Title
Sources of and Gaps in Data for Understanding Public Transit Ridership

Authors
Wasserman, Jacob L
Taylor, Brian D

Publication Date
2021-01-20

RESEARCH REPORT

Institute of Transportation Studies | UNIVERSITY OF CALIFORNIA

Sources of and Gaps in Data for Understanding Public Transit Ridership

Jacob L. Wasserman, Research Project Manager
UCLA Institute of Transportation Studies
Brian D. Taylor, PhD, FAICP, Professor of Urban Planning and Public Policy, UCLA Luskin School of Public Affairs
Director, UCLA Institute of Transportation Studies

January 2021

UCLA Institute of Transportation Studies

Report No.: UC-ITS-2020-33 | DOI: 10.17610/T66893

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UC-ITS-2020-33	N/A	N/A

4. Title and Subtitle Sources of and Gaps in Data for Understanding Public Transit Ridership	**5. Report Date** January 2021 **6. Performing Organization Code** UCLA ITS
7. Author(s) Jacob L. Wasserman, https://orcid.org/0000-0003-2212-5798 Brian D. Taylor, PhD, FAICP, https://orcid.org/0000-0002-1037-2751	**8. Performing Organization Report No.** N/A
9. Performing Organization Name and Address 337 Charles E. Young Drive East UCLA Institute of Transportation Studies Public Affairs Building 3320 Los Angeles, CA 90095-1656	**10. Work Unit No.** N/A **11. Contract or Grant No.** UC-ITS-2020-33
12. Sponsoring Agency Name and Address The University of California Institute of Transportation Studies www.ucits.org	**13. Type of Report and Period Covered** Final Report (July 2019 - January 2021) **14. Sponsoring Agency Code** UC ITS

15. Supplementary Notes

DOI: 10.17610/T66893

16. Abstract

This report presents and reviews the available sources of data on public transit riders and ridership. We intend it to be a resource for those who manage or simply wish to understand U.S. transit. In conducting this review, we consider the advantages and disadvantages of publicly available data on transit from a variety of public and private sources. We consider as well the relatively scarcer and less available sources of data on other providers of shared mobility, like ride-hail services, that compete with and complement public transit, as well as pieces we see as missing from the transit analytics pie. We conclude by discussing how data gaps both align with existing inequities and enable them to continue, unmeasured, and how the COVID-19 pandemic has made closing these gaps all the more important.

17. Key Words	18. Distribution Statement
transit, ridership, data, data gaps	no restrictions

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	21. Price
unclassified	unclassified	42	N/A

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

About the UC Institute of Transportation Studies

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

Acknowledgments

This study was made possible through California Statewide Transportation Research Program funding received by the University of California Institute of Transportation Studies from the State of California through the Public Transportation Account and the Road Repair and Accountability Act of 2017 (Senate Bill 1). The authors would like to thank the State of California for its support of university-based research and especially for the funding received for this project.

The authors would also like to thank John Gahbauer, Andrew Schouten, Eric Dasmalchi, and Hannah King at the UCLA Institute of Transportation Studies and Elizabeth Sall (UrbanLabs LLC), Ruth Miller (Caltrans), Amy Smith (Rebel Group), Vyki Englert (Compiler LA), and Gillian Gillett (Caltrans) for their insightful comments and suggestions on earlier drafts of this report.

The Institute of Transportation Studies at UCLA acknowledges the Gabrielino/Tongva peoples as the traditional land caretakers of Tovaangar (the Los Angeles basin and So. Channel Islands). As a land grant institution, we pay our respects to the Honuukvetam (Ancestors), ‘Ahiihirom (Elders) and ‘Eyoohiinkem (our relatives/relations) past, present and emerging.

Disclaimer

Institute of Transportation Studies | UNIVERSITY OF CALIFORNIA

Sources of and Gaps in Data for Understanding Public Transit Ridership

Jacob L. Wasserman, Research Project Manager
UCLA Institute of Transportation Studies
Brian D. Taylor, PhD, FAICP, Professor of Urban Planning and Public Policy, UCLA Luskin School of Public Affairs
Director, UCLA Institute of Transportation Studies

January 2021

UCLA Institute of Transportation Studies

Report No.: UC-ITS-2020-33 | DOI: 10.17610/T66893

Table of Contents

Table of Contents

Introduction **1**

Overview 1

Data on Public Shared Mobility **5**

Transit Ridership Data 5

Transit Rider Data 8

Data on Transit Service Quality 12

The General Transit Feed Specification 13

Data on Private Shared Mobility **15**

Ride-hail Data 15

Private Shuttle and Micromobility Data 17

Conclusion **19**

Data Gaps 19

Transit Data Needs Looking Forward 21

References **23**

List of Tables

Table 1. Major Datasets on Transit Ridership and on the Factors behind It.. 2

Table 2. Significant Gaps in Transit Ridership Data.. 19

Contents

Introduction

This report presents and reviews the available sources of data on public transit riders and ridership. We intend it to be a resource for those—analysts, planners, journalists, students, and more—who manage or simply wish to understand the buses, streetcars, elevateds, and subways that ply the streets and railbeds of U.S. cities each day. In conducting this review, we consider the advantages and disadvantages of publicly available data on transit from a variety of public and private sources. We consider as well data on other sources of shared mobility, like ride-hail services, that compete with and complement public transit, as well as pieces we see as missing from the transit analytics pie.

Overview

Public transit in the United States is ailing. Before the novel coronavirus (severe acute respiratory syndrome coronavirus 2 or SARS-CoV-2, the virus that causes Coronavirus Disease 2019 or COVID-19) pandemic, transit ridership fell every year between 2014 and 2019. Over that period, the country lost more than 800 million annual transit trips in total, a 7.5 percent drop (FTA, 2020b). Whereas transit use fell from 2009 to 2011 as employment and overall travel also dropped in the wake of the Great Recession, the transit ridership decline in the latter half of the 2010s occurred at a time of economic growth (FTA, 2020b and World Bank, 2020). These losses, though, pale in comparison to the staggering drops that have occurred since the onset of the COVID-19 pandemic in the spring of 2020: a plummet of 50 to as much as 94 percent over the course of just a few weeks (BTS, 2020; Transit App, 2020; Moovit, 2020; Levy and Goldwyn, 2020; and Walker, 2020). Even before the pandemic, the changes in transit ridership were uneven, varying across metropolitan areas, built environments, times of day, days of the week, trip purposes, operators, modes, and directions. Declines in ridership since the onset of the pandemic are similarly varied; buses lost fewer passengers than rail (and especially commuter rail) transit, and areas with the highest levels of pre-pandemic transit use lost more riders than areas with fewer riders. That buses and entire systems avoided the most severe losses was almost certainly not due not to better service planning but rather to their higher share of transit-dependent riders, who have fewer options but to ride during a pandemic (Timmons, 2020).

High-quality, accessible, up-to-date data are the tools with which practitioners and researchers can diagnose the causes of America’s transit ridership woes and evaluate and recommend possible cures. While the reasons behind the ridership drop since March 2020 are far more conspicuous than the causes of the preceding more gradual decline, having detailed transit data, disaggregated across a number of axes, is more important than ever to the recovery of the transit industry and the mobility of those who rely on it. Moreover, data about transit use can answer pressing questions beyond patronage declines, including analyses of transportation equity, evaluations of proposed capital and operating improvements, inquiries into the effects of private shared mobility services, and projections of emissions and pollution, among others. All of these topics rely on a growing, though still incomplete and often incompatible, set of transit data sources collected in different ways, from different sources, on different timeframes.

This white paper identifies and summarizes both the most relevant data on public transit ridership (and the internal and external factors that influence ridership) and the most significant gaps in those data. Throughout this report we consider 1) *data* relevant to public transit use (such as unlinked passenger trips), 2) the *sources*—or collectors, aggregators, and disseminators—of those data (such as the Federal Transit Administration), and 3) the range of *uses* for these data (such as analyses of transit ridership by race or income level). We focus more on the 1) data and 2) sources in this report and only occasionally reference 3) uses for the data, due primarily to scope and resource limitations. With respect to uses, the data reviewed can be used for a wide array of analyses, with respect to both the units of analysis (regions, systems, modes, riders, etc.) and time periods (at one point in time or over time), and can be used to evaluate efficiency (inputs to outputs), effectiveness (inputs or outputs to consumption), many dimensions of equity, and other aspects of performance as well.

We first review the major data sources on transit ridership itself and related operator characteristics like service supply and expenditures. Next, we discuss data on individual and household travel patterns that can be used for transit ridership research on the demand side. In addition, we review public information on further aspects of transit service: its quality and its location and schedules. We then turn to the much less accessible and systematized universe of data on private shared mobility, including ride-hail, corporate shuttles, and micromobility. We conclude with a synthesis of the widest open gaps in transit data and recommendations for addressing them. **Table 1** summarizes the most significant sources of data discussed in this white paper.

Table 1. Major Datasets on Transit Ridership and on the Factors behind It

Dataset	Data Source	Citation	Data Included	Web Address
National Transit Database	Transit operator reports to FTA, with ridership figures estimated using a variety of approved methods	FTA, 2020b	Ridership, service, finances, safety, labor, assets	transit.dot.gov/ntd
Public Transportation Fact Book	NTD data supplemented by APTA reporting	APTA, 2020a	Ridership expanded from the NTD, historical ridership	apta.com/research-technical-resources/transit-statistics
National Household Travel Survey	Semi-regular FHWA survey of U.S. households	FHWA, 2017	Household travel diary, socio-economic characteristics of travelers	nhts.ornl.gov
California Household Travel Survey	Caltrans survey of California households	Caltrans, 2012	Household travel diary, socio-economic characteristics of travelers	nrel.gov/transportation/secure-transportation-data/tsdc-california-travel-survey.html
American Community Survey data tables	The ACS, a rolling U.S. Census Bureau survey of U.S. residents	U.S. Census Bureau, 2020a	Commute characteristics, socio-economic characteristics of commuters	data.census.gov

Dataset	Data Source	Citation	Data Included	Web Address
Public Use Microdata Sample	ACS data	Ruggles et al., 2020	Commute characteristics, socio-economic characteristics of commuters	usa.ipums.org
Census Transportation Planning Products	ACS data	AASHTO, 2019	Commute flows, socio-economic characteristics of commuters	ctpp.transportation.org
LEHD Origin-Destination Employment Statistics	State employment records	U.S. Census Bureau, 2020b, 2020c	Job and residence locations, employment characteristics	onthemap.ces.census.gov lehd.ces.census.gov/data
General Transit Feed Specification data	Transit operator GTFS feeds	MobilityData, n.d.	Geographic route, stop, and schedule information	Repositories: transit.land transitfeeds.com transitwiki.org/TransitWiki/index.php/Publicly-accessible_public_transportation_data bts.gov/national-transit-map/national-transit-map-data-maps-and-apps
Nonemployer Statistics	Federal business tax records	U.S. Census Bureau, 2020d	Counts of ride-hail "establishments" (i.e., drivers)	census.gov/programs-surveys/nonemployer-statistics.html

This report focuses on established sources of data on transit use and related information collected by operators, other government entities, and other mobility providers through surveys, counts, and censuses. The growing number of "big data" sources on individual mobility gleaned from people's smartphones, GPS locations, and Internet use are beyond the scope of this paper; they can certainly provide insight and address some of the gaps in knowledge we identify. Likewise, while we discuss data on a number of external factors that influence transit, understanding transit use requires understanding land use policy, housing plans, parking requirements, road pricing, etc. (especially the dimensions of the use and management of the private automobile). We focus here on factors internal to shared mobility providers, public and private, and a salient albeit limited selection of factors beyond that influence transit use. Finally, we limit our analysis to sources of data on transit in the U.S.

Across these data sources, we identify a few major areas for improvement, in terms of better data collection, intercompatibility, centralization, and dissemination (See **Table 2** in the conclusion). Project scope and resource limitations do not permit us to systematically evaluate the range of issues and questions that can and cannot be addressed by these various data sources, though we do allude to these a few of these issues and questions in the form of examples. While we note a number of gaps in data on public transit and the external factors that influence its use, the most salient gap is data on private shared mobility. Ride-hail, micromobility, and other private mobility services have transformed the mobility landscape over the past decade and have likely contributed to declines in pre-pandemic transit use in many areas and have perhaps added riders in a few places. However, the scarcity of publicly available datasets on the use and growth of private shared mobility constitutes a significant gap in data about shared mobility in most places.

Data on Public Shared Mobility

Public shared mobility services are varied and growing, and data on them are not always easy to find. Public transit—which largely consists of buses and trains that mostly operate on fixed routes with fixed schedules, open to all riders that pay a nominal fare to ride—is the most widely known form of shared mobility in urban transportation. But these government-administered services are not the only form. There are a variety of public, quasi-public, and private van and vehicle services that carry multiple passengers. Some of these services, like airport shuttle services, may be available to all willing riders, but others are restricted to participants in social service programs or those with disabilities that require specialized transportation services mandated by the Americans with Disabilities Act. Some bus services may be entirely privately operated, under contract with a public entity or operated by a company for its employees. Since about 2009, a wide variety of app-based private transportation services have sprung on the scene, some of which have proven popular and grown at extraordinary rates. These range from private car services, like Lyft and Uber, offering door-to-door service for individuals or groups of travelers, to shared bicycle and scooter services designed to serve short trips.

This section considers data on shared mobility services operated directly by or under contract to public agencies. Most of these data are on the various public transit and paratransit modes, covering their use, supply, user characteristics, and users' perceptions. Afterwards, the next section considers data on private, mostly for-profit providers of shared mobility.

Transit Ridership Data

Established in 1974, the National Transit Database (NTD), managed by the Federal Transit Administration (FTA) of the U.S. Department of Transportation, aggregates a wide array of transit expenditure, service, and ridership data into a single data source. The NTD is a valuable resource for public transit analyses and research and contains data reported by operators—collected via a variety methods discussed below—on transit use (trips, passenger miles, etc.), service (vehicle revenue hours, vehicle revenue miles, route miles, vehicles in maximum service, etc.), finances (operating expenses and expenditures, capital expenses and expenditures, sources of revenue including fares, etc.), safety (collisions, injuries, etc.), labor (employees, wages, etc.), assets (vehicles, rights-of-way, stations, facilities, etc.), and more (FTA, 2020b).

Any recipient of federal Urbanized Area Formula Program (§5307) funds, any recipient of federal Other than Urbanized Area (Rural) Formula Program (§5311) funds, and any recipient of other Chapter 53 funds that owns, operates, or manages capital transit assets must report data to the NTD at least annually (FTA, 2020a, 2020b); the NTD also accepts voluntary reports from operators not receiving these funds (FTA, 2019). The NTD covers around 2,200 transit operators—essentially all American public agencies that operate transit, including those in all fifty states, D.C., and all populated U.S. territories—and another 900 or so companies and other entities with which those operators contract to provide transit service (FTA, 2020b). While another nearly 4,600 American non-profits operate paratransit services and do not report directly to the NTD,[1] the NTD contains data on over 99.5% of public transit trips, according to estimates from the American Public Transportation Association (APTA),

1. Though some may have their trips reported to the NTD indirectly, by public transit agencies with which they contract

including more than three-quarters of demand-response trips and nearly all other transit trips (Neff and Dickens, 2017 and FTA, 2020b).

APTA publishes its own public transit dataset, the *Public Transportation Fact Book*, which draws heavily on the NTD and supplements it slightly with reports and extrapolations from its member agencies. APTA's *Fact Book* also includes historical data on transit ridership from as early as 1890 (APTA, 2020a, 2020b). For analyses of transit prior to the mid-1990s—the earliest years of online NTD data—APTA's *Fact Book* is a valuable resource. For analyses on more recent transit trends, the two data sources are largely identical. The NTD covers a wider range of topics and has a more clearly defined universe of reporting operators; in our experience, the NTD is more suited for detailed analyses and APTA's *Fact Book* for quicker reference statistics, though both can readily serve both purposes. APTA also collects and publishes databases beyond the NTD on fares and fare products; infrastructure, rights-of-way, and real-time data systems; wage rates; and vehicle details, each based on voluntary survey responses from members (APTA, 2020c).

One major caveat in the annual time series data in the NTD (and APTA's *Fact Book* data drawn from it) concerns how years are reported. Both data sources appear, at first glance, to break down data by calendar year (FTA, 2020b and APTA, 2020a), and that is how many reports and public uses of the data treat them. For instance, the American Society of Civil Engineers (ASCE)'s *Infrastructure Report Card* uses the NTD annual time series figure for total trips in 2015 (ASCE, 2017). However, the years in the NTD and APTA data are not actually calendar years; they are instead "report years" or "reporting years." Report Year 2018—most often labeled simply as "2018" in NTD and APTA datasets—is in fact the aggregate of each agency's fiscal year ending in 2018 (FTA, 2020a, 2020b; APTA, 2020a; and Chu, 2010). Across operators in the NTD, there were ten different fiscal year end dates in Report Year 2018.[2] For almost six in ten operators, the data given for Report Year 2018 cover a timespan more than half in calendar year 2017 (FTA, 2020b).

Thus, the annual national totals in each database aggregate different periods of time for different agencies, complicating comparisons between operators. For example, NTD totals for Report Year 2020 will combine some operator fiscal years that ended before the COVID-19 pandemic or just after its onset and other operator fiscal years that occurred almost entirely during the pandemic. Given the pandemic's significant negative effects on transit use, this truly makes for an apples-and-oranges comparison. To be sure, both datasets include an explanation of "report years" in their data documentation (FTA, 2020a, 2020b and APTA, 2020a), but we suggest clearer labeling throughout the data. If possible, the NTD could also require all agencies to report calendar year or standardized fiscal year (July 1st to June 30th) totals instead of or in addition to their particular reporting year totals. For more precise annual calendar or fiscal year data, we recommend summing numbers from the NTD's monthly reports (or APTA's quarterly reports) instead. However, small[3] and rural agencies are only required to report annually; NTD monthly reports therefore only include an estimate of all small operators combined and of all rural operators combined nationally (FTA, 2020a, 2020b).

Another gap in temporally disaggregated transit data is by day of the week and time of day. In our research on transit ridership trends in California, we found (in analyzing non-NTD data) significant variation in ridership at different times and days across agencies and regions. In Greater Los Angeles, ridership declines in the 2010s

---

2. 58% of operators in Report Year 2018 ended their fiscal year on June 30, 2018, with September 30, 2018 the second-most common fiscal year end date (21% of operators) and December 31, 2018 the third-most common (20% of operators) (FTA, 2020b).

3. Operators with 30 or fewer vehicles that do not operate fixed guideway modes nor bus rapid transit (FTA, 2020a).

were distributed relatively evenly across weekdays, while in the Bay Area—especially on its commute-oriented systems—ridership losses are heavily concentrated on weekends and at off-peak times, a divergence that suggests different underlying factors at play (Taylor et al., 2020; Blumenberg et al., 2020; Wasserman, 2019; and Wasserman et al., 2020). We obtained the data to do this analysis directly from the operators (in variously compatible formats) because data broken down by time of day are not available, to our knowledge, from a central source. The NTD does include service and ridership estimates on an average weekday, Saturday, and Sunday from “full reporter” operators. However, neither small nor rural agencies are required to provide such estimates, nor are these figures included in the main annual time series spreadsheets or monthly data releases. The NTD also asks their full reporter operators for the start and end times of their A.M. peak, P.M. peak, and midday services but does not ask for ridership or service supply data for those time periods (FTA, 2020b). As operators generally collect temporally disaggregated data already, we suggest that the FTA work with its reporting operators to investigate the feasibility of collecting and including in the NTD a full set of transit use and supply data categorized by day of the week and time of day, particularly as automated data collection systems make such data easier to gather.

Across time periods, the NTD and APTA[4] datasets measure ridership as unlinked passenger trips (UPT), often called “boardings.” The way ridership is defined is of critical importance, as consumption of transit is quite often—perhaps most often—the dependent variable of analyses that draw from these datasets. Each time a passenger boards a transit vehicle counts as one UPT. When a commuter transfers between two buses, say, on a single journey from home to work, that journey counts as two UPTs. For operators that neither track nor discount transfers—instead charging the full fare for every boarding—collecting UPT data most closely matches their operations. However, many operators do offer transfer discounts or free transfers. Indeed, many subway systems—whose stations are split by fare gates into fully delineated paid and free areas—do not count boardings separately at all, since passengers transfer between trains without passing through fare gates. For example, while one such system, Bay Area Rapid Transit (BART), collects and disseminates among the most comprehensive datasets of any American transit agency (including origin-destination matrices between every pair of stations every month)—their staff nevertheless must submit UPT estimates to the NTD instead of counts, for this reason.

Not only do unlinked trips fail to capture the way many transit systems operate, their use in datasets could potentially cause misleading trends. For instance, a number of bus operators like Houston Metro have restructured their entire networks into grid-like systems (Binkovitz, 2016), wherein service along trunk routes is now more frequent and a greater share of journeys require a transfer. A resulting increase in UPTs could represent evidence of a successful redesign that has attracted more riders—but could also indicate that the same number of travelers, taking the same number of journeys, simply had to make more transfers, which research shows that riders generally dislike and seek to minimize (Iseki and Taylor, 2009). For this reason, the NTD should also explore with its reporting operators ways to manageably and reliably include measurements of linked trips as well, which, for most analyses, better reflect trends in the consumption of transit.[5] Of course, this will require many operators to estimate linked trip counts, but, as many other operators already estimate *unlinked* trip counts (not to mention passenger miles traveled and other metrics), this should not significantly change the overall reliability of the data.

---

4. Though historical data from prior to the advent of the NTD counts linked trips instead

5. The NTD once included linked trip estimates, though they were highly unreliable and were discontinued some time prior to the NTD being made accessible online.

Even if added to the NTD, these linked trip counts by individual operators may nonetheless exclude transfers between different agencies. In areas with fragmented transit operators, these transfers may make up a substantial share of transit use. Some operators do track inter-agency transfers, though, as a part of offering smart card discounts for transfers. These numbers too should be reported to the NTD where available, to give a fuller picture of transit use across regions.

Other researchers have identified further deficiencies in the NTD. Diffee (2018) points out a number of ways the NTD inconsistently classifies modes and types of rights-of-way, which creates particular problems for research on the efficacy of bus rapid transit. Likewise, NTD data on rural and tribal transit misses a number of operators, includes fewer variables than for urban operators, and only dates back to Report Year 2007 (Mattson, Godavarthy, and Hough, 2016; Godavarthy, Mattson, and Ndembe, 2014; Gan, Liu, and Alluri, 2016; and FTA, 2020b).

Naturally, the accuracy of NTD data is dependent on that of the underlying annual and monthly data that operators report. The NTD allows operators to collect ridership data using automated passenger counters (APCs), manual passenger counters, farebox data, and drivers' logs, among other methods and to either conduct full counts or surveys, the latter using a number of approved sampling and estimation methods. These different methods have varying degrees of accuracy yet are all aggregated in NTD reports. Studies by Chu (2006, 2009) have shown that patronage numbers reported to the NTD tend to be higher—in some cases up to 50% more—than figures reported to APTA.[6] Chu chalks this up to systematic sampling errors or even intentional manipulation—a number of states use NTD data in their transit grant allocation formulas, meaning agencies' NTD ridership reports are directly tied to dollars. In our research on Bay Area transit ridership, we too found differences, albeit mostly slight, between NTD data and internal operator ridership and service statistics (Wasserman et al., 2020).[7]

Transit Rider Data

For its faults, the NTD is a quite useful dataset, and often the first and best place to look for a supply-side analysis of transit in the U.S. Demand-side analysis of aggregate transit use, however, rely on a number of datasets, each of which partially overlaps and none of which alone captures transit use for all trip purposes at a geographic level more granular than the metropolitan area. However, these datasets are particularly useful for analyses of the important relationships of private automobile use and ownership with transit use.

The most comprehensive demand-side dataset on travel patterns is the National Household Travel Survey (NHTS). Administered every 5 to 8 years since 1969 by the Federal Highway Administration (FHWA), the NHTS[8]

6. Though these papers predate the FTA's 2009 *NTD Sampling Manual*, whose requirements may be more robust (FTA, 2009).

7. The data could also be unreliable at their source. In the late 1980s, one of the co-authors of this report was a regular rider of the San Francisco Municipal Railway (as it was then known) and witnessed first-hand an operator on the N Judah light rail line fail to use the (then-manual) passenger counter as patrons boarded the vehicle between Duboce and Church Streets, where the line emerges from its subway tunnel, and Ocean Beach, at the far western reaches of the Sunset District. But at the end of the line, the operator manually depressed the counter perhaps 100 times in rapid succession to produce his passenger "count" for that vehicle run.

8. Originally called the Nationwide Personal Transportation Survey

is a national stratified survey that collects a one-day travel diary and associated employment, housing, demographic, and vehicle ownership data, from over 100,000 households in 2017. The NHTS includes add-ons, or oversamples of states and other jurisdictions that fund them (FHWA, 2017 and Westat, 2019). The NHTS is the premier dataset for analyses of how households and individuals make travel decisions, including mode splits, number of trips and tours (which are chains of linked trips), and far more. The NHTS is also useful for aggregate analysis of transit ridership by allowing for demographic breakdown of who uses transit, how often, when, and for what purposes. For instance, Schouten, Blumenberg, and Taylor (forthcoming) used the NHTS to explore how changes in which groups are riding transit (composition effects) and changes in how much they ride (rate or usage effects) have affected transit ridership trends in California. They find that lower transit use among Hispanic travelers and those with limited auto access have most depressed overall transit patronage in the state. Researchers conducting analyses in California can supplement the NHTS data with the California Household Travel Survey, a similar survey that the California Department of Transportation (Caltrans) conducted within the state in 2012 and is scheduled to repeat in 2022 (Caltrans, 2012, 2017).

While the NHTS and CHTS are invaluable resources—whose uses in mobility research extend beyond public transit—their application to analyses of aggregate transit use come with certain drawbacks. First, the NHTS add-ons for specific states and regions collectively comprise the majority of the NHTS' sample (FHWA, 2009, 2017). The NHTS' coverage of areas without add-ons is somewhat limited, with most states having sample sizes too small to analyze in isolation. Moreover, two of the three most recent iterations of the NHTS, 2001 and 2009, occurred at the nadir of significant economic recessions (FHWA, 2009; Westat, 2019; and World Bank, 2020). Because NHTS data on the intermediate years are not available, the growth in per-capita transit use throughout the mid-2000s cannot fully be captured with this data source alone, especially as some of that growth was erased by 2009. Meanwhile, the most recent NHTS, 2017, was taken during a period of relative economic growth (FHWA, 2017 and World Bank, 2020). Examining trends between the 2009 and 2017 NHTSs involves comparing travel during two very different years with dissimilar economic conditions, potentially confounding research on the effects of other factors like fares, gas prices, auto ownership, etc. on transit use. Regardless of the specific conditions of the select years, NHTS (and CHTS) iterations have used different methodologies, which may also muddle comparisons over time (The 2017 NHTS differed from the 2009 iteration by calculating trip distances using Google Maps instead of self-reporting and by creating the new category of "loop trips," among other changes.) (Westat, 2019).

Moreover, the NHTS (and CHTS) do not neatly align with the NTD. Some of this is to be expected: the NHTS and CHTS, demand-side surveys, and the NTD, a supply-side near-census, differ in who provides their data, over what period the data are captured, and what share of the total population is sampled. However, some of the resulting differences are striking. For example, Californians took 39.89 transit trips each in Report Year 2009, according to the NTD (dividing by intercensal population estimates from U.S. Census Bureau) (FTA, 2020b and U.S. Census Bureau, 2020a)—almost identical to the NHTS' 2009 figure of 39.91 trips per person (FHWA, 2009). But in Report Year 2017, the NTD pegged Californians at only 33.62 annual transit trips each (FTA, 2020b and U.S. Census Bureau, 2020a), compared to 38.86 in 2017 in the NHTS (FHWA, 2017). In other words, the NTD data show a marked decline in transit use in the Golden State in the 2010s, while NHTS data present a far more modest decline (Taylor et al., 2020). For some metropolitan areas, the NTD and NHTS even disagree about the direction of ridership trends. The NTD shows a decrease from 71.63 annual transit trips per capita in Report Year 2009 to 66.18 in Report Year 2017 in the nine-county San Francisco Bay Area[9] (FTA, 2020b and U.S. Census

9. The region covered by the Bay Area's metropolitan planning organization (MPO), the Metropolitan Transportation Commission: Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, and Sonoma Counties

Bureau, 2020a), while the NHTS instead indicates an *increase* from 60.52 annual transit trips per capita in 2009 to 77.84 in 2017 in the San Francisco and San José Metropolitan Statistical Areas[10] (FHWA, 2009, 2017 and U.S. Census Bureau, 2020a). Though these time periods and geographies do not perfectly overlap,[11] the two data sources show Bay Area transit ridership moving in opposite directions.

What explains these differences? According to a Federal Highway Administration report, the 2017 NHTS differs from NTD and APTA datasets—especially by overreporting bus trips—because the 2017 NHTS overrepresented cell-phone-only households, underrepresented rural households, and changed the question wording such that transit trips with a same-mode transfer were likely overcounted (Jenkins, 2019). Because of these issues, it is somewhat difficult—though certainly still illustrative—to use the NHTS to explore recent transit ridership trends. Moreover, because the NHTS was not administered between 2009 and 2017, neither the 2010-2014 post-Great-Recession ridership revival nor the 2015-2020 pre-pandemic ridership decline can be examined in isolation using NHTS data alone. Methodological problems aside, comparing the 2009 and 2017 NHTSs simply cannot shed that much light on why transit use fell during the latter half of the 2010s, due to the dates they were taken.

Beyond the NHTS, a few data sources allow for geographic analysis of transit trips: not just who makes transit trips, but where. One of these is the American Community Survey (ACS), the primary U.S. Census Bureau product conducted between decennial censuses. Continuously administered since 2005, the ACS asks one in 40 Americans every year a comprehensive set of questions on demographics, transportation, employment, housing, and other topics. Due to the potential for a cross-tabulation to be so specific that it could breach confidentiality, the ACS is available in less-fine-grained geographic detail for one-year averages and more detail for five-year averages (U.S. Census Bureau, 2020a and Weinberger, 2019). Note that five-year ACS averages are sometimes misleadingly referred to by their final year; a reference to a “2017 ACS five-year average,” for example, likely means the 2013-2017 average and, if any year, most reflects conditions in its middle year, 2015. Standard ACS data tables are available online from the U.S. Census Bureau (U.S. Census Bureau, 2020a). A more customizable version of the ACS, the Public Use Microdata Sample, is available through the University of

10. A population-weighted average of the San Francisco-Oakland-Berkeley Metropolitan Statistical Area (MSA) (Alameda, Contra Costa, Marin, San Francisco, and San Mateo Counties; formerly known as the San Francisco-Oakland-Fremont MSA and the San Francisco-Oakland-Hayward MSA) and the San José-Sunnyvale-Santa Clara MSA (San Benito and Santa Clara Counties)

11. The NHTS classifies survey responses by metropolitan statistical area in MSAs with 1 million people or more (the top 52 MSAs) (FHWA, 2017). However, smaller MSAs and micropolitan statistical areas are not delineated, limiting the effectiveness of the NHTS for analyzing transit use in those areas (even within states with NHTS add-ons). Moreover, major transit operators like New Jersey Transit, Bay Area Rapid Transit, the Maryland Transit Administration, and Metro-North (the nation’s seventh-, eleventh-, sixteenth-, and seventeenth-largest operators, by Report Year 2018 boardings) operate across multiple MSAs (FTA, 2020b). Thus, exact NTD and NHTS geographic comparisons are often impossible. For these reasons, we broke down transit use by MPO region instead (Taylor et al., 2020). We find MPOs to be a more logical geography for many regional transit analyses.

The NTD classifies operators by their urbanized area (FTA, 2020b)—another regional geography, which, unlike MPO regions and MSAs, does not neatly align with political boundaries. In our research, we manually classified operators into MPO regions (Taylor et al., 2020; Blumenberg et al., 2020). Diffee (2018) offers a method for estimating transit service and use statistics at the MSA-level from the NTD.

Minnesota's IPUMS service (Ruggles et al., 2020), which allows users to create custom tables, though with less geographic specificity (U.S. Census Bureau, 2019, 2020e).[12]

Relevantly for public transportation research, the ACS asks about the mode each respondent takes to work, when they leave for work, how long their commute takes, and how many cars their household owns (U.S. Census Bureau, 2020a). Thus, researchers can calculate the mode share and commute time of residents (and, to a lesser extent, workers) in a given area from the ACS, break down these statistics along various demographic lines, and compare them over time. These figures are often used in academic research, evaluations of transportation projects, and reference statistics for towns and metropolitan areas, particularly because ACS data are available online in formats that obviate the need for statistical software programs.

However, what the public ACS tables do not show is commute flows.[13] In other words, while a user can look up what share of residents of a given census tract take transit to work, they cannot tell where those people work. As transit operates in networks, improvements and changes in one area affect travelers who live in many other areas, some far away. For instance, a new Transbay Tube under San Francisco Bay for BART would increase transit service and connectivity into some of the densest job clusters in California and would therefore affect resident mode share in commuter suburbs at the end of BART lines, far from the new tunnel itself.

To capture commute flows, researchers can instead turn to LODES data (U.S. Census Bureau, 2020b, 2020c). Abbreviated with a nested acronym, the LEHD Origin-Destination Employment Statistics (LODES) is a product of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) program. LODES is drawn from state employment records, which include the home and work addresses of workers and cover 95 percent of salary and wage jobs in the U.S. LODES remarkably categorizes data down to the census block level[14] (with some "fuzzing" to protect privacy) and includes information on job categories and flows between work and home as well, all accessible from the Census Bureau's *OnTheMap* website and for direct download (Graham, Kutzbach, and McKenzie, 2014; Seo et al., 2017; and U.S. Census Bureau, 2020b, 2020c). Though LODES does suffer from some flaws—large employers with multiple worksites may list all employees work addresses at one central location, and LODES excludes self-employed and informal workers (Evans, 2017; Graham, Kutzbach, and McKenzie, 2014; and Seo et al., 2017)—it is nonetheless quite useful. UCLA ITS researchers have used LODES to examine increasing commute distance and jobs-housing imbalances in California (Blumenberg and King, forthcoming; Taylor et al., 2020; and Blumenberg et al., 2020) and to provide the number of station-area jobs for stop-level analysis of the determinants of transit ridership (Wasserman et al., 2020); other researchers have also utilized LODES to explore the factors behind the recent ridership decline (Berrebi and Watkins, 2020).

Yet LODES data are missing a critical element for transit research: mode of travel. Employment records tell where people live and work, but not how they get in between the two. For that, LODES data must be analyzed in conjunction with other data sources, like the NHTS and ACS. One data source, however, does include both commute flows between home and work at reasonably granular geographies and commute mode: the Census Transportation Planning Products (CTPP) (AASHTO, 2019). Sponsored by the American Association of State Highway and Transportation Officials (AASHTO), the CTPP uses ACS data but with additional permissions and

---

12. Standard ACS tables are usually available at the census tract level (areas of 6,000 or so people), with some available by census block group (areas of 600 to 3,000 people) (U.S. Census Bureau, 2019, 2020a). PUMS data are categorized into Public Use Microdata Areas of at least 100,000 residents each (U.S. Census Bureau, 2020e).

13. Except at very large geographic scale (most the size of a county or larger) in PUMS data (Ruggles et al., 2020)

14. The smallest geographic unit used by the Census Bureau, usually actual blocks of land divided by roads.

data perturbation techniques that allow for more granular data to be released than the ACS itself. CTPP tables include home-, work-, and commute-flow-based crosstabs by mode and by other socio-economic factors (AASHTO, 2019 and Weinberger, 2019). Despite this seeming improvement on the ACS and LODES data, the CTPP is "underutilized" by researchers, according to a review of papers using the CTPP (Seo et al., 2017). For one, there are only four full vintages of CTPP data available online—1990, 2000, 2006-2010, and 2012-2016 (AASHTO, 2019)—limiting its utility for comparisons across time, especially for recent ridership changes. For another, its website is unwieldy, outdated, and underpublicized, compared to the newly revamped *data.census.gov* (U.S. Census Bureau, 2020a) and the proprietary *Social Explorer* site (Social Explorer, 2020) where the ACS is hosted.

All of these sources—ACS, PUMS, LODES, and CTPP—include information on commute trips only. However, commuting only accounts for 38.7% of transit trips and 20.2% of all trips (FHWA, 2017). Overreliance on these data sources can overrepresent the role of commuting patterns—longer distance, suburb-to-city, hub-and-spoke, etc.—at the expense of other trip types. This is particularly important in the context of recent transit ridership declines. In the Bay Area, we found that trips made at off-peak times, on weekends, in counter-commute directions, in outlying areas, and on smaller operators account for a large and disproportionate share of the whole region's patronage decline (Wasserman et al., 2020). An analysis that relied only on transit commute data would miss this trend, as commute-heavy transit trip types in the regions have remained stable or grown somewhat in recent years. High-quality data on non-commute trips, comparable across geographic areas and transit operators, is a major gap in transit data collection—and hence our understanding of transit use.

Likewise, all "household" surveys above—the NHTS, the ACS, and products derived from it—only sample households with mailing addresses (Westat, 2019; ORC Macro et al., 2002; Weinberger, 2019; and Saphores et al., 2013). The ACS does sample from overnight homeless shelters, but those experiencing homelessness unsheltered are not surveyed (ORC Macro et al., 2002). The travel patterns of people experiencing homelessness are understudied (Murphy, 2019; Jocoy and Del Casino, 2010; and Meyerhoff, Micozzi, and Rowen, 1993) and other data sources on their use of transit are highly limited. Better data collection would help further understanding of homelessness on transit systems and better help transit operators address the growing homeless crisis.

Data on Transit Service Quality

The data sources above provide a broad, if not entirely complete, picture of transit use, transit service supply, and many external factors that may influence transit ridership. However, the quality of transit service affects transit ridership as well as the quantity. Data on service quality are collected and disseminated in a much more patchwork fashion. Operators often collect service quality data, such as late timepoint arrivals, early timepoint departures, revenue miles between road calls, etc., but what exact metrics are tracked vary widely (Taylor et al., 2020). Moreover, these data are not compiled in a centralized database like the NTD. The timeliness of vehicles, the number of breakdowns and delays, crimes on vehicles and in stations, and other such factors influence patronage, yet obtaining such data often requires requests to each agency of interest or even to specific departments therein. These data are rarely comparable across operators.

In the same vein, other factors that affect passenger satisfaction—and hence ridership—can only or can best be measured by rider surveys. This is of special relevance to work on the ridership of people of color, women and non-binary people, people with disabilities, people in poverty, and other marginalized groups. Factors like harassment, perceptions of safety, and ease of access for those with impaired mobility are typically absent from other datasets and largely measured in passenger surveys (and other agency-specific data like incident reports and transit-related police statistics). Moreover, since the NHTS and other broad travel surveys generally lack the specificity to allow analysis of riders' demographics on particular operators, these rider surveys are often the only way to know the composition of who rides each system.

Yet to incorporate these aspects of service quality into analyses of transit use, researchers typically must rely on a patchwork set of passenger surveys, incident counts, etc., conducted at different times, at different intervals, with different questions or methods (Taylor et al., 2020). Again, the results of these passenger surveys and other service quality data are not collected in a central location, though the nascent Central Archive for Transit Passenger Data (CATPAD), an effort begun by Gregory Newmark (Newmark and Nixon, 2018) and now funded by FTA through the Transportation Secure Data Center (Interstate Transit Research Symposium, 2020), is designed to house such surveys.

Furthermore, surveys typically gather "stated" and not "revealed" preferences; that is, they ask respondents about their "stated" perception of transit service, preferences, and concerns but not how these relate to their "revealed" transit use behaviors (Taylor et al., 2020 and Carrel and Li, 2019). Even if researchers obtain survey data from operators, they must account for this discrepancy between stated and revealed preferences. Relatedly, operators rarely survey those who never ride or stopped riding—a much larger population than regular transit riders and a harder group to survey than riders on the system itself. Their reasons for not riding might be of great aid to agencies seeking to restore or expand ridership.

Because the specific conditions and needs of each operator differ, we conclude that suitable service quality metrics and passenger survey questions ought to vary to some degree as well. Complete uniformity in the metrics collected from operators across the country could preclude a full understanding of contexts that differ; some tailoring in metrics is needed to properly address some operators' unique circumstances. On the other hand, uniformity in reporting enables cross-agency comparisons and analyses of industry standards. One way to balance the benefit of uniformity with the benefit of tailoring in metrics collected is to develop a standard or uniform set of questions for operator peer groups, an approach to performance measurement and evaluation recommended by Gahbauer et al. (2019).[15] If the Federal Transit Administration were to annually report on trends for such a standard set of performance metrics by operator class, this would help public officials, journalists, and other interested parties better understand how transit services and use of them are changing over time. While such an annual reporting process would likely prove politically or operationally challenging in the short-term, transit agencies should at least be required to submit their existing service metrics and survey results to a central repository like CATPAD. While these statistics will not necessarily be directly comparable across agencies, having them all in one place will not only aid transportation research but would also allow for comparisons within some agencies over time.

The General Transit Feed Specification

Finally, one of the most significant changes in transportation data collection this millennium is the General Transit Feed Specification (GTFS), a standardized data format that most American transit agencies and many worldwide now use to publish route, stop, and schedule information (MobilityData, n.d.; Antrim and Barbeau, 2010; and

15. Gahbauer et al. (2019) recommend that, for the purposes of transit performance measurement, peer groups be determined by performance and system characteristics, based on the methods used in Pennsylvania and Wisconsin, which the authors identify as models.

TransitWiki, 2019b). Originally the Google Transit Feed Specification, GTFS was first used in 2005 by TriMet in Portland, Oregon to add their system into Google Maps' trip planner (Antrim and Barbeau, 2010; Roth, 2010; and TransitWiki, 2019b). Smaller systems like the Humboldt Transit Authority added their routes on GTFS in 2007, kicking off GTFS' expansion to cover over one million transit stops worldwide (Antrim, 2016). GTFS feeds are used by many applications (not just Google's), researchers, and practitioners to map transit service and measure service at a local level, as well as many ordinary travelers to plan their transit trips (Antrim and Barbeau, 2010). Transit researchers have used GTFS, for instance, to create a typology of every census tract in the country using transit supply as one of the factors (Brown et al., 2016) and to explore which segments of the population in major U.S. cities have stopped riding the bus as much in recent years (Berrebi and Watkins, 2020). A real-time (RT) add-on to GTFS (MobilityData, n.d.; TransitWiki, 2018; and Frick, Kumar, and Post, 2020), as well as comparable, less common systems based on the Representational State Transfer/Extensible Markup Language (REST/XML) format (D'Agostino, Pellaton, and Brown, 2019; and TransitWiki, 2020), allow agencies to share real-time information on transit vehicle location from on-board sensors. This too has been useful to travelers, mapping applications, and scholars. As one example, Hansen (2020) collected two months of real-time bus locations from Los Angeles County Metropolitan Transportation Authority's (Los Angeles Metro) application programming interface (API) to examine bus speed and reliability to evaluate potential bus lane corridors. Researchers are now developing further standards for "as-operated" transit data, like the Transit ITS Data Exchange Specification (TIDES) for vehicle locations, on-board passenger counts, and fare payments (Levin, 2016), currently being expanded upon in a Transit Cooperative Research Program project (L. Goldstein, 2019).

It is important to note that GTFS is not a database itself, but rather a data format—one that launched a trend of open data standards for other public data (McHugh, 2013). GTFS is designed to provide information to riders, not for data storage per se. However, GTFS data are collected at a number of online repositories, run by organizations that catalog operators' online GTFS feeds, including *TransitWiki* (TransitWiki, 2020), *Transitland* (Mapzen Foundation, n.d.), *OpenMobilityData* (MobilityData IO, n.d.), and the National Transit Map (BTS, 2018). And while GTFS data are valuable for detailed geographic analysis of transit supply and reliability, the GTFS format is not designed to track ridership itself. Though an add-on for ridership, GTFS-ride, is currently being piloted at a small scale (Carleton et al., 2019 and Oregon Department of Transportation and Oregon State University, 2019), the closest substitute for long-term ridership tracking under GTFS today is real-time measurements of crowding, which suffer from inaccuracies in APC data collection and transmission, incomplete coverage across all vehicles, and differing crowding standards across operators (Dasmalchi, 2020 and T. Dai and Taylor, 2020). Crowding data are useful to riders planning trips, but aggregating them up to produce ridership figures may result in incomplete or biased estimates. For ridership data to match with GTFS route and schedule data, researchers will need to obtain datasets from each relevant agency that, for lack of a standard specification, vary widely in their format, availability, and timespan covered.

Data on Private Shared Mobility

Despite their gaps and incompatibilities, public data sources on the supply and use of transit in America are numerous and growing. They allow researchers to analyze most internal factors under the control of operators that influence transit use, as well as many external factors like demographic change and employment patterns. But while existing, available data reveal trends in public shared mobility (i.e., public transit) and in private individual mobility that competes with it (i.e., private automobiles), relatively few datasets are available on private shared mobility (i.e., ride-hail services). While not a new phenomenon—pre-World-War-II streetcar systems were almost all for-profit enterprises, which often competed with unlicensed private jitneys (Jones, 1985)—private-sector shared mobility services have grown in use in the past two decades. These include app-based ride-hail services provided by companies like Lyft and Uber (along with traditional taxis); bikeshare and scooter rental services (along with other two- and three-wheeled device rentals); carshare; private vanpools; and corporate shuttles. Some of these services are shared serially, with a device passed from one user to another, and some shared in parallel, with multiple travelers in a vehicle at the same time. Data on how these services are used, for what types of trips, and to what extent they supplant or supplement other modes like transit represent the most significant current lacuna in data relevant to transit ridership.

Ride-hail Data

Perhaps the most well-known and most used of these private shared mobility services is ride-hail, dominated by the transportation network companies (TNCs) Uber and Lyft. Since their launch in 2009 and 2012, respectively, these firms operate in nearly every major U.S. city and abroad in Uber's case. They have delivered at least 18 billion rides worldwide as of writing (Hartmans and Leskin, 2019 and Iqbal, 2020). Ride-hail is likely having an effect on transit ridership, though the direction of this effect is theoretically ambiguous (Hall, Palsson, and Price, 2018). Ride-hail use could:

1. replace transit trips, if travelers use TNCs instead of transit, thereby decreasing transit use; and/or
2. complement transit travel, if travelers use ride-hail to connect to traditional transit at the "first mile" (from trip origins to transit stops and stations) and/or "last mile" (from transit stops and stations to final destinations), thereby increasing transit use; or
3. have no relationship with transit use, if travelers use TNCs instead of a non-transit mode (like driving or walking) or use them for trips they would otherwise not have made without ride-hail service (Taylor et al., 2020; Blumenberg et al., 2020; and Wasserman, 2019).

Ride-hail may also affect transit ridership indirectly: if ride-hail worsens traffic congestion and, consequently, reduces bus speeds, a transit rider may drive, walk, or bicycle instead. On the other hand, affordable, reliable ride-hail service could motivate some households to downsize their personal vehicle fleets and consequently make more transit trips (Taylor et al., 2020; Blumenberg et al., 2020; and Wasserman, 2019).

For reasons discussed below, ride-hail companies have resisted voluntarily sharing data on their services with government analysts and researchers, which has made it difficult to know much about their use and effects. To work around the lack of ride-hail data, researchers have taken "back-door" approaches to resolving this question. Graehler, Mucci, and Erhardt (2019); Babar and Burtch (2017); and Hall, Palsson, and Price (2018) have each used differences in the timing of ride-hail firms' launches in various metropolitan areas to analyze their effects on transit ridership; the former two find more evidence of substitution than complementarity, especially from buses. After Uber and Lyft declined the San Francisco County Transportation Authority's (SFCTA) request for trip data, SFCTA researchers used a since-closed feature of the TNCs' API to map ride-hail trips within the city, finding that many occur along major transit corridors and in areas with high transit supply (Castiglione et al., 2017). Finally, a number of surveys (Schaller, 2018; Feigon and Murphy, 2016; Circella et al., 2018; Clewlow and Mishra, 2017; Gehrke, Felix, and Reardon, 2018; Hampshire et al., 2017; Henao, 2017; New York City Department of Transportation, 2018; Rayle et al., 2016; and Dong, 2020) have asked ride-hail users how they would have made their trip if TNCs were not available, in order to quantify transit substitution rates. The results vary widely: TNCs replace anywhere from 14 to 50 percent of transit trips, depending on the operating environment and trip purpose. These surveys suffer from the same stated preference issues discussed above: respondents' actual behaviors could differ greatly from what they tell a surveyor. Moreover, such studies ask respondents to imagine travel situations in which ride-hail did not exist, leaving unchanged their current or most frequent destination, ignoring the fact that people move, take jobs, choose social venues, etc. based in part on transportation access (Wasserman, 2019).

In examining transit ridership trends, we drew from two additional data sources on ride-hail: the U.S. Census Bureau's Nonemployer Statistics and the NHTS. The former is a dataset of essentially all federally taxed business establishments with no paid employees (U.S. Census Bureau, 2020d). Most of these (87% of establishments in 2018 and 99% of establishments in the "taxi and limousine service" category that covers ride-hail) are sole proprietorships—i.e., individual ride-hail drivers who operate as single-person independent contractors under the current labor organizational model of Uber and Lyft (U.S. Census Bureau, 2018a). In other words, the Nonemployer Statistics count the number of ride-hail and taxi drivers—each one a single "establishment"—for geographies as small as counties. Note that the dataset does not distinguish between ride-hail, taxi, and other for-hire passenger drivers. However, given the decline of the traditional taxi industry over the past decade, researchers can assume that ride-hail accounts for (at least) all of the net growth in the combined sector and therefore obtain reasonable estimates of ride-hail driver numbers by subtracting out the relatively small number of establishments in the "taxi and limousine service" category before ride-hail's advent. The dataset is limited in other ways. Besides counts, its only other variable is receipts, which are grouped into a few buckets by dollar amount. And because it categorizes drivers based on their mailing address (U.S. Census Bureau, 2018b), it underestimates the number of ride-hail drivers working in urban centers, as some drivers commute in from other counties. Nonetheless, the Nonemployer Statistics are a good resource for establishing the scale of the often exponential growth of ride-hail in major metropolitan areas.

Along with the Nonemployer Statistics, the latest iteration of the NHTS tracks ride-hail trips taken on the survey day (FHWA, 2017). From the NHTS, we characterize ride-hail's market—who takes ride-hail trips, how often, when, and for what purpose—and compare it to transit's market (Taylor et al., 2020). However, the NHTS is not a longitudinal survey, so researchers cannot compare the travel patterns of specific people before and after ride-hail's establishment. Moreover, the NHTS does not have the geographic specificity necessary for most analyses of the relationship between TNCs and transit. Meanwhile, the ACS is less useful for analysis of ride-hail travel patterns, as it only includes data on respondents' regular commute mode—not ride-hail's core market—and does not include ride-hail as a listed option on the survey (U.S. Census Bureau, 2020a) (though the Census Bureau plans to pilot new language to include ride-hail in the coming years (Weinberger, 2019)).

Despite all of these data sources each providing an oblique view of the relationship between ride-hail and transit, the only true way to resolve the question of how one's use affects the other's is to analyze Uber and Lyft's proprietary trip data. Uber's online tool, Uber Movement (Uber Technologies, 2020), shows travel times between pairs of census tracts and traffic analysis zones, speeds on major roads, and heatmaps of scooter and bikeshare locations, by time of day and/or day of the week, but does not shed much light on ride-hail's relationship to transit. Because the data are presented at such an aggregated level and because volumes of trips between areas are not shown, users cannot ascertain the scale and mode share of ride-hail, its complementarity with and substitution for transit, and its change over time (Marshall, 2019 and TransitCenter, 2019). Moreover, Uber Movement only displays data for denser parts of select metropolitan areas, not smaller cities nor more outlying regions where it may well have a greater substitution effect on transit. As the research and advocacy group TransitCenter wrote in an article entitled "Accept No Substitutes for Full Trip Data from Uber and Lyft" (2019), "Uber Movement doesn't tell cities how Uber itself is affecting the transportation system." Only fuller, more disaggregated trip data—with volumes, origin-destination pairs at greater geographic specificity, and date and time information—can show that.

Researchers at Uber themselves have demonstrated the utility of such data, by, for instance, looking at Uber trip growth nearly newly opened Los Angeles Metro light rail stations to estimate complementarity between the modes (Williams, 2017). Yet Uber and Lyft keep these trip data strictly out of public view. They have granted a limited number of researchers access to select data, which those scholars have used to study, for example, the equitability of ride-hail access in Los Angeles (Brown, 2018) and the effect of a new subway line in New York City on ride-hail use (Bhatia, 2019). In fairness to the TNCs, not only are there privacy concerns over rider data, but each company could be concerned about revealing business data to competitors. And to be sure, sharing trip data would invite scrutiny from the general public and regulators. Nonetheless, a few American cities and some worldwide peers, including New York City, Chicago, Toronto, Mexico City, and Beijing, have—though negotiations and legal mandates—required TNCs to share trip data (usually on par with what taxi services share), some of which are now available to the public (Joshi et al., 2019). These arrangements have enabled novel and revealing studies of the interactions between ride-hail and transit in those cities (Barajas and Brown, 2020; Davidson, Peters, and Brakewood, 2017; and Jin, Kong, and Sui, 2019), while maintaining rider anonymity in the dataset. These cities, though, represent the exception, not the rule. In California, while TNCs must submit data to the California Public Utilities Commission (CPUC) in order to check compliance with various labor, insurance, and safety regulations, the CPUC does not make these data available even to other state agencies, much less municipalities or the public (SFCTA, 2018). This hampered our ability to analyze ride-hail's effects on transit in the Bay Area, which are likely significant but about which we could only make a circumstantial case (Blumenberg et al., 2020). Indeed, the fact that transit use nationally declined in the latter half of the 2010s as the economy grew and service supply increased leaves ride-hail as a key suspect. Without better public data, though, that suspicion remains unresolved. To be sure, policymakers must, in regulating data-sharing, balance the needs, costs, and privacy of riders and travelers, TNCs, governments, operators of other modes, and further interests; for further discussion of these concerns and models for data-sharing arrangements, see Matute, D'Agostino, and Brown (2020) and D'Agostino, Pellaton, and Brown (2019). But as TNCs continue to expand and ride-hail use likely grows with them, the need for better public data wanes too.

Private Shuttle and Micromobility Data

Beyond ride-hail, other private shared mobility services may be affecting transit ridership, though likely not to the same degree. In some American metropolitan areas, major employers operate corporate bus fleets that shuttle their employees to work. These systems date to the 1950s but have proliferated since 2000 (Singa and Margulici, 2010 and D. Dai and Weinzimmer, 2014). In the media and popular discourse, these shuttles are often associated with technology companies (though other firms also operate them); in the Bay Area, they are often called "Google buses," especially after San Francisco activists protested and blocked shuttles in 2013-2014 (De Kosnik, 2014). These shuttle systems are especially large there: collectively, they would represent the seventhlargest transit agency in the Bay Area by annual boardings. In other areas of the country, however, the degree to which corporate shuttles affect public transit is likely low, but it is difficult to determine their effects absent better public data. Surveys of Bay Area shuttle users (SFMTA, 2015 and D. Dai and Weinzimmer, 2014) show that around a third to a fifth would take public transit were shuttles unavailable, amounting to 11,000 daily trips, though the shortcomings of such rider surveys are again at play. The surveys also capture some of the indirect effects of private shuttles on public transit by asking if respondents do not own a car because of shuttles' availability and if shuttles enabled them to live somewhere different, both of which might lead people to use transit more for non-commute trips. The surveys do not, though, touch on the complex possible contribution of shuttle stops to rising rents and evictions (Goldman, 2013 and Maharawal and McElroy, 2017), which, like new transit-oriented developments around public transit (Dominie, 2012), may lead to the displacement of lower-income, more-transit-prone residents for higher-income, less-transit-prone residents.

Data on corporate shuttles, which have more operators and less geographic spread than ride-hail, are less systematically collected or discussed (Taylor et al., 2020 and Blumenberg et al., 2020). However, the Bay Area's metropolitan planning organization, the Metropolitan Transportation Commission (MTC), has conducted two surveys of private bus operators (Bay Area Council and MTC, 2016 and Bacon, 2019). While voluntary, these "shuttle censuses" are fairly comprehensive and provide the number of shuttles and rider estimates by county-to-county pairs. The San Francisco Municipal Transportation Agency (SFMTA) also collects ridership data from shuttle operators within the city as a permitting requirement (SFMTA, 2015, 2018). Beyond the Bay Area, we are unaware of any systematic data on private shuttle ridership. Even within the Bay Area, researchers and public agencies have taken creative methods to count shuttles, including using cameras along major shuttle routes (Lee, 2018), hiring bike messengers to count and follow buses in order to create a route map (Fowler, 2012), and tracking the wi-fi network names and identifiers of passing buses to determine schedules and ascertain which company is operating which often-unmarked shuttle (Poulsen, 2014).

Before the COVID-19 pandemic, bikeshare and scooter rental systems had explosive growth in major U.S. cities. Again, data on their effects on transit ridership are sparse, both because of their operation by private entities and their newness. Surveys across the country have shown that at most around ten percent of shared scooter and bikeshare trips replaced a transit trip, with higher rates of so-called micromobility and transit used in complement (Chang et al., 2019; Barnes, 2019; Portland Bureau of Transportation, 2018; and Denver Public Works, 2019). Hoping to preempt the chaos remembered from ride-hail's emergence, municipalities have rapidly developed permitting programs for micromobility services, many of which include data-sharing provisions. First developed by the Los Angeles Department of Transportation (LADOT), the Mobility Data Specification aims to be a common format for these micromobility location data, as GTFS is for transit (Open Mobility Foundation, 2020; LADOT, 2018; Owens, 2019; Matute, D'Agostino, and Brown, 2020; Marshall, 2019; and TransitWiki, 2019a). However, micromobility firms have objected to, and sued over, the collection and sharing of these data (Rana and Rundle, 2020; Marshall, 2019; and TransitWiki, 2019a).

Conclusion

Data Gaps

Transit ridership data are important to many users. Planners at transit agencies, policy analysts at government agencies that fund transit, public officials who oversee transit systems, members of the media, advocates for mobility-disadvantaged communities, private-sector mobility providers, and academic researchers all rely on accurate, up-to-date, comprehensive data on transit use to evaluate and improve transportation in the present and plan for its future. Nor are data an abstract concern for transit users. Not only do riders benefit immensely from vehicle-location data that give them in-app updates on their next bus or train's arrival (TransitWiki, 2018), they also benefit long-term from data that enable better-scheduled service and better-informed policy decisions, which can improve their travel choices and quality of life.

We list in **Table 1** above some of the most significant datasets for understanding public transit use in the U.S. These include the National Transit Database for aggregate ridership analysis, the National Household Travel Survey and products derived from the American Community Survey for exploring how individual travel decisions affect transit patronage, GTFS feeds for transit schedules and geodata, and U.S. Census Nonemployer Statistics for at least some sense of the scale of ride-hail services, as they affect transit. However, between these datasets, significant gaps remain in understanding the contours of and factors behind transit use.

In **Table 2**, we list some of the most notable gaps in transit data we have found in the course of our research. Data on public transit supply and aggregate ridership collected from operators are comparatively comprehensive, though in some cases incompatible across agencies or datasets. Collecting counts of linked trips and dividing trip counts by time of day and day of the week, though, would improve regional transit planning and would shed light on recent ridership trends. Datasets on transit riders and individual transit use have larger holes, most noticeably in information on non-commute transit trips and users. Likewise, transit operators collect passenger survey data and performance metrics related to service quality in a piecemeal fashion.

Table 2. Significant Gaps in Transit Ridership Data

Category	Gap	Recommendation
Ridership	Linked trips (transfers within or between operators)	Explore the feasibility of including linked trip counts or estimates in the NTD (potentially drawing on regional transit smartcard data); provide operators methodology and resources from FTA to do so
	Trip counts by time of day and day of the week	Collect temporally disaggregated trip counts in the NTD; provide operators methodology and resources from FTA to do so

Category	Gap	Recommendation
Riders and Travel Patterns	Riders and travel patterns in between NHTS collection years and in more granular geographic areas	Regularize the NHTS schedule every five years and encourage more states and regional planning bodies sponsor oversamples, to allow for more geographically disaggregated analyses
	Transit flows (origin-destination pairs)	Expand and improve CTPP data
	Non-commute transit trips	Adopt the NHTS recommendations above to enable better analysis of non-commute trips through the NHTS; encourage transit operators and other mobility providers to regularly survey riders on their trip purposes in ways that allow for comparisons across operators
	Transit use and travel patterns of those experiencing homelessness	Adopt better survey methods and more survey inclusion of those experiencing homelessness
Transit Service Quality	Performance metrics, passenger satisfaction, demographics of ridership on particular operators, etc.	Collect rider surveys and service metrics in a centralized database like CATPAD; establish peer group determinations; develop a small set of standard performance metrics collected and reported by FTA
	Why people do not ride transit or have given up riding	Conduct more surveys by transit operators of people beyond their own customers
	Safety, policing, fare enforcement, citations, etc. on transit and their effect on ridership	Collect incident/citation counts and reports, disaggregated by characteristics like race/ethnicity and gender, in a centralized database; improve incident reporting and data collection; survey both current and potential riders on the effects of perceptions of safety and policing on ridership
Private Shared Mobility	Ride-hail trip characteristics, especially as they complement and/or substitute for transit	Collect disaggregated, timestamped TNC trip data with origin and destination geolocations, as well as reported connections to transit; make such data available publicly in a form that adopts prudent privacy protections
	Corporate shuttle and micromobility trip characteristics, especially as they complement and/or substitute for transit	Systematically collect and make public shuttle and micromobility trip data, through the same regulations as for ride-hail

We offer some suggestions in **Table 2** for how to close outstanding gaps in transit ridership data. Overall, the scope, collection methods, and collection frequency of major data sources like the NTD and NHTS need to be improved. Data that already exist, such as passenger surveys and corporate shuttle statistics, need to be made available to users, and data that are not currently collected, like transit use by those experiencing homelessness, need to be gathered on a regular basis by transit operators or federal, state, or regional statistical bureaus. If the federal agencies behind existing data sources are unable to implement these recommendations, we suggest that states (or groups of smaller states in a region) might establish or expand their own transit data repositories. Cities and states across the country compile and disseminate data on a wide variety of topics on online open "data hubs" and "data portals"; these types of sites could host the data identified in **Table 2**. While this may not be feasible in states with fewer resources in or experience with data collection and management, California, which already conducts its own state household travel survey (Caltrans, 2012), is well-positioned to expand and centralize its collection of transit data. For instance, the state could collect transit rider surveys or linked trip counts from every operator each year and make them available online (perhaps as the "California Transit Database").

For all their gaps, what data on public transit do exist are generally public. But the scale and spread of private shared mobility, an influence on public transit, remains largely hidden from policymakers, researchers, and the public. In order to understand the future of public transit and American mobility more broadly, open data on private shared mobility is needed. To advance our understanding of public transit here in California, we recommend that the state adopt similar data collection requirements for private shared mobility to those in New York City, Chicago, or Toronto, and that the state apply them across all such services, including TNCs, corporate shuttles, and micromobility. Matute, D'Agostino, and Brown (2020) and D'Agostino, Pellaton, and Brown (2019) discuss the details and nuances of such regulations; the former flag those three cities as having data specifications that enable analysis of the interplay of transit and ride-hail, with New York City's offering the widest range of additional data. The most relevant data to collect are disaggregated, timestamped trips with origin and destination geolocations, as well as reported connections to transit, if collected via surveys or TNC applications (Matute, D'Agostino, and Brown, 2020). After establishing a more robust and open model for collecting ride-hail trip data, some of which the state currently collects to monitor regulatory compliance but keeps strictly private, California state regulators could also house the data in the same data hub as the public operator datasets above, with appropriate privacy protections. For California at least, this effort is probably best done at the state level (perhaps as a joint effort of the State Transportation Agency and the Public Utilities Commission), which has more direct and nimble power over transportation operations than the federal government without the worry of patchwork (or potentially competing) regulation at the local level.

In private and public shared mobility, gaps in data both align with existing inequities and enable them to continue, unmeasured. This is true for all categories identified in **Table 2**. For instance, knowing the demographic breakdown of ridership on different operators, as collected in currently disparate passenger surveys, can help evaluate how policy changes like modifications to fare prices and enforcement, policing, lighting and station design, etc. affect the ridership of different groups. Moreover, disaggregating trips by time of day and day of the week is vital for equity analyses of changes to off-peak service, which disproportionately carries lower-income travelers. Not coincidentally, the areas where data are most lacking correlate with the travel patterns of marginalized groups. On top of this, there are inequities between operators as well in their ability to collect and disseminate high-quality data. Smaller systems, systems with less funding, rural systems, etc. may not have the capacity to close the gaps we identify. To address this, federal and state transportation policymakers should provide funding and resources for data collection to operators or even rider groups, as they establish new data mandates; these resources should be distributed via equity criteria that prioritize systems with higher shares of marginalized riders.

Transit Data Needs Looking Forward

The COVID-19 pandemic has made closing transit data gaps all the more important. For instance, real-time estimates of ridership and crowding on vehicles might better enable social distancing and improve travelers' piece of mind. Likewise, knowing how and where use of private sector shared mobility has changed during the pandemic would allow transit to better complement it, and vice versa, during the pandemic recovery. The real-time ridership and crowding data that some transit operators have shared with their riders' points to a new era of shared mobility data collection and use. In years past, the challenge was in managing the often high costs of data collection, while today, the questions center more on how the mountains of digitally generated data can be winnowed and used to effectively plan for and regulate public and private transit services, while still protecting the confidentiality of users.

References

AASHTO (2019). Census Transportation Planning Products. *AASHTO*. Retrieved September 16, 2020, from https://ctpp.transportation.org/.

Antrim, A. (2016, May). *GTFS: What Is It? Why Does It Matter?* Presented at the Community Transportation Association of America Expo 2016, Portland, OR. Retrieved September 17, 2020, from https://www.slideshare.net/aaronantrim/ctaa-2016-portland-aaron-antrim-gtfs-what-is-it-why-does-it-matter.

Antrim, A., and Barbeau, S. (2010, February). *The Many Uses of GTFS Data—Opening the Door to Transit and Multimodal Applications*. Retrieved September 17, 2020, from http://www.locationaware.usf.edu/wp-content/uploads/2010/02/The-Many-Uses-of-GTFS-Data-%E2%80%93-ITS-America-submission-abbreviated.pdf.

APTA (2020a). Public Transportation Fact Book. *American Public Transportation Association*. Retrieved August 26, 2020, from https://www.apta.com/research-technical-resources/transit-statistics/public-transportation-fact-book/.

APTA (2020b). Ridership Report. *American Public Transportation Association*. Retrieved August 26, 2020, from https://www.apta.com/research-technical-resources/transit-statistics/ridership-report/.

APTA (2020c). Transit Statistics. *American Public Transportation Association*. Retrieved August 26, 2020, from https://www.apta.com/research-technical-resources/transit-statistics/.

ASCE (2017, January). *2017 Infrastructure Report Card: Transit*. ASCE. Retrieved August 25, 2020, from https://www.infrastructurereportcard.org/wp-content/uploads/2017/01/Transit-Final.pdf.

Babar, Y., and Burtch, G. (2017, September 25). *Examining the Impact of Ridehailing Services on Public Transit Use*. Retrieved September 30, 2019, from https://papers.ssrn.com/abstract=3042805.

Bacon, W. (2019). [Personal communication].

Barajas, J., and Brown, A. (2020, September 2). *Not Minding the Gap: Does Ride-hailing Serve Transit Deserts?* https://doi.org/10.31235/osf.io/y4jwk.

Barnes, F. (2019). *A Scoot, Skip, and a JUMP Away: Learning from Shared Micromobility Systems in San Francisco* (MURP Applied Planning Research Project). UCLA, Los Angeles. https://doi.org/10.17610/T6QP40.

Bay Area Council and MTC (2016, September 2). *2016 Bay Area Shuttle Census: First-ever Look at Emerging Regional Resource*. MTC. https://mtc.ca.gov/sites/default/files/2016%20Bay%20Area%20Shuttle%20Census.pdf.

Berrebi, S., and Watkins, K. (2020, June 1). Who's Ditching the Bus? *Transportation Research Part A: Policy and Practice*, *136*, 21–34. https://doi.org/10.1016/j.tra.2020.02.016.

Bhatia, N. (2019, January 19). *Public Transit's Effect on Ride-hail and Taxi Usage*. https://doi.org/10.2139/ssrn.3258592.

Binkovitz, L. (2016, August 16). A Year after Bus Redesign, METRO Houston Ridership Is Up. *Rice Kinder Institute for Urban Research*. Retrieved September 10, 2020, from https://kinder.rice.edu/2016/08/16/a-year-after-redesign-metro-ridership-is-up.

Blumenberg, E., Garrett, M., King, H., Paul, J., Ruvolo, M., Schouten, A., Taylor, B., and Wasserman, J. (2020, February 26). *What's behind Recent Transit Ridership Trends in the Bay Area?: Volume I, Overview and Analysis of Underlying Factors* (UCLA ITS-LA1908). UCLA ITS. https://doi.org/10.17610/T6PC7Q.

Blumenberg, E., and King, H. (forthcoming). Jobs-housing Balance Re-re-visited. *Journal of the American Planning Association*.

Brown, A. (2018). *Ridehail Revolution: Ridehail Travel and Equity in Los Angeles* (PhD diss.). UCLA, Los Angeles. https://escholarship.org/uc/item/4r22m57k.

Brown, A., Blumenberg, E., Taylor, B., Ralph, K., and Voulgaris, C. (2016, March 1). A Taste for Transit?: Analyzing Public Transit Use Trends among Youth. *Journal of Public Transportation*, *19*(1), 49–67. https://doi.org/10.5038/2375-0901.19.1.4.

BTS (2018, April 24). National Transit Map (Data, Maps, and Apps). *Bureau of Transportation Statistics*. Retrieved September 17, 2020, from https://www.bts.gov/national-transit-map/national-transit-map-data-maps-and-apps.

BTS (2020, November 28). The Week in Transportation: Selected Transportation Measures during the COVID-19 Pandemic. *Bureau of Transportation Statistics*. Retrieved December 8, 2020, from https://www.bts.gov/covid-19/week-in-transportation.

Caltrans (2012). California Household Travel Survey. *Caltrans*. Retrieved October 16, 2020, from https://dot.ca.gov/programs/transportation-planning/economics-data-management/transportation-economics/ca-household-travel-survey.

Caltrans (2017, January 15). 2010-2012 California Household Travel Survey. *National Renewable Energy Laboratory*. Retrieved October 16, 2020, from https://www.nrel.gov/transportation/secure-transportation-data/tsdc-california-travel-survey.html.

Carleton, P., Hoover, S., Fields, B., Barnes, M., and Porter, J. (2019, December 1). GTFS-ride: Unifying Standard for Fixed-route Ridership Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(12), 173–181. https://doi.org/10.1177/0361198119851748.

Carrel, A., and Li, M. (2019, April 1). Survey-based Measurement of Transit Customer Loyalty: Evaluation of Measures and Systematic Biases. *Travel Behaviour and Society*, *15*, 102–112. https://doi.org/10.1016/j.tbs.2019.01.003.

Castiglione, J., Chang, T., Cooper, D., Hobson, J., Logan, W., Young, E., Charlton, B., Wilson, C., Mislove, A., and Chen, L. (2017, June). *TNCs Today: A Profile of San Francisco Transportation Network Company Activity*. SFCTA. https://www.sfcta.org/sites/default/files/2019-02/TNCs_Today_112917_0.pdf.

Chang, A., Miranda-Moreno, L., Clewlow, R., and Sun, L. (2019). *Trend or Fad?: Deciphering the Enablers of Micromobility in the U.S.* SAE International. Retrieved January 8, 2021, from https://www.sae.org/binaries/content/assets/cm/content/topics/micromobility/sae-micromobility-trend-or-fad-report.pdf.

Chu, X. (2006, January 1). Ridership Accuracy and Transit Formula Grants. *Transportation Research Record: Journal of the Transportation Research Board*, *1986*(1), 2–10. https://doi.org/10.1177/0361198106198600101.

Chu, X. (2009). The Efficiency of Sampling Techniques for NTD Reporting. *Journal of Public Transportation*, *12*(4), 19. https://doi.org/10.5038/2375-0901.12.4.1.

Chu, X. (2010, December). *Guidebook for Using Automatic Passenger Counter Data for National Transit Database (NTD) Reporting* (NCTR778-03, FDOT BDK85 977-04, USDOT DTRS98-G-0032). National Center for Transit Research, University of South Florida. Retrieved August 27, 2020, from https://rosap.ntl.bts.gov/view/dot/18516.

Circella, G., Alemi, F., Tiedeman, K., Handy, S., and Mokhtarian, P. (2018, March). *The Adoption of Shared Mobility in California and Its Relationship with Other Components of Travel Behavior*. National Center for Sustainable Transportation. Retrieved January 8, 2021, from https://escholarship.org/uc/item/1kq5d07p.

Clewlow, R., and Mishra, G. (2017, October). *Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-hailing in the United States* (UCD-ITS-RR-17-07). UC Davis ITS. https://merritt.cdlib.org/d/ark:%252F13030%252Fm5sj6gbk/1/producer%252FUCD-ITS-RR-17-07.pdf.

D'Agostino, M., Pellaton, P., and Brown, A. (2019, August 1). *Mobility Data Sharing: Challenges and Policy Recommendations* (Issue Paper). UC Davis ITS. Retrieved September 18, 2020, from https://escholarship.org/uc/item/4gw8g9ms.

Dai, D., and Weinzimmer, D. (2014, February 1). *Riding First Class: Impacts of Silicon Valley Shuttles on Commute and Residential Location Choice* (Working Paper UCB-ITS-WP-2014-01). UC Berkeley ITS. Retrieved August 19, 2019, from https://escholarship.org/uc/item/2jr7z01q.

Dai, T., and Taylor, B. (2020, October 1). *When Is Public Transit Too Crowded, and How Has This Changed During the Pandemic?* UCLA ITS. https://doi.org/10.17610/T60S34.

Dasmalchi, E. (2020, October 1). *Using Real-time Crowding Data as a Rider Communication Strategy in the COVID-19 Pandemic*. UCLA ITS. https://doi.org/10.17610/T6W02W.

Davidson, A., Peters, J., and Brakewood, C. (2017). *Interactive Travel Modes: Uber, Transit, and Mobility in New York City*. Presented at the 96th Annual Meeting of the Transportation Research Board, Washington, D.C. Retrieved September 22, 2020, from https://trid.trb.org/view/1438587.

De Kosnik, A. (2014, November 2). Disrupting Technological Privilege. *Performance Research*, *19*(6), 99–107. https://doi.org/10.1080/13528165.2014.985117.

Denver Public Works (2019, February). *Denver Dockless Mobility Program: Pilot Interim Report*. Denver Department of Public Works. Retrieved September 22, 2020, from https://www.denvergov.org/content/dam/denvergov/Portals/705/documents/permits/Denver-dockless-mobility-pilot-update-Feb2019.pdf.

Diffee, C. (2018). *Data Preparation and Methods for Assessing U.S. Transit Ridership Trends*. Georgia Tech, Atlanta. Retrieved September 11, 2020, from https://smartech.gatech.edu/bitstream/handle/1853/60823/DIFFEE-THESIS-2018.pdf.

Dominie, W. (2012). *Is Just Growth Smarter Growth?: The Effects of Gentrification on Transit Ridership and Driving in Los Angeles' Transit Station Area Neighborhoods* (MURP Applied Planning Research Project). UCLA, Los Angeles. https://escholarship.org/uc/item/5vf5x8pg.

Dong, X. (2020, April 2). Trade Uber for the Bus?: An Investigation of Individual Willingness to Use Ride-hail versus Transit. *Journal of the American Planning Association*, *86*(2), 222–235. https://doi.org/10.1080/01944363.2019.1687318.

Evans, T. (2017). *Visualizing Commuting Patterns Using LODES Data*. Presented at Bloomington, IN. Retrieved September 16, 2020, from https://lehd.ces.census.gov/doc/workshop/2017/Presentations/TheaEvans.pdf.

Feigon, S., and Murphy, C. (2016, September 15). *Shared Mobility and the Transformation of Public Transit* (TCRP Research Report 188, Project J-11, Task 21). Shared-use Mobility Center. https://doi.org/10.17226/23578.

FHWA (2009). 2009 National Household Travel Survey. *National Household Travel Survey*. Retrieved October 2, 2019, from https://nhts.ornl.gov/.

FHWA (2017). 2017 National Household Travel Survey. *National Household Travel Survey*. Retrieved October 2, 2019, from https://nhts.ornl.gov/.

Fowler, G. (2012, October 10). Map Reveals Corporate Bus Routes Tech Workers Take. *Wall Street Journal*. Retrieved September 22, 2020, from https://online.wsj.com/article/SB10000872396390443982904578044520776395486.html.

Frick, K., Kumar, T., and Post, A. (2020, July 1). *The General Transit Feed Specification (GTFS) Makes Trip-Planning Easier—Especially During a Pandemic—Yet Its Use by California Agencies Is Uneven* (UC-ITS-2020-31). UC Berkeley ITS. https://doi.org/10.7922/G2ZG6QJZ.

FTA (2009, March 21). *National Transit Database Sampling Manual*. FTA. Retrieved September 11, 2020, from https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/The_NTD_Sampling_Manual.pdf.

FTA (2019, December). *2018 National Transit Summaries and Trends*. FTA. Retrieved August 27, 2020, from https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/ntd/data-product/134401/2018-ntst_1.pdf.

FTA (2020a, July). *National Transit Database 2020 Policy Manual: Full Reporting*. FTA. Retrieved August 26, 2020, from https://www.transit.dot.gov/sites/fta.dot.gov/files/2020-07/2020%20NTD%20Reporting%20Policy%20Manual_0.pdf.

FTA (2020b). The National Transit Database (NTD). *Federal Transit Administration*. Retrieved October 16, 2020, from https://www.transit.dot.gov/ntd.

Gahbauer, J., Lederman, J., Huang, E., Wachs, M., Matute, J., and Taylor, B. (2019, August 28). *An Assessment of Performance Measures in the Transportation Development Act*. UCLA ITS. Retrieved November 9, 2020, from https://escholarship.org/uc/item/0dk5g542.

Gan, A., Liu, K., and Alluri, P. (2016, January 1). Integrated Analysis System for Rural National Transit Database. *Transportation Research Record: Journal of the Transportation Research Board*, *2539*(1), 167–177. https://doi.org/10.3141/2539-19.

Gehrke, S., Felix, A., and Reardon, T. (2018, February). *Fare Choices: A Survey of Ride-hailing Passengers in Metro Boston* (MAPC Research Brief 1). Metropolitan Area Planning Council. Retrieved August 3, 2018, from http://www.mapc.org/wp-content/uploads/2018/02/Fare-Choices-MAPC.pdf.

Godavarthy, R., Mattson, J., and Ndembe, E. (2014, July). *Cost-Benefit Analysis of Rural and Small Urban Transit* (21177060-NCTR-NDSU03). North Dakota State. Retrieved September 11, 2020, from https://www.nctr.usf.edu/wp-content/uploads/2014/12/77060_NCTR-NDSU03.pdf.

Goldman, A. (2013). *The "Google Shuttle Effect": Gentrification and San Francisco's Dot Com Boom 2.0* (PhD diss.). UC Berkeley, Berkeley, CA. Retrieved September 22, 2020, from http://svenworld.com/wp-content/uploads/2014/01/Goldman_PRFinal.pdf.

Goldstein, L. (2019, August 8). TCRP G-18: Improving Access and Management of Transit ITS Data. *The National Academies of Sciences, Engineering, and Medicine*. Retrieved November 16, 2020, from https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4687.

Graehler, M., Mucci, R., and Erhardt, G. (2019). Understanding the Recent Transit Ridership Decline in Major U.S. Cities: Service Cuts or Emerging Modes? Presented at the 98th Annual Meeting of the Transportation Research Board, Washington, D.C. https://trid.trb.org/view/1572517.

Graham, M., Kutzbach, M., and McKenzie, B. (2014, September). *Design Comparison of LODES and ACS Commuting Data Products* (CES 14-38). U.S. Census Bureau Center for Economic Studies. Retrieved September 16, 2020, from https://www2.census.gov/ces/wp/2014/CES-WP-14-38.pdf.

Hall, J., Palsson, C., and Price, J. (2018, November). Is Uber a Substitute or Complement for Public Transit? *Journal of Urban Economics*, *108*, 36–50. Retrieved July 29, 2019, from https://www.sciencedirect.com/science/article/pii/S0094119018300731.

Hampshire, R., Simek, C., Fabusuyi, T., Di, X., and Chen, X. (2017, May 31). *Measuring the Impact of an Unanticipated Disruption of Uber/Lyft in Austin, TX* (SSRN Scholarly Paper ID 2977969). Social Science Research Network. Retrieved August 19, 2019, from https://papers.ssrn.com/abstract=2977969.

Hansen, M. (2020). *Need for Speed: Opportunities for Peak-hour Bus Lanes along Parking Corridors in Los Angeles* (MURP Applied Planning Research Project). UCLA, Los Angeles. Retrieved September 17, 2020, from https://escholarship.org/uc/item/7fv44138.

Hartmans, A., and Leskin, P. (2019, May 18). The History of How Uber Went from the Most Feared Startup in the World to Its Massive IPO. *Business Insider*. Retrieved September 27, 2019, from https://www.businessinsider.com/ubers-history.

Henao, A. (2017). *Impacts of Ridesourcing—Lyft and Uber—On Transportation Including VMT, Mode Replacement, Parking, and Travel Behavior* (Ph.D. diss.). University of Colorado at Denver, Denver. Retrieved August 7, 2018, from https://search.proquest.com/docview/1899208739/abstract/8DF6D3827B14DB6PQ/1.

Interstate Transit Research Symposium (2020, December 4). Interstate Transit Research Symposium 2020: Speaker, Panelist, and Moderator Biographies. *University of Florida Transportation Institute*. Retrieved December 14, 2020, from https://www.transportation.institute.ufl.edu/wp-content/uploads/2020/12/All-Bios-updated-12-4-2020.pdf.

Iqbal, M. (2020, August 29). Uber Revenue and Usage Statistics (2020). *Business of Apps*. Retrieved September 18, 2020, from https://www.businessofapps.com/data/uber-statistics/.

Iseki, H., and Taylor, B. (2009, November 1). Not All Transfers Are Created Equal: Towards a Framework Relating Transfer Connectivity to Travel Behaviour. *Transport Reviews*, *29*(6), 777–800. https://doi.org/10.1080/01441640902811304.

Jenkins, D. (2019, February 11). *Transit Trend Analysis: 2017 National Household Travel Survey* (FHWA NHTS Report). FHWA. Retrieved September 15, 2020, from https://nhts.ornl.gov/assets/FHWA_NHTS_Report_3A_Final_021119.pdf.

Jin, S., Kong, H., and Sui, D. (2019, April 3). Uber, Public Transit, and Urban Transportation Equity: A Case Study in New York City. *Professional Geographer*, *71*(2), 315–330. https://doi.org/10.1080/00330124.2018.1531038.

Jocoy, C., and Del Casino, V. (2010, August 1). Homelessness, Travel Behavior, and the Politics of Transportation Mobilities in Long Beach, California. *Environment and Planning A: Economy and Space*, *42*(8), 1943–1963. https://doi.org/10.1068/a42341.

Jones, D. (1985). *Urban Transit Policy: An Economic and Political History*. Englewood Cliffs, NJ: Prentice-Hall.

Joshi, M., Cowan, N., Limone, O., McGuinness, K., and Rao, R. (2019, November). *E-hail Regulation in Global Cities*. Rudin Center for Transportation. Retrieved September 22, 2020, from https://wagner.nyu.edu/files/faculty/publications/RUDIN_EHAIL_REPORT.pdf.

LADOT (2018, October 31). *Mobility Data Specification: Information Briefing*. LADOT. Retrieved September 22, 2020, from https://ladot.io/wp-content/uploads/2018/12/What-is-MDS-Cities.pdf.

Lee, W. (2018, March 1). Deprived of Data, Transit Agency Tracks Tech Shuttle Buses with Cameras. *San Francisco Chronicle*. Retrieved September 22, 2020, from https://www.sfchronicle.com/business/article/Deprived-of-data-transit-agency-tracks-tech-12717891.php.

Levin, J. (2016, August). *Establishing a Transit ITS Data Exchange Specification*. Retrieved October 14, 2020, from https://web.northeastern.edu/transitdata2016/wp-content/uploads/2016/08/Levin_John.pdf.

Levy, A., and Goldwyn, E. (2020, April 24). How U.S. Public Transit Can Survive Coronavirus: Subway and Bus Systems in the U.S. Face Financial Peril as Ridership Collapses Due to Lockdowns. To Keep Transit Alive, Here's a Playbook for Immediate and Long-term Fixes. *CityLab*. Retrieved December 8, 2020, from https://www.citylab.com/perspective/2020/04/coronavirus-public-transit-ridership-federal-funding-bus-subways/610453/.

Maharawal, M., and McElroy, E. (2017, September 22). The Anti-eviction Mapping Project: Counter Mapping and Oral History toward Bay Area Housing Justice. *Annals of the American Association of Geographers*, 1–10. https://doi.org/10.1080/24694452.2017.1365583.

Mapzen Foundation (n.d.). *Transitland*. Retrieved September 17, 2020, from https://transit.land/.

Marshall, A. (2019, May 13). Why Uber Is Fighting Cities Over Data About Scooter Trips: Los Angeles and Other Cities Want Data on Shared Scooters. Uber and Lyft Claim that Could Compromise User Privacy—And Their Own Trade Secrets. *Wired*. Retrieved January 11, 2021, from https://www.wired.com/story/why-uber-fighting-cities-data-about-scooters/.

Mattson, J., Godavarthy, R., and Hough, J. (2016, January 1). Method for Estimating Statewide Transit Needs and Investment Priorities for Rural and Small Urban Areas. *Transportation Research Record: Journal of the Transportation Research Board*, *2543*(1), 25–33. https://doi.org/10.3141/2543-03.

Matute, J., D'Agostino, M., and Brown, A. (2020, February 1). *Sharing Mobility Data for Planning and Policy Research*. UC ITS. https://doi.org/10.7922/G2ZG6QHH.

McHugh, B. (2013). Pioneering Open Data Standards: The GTFS Story. In B. Goldstein and L. Dyson (eds.), *Beyond Transparency: Open Data and the Future of Civic Innovation*. Retrieved September 17, 2020, from https://beyondtransparency.org/chapters/part-2/pioneering-open-data-standards-the-gtfs-story/.

Meyerhoff, A., Micozzi, M., and Rowen, P. (1993). Running on Empty: Travel Patterns of Extremely Poor People in Los Angeles. *Transportation Research Record: Journal of the Transportation Research Board*, *1395*, 153–160. Retrieved October 5, 2020, from http://onlinepubs.trb.org/Onlinepubs/trr/1993/1395/1395-020.pdf.

MobilityData (n.d.). *General Transit Feed Specification*. Retrieved September 17, 2020, from https://gtfs.org/.

MobilityData IO (n.d.). *OpenMobilityData*. Retrieved September 17, 2020, from https://transitfeeds.com/.

Moovit (2020). Impact of Coronavirus (COVID-19) on Public Transit Usage. *Moovit*. Retrieved December 8, 2020, from https://moovitapp.com/insights/en/Moovit_Insights_Public_Transit_Index-countries.

Murphy, E. (2019, July 3). Transportation and Homelessness: A Systematic Review. *Journal of Social Distress and the Homeless*, *28*(2), 96–105. https://doi.org/10.1080/10530789.2019.1582202.

Neff, J., and Dickens, M. (2017, February). *2016 Public Transportation Fact Book* (67th Edition). APTA. Retrieved August 26, 2020, from https://www.apta.com/wp-content/uploads/Resources/resources/statistics/Documents/FactBook/2016-APTA-Fact-Book.pdf.

New York City Department of Transportation (2018). *New York City Mobility Report: June 2018*. https://www1.nyc.gov/html/dot/html/about/mobilityreport.shtml.

Newmark, G., and Nixon, H. (2018, May). *Establishing a Central Archive for Transit Passenger Data* (WP 18-04, Project 1631). Mineta Transportation Institute, San José State. Retrieved December 14, 2020, from https://transweb.sjsu.edu/sites/default/files/1631_Newmark-Central-Archive-Transit-Passenger-Data_0.pdf.

Open Mobility Foundation (2020). FAQs. *Open Mobility Foundation*. Retrieved September 22, 2020, from https://www.openmobilityfoundation.org/faq/.

ORC Macro, Kuchak, J., Eggers, F., Thompson, D., Iachan, R., Saavedra, P., and Patel, J. (2002, September 27). *The American Community Survey: Challenges and Opportunities for HUD* (C-OPC-21762, Task Order 2). HUD. Retrieved September 16, 2020, from https://www.huduser.gov/portal/publications/pdf/ACS_FINAL_REPORT.pdf.

Oregon Department of Transportation and Oregon State University (2019). *GTFS-ride*. Retrieved October 27, 2020, from https://www.gtfs-ride.org/.

Owens, H. (2019, October 30). *MDS and Transportation Data Needs* [telephone interview].

Portland Bureau of Transportation (2018). Appendix G.1: 2018 E-scooter Pilot: User Survey Results. In Portland Bureau of Transportation, *2018 E-Scooter Findings Report*. Retrieved September 22, 2020, from https://www.portlandoregon.gov/transportation/article/700916.

Poulsen, K. (2014, March 11). Reverse Wardriving: Tracking Apple and Google Commuter Buses by Their Wi-fi Clouds. *Wired*. Retrieved September 22, 2020, from https://www.wired.com/2014/03/commuter-bus/.

Rana, P., and Rundle, J. (2020, March 25). Uber Sues Los Angeles over Data-sharing Rules. *Wall Street Journal*. Retrieved September 22, 2020, from https://www.wsj.com/articles/uber-sues-los-angeles-over-data-sharing-rules-11585104223.

Rayle, L., Dai, D., Chan, N., Cervero, R., and Shaheen, S. (2016, January 1). Just a Better Taxi?: A Survey-based Comparison of Taxis, Transit, and Ridesourcing Services in San Francisco. *Transport Policy*, *45*(Supplement C), 168–178. https://doi.org/10.1016/j.tranpol.2015.10.004.

Roth, M. (2010, January 5). How Google and Portland's TriMet Set the Standard for Open Transit Data. *Streetsblog San Francisco*. Retrieved September 17, 2020, from https://sf.streetsblog.org/2010/01/05/how-google-and-portlands-trimet-set-the-standard-for-open-transit-data/.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., and Sobek, M. (2020). IPUMS USA: Version 9.0. *IPUMS USA*. https://doi.org/10.18128/D010.V10.0.

Saphores, J., Chesebro, S., Black, T., and Bricka, S. (2013, October 9). *Exploring New Directions for the National Household Travel Survey* (Transportation Research Circular No. E-C178). Transportation Research Board. https://doi.org/10.17226/21938.

Schaller, B. (2018, July 25). *The New Automobility: Lyft, Uber, and the Future of American Cities*. Schaller Consulting. Retrieved August 7, 2018, from http://www.schallerconsult.com/rideservices/automobility.pdf.

Schouten, A., Blumenberg, E., and Taylor, B. (forthcoming). *Rating the Composition: Deconstructing the Demand-side Effects on Transit Use Changes in California*.

Seo, J., Vo, T., Lee, S., Wen, F., and Choi, S. (2017, September 29). *The CTPP Workplace Data for Transportation Planning: A Systematic Review*. Presented at the 2017 Applying Census Data for Transportation Conference, Kansas City, MO. Retrieved September 16, 2020, from https://www.researchgate.net/profile/Jung_Seo4/publication/340982797_The_CTPP_Workplace_Data_for_Transportation_Planning_A_Systematic_Review/links/5ea86a52299bf18b95820e4a/The-CTPP-Workplace-Data-for-Transportation-Planning-A-Systematic-Review.pdf.

SFCTA (2018, July). *Emerging Mobility Evaluation Report: Evaluating Emerging Mobility Services and Technologies in San Francisco*. SFCTA. Retrieved September 22, 2020, from https://www.sfcta.org/sites/default/files/2019-02/Emerging%20Mobility%20Studies_11.pdf.

SFMTA (2015, October 5). *Commuter Shuttle Pilot Program: Evaluation Report*. SFMTA. https://www.sfmta.com/sites/default/files/projects/2015/Evaluation%20Report%20-%20Oct%205%202015.pdf.

SFMTA (2018, May). *Commuter Shuttle Program: 2017 Annual Status Report*. SFMTA. https://www.sfmta.com/sites/default/files/reports-and-documents/2018/05/commuter_shuttle_program_annual_status_report.pdf.

Singa, K., and Margulici, J. (2010, March). *Privately-provided Commuter Bus Services: Role in the San Francisco Bay Area Regional Transportation Network*. California Center for Innovative Transportation. Retrieved January 11, 2021, from https://merritt.cdlib.org/d/ark:%2F13030%2Fm5gx4gbt/1/CCIT201003.pdf.

Social Explorer (2020). *Social Explorer*. Retrieved October 20, 2020, from https://www.socialexplorer.com.

Taylor, B., Blumenberg, E., Wasserman, J., Garrett, M., Schouten, A., King, H., Paul, J., and Ruvolo, M. (2020, June 29). *Transit Blues in the Golden State: Analyzing Recent California Ridership Trends* (UCLA ITS-LA1908). UCLA ITS. https://doi.org/10.17610/T67W2Z.

Timmons, M. (2020, July 14). Transit Usage Dropped 75% Nationwide Amid COVID-19—And Major Cities Are Just Starting to Recover. *ValuePenguin*. Retrieved November 3, 2020, from https://www.valuepenguin.com/2020/07/transit-usage-drop-during-covid-19.

Transit App (2020). How Coronavirus Is Disrupting Public Transit. *Transit*. Retrieved November 12, 2020, from https://transitapp.com/coronavirus.

TransitCenter (2019, May 15). Accept No Substitutes for Full Trip Data from Uber and Lyft. *TransitCenter*. Retrieved September 22, 2020, from https://transitcenter.org/accept-no-substitutes-for-full-trip-data-from-uber-and-lyft/.

TransitWiki (2018, February 2). Real-time Information. *TransitWiki*. Retrieved September 17, 2020, from https://www.transitwiki.org/TransitWiki/index.php/Real-time_information.

TransitWiki (2019a, November 14). Mobility Data Specification. *TransitWiki*. Retrieved September 22, 2020, from https://www.transitwiki.org/TransitWiki/index.php/Mobility_Data_Specification.

TransitWiki (2019b, December 2). General Transit Feed Specification. *TransitWiki*. Retrieved September 17, 2020, from https://www.transitwiki.org/TransitWiki/index.php/General_Transit_Feed_Specification.

TransitWiki (2020, January 20). Publicly-accessible Public Transportation Data. *TransitWiki*. Retrieved September 17, 2020, from https://www.transitwiki.org/TransitWiki/index.php/Publicly-accessible_public_transportation_data.

Uber Technologies (2020). *Uber Movement*. Retrieved September 22, 2020, from https://movement.uber.com/.

U.S. Census Bureau (2018a). Nonemployer Statistics: 2018. *United States Census Bureau*. Retrieved September 21, 2020, from https://www.census.gov/data/datasets/2018/econ/nonemployer-statistics/2018-ns.html.

U.S. Census Bureau (2018b, January 2). Total Nonemployer Establishments. *United States Census Bureau*. Retrieved September 21, 2020, from https://www.census.gov/quickfacts/fact/note/US/NES010218.

U.S. Census Bureau (2019, September 16). Glossary. *United States Census Bureau*. Retrieved September 16, 2020, from https://www.census.gov/programs-surveys/geography/about/glossary.html.

U.S. Census Bureau (2020a). American Community Survey and Intercensal Estimates. *Data.census.gov*. Retrieved July 30, 2020, from https://data.census.gov.

U.S. Census Bureau (2020b). Longitudinal Employer-household Dynamics: Data. *United States Census Bureau*. Retrieved October 20, 2020, from https://lehd.ces.census.gov/data/.

U.S. Census Bureau (2020c). Longitudinal Employer-household Dynamics Origin-Destination Employment Statistics (LODES). *OnTheMap*. Retrieved July 31, 2020, from https://onthemap.ces.census.gov/.

U.S. Census Bureau (2020d, July 9). Nonemployer Statistics (NES). *United States Census Bureau*. Retrieved September 21, 2020, from https://www.census.gov/programs-surveys/nonemployer-statistics.html.

U.S. Census Bureau (2020e, August 18). Public Use Microdata Areas (PUMAs). *United States Census Bureau*. Retrieved September 16, 2020, from https://www.census.gov/programs-surveys/geography/guidance/geo-areas/pumas.html.

Walker, J. (2020, April 7). In a Pandemic, We're All "Transit Dependent": Now More than Ever, Public Transportation Is Not Just about Ridership. Buses, Trains, and Subways Make Urban Civilization Possible. *CityLab*. Retrieved December 8, 2020, from https://www.citylab.com/perspective/2020/04/coronavirus-public-transit-subway-bus-ridership-revenue/609556/.

Wasserman, J. (2019). *A Time and a Place for Every Rider?: Geographic and Temporal Changes in Bay Area Transit Ridership* (MURP Applied Planning Research Project). UCLA, Los Angeles. https://doi.org/10.17610/t6kw22.

Wasserman, J., Taylor, B., Blumenberg, E., Garrett, M., King, H., Paul, J., Ruvolo, M., and Schouten, A. (2020, February 26). *What's behind Recent Transit Ridership Trends in the Bay Area?: Volume II, Trends among Major Transit Operators* (UCLA ITS-LA1908). UCLA ITS. https://doi.org/10.17610/T6JP4F.

Weinberger, P. (2019, November 14). *Census Transportation Planning Products: Training Class for CTPP Five-year Data Product*.

Westat (2019, April 23). *2017 NHTS Data User Guide*. FHWA. Retrieved September 14, 2020, from https://nhts.ornl.gov/assets/NHTS2017_UsersGuide_04232019_1.pdf.

Williams, R. (2017, June 30). How Uber Works Together with Southern California's Growing Rail Network. *Medium*. Retrieved September 22, 2020, from https://medium.com/uber-under-the-hood/how-uber-works-together-with-southern-californias-growing-rail-network-fa8b1aba03b4.

World Bank. (2020). GDP Growth (Annual %)—United States. *World Bank*. Retrieved September 15, 2020, from https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=US.