Eno
Center for
Transportation

DATA ON DEMAND

A CASE STUDY IN THE LOS ANGELES AND PUGET SOUND REGIONS

Data on Demand:

A Case Study in the Los Angeles and Puget Sound Regions

FEBRUARY 2020

Contact: publicaffairs@enotrans.org
www.enotrans.org | 202.879.4700

Authors

Alice Grossman, *Senior Policy Analyst, Eno Center for Transportation*
Paul Lewis, *Vice President of Policy and Finance, Eno Center for Transportation*

Acknowledgments

The report authors would like to thank Romic Aevaz, the staff at LA Metro, King County Metro, and Sound Transit as well as representatives from Lyft and Via for their review and input. Thank you also to the research collaborators, Anne Brown, Mark Hallenbeck, and Michael Manville for their assistance in project development.

About the Eno Center for Transportation

The Eno Center for Transportation is an independent, nonpartisan think tank whose vision is for an American transportation system that fosters economic vitality and improves the quality of life for all. The mission of Eno is to shape public debate on critical multimodal transportation issues and build an innovative network of transportation professionals. As an organization, Eno values integrity, independence, objectivity, quality, and relevance. These core values are reflected in everything we do.

Table of Contents

About the Federal Mobility on Demand Project

Mobility on demand (MOD) refers to transportation services that can be hailed in real-time for an impending trip. MOD integrates data such as location tracking and traffic conditions, with user-entered destination and payment information. Though most MOD services are designed for users to interface using a smartphone, MOD can be requested through a web browser or call center, which can increase accessibility and equity of the service for people without access to a smartphone, people vision impairments, people who require non-English communication, and others. While MOD is not a new concept, recent technological advancements facilitate its deployment in a new way. Its role in the future of transit systems is yet to be determined.

In May 2016, the Federal Transit Administration (FTA) announced $8 million in funding for its Mobility on Demand Sandbox Demonstration Program. The program is part of FTA's support of transit agencies, government entities, educational institutions, and communities as they experiment with on-demand mobility tools such as smart phone applications and shared mobility services to augment and enhance existing transit agency services. MOD Sandbox was developed to test new ways to encourage multimodal, integrated, automated, accessible, and connected transportation. Among the key features of the program is its focus on local partnerships and demonstrated solutions in real-world settings.

Some of the eligible activities applicants could propose to advance MOD and transit integration were new business models for planning and development, the acquisition of new equipment, services, software and hardware, and operation of the project in a real-world setting. Eligible partners included public transportation providers, state and local departments of transportation, federally recognized Indian tribes, private for- and not-for-profit organizations, transportation service operators, state or local government entities, consultants, research institutions and consortia, and not-for-profit industry organizations. In October 2016, 11 projects were selected for funding (See Appendix A).

The largest project awarded was a two-region partnership between Los Angeles and the Puget Sound Region. The Los Angeles County Metropolitan Transportation Authority (LA Metro) collaborated with King County, Washington Metro Transit (King County Metro) and the Central Puget Sound Regional Transit Authority (Sound Transit) on a project to contract with a transportation network company (TNC) to provide first/last mile service to select transit stations near disadvantaged communities. This proposal included evaluation and reporting by the Eno Center for Transportation and local research universities. The FTA awarded the team a grant of $1.35 million for the pilot and corresponding research.

The stated overall goal of the Los Angeles/Puget Sound project is to: 1) define how TNC services can be aligned with existing transit service to serve an effective first-mile/last-mile solution; 2) define how key partners can cost-effectively ensure equal access for individuals with disabilities and low incomes; 3) demonstrate payment integration across transit operator and TNC platforms, specifically to enable service to lower income and unbanked populations.

1. Introduction

Transit agencies across the country are testing mobility-on-demand (MOD) projects in order to evaluate whether or how such pilots could be part of their long-term service plans. This includes integrating private options into agencies existing trip-planning applications, using a private operator for on-demand first/last mile (FMLM) connections to transit stops, and providing real-time flexible route service to replace underutilized transit routes. Importantly, these private companies have access to detailed data about how their network operates within the overall transportation system. Public access to those data is crucial for good service planning, operations, accounting, and evaluation.[1]

Accurate, granular data are necessary to assess MOD projects and pilots and to measure progress towards identified goals and objectives. Agreements to share data between private and public partners are the backbone of successful MOD projects since they outline agreed-upon parameters for data ownership, access, storage, and usage. Starting with research and service goals for MOD projects, as well as the data needed to meet those goals, helps agencies tackle the challenges associated with collecting and sharing data by setting clear needs and priorities.

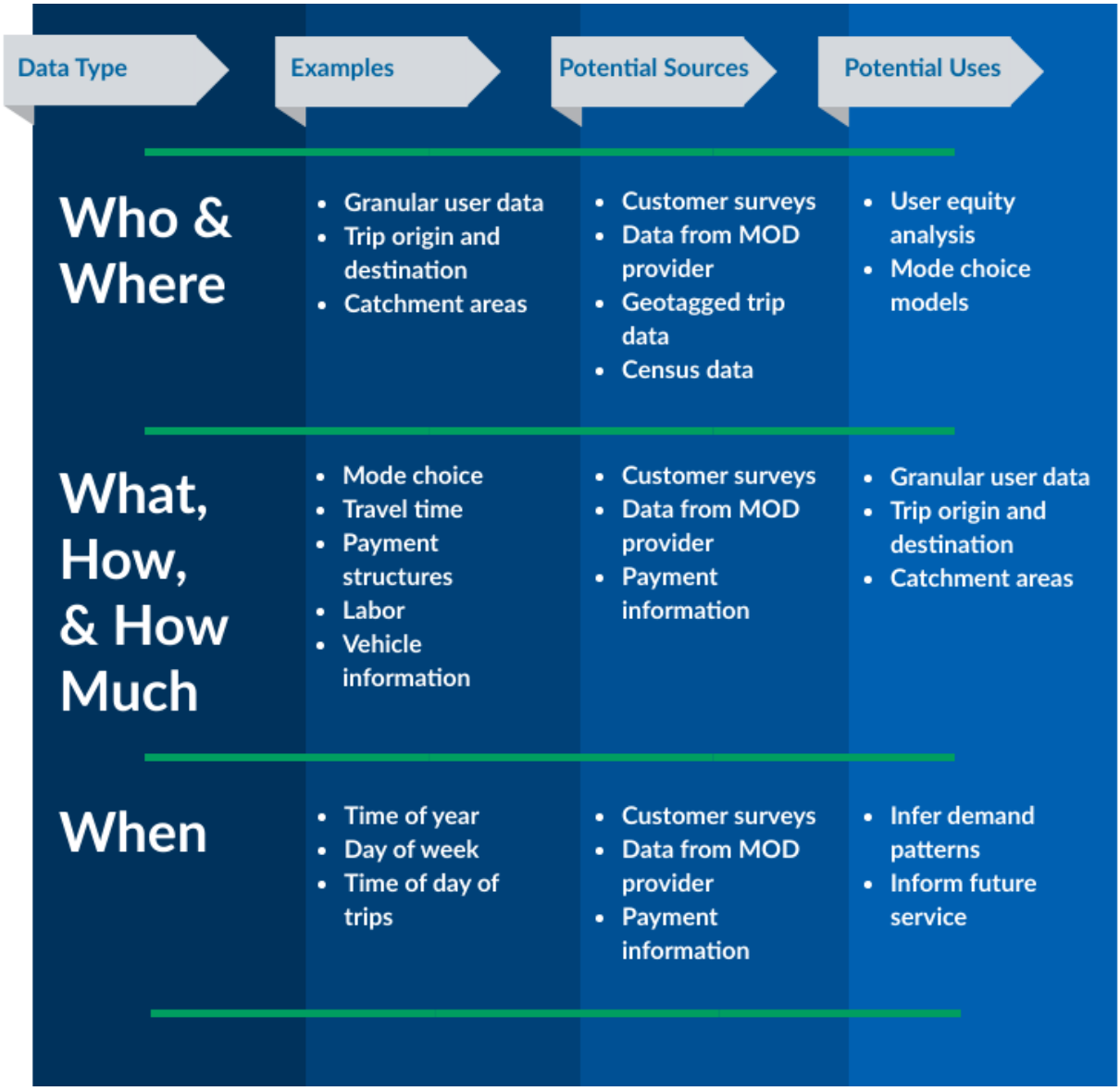
Many MOD providers are relatively new private companies often referred to as transportation network companies (TNCs).[2] Differing goals, organizational structures, requirements, and service types between public transit agencies and private MOD providers must be understood and addressed to provide useful, coordinated projects. As more pilots and partnerships emerge, so are a useful set of best practices and lessons learned for how data should be shared between transit agencies and MOD providers.

Combining existing data as well as new data sets—such as observed and survey data—can provide a robust project evaluation framework. Quantitative evaluation, in conjunction with qualitative evaluation, helps inform planning, funding, partnering, and service decision-making to reach agency goals. In order to measure progress, Figure 1 shows examples of what types of data can be collected from various sources and how they can be used in the context of evaluating a transportation program that provides enhanced transit station access.

This report examines the data needs that agency staff need to consider when developing a MOD agreement with private providers. Background information on elements of data sharing in this context includes general state of the practice, and challenges and opportunities for transit agencies. The FTA Mobility on Demand Sandbox project in the Los Angeles and Puget Sound regions serves as a case study with robust data sharing between multiple parties. Both general information and the case study help provide examples of various levels of success in developing and implementing these types of collaborations. The purpose is to inform transit

agencies, private MOD providers, and researchers of elements to consider when developing data sharing agreements.

Figure 1: Data Sharing Agreement Considerations



2. Background

2.1 Data Sources and Use

Public transit agencies have long used various types of data for service planning, project prioritization, and equity analysis, among other functions. These data are often collected by the transit agency through fare collection or with rider intercept surveys, but sometimes are gathered from external sources which may be publicly available such as the federal government's National Household Travel Survey

(NHTS), or may require an agreement between the transit agency and another party for data from toll payments or cell phone location. Planning, accounting, and evaluating needs for MOD projects can draw upon a range of data sources from transit agencies, mobility providers, rider surveys, and publicly available data such as from the U.S. Census.

Understanding what data are available and how they can be used helps agencies assess formats and the appropriate levels of aggregation to look or ask for when developing mobility partnerships. Data can come from various sources and fall into one of four categories:

- *Internal transit agency-collected data* – Transit agencies collect data on assets, service, fares, and ridership. Data such as boardings and alightings – potentially measured through fare payment data (i.e. tapping in and/or out) or automated passenger counters – and other internal data can shed light on a MOD pilot's impact. Agencies aiming to boost ridership, improve service quality, or increase social equity, can compare impacts of the MOD service by drawing upon their own ridership data to track and evaluate the extent to which (or whether) they have met these goals.
- *External data sets* – Publicly available data such as Census data, the National Transit Database (NTD), land use or zoning data provides additional context and elements for analysis. Private sources such as cell phone location data or credit card/marketing data is also potentially useful and readily available for purchase.[3]
- *MOD provider-collected data* - In projects where an external partner provides trips, those partners have detailed information including counts of trips taken and the number of unique riders. MOD providers typically have data at a fine scale given their use of GPS equipped smartphones for reserving and paying for trips, including origin-destination information, time stamps, and trip distance, that can be used to evaluate where MOD service is being utilized, whether the service is achieving geographic equity goals, what the impact is on vehicle miles traveled (VMT), and what wait times are (if any). These data can be used to compare MOD provider service performance to existing transit or other services.
- *Project user survey data* – Some types of data are not available at the agency or from the MOD provider and must be asked of riders directly. Rider surveys always have inherent survey bias, but can allow agencies to evaluate the effects of MOD service integration on rider behavior and mode choice.[4] For example, the Dallas Area Rapid Transit (DART) implemented a partial integration of MOD and other mobility providers apps into their existing apps with a goal to improve rail station access. That pilot uses rider surveys to evaluate perceptions, use, and connections to DART services.[5] Rider surveys that capture demographic information of riders (i.e. income, disabilities, ethnicity, payment method) are also a useful means of providing data for use

in evaluating whether equity goals of a project have been met. For example, if a pilot improves transit access among lower income riders, the unbanked, or passengers with physical disabilities.

Ultimately, data sources fit into different categories depending on governance and service structure and combining them is helpful to better understand whether agencies are meeting their project's goals. Surveys and MOD user data tied to existing transit fare payment data can help assess whether ridership is largely driven by existing public transit users or by attracting new riders. Survey data combined with agency fare payment user information can also inform income and sociodemographic equity analyses.

MOD Partnership Data Example: Pinnellas County, FL

Origin-destination data tied to fare payment can reveal whether riders in targeted areas are utilizing MOD services and inform adjustment of geographic boundaries. For example, during one of the first transit agency MOD pilots to directly subsidize TNC trips, Pinellas Suncoast Transit Agency (PSTA) in Florida partnered with ride hail companies to provide riders with trips to and from stations. In the first phase of the pilot, the agency set a 400-foot drop-off zone around eligible bus stops for the MOD trips. After which PSTA and the private company noticed that the boundary sometimes led drivers to drop passengers off on potentially dangerous, high-traffic streets. Based on the data, PSTA expanded the boundary to 800 feet.[6] The PSTA example underscores the value trip data (particularly geospatial information) have in program evaluation and ongoing planning. In cases where individual trip-level data is not shared, highly aggregated origin-destination data could prevent officials from drawing meaningful conclusions about travel patterns.[7]

Data from all partners are also critical for auditing, compliance, and accounting purposes. Information on trips provided, along with origin-destination and time stamp information, allows transit agencies to ensure that the provider is following the terms of their contract and providing trips as agreed. Some projects might set targets for companies and agencies to work towards, such as average wait times or total number of users per driver hour. Information on the number of wheelchair-accessible vehicles (WAVs) available, or the number of WAV trips requested and fulfilled, are often also critical to evaluation. In cases where an agency subsidizes trips, information about the cost of each (MOD trips and transit trips) allows agencies to accurately conduct cost-per-rider comparisons between fixed-route or paratransit services and the pilot program. Agencies are also able to measure and monitor 1) budgeted expenditures for subsidized rides, 2) the occurrence, use and value from surge pricing, 3) passengers-per-vehicle-per-hour, and 4) costs between shared versus non-shared rides.[8]

2.2 Data Formats, Access, and Storage

When data are coming from multiple sources, an important consideration lies in the compatibility and usefulness of formats, as well as methods for storage and access. Data ownership is often made clear in advance of data sharing, and many options exist for accessing data owned by others. Format and standardization vary among different data sources; for example, transit agencies report general transit feed specification (GTFS) data inconsistently and MOD providers collect and share different variables at various levels of aggregation. Understanding what data are available and how they will be used can help agencies assess formats and levels of aggregation.

The owner of a particular set of data has the ability to control access to it, to the extent desired as applicable under state and Federal laws. The data can then be shared through licensing or other agreements. In some cases, institutions, public or private, with strict data controls may not allow any transfer of data outside of their organization. In these cases, researchers or partners who wish to access raw data might have to physically travel to the facility where the owner of the data allows access to see and manipulate information in their presence. This practice can occur with private or public companies, is typical for audits of confidential data, and is common with governmental data sets, especially in settings where a department does not have the most up-to-date software or computing system to better facilitate secure data sharing.

Slightly less restrictive are data sharing agreements that allow for data access by way of login access at an off-site confidential data room or through a portal hosted by and belonging to the company that owns the data. For example, existing for-purchase aviation datasets typically grant users access to an online database manager. Users can run analyses through the portal but not download any raw data.[i] With these types of arrangements, access to the portal is only accessible to specified users and expires on a certain date and the data can no longer be accessed or manipulated.

Other models allow for downloading data sets, which gives the user a raw data file that they can import into any analysis software such as R, SPSS, SAS, or Excel. Some agreements allow the user to store and maintain that data for as long as they wish, provided they follow specific privacy and security stipulations such as removing unique identifiers, aggregating data, and maintaining secure server and computer access. Comprehensive data sharing agreements address these aspects to make sure that the agency will be able to use the data in the way that they intend for analysis communicated both internally and externally.

---

[i] Examples include datasets from Sabre Corporation and OAG.

2.3 Staff and Data Expertise

Depending on their size, and given the competition with high-paid opportunities available in the private sector, public entities like transit agencies often lack the necessary resources to hire in-house data scientists.[9] They may also lack the necessary software, servers, hardware and other infrastructure to store, process, and analyze large, complex datasets.[10] As a result, agencies need to assess data processing capabilities at their disposal either within the agency or contracted out when planning a MOD project. Factors to consider include staff ability to craft data collection strategies and performance monitoring plans, perform data collection, craft data sharing agreements, develop and administer surveys, manage data portals, protect user privacy, and analyze data for project evaluation and system management.

In cases where agencies do have robust capacity, staff may be able to do much of the data collection, management, and analysis in-house. In other cases where an agency lacks expertise and data management capacity, or where public records laws or privacy considerations preclude the agency from housing proprietary data from an external partner, agencies may wish to rely on outside entities like think tanks, universities, or data analytics firms.[11]

Agencies can draw upon external expertise by including an outside evaluation team in the project planning and contracting phase, particularly when developing evaluation plans and data sharing agreements. Outlining specific data sources and evaluation metrics at the start of a pilot can also help agencies identify which analyses can be completed in house, and which should be outsourced.[12] For example, agencies may task outside entities with administering rider surveys, managing and updating a data portal or other data repository, inspecting data quality, or providing more sophisticated data analytics such as a dashboard that shows current summary statistics.

While incorporating third parties or an external research team into a project may allow agencies to draw upon outside data management expertise, it can be expensive. It also requires further coordination of data sharing agreements across multiple entities, each with varying goals and privacy restrictions, potentially adding further complexity to the contracting process.[13] Managing data access, usage, and rights among a broader range of users is complicated, particularly in cases where sensitive or proprietary data from an agency or private mobility provider are being shared.

2.4 Privacy and Cybersecurity

Data collection, storage, and sharing introduce the need for data sharing agreements to consider personal privacy, public cybersecurity, and company trade secrets. Privacy considerations play a critical role in the collection, sharing, management, and use of personal mobility data. Even when personal identifiers are

removed from data sets, they can contain information that easily allows for re-identification of individual users. Studies of mobility data with course hourly location information for individuals can lead back to unique identification of individuals 95 percent of the time.[14] New mobility technologies increasingly generate a wealth of user data with granular location information collected multiple times a minute, particularly with trip planning and booking apps, which create new challenges in data privacy and ownership. The lack of clear ownership policy in the United States and the patchwork of data privacy laws makes it difficult to navigate the sharing and management of personal mobility data on a national level.[15]

MOD providers have been hesitant to share detailed trip data out of concern for protecting their customers' privacy as well as their own trade secrets. On the customer privacy front, just a few cases of detailed origin-destination data can reveal the identities of users.[16] On the trade secret front, companies prefer to release data only if it cannot reveal insights on trip routing and other algorithms that would compromise information on their market share, venture capital fundraising opportunities, pricing or internal operations. In most cases, MOD providers require the use of non-disclosure agreements (NDA) that legally limit who can access their data and what those with access to the data can release to the public.

Fears of compromising sensitive user and internal proprietary data stem from both Public Records Act (PRA)/Freedom of Information Act (FOIA) and cybersecurity concerns. Private entities may be concerned that data shared with an agency could be made public through a PRA or FOIA request, which is unique to public entities, or leaked/compromised through a data breach, which could happen in any variety of settings.

Cybersecurity considerations, particularly around data storage, are also a critical part of data sharing negotiations. The persistence of hacking, data breaches, and other cyber-attacks on both public and private organizations has only reinforced the need for strong cybersecurity infrastructure to prevent the unauthorized disclosure or theft of sensitive mobility information or the potential nefarious use of technology or data.

Data sharing agreements can address access and ownership rights to facilitate assessment, evaluation, and good stewardship of sensitive data. Agreements can include specifications on levels of aggregation to be collected and reported, frequency of the collection, limitations on how long raw data can be stored, implementation of ongoing security audits, data breach reporting provisions, and/or restriction of access to sensitive data. Provisions on format access, and storage can protect sensitive data with strong cybersecurity and privacy protections (See Section 2.2).[17]

2.5 Laws and Regulations

Federal, state, local, and agency specific laws and regulations between transit agencies and MOD providers play a role in data sharing requirements and barriers.

At the federal level, the United States has few limitations on the collection and use of location-based data. Some proposals for federal legislation are modeled after the European Union's General Data Protection Regulation (GDPR), which has strict data ownership, consent, and handling requirements.[18] While there is currently no national data privacy law, the Federal Trade Commission has taken steps to regulate, stating that location-based data should be considered sensitive material and require disclosures and affirmative consent before collection.[19] The state of data privacy regulation in the United States will continue to evolve as Congress and state governments consider new proposals and rules that clarify rights and requirements around data privacy, ownership, and management.

In the absence of federal leadership, states developed their own regulatory environments for data. Within the patchwork of state data privacy laws, California's Consumer Privacy Act, enacted in 2018, stands out. The law deems geolocation data that is or can be linked to an individual as protected personal information, and gives users the right to learn how their data are being processed. The law also allows users to direct companies to delete or stop selling their personal location data under certain circumstances.[20]

Trade secret and public records laws vary across municipalities and states, particularly in regard to information that is exempt from disclosure. On the state level, a 2014 example where Uber and Lyft agreed to share zip-code level trip data with Seattle rose to the state supreme court level.[21] The Court decided whether this data constituted a trade secret under Washington State's Uniform Trade Secrets Act (UTSA) when a researcher requested access to the data. The Court ultimately ruled that the aggregated data would not cause the company irreparable harm, and that state's public records law superseded the UTSA and did not prevent disclosure of the data.[22]

Creating quality data sharing agreements may involve working with the state legislature to exempt personally identifiable information and other forms of electronic data from app users from disclosure laws, as DART in Texas and TriMet in Oregon have done. Or agencies can include specific legal language that outlines which data are deemed trade secrets or personally identifiable information and exempt from disclosure under state public records laws, as LA Metro has done in California.[23]

Other regulations impacting data sharing include FTA reporting requirements. Agencies that receive grants from the FTA are required to report specific trip information to the National Transit Database (NTD), including when agencies

contract services out to private operators. However, pilot projects funded by FTA such as the MOD Sandbox program, are exempt from this reporting requirement. The variables needed for NTD reporting typically guide agency's data needs and planning. TNCs and other shared mobility providers are currently not formally included in NTD reporting requirements, though the FTA appears likely to modify requirements to allow for more accurate trip reporting and performance management.[24]

3. Case Study: Los Angeles and Puget Sound MOD Pilot

Data sharing was a key part of the Los Angeles and Puget Sound MOD Pilot project. The LA Metro, Sound Transit, and King County Metro service partnerships with the MOD provider Via required robust data sharing agreements 1) to allow the transit agencies and Via to evaluate and adjust service throughout the year-long pilot, and assess potential for extension of the pilot and 2) for the associated research team to be able to assess access and equity impacts of the pilot. Two specific aspects of the pilot made the data sharing agreement particularly critical and complex.

First, the program included a heavy element of analysis. The stated overall goal of the LA Metro project, as specified in the original Coorporative Agreement for the grant award, is to: 1) define how and if MOD services can be aligned with existing transit service to serve as an effective FMLM solution; 2) define how key partners can cost-effectively ensure equal access for individuals with disabilities and low incomes; and 3) demonstrate payment integration across transit operator and TNC platforms, specifically to enable service to lower income and unbanked populations.[25] Research questions related to these goals include a need to understand how people access stations before and during pilot implementation, what types of people access the stations before and during pilot implementation, and why, when, and at what costs these trips happen.

The second factor that made a data sharing agreement both critical and complex was the inclusion of an external research team to lead efforts on data needs and to analyze the project from an external perspective. That team included the Eno Center for Transportation, the University of California-Los Angeles, the University of Washington, and later the University of Oregon. As the project evolved, changes and additions were made to the project team, but the core research team remained the same.

3.1 Negotiations

Data sharing negotiations between the transit agencies, MOD provider, and research team began prior to the award of the federal MOD Sandbox Grant for the project. In the initial grant application, LA Metro named Lyft as the MOD provider, and both parties were under a verbal agreement to share all of the data required to successfully evaluate the project. But ultimately, the vision for the project held by

the transit agencies differed too greatly from the vision by Lyft. This was in large part due to differences with respect to data sharing, provision of multi-lingual call center support, and delivery of wheelchair accessible vehicle service as part of service delivery. The partnership between Lyft and the transit agencies dissolved. After a streamlined competitive process, Via was selected to be the MOD provider to deliver on intent of the Cooperative Agreement or grant award. LA Metro entered into negotiations for a contract with Via for the LA County pilot deployment, and King County Metro entered into negotiations with Via for the Puget Sound pilot deployment.

After a complex process, Via was selected to be the new MOD provider. LA Metro and King County Metro entered into separate negotiations with Via for their pilot deployment.[26] Data sharing negotiations first progressed with informal discussions, followed by a signed "Term Sheet" that outlined the main data points requested in November 2017. The final data sharing agreement was part of the contracts signed by Via and the agencies in December 2018.[27]

Much of the negotiation for access to data with both Lyft and Via revolved around types of data and levels of aggregation. In the end, most reported variables involved negotiation and compromise between Via and the research team. For example, Via initially preferred to report many data fields—such as passenger wait time from the time of ride requested to actual pick-up—as averages, but agreed to report these data on a trip-level basis for better analysis. Furthermore, requests from the transit agencies for real-time data from Via ended with a compromise of weekly data uploads. As part of the negotiations, Via also raised concerns about personally-identifiably information and trade secrets.

The transit agencies decided to enter into a Term Sheet with Via, one for each pilot deployment, in order to clarify the intent of the partnership before entering into a full contract negotiation process. The transit agencies initially identified data needs in the Term Sheets that Via simply did not collect. For example, the agencies hoped for demographic information regarding service for riders, including people with disabilities to be able to use this to learn from the pilot and be able to better plan services in the future. However, Via does not collect demographic data on specific service elements such as users who put their wheelchair or other mobility devices in the trunk or in the vehicle, or specific trainings of working with people with disabilities of their contracted driver partners.

3.2 Term Sheet

Although not legally binding, the Term Sheets with Via were helpful in moving forward with the broader contract negotiations and ensuring that both parties agreed to data sharing in writing. Included were provisions that Via would own all of the Via data collected through their platform and give the agencies permission to use it for planning purposes. Setting this clear expectation allowed for both parties

to continue discussions around data sharing knowing that Via would continue to own their data, and the agencies would be able to access the data as desired. The Term Sheets set a clear understanding of expectations before investing more time in negotiating.

A key component of the Term Sheets was indicating therein what data points were requested and how the agencies planned to put them to use or, in other words, why they were needed. All of the requested data relate directly to project goals and research questions from the three transit agencies and research partners. Quantitative data allow for measurement of success in relation to the goals of the researchers and project sponsors. The Term Sheet set the broad expectation of Via sharing data with the transit agencies. It also listed specific variables and the level of granularity for each variable as well as the purpose of the data request relating each variable back to project goals, as summarized in Appendix B Table 1. The precise method of data access was not established until later, but the Term Sheet set a mutual agreement to base further discussions.

The Term Sheet also indicated the level of data aggregation for each field or area. The Term Sheet was a key step to securing the relationship between Via and the transit agencies by having a preliminary signed data sharing agreement. It also provided transparency to all parties involved by justifying each data request from the transit agencies. The Sheet was developed and signed by only the agencies and not the outside research team.

3.3 Content and Actual Data Received

The final contracts between LA Metro, King County Metro, and Via included a scope of work (SOW) with the data sharing requirements.[28] The SOW stipulates levels of aggregation and, in some cases, what units or variable categories should exist. Only some of the data shared are defined as trade secret by Via, and the SOW also specifies exactly which data are to be considered trade secret, and at what level of aggregation. The data included in the contract are compared with actual data received and data specified in the Term Sheet in Appendix B Table 2.

Even with specific requests in the contract, the final data received did not exactly match the contractual language and the contracts did not exactly match the original Term Sheets. For example, vehicle dwell time to calculate greenhouse gas emissions were not included in the final contracts. The number of times that a wheelchair or scooter user asked to get out of their chair for transport and the number of times they did not also were not included in the final contracts or provided nor were the number of drivers with ADA sensitivity and wheelchair securement training. The latter data were not included because they are not typically collected by Via, and negotiations led to their omission. Other data fields were added such as Station ID in Los Angeles when more stations within overlapping zones were added to the pilot. User comments through the Via app were also added when the transit

agencies learned that they were being collected anyway and Via agreed to share those data as had been specified earlier in the Term Sheets.

Some of the initial data requests that were not included—such as the omission of the vehicle year along with make and model used—slightly limit analysis but are not essential to meet the main goals of the project. Furthermore, in the Puget Sound region, all vehicles are the same make and model and year due to the structuring of the contract, eliminating the need for this variable.[29] The goals of measuring access to transit do not relate to vehicle year, but an additional interesting analysis of project-related emissions becomes less accurate without detailed vehicle information.

Via also shared data with the project team through there are collection system. In the Puget Sound region, integrating ORCA payment meant that fare collection data would already be owned by the transit agencies. In the Los Angeles region, since fare integration for the pilot was not possible, LA Metro asked Via to have riders enter their TAP card identification number when they signed up for the service. This allowed for LA Metro to track riders' interactions with the Via platform. However, once Via waived fares for rides early in the pilot, they received LA Metro's approval to remove the request for TAP identification number, thus lessening the ability to connect individual travel behavior on Metro and on the Via service.

3.4 Data Formats, Access, and Storage

The research team played a major role in shaping the data access and usage practices for the project. Initial discussion around using the University of Washington Transportation Data Center seemed like a logical use of a third party data warehouse to protect the data and provide easy access for researchers, but unfortunately the Center was not operational in time.[30]

Instead, Via took on data storage and access duties and provided data such that each member of the research team can access the data sets through a secure portal. The portal updates new data every week, giving the transit agency staff and researchers access to up-to-date data at any time. The contract between Via and King County Metro for the pilot deployment in the Puget Sound region also included a provision for data visualization provided by Via, which can be accessed through a portal for ease of the user. LA Metro initially preferred the raw data, but when it became clear that Via was open to providing both raw data and data visualization, LA Metro asked for this and Via complied.

Due to the nature of the data storage through Via, when researchers have technical difficulties with the portal, the problem must be solved through Via's technical support as researchers cannot send data over email or other less secure means. This can add time to analysis. Standard elements to data sharing such as shared data dictionaries are necessary for accurate interpretation and analysis of data, but were

not specifically included in the initial negotiations and had to be requested from Via later. This problem did not surface with transit agency data, as these data were already familiar to the researcher team. Difficulties with the transit data surfaced in other ways, such as through challenges linking fare collection data sets and vehicle data which are separated for privacy protection but must be joined for detailed analysis (with published results aggregated to eliminate personally identifiable information).

Via also provides the Puget Sound region with a table that lists the key performance indicators (KPIs) aggregated weekly as specified in the contract, including number of trips, number of passengers, average ride rating, percent demand met, call center percent demand met, WAV percent demand met, prepaid card trips, and vehicle utilization (passengers per driver per hour). The Puget Sound agencies and Via meet monthly to review KPIs, and a provided data table helps facilitate these conversations. LA Metro calculates the KPIs on their own with the provided data and also reviews the measures with Via on a monthly basis.

Fare payment systems vary greatly in the Puget Sound and Los Angeles regions, affecting the final outputs of the data. While the use of regional transit fare cards is well established in both, the MOD payment methods in Puget Sound include the ORCA card, Transit Go Ticket, and credit/debit card. However, the Los Angeles region does not use their conventional transit payment method, the TAP card, and instead users pay via debit or credit card through the Via app. In both regions, riders can create new accounts and pay over the phone for call center requests. This allows for additional travel information tied to individual users in the Puget Sound Region from transit agency fare payment data. The disparity in fare payment methods between the two regions in fare payment data results in an ability for more detailed person level analysis in the Puget Sound region, but all project goals and research questions are addressable in LA County as well in part by supplementing Via and transit agency data with survey and census data. Due to the differences in fare integration, research questions related to mode shift and demographics can only be answered in Los Angeles through survey data, while the Puget Sound region has more comprehensive data from the ORCA cards.

3.5 Staff and Data Expertise

LA Metro, Sound Transit, and King County Metro are all large agencies with substantial expertise amongst agency staff. However, the MOD Sandbox application included participation of a research team to provide outside expertise. Outsourcing data science and analysis allowed the transit agencies to focus on project management and operations as well as research elements while maintaining their data and analysis needs. While including more organizations added complexities, building a research team of institutions and individuals with years of experience in working with sensitive transportation data and previous experience with transit

and MOD provider data helped with the development of a data sharing agreement and conducting valuable analyses.

The research team participated in early discussions led by the transit agencies on goal identification, related performance areas, and specific measures needed to assess those performance areas. Then, the project partners identified data needed to track those measures. Discussions with Via occasionally included all project partners such that the FTA, all three transit agencies, and the Eno/university research team could all assess measurable outcomes related to their goals.

Incorporating researchers with high levels of expertise and experience with transportation data from the inception of the project enabled the research team to provide input and shape the project to facilitate the best experimental design possible. The research team was able to help shape the final data sharing agreement and coordinate survey development with existing transit data. The team also ensured Via that their public university institutions handle sensitive data all the time, and there were measures in place to protect public records requests from opening up sensitive data. Without the experience and persistence of the research team, the final data sharing agreement might have been sub-optimal and a robust ridership and equity analysis would not have been possible.

3.6 Privacy and Cybersecurity

During negotiations with both Lyft and Via, LA Metro maintained that they could access and store data that contained a potential trade secret for the MOD provider and were able to protect that trade secret in the possibility of public records requests. In the Puget Sound Region, some publicly-held data is protected by way of a Washington State statute that protects fare media, including location, time, and other elements of trip and ridership data.

The supplemental survey data is a result of collaboration between the research team and the transit agencies to conduct ridership intercept surveys as well as with Via to conduct online surveys to Via users. The intercept survey data are collected, owned, and stored by the transit agencies, and shared with the research team as requested by sending downloadable files without identifying information of respondents. The online survey data collected are owned by the Eno Center for Transportation and shared with the research team by direct access to the survey platform.

3.7 Laws and Regulations

The MOD Sandbox Program was announced in 2016 using 49 U.S.C. 5312, Public Transportation Innovation funds. While the request for proposals did not specify data reporting or performance measurement requirements, it did note that projects would be awarded based on multiple criteria including support for data collection and evaluation.[31] Many projects, including those in the Los Angeles and Puget

Sound regions, developed their own performance measures with unique data needs for project evaluation. The eleven MOD Sandbox projects are evaluated on common performance measures developed by an external evaluator who is overseen by the FTA, thus requiring specific data needs as requested by the external evaluators.

Beyond the federal legal framework, the project team had to work within the requirements of each state, locality, and agency policies. LA Metro, Sound Transit, and King County Metro all engaged their in-house legal teams to review existing legislation at the local and state levels. The agencies in Washington took into account a recent state supreme court case preempting the protection of trade secret with public records release requirements by connecting all data to fare data which is protected under state law.[32] In California, privacy laws and trade secret protection laws allowed for the organizations to share data as requested. California's consumer privacy laws allow for sharing of de-identified or aggregated data, as already underway in the project.[33]

4. Recommendations

Data sharing agreements are essential to successful MOD pilots to assess, evaluate, and inform future decision-making. Goals, laws, and context will vary from project to project, but the lessons from the LA Metro and Puget Sound MOD project have broad applicability. The following recommendations provide guidance to FTA, private contractors, research institutions, and agency staff when developing data sharing agreements.

Relate data requests directly to a measure that will help assess or evaluate the project in relation to a clear goal. This will help ensure that every data request is useful and the time put into negotiating and developing the data sharing agreement as well as collecting, storing, and sharing the data is all worthwhile. Agencies should be firm about what data they are requesting and why. Targeting data requests also prevents asking for data for the sake of data.

Set data ownership, access, and sharing policies up front for every data set. These parameters do not have to be the same for all data sets. Allowing for flexibility to provide data only as needed can have a positive impact on relationships by not "over asking" for data. This also positively impacts data storage, access, and analysis by minimizing the size of files and excess information.

Agree in writing to preliminary arrangements such as a Term Sheet early in the process. A Term Sheet lays out mutually agreed basic expectations for data ownership and access before spending time on further negotiations. Coming to interim agreements around difficult topics such as data sharing can help clarify expectations and serve as an outline for detailed agreements later on.

Engage with research experts when deploying a MOD pilot. A MOD pilot is only successful if an agency is able to learn about the service and whether or how it can be implemented in long term service plans or scaled to additional times or locations. Additional data science capacity for pilot projects can be outsourced to expert researchers who have expertise in data collection, access, and analysis. Research institutions such as universities have robust privacy protections in place and individuals have years of experience in responsibly using personal data for analysis.

Understand state, local, and agency legal constraints. Differing agency policies and state and local laws can prompt the need for differing data sharing policies. Protecting against public records requests in California and Washington led to the need to assure Via that the trade secrets would still be protected under a separate law protected fare data. Other states might have different contexts and legal teams at agencies should be part of the process.

Examine all possible data sources. MOD data is hardly the only type of data available to agencies. Administering surveys and matching MOD data with agency and external information can provide unique insights. An important consideration, however, is that surveys are expensive and require significant time to implement correctly.

5. Conclusion

Mobility on Demand partnerships between public agencies and private providers represent a bold experiment in enhancing certain aspects of public transit services. Data sharing agreements and publication of final results are critical for the transit industry to understand if, when, and how to implement these services as part of standard operating plans. Agencies can take careful steps to ensure that they receive robust and useful data while still protecting the privacy of riders and corporate partners.

The Puget Sound and Los Angeles regions provide valuable insights into how to prepare for, write, and execute a successful data sharing agreement. While many aspects of those projects were unique, agencies around the country are deploying similar MOD pilots and services and can draw from the lessons learned in this pilot.

Appendix A

Fiscal Year 2016 Mobility on Demand (MOD) Sandbox Program Projects

Project Sponsor	Description	Funding
Regional Transportation Authority of Pima County, Arizona	The Adaptive Mobility with Reliability and Efficiency project, integrating fixed route, subscription based ride-sharing and social carpooling services into an existing data platform to provide affordable, convenient and flexible service. The project augments transit by addressing first mile/ last mile issues and congestion mitigation by incorporating shared ride-on-demand services, integrated open payment systems and advanced traveler information systems.	$669,158
Valley Metro Rail, Inc., Phoenix	A smart phone mobility platform that integrates mobile ticketing and multimodal trip planning. The network will include a range of mobility providers, including ride-hailing, bike sharing, and car-sharing companies, allowing all levels of income, age and people with disabilities to have access to an integrated, connected multimodal transportation system.	$1,001,000
City of Palo Alto, California	The Bay Area Fair Value Commuting Demonstration project, which aims to reduce single-occupant vehicle driving from 75% to 50% in the Bay Area. The project includes commuter trip reduction software, a mobility aggregation multimodal trip planning app, workplace parking rebates and analytics to compare commutes.	$1,085,000
Los Angeles County Metropolitan Transportation Authority	A two-region mobility on demand partnership with the car-sharing company, Lyft*, in Los Angeles and Puget Sound. The project will explore the viability of first/last mile solutions for trips originating and ending at select transit stops. Customers can use the Lyft* app or call a dispatcher phone number, providing equity to lower income individuals. (*Partnership changed from Lyft to Via since announcement.)	$1,350,000
San Francisco Bay Area Rapid Transit	An integrated carpool to transit program that will help users find carpool matches as well as match them to their transit destinations. The project will provide a seamless way to reserve and pay for in-demand parking spaces at BART stations, allow preferential parking for carpoolers while increasing transit ridership by improving access to BART stations. The software will include ways to identify drivers with wheelchair-accessible vehicles.	$358,000

Project Sponsor	Description	Funding
Pinellas Suncoast Transit Authority, Florida	For the Paratransit Mobility on Demand Demonstration, a set of partnerships with a taxi company, a paratransit service and a car-sharing company to develop a model to provide more cost-effective on-demand door-to-door paratransit service. The project will feature a central dispatch software that provides users with a selection of transportation service providers based on an estimated time of pickup, available payment types, and physical limitations.	$500,000
Chicago Transit Authority	A project that will incorporate the local bike sharing company, Divvy, a 580-station bike share service, into CTA's existing transit trip planning app so users can identify the availability of bikes or docking stations near their transit stops, and pay for bike rentals.	$400,000
Tri-County Metropolitan Transportation District, Oregon	An Open Trip Planner Share Use Mobility project that will create a platform integrating transit and shared-use mobility options. TriMet will build on its existing trip planning app to incorporate shared use mobility options and more sophisticated functionality and interfaces, including data sharing for shared-use mobility providers. By integrating data, the project will allow users to plan trips that address first/last mile issues while traveling by transit.	$678,000
Dallas Area Rapid Transit	A project that integrates ride-sharing services into its GoPass ticketing app to solve first and last mile issues. This project will combine traveler applications to create an integrated, multimodal application that leverages ride-sharing services. The project will improve ease of access to DART stations, particularly in non-walkable areas not well served by transit.	$1,204,000
Vermont Agency of Transportation	A statewide transit trip planner that will enable flex-route, hail-a-ride, and other non-fixed-route services to be incorporated in mobility apps. The online trip planner for both fixed and flexible transit services particularly benefits non-traditional rural transit system users, allowing universal access to transit information, including to people with disabilities.	$480,000
Pierce County Public Transportation Benefit Area Corporation	The Limited Access Connections project, an initiative connecting Pierce Transit local service, Sound Transit/Sounder regional service, and local ride-share companies in order to increase regional transit use. By providing first/last mile service in and between traditional zones, guaranteed rides home, and rides to park-and-ride lots, the project will extend service hours and provide access to transit for riders who have limited transit options.	$205,922

Source: Federal Transit Administration

Appendix B

Table 1: Data and Rationale in Term Sheet for Data Via Agreed to Provide Transit Agencies

Data Type	Reason
Unique passenger ID	1) Allows researchers to determine connections people make to transit. Will connect passenger ID to TAP/ORCA ID. Will also allow researchers to figure out transfer time and total travel time. 2) Enables researchers to differentiate between one person taking ten trips or ten people taking one trip.
Requested pick-up location, time, and day of week	Pick-up location 1) Allows for selection of Via trips that begin/end around transit stations 2) Allow calculation of distances traveled between origin/destination Time of day and day of week 1) Allows for identification of travel during peak hours 2) Allows for identification of if people are using Via when transit has stopped running or runs very infrequently
Requested drop-off location, time, and day of week	
Travel Time	1) Allows for evaluation of level of service in comparison to other mobility options 2) Enables evaluation of GHG emissions; If the project demonstrates a positive environmental impact, it will have a higher likelihood of receiving continued funding.
Speed	
Vehicle Type	
Trip Cost: full cost and subsidized amount	1) Shows what people were spending on trips before and after the program 2) Connects to overall trip costs when considering transit connections 3) Informs partnership business model
Trips that are shared by two or more passengers with unique locations	Allows for understanding of the impact on SOV travel, VMT, and congestion
Non-Revenue miles traveled	Allows for understanding of the impact on SOV travel, VMT, and congestion
Vehicle dwell time	Enables calculation of VMT
Number, date, type (WAV), and time of fulfilled rides	1) Shows regulatory compliance 2) Allows for auditing WAV rides which will be paid for with a higher subsidy 3) Gives understanding of the effectiveness of this model to serve the needs of transit customers and identify ways to improve it
Number, date, type (WAV), and time of unfulfilled rides	
Number, date, type (WAV), and time of declined rides	
Number, date, type (WAV), and time of canceled rides	
Average wait time, non-WAVs	1) Shows regulatory compliance 2) Allows for auditing WAV rides which will be paid for with a higher subsidy
Average wait time, WAVs	

Data Type	Reason
Number of times that a wheelchair or scooter user asked to get out of their chair for transport	3) Gives understanding of the effectiveness of this model to serve the needs of transit customers and identify ways to improve it
Number of times that a wheelchair or scooter transported in their own mobility device (not asked to transfer)	
Drivers working who have received ADA sensitivity/wheelchair securement training	
Make and model of vehicles providing service/number of people in each vehicle	To evaluate energy use and fuel consumption
User rating of experience and comments	To understand the customer experience and build political support
Rider Contact Information *(Puget Sound only)*	To enable survey deployment for additional data collection

Source: Term Sheets between LA Metro and Via for Metro's MOD Partnership with Via as well as between King County Metro and Via for Via to Transit

Table 2: Data Sharing Agreement in Final Contracts – Trip Level Data

Data Field	Response Type	Description	Actual Data Received (LA)	Actual Data Received (PS)	Mismatch?
Passenger ID *	3) ID number	A unique passenger identification number that contains no personally identifiable information	Passenger ID	Passenger ID	N/A
Vehicle make, model, year *	Categorical	Information of vehicle utilized to transport Passenger	Vehicle make, model	Vehicle make, model	unknown
TAP/ORCA ID *	ID number	ORCA transit card number	N/A	ORCA ID	Absence of TAP payment integration
Zone ID	Categorical	Which of the catchment areas did the ride originate in	Zone ID	Zone ID	N/A
N/A	N/A	Name of station the ride originated or ended in	Station ID	N/A	Unknown
Request pick-up location latitude*	Numerical (3 digits after decimal point)	Latitude of requested pick-up location	Request pick-up location latitude	Request pick-up location latitude	N/A
Request pick-up location longitude*	Numerical (3 digits after decimal point)	Longitude of requested pick-up location	Request pick-up location longitude	Request pick-up location longitude	N/A
Request drop-off location latitude*	Numerical (3 digits after decimal point)	Latitude of requested drop-off location	Request drop-off location latitude	Request drop-off location latitude	N/A
Request drop-off location longitude*	Numerical (3 digits after decimal point)	Latitude of requested drop-off location	Request drop-off location longitude	Request drop-off location longitude	N/A
Request pick-up date/time*	Numerical (YYYY-MM-DD HH:MM)	The time stamp that the request is made	Request pick-up date/time	Request pick-up date/time	N/A
Estimated response time communicated to passenger	Numerical (HH:MM)	The estimated response time communicated to the passenger after the Driver Partner is dispatched	Estimated response time communicated to passenger	Estimated response time communicated to passenger	N/A
Actual wait time to passenger before pick-up*	Numerical (HH:MM)	This is the actual amount of time the passenger spent waiting to be picked up by Driver Partner	Actual wait time to passenger before pick-up	Actual wait time to passenger before pick-up	More significant figures than requested
Actual pick-up date and time*	Numerical (HH:MM)	The time stamp when the trip starts with the passenger	Actual pick-up date and time	Actual pick-up date and time	More significant figures than requested

Data Field	Response Type	Description	Actual Data Received (LA)	Actual Data Received (PS)	Mismatch?
Actual drop-off date and time*	Numerical (HH:MM)	The time stamp when the trip ends with the passenger	Actual drop-off date and time	Actual drop-off date and time	More significant figures than requested
Origin to destination distance	Numerical (Miles)	Actual distance of travel of the vehicle in order to deliver passenger from origin to destination	Origin to destination distance	Origin to destination distance	More significant figures than requested
Average travel speed of ride*	Numerical (Miles per hour)	Average miles per hour travel speed during the passenger's trip	Average travel speed of ride	Average travel speed of ride	More significant figures than requested
Trip cost charged to paying passenger*	Numerical (Cents)	Total cost of the trip charged to the paying passenger	Trip cost charged to paying passenger	Trip cost charged to paying passenger	N/A
Number of guests with requesting passenger (if any)	Numerical (people)	Number of guests per passenger ID (if any)	Number of guests with requesting passenger (if any)	Number of guests with requesting passenger (if any)	N/A
Accessible Vehicle ride requested *	Binary	Indicates Yes or No on whether the passenger requested an accessible vehicle	Accessible Vehicle ride requested	Accessible Vehicle ride requested	N/A
Accessible Vehicle ride provided *	Binary	Indicate Yes or No on whether Contractor provided an accessible vehicle ride	Accessible Vehicle ride provided	Accessible Vehicle ride provided	N/A
Trip outcome *	Categorical (completed, rider cancelled, driver cancelled, no show, other)	Indicate whether the trip was completed, rider cancelled, Driver Partner cancelled, or the passenger was a no-show	Trip outcome	Trip outcome	N/A
Trip Cancellation or no-show timestamp. Rounded to the nearest minute (YYYY-MM-DD HH:MM) *	Categorical (date)/Numerical (time)	for cancelled or no-show trips	Trip Cancellation or no-show timestamp. Rounded to the nearest minute	Trip Cancellation or no-show timestamp. Rounded to the nearest minute	N/A

Source: Final Contracts between LA Metro and Via for Metro's MOD Partnership with Via and King County Metro and Via for Via to Transit
** = in Term Sheet*

Table 3: Data Sharing Agreement in Final Contracts – App and Call Center Data (aggregated weekly)

Data Field	Response Type	Description	Actual Data Received (LA)	Actual Data Received (PS)	Mismatch?
New Via accounts created	4) Numerical (Number of new Via accounts registered)	Indicates the number of new Via accounts registered in the Los Angeles or Puget Sound area per week	New Via accounts created	New Via accounts created	N/A
Number of customer service inquiries submitted to Contractor through app	Numerical (number of customer service inquires submitted to Via)	Indicates the number of customer service inquires submitted to Contractor per week	Number of customer service inquiries submitted to Contractor through app	Number of customer service inquiries submitted to Contractor through app	N/A
Ride request source (App)	Numerical (number of rides requested through the app each week)	Indicates the number of trip requests made using Via's app	Number of rides requested through the app	Number of rides requested through the app	N/A
Ride request source (Call Center)	Numerical (number of rides requested through the call center each week)	Indicates the number of trip requests made using the call center	Number of rides requested through the call center	Number of rides requested through the call center	N/A
Frequency of proposals not booked	Numerical	Gives the number of times a requested ride proposal is not booked	Frequency of proposals not booked	Frequency of proposals not booked	N/A
Number of calls received by call center	Numerical	Indicates the number of customer service inquires submitted to Contractor per week	Number of calls received by call center	Number of calls received by call center	N/A
Number of rides dispatched through call center	Numerical	Indicate the number of rides dispatched through the call center on a weekly basis	Number of rides dispatched through call center	Number of rides dispatched through call center	N/A

Source: Final Contracts between LA Metro and Via for Metro's MOD Partnership with Via and King County Metro and Via for Via to Transit
** = in Term Sheet*

Table 4: Data Sharing Agreement in Final Contracts – Vehicle Data (aggregated weekly)

Data Field	Response Type	Description	Actual Data Received (LA)	Actual Data Received (PS)	Mismatch?
Vehicle make, model, year*	Categorical	Information of vehicle utilized to transport Passenger	Vehicle make, model	Vehicle make, model	unknown
Date/time of beginning of shift	Numerical (YYYY-MM-DD HH:MM)	Date and time of the start of the shift for the vehicle	Date/time of beginning of shift	Date/time of beginning of shift	N/A
Date/time of end of shift	Numerical (YYYY-MM-DD HH:MM)	Date and time of the end of the shift for the vehicle	Date/time of end of shift	Date/time of end of shift	N/A
Non-revenue miles driven while on shift*	Numerical (Miles)	Total amount of miles the vehicle drove without any passengers on board during the shift	Non-revenue miles driven while on shift	Non-revenue miles driven while on shift	More significant figures than requested
Revenue miles driven while on shift	Numerical (Miles)	Total amount of miles the vehicle drove with at least one paying passenger on board during the shift	Revenue miles driven while on shift	Revenue miles driven while on shift	Includes non-paying passengers as many rides are free. More significant figures than requested
Vehicle miles driven with 1, 2, 3, …, 8 bookings on board during that shift	Numerical (Miles)	This is the miles driven by count of passengers in the vehicle at a time	Vehicle miles driven with 1, 2, 3, …, 8 bookings on board during that shift	Vehicle miles driven with 1, 2, 3, …, 8 bookings on board during that shift	More significant figures than requested
PMT:VMT on an hourly basis per vehicle	Numerical (ratio)	Comparison of passenger miles travelled versus vehicle miles travelled on an hourly basis per vehicle during each shift.	PMT:VMT on an hourly basis per vehicle	PMT:VMT on an hourly basis per vehicle	N/A

Source: Final Contracts between LA Metro and Via for Metro's MOD Partnership with Via and King County Metro and Via for Via to Transit
** = in Term Sheet*

Table 5: Data Sharing Agreement in Final Contracts – Ridership (aggregated weekly)

Data Field	Response Type	Description	Actual Data Received (LA)	Actual Data Received (PS)	Mismatch?
New riders	Numerical	Indicate the number of customers who took their first rides on a weekly basis	New riders	New riders	N/A
Unique active riders	Numerical	Active is considered to be someone who has used Pilot service at least once in the week. This does not count multiple trips made by the same rider	Unique active riders	Unique active riders	N/A
Unique repeat riders	Numerical	A repeat rider is considered someone who has used Pilot service more than once in a week.	Unique repeat riders	Unique repeat riders	N/A
Passengers per vehicle per hour	Numerical (ratio)	Indicates the average number of passengers per vehicle per hour	Passengers per vehicle per hour	Passengers per vehicle per hour	N/A

Source: Final Contracts between LA Metro and Via for Metro's MOD Partnership with Via and King County Metro and Via for Via to Transit
** = in Term Sheet*

Table 6: Data Sharing Agreement in Final Contracts – Pick-up drop off by week (trip level)

Data Field	Response Type	Description	Actual Data Received (LA)	Actual Data Received (PS)	Mismatch?
Pick-up or drop-off	categorical	Indicates whether the below latitude/longitude is in reference to a pick-up or drop-off	Pick-up or drop-off	Pick-up or drop-off	N/A
Pick-up or drop-off location latitude	Numerical (3 digits after decimal point)	Indicates the actual pick-up and drop-off locations for the previous week where customers were picked up and dropped off. Origins and destinations do not need to be paired. This data point is to determine whether the pick-ups and drop-offs are at accessible locations	Pick-up or drop-off location latitude	Pick-up or drop-off location latitude	N/A
Pick-up or drop-off location longitude	Numerical (3 digits after decimal point)	Indicate the actual pick-up and drop-off locations for the previous week where customers were picked up and dropped off. Origins and destinations do not need to be paired. This data point is to determine whether the pick-ups and drop-offs are at accessible locations	Pick-up or drop-off location longitude	Pick-up or drop-off location longitude	N/A

Source: Final Contracts between LA Metro and Via for Metro's MOD Partnership with Via and King County Metro and Via for Via to Transit
** = in Term Sheet*

Table 7: Data Sharing Agreement in Final Contracts – Daily Driver Hours

Data Field	Response Type	Description	Actual Data Received (LA)	Actual Data Received (PS)	Mismatch?
Driver hours	Numerical (to 1 billionth of 1 hour decimals)	Total number of hours driven daily	Driver hours	Driver hours (through invoicing process)	In invoicing requirements for LA only to align payment with utilized driver hours.

Source: Final Contracts between LA Metro and Via for Metro's MOD Partnership with Via and King County Metro and Via for Via to Transit
** = in Term Sheet*
Note: Via also provides Puget Sound with a table denoting driver hours per day for easier tracking of labor.

Endnotes

[1] Prashanth Gururaja and Rudy Faust. "Objective-Driven Data Sharing for Transit Agencies in Mobility Partnerships," Shared Use Mobility Center, 2019.
[2] Terra Curtis, Meg Merritt, Carmen Chen, David Perlmutter, Dan Berez, and Buffy Ellis. "Partnerships between Transit Agencies and Transportation Network Companies (TNCs)," TCRP 204. National Academy of Sciences. 2019.
[3] Sara Khoeini. "Modeling Framework for Socio-Economic Analysis of Managed Lanes," Dissertation presented to the Georgia institute of Technology. 2014.
[4] Moshe Ben-Akiva and Steven R. Lehman. Discrete Choice Analysis. MIT Press. Cambridge Massachusetts. 1985.
[5] Gustave Cordahi, Susan Shaheen, and Elliot Martin. "MOD Sandbox Demonstrations Independent Evaluation Dallas Area Rapid Transit – First and Last Mile Solution Evaluation Plan," 2019.
[6] Colin Murphy, Kevin Karner, and Zak Accuardi. "When Uber Replaces the Bus: Learning from the Pinellas Suncoast Transit Authority's Direct Connect Pilot", *Shared Use Mobility Center*, 2019.
[7] Adam Vaccaro. "Highly touted Boston-Uber partnership has not lived up to hype so far," Boston.com, *June 16, 2016*.
[8] Prashanth Gururaja and Rudy Faust. "Objective-Driven Data Sharing for Transit Agencies in Mobility Partnerships," Shared Use Mobility Center, 2019.
[9] Robert George, "The data science skill shortage in the public sector," GovLoop, February 15, 2016.
[10] Adie Tomer and Ranjitha Shivaram. "Modernizing Government's Approach to Transportation and Land Use Data," *The Brookings Institution*, July 20, 2017.
[11] "Private Mobility, Public Interest", *TransitCenter*, September 20, 2016
[12] Alice Grossman and Paul Lewis (2019). Contracting for Mobility: A Case Study in the Los Angeles and Puget Sound Regions. Eno Center for Transportation. Washington D.C., 2019.
[13] ibid.
[14] Yves-Alexandre De Montjoye, César .A. Hidalgo, Michel Verleysen, and Vincent D. Blondel (2013). "Unique In the Crowd: The Privacy Bounds of Human Mobility." *Scientific Reports* vol. 3 (2013):1376. Doi:10.1038/srep01376
[15] Johanna Zmud, Melissa Tooley, Matthew Miller "Data Ownership Issues in a Connected Car Environment: Implications for State and Local Agencies", *Texas A&M Transportation Institute*, November 2016.
[16] Yves-Alexandre De Montjoye, César .A. Hidalgo, Michel Verleysen, and Vincent D. Blondel (2013). "Unique In the Crowd: The Privacy Bounds of Human Mobility." *Scientific Reports* vol. 3 (2013):1376. Doi:10.1038/srep01376
[17] University of Washington Transportation Data Collaborative, Homepage, Seattle, WA, 2019; Shared Streets, Homepage, NACTO
[18] UK Information Commissioner's Office. Overview of the General Data Protection Regulation (GDPR), October 20, 2017
[19] The United States of American Before the Federal Trade Commission In the Matter of BLU PRODUCTS, INC., a corporation; and SAMUEL OHEV-ZION, individually and as owner and President of BLU PRODUCTS, INC. April, 2018.; Paige M. Boshell. "The Power of Place: Geolocation Tracking and Privacy," *Business Law Today,* March 25, 2019
[20] Paige M. Boshell. "The Power of Place: Geolocation Tracking and Privacy," *Business Law Today,* March 25, 2019; Cal. Civ. Code § 1798.140(o)(1)(G) (2018)
[21] Supreme Court of the State of Washington. Lyft, INC. and Raster, LLC, City of Seattle, And Jeff Kirk, Respondents, Appellant, Defendant. May 31, 2018.

[22] Alice Grossman. "Zip Code Trip Data Ruled Publicly Accessible in Washington State," *Eno Transportation Weekly,* Washington D.C., June 8, 2018; The Supreme Court of the State of Washington, Lyft Inc. and Raiser, LLC versus the City of Seattle and Jeff Kirk, Seattle, WA, May 31, 2018.
[23] Prashanth Gururaja and Rudy Faust. "Objective-Driven Data Sharing for Transit Agencies in Mobility Partnerships," Shared Use Mobility Center, 2019.
[24] Government Accountability Office. "Public Transit Partnerships: Additional Information Needed to Clarify Data Reporting and Share Best Practices", *Government Accountability Office,* July 2018
[25] USDOT, Executed Contract with LA Metro, FAIN: CA-2017-018-00, January, 2017.
[26] Alice Grossman and Paul Lewis, "Contracting for Mobility: A Case Study in the Los Angeles and Puget Sound Regions," Eno Center for Transportation, 2019.
[27] ibid.
[28] ibid.
[29] ibid.
[30] UW Transportation Data Collaborative, University of Washington, 2019
[31] FTA, Public Transportation Innovation Funding Opportunity; Mobility on Demand (MOD) Sandbox Demonstration Program, Federal Register Volume 81, Number 85 Docket Number 2016-10320. May, 2016.
[32] The Supreme Court of the State of Washington, Lyft Inc. and Raiser, LLC versus the City of Seattle and Jeff Kirk, Seattle, WA, May 31, 2018.
[33] Squire Patton Boggs. The "Final" California Consumer Privacy Act. October 17, 2019.

Eno
Center for Transportation

www.enotrans.org

@EnoTrans

/EnoCenterForTransportation

bit.ly/EnoLinkedIn

/EnoTransportation

@EnoTrans

1629 K Street, NW
Suite 200
Washington, DC 20006

CONTACT US:
publicaffairs@enotrans.org
202-879-4700