RESEARCH REPORT

Institute of Transportation Studies | UNIVERSITY OF CALIFORNIA

Where Ridehail Drivers Go Between Trips: Trading off Congestion and Curb Availability?

Adam Millard-Ball, Associate Professor, Department of Urban Planning, UCLA
Liwei Liu, PhD Candidate, Economics Department, UC Santa Cruz
Whitney Hansen, PhD Student, Environmental Studies Department, UC Santa Cruz
Drew Cooper, Senior Transportation Modeler, San Francisco County Transportation Authority
Joe Castiglione, Deputy Director for Technology, Data, and Analysis, San Francisco County Transportation Authority

June 2021

UC SANTA CRUZ

Report No.: UC-ITS-2020-15 | DOI: 10.7922/G2QV3JTZ

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UC-ITS-2020-15	N/A	N/A

4. Title and Subtitle	5. Report Date
Where Ridehail Drivers Go Between Trips: Trading off Congestion and Curb Availability?	June 2021
	6. Performing Organization Code UC Santa Cruz

7. Author(s)	8. Performing Organization Report No.
Adam Millard-Ball, https://orcid.org/0000-0002-2353-8730 Liwei Liu, https://orcid.org/0000-0002-2236-5081 Whitney Hansen, https://orcid.org/0000-0002-6986-0645 Drew Cooper, https://orcid.org/0000-0002-7240-1845 Joe Castiglione, https://orcid.org/0000-0002-3931-6286	N/A

9. Performing Organization Name and Address	10. Work Unit No.
Environmental Studies Department UC Santa Cruz 1156 High St Santa Cruz, CA 95064	N/A
	11. Contract or Grant No. Same number as box 1.

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
The University of California Institute of Transportation Studies www.ucits.org	Final Report (September 2019 – January 2021)
	14. Sponsoring Agency Code UC ITS

15. Supplementary Notes
DOI:10.7922/G2QV3JTZ

16. Abstract
We analyze what ridehail drivers do when out of service between paid trips. We use a dataset of 5.3 million trips in San Francisco and partition each out of service trip into cruising, repositioning, and parking segments. We find that repositioning accounts for nearly two-thirds (63%) of the time between trips, with cruising and parking accounting for 23% and 14% respectively (these figures exclude short trips). Our regression models suggest that drivers tend to make reasonable choices between repositioning and parking, heading to high-demand locations based on the time of day. However, we also find suggestive evidence of racial bias, supporting previous studies of both taxis and ridehailing that indicate that drivers tend to avoid neighborhoods with high proportions of residents of color.

17. Key Words	18. Distribution Statement
Ridesourcing, professional drivers, regression analysis, data analysis, traffic congestion, travel demand, parking, social equity	No restrictions.

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	21. Price
Unclassified	Unclassified	39	N/A

Form Dot F 1700.7 (8-72) Reproduction of completed page authorized

About the UC Institute of Transportation Studies

The University of California Institute of Transportation Studies (UC ITS) is a network of faculty, research and administrative staff, and students dedicated to advancing the state of the art in transportation engineering, planning, and policy for the people of California. Established by the Legislature in 1947, ITS has branches at UC Berkeley, UC Davis, UC Irvine, and UCLA.

Acknowledgments

This study was made possible with funding received by the University of California Institute of Transportation Studies from the State of California through the Public Transportation Account and the Road Repair and Accountability Act of 2017 (Senate Bill 1). The authors would like to thank the State of California for its support of university-based research, and especially for the funding received for this project. The authors would also like to thank Bhargava Sana for contributing ideas and helping with traffic data; Christo Wilson and Alan Mislove at Northeastern University for capturing the original data set; and Mark Garrett for editorial assistance.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the State of California in the interest of information exchange. The State of California assumes no liability for the contents or use thereof. Nor does the content necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification, or regulation.

Institute of Transportation Studies | UNIVERSITY OF CALIFORNIA

Where Ridehail Drivers Go Between Trips: Trading off Congestion and Curb Availability?

Adam Millard-Ball, Associate Professor, Department of Urban Planning, UCLA
Liwei Liu, PhD Candidate, Economics Department, UC Santa Cruz
Whitney Hansen, PhD Student, Environmental Studies Department, UC Santa Cruz
Drew Cooper, Senior Transportation Modeler, San Francisco County Transportation Authority
Joe Castiglione, Deputy Director for Technology, Data, and Analysis, San Francisco County Transportation Authority

June 2021

UC SANTA CRUZ

Report No.: UC-ITS-2020-15 | DOI: 10.7922/G2QV3JTZ

Table of Contents

Table of Contents

List of Tables

List of Figures

Executive Summary

Executive Summary

App-based ridehailing services such as Uber and Lyft have revolutionized travel in urban centers in recent years. A large number of studies show that ridehailing improves mobility but increases vehicle travel by inducing more trips and shifting trips away from walking and public transit. The impacts of ridehailing, however, depend not only on the travel decisions of riders — the number of trips they take and their choice of mode — but also on the choices of ridehail drivers when they are out of service between paid rides. These choices about where to go between trips not only determine the overall impacts of ridehailing on vehicle travel and associated congestion and pollution, but also on parking availability.

In this report, we quantify the choices that ridehail drivers make between paid trips. We develop a method to partition each trip into **cruising** (circling around while looking for a fare), **repositioning** (moving to another location in anticipation of the next ride request), and **parking** segments, and apply it to a dataset of 5.3 million trips in San Francisco. We also develop a regression model to quantify the factors associated with driver choices.

We find that the average out-of-service segment between paid trips lasts 4.1 minutes, during which time drivers travel 1.0 km (0.6 miles). The average in-service vehicle trip length is 4.2 km (2.6 miles), meaning that the out-of-service portion accounts for 19 percent of ridehail vehicle travel. Our estimated proportion of 19 percent is lower than the roughly 40 percent typically cited in the literature, but our data excludes deadheading between a driver's home and when they turn on the ridehail app for the first time, and the portion of the trip between accepting a ride request and picking up the passenger. High demand and short distances within San Francisco may also account for our lower estimate.

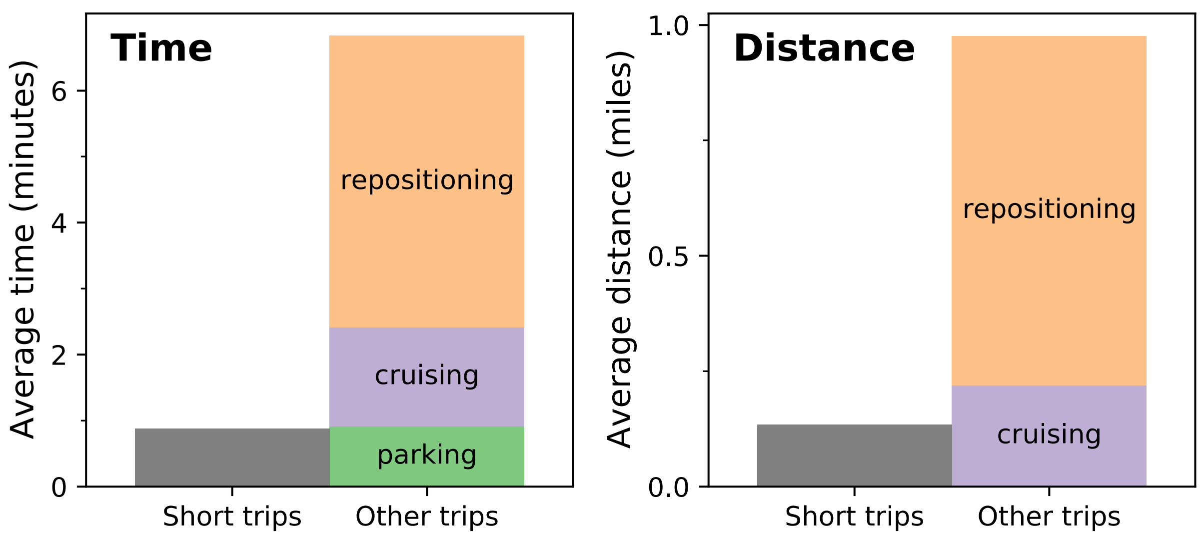

Figure ES-1. Driver behavior when out-of-service

Repositioning accounts for nearly two-thirds (63 percent) of out-of-service time, with parking accounting for a further 14 percent (these figures exclude short trips). Both repositioning and parking can represent rational behavior on the part of drivers seeking to minimize downtime and maximize revenue from their next trip. Indeed, our regression models suggest that drivers tend to make reasonable choices between repositioning and parking, heading to high-demand locations based on the time of day. For example, we find they tend to reposition away from dense residential neighborhoods in the afternoon and evening when demand is likely to be higher in other areas, but stay within those neighborhoods in the morning and at night. However, we also find suggestive evidence of racial bias, supporting previous studies of both taxis and ridehailing that indicate that drivers tend to avoid neighborhoods with high proportions of residents of color.

From a public interest perspective, it is unclear whether a city should favor repositioning or parking. More time spent parked means less vehicle travel, but potentially greater impacts on the availability of curbspace for parking, and potentially for drop-offs and deliveries as well. More repositioning, on the other hand, decreases the pressure on curb space and allows the fleet to operate more efficiently by moving to higher-demand areas, but at the cost of more vehicle travel and the associated congestion and pollution.

Cruising accounts for 23 percent of the time between trips. Compared to repositioning and parking, cruising would seem to offer little advantage to a ridehail driver in most parts of the city, at least if parking is readily available. Cruising also exacerbates congestion and pollution. Explaining why ridehail drivers cruise is beyond our ability to answer with the present dataset, and future research might usefully probe driver decision-making processes. In some cases, a lack of available curb space or high levels of parking enforcement may be the cause. Possibly, drivers believe that they can game the ridehail firms' matching system by driving around to be closer to potential passengers, and thus being allocated a trip. Alternatively, psychological factors may be at work. More experienced drivers cruise less, suggesting that drivers learn over time that cruising is a suboptimal strategy.

Cruising might be partly reduced through tweaks to driver-facing ridehail apps, prompting drivers to find a safe place to park while waiting for their next ride. Cities, meanwhile, might consider how ridehailing can take advantage of curbspace in front of residential driveways and other curb cuts that are used only occasionally. Our results suggest that some ridehail drivers already park in front of driveways on an informal basis, as they can quickly move if a resident needs to access their garage.

Ultimately, however, revising fee structures to be distance- and time-based, regardless of whether a passenger is in the vehicle, may be the most effective way for cities to reduce the external costs of ridehailing including congestion and pollution. Ridehail firms would pay these fees, and determine whether and how to pass them on to passengers. In addition, place-based time charges might be used as a proxy for parking fees, and to encourage drivers to park in locations where they do not compete with other curbspace users. Many cities already levy ridehail fees or taxes on a per-trip or percentage basis, but these charges only apply to the in-service portion of a trip. To more comprehensively address pollution, congestion, and other externalities caused by ridehailing, policy makers need to extend these policies to encompass what drivers do between trips.

Contents

Introduction

App-based ridehailing services, often known as Transportation Network Companies (TNCs), have revolutionized travel in urban centers in recent years. TNC firms such as Uber and Lyft often provide more abundant, reliable, and cheaper service than taxis, their closest competitor (Brown & LaValle, 2021), leading to rapid growth in ridership. Within San Francisco, for example, ridehailing accounted for 15 percent of vehicle trips in 2016 (SFCTA, 2017).

A large number of studies have analyzed the consequences of ridehailing for travel behavior and congestion. The most common finding is that ridehailing induces users to make more trips, and that it shifts trips away from private cars, walking, and public transit (Rayle et al., 2016; Hampshire et al., 2017; Clewlow & Mishra, 2017; Gehrke et al., 2019; Babar & Burtch, 2020; Bradley et al., 2021). In San Francisco, ridehailing has been the largest contributor to increased congestion in recent years (Erhardt et al., 2019). However, in some cases, ridehailing can complement transit use by filling gaps in the reach of scheduled bus and rail services (Hall et al., 2018). It can also improve mobility, particularly in neighborhoods where car ownership is low (Brown, 2019a) and for older adults (Leistner & Steiner, 2017).

Less attention, however, has been paid to the strategies of ridehailing drivers, and in particular what they do when out of service between paid rides. Most analyses focus on the paid, with-passenger portion of a ridehail trip, but out-of-service travel, such as deadheading to the next trip and cruising while waiting for a trip request, may have major consequences for the environment and congestion. Driver choices regarding whether and where to park while waiting for the next trip also affect curbspace and parking availability. Thus, understanding out-of-service travel is important for developing municipal policies for regulating and pricing ridehail services, such as congestion surcharges, and for allocating and pricing curbspace (Strong, 2015; Li et al., 2019; Marsden et al., 2020).

In this report, we quantify the choices that ridehail drivers make between paid trips. We develop a method to partition each trip into cruising, repositioning, and parking segments, and apply it to a dataset of 5.3 million trips in San Francisco. We find that while almost all trips involve repositioning (traveling to another location in anticipation of receiving a trip request), a surprising portion (29 percent) entail at least some cruising. We develop a regression model to quantify the factors associated with driver choices, and find that ridehail drivers appear to reposition to neighborhoods where ridehail demand is high, but the model also suggests that drivers avoid neighborhoods with high proportions of residents of color.

The rapid growth of ridehailing mean that our findings are relevant to policymakers dealing with present-day transportation challenges. However, our results also provide a preview of what might be expected in a future with autonomous vehicles, whose transportation and environmental consequences may bear many parallels to those of ridehailing.

Driver behavior: comparing taxis and ridehail

Taxi drivers in large cities often cruise along busy streets in search of a street hail, or reposition to major trip generators such as airports and hotels. In New York City, for example, cruising and repositioning account for 44 percent of miles driven by taxicabs, with an average of 2.9 miles of out-of-service travel between trips (Abrams et al., 2007, p. 124). Driving around rather than waiting at a taxi stand may be rational from the driver's perspective, as it makes the vacant taxicab visible to prospective passengers, but its impacts from a social welfare perspective are mixed. On the one hand, a ready supply of available taxis reduces wait times for passengers, but cruising taxis are highly visible contributors to congestion. Thus, limiting cruising has often been a key goal of taxicab regulators and a justification for limits on the number of taxicabs (Shreiber, 1975; Yang et al., 2005; Abrams et al., 2007).

Another long-standing regulatory challenge has been to ensure the availability of taxis in low-income neighborhoods and communities of color, which typically experience longer wait times. Drivers often decline to accept calls for service to such neighborhoods, and also tend to reposition away from them after dropping off a passenger due to perceptions of lower demand, fears for their personal safety, and racial profiling (Davis, 2003; Ingram, 2003; Brown, 2019b). Regulatory responses have included enforcement "stings," but also programs such as New York City's "green cabs," which can only pick up passengers outside of the high-demand areas of Lower Manhattan and the airports (King & Saldarriaga, 2018).

To what extent do these findings translate from taxis to ridehailing? Both sets of drivers should seek to maximize the expected net revenue from their next paid trip, and thus minimize wait time and the distance driven while cruising or deadheading. The options open to taxi and ridehail drivers are also similar. They can park (or equivalently, wait at a taxi stand), cruise around while remaining in the same general neighborhood, or reposition to a different neighborhood where they expect demand to be higher. Cruising and repositioning are often conflated as "deadhead" trips in the literature (e.g. Henao & Marshall, 2019; Nair et al., 2020), but conceptually the two categories are distinct.

While the options of taxi and ridehail drivers may be similar, their optimal strategies are likely to be considerably different because their costs and sources of information differ in four main respects. First, while taxi drivers must normally be conspicuous to passengers hailing a taxi on the street,[1] most cities prohibit ridehail drivers from accepting street hails, and in any case the app-based system used by ridehail firms renders such visibility unnecessary. Second, a first-in, first-out rule typically applies at taxi stands at hotels, airports, and other major trip generators. In contrast, ridehail drivers are subject to the opaque methods that ridehail firms use to match drivers with passengers, and the incentives that the firms use to encourage drivers to head to specific locations and to start or extend their shifts. Third, while taxi drivers might rely on heuristics or experience to identify high-demand locations, ridehail drivers have access to real-time information on demand patterns through their smartphone app. Fourth, taxi drivers may have lower costs for repositioning if, as in cities such as San Francisco, they have access to bus lanes or dedicated taxi stands.

As a result, one would expect ridehail drivers to cruise less frequently than taxi drivers. While cruising may be rational for taxi drivers, it is less reasonable for ridehail drivers unless parking is limited or expensive. For a ridehail driver, parking is likely to provide similar prospects to cruising in terms of obtaining the next paid ride, without the costs of fuel and vehicle wear and tear. Since drivers can easily move if and when an enforcement officer arrives, they have little need to pay for

[1] This discussion focuses on street hail taxis, rather than telephone dispatch systems which are more similar to ridehailing in the incentive structures that they provide to drivers.

parking either. The relative advantages of repositioning for taxis and ridehailing, in contrast, are not intuitively clear, but one might expect shifts in the destinations and times of repositioning. Given the dynamic information available to ridehail drivers, they might be expected to reposition to a broader range of destinations, not just the hotels and airports that are obvious sources of demand for taxis (Dempsey, 1996; Schaller, 2007).

Little empirical work, however, exists to support or refute these hypotheses. Data sharing by ridehail firms such as Uber and Lyft has been extremely limited, meaning that most researchers have focused on the paid portion of the trip which is easier to observe through field or household surveys (e.g. Grahn et al., 2019; Brown & LaValle, 2021). Out-of-service behavior is harder to identify, and often, the distance driven while cruising and repositioning is simply assumed (e.g. Tirachini & Gomez-Lobo, 2020) or simulated based on assumptions of rational driver behavior (e.g. Komanduri et al., 2018; Gurumurthy et al., 2020). In almost all travel demand models, the vehicle dematerializes after dropping off a passenger, only to reappear on the network at the start of the next paid trip.

Among the exceptions, Henao and Marshall (2019) find that out-of-service travel accounts for 41 percent of the miles driven by ridehail drivers, but this estimate is based on data from a single driver — the first author. Several studies use a dataset released by RideAustin to impute out-of-service travel based on pick-up and drop-off locations. While the actual paths taken by drivers are uncertain, the data indicate that 37 to 45 percent of total miles driven were by out-of-service vehicles (Komanduri et al., 2018; Wenzel et al., 2019). In San Francisco, preliminary analysis of the dataset used for this study yielded a figure of 19-21 percent, depending on the day, compared to 44-46 percent for taxicabs (SFCTA, 2017). In California as a whole, analysis of data provided by ridehail firms (under a legal requirement) indicates that out-of-service travel accounts for 39.5 percent of miles driven (CARB, 2019). Geographically, the broadest estimates are made by Cramer and Krueger (2016) using proprietary data provided by Uber; they find that out-of-service travel accounts for 39 percent of miles by Uber drivers across five major cities. Finally, a study commissioned by Uber and Lyft puts the proportion of out-of-service travel at 38-46 percent in a different set of six metropolitan regions (Fehr and Peers, 2019). Their breakdown indicates that 28-37 percent of the distance is driven while waiting for a ride request, and 9-10 percent while driving to the pick-up location after accepting a request.

With the exception of the San Francisco study, which is an outlier perhaps in part due to its exclusion of the distance between accepting a ride and picking up the passenger, these estimates are remarkably consistent. They suggest that out-of-service travel by ridehail vehicles is substantial at about 40 percent of the total distance driven. This consistency comes in spite of different methodologies, data sources, and scopes — for example, whether they consider travel between a driver's home and the first activation of the ridehail app, or whether they consider cruising or assume shortest-path travel distances. Surprisingly, estimates of out-of-service travel for ridehail services are not much less than those for taxis, in spite of the information advantages held by the former.

Studies of racial equity, meanwhile, suggest that discrimination still exists in the ridehail market, although perhaps to a lesser extent than with conventional taxis. At the individual level, a field audit that requested rides in Boston found that cancellations doubled when using an African American-sounding name rather than a white-sounding name (Ge et al., 2020). On the other hand, another study indicates that aggregate wait times for ridehailing requests in Seattle are no worse in neighborhoods with a higher proportion of people of color, after controlling for residential and employment densities and average income (Hughes & MacKenzie, 2016).

Research Approach

Ridehail Data

We used a unique dataset of 5.3 million ridehail trips in San Francisco from November 12, 2016 through December 21, 2016, compiled by researchers at Northeastern University by querying the Uber and Lyft Application Programming Interfaces (APIs) which give access to vehicle locations. Further details of data acquisition and processing are elaborated in Cooper et al. (2018). The dataset has been used in several empirical analyses, most notably an assessment of the congestion impacts of ridehailing in San Francisco (Erhardt et al., 2019), and a profile of TNC activity in San Francisco (SFCTA, 2017). However, these analyses focus on the occupied portions of the rides, rather than the out-of-service portions on which we focus here.

Each trip in the dataset consists of a sequence of points with geographic coordinates and a timestamp. On average, the points are 3.0 seconds apart. We cleaned the dataset to drop points with invalid coordinates, restricted the dataset to trips within the city of San Francisco, and excluded shared (e.g., Lyft Line) and delivery (e.g., UberEats) trips. Note that the dataset only includes points when the ridehail app is turned on and the driver is available to accept a ride.

We map-matched each trip to the OpenStreetMap road network in order to provide more accurate estimates of driving distances that are not affected by irregularities in the GPS trace. We used a three-stage process: (i) matching GPS points to OpenStreetMap (OSM) links using Mapillary's publicly available map-matching algorithm,[2] (ii) dropping links where the preceding and succeeding links directly connect, in order to eliminate out-and-back detours down side streets, and (iii) interpolating gaps in the link sequence using the turn-restricted shortest path function in the pgRouting software package.

Classification of Behavior

We classified each point[3] as short, parking, cruising, or repositioning as follows:

Short points are those on trips where either (i) there are fewer than six GPS points or (ii) the trip duration is less than two minutes. For these trips, it was not possible to determine the driver's intent. Except where indicated, short trips are excluded from the subsequent analysis.

Parking points are defined as a cluster of points within any three-minute interval where at least 90 percent of the points are within 7.5 meters of each other. After identifying these clusters, each point within the cluster was classified as parking, and the parking location was defined as the closest point to the centroid of the cluster. To avoid classifying vehicles stuck in congested traffic as parked, we created exceptions where time- and location-specific traffic speeds (obtained from

[2] Code is available at https://github.com/caomw/map_matching
[3] We did not classify the first point in each GPS trace, because the classification of each point is based on the driver's behavior between that point and the previous point.

INRIX) were less than three mph, or where the GPS point was on a freeway. In these instances, the parking classification was not applied.

Cruising points are those that involve circling or backtracking. We first identified cruising at the trip level using the definition in Weinberger et al. (2020) — where the actual (map-matched) distance is at least 200 meters longer than the shortest-path network distance. Within each cruising trip, however, the driver may not be cruising the entire time. Therefore, we identified the cruising portion of each trip as a function of the path of the squared displacement — the squared (Euclidean) distance from each point to the origin. This metric is often used in movement ecology studies to distinguish the movements of individual animals, such as deer collared with a GPS tracker, and can distinguish between migratory, non-migratory, and dispersing behavior (Killeen et al., 2014; Singh et al., 2016).

Specifically, if we plot the squared displacement over time, a positive slope indicates that the driver is moving away from the origin. A negative slope shows that the driver is returning towards their origin (i.e., the start of the out-of-service segment). After smoothing the standardized slope,[4] consecutive points with a slope of +1 form a positive segment, and consecutive points with a slope of -1 form a negative segment. We therefore classified a point as cruising if the trip involves cruising per the definition above and either (i) the point is on a negative segment, or (ii) the point is on a positive segment, but its squared displacement is offset by a subsequent negative segment. Figure 1 provides an example.

Repositioning points constitute the remainder of the data set. In other words, all other points (i.e., those that are not classified as parking, cruising, or short) were classified as repositioning.

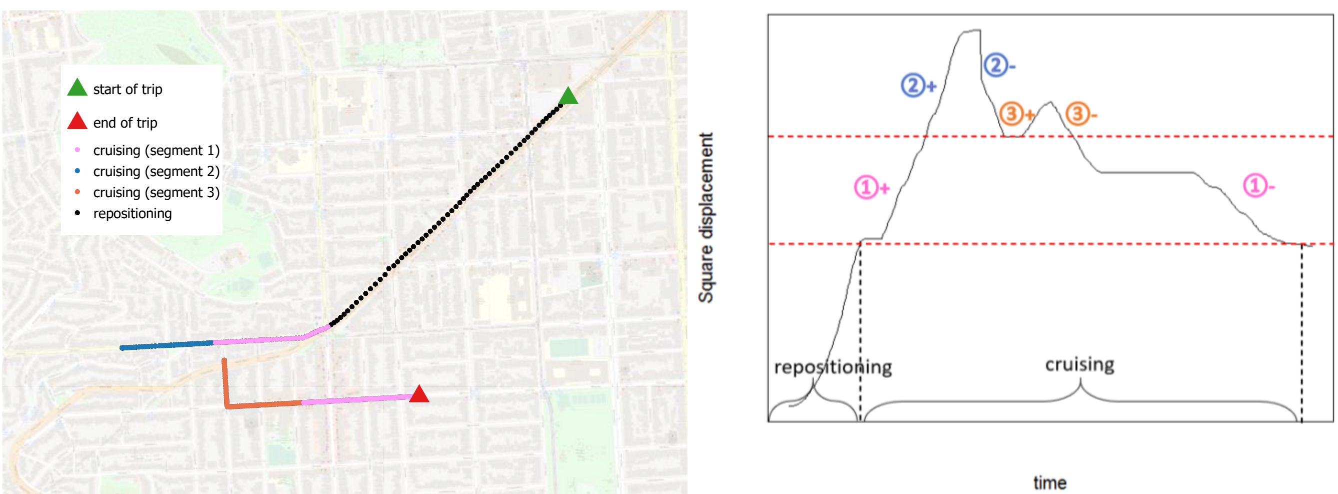

Figure 1. Example of cruising and repositioning segments

The driver's route is shown in the left panel, with the right panel showing how squared displacement changes over the route. The first segment (marked in black) is classified as repositioning. The subsequent segments are classified as cruising, either because

---

[4] The smoothing method considers five points before and after the current point. If these points have the same normalized slope of square displacement, and the current point is also within 30 seconds of these points, we assign the same value of the normalized slope of these 10 points to the current point.

they have negative squared displacement, or because a positive segment is offset by a subsequent negative segment, as shown by the three pairs of segments labeled in the figure as (1+) and (1-), (2+) and (2-), and (3+) and (3-).

Other Data Sources

Each GPS point in the database was associated with a number of independent variables as shown in Table 1. These variables, such as household density, age of population, income, and race and ethnicity, mostly reflect the characteristics of the Transportation Analysis Zone (TAZ), the geographical unit used in transportation analysis by the San Francisco County Transportation Authority (SFCTA), in which the point is located. We produced a weighted average for each point by aggregating the values of the variables for the TAZ containing the point and neighboring TAZs where the TAZ values were weighted using a distance decay function. This smoothing algorithm avoids abrupt changes in the values of the variables at TAZ boundaries, and also reflects how drivers are likely to perceive gradual changes in neighborhood demographics and parking supply. There are 981 TAZs in San Francisco, with a mean surface area of 0.12 km^2. We added lagged dependent variables (indicating prior driver behavior) and time of day and day of week variables for each point. For Lyft trips, we also calculated driver experience, measured as the number of trips by that particular driver observed in the dataset. (The Uber API does not provide a persistent driver identifier.)

Regression Analysis

We used multinomial logistic regression to estimate the effects of the various variables in Table 1 on the driver's decisions to reposition, cruise, or park. The model tested whether these decisions could be explained based on any of the TAZ level variables in Table 1 or differed based on the time or day of the week. Details of the regression modeling are provided in the Appendix.

Table 1. Descriptive Statistics

Description	Mean	Standard Deviation	Data Source
TAZ-level variables			
Household density, calculated by the number of households divided by TAZ area	26.827	20.276	SFCTA
Share of the population age 62 or over	0.180	0.068	SFCTA
Fraction of population of working age (aged 20-64)	0.678	0.151	SFCTA
Total employment density, calculated as the total employment divided by TAZ area	99.428	175.733	SFCTA
Service and visitor employment density, calculated as the sum of service and visitor employment divided by TAZ area	20.802	31.453	SFCTA
Fraction of high-income households, calculated as the fraction of households in the highest and second highest income quantiles	0.391	0.137	SFCTA
On-street parking capacity (number of spaces)	238.156	138.104	SFCTA
Off-street public parking capacity (number of spaces)	115.835	144.410	SFCTA
Off-street private parking capacity (estimated number of spaces)	338.821	194.098	SFCTA
Fraction of population that is Hispanic / Latino	0.123	0.098	US Census
Fraction of population that is African American	0.059	0.062	US Census
Fraction of population that is white	0.536	0.172	US Census
Other variables			
Driver experience, calculated by counting the number of trips for each driver [only available for the Lyft subsample]	212.283	139.593	Calculated
Day of week, a categorical variable: "weekday" for Monday to Thursday, "fri" for Friday, "weekend" for Saturday and Sunday	-	-	Calculated
Time of day, binary variables indicating 6 periods: "ea" for 3-6 am, "am" for 6-9 am, "md" for 9 am-3:30 pm, "pm" for 3:30-6:30 pm, "ev" for 6:30 pm-0:00 am, "night" for 0-3 am.	-	-	Calculated
Lag cruising. Dummy variable that is 1 if the previous point is classified as cruising (only for point level regression)	-	-	Calculated
Lag parking. Dummy variable that is 1 if the previous point is classified as parking (only for point level regression)	-	-	Calculated

Results

Classification of driver behavior

We begin by presenting the broad patterns of driver behavior in terms of the choices between parking, cruising, and repositioning. Table 2 shows the percentage of time and distance driven in each of the categories. Repositioning accounts for the majority of the time and distance traveled by out-of-service drivers, and almost all trips involve at least a small amount of repositioning. Perhaps surprisingly given the fuel and wear-and-tear costs of cruising, more time is spent cruising than parking, and the average trip cruises for nearly half a kilometer.

As shown in Table 2, the average out-of-service distance traveled is 0.98 km (0.6 miles). The average in-service vehicle trip length is 4.2 km (2.6 miles), based on a previous analysis of the same dataset (SFCTA, 2017). Therefore, the out-of-service portion accounts for 19 percent of ridehail vehicle travel. Note that this is a lower bound estimate as it excludes travel before and after the driver activates the app, and the portion of the out-of-service trip between accepting a ride request and picking up the passenger.

Drivers for the two ridehail firms operating in San Francisco — Uber and Lyft — spend almost identical proportions of their time across the three categories of parking, cruising, and repositioning. However, out-of-service trips are longer for Lyft drivers (5.5 minutes and 1.35 km, compared to 3.6 minutes and 0.86 km for Uber drivers). Lyft drivers also have a smaller proportion of short trips (36 percent compared to 46 percent for Uber). Since Uber accounts for three-quarters of the trips in our sample, it is likely that economies of scale lead to their drivers obtaining a paid fare more quickly, reducing the amount of out-of-service travel required.

There is surprisingly little geographic variation in the three behaviors across the city (Figure 2). Drivers finding themselves in the ring of dense residential neighborhoods around the downtown core are more inclined to park rather than reposition or cruise, but the effects are not strong. Northeastern San Francisco — the densest part of the city — accounts for the largest share of out-of-service time (Figure 2) and trip starts and ends (Figure 3). There is a noticeable concentration of trip starts on freeway corridors, perhaps reflecting drivers turning on their app as they enter the city. Otherwise, there is no obvious geographic pattern in the number of out-of-service trip ends minus the number of trip starts (net trip flows), with Figure 3 showing a patchwork quilt across the city. The exception is along freeways, where for obvious reasons there is a net movement away from these facilities.

Table 2. Classification of out-of-service driver behavior

	% of trips	Mean time per trip	Mean distance per trip
Excluding short trips			
Parking	10%*	0.9 mins (14%)	N/A
Cruising	29%*	1.5 mins (23%)	0.35 km (22%)
Repositioning	86%*	4.1 mins (63%)	1.22 km (78%)
All trips	100%	6.6 mins	1.57 km
Including short trips			
Short	44%	0.9 mins	0.22 km

	% of trips	Mean time per trip	Mean distance per trip
Not short	56%	6.6 mins	1.57 km
All trips	100%	4.1 mins	0.98 km

* Trips may include multiple behaviors. This column counts trips that have at least one point in a given behavior.

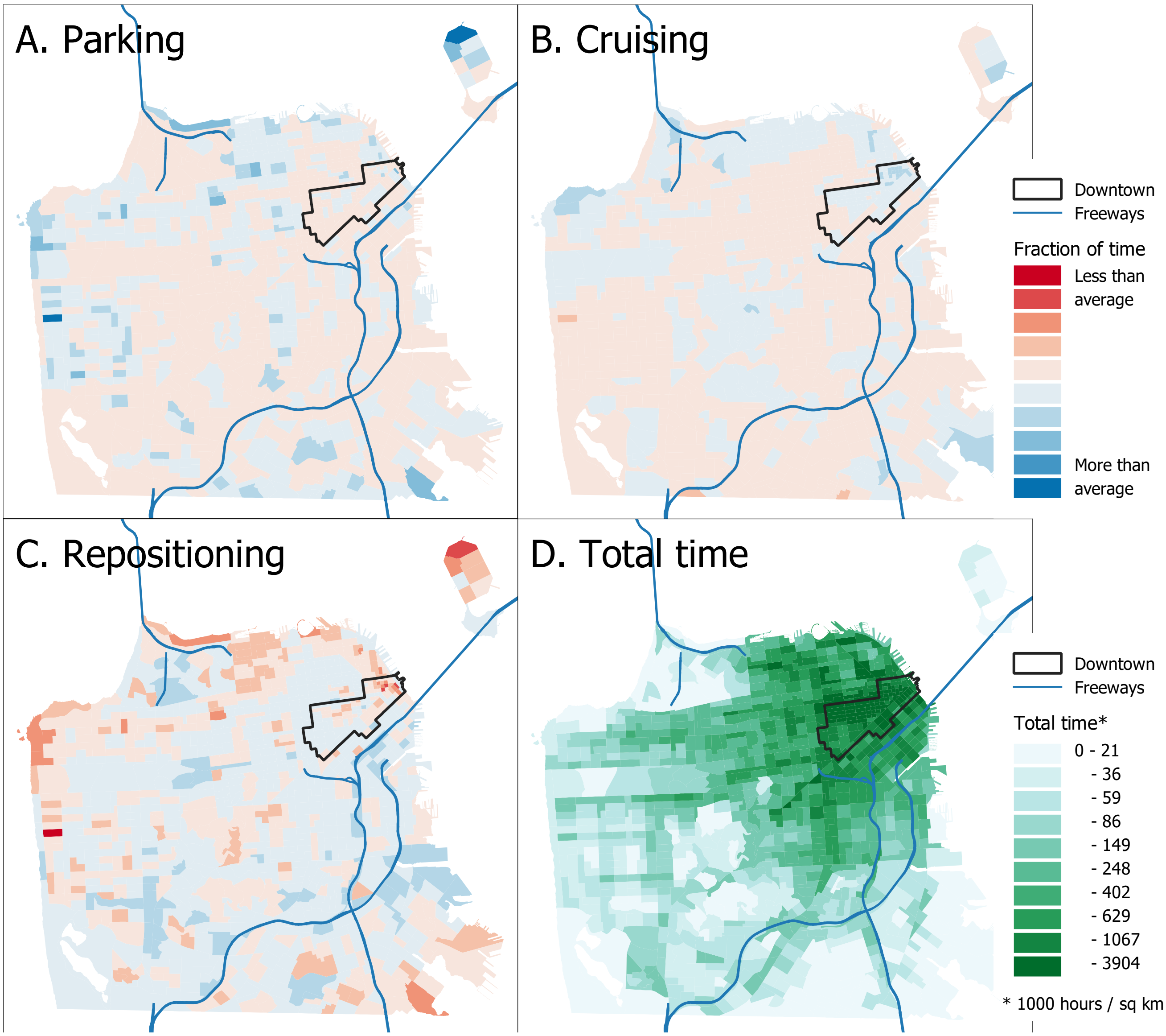

Figure 2. Geographic patterns in parking, cruising, parking, and repositioning

Panels A, B and C show the fraction of time within each TAZ spent parking, cruising, and repositioning respectively. Each category spans a ten percentage point range (e.g. 40-50% below average, 30-40% below, etc.) Panel D shows the distribution of out-of-service time across the city, normalized to land area and expressed as thousand hours per square kilometer.

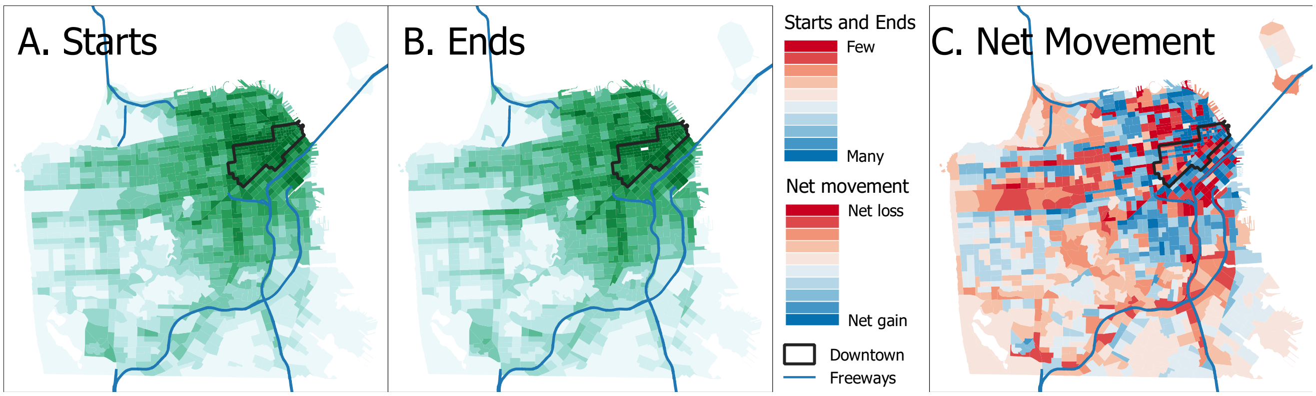

Figure 3. Net out-of-service flows

Panels A and B and C show the number of out-of-service trip starts and trip ends in each TAZ, normalized to area and express as deciles. Panel C shows the net movement, with red-shaded TAZs having more trip starts than ends (a net movement away) and blue-shaded TAZs having more trip ends than trip starts.

Parking

We now consider the characteristics of parking events. The map in Figure 4 (left panel) shows a concentration in the inner ring of dense residential neighborhoods. Within this general area, however, drivers find a range of parking options. Off-street parking is most visibly concentrated in grocery store surface parking lots, gas stations, and similar locations, where drivers may be able to linger for a short time before being moved on by security staff or parking attendants. On-street parking is spread more diffusely, but concentrations are evident along neighborhood commercial corridors. In some cases, ridehail drivers park on blocks where driveways, fire hydrants, loading zones, or other restrictions preclude parking for regular vehicles, but mean that curb space is readily usable by ridehail drivers who can quickly move if needed. These concentrations are most visible in an interactive online version of the parking map (right panel of Figure 4), available at https://tncparking.sfcta.org.

Overall, almost all the time spent parking (93 percent of the total duration) occurs on-street. Non-metered on-street spaces (both legal and illegal) account for the majority of ridehail parking, with the largest share (31 percent) occurring on residential streets (Table 3 and Figure 5). Parking at meters accounts for just over one-third of the aggregate time spent parked, but given that most drivers do not park at all while out-of-service, this amounts to only 12 seconds in the average trip. Thus, on a per-trip basis, the impact on parking availability is minimal, as is the revenue loss to the City (less than half a cent). However, given the 1.2 million ridehail trips per week in late 2016 (SFCTA, 2017), aggregate meter revenue amounts to more than $200,000 per year, based on the typical meter rate of $2.50 per hour. This calculation also excludes time spent while loading or unloading passengers at meters, and stays of less than three minutes (the minimum length of a parking event in the analysis).

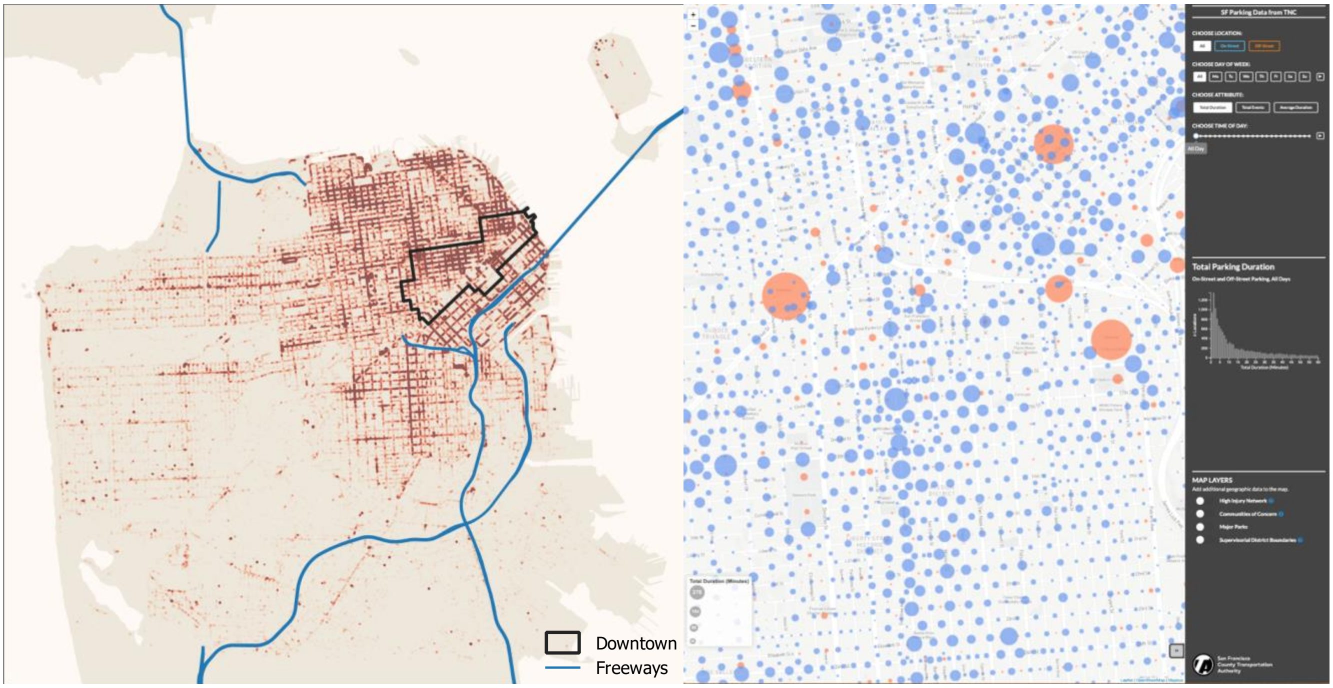

Figure 4. Concentrations of parking locations

Each location is weighted by the length of time parked. The right panel shows a screenshot from the interactive online map available at https://tncparking.sfcta.org. Blue symbols denote on-street parking, and red symbols denote off-street parking, with a gas station and surface lots at two grocery stores being readily apparent.

Table 3. Time spent parked (hours per week)

	Metered hours*	Non-metered hours
Parking meter	1,562 (37%)	2,281 (37%)
On-street: residential	1,303 (31%)	1,918 (31%)
On-street: other	992 (24%)	1,557 (25%)
Off-street	353 (8%)	368 (6%)
Total	4,209	6,124

** Typically Mon-Sat 9am-6pm. A total of 37 percent of trips in our dataset take place during metered hours.*
Includes trips with no parking events. Time spent parked is calculated on a per-trip basis, and scaled up to a weekly aggregate based on 1.2 million ridehail trips per week in late 2016 (SFCTA, 2017).

Parking meter locations are defined as those within 10 meters of a parking meter. We used OpenStreetMap (OSM) to identify residential streets and off-street parking locations (signified by a "service" classification in OSM, which typically consists of access roads or parking aisles in surface lots). We were unable to identify parking garages.

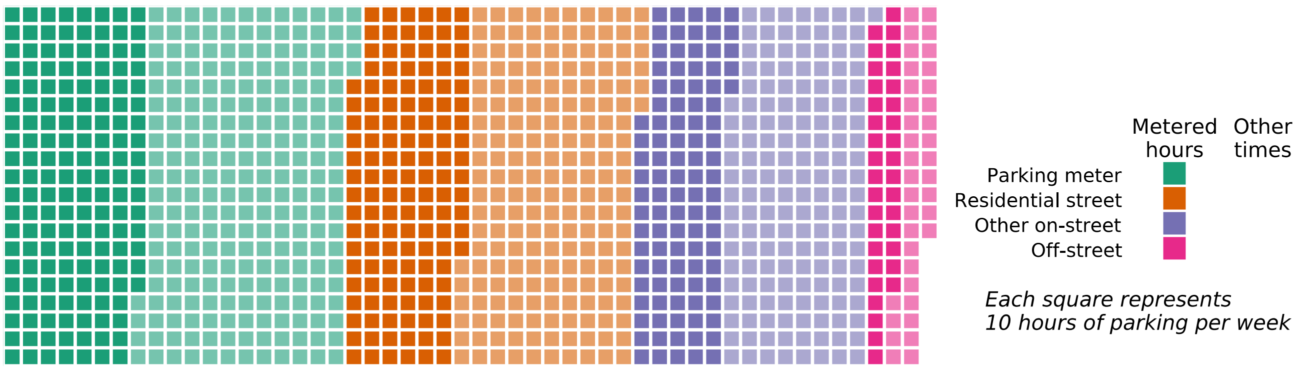

Figure 5. Distribution of time spent parked

Determinants of driver behavior

We now consider the associations between neighborhood characteristics and a driver's decision to park, cruise, or reposition, using the logistic regression models discussed in the Research Approach section. Two coefficients are attached to each variable, indicating the associated change in the probability of repositioning and cruising respectively, compared to a baseline behavior of parking. All coefficients are shown in Table A-1 and, with the confidence intervals graphically represented, in Figure A-2.

The variables are standardized, and so each coefficient represents the effect of a one-standard deviation change. A positive sign indicates that that behavior is more likely compared to parking, and a negative sign that it is less likely. For example, drivers are less likely to reposition away from TAZs with a high proportion of White residents (coefficient of –0.059), and slightly less likely to cruise (–0.005), compared to parking.

Given the large sample size, most of the coefficients are statistically significant at conventional levels. However, they are hard to interpret given that there are three separate behaviors (parking, cruising, and repositioning); and interaction terms that allow our density coefficients to vary by time of day and day of week (right columns in Table A-1). Therefore, Figure 6 plots the effects of each variable in terms of the probabilities of each behavior. Several findings emerge from these analyses.

Ridehail drivers tend to reposition away from neighborhoods with more parking, especially on-street parking as shown in Figure 6A. This perhaps indicates that individuals might choose to drive themselves to neighborhoods with plentiful parking, meaning less demand for ridehail services in these areas. This demand-side effect appears to outweigh the advantage to ridehail drivers of readily available parking.

Drivers also tend to reposition away from neighborhoods with a higher proportion of residents of color, and do the opposite in neighborhoods with more White residents (Figure 6C-E).[5] These findings provide suggestive evidence that drivers avoid neighborhoods with more people of color, supporting the findings of the earlier research on both ridehail and taxi drivers discussed above.

As seen in Figure 6G-I, the effects of density are somewhat counterintuitive. Drivers are more likely to reposition away from neighborhoods with higher residential or service employment density, even though these types of neighborhoods might be expected to generate more ridehail trips, whether due to the presence of bar and restaurant customers or the lower car ownership rates seen in dense residential neighborhoods. In contrast, drivers are less likely to reposition away from neighborhoods with a higher density of non-service employment. However, a more intuitive picture emerges when we consider how the effects of density change over the course of the day and week, through the interaction terms in the regression model. As illustrated in Figure 7, while there is little change in the effect of density throughout the week, there are strong time-of-day effects. Drivers are more likely to reposition away from dense residential neighborhoods in the afternoon and evening, and less likely to do so in the morning and at night, presumably when more people are at home to request ridehail trips. The opposite patterns are seen with employment density (but not service and visitor employment density), with drivers more likely to reposition away from job-rich areas in the mornings, presumably when potential customers are traveling from home to work. In addition to perceptions of demand, lack of parking and traffic congestion may also be factors that affect repositioning decisions.

Figure 6F also plots the effects of driver experience (estimated using the Lyft subsample only). Full-time drivers are less likely to cruise and more likely to reposition, suggesting that they are more aware of areas of high demand.

[5] The baseline category in the regressions is the fraction of Asian residents. The negative coefficient for the fraction of White residents shows that drivers are less likely to reposition away from a TAZ if it has a higher fraction of White residents compared to Asian residents. While the coefficient for Black residents is also negative, it is much smaller than that for White residents. The positive coefficient for Latinx residents means that repositioning away is even more likely, again compared to the baseline of Asian residents.

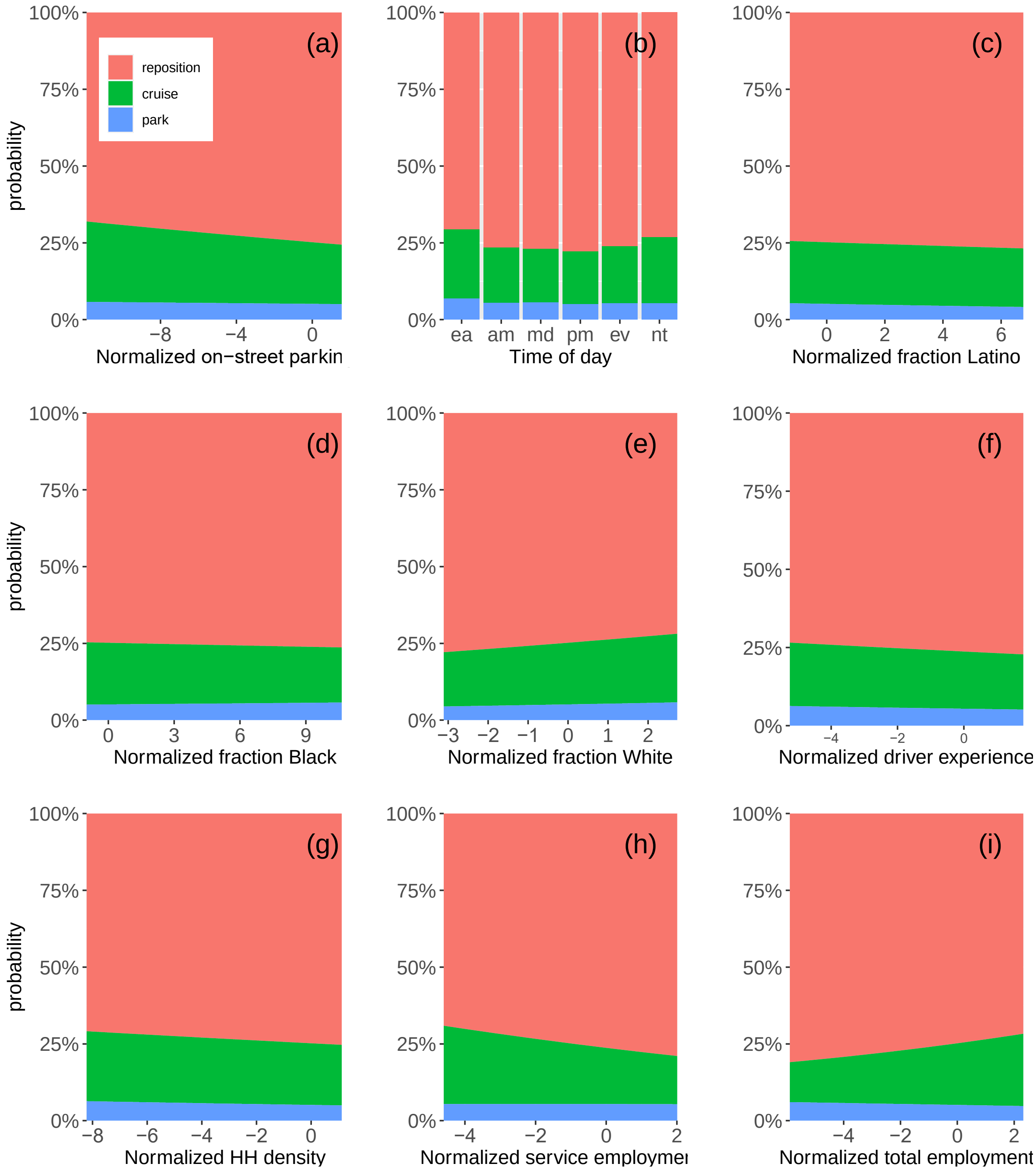

Figure 6. Probability of specific behaviors given change in key variables

The plots show the probability of repositioning, cruising, and parking against changes in several key independent variables, which are normalized so that the x-axis indicates standard deviations from the mean. All other variables are held at their means. For example, the upper-left plot shows that repositioning is the most common behavior, but even more so in high-residential density neighborhoods. As the prevalence of repositioning increases with density, that of cruising declines, while parking remains at similar levels.

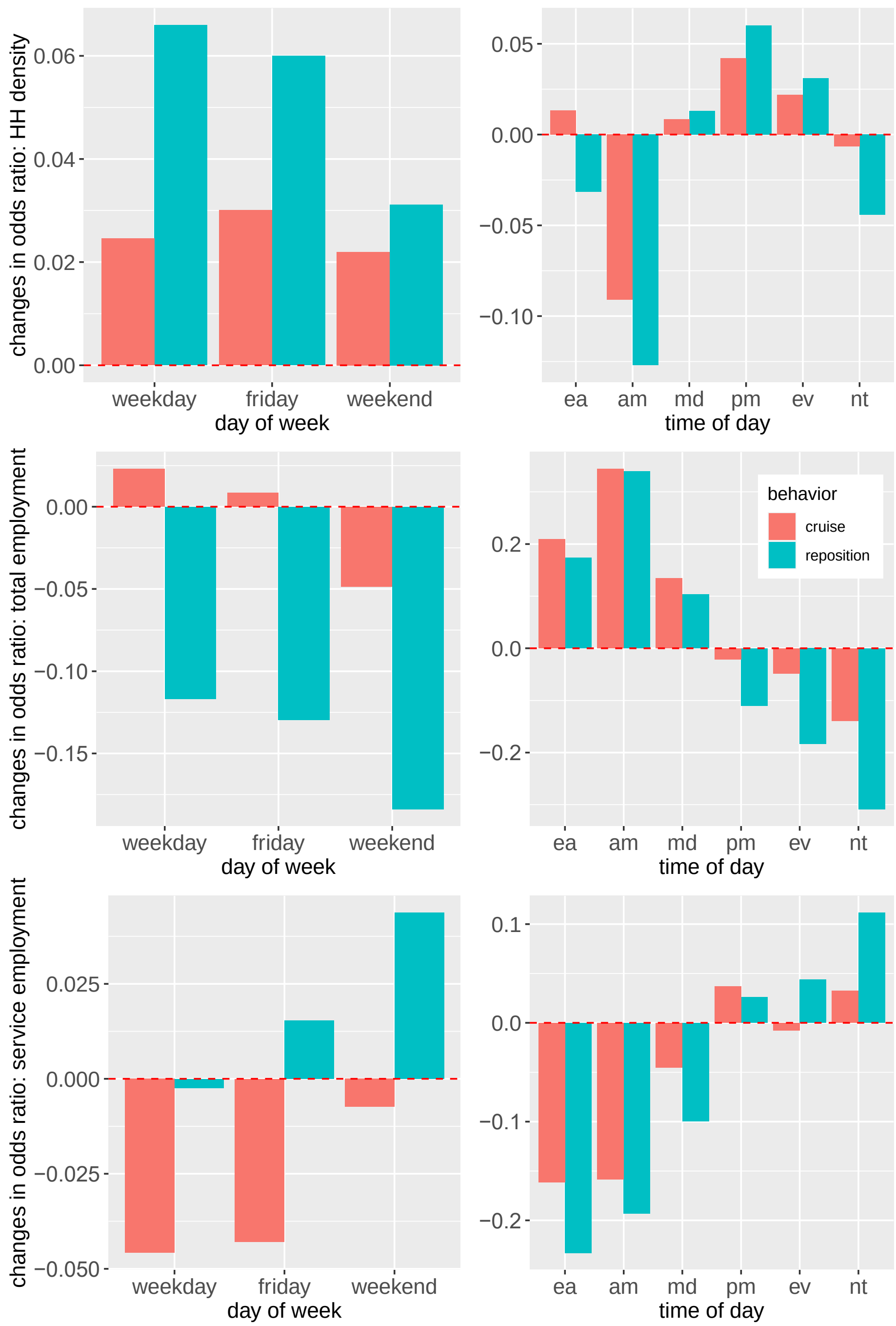

Figure 7. Effect of density on driver behavior by day of week and time of day

The plots show how the changes in probability (measured by odds ratios) for residential, employment, and service/visitor density vary with the time of day and day of week. Positive changes mean that the probability of cruising (red bars) or repositioning (blue bars) increases more than the baseline probability of parking, and vice versa. For example, the impact of household density is similar on weekdays and weekend days (top left plot). But drivers are more likely to cruise in and reposition away from higher-density neighborhoods in the afternoon and evening, and less likely at night and in the mornings (top right plot).

Conclusions

The choices made by ridehail drivers about where to go between trips determine the overall impacts of ridehailing on vehicle travel and associated congestion and pollution, as well as on parking availability. There is a tradeoff between the two — more time spent parked means less vehicle travel, but potentially greater impacts on the availability of curbspace for parking, drop-offs and deliveries. More out-of-service vehicle travel, on the other hand, decreases the pressure on curb space and allows the fleet to operate more efficiently through repositioning to higher-demand areas, but at the cost of more vehicle travel, congestion and pollution.

Such tradeoffs between parking demand and vehicle travel would also apply to future autonomous vehicles, as demonstrated by Kondor et al. (2020) in the Singapore context. For a given number of trips, the more that the deployment of autonomous vehicles lowers parking demand, the greater the distance driven by out-of-service vehicles. In this report, we provide the first analysis of how ridehail drivers make these tradeoffs using a dataset of 5.3 million out-of-service trips in San Francisco.

We find that the average out-of-service segment between paid trips lasts 4.1 minutes, during which time drivers travel 1.0 km (0.6 miles). The average in-service vehicle trip length is 4.2 km (2.6 miles), meaning that the out-of-service portion accounts for 19 percent of ridehail vehicle travel. Our estimated proportion of 19 percent is lower than the roughly 40 percent typically cited in the literature, but our data excludes deadheading between a driver's home and when they turn on the ridehail app for the first time, and the portion of the trip between accepting a ride request and picking up the passenger. High demand and short distances within San Francisco may also account for our lower estimate.

We classify points on each out-of-service trip as cruising, repositioning, or parking. Both repositioning and parking can represent rational behavior on the part of drivers seeking to minimize downtime and maximize revenue from their next trip. Indeed, our regression models suggest that drivers tend to make reasonable choices between repositioning and parking, heading to high-demand locations based on the time of day. For example, they reposition away from dense residential neighborhoods in the afternoon and evening when demand is likely to be higher in other areas, but stay within those neighborhoods in the morning and at night. However, we also find suggestive evidence of racial bias, supporting previous studies of both taxis and ridehailing (Ingram, 2003; Ge et al., 2020) that indicate that drivers tend to avoid neighborhoods with high proportions of people of color.

While cruising by traditional taxicabs makes them visible to potential passengers, it would seem to offer little advantage to a ridehail driver who can simply park instead. Therefore, perhaps our most surprising finding is that cruising accounts for 23 percent of the out-of-service time and 22 percent of the out-of-service distance driven by ridehail drivers (excluding short trips). Cruising in lieu of parking means that the impacts on curb occupancy and meter revenue loss are smaller than might be expected, but those on congestion, pollution, and the other consequences of vehicle travel are greater.

Why do ridehail drivers cruise? This question is beyond our ability to answer with the present dataset, and future research might usefully probe driver decision-making processes. In some cases, a lack of available curb space or high levels of parking enforcement may be the cause. Possibly, drivers believe that they can game the trip allocation system by driving around to be closer to potential passengers, and thus being allocated a trip. Alternatively, psychological factors may be at work. More experienced drivers cruise less, suggesting that drivers learn over time that cruising is a suboptimal strategy.

Cruising might partly be reduced through tweaks to driver-facing ridehail apps, prompting drivers to find a safe place to park while waiting for their next ride. Cities, meanwhile, might consider how ridehailing can take advantage of curbspace in front of residential driveways and other curb cuts that are used only occasionally. Some ridehail drivers already park in front of driveways on an informal basis, as they can quickly move if a resident needs to access their garage.

Ultimately, however, revising fee structures to be distance- and time-based, regardless of whether a passenger is in the vehicle, may be the most effective way for cities to reduce the external costs of ridehailing including congestion and pollution. Ridehail firms would pay these fees, and determine whether and how to pass them on to passengers. In addition, place-based time charges might be used as a proxy for parking fees, and to encourage drivers to park in locations where they do not compete with other curbspace users. Many cities already levy ridehail fees or taxes on a per-trip or percentage basis, but these charges only apply to the in-service portion of a trip. To more comprehensively address pollution, congestion, and other externalities caused by ridehailing, policy makers need to extend these policies to encompass what drivers do between trips.

References

Abrams, R., Harris, S., Marton, D., Mauldin, C., Millard-Ball, A., Sawhney Mittal, A., Rothman, E., & Weinberger, R. (2007). *Taxi 07: Roads Forward.* Design Trust for Public Space and New York City Taxi & Limousine Commission.

Babar, Y., & Burtch, G. (2020). Examining the Heterogeneous Impact of Ride-Hailing Services on Public Transit Use. *Information Systems Research*, *31*(3), 820–834. https://doi.org/10.1287/isre.2019.0917

Bradley, M., Greene, E., Sana, B., Cooper, D., Castiglione, J., Israel, S., & Coy, C. (2021). *Results of the First Large-scale Survey of TNC Use in the Bay Area*. Paper presented at Transportation Research Board Annual Meeting, Washington, DC.

Brown, A. (2019a). Redefining Car Access: Ride-Hail Travel and Use in Los Angeles. *Journal of the American Planning Association*, *85*(2), 83–95. https://doi.org/10.1080/01944363.2019.1603761

Brown, A. (2019b). Prevalence and Mechanisms of Discrimination: Evidence from the Ride-Hail and Taxi Industries. *Journal of Planning Education and Research*, 0739456X1987168. https://doi.org/10.1177/0739456X19871687

Brown, A., & LaValle, W. (2021). Hailing a change: Comparing taxi and ridehail service quality in Los Angeles. *Transportation*, *48*, 1007–1031. https://doi.org/10.1007/s11116-020-10086-z

CARB. (2019). *SB 1014 Clean Miles Standard. 2018 Base-year Emissions Inventory Report*. California Air Resources Board. https://ww2.arb.ca.gov/sites/default/files/2019-12/SB%201014%20-%20Base%20year%20Emissions%20Inventory_December_2019.pdf

Clewlow, R. R., & Mishra, G. S. (2017). *Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States* (UCD-ITS-RR-17-07). Institute of Transportation Studies, UC Davis.

Cooper, D., Castiglione, J., Mislove, A., & Wilson, C. (2018). Profiling Transport Network Company Activity using Big Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2672*(42), 192–202. https://doi.org/10.1177/0361198118798459

Cramer, J., & Krueger, A. B. (2016). Disruptive Change in the Taxi Business: The Case of Uber. *American Economic Review*, *106*(5), 177–182. https://doi.org/10.1257/aer.p20161002

Davis, D. L. (2003). Why Hailing a New Idea about Public Accommodation Laws May Be Easier Than Hailing a Taxi. *Valparaiso University Law Review*, *37*, 929–981.

Dempsey, P. S. (1996). Taxi Industry Regulation, Deregulation and Reregulation: The Paradox of Market Failure. *Transportation Law Journal*, *24*(1), 73–120.

Erhardt, G. D., Roy, S., Cooper, D., Sana, B., Chen, M., & Castiglione, J. (2019). Do transportation network companies decrease or increase congestion? *Science Advances*, *5*(5), eaau2670. https://doi.org/10.1126/sciadv.aau2670

Fehr and Peers. (2019). *Estimated TNC Share of VMT in Six US Metropolitan Regions*. https://www.fehrandpeers.com/what-are-tncs-share-of-vmt/

Ge, Y., Knittel, C. R., MacKenzie, D., & Zoepf, S. (2020). Racial discrimination in transportation network companies. *Journal of Public Economics*, *190*, 104205. https://doi.org/10.1016/j.jpubeco.2020.104205

Gehrke, S. R., Felix, A., & Reardon, T. G. (2019). Substitution of Ride-Hailing Services for More Sustainable Travel Options in the Greater Boston Region. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(1), 438–446. https://doi.org/10.1177/0361198118821903

Grahn, R., Harper, C. D., Hendrickson, C., Qian, Z., & Matthews, H. S. (2019). Socioeconomic and usage characteristics of transportation network company (TNC) riders. *Transportation*. https://doi.org/10.1007/s11116-019-09989-3

Gurumurthy, K. M., de Souza, F., Enam, A., & Auld, J. (2020). Integrating Supply and Demand Perspectives for a Large-Scale Simulation of Shared Autonomous Vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, *2674*(7), 181–192. https://doi.org/10.1177/0361198120921157

Hall, J. D., Palsson, C., & Price, J. (2018). Is Uber a substitute or complement for public transit? *Journal of Urban Economics*, *108*, 36–50. https://doi.org/10.1016/j.jue.2018.09.003

Hampshire, R. C., Simek, C., Fabusuyi, T., Di, X., & Chen, X. (2017). *Measuring the Impact of an Unanticipated Suspension of Ride-Sourcing in Austin, Texas* [Working Paper]. https://learn.sharedusemobilitycenter.org/wp-content/uploads/policy-documents-4/SSRN-id2977969.pdf

Henao, A., & Marshall, W. E. (2019). The impact of ride-hailing on vehicle miles traveled. *Transportation*, *46*(6), 2173–2194. https://doi.org/10.1007/s11116-018-9923-2

Hughes, R., & MacKenzie, D. (2016). Transportation network company wait times in Greater Seattle, and relationship to socioeconomic indicators. *Journal of Transport Geography*, *56*, 36–44. https://doi.org/10.1016/j.jtrangeo.2016.08.014

Ingram, J. D. (2003). Racial and Ethnic Profiling. *Thurgood Marshall Law Review*, *29*, 55–88.

Killeen, J., Thurfjell, H., Ciuti, S., Paton, D., Musiani, M., & Boyce, M. S. (2014). Habitat selection during ungulate dispersal and exploratory movement at broad and fine scale with implications for conservation management. *Movement Ecology*, *2*(1), 15. https://doi.org/10.1186/s40462-014-0015-4

King, D., & Saldarriaga, J. (2018). Spatial regulation of taxicab services: Measuring empty travel in New York City. *Journal of Transport and Land Use*, *11*(1). https://doi.org/10.5198/jtlu.2018.1063

Komanduri, A., Wafa, Z., Proussaloglou, K., & Jacobs, S. (2018). Assessing the Impact of App-Based Ride Share Systems in an Urban Context: Findings from Austin. *Transportation Research Record: Journal of the Transportation Research Board*, *2672*(7), 34–46. https://doi.org/10.1177/0361198118796025

Kondor, D., Santi, P., Le, D.-T., Zhang, X., Millard-Ball, A., & Ratti, C. (2020). Addressing the “minimum parking” problem for on-demand mobility. *Scientific Reports*, *10*, 15885. https://doi.org/10.1038/s41598-020-71867-1

Leistner, D. L., & Steiner, R. L. (2017). Uber for Seniors?: Exploring Transportation Options for the Future. *Transportation Research Record*, *2660*, 22–29. https://doi.org/10.3141/2660-04

Li, S., Tavafoghi, H., Poolla, K., & Varaiya, P. (2019). Regulating TNCs: Should Uber and Lyft set their own rules? *Transportation Research Part B: Methodological*, *129*, 193–225. https://doi.org/10.1016/j.trb.2019.09.008

Marsden, G., Docherty, I., & Dowling, R. (2020). Parking futures: Curbside management in the era of 'new mobility' services in British and Australian cities. *Land Use Policy*, *91*, 104012. https://doi.org/10.1016/j.landusepol.2019.05.031

Nair, G. S., Bhat, C. R., Batur, I., Pendyala, R. M., & Lam, W. H. K. (2020). A model of deadheading trips and pick-up locations for ride-hailing service vehicles. *Transportation Research Part A: Policy and Practice*, *135*, 289–308. https://doi.org/10.1016/j.tra.2020.03.015

Rayle, L., Dai, D., Chan, N., Cervero, R., & Shaheen, S. (2016). Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transport Policy*, *45*, 168–178. https://doi.org/10.1016/j.tranpol.2015.10.004

Schaller, B. (2007). Entry controls in taxi regulation: Implications of US and Canadian experience for taxi regulation and deregulation. *Transport Policy*, *14*(6), 490–506. https://doi.org/10.1016/j.tranpol.2007.04.010

SFCTA. (2017). *TNCs Today: A Profile of San Francisco Transportation Network Company Activity*. San Francisco County Transportation Authority. https://www.sfcta.org/projects/tncs-today

Shreiber, C. (1975). The Economic Reasons for Price and Entry Regulation of Taxicabs. *Journal of Transport Economics and Policy*, *9*(3), 268–279. http://www.jtep.org

Singh, N. J., Allen, A. M., & Ericsson, G. (2016). Quantifying Migration Behaviour Using Net Squared Displacement Approach: Clarifications and Caveats. *PLOS ONE*, *11*(3), e0149594. https://doi.org/10.1371/journal.pone.0149594

Strong, C. K. (2015). When Apps Pollute: Regulating Transportation Network Companies to Maximize Environmental Benefits. *University of Colorado Law Review*, *86*, 1049–1094.

Tirachini, A., & Gomez-Lobo, A. (2020). Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile. *International Journal of Sustainable Transportation*, *14*(3), 187–204. https://doi.org/10.1080/15568318.2018.1539146

Weinberger, R. R., Millard-Ball, A., & Hampshire, R. C. (2020). Parking search caused congestion: Where's all the fuss? *Transportation Research Part C: Emerging Technologies*, *120*, 102781. https://doi.org/10.1016/j.trc.2020.102781

Wenzel, T., Rames, C., Kontou, E., & Henao, A. (2019). Travel and energy implications of ridesourcing service in Austin, Texas. *Transportation Research Part D: Transport and Environment*, *70*, 18–34. https://doi.org/10.1016/j.trd.2019.03.005

Yang, H., Ye, M., Tang, W. H., & Wong, S. C. (2005). Regulating Taxi Services in the Presence of Congestion Externality. *Transportation Research: Part A: Policy and Practice*, *39*, 17–40. doi: 10.1016/j.tra.2004.05.004

Appendix

This appendix provides more technical details of the regression models and supplementary results.

We used multinomial logistic regression to estimate the effects of the various variables in Table 1 on the driver's decisions to reposition, cruise, or park. The model tested whether these decisions could be explained based on any of the TAZ level variables in Table 1 or differed based on the time or day of the week.

To avoid serial correlation of the error terms, we downsampled the data to 1-minute resolution. The downsampled dataset is about 5 percent of the full dataset. Because the distributions of most non-ratio numeric covariates are right-skewed, we applied a log transformation on the non-ratio covariates. This can further avoid serial correlation and strong effects from extreme values. Since the magnitudes of covariates have a large variation, we also normalized all numeric covariates by subtracting the mean and then dividing by the standard deviation.

We also tested the robustness of our results to key modeling assumptions in two ways. First, we used a nested logistic regression to model a process where drivers first choose between repositioning and remaining in the same area, and if the latter, choosing between cruising and parking. The hypothesis is that with low demand, drivers would prefer to reposition to another place, while with high demand the driver would choose between parking and cruising. Second, we aggregated the point level data to the TAZ level with different times of day and days of week, and then ran a fractional multinomial logistic regression of the ratio of points for each behavior on the covariates. Fractional logistic models are designed for aggregate data where the dependent variable is a proportion, rather than a binary or categorical outcome.

Figure A-1 plots the regression residuals, while Table A-1 and Figure A-2 present our main regression results. Tables A-2 through A-5 present alternative regression specifications to demonstrate the robustness of the results to alternative modeling assumptions.

These plots show the serial correlations of the residual terms of the regressions. The left plot is for the regression results reported in Table A-4, which indicate strong serial correlation that may affect the validity of the model. The right plot shows the residuals for our main regression with 1-minute resolution, as reported in Table A-1. The plot shows there is no strong serial correlation and thus the effectiveness of downsampling to one-minute resolution.

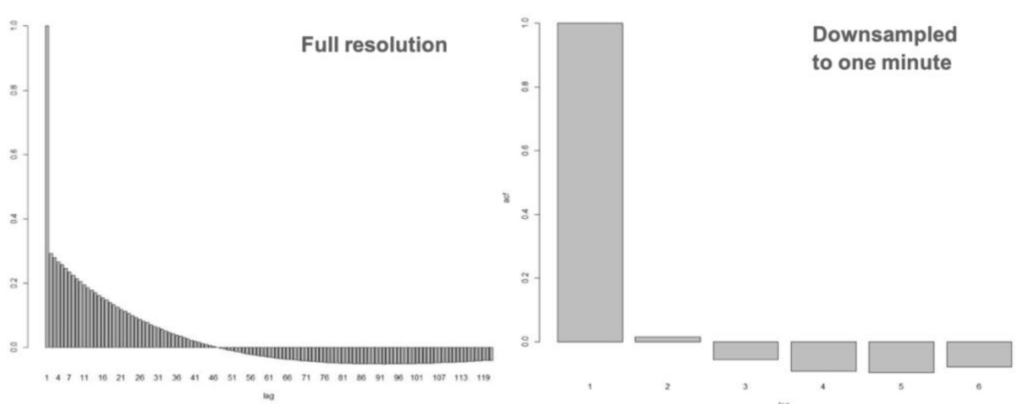

Figure A-1. ACF Plots of Regression Residuals

Table A-1. Regression coefficients (main specification)

Covariate	Reposition Coefficient	Cruise Coefficient
(Intercept)	3.822***	1.736***
Demographic and neighborhood variables		
Fraction age 62+	−0.048***	−0.037***
HH density	0.031***	0.022***
Fraction of working age	−0.013***	−0.051***
Employment density	−0.184***	−0.049***
Service and visitor employment density	0.044***	−0.007
Fraction high income HHs	−0.110***	−0.086***
On-street parking capacity	0.018***	−0.012***
Off-street parking capacity (public)	0.011***	−0.018***
Off-street parking capacity (residential)	0.013**	0.014**
Fraction Latinx residents	0.037***	0.026***
Fraction African-American residents	−0.009***	−0.021***
Fraction White residents	−0.059***	−0.005
Driver experience (Lyft subsample only)		
Driver experience	0.035***	0.008**
Time and day of week variables		
Time period: early AM	−0.303***	−0.097***
Time period: AM	−0.086***	−0.119***
Time period: midday	−0.092***	−0.125***
Time period: PM	0.061***	−0.040***
Time period: night	0.003	0.133***
Friday	0.063***	−0.023***
Mon-Thurs	−0.065***	−0.057***
Lagged dependent variables		
Lag cruise	−1.378***	1.769***
Lag park	−6.892***	−5.064***

*A positive coefficient for repositioning indicates that the driver is more likely to reposition **away** from a TAZ.*

Baseline (omitted) categories are the fraction of Asian residents, weekend days, and the evening time period.

Covariate	Reposition Coefficient	Cruise Coefficient
Interaction: with HH density		
Weekday	0.035***	0.003
Friday	0.029***	0.008
Time period: early AM	−0.063***	−0.009
Time period: AM	−0.158***	−0.113***
Time period: midday	−0.018***	−0.013*
Time period: PM	0.029***	0.020**
Time period: night	−0.075***	−0.029**
Interaction: with employment density		
Weekday	0.067***	0.071***
Friday	0.054***	0.057***
Time period: early AM	0.358***	0.257***
Time period: AM	0.524***	0.392***
Time period: midday	0.287***	0.183***
Time period: PM	0.073***	0.027*
Time period: night	−0.125***	−0.091***
Interaction: with service/visitor employment density		
Weekday	−0.046***	−0.038***
Friday	−0.028*	−0.036**
Time period: early AM	−0.277***	−0.154***
Time period: AM	−0.237***	−0.151***
Time period: midday	−0.143***	−0.038***
Time period: PM	−0.018	0.044***
Time period: night	0.068***	0.040*

*** $p<0.01$ ** $p<0.05$ * $p<0.1$

Residual Deviance: 6957399 on 15002084 degrees of freedom
Log-likelihood: −3478700 on 15002084 degrees of freedom

The estimates are from our preferred model (the point-level multinomial logistic regression at one-minute resolution). All coefficients are estimated from the full dataset, except that for driver experience, which is estimated from the Lyft subsample. For computational reasons, we use a random 40% subsample of the full dataset.

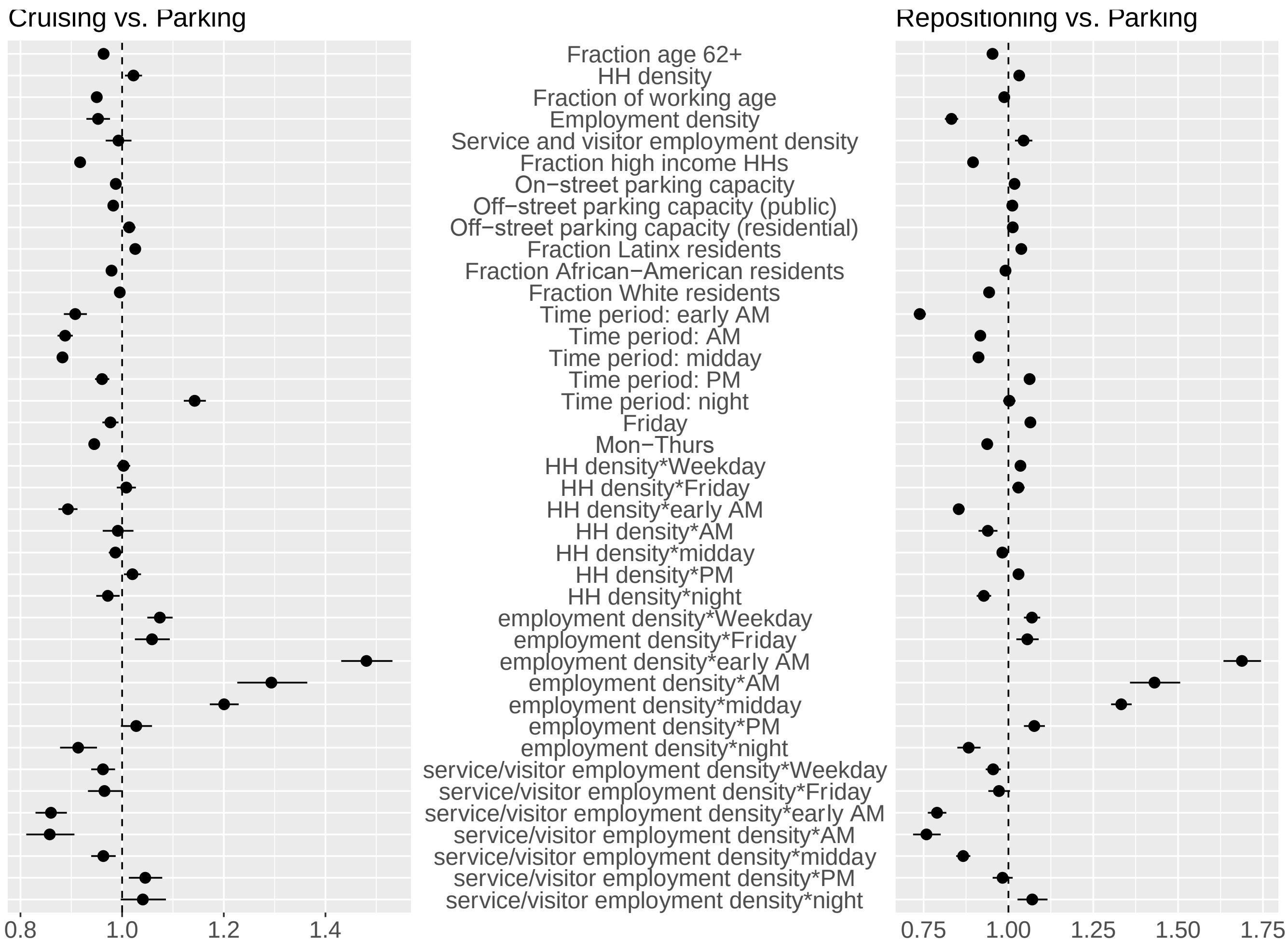

Figure A-2. Confidence intervals for regression coefficients

Note that the chart omits the lag behavior coefficients, which are much larger than the other covariates

Table A-2. Regression coefficients (Lyft Subsample)

Covariate	Reposition Coefficient	Cruise Coefficient
(Intercept)	3.670***	1.484***
Demographic and neighborhood variables		
Fraction age 62+	−0.042***	−0.052***
HH density	0.012	−0.021
Fraction of working age	−0.040***	−0.098***
Employment density	−0.148***	−0.061***
Service and visitor employment density	0.097***	−0.031
Fraction high income HHs	−0.101***	−0.040***
On-street parking capacity	0.004	−0.029***
Off-street parking capacity (public)	−0.008	−0.044***
Off-street parking capacity (residential)	−0.0001	0.039***
Fraction Latinx residents	0.032***	−0.0001
Fraction African-American residents	−0.018***	−0.056***
Fraction White residents	−0.038***	−0.001
Driver experience (Lyft subsample only)		
Driver experience	0.035***	0.008**
Time and day of week variables		
Time period: early AM	−0.313***	−0.044*
Time period: AM	0.003	−0.035**
Time period: midday	−0.027***	−0.101***
Time period: PM	0.102***	0.010
Time period: night	−0.015	0.175***
Friday	0.078***	−0.042***
Mon-Thurs	−0.034***	0.071***
Lagged dependent variables		
Lag cruise	−1.187***	2.234***
Lag park	−6.657***	−5.043***

*A positive coefficient for repositioning indicates that the driver is more likely to reposition **away** from a TAZ.*

Covariate	Reposition Coefficient	Cruise Coefficient
Interaction: with HH density		
Weekday	0.0003	−0.015
Friday	0.025	0.019
Time period: early AM	0.005	0.0005
Time period: AM	−0.114***	−0.112***
Time period: midday	0.030***	0.005
Time period: PM	0.042***	0.040***
Time period: night	−0.067***	−0.042*
Interaction: with employment density		
Weekday	0.108***	0.071***
Friday	0.093***	0.045
Time period: early AM	0.388***	0.462***
Time period: AM	0.403***	0.470***
Time period: midday	0.226***	0.219***
Time period: PM	0.056**	0.026
Time period: night	−0.120***	−0.042
Interaction: with service/visitor employment density		
Weekday	−0.051**	−0.049**
Friday	−0.041	−0.024
Time period: early AM	−0.342***	−0.279***
Time period: AM	−0.119***	−0.118***
Time period: midday	−0.072***	−0.009
Time period: PM	0.011	0.087***
Time period: night	0.062*	0.030

*** $p<0.01$ ** $p<0.05$ * $p<0.1$

Estimates are from a point-level multinomial logistic regression at one-minute resolution, as in Table 4. However, the data is limited to the Lyft subsample, in order to capture the effects of driver experience. For computational reasons, we estimate using a random 2% sample of trips.

Table A-3. Regression coefficients (Nested Logistic Regression)

Covariate	Reposition Coefficient	Cruise Coefficient
(Intercept)	2.596***	0.396***
Demographic and neighborhood variables		
Fraction age 62+	−0.025***	−0.009***
HH density	0.012	0.002
Fraction of working age	0.019***	−0.013***
Employment density	−0.152***	−0.018***
Service and visitor employment density	0.043***	0.002
Fraction high income HHs	−0.053***	−0.024***
On-street parking capacity	0.019***	−0.001
Off-street parking capacity (public)	0.021***	−0.001
Off-street parking capacity (residential)	0.008	0.002
Fraction Latinx residents	0.022***	−0.007***
Fraction African-American residents	0.003	−0.007***
Fraction White residents	−0.056***	−0.002
Time and day of week variables		
Time period: early AM	−0.245***	−0.019***
Time period: AM	−0.032***	−0.039***
Time period: midday	−0.029***	−0.035***
Time period: PM	0.061***	−0.017***
Time period: night	−0.101***	0.036***
Friday	0.068***	−0.003
Mon-Thurs	−0.050***	−0.010***
Lagged dependent variables		
Lag cruise	−2.850***	0.399***
Lag park	−5.727***	−1.145***

A positive coefficient for repositioning indicates that the driver is more likely to reposition ***away*** *from a TAZ.*

Covariate	Reposition Coefficient	Cruise Coefficient
Interaction: with HH density		
Weekday	0.037***	0.002
Friday	0.016*	−0.001
Time period: early AM	−0.075***	−0.001
Time period: AM	−0.102***	−0.031***
Time period: midday	−0.015**	0.001
Time period: PM	0.023***	0.007*
Time period: night	−0.029***	−0.001
Interaction: with employment density		
Weekday	0.033***	0.016***
Friday	0.016	0.007
Time period: early AM	0.225***	0.090***
Time period: AM	0.311***	0.116***
Time period: midday	0.175***	0.055***
Time period: PM	0.038***	0.005
Time period: night	−0.067***	−0.022**
Interaction: with service/visitor employment density		
Weekday	−0.042***	−0.010*
Friday	−0.004	−0.001
Time period: early AM	−0.197***	−0.061***
Time period: AM	−0.138***	−0.044***
Time period: midday	−0.113***	−0.018***
Time period: PM	−0.033**	0.012
Time period: night	0.027	0.009

Within-nest correlation: 0.227***

*** $p<0.01$ ** $p<0.05$ * $p<0.1$

The estimates are from a nested logistic regression, which first models the decision to stay in the neighborhood ("park or cruise," nest one) vs reposition away to a different neighborhood (nest two). Then, for the first nest, it models the decision to park vs cruise.

Table A-4. Regression coefficients (fractional logistic regression at the TAZ level)

Covariate	Reposition Coefficient	Cruise Coefficient
(Intercept)	2.468***	1.558***
Demographic and neighborhood variables		
Fraction age 62+	−0.106***	−0.092***
HH density	−0.105***	−0.019***
Fraction of working age	−0.014***	−0.093***
Employment density	−0.519***	−0.241***
Service and visitor employment density	0.069***	−0.088***
Fraction high income HHs	−0.201***	−0.120***
On-street parking capacity	0.106***	0.015***
Off-street parking capacity (public)	0.037***	−0.014***
Off-street parking capacity (residential)	0.114***	0.114***
Fraction Latinx residents	0.084***	0.054***
Fraction African-American residents	−0.043***	−0.082***
Fraction White residents	−0.083***	−0.028***
Time and day of week variables		
Time period: early AM	−0.584***	−0.512***
Time period: AM	−0.275***	−0.515***
Time period: midday	−0.272***	−0.526***
Time period: PM	−0.020***	−0.340***
Time period: night	−0.049***	−0.228***
Friday	0.050***	−0.058***
Mon-Thurs	−0.219***	−0.182***

A positive coefficient for repositioning indicates that the driver is more likely to reposition ***away*** *from a TAZ.*

Covariate	Reposition Coefficient	Cruise Coefficient
Interaction: with HH density		
Weekday	0.088***	0.033***
Friday	0.058***	0.012***
Time period: early AM	−0.070***	−0.041***
Time period: AM	−0.236***	−0.224***
Time period: midday	0.033***	−0.042***
Time period: PM	0.126***	0.036***
Time period: evening	0.103***	0.021***
Interaction: with employment density		
Weekday	0.121***	0.112***
Friday	0.033***	0.026***
Time period: early AM	0.931***	0.690***
Time period: AM	1.106***	0.867***
Time period: midday	0.600***	0.411***
Time period: PM	0.239***	0.137***
Time period: evening	0.122***	0.089***
Interaction: with service/visitor employment density		
Weekday	−0.094***	−0.082***
Friday	−0.005***	−0.001
Time period: early AM	−0.598***	−0.329***
Time period: AM	−0.467***	−0.299***
Time period: midday	−0.280***	−0.071***
Time period: PM	−0.067***	0.075***
Time period: evening	−0.045***	0.002*

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

The estimates are from a fractional logistic regression model. Points are aggregated to the TAZ level, and the model estimates the proportion of points according to each behavior. Note that the lagged behaviors cannot be captured using this approach.

Table A-5. Regression coefficients (full resolution)

Covariate	Reposition Coefficient	Cruise Coefficient
(Intercept)	7.000***	2.408***
Demographic and neighborhood variables		
Fraction age 62+	−0.027*	−0.019
HH density	−0.003	−0.003
Fraction of working age	0.0003	−0.045**
Employment density	−0.184***	−0.06
Service and visitor employment density	0.025	0.04
Fraction high income HHs	−0.133***	−0.089***
On-street parking capacity	0.047***	0.034*
Off-street parking capacity (public)	−0.007	−0.028*
Off-street parking capacity (residential)	0.008	0.014
Fraction Latinx residents	0.057***	0.043***
Fraction African-American residents	−0.017	−0.032**
Fraction White residents	−0.031**	0.003
Time and day of week variables		
Time period: early AM	−0.288***	−0.089*
Time period: AM	−0.085**	−0.108***
Time period: midday	−0.095***	−0.113***
Time period: PM	0.027	−0.036
Time period: night	0.003	0.101**
Friday	0.048	0.008
Mon-Thurs	−0.077***	−0.095***
Lagged dependent variables		
Lag cruise	−3.518***	4.658***
Lag park	−12.598***	−8.211***

*A positive coefficient for repositioning indicates that the driver is more likely to reposition **away** from a TAZ.*

Covariate	Reposition Coefficient	Cruise Coefficient
Interaction: with HH density		
Weekday	0.032	−0.007
Friday	−0.020	−0.046
Time period: early AM	−0.044	0.007
Time period: AM	−0.128***	−0.066
Time period: midday	0.039	0.041
Time period: PM	0.076**	0.051
Time period: night	−0.094*	−0.054
Interaction: with employment density		
Weekday	0.038	0.032
Friday	−0.054	−0.098
Time period: early AM	0.531***	0.387***
Time period: AM	0.590***	0.576***
Time period: midday	0.322***	0.257***
Time period: PM	0.017	0.016
Time period: night	−0.053	0.054
Interaction: with service/visitor employment density		
Weekday	−0.025	−0.012
Friday	0.098	0.129*
Time period: early AM	−0.368***	−0.263**
Time period: AM	−0.299***	−0.337***
Time period: midday	−0.206***	−0.116**
Time period: PM	−0.029	0.015
Time period: night	0.005	−0.051

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Estimates are from a point-level multinomial logistic regression at one-minute resolution, as in Table 4. However, the full dataset is used, rather than downsampling to one-minute resolution. Downsampling mitigates the problems with serial correlation (see Figure A-1), but means that estimates are less precise (i.e., confidence intervals are wider).