Do transportation network companies increase or decrease transit ridership? Empirical evidence from San Francisco

Gregory D. Erhardt¹ · Richard Alexander Mucci¹ · Drew Cooper² · Bhargava Sana² · Mei Chen¹ · Joe Castiglione²

Accepted: 21 January 2021 / Published online: 28 February 2021
© The Author(s) 2021

Abstract

Transportation network companies (TNCs), such as Uber and Lyft, have been hypothesized to both complement and compete with public transit. Existing research on the topic is limited by a lack of detailed data on the timing and location of TNC trips. This study overcomes that limitation by using data scraped from the Application Programming Interfaces of two TNCs, combined with Automated Passenger Count data on transit use and other supporting data. Using a panel data model of the change in bus ridership in San Francisco between 2010 and 2015, and confirming the result with a separate time-series model, we find that TNCs are responsible for a net ridership decline of about 10%, offsetting net gains from other factors such as service increases and population growth. We do not find a statistically significant effect on light rail ridership. Cities and transit agencies should recognize the transit-competitive nature of TNCs as they plan, regulate and operate their transportation systems.

Keywords Public transportation · Transit ridership · Transportation network company (TNC) · Ride-hail · Ride-source · Ride-share · Mobility-as-a-Service (MaaS) · Uber · Lyft

Introduction

Uber launched in San Francisco in 2010, and Lyft followed in 2012, providing on-demand car rides booked through a smartphone app. These ride-hailing services, or transportation network companies (TNCs) resemble a taxi service, but are typically provided by “gig workers” driving their personal vehicles. TNCs have since expanded rapidly, serving more than 250 metro areas in the US and 700 globally. Most trips in the US are in nine large metro areas (Schaller 2018), and are further concentrated in the downtown cores of those metro areas (Schaller 2018; Feigon and Murphy 2018; Fehr and Peers 2019). These are the same areas where public transit ridership is highest, raising the question of how these two

✉ Gregory D. Erhardt
greg.erhardt@uky.edu

¹ Department of Civil Engineering, University of Kentucky, Lexington, KY, USA

² San Francisco County Transportation Authority, San Francisco, CA, USA

transportation services interact. Specifically, do TNCs increase or decrease public transit ridership?

TNCs could complement public transit and increase ridership in several ways. First, TNCs could serve as first and last mile feeders to transit and increase transit ridership by increasing the number of people who can access it. Second, TNCs could complement transit by providing service to locations and at times-of-day not well served by transit, making journeys without a private car more feasible (Feigon and Murphy 2016). Third, TNCs have the potential to reduce car ownership, with lower car ownership leading to with higher transit use.

Conversely, TNCs may reduce public transit use by competing for the same riders. TNCs offer door-to-door service in private vehicles at a cost lower than traditional taxis primarily in areas already well-served by transit. In this research, we use the best available data to measure the overall net effect of TNCs on transit ridership in San Francisco, considering the timing and location of TNC trips and comparing to what would happen without TNCs.

Previous research and contribution of this work

Past studies of the effect of TNCs on transit ridership produced variable results. Several studies suggest TNCs either increase transit ridership or complement public transit (Hall et al. 2018; Feigon and Murphy 2016, 2018; Rayle et al. 2016; Alemi and Rodier 2017). Other studies found an insignificant effect on transit ridership or one that varies for different years of analysis (Boisjoly et al. 2018; Malalgoda and Lim 2019; Nelson and Sadowsky 2019). Among those studies with significant results, most found that TNCs reduce bus ridership (Babar and Burtch 2020; Clewlow and Mishra 2017; Doppelt 2018; Li 2019; Henao 2017; Gehrke et al. 2018; Graehler et al. 2019) and increase commuter rail ridership (Babar and Burtch 2020; Clewlow and Mishra 2017; Doppelt 2018). The effect on light rail and subway ridership depends on the study.

Few cities have data on TNC use, which limits research and may contribute to these mixed results. Instead, several previous studies used proxies such as the start date of TNC operations to identify the correlation with transit ridership trends at the scale of metropolitan areas (Hall et al. 2018; Boisjoly et al. 2018; Malalgoda and Lim 2019; Babar and Burtch 2020; Nelson and Sadowsky 2019; Doppelt 2018; Graehler et al. 2019). Others used travel surveys to observe individual travel patterns, including who uses TNCs and what other modes they use (Rayle et al. 2016; Feigon and Murphy 2016; Henao 2017; Clewlow and Mishra 2017; Gehrke et al. 2018; Circella et al. 2019; Young and Farber 2019; Grahn et al. 2019; Dong 2020). A third set of studies takes a more theoretical approach, running models or scenarios to test possible effects rather than estimating the effects directly from the data (Kessler 2017; Alemi and Rodier 2017; Martinez and Viegas 2017).

Our research takes a different approach: we used a unique data set scraped from the Application Programming Interfaces (APIs) of two TNCs to infer where within a city TNC trips occur and at which times-of-day (Cooper et al. 2018). These detailed data are important, because the location and timing of TNC trips directly affects the modes with which they compete (Roy et al. 2020). Previous work in Toronto and at San Francisco airport leveraged detailed TNC data (Li 2019; Sturgeon 2019; Young et al. 2020), while other work that used spatially detailed transit ridership data but did not consider TNC data (Mucci and Erhardt 2018; Berrebi and Watkins 2020).

TNC trips concentrate in city centers where transit ridership is highest (Schaller 2018; Feigon and Murphy 2018; Fehr and Peers 2019), and in the biggest cities during peak travel times (Gehrke et al. 2018; San Francisco County Transportation Authority 2017). The same travelers often make trips by both transit and TNC (Feigon and Murphy 2016; Zhang and Zhang 2018). These relationships do not establish either complementarity or competition. In fact, cross-sectional revealed preference data are inherently limited in their ability to distinguish between complementary and competitive trips because both would occur in similar places and times and among the same travelers.

Instead, we must also consider what would have happened without TNCs. TNC rider surveys asking this question directly provide insight, suggesting between 15 and 42% of TNC trips otherwise would have been via transit (Gehrke et al. 2018; Henao 2017). However, they do not capture potential indirect effects, such as lifestyle and car-ownership changes, and we do not know the quality transit trips they replace. We address this issue using a before-and-after approach.

Specifically, we compared San Francisco bus and light rail ridership in 2010 before the introduction of TNCs to ridership in 2015 when TNCs were widespread. We estimated panel data regression models of transit ridership as a function of TNCs, controlling for household and employment growth, transit service changes and other relevant factors. These models represent average weekday conditions for 981 Traffic Analysis Zones (TAZs) and 4 times-of-day. To validate the panel model findings, we also estimated time-series models of bus ridership over the same period. Together, these models let us infer the net transit ridership change attributable to TNCs and to each of our control factors.

Observations

The San Francisco Municipal Transportation Agency (SFMTA) operates the Muni bus and Muni Metro light rail systems (Fig. 1), which we study here, as well as cable cars and historic streetcars. Muni serves the City and County of San Francisco, at the center of the larger 9-county Bay Area. Several other transit agencies operate in the larger Bay Area, including the Bay Area Rapid Transit (BART) and Caltrain which serve regional trips to/from San Francisco.

Between 2010 and 2015, average weekday Muni bus ridership increased 0.3% from 486,000 to 488,000 and average weekday rail ridership increased 13.1% from 150,000 to 170,000 (SFMTA n.d.). Population and employment grew, and transit service expanded during this period, so it is surprising that ridership does not grow more, particularly on the bus system. Competition with TNCs may explain the lack of ridership growth, but other factors may contribute too. In the remainder of this section, we consider possible determinants of transit ridership change and examine how they change in San Francisco.

Households and employment

From 2010 to 2015, San Francisco grew substantially: population increased 7.5%, from 809,000 to 869,000 and the number of household grew 5.8% from 338,000 to 357,000 (U.S. Census Bureau n.d.). Employment grew 24.8%, from 555,000 to 692,000 (U.S. Bureau of Labor Statistics n.d.). Population and employment increases should transit ridership (Gomez-Ibanez 1996; Kain and Liu 1999; Taylor et al. 2009).

Fig. 1 San Francisco transit map (Howard 2009)

Because employment grew faster than population, more workers commute from elsewhere in the Bay Area. This changing jobs-housing balance may moderate the effect of growing employment on Muni, which primarily serves San Francisco residents (Blumenberg et al. 2020).

Socioeconomic factors

Socioeconomic changes may either increase or decrease transit ridership. Low-income riders often comprise a large share of transit riders (Manville et al. 2018; Grahn et al. 2019) because those riders often do not own a car, and because they may be less willing or able to pay for parking and other costs associated with driving. San Francisco became wealthier between 2010 and 2015 as inflation-adjusted median household income grew from \$70,800 to \$87,448 (U.S. Census Bureau n.d.). However, income growth may relate to higher employment or more car ownership, so the net income effect could go in either direction (Gomez-Ibanez 1996; Taylor et al. 2009).

Transit riders are younger on average than the general population (Grahn et al. 2019), but in San Francisco, the adult population aged over the analysis period, with a smaller share of individuals in the 20–44 age group and more in the 45–64 and 65+ age groups (MTC n.d.).

Transit service

Increased transit service increases transit ridership (Gomez-Ibanez 1996; Kain and Liu 1999; Taylor et al. 2009). Muni provided 11% more bus service miles over this period, but 5.5% fewer rail service miles (GTFS n.d.). While the physical configuration of light rail stations remained the same, the Muni Forward initiative included skip-stop Rapid service and other operational improvements along high-ridership bus corridors. Muni also missed fewer due to driver absence or mechanical issues, delivering 96.9% of scheduled service in 2010 and 99.6% in 2015 (SFMTA n.d.). The overall customer satisfaction with Muni service increased over this time period as well: among adult San Francisco residents who had used Muni in the previous 6 months, 66% rated the service as good or excellent in 2015 compared to 52% in 2010 (Corey and Galanis Research 2016).

Transit fare

Transit fare increases reduce transit ridership (Pratt 2013). In 2010, the cash fare on Muni was \$2.00 per trip. It increased to \$2.25 per trip in 2014, with associated increases to the cost of monthly passes and other fare options (SFMTA n.d.). When adjusted for inflation, this reflects a 3.6% increase in the cash fare.

Gas price

Lower gas prices reduce the cost of driving and may reduce transit use (Nowak and Savage 2013). Gas in San Francisco cost \$3.20 per gallon in 2010 and \$2.90 per gallon in 2015, a 16% inflation-adjusted decline (Energy Information Administration n.d.).

Car ownership

Travelers who do not own a car are an important market of transit riders (Taylor et al. 2009; Grahn et al. 2019). In contrast to the idea TNCs reduce car ownership, car ownership increased slightly in San Francisco: 30.6% of households owned 0 vehicles in 2010, compared to 30.8% in 2015 (U.S. Census Bureau n.d.).

Transfers from regional transit

Because San Francisco employment grew faster than households, more people commuted from elsewhere in the Bay Area. Many of these travelers arrived on BART or Caltrain, then transferred to Muni reach their destination.

Bike share

Bike share systems proliferated over this period and may reduce bus ridership (Campbell and Brakewood 2017; Graehler et al. 2019). Bay Area Bike Share launched in August 2013, with docks in San Francisco, San Jose, Palo Alto, and Mountain View. However, the bike share system is smaller than in other large cities: in October 2015 it served an average of only 1,200 trips per day in San Francisco (BABS n.d.), equivalent to only 0.25% of bus ridership and 0.8% of rail ridership.

TNCs

The *TNCs Today* report profiled TNC activity in San Francisco in fall 2016 (San Francisco County Transportation Authority 2017). On average weekdays, 170,000 TNC vehicle trips both started and ended in San Francisco, representing 15% of the total intra-San Francisco vehicle trips and 12 times the number of taxi trips. Trips averaged 2.6 miles for the portion with a passenger. They concentrated in the densest part of San Francisco, and during the morning and evening peak travel hours. About 40% of TNC vehicle miles traveled (VMT) is travel without a passenger, or deadheading (Fehr and Peers 2019; California Air Resources Board 2019).

TNCs carry an average of 1.55 passengers per trip, with little difference in occupancy between pooled and non-pooled trips (California Air Resources Board 2019). This occupancy implies about 260,000 TNC person trips in San Francisco on an average weekday in 2016, compared to 488,000 Muni bus boardings and 170,000 Muni light rail boardings.

We did not directly observe TNC growth in San Francisco after 2016, but can look to New York for an analogy, where open data from the Taxi and Limousine Commission provides a time-series of the number of TNC trips per day (TLC n.d.). Table 1 summarizes these values for values in October of each year, showing rapid growth until 2019 when New York implemented a cap on the number of ride-hail vehicles allowed to operate in the city. By 2018, TNCs generated 12–14% of VMT in San Francisco (Fehr and Peers 2019). This is about double the VMT share attributed to TNCs estimated in 2016 by the *TNCs Today* report, and consistent with growth rate observed in New York.

Table 1 Growth of TNC trips in New York

	TNC trips per day	Year-over-year growth (%)
2015	163,178	N/A
2016	310,749	90.4
2017	468,318	50.7
2018	680,433	45.3
2019	682,635	0.3

Methodology and data

Because we could not hold the above factors constant in an experimental setting, we instead controlled for them statistically. Doing so let us to compare the actual ridership change before and after the introduction of TNCs to the expected change without TNCs. We hypothesize that if TNCs increase transit ridership, then actual ridership would be higher than predicted by the remaining factors. Conversely, if TNCs decrease transit ridership, then actual ridership would be lower than otherwise expected. Further, we expect the difference between actual and expected ridership in either direction to be greatest in locations, and at times with more TNC activity.

To test these hypotheses, we estimated a fixed-effects panel data model of zonal-level transit ridership by time-of-day. The estimated coefficients give the elasticity of ridership with respect to changes each variable, including TNCs. We applied those elasticities to the observed change in each variable to determine the ridership change attributable to each variable.

While the panel models capture the changes both by location and by time-of-day, they do not separate the effect of factors that change uniformly between those two periods, such as gas price and transit fare. To break out those terms and check the panel model results, we also estimated a regression on time-series model of system-wide Muni bus ridership.

We assess the consistency of both models when presenting our results. By specifying both models with transit ridership as the dependent variable, we assume the emergence of TNCs leads to changes in transit ridership, which is reasonable because TNCs do not appear to enter a market in an effort to capitalize on pre-existing transit ridership trends (Hall et al. 2018). Both models rely on the same data, aggregated to different resolutions.

Data

We compiled data about each of the categories discussed above and assembled transit accessibility measures. We prepared the same data in two different formats to support the two model structures. For the panel models, each entity is a unique combination of TAZ, time-of-day and mode, observed for two time periods. For the time-series models, each observation is for San Francisco as a whole, in each month from June 2009 through June 2016. Table 2 summarizes the data and their sources, which we discuss next.

Transit ridership

Bus ridership data comes from Automated Passenger Counters (APCs) installed on a sample of about 20% of the fleet and integrated with an Automated Vehicle Location (AVL) system (SFMTA n.d.). To calculate the total ridership we linked the APC data the GTFS data and expanded using the ratio of scheduled (bus) trips to observed (Erhardt et al. 2017). We adjusted the ridership to account for the percent of the scheduled service delivered. This process produced monthly average weekday bus boardings and alightings at each bus stop, on each route, by time-of-day, starting in June 2009.

In 2016, SFMTA began transitioning to a new APC system. To avoid the potential for inconsistent measurements between the old and new equipment, we stop the data series in June 2016. The panel models use the average of October and November of each year (2010 and 2015), aggregated to the TAZ level for all TAZs in San Francisco. The time-series models use the total system ridership by month. Because the panel model

Table 2 Data fields and sources

Category	Variable	Description	Source
Transit ridership	Bus ridership	Muni bus ridership, average of boardings and alightings	Expanded APC/AVL
	Rail ridership	Muni light rail ridership, average of boardings and alightings	Manual counts
Households and employment	Households	Households	ACS/SF-CHAMP
	Population	Population	ACS/SF-CHAMP
	Employment	Number of employees at work location	QCEW/SF-CHAMP
	Workers	Total employed residents at home location	ACS/SF-CHAMP
	Employment rate	Share of population that is employed (workers/population)	ACS
Socio-economic factors	Share of households by income quartile	Share of households in each TAZ that are in the 1st, 2nd, 3rd and 4th income quartiles. Quartiles are defined based on distributions in the 9-county Bay Area	Metropolitan Transportation Commission
	Median household income	Median household income in 2010 dollars	ACS
	Share of population by age group	Share of population in each TAZ that is age: 0–4, 5–19, 20–44, 45–64, and 65 +	Metropolitan Transportation Commission
	Service miles	Service miles delivered	GTFS/SFMTA
Transit service	Number of routes	Number of bus/rail routes	GTFS
	Stops	Number of bus/rail stops	GTFS
	Trip stops	Number of bus/rail trip stops. A trip stop is one transit vehicle making a stop	GTFS
	Competing trip stops	The number of competing bus/rail trip stops in nearby TAZs	GTFS
	Runspeed	Average vehicle running speed (mph). Excludes time when vehicle is stopped, and doors are open. Bus value measured from APC/AVL data, rail value as reported in GTFS	Expanded APC/AVL
On-time performance	On-time performance	The share of stops that are made along a transit segment that are no more than 1 min early and 5 min late	Expanded APC/AVL

Table 2 (continued)

Category	Variable	Description	Source
Accessibility	Accessibility	The households plus employment that can be reached within 30 min on bus or on rail	GTFS/SFCTA
Transit fare	Cash fare	Cash fare in 2010 dollars	SFMTA press releases/news reports
Gas price	Gas Price	The cost per gallon for gasoline in 2010 dollars	Energy Information Administration
Car ownership	Share of 0-Car Households	The share of households that have 0 vehicles	ACS
Transfers from regional transit	BART ridership	Average of BART boardings and alightings	Faregate counts
	Caltrain ridership	Average of Caltrain boardings and alightings	Ridership counts
TNCs	TNC pick-ups and drop-offs	Average of TNC pick-ups and drop-offs in TAZ	TNC API data-scraping
	Years since TNC launch	The number of years since TNCs started operating	TNC Websites

excludes a small number of stops outside San Francisco, the total ridership differs slightly between the two tabulations.

Light rail vehicles do not have APC equipment. Instead SFMTA provided manual counts of the light rail boardings and alightings at each rail station, on each route, by time-of-day, for average weekday conditions in 2009 and 2016. They also provided route-level ridership estimates for 2010 and 2015 which we used to scale the station-level counts to those years. Because the data were not collected monthly, we could not estimate time-series models of rail ridership.

To ensure that ridership on a run was counted in the same time period, we defined times-of-day according to the transit vehicles departure from its first stop. We counted ridership on vehicles departing their first stop prior to 3 a.m. with the overnight period for the day before and counted ridership on vehicles departing their first stop 3–6 a.m. with the AM peak ridership because many of those runs overlap into the 6–9 a.m. period.

Households and employment

We to households, population and workers the American Community Survey (ACS) (U.S. Census Bureau n.d.) and interpolated these annual data to monthly county level estimates. We take employment from the Quarterly Census on Employment and Wages (QCEW), which provides monthly estimates of employment by industry for every county in the United States (U.S. Bureau of Labor Statistics n.d.). For the panel model, we use the same TAZ-level socio-economic data that are input to San Francisco's travel demand model, SF-CHAMP. These data are derived from several of the same data sources, but also consider Planning Department data and regional estimates from the Metropolitan Transportation Commission (MTC).

Because people may walk a few blocks before boarding a transit vehicle, transit ridership in a TAZ is influenced not only by households and employment in that TAZ, but also those in nearby TAZs. Therefore, we applied a smoothing process that considers the values in TAZs within walking distance using a decay function (Zorn et al. 2011). We calculated the smoothed values as:

$$X'_i = \sum_{j \in J} X_j * e^{-11.8 * D_{ij}}$$

where X'_i is the smoothed value of a variable X at TAZ i ; X_j is the unsmoothed value for TAZ j ; D_{ij} is the distance from i to j in miles; and J is the set of all TAZs within 0.75 miles. We subsequently refer to them as referencing “Nearby TAZs”.

Socioeconomic factors

We obtained socioeconomic data from two sources: the ACS and MTC. The ACS provided annual county-level estimates of the median household income. MTC provided TAZ level estimates of the number of households by income quartile and the population by age group. We apply the smoothing function described above to the TAZ-level estimates for use in the panel models.

Transit service

We assembled transit service data using the same software tool used to compile bus ridership and adjusted for the percent of service delivered (Erhardt et al. 2017). We derived runspeed and on-time performance from the APC and AVL data and averaged across transit vehicles. We calculated rail runspeed from the scheduled speed, and did not have on-time performance data for rail.

For the time-series models, we aggregated all values to system-wide totals. For the panel models, we aggregated to TAZs based on the stops in each TAZ. Because we modeled the ridership in each TAZ, we also tracked the service in the same TAZ without the smoothing employed for household and employment data. The exception is for the competing trip stops metric, which employs the same distance decay function while excluding any stops in the same TAZ. This means that it measures the amount of other nearby transit service, allowing us to capture the possibility that ridership in a TAZ decreases if service in a neighboring TAZ improves, causing people to walk there instead.

Accessibility

Transit ridership is a function not only of the households and employment near a transit stop, but also of the access provided to households and employment at the far end of a transit route. Providing such accessibility can be viewed as a core purpose of the transportation system, and an important element of transportation planning (Levinson et al. 2017). Therefore, we defined an accessibility metric as the number of jobs plus households that can be reached within 30 min by boarding a transit stop within the specified TAZ and within the relevant time period.

We calculated accessibility separately for bus and rail using the UrbanAccess package (Blanchard and Waddell 2017), which uses an OpenStreetMap walk network and GTFS transit specification to calculate the fastest transit paths through the network accounting for walk time, wait time and in-vehicle time. For bus, we used UrbanAccess to calculate the travel time from every Muni bus stop to every Census block. We required the first boarding to be at the bus stop specified and allowed transfers to other bus routes. Then we aggregated the stop-to-Census block matrix to a TAZ-to-TAZ matrix, retaining the minimum travel time as we aggregated. Then we summed the total households and total jobs for all TAZs that could be reached from a given origin within 30 min. We applied an equivalent process to calculate rail accessibility, but for rail we allowed transfers to other rail lines and to Muni bus routes. The first boarding must be on rail because we related it directly to the station-level boarding counts, but bus could be used as a feeder mode at the other end similar to how SF-CHAMP treats rail in mode choice.

Transit fare

To reflect the cost of transit, we measured the full cash fare (SFMTA n.d.), adjusted for inflation to constant 2010 dollars.

Gas price

Gas price data measured the average cost of a gallon of gasoline in San Francisco (Energy Information Administration n.d.) and adjusted for inflation to constant 2010 dollars.

Car ownership

We measured car ownership as the share of households owning zero vehicles. In the time-series models it we take it from the 1-year ACS data. In the panel models we take it from the 5-year ACS data for 2006–2010 and 2011–2015. We calculate both zero-vehicle households and total households at the TAZ level, then applied the distance decay function to smooth them, and used the smoothed values to re-calculate the share of zero-care households in nearby TAZs.

Transfers from regional transit

BART and Caltrain ridership reports provided the average weekday boardings and alightings at each station by time-of-day (BART n.d.; Caltrain n.d.). We averaged the boardings and alightings, allocated them to TAZs, and applied the smoothing function because people may transfer from those stations to Muni stops in nearby TAZs.

TNCs

California Public Utilities Commission (CPUC) regulates TNCs in California and collects anonymized data for every TNC ride request and trip made in the state (Dobush 2019). However, the CPUC declined to share those data with other local public agencies, including the San Francisco County Transportation Authority (SFCTA), stating that the sharing of such data is “not in the public interest.” Our requests directly to TNCs for anonymized and aggregated data to support this research were met with an offer to share detailed trip-level data specifically for trips to and from rail stations. Instead, we relied on a data set of TNC use that was independently scraped from the Application Programming Interfaces (APIs) of the two largest TNCs. For a 6-week period in fall 2016, these data provide second-by-second traces of available TNC vehicles which are used to infer passenger-serving trips, pick-up locations, drop-off locations, and both in-service and out-of-service volumes (Cooper et al. 2018). The same data were previously used to profile TNC activity in San Francisco and to analyze the effects of TNCs on traffic congestion (San Francisco County Transportation Authority 2017; Erhardt et al. 2019). They can be visualized and downloaded from <http://tncstoday.sfcta.org/>.

For this research, we tabulated the average of TNC pick-ups and drop-offs by TAZ and time-of-day for average weekday conditions and the smoothing function, as above. This provided a snapshot of TNC use in fall 2016. Because TNCs did not start operating until October 2010, we assume that TNC use is negligible in 2010 and assign zero to the TNC volume fields for that year. There is a discrepancy between the year of TNC data collection (2016) and the second year of TAZ-level transit data (2015) because by fall of 2016 the transition to new APCs is substantial, leading to concerns about transit data consistency. For this reason, we analyzed fall 2015 conditions, and assumed that while TNC use

is likely higher in 2016 than 2015 that the spatial pattern is likely to be similar. While the coefficient on TNC use may be lower than if we had less measured TNC use in 2015, we expect the resulting effect on transit ridership to be correct. While we expect the cost of TNCs to be important in determining how many people switch from transit we do not have data on TNC fares in San Francisco.

For the time-series models, we do not have a continuous measure of TNC use, but we expect that it grows over time. Therefore, we use the number of years since the introduction of TNCs (in decimal form) as a proxy, consistent with previous research (Graehler et al. 2019).

Panel data models

To estimate the sensitivity of ridership to changes in TNCs and each control variable, we estimated a fixed-effects panel data regression model (Greene 2003), a method used elsewhere to study changes in transit ridership (Li 2019; Berrebi and Watkins 2020). Panel data refers to data where the same entities are observed at multiple points in time. In our case, an entity is a unique combination of TAZ, time-of-day and mode, each observed in both 2010 and 2015. The model takes the form:

$$\log(Y_{i,t}) = \beta * X_{i,t} + TE_t + FE_i + \epsilon_{i,t}$$

where $Y_{i,t}$ is the transit ridership for entity i at time period t , β is a vector of estimated model coefficients, $X_{i,t}$ is the set of exogenous attributes of entity i at time t , TE_t is an estimated time effect that represents a constant change across all entities between the 2 years, FE_i is an estimated fixed-effect on each entity, and $\epsilon_{i,t}$ is the error term.

The fixed-effect term ensures that omitted variables do not bias the remaining coefficient estimates if they do not change over time, such as a TAZ's location. Conversely, the time-effect absorbs omitted variables that change uniformly over time, such as gas price. The strength of the fixed effects model is estimating the effects of variables that change at different rates on different entities. For example, we know that the change in TNC pick-ups and drop-offs is not uniformly distributed in space or by time-of-day, so the fixed-effects model is ideal to capture their influence on transit ridership.

For most variables, we applied a log transformation to $X_{i,t}$ such that we can interpret the β coefficients as elasticities. We applied the elasticities to calculate the net contribution of each variable to the change in ridership, holding all other variables constant.

Model estimation

Our data include 981 TAZs, 4 times-of-day: 3:00–8:59 a.m., 9:00 a.m.–3:59 p.m., 4:00–6:59 p.m., and 7:00 p.m. to 2:59 a.m., and two modes: bus and rail, for 7848 possible entities. We observe each in two time periods: fall 2010 and fall 2015, for 15,696 total records. We exclude 8215 records with no transit stops, 35 with inconsistent accessibility data, 8 with inconsistent rail ridership counts between the 2 years, and 80 in which there are stops in 1 year but not the other. This results in 7358 total records or 3679 matched entities. We used the python linearmodels package to estimate a model where the dependent variable is the log of bus or rail ridership.

Table 3 shows the panel model estimation results. The columns provide a description of the exogenous variables, any transformations applied to those variables, any smoothing,

Table 3 Estimation results for panel models of Muni bus and light rail ridership

<i>Model structure</i>					
Dependent variable	Log (bus or rail ridership)				
Model form	Fixed-effects panel data regression				
Description	Transformation	Smoothing	Coefficient	SE	T-Stat
<i>Exogenous variables</i>					
<i>Accessibility</i>					
Accessibility (jobs + HH w/i 30 min by transit)	Log	TAZ only	0.1324	0.0445	2.98
<i>Transit service</i>					
Number of routes serving TAZ	Log	TAZ only	0.3146	0.0477	6.60
Trip stops: bus	Log	TAZ only	0.9816	0.0441	22.25
Trip stops: light rail	Log	TAZ only	0.2242	0.0736	3.05
On-time performance (share of vehicles + 1 to – 5 min)		TAZ only	0.3317	0.0574	5.77
Competing bus trip stops	Log	Nearby TAZs	– 0.4935	0.0869	– 5.68
<i>Transfers from regional transit</i>					
BART + Caltrain, average of Ons + offs	Log	Nearby TAZs	0.3121	0.0771	4.05
<i>Income</i>					
Share of households in lowest income quartile		Nearby TAZs	2.1997	0.6250	3.52
Share of households in highest income quartile		Nearby TAZs	– 0.7763	0.3944	– 1.97
<i>TNCs</i>					
Average of TNC ons + offs: bus	Log	Nearby TAZs	– 0.0254	0.0047	– 5.42
Average of TNC ons + offs: rail	Log	Nearby TAZs	0.0111	0.0086	1.29
<i>Unexplained change</i>					
Systemwide changes (time effect)		Everywhere	0.0139	0.0167	0.83
<i>Model statistics</i>					
Number of entities			3679		
Number of time periods			2		
Log likelihood			1909.9		
R-squared between groups			0.853		
R-squared within groups			0.326		

the coefficient and the t-statistic. We grouped the exogenous variables by category, corresponding to the observations above. The r-squared between groups (measured cross-sectionally) is 0.853 and the r-squared within groups (measured over time) is 0.326.

Accessibility is positive and significant, as are several measures of transit service: the number of routes serving a TAZ, trip stops and on-time performance. The trip stops variable is segmented by mode, and the results show that ridership is more sensitive to bus trip stops than to rail trip stops, consistent with previous literature (Pratt 2013). The number of routes

serving a TAZ changes only when the structure of the transit system changes. The number of competing bus trip stops is negative and significant, which means that if bus service is added to a nearby TAZ, then ridership in the current TAZ will reduce because some travelers may divert to the nearby TAZ instead. Increased boardings and alightings at BART and Caltrain stations are positively correlated with Muni ridership because people may transfer to and from those systems. Proximity to low-income households is associated with higher transit ridership, while proximity to high-income households is associated with lower ridership. The time-effect is a constant applied only to the year 2015. The estimated value of 0.0139 suggests that there is a net 1.39% ridership increase beyond what the other variables explain.

One TNC coefficient applies only to bus observations, and a second applies only to rail observations. The effect of TNCs on bus ridership is negative and significant, suggesting that more TNCs lead to lower bus ridership. The effect of TNCs on rail ridership is positive, but statistically insignificant.

We could not estimate a significant and logical coefficient on car ownership or age, we tested several options for capturing transit service changes, and we found that using accessibility fit better than including households and employment separately.

Model application

Next, we applied the estimated coefficients to calculate each variables' contribution transit ridership change. We applied the model separately for each entity, then sum across entities. We tabulated these results separately for bus and rail, as reported in Tables 4 and 5. For

Table 4 Contributions to Muni bus ridership change derived from panel models

	Coefficient	Average value		Ridership change	
		2010	2015	Absolute	Percent
<i>Accessibility</i>					
Accessibility (jobs + HH w/i 30 min by transit)	0.1324	489,186	574,305	9692	2.0
<i>Transit service</i>					
number of routes serving TAZ	0.3146	2.6	2.8	4833	1.0
Trip stops: bus	0.9816	110.7	119.2	35,416	7.4
Ontime performance (share of vehicles + 1 to – 5 min)	0.3317	0.62	0.62	– 670	– 0.1
Competing bus trip stops	– 0.4935	22.8	26.9	– 7758	– 1.6
<i>Transfers from regional transit</i>					
BART + Caltrain, average of ons + offs	0.3121	82.3	87.0	5923	1.2
<i>Income</i>					
Share of households in lowest income quartile	2.1997	0.31	0.32	3600	0.8
Share of households in highest income quartile	– 0.7763	0.26	0.25	966	0.2
<i>TNCs</i>					
Average of TNC Ons + offs: bus	– 0.0254	0.0	50.7	– 41,244	– 8.6
<i>Unexplained change</i>					
Systemwide changes	0.0139	0.0	1.0	6720	1.4
Other factors				– 15,835	– 3.3
Total ridership		479,248	480,892	1644	0.3

Table 5 Contributions to Muni light rail ridership change derived from panel models

Description	Coefficient	Average value		Ridership change	
		2,010.00	2,015.00	Absolute	Percent
<i>Accessibility</i>					
Accessibility (Jobs + HH w/i 30 min by transit)	0.1324	433,166	512,703	3182	2.1
<i>Transit service</i>					
Number of routes serving TAZ	0.3146	1.5	1.6	2236	1.5
Trip stops: light rail	0.2242	74.0	69.1	– 1803	– 1.2
<i>Transfers from regional transit</i>					
BART + Caltrain, average of ons + offs	0.3121	0.0	0.0	3057	2.0
<i>Income</i>					
Share of households in lowest income quartile	2.1997	0.29	0.30	744	0.5
Share of households in highest income quartile	– 0.7763	0.26	0.26	292	0.2
<i>TNCs</i>					
Average of TNC ons + offs: rail	0.0111	0.0	24.6	5736	3.8
<i>Unexplained change</i>					
Systemwide changes	0.0139	0.0	1.0	2113	1.4
Other factors				4149	2.8
Total ridership		150,681	170,387	19,706	13.1

each variable, the table shows a description of the variable, the associated coefficient, the average value of that variable in 2010 and 2015, and the resulting net ridership change. The “unexplained change” category includes both the systemwide changes from the time-effect constant, and any remaining difference between observed and modeled ridership.

In Table 4, we observe that bus ridership increase by 0.3% from 2010 to 2015. The net ridership loss to TNCs offsets net gains from better accessibility, service expansion, more regional transfers, and income changes. In Table 5, we observe that rail ridership increases by 13.1% from 2010 to 2015. On average, the number of rail trip stops decreases over this period, but changes to the route structure, such as running short-turn trains in the core corridor offset these overall reductions. The results show that several sources contribute to rail ridership gains, with no single factor dominating.

Examining the panel model results in more detail, Table 6 shows the contributions to bus and rail ridership change by time-of-day. In these tables, we combined the results by category, rather than reporting individual variables. We observe that bus ridership increases in the AM and PM peaks, while decreasing in the mid-day and night periods, and rail ridership increases the most in the AM and PM peak periods. Most variables contribute consistently across times-of-day, except for overnight service cuts to rail that explain the lack of growth in that time period. The unexplained change also differs by time-of-day, with higher values in the peaks and lower values overnight.

Figure 2 shows how the bus results vary by geographic district. Each sub-plot shows the net effect for variables in that category. The effect of TNCs on bus ridership is most concentrated in Downtown and the surrounding districts, where TNCs are most prevalent and they reduce bus ridership by up to 13%. The TNC effect is progressively less as radiating outwards from the core towards the less dense and more residential neighborhoods. Overall bus ridership increases notably in two areas. In SoMa, a combination of service increases, accessibility growth and more transfers from regional transit lead to more ridership. In

Table 6 Contributions to Muni bus and rail ridership change by time-of-day from panel model: 2010 to 2015

Time-of-day	Households & employment (%)	Transit service (%)	Transfers from regional transit (%)	Income (%)	TNCs (%)	Unexplained change (%)	Total (%)
<i>Bus ridership change due to</i>							
3:00–8:59 a.m.	2.3	6.4	1.7	1.0	– 7.3	1.2	5.3
9:00 a.m.–3:59 p.m.	1.9	7.1	0.6	1.0	– 9.1	– 2.9	– 1.3
4:00–6:59 p.m.	2.1	5.1	1.6	0.9	– 8.3	– 1.1	0.3
7:00 p.m.–2:59 a.m.	1.8	8.3	2.1	1.0	– 9.8	– 5.4	– 2.0
Total	2.0	6.6	1.2	1.0	– 8.6	– 1.9	0.3
<i>Rail ridership change due to</i>							
3:00–8:59 a.m.	2.5	0.5	2.7	0.8	3.3	5.6	15.3
9:00 a.m.–3:59 p.m.	1.9	0.7	1.3	0.8	4.0	1.9	10.6
4:00–6:59 p.m.	2.1	3.5	2.5	0.5	3.7	8.2	20.5
7:00 p.m.–2:59 a.m.	2.1	– 9.1	2.3	0.7	4.5	0.2	0.8
Total	2.1	0.3	2.0	0.7	3.8	4.2	13.1

Fig. 2 Contributions to Muni bus ridership by district from panel model: 2010 to 2015

Bayshore, there is a substantial unexplained ridership increase. We investigated this result and found that the ridership gains are concentrated in zones served by the 9R San Bruno Rapid Bus route. Between 2010 and 2015, SFMTA converted this route from local/limited service to rapid service and implemented several operational changes to improve travel time and reliability. It may be that those changes a positive ridership beyond what the model captures, which would be encouraging because SFMTA has since converted other routes as part of the Muni Forward project.

Figure 3 shows how the contributions to rail ridership change vary by geographic district. Rail ridership gains due to TNCs are concentrated in the downtown area where TNC use is highest and lowest in the outlying areas. The unexplained change shows a similar pattern, but the magnitude of the unexplained change is greater. Overall, we observe that

Fig. 3 Contributions to Muni light rail ridership by district from panel model: 2010 to 2015

Muni light rail ridership grows the most in SoMa and Downtown. Specifically, ridership in the Market Street tunnel, where the light rail lines converge into a subway, grows by 55% between 2010 and 2015, and the model only explains a portion of this strong growth. Conversely, the Hills District and Outer Mission lose ridership, which may be partly due to construction at the Balboa Park BART station affecting nearby light rail stations.

Time-series models

The time-series model serves as a check on the panel model using a different data structure and a different model form. Each observation is the average weekday Muni bus ridership in San Francisco in 1 month. We did not model rail because we do not have monthly rail ridership data. The time-series model is a regression model based on data with a 12-month difference. This means that we are modeling the change in bus ridership from the same month 1-year prior as a function of the change in any exogenous variables from the same month 1-year prior. Excluded factors that affect transit ridership but are constant fall out of the model, and the 12-month differencing absorbs seasonality. The model takes the form:

$$\log(Y_t) - \log(Y_{t-12}) = \beta * (X_t - X_{t-12}) + \epsilon_t$$

where Y_t is the Muni bus ridership on an average weekday in month t , β is a vector of estimated coefficients, X_t is a vector of exogenous variables and ϵ_t is the error term of the regression model. In most cases we take the log of the exogenous variables as well ($\log(X_t) - \log(X_{t-12})$), for some variables we also test the untransformed variable ($X_t - X_{t-12}$). We estimated the models in R and confirmed that no serial auto-correlation exists for the errors via a Box-Pierce test. As with the panel model, we applied the estimated coefficients to the change in each exogenous variable to calculate each variable's contribution transit ridership change. We refer to any residual change as “unexplained”.

Model estimation

Table 7 shows the model estimation results for the regression on time-series model, including the variables, any transformations, the coefficients and t-statistics. The model achieves an R-squared of 0.563. With only 85 time periods after differencing, some variables are statistically insignificant, but we retain them in the model specification when their coefficients are logical. Specifically, ridership increases with more households, higher gas prices, and a higher share of 0-car households and retaining these three variables slightly improves the model fit. The remaining variables are significant with logical signs, including positive coefficients for the employment rate and service miles, and a negative coefficient on cash fare.

The years since TNC launch serves as a proxy for the TNC volume, which we expect to increase as the TNC market share grows. Its coefficient is negative and significant. The value of -0.0226 suggests that bus ridership is net 2.26% lower for each year after the TNC launch, and this effect compounds over time.

We tested and decided against other specifications for the time-series models. These included several specifications aimed at better capturing the economic growth that occurred over this period. Including the employment rate (the share of the population that is employed) produced a model with a better fit than employment, possibly because many

Table 7 Estimation results for time-series model of Muni bus ridership

<i>Model structure</i>				
Model form	Regression on time-series data			
Differencing	12-month 1st difference			
Dependent variable	$\log(Y_t) - \log(Y_{t-12})$			
	Where Y_t = average weekday Muni bus ridership in month t			
Description	Transformation	Coefficient	SE	T-Stat
<i>Exogenous variables (X_t)</i>				
Households and employment				
Households	$\log(X_t) - \log(X_{t-12})$	0.2545	0.3991	0.64
Employment rate (workers/population)	$\log(X_t) - \log(X_{t-12})$	1.1118	0.3707	3.00
Transit service				
Service miles, 1000 s	$\log(X_t) - \log(X_{t-12})$	0.5541	0.0957	5.79
Transit fare				
Cash fare (2010 USD)	$\log(X_t) - \log(X_{t-12})$	- 0.2730	0.0684	- 3.99
Gas price				
Gas price (2010 USD)	$\log(X_t) - \log(X_{t-12})$	0.0282	0.0348	0.81
Car ownership				
Share of 0-Car households	$\log(X_t) - \log(X_{t-12})$	0.1346	0.1764	0.76
TNCs				
Years since TNC launch	$X_t - X_{t-12}$	- 0.0226	0.0100	- 2.26
<i>Model statistics</i>				
Number of time periods		85		
Log-likelihood		163.4		
AIC		- 310.79		
R-squared		0.563		
Box–Pierce test p value		0.19		

employees in San Francisco commute in from neighboring counties, while Muni serves primarily trips within the county. We tried substituting the years since the Great Recession for the years since TNC launch and including the median household income. In both cases the variables are not significant at a 95% level and the increases in these terms are correlated with the years since TNC launch causing the TNC coefficient to be more negative. Given that these economic variables did not clearly improve the model, we opted for the specification shown in Table 7 for its more conservative estimate of the TNC effect.

Model application

We applied the estimated coefficients to the change in the value of each variable, as shown in Table 8. For consistency with the panel models, we used an average of October and November 2010 and an average of October and November 2015. The total ridership is slightly higher than in the panel models because Table 8 includes ridership for some Muni stops that fall across the border with San Mateo County.

Table 8 Contributions to Muni bus ridership change derived from time-series models

Description	Coefficient	Value		Ridership change	
		2010	2015	Absolute	Percent
<i>Households and employment</i>					
Households	0.2545	337,964	357,437	6985	1.4
Employment rate (workers/population)	1.1118	0.55	0.58	29,925	6.2
<i>Transit service</i>					
Service miles, 1000 s	0.5541	49.6	56.1	34,680	7.1
<i>Transit fare</i>					
Cash fare (2010 USD)	− 0.2730	1.99	2.06	− 4654	− 1.0
<i>Gas price</i>					
Gas price (2010 USD)	0.0282	3.19	2.65	− 2499	− 0.5
<i>Car ownership</i>					
Share of 0-Car households	0.1346	0.31	0.31	609	0.1
<i>TNCs</i>					
Years since TNC launch	− 0.0226	0.0	5.0	− 52,538	− 10.8
<i>Unexplained change</i>					
Other factors				− 10,879	− 2.2
Total ridership		486,431	488,060	1629	0.3

Table 8 shows that the main factors contributing to net bus ridership increases are more service miles, a higher employment rate, and more households. Together, we would expect these three factors to result in a 14.7% bus ridership increase. However, bus ridership only increases by 0.3%, so other factors offset the growth, the biggest of which is TNCs, which result in a net 10.8% ridership decrease. Higher fares, lower gas prices, and other unexplained factors contribute smaller net decreases.

The bivariate area plots in Fig. 4 show each category's contribution to bus ridership change, similar to plots used previously to examine transit ridership trends (Erhardt 2016). The black line shows the observed bus ridership by month, with seasonality contributing to the peaks and valleys. When changes to a variable result in a net transit ridership decrease, we shaded the area red. When changes to a variable result in a net transit ridership increase, we shaded the area green. For both cases, we measured the variable change relative to October 2010 and calculated the ridership change using the estimated model coefficient. If not for changes to that value relative to October 2010, we would expect ridership to follow the top of the red area or the bottom of the green area instead.

In the first plot, we observe that households and employment increase bus ridership starting in late 2012 as economic growth starts to increase. In the next panel, the red areas show that service cuts result in ridership losses in mid-2010 and in spring 2012, but service increases increase ridership substantially starting in January 2015. Changes in the fare, fuel price and car ownership have only minor ridership effects. The bottom right panel shows that TNCs decrease bus ridership noticeably starting in about 2012 and growing over time. If not for the loss of ridership to TNCs, bus ridership would continue to grow throughout this period.

Fig. 4 Contributions to Muni bus ridership change by month from time-series model

Discussion

This research examined the question of whether TNCs increase or decrease transit ridership, considering San Francisco between 2010 and 2015. Having analyzed the contributions to transit ridership change using two models, we consider the consistency of the results and discuss those results in a broader context.

Summary and context

Table 9 shows the contributions to bus and rail ridership change for both the panel and the time-series models. The models show that TNCs result in net bus ridership decreases of 8.6% to 10.8%. Both models show that transit service increases result in about 7% more bus ridership. The effect of households and employment in the time-series model is greater

Table 9 Contributions to Muni ridership change: 2010 to 2015

Category	Bus		Rail
	Panel model estimate (%)	Time-series model estimate (%)	Panel model estimate (%)
Households and employment/accessibility	2.0	7.6	2.1
Transit service	6.6	7.1	0.3
Transit fare	n/a	– 1.0	n/a
Gas price	n/a	– 0.5	n/a
Car ownership	n/a	0.1	n/a
Transfers from regional transit	1.2	n/a	2.0
Income	1.0	n/a	0.7
TNCs	– 8.6	– 10.8	3.8
Unexplained change	– 1.9	– 2.2	4.2
Total ridership	0.3	0.3	13.1

than the effect of accessibility in the panel model. The remaining factors have a modest effect on bus ridership in either direction. Of the 13.1% rail ridership increase, we estimate that 2.1% is attributable to higher accessibility, 0.3% to increased rail service, and 2.0% from more regional transit transfers. The model suggests that 3.8% of the ridership increase is attributable to TNCs, although this estimate is statistically insignificant. There is something driving the rail ridership increase beyond what we can measure with this model: 4.2% of the growth remains unexplained.

To evaluate whether this estimate of bus ridership loss to TNCs, we compared it to a back-of-the-envelope calculation:

- In 2016 there were 170,000 TNC vehicle trips with both ends in San Francisco (San Francisco County Transportation Authority 2017),
- In New York in 2015, TNC trips were 52% of their 2016 value (TLC n.d.),
- TNCs carried 1.55 passengers per vehicle (California Air Resources Board 2019),
- 30% of TNC passengers in the Bay Area would have taken transit if TNCs were not available (SANDAG 2019),
- 75% of transit boardings within San Francisco are on bus, and
- Muni averaged 1.25 bus boardings per linked trip (SFMTA 2017).

These numbers suggest a 2015 net bus ridership loss to TNCs of 39,000 bus boardings ($170,000 * 0.52 * 1.55 * 0.30 * 0.75 * 1.25$), or about an 8% net ridership loss. An equivalent calculation for light rail would suggest a ridership loss due to TNCs.

Our results reinforce the majority of the existing literature in the conclusion that TNCs decrease bus ridership (Babar and Burtch 2020; Clewlow and Mishra 2017; Doppelt 2018; Li 2019; Henao 2017; Gehrke et al. 2018; Graehler, Mucci, and Erhardt 2019), and add little insight to the various findings on the effect on rail ridership. The difference may relate to trip length. Bus trips and TNC trips both average 2–3 miles in length, so may compete more directly. On such a short trip a TNC may substitute for the whole trip to avoid a transfer. Conversely, rail systems often serve much longer trips, so may reach more potential ridership with better first and last mile access. Several studies show slightly higher commuter rail ridership (Babar and Burtch 2020; Clewlow and Mishra 2017; Doppelt 2018),

with little agreement on the effect on light rail and subway. In part, this may reflect the diversity of rail systems and the markets they serve.

Several of the existing studies use a panel data approach to analyze multiple cities at the level of a city or metropolitan area (Hall et al. 2018; Boisjoly et al. 2018; Malalgoda and Lim 2019; Babar and Burtch 2020; Nelson and Sadowsky 2019; Doppelt 2018; Graehler et al. 2019). Such an approach is complementary to our own analysis, but does not account for where and when TNC trips occur within a city, or how many there are. In particular, treating TNC presence as a binary variable is limiting because we know that the number of TNC trips grow over time—that is why our time-series model uses years since TNCs started instead of a binary flag. One multi-city model (Graehler et al. 2019) took a similar approach and found that TNCs reduce bus ridership by 1.7% per year (8.1% by 2015) compared to our estimate of 2.2% per year.

Studies in two other locations used detailed TNC data. A Toronto study found TNCs reduce bus ridership, but were a much smaller share of total trips than in San Francisco (Li 2019). Another study found TNCs reduced BART ridership to San Francisco airport (Sturgeon 2019).

The remaining studies that suggest that TNCs increase transit ridership do so based on the observations that TNC trips are made close to transit stops, or that the same individuals make both TNC trips and transit trips (Feigon and Murphy 2016, 2018; Rayle et al. 2016; Zhang and Zhang 2018). That is flawed logic, because the same observations—TNC and transit trips proximate in location and made by the same people—may also lead to higher substitution. To distinguish between these possible conclusions, we must consider the change as we have done here, and cannot rely on cross-sectional data.

Limitations

The strength of this work is examining the effect of TNCs and transit ridership at a detailed spatial level, and by time-of-day within a city. However, it studies a single city in a single year, so we consider if it can be generalized. Total bus ridership in the US has been declining since 2012, and rail ridership has been declining since 2014, with the decline including both large and small cities, suggesting the reason for that decline may be shared broadly (NTD n.d.). Because TNCs arrived in San Francisco earlier than in many other cities, it may be a leading indicator of the effect elsewhere. We may expect TNCs to further decrease transit ridership as TNC use grows. The time-series model coefficients suggest TNCs reduce bus ridership in San Francisco by 2.3% per year, implying a net ridership loss of 19% in 2019 versus 10% in 2015. While we lack data on TNC growth in San Francisco, New York data show TNC trips increase fourfold between 2015 and 2018 before leveling off (Table 1). If San Francisco followed a similar trajectory, we may reasonably expect a larger effect. Since 2015, some transit agencies have begun partnering with TNCs (Curtis et al. 2019; Schwieterman et al. 2018), which may further affect ridership. Two other limitations of this study should be noted.

Previous research has shown employment is among the most important drivers of transit ridership, but after testing several model specifications we found the employment effect to be modest (Gomez-Ibanez 1996; Kain and Liu 1999; Chen et al. 2011). This result may be due to a jobs-housing imbalance, or it may be because employment growth in San Francisco is in the same parts of San Francisco that TNCs are most active, making it difficult to estimate the magnitude of their offsetting effects. If the models captured a larger positive employment effect, then the TNC effect might be more negative.

The panel model only explains about two-thirds of the observed light rail ridership growth. Light rail ridership increases the most in the Market Street tunnel, where all the lines converge into a subway through downtown, versus operating in mixed traffic elsewhere. The unexplained increase may relate to car traffic restrictions on Market Street implemented in 2015 or to worsening traffic congestion giving the exclusive guideway a greater travel time advantage.

Future research

Future research could address the limitations of this study and further enhance our understanding of the relationship between transit ridership and TNCs and appropriate policy responses. First, researchers can study the effect in more cities and operating contexts. For example, a similar spatially detailed study of transit ridership decline in 4 US cities could be extended to explicitly consider TNCs (Berrebi and Watkins 2020). However, such a study would require detailed TNC trip data, highlighting the importance of making those data available to cities and to researchers. Second, analyzing a richer data set, such as the smartphone-enabled household travel survey recently collected in California would provide further insight, such as the characteristics of TNC users and how TNCs fit in larger trip chains (SANDAG 2019). Finally, cities are considering strategies in response to these trends that include congestion pricing, TNC user fees, dedicated bus lanes, and free or reduced transit fares. Researchers should evaluate the effectiveness, costs and benefits of those possible responses.

Conclusions

Our panel model analysis showed bus ridership in 2015 is 8.6% lower than expected given changes in control factors, and this net difference is highest in the locations and at times-of-day with more TNC pick-ups and drop-offs. Both support the hypothesis that TNCs decrease transit ridership. The results of a separate regression on time-series model show bus ridership in 2015 is 10.8% lower than expected given control factor changes. These net ridership losses due to competition with TNCs offset the net gains from household and employment growth combined with transit service expansions. Given this evidence, we conclude that in 2015, TNCs reduced bus ridership in San Francisco by about 10%. We do not find a statistically significant relationship between TNCs and Muni light rail ridership.

It is logical that TNCs are more competitive with than complementary to bus transportation because both tend to serve relatively short trips in city centers where parking is difficult, and many travelers arrive without a personal car. While some authors have suggested that this co-location of TNC and transit trips is evidence of complementarity, our results show that net bus ridership losses are most severe in locations with high TNC activity. While individual travelers who switch from bus to TNC may benefit, such a shift may lead to negative externalities such as more traffic congestion for other road users, including those who remain on the bus (Erhardt et al. 2019). This outcome raises important equity questions because those who remain on the bus may be lower income or otherwise have different socio-economic characteristics.

While these results are specific to San Francisco, cities throughout the US have experienced bus ridership declines of 12–18% rail ridership declines since TNCs have emerged (Watkins et al. 2020). Our research adds to a growing consensus that TNCs reduce bus

ridership. These results provide better understanding of the competitive relationship between TNCs and transit that is important as cities aim to provide a transportation system that is efficient and equitable.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11116-021-10178-4>.

Acknowledgements Thank you to Michael Batty, Elsa Arcaute, Simon Berrebi, Kari Watkins and the staff at the San Francisco Municipal Transportation Authority for discussing this work. Kaya Tollas assisted with calculating accessibility metrics.

Author contributions The time-series analysis presented here originated in Greg Erhardt's Ph.D. thesis (Erhardt 2016) and the panel analysis presented here originated in Alex Mucci's Master's thesis (Mucci 2017). Both have since been expanded and revised. Drew Cooper processed the TNC data. Bhargava Sana provided the visualization. Mei Chen and Joe Castiglione provided direction and guidance throughout, including review of the data, analysis, and writing, with Joe Castiglione initiating the project. All authors contributed intellectually to the design and interpretation of the research.

Funding This work was funded internally by the San Francisco County Transportation Authority.

Code availability Scripts for model estimation are available upon request.

Availability of data and materials The estimation data for both models are included with this manuscript as supplementary materials.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest or competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alemi, F., Rodier, C.: Simulation of Ridesourcing Using Agent-Based Demand and Supply Regional Models: Potential Market Demand for First-Mile Transit Travel and Reduction in Vehicle Miles Traveled in the San Francisco Bay Area. UC Davis Research Report. National Center for Sustainable Transportation: UC Davis Research Report (2017). <https://escholarship.org/uc/item/3h1550wm>
- Babar, Y., Burtch, G.: Examining the heterogeneous impact of ride-hailing services on public transit use. *Inf. Syst. Res.* (2020). <https://doi.org/10.1287/isre.2019.0917>
- BABS. n.d.: Bay Area Bike Share Open Trip Data. Accessed 13 June 2017. <http://www.bayareabikeshare.com/open-data>
- BART. n.d.: BART Monthly Ridership Reports. Accessed 16 Oct 2020. <https://www.bart.gov/about/reports/ridership>
- Berrebi, S.J., Watkins, K.E.: Who's ditching the bus? *Transp. Res. Part A Policy Pract.* **136**(June), 21–34 (2020). <https://doi.org/10.1016/j.tra.2020.02.016>
- Blanchard, S.D., Waddell, P.: Urban access: generalized methodology for measuring regional accessibility with an integrated pedestrian and transit network. *Transp. Res. Rec. J. Transp. Res. Board* **2653**(1), 35–44 (2017). <https://doi.org/10.3141/2653-05>

- Blumenberg, E., Garrett, M., King, H., Paul, J., Ruvolo, M., Schouten, A., Taylor, B.D., Wasserman, J.: What's behind Recent Transit Ridership Trends in the Bay Area? UCLA Institute for Transportation Studies (2020). <https://doi.org/10.17610/T6JP4F>
- Boisjoly, G., Grisé, E., Maguire, M., Veillette, M.-P., Deboosere, R., Berrebi, E., El-Geneidy, A.: Invest in the ride: a 14 year longitudinal analysis of the determinants of public transport ridership in 25 North American cities. *Transp. Res. Part A Policy Pract.* **116**(October), 434–445 (2018). <https://doi.org/10.1016/j.tra.2018.07.005>
- California Air Resources Board: SB 1014 Clean Miles Standard 2018 Base-Year Emissions Inventory Report (2019). <https://ww2.arb.ca.gov/resources/documents/2018-base-year-emissions-inventory-report>
- Caltrain. n.d.: Caltrain Annual Ridership Counts. Accessed 16 Oct 2020. <https://www.caltrain.com/about/statsandreports/Ridership.html>
- Campbell, K.B., Brakewood, C.: Sharing riders: how bikesharing impacts bus ridership in New York City. *Transp. Res. Part A Policy Pract.* **100**(June), 264–282 (2017). <https://doi.org/10.1016/j.tra.2017.04.017>
- Chen, C., Varley, D., Chen, J.: What affects transit ridership? A dynamic analysis involving multiple factors, lags and asymmetric behaviour. *Urban Stud.* **48**(9), 1893–1908 (2011). <https://doi.org/10.1177/0042098010379280>
- Circella, G., Lee, Y., Alemi, F.: Exploring the Relationships among Travel Multimodality, Driving Behavior, Use of Ridehailing and Energy Consumption. UC Davis Research Report. UC Davis (2019). <https://escholarship.org/uc/item/31v7z2vf>
- Clewlow, R.R., Mishra, G.S.: Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States. UC Davis Research Report UCD-ITS-RR-17-07. UC Davis (2017)
- Cooper, D., Castiglione, J., Mislove, A., Wilson, C.: Profiling TNC Activity Using Big Data. In: Transportation Research Board Annual Meeting (2018)
- Corey, C., Galanis Research: S.F. Municipal Transportation Agency Ridership Survey. San Francisco, California: San Francisco Municipal Transportation Agency (2016). <https://www.sfmta.com/reports/2016-muni-customer-satisfaction-survey>
- Curtis, T., Meg, M., Carmen, C., David, P., Dan, B.: Partnerships Between Transit Agencies and Transportation Network Companies (TNCs). Transit Cooperative Research Program (2019). file:///C:/Users/Alex's%20Laptop/Downloads/25576.pdf
- Dobush, G.: Uber has troves of data on how people navigate cities. Urban planners have begged, pleaded, and gone to court for access. Will they ever get it?" *Marker*, September 9, 2019 (2019). <https://marker.medium.com/ubers-real-advantage-is-data-e54984ff524c>
- Dong, X.: Trade Uber for the bus?: An investigation of individual willingness to use ride-hail versus transit. *J. Am. Plan. Assoc.* **86**(2), 222–235 (2020). <https://doi.org/10.1080/01944363.2019.1687318>
- Doppelt, L.: Need a Ride? Uber Can Take You (Away from Public Transportation). MPP Thesis, Washington, D.C.: Georgetown University (2018)
- Erhardt, G.D.: Fusion of Large Continuously Collected Data Sources: Understanding Travel Demand Trends and Measuring Transport Project Impacts. PhD Thesis, University College London (UCL) (2016)
- Erhardt, G.D., Lock, O., Arcaute, E., Batty, M.: A big data mashing tool for measuring transit system performance. In: Thakuriah, P., Tilahun, N., Zellner, M. (eds) *Seeing Cities Through Big Data*. Springer Geography, pp. 257–78. Springer, Berlin (2017). https://doi.org/10.1007/978-3-319-40902-3_15
- Erhardt, G.D., Roy, S., Cooper, D., Sana, B., Chen, M., Castiglione, J.: Do transportation network companies decrease or increase congestion? *Sci. Adv.* **5**(5), 670 (2019). <https://doi.org/10.1126/sciadv.aau2670>
- Energy Information Administration. n.d.: Gasoline and Diesel Fuel Update. Accessed 5 Nov 2014. <http://www.eia.gov/petroleum/gasdiesel/>
- Fehr & Peers: Estimated TNC Share of VMT in Six US Metropolitan Regions (2019). <https://drive.google.com/file/d/1FIUskVkj9lsAnWJQ6kLhAhNoVLjfFdx3/view>
- Feigon, S., Murphy, C.: Shared Mobility and the Transformation of Public Transit. TCRP 188. Washington, D.C.: Transportation Research Board. <https://doi.org/10.17226/23578> (2016)
- Feigon, S., Murphy, C.: Broadening Understanding of the Interplay Between Public Transit, Shared Mobility, and Personal Automobiles. TCRP 195. Washington, D.C.: Transportation Research Board (2018). <https://doi.org/10.17226/24996>
- Gehrke, S., Felix, A., Reardon, T.: Fare Choices Survey of Ride-Hailing Passengers in Metro Boston. Metropolitan Area Planning Council (2018). <https://www.mapc.org/farechoices/>
- Gomez-Ibanez, J.A.: Big-city transit, ridership, deficits, and politics. *J. Am. Plann. Assoc.* **62**(1), 30 (1996)
- Graehler, M., Mucci, R.A., Erhardt, G.D.: Understanding the recent transit ridership decline in major US cities: service cuts or emerging modes? In: Transportation Research Board Annual Meeting. Washington (2019)

- Grahn, R., Harper, C.D., Hendrickson, C., Qian, Z., Scott Matthews, H.: Socioeconomic and usage characteristics of transportation network company (TNC) riders. *Transportation* (2019). <https://doi.org/10.1007/s11116-019-09989-3>
- Greene, W.H.: *Econometric Analysis*, 5th edn. Prentice Hall, Upper Saddle River (2003)
- GTFS. n.d.: GTFS Data Exchange—San Francisco Municipal Transportation Agency. Accessed 15 July 2014. <http://www.gtfs-data-exchange.com/agency/san-francisco-municipal-transportation-agency/>
- Hall, J.D., Palsson, C., Price, J.: Is Uber a substitute or complement for public transit? *J. Urban Econ.* **108**(November), 36–50 (2018). <https://doi.org/10.1016/j.jue.2018.09.003>
- Henao, A.: Impacts of Ridesourcing—Lyft and Uber—on Transportation Including VMT, Mode Replacement, Parking, and Travel Behavior.” PhD Thesis, University of Colorado at Denver (2017). <http://search.proquest.com/openview/5486ff6cc229889a3cdf2df1cd3993cb/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Howard, D.: San Francisco Transit Map. Dannyman.Toldme.Com (blog) (2009). <https://dannyman.toldme.com/2009/11/03/san-francisco-transit-map/>
- Kain, J.F., Liu, Z.: Secrets of success: assessing the large increases in transit ridership achieved by Houston and San Diego transit providers. *Transp. Res. Part A Policy Pract.* **33**(7), 601–624 (1999)
- Kessler, M.L.: How Transportation Network Companies Could Replace Public Transportation in the United States. MS Thesis, University of South Florida (2017)
- Levinson, D.M., Marshall, W., Axhausen, K.: *Elements of Access*. Network Design Lab (2017). Isbn Softcover: 9781389067617. <https://www.blurb.com/b/8469315-elements-of-access>
- Li, W.: Empirical Analysis of the Relationship Between Ride-Hailing and Public Transit in Toronto. MS Thesis, University of Toronto (2019)
- Malalagoda, N., Lim, S.H.: Do transportation network companies reduce public transit use in the U.S.? *Transp. Res. Part A Policy Practice* **130**(December), 351–372 (2019). <https://doi.org/10.1016/j.tra.2019.09.051>
- Manville, M., Taylor, B.D., Blumenberg, E.: Falling transit ridership: California and Southern California. Prepared for the Southern California Association of Governments. UCLA Institute for Transportation Studies (2018). https://www.scag.ca.gov/Documents/ITS_SCAG_Transit_Ridership.pdf
- Martinez, L.M., Viegas, J.M.: Assessing the impacts of deploying a shared self-driving urban mobility system: an agent-based model applied to the city of Lisbon, Portugal. *Int. J. Transp. Sci. Technol.* **6**(1), 13–27 (2017). <https://doi.org/10.1016/j.ijtst.2017.05.005>
- MTC. n.d.: Metropolitan Transportation Commission Transportation Analysis Zone Data
- Mucci, R.: Transportation Network Companies: Influencers of Transit Ridership Trends. MS Thesis, University of Kentucky (2017). https://uknowledge.uky.edu/ce_etds/61
- Mucci, R.A., Erhardt, G.D.: Evaluating the ability of transit direct ridership models to forecast medium-term ridership changes: evidence from San Francisco. *Transp. Res. Rec.* (2018). <https://doi.org/10.1177/0361198118758632>
- Nelson, E., Sadowsky, N.: Estimating the impact of ride-hailing app company entry on public transportation use in Major US Urban areas. *The B.E. J. Econ. Anal. Policy* (2019). <https://doi.org/10.1515/bejeap-2018-0151>
- Nowak, W.P., Savage, I.: The cross elasticity between gasoline prices and transit use: evidence from Chicago. *Transp. Policy* **29**, 38–45 (2013). <https://doi.org/10.1016/j.tranpol.2013.03.002>
- NTD. n.d.: National Transit Database. Federal Transit Administration. Accessed 1 Aug 2018. <https://www.transit.dot.gov/ntd>
- Pratt, R.H.: Traveler Response to Transportation System Changes. Chapter 1-Introduction. Transit Cooperative Research Program TCRP Report 95. Washington, D.C.: Transportation Research Board (2013). <http://www.trb.org/Publications/TCRPReport95.aspx>
- Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S.: Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transp. Policy* **45**(1), 168–178 (2016)
- Roy, S., Cooper, D., Mucci, A., Sana, B., Chen, M., Castiglione, J., Erhardt, G.D.: Why is traffic congestion getting worse? A decomposition of the contributors to growing congestion in San Francisco. *Case Stud. Transp. Policy* (2020). <https://doi.org/10.1016/j.cstp.2020.09.008>
- San Francisco County Transportation Authority. TNCs Today: A Profile of San Francisco Transportation Network Company Activity (2017). http://www.sfcta.org/sites/default/files/content/Planning/TNCs/TNCs_Today_061317.pdf
- SANDAG. California Statewide Ridehail Survey. San Diego Association of Governments, Southern California Association of Governments and Metropolitan Transportation Commission (2019). <https://www.sandag.org/index.asp?classid=13&projectid=540&fuseaction=projects.detail>
- Schaller, B.: *The New Automobility: Lyft, Uber and the Future of American Cities*. Schaller Consulting, Brooklyn (2018)

- Schwieterman, J.P., Livingston, M., Van Der Slot, S.: Partners in Transit: A Review of Partnerships Between Transportation Network Companies and Public Agencies in the United States. Depaul University (2018). https://las.depaul.edu/centers-and-institutes/chaddick-institute-for-metropolitan-development/research-and-publications/Documents/Partners%20in%20Transit_Live1.pdf
- SFMTA. Systemwide On-Board Survey. San Francisco Municipal Transportation Agency (2017). <https://mtc.ca.gov/our-work/plans-projects/other-plans/regional-onboard-survey-program>
- SFMTA. n.d. San Francisco Municipal Transportation Agency Ridership and Service Data
- Sturgeon, L.R.: The Impact of Transportation Network Companies on Public Transit: A Case Study at the San Francisco International Airport. BA Thesis, Scripps College (2019). https://scholarship.claremont.edu/cgi/viewcontent.cgi?article=2287&context=scripps_theses
- Taylor, B.D., Miller, D., Iseki, H., Fink, C.: Nature and/or nurture? Analyzing the determinants of transit ridership across US Urbanized areas. *Transp. Res. Part A Policy Pract.* **43**(1), 60–77 (2009). <https://doi.org/10.1016/j.tra.2008.06.007>
- TLC. n.d.: New York City Taxi & Limousine Commission Trip Record Data. Accessed 16 Oct 2020. <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>
- U.S. Bureau of Labor Statistics. n.d.: Quarterly Census of Employment and Wages
- U.S. Census Bureau. n.d.: American Community Survey. Accessed 5 Nov 2014. <http://www.census.gov/acs/www/>
- Watkins, K. E., Berrebi, S., Erhardt, G., Brakewood, C.: Recent decline in public transportation ridership: analysis, causes, responses: interim report: Phase I." TCRP Project A-43. Transit Cooperative Research Program. Transportation Research Board (2020)
- Young, M., Farber, S.: The who, why, and when of Uber and other ride-hailing trips: an examination of a large sample household travel survey. *Transp. Res. Part A Policy Pract.* **119**(January), 383–392 (2019). <https://doi.org/10.1016/j.tra.2018.11.018>
- Young, M., Allen, J., Farber, S.: Measuring when Uber behaves as a substitute or supplement to transit: an examination of travel-time differences in Toronto. *J. Transp. Geogr.* **82**(January), 102629 (2020). <https://doi.org/10.1016/j.jtrangeo.2019.102629>
- Zhang, Y., Zhang, Y.: Exploring the relationship between ridesharing and public transit use in the United States. *Int. J. Environ. Res. Public Health* **15**(8), 1763 (2018). <https://doi.org/10.3390/ijerph15081763>
- Zorn, L., Sall, E., Bent, Z.: Exploring parking pricing for congestion management using the SFCTA activity-based regional pricing model (2011). <https://trid.trb.org/View/1093088?ajax=1>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Gregory D. Erhardt is an assistant professor of Civil Engineering at the University of Kentucky. He has worked for 20 years developing and applying travel demand models in government, consulting and academia, both in the US and in the UK. He specializes in developing more sophisticated modeling tools to forecast the effects of proposed transportation projects, analyzing new and emerging data sources to understand the effects of past projects, and applying both to improve policy and planning decisions.

Richard Alexander Mucci is a transportation engineer at the Kentucky Transportation Cabinet and pursuing a Civil Engineering PhD degree at the University of Kentucky. He has a masters degree from the University of Kentucky and has worked as a researcher in the university's Transport lab for the last 5 years. He specializes in analyzing the effects of new and emerging transportation modes.

Drew Cooper is a Senior Transportation Modeler in the Technology, Data & Analysis (TD&A) division of the San Francisco County Transportation Authority (SFCTA). He led the data processing necessary to infer TNC trips from the scraped data that form the basis of this study and several others. He holds a bachelors degree in mathematics from the University of Texas and a master's degree in transportation engineering from the University of Washington.

Bhargava Sana is a Senior Transportation Modeler in the Technology, Data & Analysis (TD&A) division of the San Francisco County Transportation Authority (SFCTA). He has over 10 years experience in travel demand modeling. He specializes in the development and calibration of activity-based travel demand models. Bhargava enjoys building software tools for making day-to-day modeling tasks efficient and has lately focused on developing web-based interactive visualization tools for transportation data.

Mei Chen is a professor of civil engineering at University of Kentucky. Her research is in the area of transportation planning and modeling. She has led a number of studies involving integrating emerging data into safety, incident management and various planning applications.

Joe Castiglione is the Deputy Director for Technology, Data & Analysis at the San Francisco County Transportation Authority. He has been developing, refining and applying advanced transportation supply and demand models for over 20 years, including the first activity-based travel model used extensively in practice in the US, and has subsequently led efforts to integrate these activity-based models with dynamic traffic models. Recently, he has been leading the Transportation Authority's research into the effects of Transportation Network Companies in San Francisco on congestion, transit ridership, and equity.