PROFILING TNC DEADHEADING: SEARCHING VS. DISPATCHING

Priyanka Paithankar, Ph.D.
Department of Civil, Architectural, and Environmental Engineering
The University of Texas at Austin, 301 E. Dean Keeton St, Austin, TX 78712
priyanka.paithankar@utexas.edu

Kara M. Kockelman, Ph.D., P.E. (Corresponding Author)
Department of Civil, Architectural, and Environmental Engineering
The University of Texas at Austin, 301 E. Dean Keeton St, Stop C1761, Austin, TX 78712
kkockelm@mail.utexas.edu

Total Pages: 13

STRUCTURED ABSTRACT

Objectives: This study models and distinguishes two components of deadheading: driving while a driver awaits a new rider (P1), and while the vehicle drives to a newly assigned rider (P2). Analyzing these separately is essential to developing interventions to reduce empty vehicle-miles travelled (eVMT) and its negative effects.

Methods: Analysis of 769,806 Uber trips across California from September 2019-August 2020 with demographic and built-environment data. Three models estimated: fractional-response logit for the total empty share of miles driven, a two-part model for searching distance in miles, and a least-squares regression for mean dispatching distance in miles.

Findings: eVMT accounts for 29.6% of vehicle-miles driven, with dispatching (16.4%) exceeding searching (13.2%). Pooled trips run 10.1% empty versus solo trips at 31.4%, reducing search probability by 29 percentage points ($p<0.001$). Employment density governs dispatching distance: one-standard-deviation increase reduces pickup distance by 82% of mean. Strong spatial autocorrelation (Moran's I = 0.839) shows clustering of high and low deadhead zones.

Novelty: This comprehensive study separately measures and explains searching vs. dispatching deadheading determinants using trip data. It shows that these components respond to different factors, requiring distinct policy approaches.

Practical Applications: Searching miles can be reduced through pooling incentives and optimized matching. Dispatching miles are governed primarily by employment density. Targeted interventions in high-deadhead zones, which are identified through spatial analysis, are most effective. Findings inform TNC operations and transportation planning strategies to reduce empty miles and environmental costs.

BACKGROUND

Empty vehicle-miles traveled (eVMT) is a hidden cost of the ride-hailing services operated by transportation network companies (TNCs). The passenger-miles Uber, Lyft, DiDi, Waymo, ApolloGo, AVRide and other fleets serve currently require some driving without passengers: to pick up a new customer, to favorably reposition after a drop-off, and to head somewhere to be cleaned, recharge and/or refuel (Wu and MacKenzie, 2021; Henao and Marshall, 2019). Past studies' estimates of deadheading vary widely, from about 20% to over 40% of total miles depending on the city, time period, and methodology (Kontou et al., 2020; Wu and MacKenzie, 2021). Such empty driving has two distinct components: The first is the searching component (Period 1 or P1), when a driver waits for a dispatch and may cruise, reposition toward expected demand, or stay parked (Ashkrof et al., 2024; Xu et al., 2023). The second is the dispatching component (Period 2 or P2), when the platform has matched the driver to a rider and the vehicle travels to the pickup point. P1 reflects driver strategy, platform incentives, and the spatial pattern of latent demand, while P2 reflects matching-algorithm efficiency (Castillo et al., 2025; Vazifeh et al., 2018; Xu et al., 2023) and the distance between available drivers and riders (Henao and Marshall, 2019; Vazifeh et al., 2018).

Despite such differences, most prior work combines both phases of empty driving or examines just one. Several studies estimated eVMT using trip-level data (Kontou et al., 2020; Henao and Marshall, 2019), while others modeled using simulation environments (Zhu et al., 2021; Zhang et al., 2024). Past research on matching efficiency explored how supply-demand ratios and surge pricing affect wait times and driver usage (Zhu et al., 2021; Ashkrof et al., 2020, 2024; Zhang et al., 2024) but rarely showed how such changes the searching and dispatching miles drivers actually travel. The built environment literature linked ride-hailing trip generation to population density, land use mix, and transit access (Chang et al., 2024; Mucci and Erhardt, 2024). However, these studies did not examine whether these factors differently predict searching versus dispatching miles. Similarly, COVID-19 produced large shifts in TNC demand (Du and Rakha, 2020; Chang and Miranda-Moreno, 2022; Brown and Williams, 2023), but the effects on P1 versus P2 composition remained unexplored.

This study addresses these gaps by separating searching, dispatching, and passenger-carrying distances and modeling the determinants of each. The analysis uses a 0.5% random sample of about 157 million Uber trips spanning September 2019 through August 2020, a period that captures both normal operations and the onset of COVID-19. The remainder of this paper is organized as follows. The next section describes the data and key variables, followed by a descriptive analysis of how deadheading varies with demand-supply balance, time of day, COVID-19, and geography. The methodology section presents three models: a fractional response model for total eVMT, a two-part model for searching distance (P1), and a zip-level regression for dispatching distance (P2). The results section reports their estimates, and the paper concludes with key findings, policy implications, limitations, and directions for future research.

DATA

This study uses 769,806 trip records drawn as a 0.5% random sample from about 157 million Uber trips in California between September 2019 and August 2020. Each trip record contains distances for the three trip components (P1 searching, P2 dispatching, and P3 passenger-carrying), along with match time, rider wait time, fare, fare surge indicator, service type (UberX, UberPool request, UberPool matched), and timestamp. P1 is measured from the moment the driver activates the app, or drops off a previous passenger, until a new request is accepted. It therefore captures every mile driven with the app on but no assignment, including repositioning toward expected demand and also personal or vehicle-service driving such as meal, restroom, refueling, and maintenance trips. The data do not distinguish these motives, so P1 should be read as all non-assigned driving during an active shift rather than as search behavior alone.

UberPool trips drive 17.5% empty on average (P1 = 0.88 mi, P2 = 0.62 mi), compared to 31.4% for UberX trips (P1 = 1.65 mi, P2 = 1.29 mi). This 14-percentage-point gap reflects two advantages of pooling: shared rides generate longer passenger legs, and pooled matching places drivers closer to their next pickup. Average P1 distance is 1.56 miles, but the median is just 0.16 miles (Table 1). The average eVMT per-trip is 29.6% of which 13.2% is search share and 16.4% is a dispatch share. At the aggregate fleet level, the eVMT share is 28.4%, splitting into 16.0% searching and 12.4% dispatching. The two values differ because short trips carry high per-trip eVMT ratios but few trip miles, while in long trips, the aggregate gives more weight to long trips. These shares describe app-on driving only. Miles driven before a driver activates the app, after deactivating it, or while off-app for refueling or cleaning are not recorded, so the figures here are a lower bound on the full vehicle-miles cost of a driving shift. The rider match time (the duration between a rider request and a driver assignment), is very short (11 and 44 seconds on average), so most of the 6.7-minute rider wait is taken up by the driver actually traveling to the pickup point rather than by the matching algorithm itself. The matching is generally tight: 90% of pickup distances stay within 2.81 miles, and only 7.0% of trips have a match time above 2 minutes. However, 3.0% of trips have P2 distances above 5 miles.

TABLE 1. Summary Statistics of Trip-Level Variables (n = 769,806 Uber Trips, California, September 2019 to August 2020).

Variable	Mean	Median	Std. Dev.	Min	Max
Searching Miles (P1)	1.556	0.160	3.497	0.000	30.730
Dispatching Miles (P2)	1.208	0.750	1.392	0.000	9.960
Passenger Miles (P3)	6.959	4.370	7.188	0.000	48.250
Empty Miles (P1 + P2)	2.764	1.400	3.938	0.000	39.120
Total Miles (P1 + P2 + P3)	9.722	6.950	8.553	0.010	80.660
Total eVMT (%)	29.6	24.2	23.6	0.000	100.0
Searching eVMT (%)	13.2	2.7	20.0	0.000	100.0
Dispatching eVMT (%)	16.4	11.6	15.7	0.000	100.0
Match Time (min)	0.726	0.183	1.884	0.000	56.933
Rider Wait Time (min)	6.700	5.583	4.583	0.167	102.90
Trip Duration (min)	16.10	12.80	11.91	0.020	409.95
Total Fare ($)	15.68	11.80	12.29	0.01	310.63
Tip ($)	0.59	0.00	1.70	0.00	100.00
Fare per Mile ($/mi)	3.54	2.64	3.00	0.00	99.83
Surge Pricing (1 = yes)	0.137	0.000	0.344	0.000	1.000
UberPool Request (1 = yes)	0.128	0.000	0.334	0.000	1.000
UberPool Matched (1 = yes)	0.082	0.000	0.274	0.000	1.000
UberX Trip (1 = yes)	0.872	1.000	0.334	0.000	1.000
Vehicle Occupancy (riders)	1.087	1.000	0.282	1.000	2.000
Weekend Trip (1 = yes)	0.461	0.000	0.498	0.000	1.000
Post-COVID Trip (1 = yes)	0.155	0.000	0.362	0.000	1.000

Each trip was linked to its pickup Zip Code Tabulation Area (ZCTA), and three external sources provide zip-level characteristics. The 2020 American Community Survey (ACS) 5-year estimates supply demographic and socioeconomic variables, including population, income, age, race, renter share, poverty rate, vehicle ownership, and commute mode shares. The U.S. Census TIGER/Line files provide zip-code area, road and intersection density. Population and job density are calculated by dividing ACS counts by zip-code area. The final sample retains 939 California zip codes with at least 50 trips; road and intersection density are available for only 114 of these zip codes and are therefore used only in a robustness check rather than the main model. Table 2 presents zip-level summary statistics. On average, zip code has a median household income of $91,208, ranging from $17,270 to $250,001. The non-white population's share averages to 43.3%, median age is 38.2 years, and the poverty rate averages 12.0%. Renter households make up 45.3% of the typical zip code, and zero-vehicle households average just 7.5%. The average zip code covers 24.1 square miles, ranging from 0.02 sq. mi. in dense urban areas to 432 sq. mi. in rural ones. Population density averages 5,986 persons per square mile, peaking at 55,723 in downtown San Francisco, while job density peaks at 30,974 in central business districts.

TABLE 2. Summary Statistics of Zip-Level Demographic and Built-Environment Variables (n = 939 California ZCTAs).

Variable	Mean	Median	Std. Dev.	Min	Max
Total Population (persons)	37,456	35,145	20,634	0	110,750
Median Household Income ($)	91,208	84,891	38,574	17,270	250,001
Median Age (years)	38.2	37.5	7.6	19.2	75.4
Share Non-White (%)	43.3	41.4	22.6	0.0	90.9
Share Renter Households (%)	45.3	43.3	20.0	0.0	100.0
Poverty Rate (%)	12.0	10.0	8.4	0.0	100.0
Zero-Vehicle Households (%)	7.5	5.2	8.5	0.0	100.0
Drive-Alone Commute Share (%)	70.5	74.5	15.6	0.0	100.0
Transit Commute Share (%)	5.0	2.4	8.1	0.0	66.0
ZCTA Area (sq. mi.)	24.1	7.9	49.0	0.02	432.0
Population Density (persons/sq. mi.)	5,986	4,351	6,712	0	55,723
Job Density (jobs/sq. mi.)	2,970	2,080	3,684	0	30,974
Road Density (mi/sq. mi.)	14.6	14.5	6.5	0.0	38.3
Intersection Density (per sq. mi.)	217	194	149	0	812
Total Road Length (mi)	165.0	135.6	108.7	0.1	612.3

Demand-Supply Balance

Table 3 presents how eVMT varies with the demand-supply (D/S) ratio, defined as the number of rider requests relative to active drivers in a given zip code and hour. The system operates close to balance on average (mean D/S = 1.06, median = 1.00). Empty driving is lowest under mild undersupply (D/S = 1.25 to 1.5, eVMT = 24.7%), where nearly every driver has a queued request waiting. Under slight oversupply (D/S = 0.7 to 0.9), eVMT rises to 25.8% as idle drivers pile up searching miles (P1 = 1.05 mi, P2 = 0.67 mi, rider wait = 4.2 min). Under severe undersupply (D/S > 2), the platform must dispatch increasingly distant drivers to meet excess demand, pushing P1 to 3.71 miles and rider wait to 8.2 minutes.

TABLE 3. Deadhead Components across Demand/Supply (D/S) Ratio Bins.

D/S Bin	N	P1 (mi)	P2 (mi)	Total eVMT	Search eVMT	Dispatch eVMT	Pax-Wait (min)
0.5 - 0.7	264	2.36	0.78	30.0%	18.7%	11.3%	4.9
0.7 - 0.9	4,688	1.05	0.67	25.8%	13.2%	12.7%	4.2
0.9 - 1.0	181,964	1.90	1.87	34.9%	13.9%	21.0%	6.6
1.0 - 1.25	542,679	1.43	1.00	28.2%	13.0%	15.2%	5.3

1.25 - 1.5	26,630	1.17	0.73	24.7%	11.4%	13.2%	5.3
1.5 - 2.0	9,981	2.31	1.32	25.6%	12.5%	13.0%	6.9
> 2.0	3,598	3.71	2.32	25.9%	14.1%	11.9%	8.2

eVMT varies considerably across the day (see Figure 1). Total eVMT ranges from 26.5% at 5 PM to 35.6% at 3 AM. Searching (P1) peaks during midday when demand is moderate and drivers may cruise between dispersed trip requests, while dispatching (P2) peaks during late-night hours when demand intensifies in CBDs but available drivers are sparse. During the evening peak, 44.3% of trips have P1 approximately equal to zero, indicating that high demand enables continuous dispatching. Moreover, the COVID-19 onset in March 2020 produced a large shift in deadheading. Pre-COVID (September 2019 through February 2020), the eVMT averaged 28.4% (P1 = 13.3%, P2 = 15.1%) across about 3,596 trips per day (see Figure 2). During the transition month of March 2020, eVMT spiked to 41.0%, driven almost entirely by a surge in searching miles (P1 = 26.3%) as demand reduced but drivers stayed on the platform. During lockdown (April through May), eVMT reached 38.6%, but the composition shifted: P1 returned to 15.1% while P2 rose to 23.5%. Figure 3 shows visible clusters of low eVMT in urban cores (San Francisco, Los Angeles, San Diego) surrounded by higher eVMT in suburban and rural ZCTAs. Zip-level eVMT shows strong positive spatial autocorrelation. Moran's I, computed using a queen-contiguity spatial weights matrix on the 939 California ZCTAs, is 0.839, indicating that zip codes with high eVMT tend to be located near other high-eVMT zip codes, and vice versa.

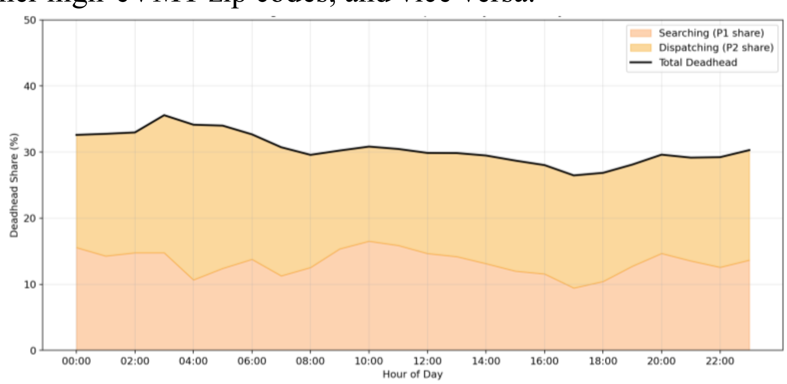

FIGURE 1. Deadhead Ratio by Hour of Day, Decomposed into Searching and Dispatching Components.

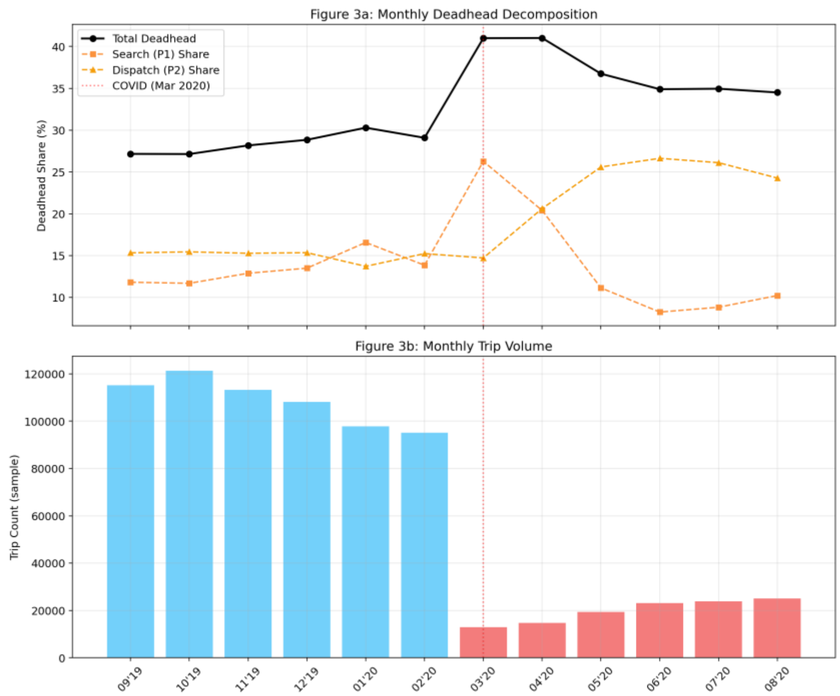

FIGURE 2. Daily Deadhead Ratio around the COVID-19 Onset (March 15, 2020).

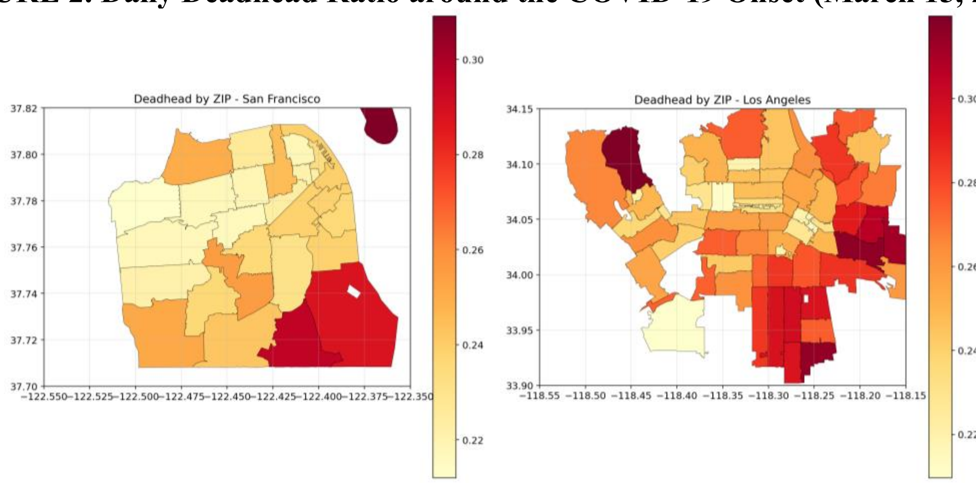

FIGURE 3. Zip-Level Map of Total Deadhead Ratio across California.

METHODOLOGY

This study employs three models, one for each component of empty travel: (i) a fractional response model for the total eVMT share, (ii) a two-part model for the zero-inflated searching distance (P1), and (iii) a least-squares regression for the mean dispatching distance (P2). The first captures the overall determinants of empty driving; the second separates whether a driver searches at all from how far they search when they do; and the third uses variations across pick-up zip codes to explain how far drivers must travel to reach an assigned pickup. Since eVMT showed strong spatial autocorrelation across zip codes, all trip-level standard errors are clustered by pickup zip code.

Fractional response model for eVMT (P1+P2)

The eVMT share takes values between 0 and 1, ruling out Ordinary least square regression (OLS), which can predict outside the unit interval. Following Papke and Wooldridge (1996), a fractional response model is estimated as a generalized linear model with a logit link and a binomial variance function, fit by Bernoulli quasi-maximum likelihood.

$$E[DH_i|X_i] = G(\alpha+X_i'\beta), \quad G(z) = \exp(z)/[1+\exp(z)] \quad (1)$$

where $DH_i = (P1+P2)/(P1+P2+P3)$ is the trip-level eVMT share, $G(\cdot)$ is the logistic CDF, and X_i is the covariate vector. X includes log revenue trip miles (P3); service-type indicators (UberPool requested, UberPool matched); a surge indicator; time-of-day (AM-peak 7-9, PM-peak 16-18, night 22-4) and weekend indicators; COVID-19 period indicators (transition: March 2020; lockdown: April-May 2020; recovery: June-August 2020, with pre-COVID as the base); and pickup-zip characteristics merged onto each trip (log population density, log job density, log median household income; shares of non-white, renter, below-poverty, and zero-vehicle households; transit and drive-alone commute shares; and median age). The parameters (α, β) are estimated by Bernoulli quasi-maximum likelihood (QMLE), though the true distribution of DH is not Bernoulli, Papke and Wooldridge (1996) showed that this QMLE is consistent and asymptotically normal for any fractional outcome.

Two-part model for searching distance (P1)

P1 cannot be modeled like total eVMT for two reasons: over a third (35.5%) of trips have $P1 \approx 0$ (the driver was dispatched almost immediately), and P1 is a non-negative continuous distance rather than a bounded share. A two-part model accommodates these two features. The first part is a binary-choice model for whether any searching occurs. Define $S_i = 1\{P1_i > 0\}$; S_i follows Bernoulli(π_i) with

$$\pi_i = \Pr(S_i = 1 \mid X_i) = \Lambda(\alpha_1+X_i'\beta_1) \quad (2)$$

where $\Lambda(\cdot)$ is the logistic CDF and X_i is the same covariate vector as in (1). The parameters are estimated by logit maximum likelihood on all 769,806 trips. The second part models how far the driver searches when searching does occur. Conditional on $P1 > 0$, log P1 is modeled as a linear function of the same covariates:

$$\log P1_i = \alpha_2 + X_i'\beta_2 + \varepsilon_i, \quad E[\varepsilon_i \mid X_i, P1_i > 0] = 0 \quad (3)$$

where ε_i is the regression error and X_i is the same covariate vector as in (1). The parameters are estimated by ordinary least squares on the 496,810 trips where $P1 > 0$.

Zip-level OLS for dispatching distance (P2)

P2 varies primarily across zip codes rather than individual trips, driven by differences in activity density and street connectivity, so it is modeled at the zip-code level across 927 California ZCTAs:

$$P2_z = \alpha + X_z'\beta + \varepsilon_z \quad (4)$$

where $P2_z$ is the mean dispatching distance (miles) in zip code z, ε_z the error, and X_z the vector of standardized zip-level predictors: log population and job density, drive-alone and transit commute shares, renter share, zero-vehicle share, share non-white, median age, and zip-code area (densities log-transformed before standardizations). The model is fit by OLS on the 927 ZCTAs with HC1 robust SEs and backward elimination at $p<0.10$ (Hocking 1976); each coefficient is the change in P2 (miles) per 1-SD increase.

RESULTS

Total eVMT (P1+P2)

The fractional response model explains the share of trip-level variation in eVMT (Table 4). A matched pool trip lowers eVMT by 18.3 percentage points (pp), while just requesting a pool lowers it by 2.4 pp. A 1% rise in revenue-miles cuts eVMT by 0.18 pp, reflecting fixed dispatching costs spread over longer revenue legs. Time-of-day swings the share by approximately ±3 points (higher at night, lower during the PM peak), and the COVID transition, lockdown, and recovery raised eVMT by 8.2, 6.5, and 5.5 points relative to pre-COVID. The denser job markets and renter-heavy neighborhoods reduce eVMT, while car dominated zip codes with larger non-white population shares show more empty driving.

TABLE 4. Fractional Response (GLM Logit) Model for Total eVMT (N = 755,771 trips; Pseudo R^2 = 0.11)

Variable	Coef.	Marg. Effect (pp)
Intercept	2.830***	
Log (revenue miles, P3)	-0.974***	-18.1
UberPool request	-0.127***	-2.4
UberPool matched	-0.988***	-18.3
Surge pricing indicator	-0.126***	-2.3
Weekend indicator	-0.064***	-1.2
AM peak (7–9)	0.043***	+0.8
PM peak (16–18)	-0.157***	-2.9
Night (22–4)	0.154***	+2.9
COVID transition (Mar 2020)	0.440***	+8.2
COVID lockdown (Apr–May)	0.349***	+6.5
COVID recovery (Jun–Aug)	0.298***	+5.5
Log (population density)	0.182	+3.4
Log (job density)	-0.305**	-5.7
Log (median HH income)	-0.068	-1.3
Share non-white	0.154**	+2.9
Share renter	-0.753***	-14.0
Poverty rate	-0.258	-4.8
Zero-vehicle share	0.126	+2.3
Transit commute share	0.016	+0.3
Drive-alone share	0.707***	+13.1
Median age	-0.011***	-0.2

Note: Significance: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Searching distance (P1)

The two-part model separates searching into whether any searching occurs (Part 1) and how far the driver searches when it does (Part 2). Part 1 results (Table 5) show that a successful pool match cuts the probability of any searching by 29 pp, fare surge by 13 pp, and a weekend trip by 4 pp. Pooled trips almost always skip the cruising phase because matched trips are dispatched directly from a queue. Longer passenger trips marginally raise the probability of any searching, likely because longer trips end in lower-density areas where the next dispatch takes longer. Part 2, conditional on searching, shows that surge shortens conditional searching distance by about 27%, weekend trips by 10%, and a pool request by about 3%. The COVID transition month raised both the probability (+21 pp) and the conditional distance (+104%). Lockdown and recovery however reduced the probability of searching by 6 and 17 pp respectively, and yet conditional distances remained above pre-COVID levels at 38% and 12% higher.

TABLE 5. Two-Part Model for Searching Distance (P1)

	Y = Pr(P1>0), N = 769,806, pseudo R^2 = 0.057		Y= log(P1) given P1>0, N = 496,810, R^2 = 0.019.	
Variable	**Coef.**	**Marg. Eff. (ΔPr, pp)**	**Coef.**	**Marg. Eff. (%Δ P1)**
Intercept	0.898***	—	-0.533***	—
Log(P3) trip distance	0.090***	+1.9	0.175***	+19.1%
UberPool request	-1.372***	-29.1	-0.033*	-3.2%
Surge pricing indicator	-0.612***	-13.0	-0.311***	-26.7%
Weekend indicator	-0.189***	-4.0	-0.105***	-10.0%
Log(population density)	-0.033	-0.7	-0.284	-24.7%
COVID transition (March 2020)	1.012***	+21.4	0.711***	+103.6%
COVID lockdown (April to May)	-0.276***	-5.9	0.320***	+37.7%
COVID recovery (June to August)	-0.785***	-16.6	0.115***	+12.2%

Note: Significance: * p<0.05, ** p<0.01, *** p<0.001.

Dispatching distance (P2)

The least-square regression explains 68% of the variation in average dispatching distance across zip codes (Table 6), where all predictors are standardized. A one-standard-deviation increase in job density reduces dispatching distance by 1.6 miles, or 82% of the sample mean of 1.96 miles, as concentrated employment keeps riders and drivers in close proximity. As residential population density and job density are strongly correlated, job density absorbs the activity-density effect and residential density picks up the residual. Among sociodemographic characteristics, zip codes with higher drive-alone, transit, and zero-vehicle shares have longer dispatching distances, while renter-heavy, older, and higher non-white share zip codes have shorter ones.

TABLE 6. Least-Squares Regression for Dispatching Distance (N = 927 ZCTAs, R^2 = 0.68).

Variable	**Coef. (Δ P2 per 1 SD, mi)**	**Std. β (effect size)**	**% of mean P2**
Constant (mean P2)	1.962***	—	—
Log (population density)	1.138***	1.272	+58.0%
Log (job density)	-1.607***	-1.797	-81.9%
Median age	-0.258***	-0.289	-13.1%
Share non-white	-0.143***	-0.160	-7.3%
Share renter	-0.456***	-0.510	-23.2%
Zero-vehicle share	0.180***	0.201	+9.2%
Transit commute share	0.100**	0.112	+5.1%
Drive-alone share	0.284***	0.318	+14.5%
Area (sq. mi)	-0.135***	-0.151	-6.9%

Note: Significance: * p<0.05, ** p<0.01, *** p<0.001.

CONCLUSION

This study finds that empty driving in ride-hailing stems from two distinct sources that existing research has largely overlooked. This study separates TNC deadheading into its searching and dispatching components using 769,806 Uber trips from a 0.5% sample of California's 157 million Uber trips between September 2019 and August 2020. This study shows that Uber trips drive 29.6% of vehicle-miles empty on average, with 13.2% from searching and 16.4% from dispatching. Searching miles respond primarily to platform matching and pricing decisions, while dispatching miles are influenced by land use patterns and street network structure. Dispatching, in contrast, is influenced largely by the built environment. On average, 29.6% of vehicle miles are driven empty, with dense urban areas generating far less empty driving than suburban and rural areas. Future research should extend this work in several directions. First, the sample covers 0.5% of Uber trips in a single state, and while the large sample size produces precise estimates, the results may not apply to other platforms such as Lyft, which uses a different matching algorithm, or to areas outside California. Access to real-time matching algorithm data can further clarify how platforms generate dispatching miles and how platform strategic changes might reduce them. Future research can quantify the distinct environmental costs of searching and dispatching, as the low-speed cruising of searching and the directed travel of dispatching produce different emission rates per mile.

ACKNOWLEDGMENTS

The authors used Claude, a generative artificial intelligence (AI) tool made by Anthropic, while preparing this paper. It helped edit the writing, reorganize the methodology and results sections and write the Python scripts that estimated the models. All modeling decisions, specifications, and interpretations are the authors' own. All reported estimates were produced by the statsmodels library in Python, applied to the trip data described in the Data section; no data, results, or references were generated by an artificial intelligence tool. The authors verified all cited sources against publisher records and take responsibility for the accuracy and integrity of the work.

REFERENCES

Ashkrof, P., Correia, G. H. de A., Cats, O., & van Arem, B. (2020). Understanding ride-sourcing drivers' behaviour and preferences: Insights from focus groups analysis. Research in Transportation Business & Management, 37, 100516. https://doi.org/10.1016/j.rtbm.2020.100516

Ashkrof, P., Correia, G. H. de A., Cats, O., & van Arem, B. (2024). On the relocation behavior of ride-sourcing drivers. Transportation Letters, 16(4), 330–337. https://doi.org/10.1080/19427867.2023.2192581

Brown, A., & Williams, R. (2023). Equity implications of ride-hail travel during COVID-19 in California. Transportation Research Record, 2677(4), 1–14. https://doi.org/10.1177/03611981211037246

Castillo, J. C., Knoepfle, D., & Weyl, E. G. (2025). Matching and pricing in ride hailing: Wild goose chases and how to solve them. Management Science, 71(5), 4377–4395. https://doi.org/10.1287/mnsc.2022.00096

Chang, A., & Miranda-Moreno, L. (2022). Ride-hailing through the COVID-19 pandemic in New York City. Findings, March. https://doi.org/10.32866/001c.33160

Chang, A. Y. J., Wang, X., Sharafi, M., Miranda-Moreno, L., & Sun, L. (2024). Headwind or tailwind? The evolution of bike-sharing and ride-hailing demand during the COVID-19 pandemic. Journal of Transport Geography, 118, 103944. https://doi.org/10.1016/j.jtrangeo.2024.103944

Du, J., & Rakha, H. A. (2020). COVID-19 impact on ride-hailing: The Chicago case study. Findings, October. https://doi.org/10.32866/001c.17838

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

Hocking, R. R. (1976). The analysis and selection of variables in linear regression. Biometrics, 32(1), 1–49. https://doi.org/10.2307/2529336

Kontou, E., Garikapati, V., & Hou, Y. (2020). Reducing ridesourcing empty vehicle travel with future travel demand prediction. Transportation Research Part C: Emerging Technologies, 121, 102826. https://doi.org/10.1016/j.trc.2020.102826

Mucci, R. A., & Erhardt, G. D. (2024). Understanding ride-hailing sharing and matching in Chicago using travel time, cost, and choice models. Transportation Research Record, 2678(2), 293–306. https://doi.org/10.1177/03611981231173636

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics, 11(6), 619–632. https://doi.org/10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1

Vazifeh, M. M., Santi, P., Resta, G., Strogatz, S. H., & Ratti, C. (2018). Addressing the minimum fleet problem in on-demand urban mobility. Nature, 557(7706), 534–538. https://doi.org/10.1038/s41586-018-0095-1https://doi.org/10.1016/j.trb.2021.06.014

Wu, X., & MacKenzie, D. (2021). Assessing the VMT effect of ridesourcing services in the US. Transportation Research Part D: Transport and Environment, 94, 102816. https://doi.org/10.1016/j.trd.2021.102816

Xu, M., Yue, P., Yu, F., Yang, C., Zhang, M., Li, S., & Li, H. (2023). Multi-agent reinforcement learning to unify order-matching and vehicle-repositioning in ride-hailing services. International Journal of Geographical Information Science, 37(2), 380–402. https://doi.org/10.1080/13658816.2022.2119477

Zhang, J., Hu, L., Li, Y., Xu, W., & Jiang, Y. (2024). Dynamic joint decision of matching parameters and relocation strategies in ride-sourcing systems interacting with traffic congestion. Transportation Research Part C: Emerging Technologies, 161, 104524. https://doi.org/10.1016/j.trc.2024.104524

Zhu, Z., Ke, J., & Wang, H. (2021). A mean-field Markov decision process model for spatial-temporal subsidies in ride-sourcing markets. Transportation Research Part B: Methodological, 150, 540–565. https://doi.org/10.1016/j.trb.2021.06.014