Jackson Zeng
Seth Reichert
Purva Kapshikar

Bay Area Transit Access and Disconnectivity

Introduction

The San Francisco Bay Area comprises nine counties which surround a large tidal estuary. Due to the unique geographical challenge of traversing the San Francisco Bay, the region is well connected by numerous freeways and eight bridges. Although the region is served by several transit agencies, transit use only makes up a fraction of the travel mode distribution. However, certain trip purposes, such as commuting to work — and commuting to San Francisco, in particular — are significantly done using transit.

Much of the development and improvement in transit in the Bay Area has occurred after 2000. In 2000, about 58% of all intraregional trips were done by driving alone, 24% by carpooling, 13% by taking transit, 2% by walking, and 2% by biking (MORPACE International, Inc., 2013). Of course, at this time, various transit agencies were fairly disconnected in terms of geographical station and stop location and fare payment.

We believe that Bay Area residents' mode choice, and large reliance on driving, has been influenced by the disconnectivity of the transit network. We hope to evaluate the 2000 Bay Area transit network and how changes we manually introduce in this network may influence mode choice. This may also reaffirm or impugn decisions made by these transit agencies in changing service or coverage since 2000.

Our primary dataset is the Bay Area Travel Survey (BATS), a regional household travel survey periodically conducted in the San Francisco Bay Area by the Metropolitan Transportation Commission (MTC). The BATS from the 2000 calendar year requested respondents to provide information on all their travel activities within a two-day period, with over 15,000 households participating. The survey had six alternatives in its choice set: drive alone, shared ride, walk, bike, walk to transit, and drive to transit. It also tracked respondents' mode choices for trips specified by origin and destination Transportation Analysis Zones (TAZs), a MTC-defined geographic unit used in MTC travel models.

Literature Review

Driving is clearly the dominant mode of transportation across the United States. Many people have gradually become more reliant on their personal vehicles over time, as having a car actually encourages people to alter their activity patterns and choices to increase or optimize use of their car (Gould et al., 1998)

Driving also offers convenience for certain groups of the population. For instance, women with children are more likely to drive to work, regardless of their income levels (Rosenbloom & Burns, 1994). This is because many working mothers rely on their personal vehicles to carry out several domestic or childcare responsibilities as well (Rosenbloom & Burns, 1994). This may be especially relevant to the Bay Area context as the employment rate of women is nearly 20 percentage points higher than the national average (Castro et al., 2022; U.S. Department of Labor, 2022). Those who prioritize or need to prioritize flexibility in their travels also are more likely to be car users (Şimşekoğlu et al., 2015). Driving may also be tied to obstacles and preferences that are culture-specific; for example, immigrant women are less likely to have licenses, and many instead rely on carpooling (Blumenberg & Smart, 2010). Transit use is also greater amongst new immigrants than native-born Americans, though new immigrants still prefer to carpool more than using public transit, likely to leverage social and family ties to help with adjusting to life in a new country (Blumenberg & Smart, 2010).

Although driving is the most common mode choice, improvements to transit quality and service could attract more riders. For example, one survey from RailCorp, an Australian rail company, suggested that, in exchange for a ten percent increase in train cleanliness, transit users were willing to pay up to about $0.07 USD per minute or have a 0.26-minute increase in their onboard travel time (Litman, 2008). Some other changes respondents were willing to spend more on (in fare or time) included ease of train boarding, quietness, and improved on-train announcements (Litman, 2008). The time and effort that is spent waiting for a transit vehicle increases a traveler's perceived burden of travel (Iseki & Taylor, 2010). Travelers tend to value frequent, reliable service more than the physical characteristics of or amenities offered by a transit facility — especially when considering their own personal safety (Iseki 2010). Transit ridership may also increase from better street network connectivity to transit (Ewing & Cervero, 2010).

Transit use may increase due to increased coverage as well. In one study, a 10 percent increase in population covered by transit increased transit use by 5.9 percent and a 10 percent increase in job accessibility by transit increased transit use by 6.6 percent (Zuo et al., 2020). Accessibility to transit, also known as "to-transit accessibility," has conventionally been estimated based on characteristics such as the transit service availability within a neighborhood or travel time and distance to transit stops and stations (Yang et al., 2020). Models from previous studies have used a quarter mile from bus stops and a half mile from rail stations as the standard maximum walking distance for a stop or station to be deemed accessible (El-Geneidy et al., 2014; Guerra et al., 2012; Horner & Murray, 2004). Some municipal reports and policies used a 1-mile buffer as the threshold for accessibility (Ratner & Goetz, 2013). Meanwhile, results from other research have also suggested that riders may even be willing to travel as far as two miles to access their nearest rail station (Houston, Boarnet, et al., 2014). Due to the variability, some studies have even chosen to evaluate accessibility using several different distance thresholds (Kwoka, 2015).

We were curious about how transit access in the Bay Area might have influenced travel mode preferences and transit agency decisions to expand service, yet we could not find existing literature about this subject. As a result, we hope to explore how better access to fixed guideway transit may influence Bay Area residents' choice to take transit instead of drive.

Methods

Data Sources

Our primary dataset for model development was the BATS 2000. We supplemented this data using MTC Plan Bay Area 2040 forecasts for the year 2005 and joined the BATS data with MTC land use, demographic, and employment data (published for TAZs) to provide additional covariates for our utility specifications. We also calculated the percentage of the area in TAZs that had walkable access to fixed guideway transit service in 2000 as detailed below.

In 2000, the three major urban transit agencies in the Bay Area based on unlinked passenger trips were San Francisco Municipal Railway (MUNI), Bay Area Rapid Transit (BART), and Alameda-Contra Costa Transit (AC Transit) (Bureau of Transportation Statistics, 2012). However, due to the challenges of obtaining 2000 data on specific routes, coverage, and frequency, we have chosen to focus on the impact that having access to fixed guideway service (typically rail, streetcar, trolley, and BRT) has on mode choice. Transit stop locations for these services for all agencies in the Bay Area under 2022 conditions are provided by MTC (Metropolitan Transportation Commission, 2022). These provided stop locations were manually restricted to 2000 conditions for all agencies using a variety of archived transit maps, and timelines of service expansions. For example, a map of the BART system in 2000 can be seen in Figure 1 below. The final selected stops, shown by agency are shown in Figure 2.

Figure 1: 2000 BART System Map

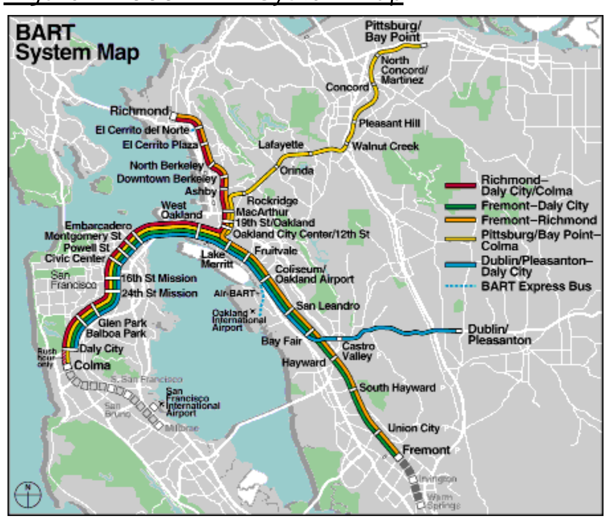
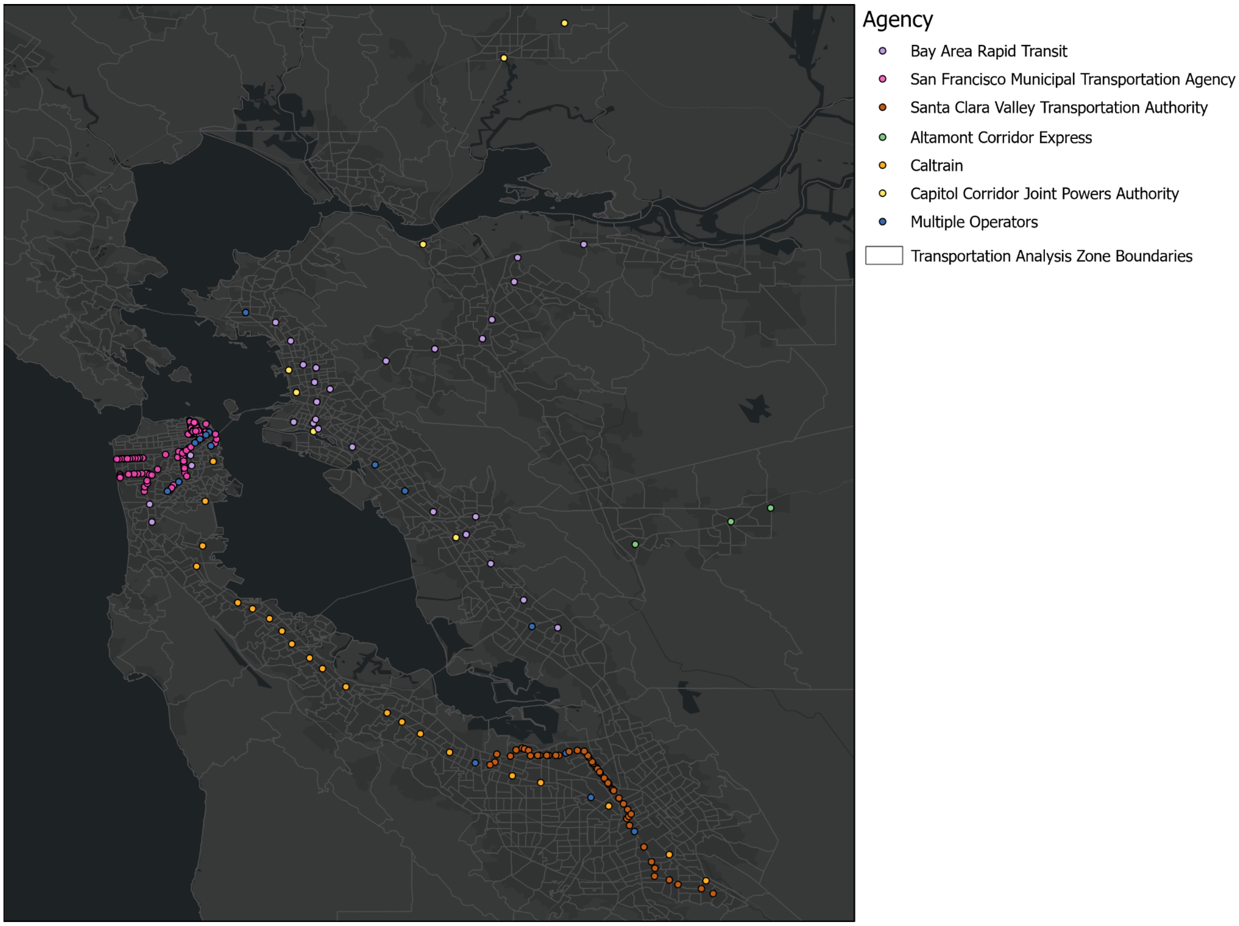


Figure 2: Fixed Guideway Transit Stops, 2000 Conditions



Data Processing

We used the transit stops shown above to evaluate TAZ transit access by creating ½ mile buffers around each stop (approximately a 10 minute walk, and consistent with previous research as described above) and calculating the percentage area of each TAZ that was within these buffers.

We further processed a number of other variables from the MTC Bay Area Plan 2040, and while not all were used in our model specifications, we processed them in order to have greater flexibility in developing our models. The additional variables were processed as follows under the categories of Land Use, Employment, and Household Income.

- Land Use
  - Area Type: Core, Central Business District, Urban Business, Urban, Suburban, or Rural. Coded as dummy variables.
  - Primarily Residential vs Commercial Uses: Coded as 1 if there were more acres of residential than commercial land, 0 if not.
  - Primarily Multifamily Housing vs Single Family: Coded as 1 there were multi-family dwelling units than single family, 0 if not.
  - Average Travel Time to Vehicle Storage Location: Used as provided.

- Short-term Parking Costs: Used as provided.
- Long-term Parking Costs: Used as provided.

- Employment
  - Employment Density: Calculated as the total number of employees over the area in acres.
- Household Income
  - Most prevalent household income quartile: Provided data was given as the number of households in a given household income quartile. For each TAZ the income quartile with the most households was selected and this was coded to a dummy variable.

Model Structure

We chose a nested logit model for our final specification. We tried three different variations of nesting structure: first, a model with a nest for transit modes (drive to transit and walk to transit) and a nest for active modes (walk and bike), and then a model with just a nest for walking modes (walk and walk to transit). Both of these nesting structures had scale parameters that were not significantly different from 1. However, our model that had a nest for solely driving modes (drive alone and shared ride in one nest, with the remaining four variables outside of the nest) had a scale parameter of $\mu = 1.43$ and passed a t-test at a significance level of 0.05. As a result, we chose this as our final nesting structure.

Model Development

Our initial model development was done using only the variables in the BATS dataset. We used alternative-specific coefficients for the pair of driving modes (drive alone and shared ride), transit modes (drive to transit and walk to transit), and active modes (walk and bike) for travel time. This is because travelers likely view their travel time spent driving equivalently, whether they are a passenger or the driver of the vehicle. Similarly, when walking to transit or driving to transit, the main mode of travel is still transit, and different passengers would likely value the travel time of taking transit equivalently as well. This was also justified by a likelihood ratio test, which achieved a test statistic of 176.560. This statistic follows a χ^2 distribution with 2 degrees of freedom, and is larger than 5.991, the critical value at the chosen significance level of 0.05. We used alternative-specific coefficients for each driving mode individually, but a single coefficient for the two modes that involved transit. This is because carpooling can significantly reduce the cost of traveling and this can impact the resulting mode choice utility, while driving and walking to transit likely have a similar cost. This was justified by a likelihood ratio test that had a test statistic of 1236.760, which is also greater than the critical value of 5.991.

After selecting variables and a model based solely on the BATS dataset, we began incorporating variables from the MTC Plan 2040 datasets. We specifically tested combinations of residential density, employment density, and transit access for origin and destination TAZs and found that they held predictive power. Initially these were tested in the utility function specifications for the transit trips with generic parameters, but we found that employment and residential densities at the origin

TAZs did not have statistically significant parameters. We then explored making alternative specific parameters for transit access because it is likely that those driving and walking to and from transit would have different thresholds for how far they might travel to use transit, and value their proximity to transit differently. In doing so we found that the parameter for transit access at the origin TAZ for driving to transit was not statistically significant, so we removed it from our model. This is likely because, for driving to transit, level of access within walking distance at the destination matters, but access at the origin matters less since travelers are driving to the transit station. In this case a further distance might be more indicative in a future mode. We further tested the generic versus alternative specific specifications of transit access and found that there was a statistically significant difference in the log likelihoods to justify keeping alternative specific parameters.

We finally tested nested logit structures as shared above and arrived at the model described below.

Results

Our final model is a nested logit model with the previously described structure and the following utility specifications:

1. $U_{DA} = \beta_{DA} + \beta_{travel\ time,\ car} travel\ time_{DA} + \beta_{cost,\ DA} cost_{DA} + \beta_{dist,\ car} dist_{car} + \epsilon_{DA}$
2. $U_{SR} = \beta_{SR} + \beta_{travel\ time,\ car} travel\ time_{SR} + \beta_{cost,\ SR} cost_{SR} + \beta_{dist,\ car} dist_{car} + \epsilon_{SR}$
3. $U_{walk} = \beta_{walk} + \beta_{travel\ time,\ active} travel\ time_{walk}$
4. $U_{bike} = \beta_{travel\ time,\ active} travel\ time_{bike}$
5. $U_{WT} = \beta_{WT} + \beta_{travel\ time,\ transit} travel\ time_{WT} + \beta_{cost,\ transit} cost_{WT} + \beta_{wait\ time} wait\ time_{WT}$
   $+ \beta_{access/egress\ time} (access\ time_{WT} + egress\ time_{WT})$
   $+ \beta_{res\ density,\ dest}\ res\ density_{dest} + \beta_{empl\ density,\ dest}\ empl\ density_{dest}$
   $+ \beta_{transit\ access,\ origin} transit\ access_{origin} + \beta_{transit\ access,\ dest,\ WT} transit\ access_{dest} + \epsilon_{WT}$
   $+$
6. $U_{DT} = \beta_{DT} + \beta_{travel\ time,\ transit} travel\ time_{DT} + \beta_{cost,\ transit} cost_{DT} + \beta_{access\ dist,\ transit} access\ dist_{DT}$
   $+ \beta_{wait\ time} wait\ time_{DT} + \beta_{access/egress\ time} (access\ time_{DT} + egress\ time_{DT})$
   $+ \beta_{empl\ density,\ dest}\ empl\ density_{dest} + \beta_{transit\ access,\ dest,\ DT} transit\ access_{dest} + \epsilon_{DT}$

(Note: DA = drive alone, SR = shared ride, WT = walk to transit, DT = drive to transit)

The final parameter estimates are in Table 1 below, with their robust standard errors listed in parentheses under each. All estimates were significant at a significance level of 0.05.

Table 1: Final Parameter Estimates with Robust Standard Errors

<table>
<tr><th>Variable</th><th>Drive alone</th><th>Shared ride</th><th>Walk</th><th>Bike</th><th>Walk to transit</th><th>Drive to transit</th></tr>
<tr><td>Constant</td><td>3.251
(0.122)</td><td>2.445
(0.130)</td><td>2.208
(0.150)</td><td></td><td>2.546
(0.174)</td><td>1.684
(0.174)</td></tr>
<tr><td>Travel time</td><td colspan="2">-0.016
(0.002)</td><td colspan="2">-0.021
(0.003)</td><td colspan="2">-0.023
(0.001)</td></tr>
<tr><td>Cost</td><td>-0.972
(0.0327)</td><td>-2.015
(0.068)</td><td></td><td></td><td colspan="2">-0.290
(0.023)</td></tr>
<tr><td>Travel distance</td><td colspan="2">0.128
(0.008)</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Wait time</td><td></td><td></td><td></td><td></td><td colspan="2">—0.070
(0.005)</td></tr>
<tr><td>Access + egress time</td><td></td><td></td><td></td><td></td><td colspan="2">-0.016
(0.002)</td></tr>
<tr><td>Residential density at destination</td><td></td><td></td><td></td><td></td><td>0.009
(0.004)</td><td></td></tr>
<tr><td>Employment density at destination</td><td></td><td></td><td></td><td></td><td colspan="2">0.003
(0.0002)</td></tr>
<tr><td>Transit access at origin</td><td></td><td></td><td></td><td></td><td>0.006
(0.001)</td><td></td></tr>
<tr><td>Transit access at destination</td><td></td><td></td><td></td><td></td><td>0.007
(0.001)</td><td>0.013
(0.001)</td></tr>
<tr><td>Access distance</td><td></td><td></td><td></td><td></td><td></td><td>-0.093
(0.005)</td></tr>
</table>

This model is consistent in the signs of the parameters; if cost or time for a certain mode increases, the probability of choosing that mode decreases. The rho bar square value is 0.394, which was the highest across all models we constructed.

Our guiding question for this analysis was whether increased access to transit would influence people's decisions to take transit. We can consider this impact by evaluating one fixed guideway transit system, BART. Since 2000, BART has expanded its stations to the South Bay, as seen in Figure 3 below.

Figure 3: BART System Map as of December 2022

The stations of Warm Springs/South Fremont, Milpitas, and Berryessa/North San Jose have been added in recent years. We considered how transit access (as we have defined it above) has changed for the TAZs served by these new BART stations. The following table shows the transit access for these TAZs.

Table 2: Transit Access in 2022 for TAZs with New BART Stations

TAZ	Transit Access in 2022 (%)
605	2.683
753	5.385
597	5.546
600	8.102
593	8.552
607	12.555
621	16.947
602	17.401

608	24.363
601	25.313
606	29.638
752	29.764
598	45.958

All of these TAZs had a transit access of 0 percent in 2000, so they all experienced an increase in transit access in the last 22 years. As our coefficients for the transit access variable are positive (0.006 for origin TAZ for walking to transit, 0.007 for destination TAZ for walking to transit, and 0.013 for destination TAZ for driving to transit), this indicates that an increase in transit access will lead to an increase in the utility of taking transit. Transit access is about equally important at both origins and destinations for those walking to transit, and understandably not too important at the origin for those driving to transit, as they can presumably drive a longer distance to access it. However, it is interesting that transit access is more important in destination TAZs for those driving to transit, perhaps because people who drive to transit may prefer to drive to their destination instead if the last-mile trip between transit and their destination is longer due to less transit access. This also reaffirms BART's decision to expand service to south Fremont and San Jose, as better transit access means higher probabilities of transit use.

Discussion and Conclusions

Our model is sensitive to a number of meaningful factors impacting mode choice and addresses a great deal of the variation in our dataset. In developing our model we also identified interesting caveats to the predictive power of land use and transit access variables, notably discerning that land use factors at destination TAZs were statistically significant, while most factors at the origin were not. This in particular warrants potential future study as to why this may be the case.

Additionally, future work should further explore different measures of transit proximity and access, including exploring the following alternative proximities:

- ¼-mile walk shed: This equates to a 4-5 minute walk. Using a quarter-mile walk shed may result in a greater proportion of transit riders within each TAZ. However, since we used stops and stations for fixed guideway transit in our analysis, and since these systems typically allow transit users to travel faster and farther than traditional bus lines, it is possible that those who decide to take transit may be willing to make farther trips to access a fixed guideway transit stop or station, knowing that it can take them farther distances or to their destination in a shorter amount of time than a traditional bus. If this were the case, to-transit accessibility based on distance may be underestimated.
- 1-mile walk shed: This buffer may capture more people within each TAZ who did not take transit, as the walk shed covers a larger radius than that in our analysis. As a result, this may

perhaps lead to assumptions that transit is not very accessible within that TAZ that may be controversial.
- 15-minute walk shed: A 15-minute walking distance translates to about 0.8-0.9 miles in distance. Similarly, we may capture more people who did not take transit, but consequently, this may lead to more conservative predictions of to-transit accessibility.

Also, while our analysis utilized data under transit conditions from 2000 (and as such externally available variables were hard to come by), current and future analyses should seek to utilize "high quality transit access" (access to transit systems with less than 15 minute average headways) as a better metric of transit service.

Ultimately, our model shows that transit access, residential density, and employment density all play a role in transit mode share, yet those who walk and those who drive to transit are distinct groups who value these factors differently. As the transit service in the Bay Area has expanded (and continues to expand), we anticipate that there would have been an increase in transit usage and hope that there are opportunities to further study how these developments have impacted travel behavior.

Citations

Blumenberg, E., & Smart, M. (2010). Getting by with a little help from my friends…and family: Immigrants and carpooling. *Transportation*, *37*(3), 429–446. https://doi.org/10.1007/s11116-010-9262-4

Bureau of Transportation Statistics. (2012). *Table 4-3: Major Urban Transit Agencies in California: 2000*. https://www.bts.gov/archive/publications/state_transportation_statistics/california/table_04_03

Castro, I., Huang, M., & Henderson, J. (2022, July 27). *Do All Bay Area Residents Have Access to Employment Opportunities?* Bay Area Equity Atlas. https://bayareaequityatlas.org/employment-opportunities

El-Geneidy, A., Grimsrud, M., Wasfi, R., & Tetreault, P. (2014). New evidence on walking distances to transit stops: Identifying redundancies and gaps using variable service areas. *Transportation*, *41*, 193–210.

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Gould, J., Golob, T. F., & Barwise, P. (1998). Why Do People Drive to Shop? Future Travel and Telecommunications Tradeoffs. *Institute of Transportation Studies*, 26.

Guerra, E., Cervero, R., & Tischler, D. (2012). Half-Mile Circle: Does It Best Represent Transit Station Catchments? *Transportation Research Board*, *2276*(1). https://journals.sagepub.com/doi/10.3141/2276-12

Horner, M. W., & Murray, A. T. (2004). Spatial Representation and Scale Impacts in Transit Service Assessment. *Environmental and Planning B: Urban Analytics and City Science*, *31*(5). https://journals.sagepub.com/doi/10.1068/b3046

Houston, D., Boarnet, M. G., Ferguson, G., & Spears, S. (2014). Can compact rail transit corridors transform the automobile city? Planning for more sustainable travel in Los Angeles. *Urban Studies*, *52*(5). https://journals.sagepub.com/doi/10.1177/0042098014529344

Iseki, H., & Taylor, B. (2010). Style versus Service? An Analysis of User Perceptions of Transit Stops and Stations. *Journal of Public Transportation*, *13*(3), 23–48. https://doi.org/10.5038/2375-0901.13.3.2

Kwoka, G. J., Boschmann, E. E., & Goetz, A. R. (2015). The impact of transit station areas on the travel behaviors of workers in Denver, Colorado. *Transportation Research Part A: Policy and Practice*, *80*, 277–287. https://doi.org/10.1016/j.tra.2015.08.004

Litman, T. (2008). Valuing Transit Service Quality Improvements. *Journal of Public Transportation*, *11*(2). https://doi.org/10.5038/2375-0901.11.2.3

Metropolitan Transportation Commission. (2022). *Existing and Planned Fixed-Guideway Transit Stations (2022)* [Map]. https://opendata.mtc.ca.gov/datasets/existing-and-planned-fixed-guideway-transit-stations-2022/explore?location=37.447470%2C-122.264838%2C7.97

MORPACE International, Inc. (2013). *Bay Area Travel Survey, 2000: Version 1* [Data set]. ICPSR - Interuniversity Consortium for Political and Social Research. https://doi.org/10.3886/ICPSR34805.V1

Ratner, K. A., & Goetz, A. R. (2013). The reshaping of land use and urban form in Denver through transit-oriented development. *Cities*, *30*, 31–46. https://doi.org/10.1016/j.cities.2012.08.007

Rosenbloom, S., & Burns, E. (1994). *Why Working Women Drive Alone: Implications for Travel Reduction Programs*. https://escholarship.org/uc/item/4x17v3f1

Şimşekoğlu, Ö., Nordfjærn, T., & Rundmo, T. (2015). The role of attitudes, transport priorities, and car use habit for travel mode use and intentions to use public transportation in an urban Norwegian public. *Transport Policy*, *42*, 113–120. https://doi.org/10.1016/j.tranpol.2015.05.019

U.S. Department of Labor. (2022, November). *Labor Force Status of Women & Men November 2022*. https://www.dol.gov/agencies/wb/data/widget

Yang, L., Chau, K. W., Szeto, W. Y., Cui, X., & Wang, X. (2020). Accessibility to transit, by transit, and property prices: Spatially varying relationships. *Transportation Research Part D: Transport and Environment*, *85*, 102387. https://doi.org/10.1016/j.trd.2020.102387

Zuo, T., Wei, H., & Chen, N. (2020). Promote transit via hardening first-and-last-mile accessibility: Learned from modeling commuters' transit use. *Transportation Research Part D: Transport and Environment*, *86*, 102446. https://doi.org/10.1016/j.trd.2020.102446