Institute of Transportation Studies
UNIVERSITY OF CALIFORNIA

Sharing Mobility Data for Planning and Policy Research

Juan Matute, Mollie Cohen D'Agostino, Austin Brown
Reviewers/Contributors: Susan Shaheen, Elliot Martin, Dillon Fitch, Jacob Wasserman, Bo Liu, Marcel Moran, Jane MacFarlane

February 2020

Introduction

A California Public Utilities Commission (CPUC) rulemaking and possible legislative action in 2020 could affect data sharing requirements, with implications for shared mobility providers. The purpose of this brief is to inform this regulatory and legislative decision-making.

We solicited policy and planning questions and data needs for shared mobility from within the University of California Institute of Transportation Studies research network. We defined shared mobility as including shared mobility devices, such as e-bikes and e-scooters, and transportation network companies (TNCs). We evaluated whether data shared in accordance with each of six mobility data specifications could be used to support analyses that would answer these questions. We then defined three approaches to data sharing and analysis to address these and other questions, presenting the advantages and disadvantages of each.

This brief does not address the full breadth of the questions raised in the CPUC rulemaking nor does it introduce the complexities of this topic.[1] Beyond the scope of this brief are issues of user privacy, the legal authority for sharing data, and contractual or requirements for each possible model of data sharing and analysis.

Definitions

Aggregate data: summary statistics, information based on analysis of data, less likely to be reidentified. The June 19, 2019 version of AB 1112 defines aggregated data as " data that relates to a group of trips, from which the start points, stop points, routes, and times of individual trips have been removed and that cannot be used, or combined with other information to isolate details of an individual trip."

Disaggregated data: unprocessed or minimally-processed "big" data; may contain personally-identifiable information or can be used in conjunction with other publicly-identifiable information to identify individuals even if personally-identifiable information is not released with the data. Disaggregated data includes "individual trip data" from the June 19, 2019 version of AB 1112.

Deidentified data: data processed to mitigate the potential disclosure of personally-identifiable information. The June 19, 2019 version of AB 1112 defines deidentified data as data "that cannot reasonably identify, relate to, describe, be capable of being associated with, or be linked, directly or indirectly, to a particular user, provided that an entity that uses deidentified data meets all of the following criteria: (1) Has implemented technical safeguards that prohibit reidentification of the user to whom the data may pertain. (2) Has implemented business and security processes that specifically prohibit reidentification of the data. (3) Has implemented business and security processes to prevent inadvertent release of deidentified. data. (4) Makes no attempt to reidentify the information.

Background

City, regional, and state mobility planning and regulatory authorities have much to gain from access to data from shared mobility providers.[2] But to date, significant challenges in collecting, distributing, and analyzing the data have kept these insights "locked up." Some examples of the ways (and the areas) in which shared mobility data would serve the public interest include:

- **Planning:** by enabling a better understanding of regional travel demand.
- **Policy Research:** by improving assessments of pooling incentive programs and the distribution, availability, and use of shared mobility in diverse neighborhoods.
- **Operations:** by facilitating better management of traffic events and emergencies.
- **Monitoring and enforcement:** by making compliance enforcement easier.

This brief focuses on the applications of shared mobility data in planning and policy research. Although these data could be used in a variety of other ways, i.e., for enforcement and operations, such uses are beyond our scope.

Specific data needs will differ by application and geographic scale of interest. For example, state planning entities and regional planners can likely conduct most long-range planning activities with annual and aggregated data. However, city and state-level regulatory authorities could benefit from more granular route and path data for planning and policy to respond to emerging trends and challenges.

City planners could also use mobility data to build "digital twins", or complex digital replicas, of their cities that allow them to model planning scenarios. For example, a city hosting a large sporting event could employ a digital twin model to consider different curbside pickup points and street lane closures that would improve traffic around the event.[3]

Planning and Policy Research Applications

The table below (Table 1) outlines potential policy and planning analyses, and lists the specific data needed for each analysis. Table 2 lists six existing and proposed data specifications, and Table 3 maps whether each of the data specifications could be used to support the analysis in Table 1.

Table 1. Data requirements for public interest planning and policy research applications of mobility data

#	Application	Data Requirements
Analyzing changes in bicycle usage at new and enhanced facilities		
1	Understand the effects of new bike facilities on usage and attraction from other routes. It can also be used for ex-post evaluation of Active Transportation Program investments and for use in modeling efforts to inform decisions about targeting future investments. Finally, it can be used to identify areas of high micromobility volumes that have substandard bikeway facilities.	• Disaggregated data on trip route polyline with timestamps for trip begin and end • Alternatively: aggregation of trip route to street segments/directions and hour of day/date. However, this can obscure attraction of users from parallel routes (that may be higher stress) • Provider-reported data should be supplemented by a screenline count that measures the proportion of data-reporting vehicles as percentage of total vehicles.
Improving curb space management		
2	Understand shared mobility service demand for curb space to inform the reallocation of curb zones. These data can be collected from all types of curb-using vehicles and devices including shared scooters and bicycles, transportation network companies, and courier network service companies (e.g. Doordash, Grubhub).	• Disaggregated timestamped data for when vehicle or device arrives at curb or leaves curb for all events including trips and non-trips • Also possible with aggregation of trip start and end locations to blockface and hour of day/day of week
Understanding legislation's effects on shared mobility workers		
3.1	Assess the effects of AB 5 and other enacted legislation on TNC drivers' hours of work, number and type of trips provided.	• Unique identifier information for the TNC driver accompanying timestamped data for period 1 logon and logoff and period 3 (trips). See footnote for CPUC-defined TNC trip periods.
3.2	Assess the effects of AB 5 and other enacted or proposed legislation on TNC driver's employment status.	• Data on W-2 employee status
3.3	Assess the earnings and expenses of TNC drivers.	• Unique identifier information for the TNC driver accompanying timestamped data for period 1 logon and logoff and period 3 (trips) • Driver pay information, including tips

Table 1. Data requirements for public interest planning and policy research applications of mobility data (Continued)

#	Application	Data Requirements
		• Vehicle make and model information • Total vehicle miles traveled (VMT) and vehicle hours traveled (VHT) • Ideally, the analysis would require information on vehicle lease payments (if leased directly from TNC or TNC partner)
Assessing equity and availability of shared mobility across neighborhoods		
4.1	Compare average wait times, proportion of fulfilled requests, total requests, and device/service availability by neighborhood.	• Request timestamp and geolocation • Request status, including wheelchair-accessible vehicle (WAV) request and status • Trip start geolocation • Vehicle/device available location (tract or more precise) and time (hour or more precise)
4.2	Assess differences in the costs of trips requested in certain neighborhoods.	• Start geolocation or census tract for a trip • Total cost or fare for a trip
Evaluating policy choices		
5.1	Model routes and pricing for dynamic roadway pricing studies considering cordon pricing, high-occupancy toll lane deployment and expansion, and toll adjustments.	• Disaggregated, timestamped, geolocated data on vehicle route polylines (breadcrumbs), with timestamp for all ride trip periods (TNC period 3) and non-trip periods (periods 0 through 2) • Cost of trip:fare paid or cost to user • Any congestion or tolling surcharge assessed • Can be conducted less accurately with trip start and end census tracts
5.2	Create vehicle and device environmental and efficiency metrics including passengers per vehicle, vehicle miles traveled, and the proportion of passenger-serving vehicle or device miles traveled.	• Shared rides requested and fulfilled • VMT for passenger-serving trip(s) • VMT for other activity or support vehicles for mobility devices
5.3	Assess vehicle electrification and estimate emissions (e.g. for SB 1014 compliance).	• Vehicle powertrain type in addition to environmental and efficiency metrics in 5.2 • Alternatively:vehicle make and model, or VIN or license plate data to look up make and model. • Vehicle speed from telemetry data is not not required but would lead to more accurate estimates of emissions

Table 1. Data requirements for public interest planning and policy research applications of mobility data (Continued)

#	Application	Data Requirements
Assessing regulatory compliance		
6	Assess provider compliance with permit conditions or regulations. Data can be used to assess compliance on several dimensions, including compliance rates specific to a particular provider or neighborhood.	Disaggregated timestamped, geolocated data with event information on all regulated activities, such as: • wheelchair-accessible vehicles (WAVs) trip requests, pickups, and dropoffs (SB 1376) • Vehicles or devices entering geofenced restricted or fee-assessed areas, such as volume-capped zones, safety exclusion areas, or airports.
Evaluating safety		
7.1	Evaluate shared mobility driver safety records, perhaps in comparison with all drivers	• Vehicle code violations, as reported by law enforcement to the DMV with personally-identifiable information.
7.2	Identify locations where vehicles or mobility devices travel at speeds exceeding the posted speed limit for a road or bikeway path, where unsafe speeds cause dangerous conditions.	• Telemetry data on vehicle location and speed, aggregated to data on observed speeds on specific links of the roadway networks but without trip-end, vehicle, or operator identifiers.
7.3	Identify specific individuals in TNC services that drive in excess of posted speed limits. This is not a legal enforcement option in California.	• Telemetry data with vehicle geolocation and speed • VINs of vehicles traveling in excess of posted speed limits for a roadway segment. May also include driver's identity.
Assessing transit ridership trends		
8.1	Assess the substitution of new mobility for transit (when new mobility trips compete with transit routes) and complementarity of new mobility with transit (when new mobility trips are used to go to or from transit stations). Assess new mobility's effects on aggregate transit ridership trends and specifically whether new mobility is a cause of the recent transit ridership decline. If so, to what degree and in what areas?	In addition to timestamped data for origin and destination presently available to public transportation operators: • TNC riders reported connections to transit (via surveys) • Disaggregated, timestamped data on trip start and end geolocations for trips which begin or end within a specified distance of a public transit stop or bus zone
Tracking transportation system performance		
9.1	Assess impacts of TNC vehicles on traffic congestion and total vehicle volumes in an area.	• Disaggregated, timestamped, geolocated data on vehicle route polylines (breadcrumbs), with begin/end timestamp for all TNC trip periods

Table 1. Data requirements for public interest planning and policy research applications of mobility data (Continued)

#	Application	Data Requirements
		• Vehicle speed from telemetry data can improve accuracy of speed estimates • Requires additional speed and position data from non-TNC vehicles
9.2	Compare the speeds, prices, reliability, etc. of various new mobility services, compared to other modes like transit and solo driving	• Disaggregated, timestamped, geolocated data on vehicle route polylines, with timestamp (breadcrumbs) for all ride trip period or tract-level information on origin, destination and VMT and duration/VHT • Trip fare or cost information
Identifying ZEV TNC charging needs		
10.1	Estimating charging infrastructure needs for long-range TNC vehicles by looking at driver charging behavior	• Periodic data on vehicle state of charge, with timestamp • VMT between charging events • Route polyline data prior and after to charging event (allow inference of route deviations for charging. • Frequent data on state-of charge during charging events, with timestamp • State-of-charge event timestamp
10.2	Analyzing temporal and spatial trends of electricity demand from the shared mobility industry and its potential impacts on the power grids	• Periodic data on vehicle state of charge during charge events • Latitude and longitude at onset of charge event • Charge event begin and end timestamp
10.3	Exploring the impact of “range anxiety” on the provision of services by analyzing changes in state of charge and trip making.	• Route breadcrumbs with state of charge: timestamped latitudes and longitudes with state-of-charge level to form a polyline • Charge event begin and end timestamp • Charge event state-of-charge

Timestamp means datetime stamp

Mobility Data Sharing Specifications

Table 2 (below) outlines six mobility data sharing specifications. All but two are currently in use. While the specifications of the Electronic Frontier Foundation (EFF) and the City and County of San Francisco (SF) are not yet in use, they have been proposed to the CPUC Rulemaking R12-12-011, concerning Transportation Network Companies Comments.

The CPUC is currently considering data collection and sharing from Transportation Network Companies (TNCs).[5] The CPUC has sought comment on data collection protocols, reporting requirements (e.g., whether data will be collected monthly, quarterly or annually), the level of data granularity that will be required (e.g., whether data will be collected at ZIP code, census tract, block level, etc.), what parties will be able to access the data collected (e.g., cities, transit agencies, researchers), who will be responsible for analyzing the data (e.g., CPUC, cities, researchers, etc).

The remaining specifications in Table 3 are used in practice. Three are TNC data specifications currently utilized by local regulatory agencies in North America: City of Chicago, Toronto, and New York City Taxi and Limousine Commission. And the sixth, the Open Mobility Foundation's Mobility Data Specification (OMF MDS) is currently used only sharing of data by shared mobility device providers, but the specification can be adapted to transportation network companies. The Open Mobility Foundation's Mobility Data Specification was originally developed by the City of Los Angeles.

Table 2: Six Mobility Data Specifications

		Electronic Frontier Foundation (as proposed to CPUC)	City and County of San Francisco (as proposed to CPUC)	Open Mobility Foundation/ Los Angeles Department of Transportation Mobility Data Specification	Chicago TNP Reporting Manual	Toronto PTC Data	New York TLC data requirements for high-volume for-hire services
REQUEST	Request time		• Datetime stamp			• Datetime stamp (only prior to April 2017)	• Request datetime stamp • Cancel datetime stamp (if cancelled)
	Requestor location		• Census Tract				• Request latitude/ longitude • Request address • Intended destination latitude/longitude • Intended destination address
	Request additional info		• Wheelchair accessibility request • Pooled/shared service request • Accepted? • Reason and datetime rejected *(if rejected)*		• Pooled/shared service request		• Request method: app, website, phone, other • Request outcome - completed, no accepts, passenger cancelled, passenger no show, driver cancelled • Wheelchair accessibility • Unique ride ID • Provider TLC license number • Pooled/shared service request • Wheelchair accessibility request
PERIOD 1 - LOGGED ON	Period 1 Time		• Start datetime stamp	• Start datetime stamp	• Start (log on) datetime stamp • End (log off) datetime stamp		• Start (log on) datetime stamp • End (log off) datetime stamp
	Period 1 Location		• Start Tract	Telemetry data with • Device ID • GPS coordinates (lat/lon) • Altitude • Heading • Speed • GPS accuracy (3 measures) • Timestamp of last telemetry report			• Route breadcrumbs via datetime stamped lat/lon • Vehicle entered or exited congestion zone, w/ datetime stamped lat/lon of event
	Period 1 additional info		• VMT • VHT		• Driver's license number • Driver's license state		• Total VMT session • Reason for logging off (log off or left service area) • Unique session ID • Driver TLC license number
PERIOD	Period 2 time		• Start datetime	• Start datetime stamp			
	Period 2 location		• Start tract	Telemetry data with • Device ID • GPS coordinates (lat/lon) • Altitude • Heading			• Vehicle entered or exited congestion zone, w/ datetime stamped lat/lon of event

Table 2: Six Mobility Data Specifications (Continued)

		Electronic Frontier Foundation (as proposed to CPUC)	City and County of San Francisco (as proposed to CPUC)	Open Mobility Foundation/ Los Angeles Department of Transportation Mobility Data Specification	Chicago TNP Reporting Manual	Toronto PTC Data	New York TLC data requirements for high-volume for-hire services
2 - EN ROUTE				• Speed • GPS accuracy (3 measures) • Timestamp of last telemetry report			
	Period 2 additional info		• VMT • VHT				• Total VMT for session
PERIOD 3 - TRIP	P3 locations	• Start Tract • End Tract	• Start tract • End tract	• Start GPS coordinate (lat/lon) with altitude and accuracy • Route polyline, with accuracy information	• Start census block • End census block • (*optional* by arrangement: lat/lon)	• Start intersection (or municipality if outside Toronto) • End intersection (or municipality if outside Toronto)	• Start lat/lon (5 decimal places) for each unshared trip or passenger in a shared ride • Route breadcrumbs via datetime stamped lat/lon • Vehicle entered or exited congestion zone, w/ datetime stamped lat/lon of event
	P3 times	• Start day of week • Start hour of day • End day of week • End hour of day	• Start datetime • End datetime	• Start datetime • End datetime	• Start datetime • End datetime	• Start hour • Trip duration	• Start (pickup) datetime (distinguished from arrival on scene) for each unshared trip or passenger in a shared ride • End (pickup) datetime for each unshared trip or passenger in a shared ride • Trip duration
	P3 Additional Info	• Pooled/ Non-pooled	• VHT (minutes) • VMT • Fare • Tip • # Passengers	• Trip ID • Parking verification URL • Cost of trip (sticker price, actual cost, currency) • VMT • Duration (minutes)	• Driver's license number • Driver's license state • VIN • VMT • Total cost • Total fare • Tip • Taxes and fees • Other charges • Shared trip ID (if shared)	• VKT (VMT) • Shared trip ID (if shared)	• Route ID (in addition to Ride ID for shared rides)t • TLC License number of vehicle • TLC License number of driver • TLC Provider ID • VMT for trip (plus aggregate of total passenger VMT per session) • Base fare • Tolls • Black Car Fund surcharge • Sales tax • Congestion surcharge • Tips • Driver pay • Number of passengers • Shared/pooled trip match successful • Wheelchair accessible vehicle

Table 2: Six Mobility Data Specifications (Continued)

	Electronic Frontier Foundation (as proposed to CPUC)	City and County of San Francisco (as proposed to CPUC)	Open Mobility Foundation/ Los Angeles Department of Transportation Mobility Data Specification	Chicago TNP Reporting Manual	Toronto PTC Data	New York TLC data requirements for high-volume for-hire services
Vehicle Information		• VIN • Make • Model • Propulsion • Wheelchair accessibility	• Device ID • Provider ID • VIN • Vehicle type • Propulsion type (human, electric assist, electric, combustion) • Year • Make/manufacturer • Model • State of charge	• VIN • License plate number • License plate state • Make • Model • Color • Model year • Driver's license and issuing state • Wheelchair accessibility • Inspection date	• Type of service, including wheelchair accessible vehicles	• Vehicle license plate • Type of service, including wheelchair accessible vehicles
Operator Information				• Driver's name • Driver's license number • State issuing driver's license • Driver's start date with provider • Driver's end date with provider		• TLC driver's license number • Was trip conducted by W-2 employee? • VHT and VMT spent outside city boundaries
Additional Info			• Vehicle status updates for maintenance, retrieval, etc.		• Trip status - Completed or cancelled (only until	
			• Last vehicle event (used for querying activity for a specific vehicle)		April 2017)	
Notes	• EFF proposal would only report combinations of day/hour, begin tract, and end tract for which at least 3 observations exist.		• MDS is intended to be used for all mobility services was not developed specifically for TNCs, but can be adapted for TNCs. This table represents the author's adaption of MDS data for TNC trip periods.	• Data is reported quarterly • Applies only to trips starting or ending in the City of Chicago • Some fields from Vehicles, Drivers, and Trips published publicly		• Data is submitted bimonthly or monthly • Trips leaving New York City subject to reduced reporting requirements • High-volume for-hire services also produce reports on employee pay

Which Analyses Questions Can Be Supported Answered With Data Provided Under Each Specification?

Table 3 (below) shows whether specific analyses could be supported by each data specification, as well as whether the question is applicable to shared mobility devices (SMD), transportation network companies (TNC), or both. Data produced under the Electronic Frontier Foundation's proposal to the CPUC would not suffice to answer any of the planning and policy research questions listed in Table 1. Data produced under the New York City Taxi and Livery Commission's Standards for high-volume for-hire vehicles and the Open Mobility Foundation's Mobility Data Specification would support most of the analyses listed. However, only the latter could be used for real-time compliance assessment, enforcement, and operational command and control.

Data Sharing Requirements of AB 1112

Proposed legislation (AB 1112) differentiates between disaggregated individual trip data and aggregate trip data but does not specify which data may be required. The City and County of San Francisco (as proposed to CPUC), Open Mobility Foundation/Los Angeles Department of Transportation Mobility Data Specification, Chicago TNP Reporting Manual, Toronto PTC Data, and New York TLC data requirements for high-volume for-hire services all require sharing of data defined as "individual trip data" in the 6/19/19 version of AB 1112, which would prohibit sharing of individual trip data with local authorities.

Table 3: Sufficiency of Data Specification for Analyses

Analysis SMD= shared mobility device TNC = transportation network company	EFF	SF	OMF MDS	Chicago	Toronto	New York
✓ = sufficient; ○ = possibly or partially sufficient						
1: Bicycle usage (SMD)			✓			✓
2: Curb management (SMD, TNC)			✓		○	✓
3.1: Labor - hours and trips (TNC)			○	✓		✓
3.2: Labor - employment status (TNC)						✓
3.3: Labor - driver pay (TNC)						✓
4.1: Neighborhood availability (SMD, TNC)		✓	✓	✓	✓	✓
4.2: Neighborhood cost (SMD, TNC)		✓	✓	✓	✓	✓
5.1 Congestion pricing (SMD, TNC)		✓	✓	✓		✓
5.2: Sustainability (SMD, TNC)		✓	✓	✓	✓	✓
5.3: Vehicle electrification (TNC)		✓	✓	✓	✓	✓
6: Regulatory compliance (SMD, TNC)		○	✓	✓	○	✓
7.1: Driver's safety records						
pots (SMD, TNC)			✓			
7.3 Speed violators (SMD, TNC)			✓			
8.1: Transit ridership (SMD, TNC)			✓	✓	✓	✓
9.1 TNCs and Congestion (TNC)			✓			✓
9.2 Modal comparison (SMD, TNC)		✓	✓	✓	✓	✓
10.1 ZEV infrastructure (TNC)			✓			
10.2 ZEV grid demand (TNC)			✓			
10.3 ZEV TNC driver behavior (TNC)			✓			
Total Answerable Example Questions (Possibly ○ valued at ½)	0	6.5	16.5	8	7	14

Three Models for Sharing and Analyzing Data

Shared mobility providers collect and retain more data than would be used in the analyses listed in Table 1. Such data might include personal information on users (e.g., name, address, payment information), app usage information (e.g. geolocation and timestamp for each occurrence the app is opened, even if services are not requested), and operational information (e.g., information on the party which last inspected or maintained a mobility device or vehicle).

This section presents three alternative models a regulatory authority can specify for access to disaggregated data, analysis, and sharing of aggregated statistics and information. Each model varies in the extent to which potentially personally identifiable information is excised, resulting in some differences in what types of analysis are possible. However, some models can limit the utility of the analysis to end users and preclude the ability to perform newly-identified, retrospective analyses.

1) Mobility Provider Leads Analyses

In this model, the shared mobility provider conducts the analysis of data collected from its users, following this process:

1. A regulatory agency, with possible input from other agencies and researchers, specifies the data requirements, the analysis, and the methodology and requires regular reporting of aggregated analysis as a condition of permitting or licensing.
2. The shared mobility provider uses their collected, disaggregated data and uses them to conduct the necessary analyses.
3. The shared mobility provider aggregates the results and shares them with the regulatory agency and/or public.

Pros:

- Potentially personally-identifiable data that are collected by the shared mobility provider are not transferred to another party.
- Shared mobility provider has the greatest level of experience with the disaggregated business data and is likely to have the internal capacity to perform the analysis.

Cons:

- Requires that the regulatory agency trust the accuracy of the data and analysis performed by the mobility provider.
- The agency's pre-specification of data and analysis needed (as a condition of regulatory approval) could result in limited datasets that preclude future analyses.
- Providers may be reluctant to accurately self-report aggregated statistics that could regulatory agencies could use to assess compliance violations or revoke permits.

2) Regulatory Agency Leads Analyses

A regulatory agency, such as a state or city authority compels shared mobility service providers to report certain data as a condition of licensing or permitting. In this model, analyses of data would follow this process:

1. The regulatory agency specifies the data needed for analysis and requires shared mobility providers to retain and report these data.
2. The shared mobility provider sends the disaggregated data to the regulatory agency and the agency conducts the analysis.
3. The regulatory agency may choose whether or not to share data with third parties under agreements or to publish all or a subset of the disaggregated data.
4. The regulatory agency uses aggregated results internally and may share all or a subset of aggregated results with other public agencies or with the public.[6]
5. The regulatory agency may retain data for future purposes, including analyses not identified at the time the agency specified the data sharing requirements.

Pros:

- The regulatory agency has increased flexibility in using data to analyze new questions not previously specified.
- Access to disaggregated data allows agencies to compare provided data with field observations in audits, which increases trust in the data and analysis.
- The regulatory agency gains experience analyzing disaggregate data, which may enhance its capacity to use this and other data for decision-support.
- The regulatory agency may use disaggregated data for compliance and control purposes, including in real-time.

Cons:

- Requires that the regulatory agency trust that disaggregated data provided is complete and accurate.
- Certain disaggregated data published could be used to identify individuals, even if the published data does not include personally-identifiable information.
- Retrospective analysis of new questions is dependent on data previously reported.
- Regulatory agency may not have the internal capacity to perform the analysis, although this can be alleviated by sharing disaggregated data with a third party under contract with the regulatory agency.

3) Third Party Brokers Data and Conducts Analysis

Third parties can serve as data brokerages for disaggregated mobility data and perform analytical services for public agencies on disaggregated mobility data. A successful data broker must develop an ongoing, trusting relationship with the shared mobility provider in order to gain access to the data that their public agency clients need for planning and policy analysis, and data that has not been pre-specified by the brokerage or regulatory agencies. In this model, public agencies do not obtain direct access to disaggregate data.

Examples of third party brokerages are SharedStreets, Remix, and the App-Based Transportation Clearinghouse. SharedStreets is a project of the 501(c)(3) Open Transport Partnership and is backed by the 501(c)(3) National Association of City Transportation Officials (NACTO), an association of large cities and transit agencies. SharedStreets presently brokers disaggregated data from TNCs to visualize pickup and dropoff locations. Remix is a private company that has received investment from venture capital firms and provides analytical services to regulatory agencies which obtain and share disaggregated data or instruct mobility providers to share data directly with Remix. The App-Based Transportation Clearinghouse is a product of the Airport Research and Development Foundation and is used by Airports with non-exclusive regulatory authority over TNCs. A data trust or university may also serve as a third party.

For analyses conducted by a third party:

1. The regulatory agency may require regular reporting of specified data to a brokerage as a condition of permitting or licensing, except where that data may be voluntarily provided.
2. The regulatory agency, with potential input from other agencies and researchers, specifies the questions, data, or specific requirements. and methodology needed to perform a requested analysis.
3. The third party conducts analysis.
4. The third party shares results with public agency clients, or may share results publicly if requested by the public agency client.

Pros:

- Third party may be identified by mutual agreement of the shared mobility provider and the regulatory agency.
- If the third party serves multiple public agency clients, it can amortize fixed costs to develop analytical tools and staff expertise over many end users to lower overall costs and use and share best practices in mobility data collection and analysis with those clients.
- A third party serving many public agency clients can develop a stronger working relationship with the shared mobility provider than any individual client may be able to develop on their own.

Cons:

- Requires that regulatory agency trust the accuracy of data and analysis performed by the third party brokerage.
- If data sharing is not required by regulation or contract, there may be limits to data that a mobility provider voluntarily share with the data brokerage, which could hinder some analysis.
- The public agency does not develop the capacity to analyze disaggregated data, which may limit its ability to understand the applicability of aggregated analysis to decision-making.
- There may be issues surrounding both long-term control of mobility data by one brokerage and issues deploying multiple brokerages.

Conclusion

The twenty analyses specified in this brief are a small sample of the myriad planning and policy research analyses that shared mobility data could support. Whereas the analyses specified here focus on historic trends, shared mobility data could enable superior forecasts and modeling to answer future questions and address new challenges. We encourage decision-makers to consider a data sharing specification and model that allows analysis to be done promptly as new issues are identified. This could be accomplished by requiring that more expansive set of data be shared with the regulatory agency or a third party or by requiring shared mobility providers to retain data for future use and analysis.

[1]For additional background information refer to the following issue paper, D'Agostino, Mollile, Paige Pellaton, Austin Brown (2019) Mobility Data Sharing: Challenges and Policy Recommendations. Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-19-19

[2]Footnotes. "Shared mobility providers" refers to Transportation Network Companies such as Uber and Lyft, as well as companies that offer scooter- and bike-sharing. Public transit is not included in this definition.

[3]Bliss, Laura. (2019, July 31). Why Real-Time Traffic Control Has Mobility Experts Spooked. https://www.citylab.com/transportation/2019/07/digital-twin-mobility-data-standard-city-real-time-traffic/593914/

[4]Period 0 - Deadheading: TNC driver is going to/from TNC service without app open. Period 1 - Fishing: The app is open and the driver is waiting for a match. Period 2 - Enroute: Match accepted but passenger not yet picked up. Period 3 - Passenger Trip: Passenger is in vehicle

[5]Under CPUC Proceeding 12-12-011 a Scoping Ruling was issued by the Commission on October 25, 2019 asking for public comment for suggested changes to Decision.§ 13-09-045, which established protocols for annual and confidential reporting directly to the CPUC To date, as of January 30, 2020.

[6] An example of a publicly-accessible dashboard of aggregated data is available from the New York City Taxi and Limousine Commission and a private individual's NYC Taxi & Ridehailing Dashboard).

Research presented in this policy brief was made possible through funding received by the University of California Institute of Transportation Studies (UC ITS) from the State of California via the Public Transportation Account and the Road Repair and Accountability Act of 2017 (Senate Bill 1). The 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, the UC ITS has branches at UC Berkeley, UC Davis, UC Irvine, and UCLA.

DOI: 10.7922/G2ZG6QHH