.S. for First Time since Pandemic | Reuters, 2021.
26. Lyft, *Shared Rides Return to More Cities*, 2022.

27. Bellan, R., Lyft Might Drop Shared Rides, Stay Focused on Basics under New CEO. *TechCrunch*, 2023.
28. Martin, C., Granting Shares for Fares: An Uber Rival's Play for Drivers. *The New York Times*, 2016.
29. Via, *Public Mobility Solutions. | Via Transportation*, 2024.
30. Uber, *UberX Share | Share Your Ride to Save*, 2024.
31. Dotan, T., Uber's Cheap and Popular Pool Ridesharing Service Lost the Company as Much as $1 Million a Week in San Francisco Alone. *Business Insider*, 2021.
32. Liu, H., S. Devunuri, L. Lehe, and V. V. Gayah, Scale Effects in Ridesplitting: A Case Study of the City of Chicago. *Transportation Research Part A: Policy and Practice*, Vol. 173, 2023, p. 103690.
33. Abkarian, H., S. Hegde, and H. S. Mahmassani, Does Taxing TNC Trips Discourage Solo Riders and Increase the Demand for Ride Pooling? A Case Study of Chicago Using Interrupted Time Series and Bayesian Hierarchical Modeling. *Transportation Research Record*, Vol. 2677, No. 1, 2023, pp. 212–223.
34. Macdonald, A., *Uber Newsroom*, 2022.
35. Davalos, J., You'll No Longer Be Able to Take a Carpool Ride with Lyft to Save Money. *Fortune*, 2023.
36. Wooldridge, J. M., Introductory Econometrics: A Modern Approach, 2012.
37. Romer, P., *The Trouble With Macroeconomics*, 2016.