Inequitable Inefficiency: A Case Study of Rail Transit Fare Policies

Abstract
Transit fare equity research overwhelmingly measures equity based on disparity in the fare paid for travel without consideration of the costs of service delivery. Research also ignores the cost-sharing nature of transit — that as more riders consume it, the average cost per rider declines. Together, this leaves an incomplete understanding about who receives more subsidy.

I measure equity by analyzing spatial and temporal *cost recovery variability* of two rail systems, BART in the San Francisco Bay Area and MARTA in Atlanta. I scale origin-destination trip cost recoveries to stations and operating time periods. I find that travel associated with outlying areas and off-peak times receive more subsidy. I further find that subsidy patterns are marginally progressive; they positively correlate with select disadvantaged socioeconomic groups. I offer ideas on why these findings appear divergent from past research.

Keywords
Transit, equity, fare policy, transportation finance

1. Introduction
Transportation finance policy in the United States overwhelmingly rests on a principle that travel is a public good such that it is a public policy interest to provide travel-supportive infrastructure and services. As a result, travelers do not proportionally pay for these costs of their travel, if at all. This pertains regardless of mode. The *N*th driver on Boston's Big Dig (Interstate 93) or New York City's Major Deegan Expressway (Interstate 87) does not pay one-*N*th the costs of the maintenance of these roads, nor the marginal costs of additional capacity for and traffic impacts of their travel. Ditto for the *N*th transit rider in the Bay Area Rapid Transit District's (BART) Transbay Tube, the *N*th bicyclist using Utrecht Centraal's bicycle parking facility, and so forth. While this is a prima facie outcome, surprisingly little research evaluates how transport subsidies — the difference between the cost of providing transport and what travelers pay for the transport they consume — are distributed across space and time. Are travelers in different locations or at different times subsidized more than others? If so, this suggests that certain travel and development patterns (e.g., peak-period travel, sprawl, etc.) may be facilitated through largely opaque and poorly documented transport subsidies.

Research on the incidence of transport subsidies is especially limited for public transit. This may be due to its dual policy objectives that inherently foster subsidization, including serving as a transportation lifeline for those who cannot or choose not to drive, and an alternative to driving that can reduce the negative externalities of driving. Indeed, luring travelers from driving to transit with capital and operating investments paid by others is a principal method of achieving the second objective. And internalizing the costs of transit service provision flies in the face of the first objective, as captive riders would not be able to afford much travel at all. As a result, whereas highway operations and maintenance nationwide were financed 54% through user fees in 2018 (United States Department of Transportation, Federal Highway Administration, 2019), transit operating costs were financed 33% through fares in the same year (United States Department of Transportation, Federal Transit Administration, 2019).

Understanding who benefits most from transit subsidies can help inform the effectiveness of transit subsidy and investment practices. Are subsidies being allocated to places and times with high propensity for achieving the desired travel mode shift or serving captive riders? Are subsidies being made in a cost-effective way wherein the subsidy per benefit (e.g., rider, passenger-mile, car-to-transit mode shift, etc.) is minimized? In addition, do the implications of subsidy patterns warrant reconsideration of transit (or transportation) finance policies? For example, if achieving the goals of transit necessitates concentrating subsidies in cost-ineffective, low-density areas that are distant from urban activity centers, this suggests that transit subsidies reinforce sprawl patterns that are a common source of transportation and land use concern.

In the following analysis, I evaluate the location and temporal distribution of cost-recovery patterns of the BART and Metropolitan Atlanta Rapid Transit Authority (MARTA) rail networks. Noting that transit generally does not recover its costs through fares, and so is broadly inefficient, any variability in cost recovery of a network makes it *inequitably inefficient*. Further, I test whether trips' lengths or orientation around the network's urban core have a greater influence on subsidy patterns. Ceteris paribus, those who travel more receive more benefit and cost more to serve, so fares that do not scale with trip length will generate disparities. This motivates many fare equity studies that employ the market equity principle (Taylor and Norton, 2009). However, by considering only what users pay — the so-called pricing axis (Mallett, 2023a) — these studies ignore costs, including how they are distributed across a network and the cost-sharing nature of transit. That is, different speeds of travel, levels of service output, and capital intensity will affect costs, and as more travelers use a service, the cost per user will decrease. I account for both by using highly granular cost allocation, origin-destination (OD) ridership, and fare data to examine how cost recovery varies across space and time. To ensure no impact of the COVID-19 pandemic, fiscal year 2019 (FY19) — July 1, 2018, to June 30, 2019 — is the study period.

I hypothesize that travelers in outlying areas, travelers who travel longer distances, and travelers who travel outside of the weekday peak period each pay a lower share of their costs, but that the middle hypothesis is attenuated with fare policies that are distance based. In other words, I hypothesize that non-peak-period travel, long-distance travel, and travel to/from suburban and exurban areas are disproportionately subsidized relative to their counters, but that distance-based fare policies can reduce this effect for long-distance travel.

In the following section, I discuss research to date on the incidence of transit subsidies and fare equity, noting key distinctions in studies that do and do not account for costs. I then contextualize BART and MARTA and explain the basis for selecting them as case studies. While some general data about the agencies and travel patterns they serve are discussed here, I reserve the data used in my analysis for the sections that succeed, namely, Data and Methods and Descriptive Statistics. Finally, I summarize results and discuss their implications for planning and policy. Because I perform two parallel analyses — spatial and temporal — some sections are subdivided accordingly.

2. Transit Expenditure and Pricing Research

Transit subsidies are generally categorized as either supply side or demand side. Supply-side subsidies cover the costs of service provision that exceed income generated from fares, while demand-side subsidies are offered to riders to reduce the fares they pay. As reasoned, the incidence of both subsidies "trickles down" to riders; by partially covering operating costs, supply-side subsidies have the potential to reduce the fares needed to cover costs, while demand-side subsidies directly reduce fares paid by riders. However, supply-side subsidies have been shown to disproportionately benefit unionized transit workers and enable the industry to suffer from Baumol's Cost Disease (Jones, 1985; Wachs, 1989; Pickrell, 1985; Sarriera and Salvucci, 2016; Sarriera et al., 2018). At various times, as transit subsidies increased with the goal of increasing service levels, service levels were either maintained or reduced while unionized worker salaries were increased (Wachs, 1989). Sarriera and Salvucci (2016) and Sarriera et al. (2018) suggest this is driven by the industry being stagnant in technological and efficiency advancements relative to other industries, resulting in wage growth outpacing labor productivity. And according to Jones (1985), because there are few promotion opportunities for operators and mechanics, across-the-board salary increases are the primary way frontline workers can increase their incomes. As for whether and how subsidies trickle down to users, Serebrisky et al. (2009) conduct a critical literature review and surmise that most supply-side subsidy programs are socioeconomically neutral or regressive, while few demand-side subsidies are effective.

Apart from these top-level considerations, user equity of different transit fare structures has also been extensively studied. Methods have converged to emphasize what I call the "pricing axis" in expenditure and pricing equity (Mallett, 2023a). That is, while there are costs of service provision and the price paid for service, fare equity studies overwhelmingly consider only disparities in what users pay. Studies that considered costs were most popular in the 1980s as transit agencies migrated from using time- and distance-variant fare structures to flat-rate fare structures. Researchers and policymakers sought to understand the equity and efficiency implications of the transition. To account for cost variability, studies of the era employed cost allocation models in which input-output relationships are used to allocate categories of costs to service outputs (i.e., occasions of costs). For example, power supply costs may be allocated proportionally to vehicle-miles and labor costs to vehicle-hours. Once all costs are allocated to different dimensions of a network, such as time periods or service routes, the result can explain disparities in expenditure patterns. Studies of this time (hereafter, "earlier research") included those that considered only the expenditure axis (e.g., Cherwony and Mundle, 1978; Cherwony and Mundle, 1980; Taylor et al., 2000) and those that considered both pricing and expenditure (e.g., Cervero, 1981; Hodge, 1988; Parody et al., 1990; Reilly, 1977). More recent studies (hereafter, "newer research") have almost exclusively considered the pricing axis only (e.g., Bandegani and Akbarzadeh, 2016; Farber et al., 2014; Nuworsoo et al., 2009)

With few exceptions, earlier research finds that it is more costly to serve peak-period travel than off-peak-period travel in both net and gross terms — meaning, respectively, whether fares are or are not accounted for in the analysis. Earlier research also finds that urban *riders* subsidize suburban *riders*, but the collective of suburban *riders and taxpayers* subsidizes the collective of urban *riders and taxpayers*. On the latter point, Hodge (1988) found that farebox recovery flows from urban areas to suburban areas, but that when the tax base is accounted for, the net flow is reversed. Cervero (1981), in research also documented in Cervero and Wachs (1982), finds that it is more costly to provide peak-period and suburban service, that suburban trip lengths are longer and generate less net revenue per passenger-mile, and that this all makes flat-rate fares cost recovery inequitable. He further finds that flat-rate fares are regressive because lower-income, minority riders make disproportionately shorter, urban, and off-peak trips.

Of the earlier research, only Reilly (1977) found that the peak period costs less to serve in net terms. A deciding factor of this divergence appears to be if and how fixed and semi-fixed asset costs are included in the cost allocation. Cervero (1981) and Parody et al. (1990) allocate asset costs 85% to the peak period and 15% to the off-peak period in evaluating cost recovery variability; Taylor et al. (2000) use a marginal allocation method, derived from the Bradford Bus

Study (Savage, 1989), based on the share of transit vehicles in use during different time periods; and Reilly (1977) does not allocate asset costs. Key limitations of earlier research include the use of, at best, weakly disaggregate data — such as allocating costs and fare revenues to just two time periods (e.g., Cervero, 1981; Reilly, 1977; Parody et al., 1990), bus yard “cost centers” (e.g., Cervero, 1981), or “urban” and “suburban” labeled geographies (e.g., Hodge, 1988) to evaluate temporal and spatial variability of costs and subsidies — and their near-exclusive focus on bus transit. I elaborate more on the findings of these studies and the development of cost allocation methods in Mallett (2023b).

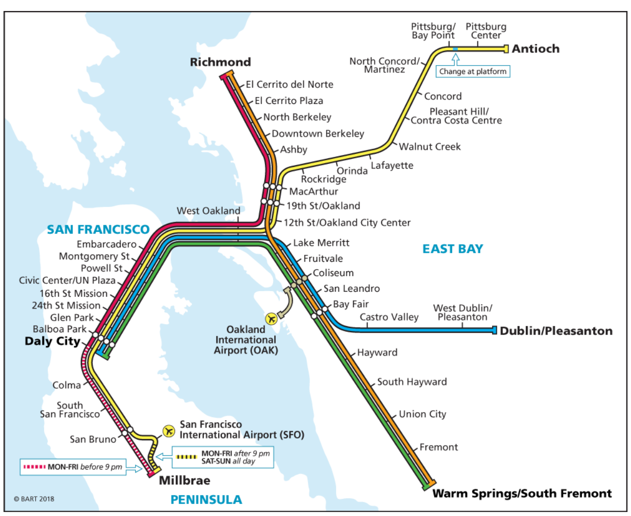
Since the early research era, data granularity has greatly improved, which can allow for much more informative findings about how costs and farebox recovery vary by time and location. However, as interest in fare equity has gained traction, newer research overwhelmingly ignores costs. Bandegani and Akbarzadeh (2016) merely test if distance-based fares are more equitable at charging riders proportionally to the amount of travel they consume. The answer is self-evident. Farber et al. (2014) use highly granular geographic data to test whether the impact of the Utah Transit Authority converting from flat to distance-based fares would vary by socioeconomic groups of riders. They find that the socioeconomic impact differential will vary by location since trip patterns of socioeconomic groups vary by location. And Nuwusoo et al. (2009) analyzed the equity impacts of various possible changes to the Alameda–Contra Costa Transit District’s (AC Transit) fare structure, most notably the implications of charging per trip segment or per trip inclusive of transfers. They found the former disproportionately burdens marginalized communities. Apart from my cost and cost-effectiveness study that is a precursor to this research, I found just two newer studies that consider costs. However, they rely on past study findings to inform their analysis (Brown, 2018) or continue to use highly aggregate time periods (Brown, 2018) and geographies (Iseki, 2016). Brown (2018) uses the 45% peak-to-base net cost ratio from Parody et al. (1990) based on 1983 national aggregate data and 2012 California Household Travel Survey data to conclude that the transit fare structure in Los Angeles is regressive and would be more equitable under a time- and distance-variant structure. Iseki (2016) conducted a similar study as Hodge (1988) focused on Toledo, Ohio, and found similar results. Other newer studies are international and similarly tend to assess fare equity based on pricing alone, though some highlight that distance-based fares are regressive relative to flat fares in countries other than the United States (e.g., Rubensson et al., 2020; Zhao and Zhang, 2019).

In sum, past literature suggests that, when asset costs are accounted for, the peak period is subsidized more than the off-peak period, but that the pattern is reversed when these costs are not accounted for; flat-rate fare structures are socioeconomically regressive because they allow persons who consume longer and more suburban trips — who tend to be white and wealthy — to pay both less per mile and less of their trip costs; and urban areas pay a higher share of travel costs through fares than suburban areas, but a lower share on net when tax source revenues are accounted. Among missing elements in the literature are more thorough analyses of transit modes other than bus transit, as well as the use of available granular data to evaluate temporal and spatial variability of cost recovery patterns more precisely. For example, transit operators’ time periods often are more nuanced than merely peak and base, and spatial variability is more varied than geopolitical boundaries. Most significantly, past research on spatial fare equity (e.g., Cervero, 1981) does not account for the cost-sharing nature of transit — that is, how much the cost of each unique mile of travel is shared among the consumers of that unique mile of travel. Some game theory literature on how transit riders may change travel patterns in a cost-sharing scheme (Rosenthal, 2017), as well as research on private-sector “collaborative transportation” concepts (e.g., Frisk et al., 2010; Guajardo and Rönnqvist, 2016), exists, but no research on transit cost recovery and fare equity accounts for this.

3. About BART and MARTA

BART and MARTA are 1970s-era regional rapid rail transit systems. They are distinct from metro rail or subway services, like the New York City Subway and the Massachusetts Bay Transit Authority's T system, as well as traditional commuter rail services, like Caltrain on the San Francisco Bay Peninsula and Metra in Chicago. Metro rail systems are characterized by being high-frequency, frequent-stop, fully grade separated, and catering to a principally urban travel base. By contrast, commuter rail services cater to long-haul suburban and exurban commuters, usually have more spacing between stops, extensively scale their services to cater to peak-period travel, are not subject to grade separation, and generally converge all routes around a downtown "union" station. Regional rapid rail transit systems are effectively a blend of these models. They bring the benefits of automated train control and rapid electric propulsion of metro rail services to longer-haul *regional* travel. With automation, trains can run at routine speeds and headways, and, with electrically propelled high speeds and acceleration, the time expense of stopping is significantly less than for commuter rail services that often must employ skip-stop service to overcome the cost.

I selected BART and MARTA as case studies because of their comparable operating characteristics, coupled with their having distinct fare structures — distance based and flat rate, respectively. Thus, going into the research, I theorized that any difference in findings between the two agencies would be explained by fare structure. Furthermore, the two agencies have highly granular trip and operating data, making them viable candidates for the research. For example, the entry, exit, and fare of every trip made, as well as the run time and railcar length of every scheduled train run, are recorded. The Washington Metropolitan Area Transit Authority (WMATA) was to be included and would have added an additional fare type for comparison: time-variant fares. However, agency staff did not provide all data needed for inclusion. **Figure 1** shows the system maps in use by BART and MARTA during most of FY19.



A: BART (effective July 1, 2018, to February 10, 2019)

B: MARTA (Effective FY19)

Figure 1: Agency System Maps

During FY19, MARTA's flat-rate fare was $2.50 per trip. However, it offered various discounts, including through transfer agreements with other transit operators, cooperative arrangements with area employers, multi-day passes, and more. At BART, base fares are distance based but follow a stepwise function. Riders pay a minimum fare for the first six miles of travel. Beyond six miles, riders pay for the first six miles plus a rate per mile up to 14 miles; beyond 14 miles, riders pay for the first 14 miles plus a lower cost per mile greater than 14 miles. Accordingly, although riders pay more for every additional mile traveled beyond six, longer-distance travel is discounted on a per-mile basis. There are also various fees, including a fee for use of the Transbay Tube, for travel to or from San Mateo County stations, and for travel to or from San Francisco and Oakland international airports; as well as various discount programs, including for senior and disabled riders, youth, and high-value discount tickets. I account for these many fees and discounts in OD pair cost recoveries by using the weighting average fare paid for the OD trip.

Table 1 provides agency profile information of the BART and MARTA rail networks, effective during FY19. Unless otherwise noted, this comes from Mallett (2022). Notably, although BART has a network about 2.3 times the size of MARTA's, it generated 3.5 times as many vehicle revenue-miles and 3.9 times as many trip-miles, but only 1.9 times as many trips.

	BART	MARTA
Miles (Mainline)	109.4 miles	48 miles
Stations	48 stations	38 stations
Fare Structure	Distance-based	Flat rate
Net Fare Revenue	$448,688,735	$58,576,496
Gross Costs	$847,799,127	$237,992,718
Farebox Recovery Ratio	52.9%	24.6%
Vehicle Revenue-Miles*	78M vehicle-miles	23M vehicle-miles

Annual Trips*	125M trips	65M trips
Annual Passenger-Miles*	1.756B passenger-miles	450M passenger-miles

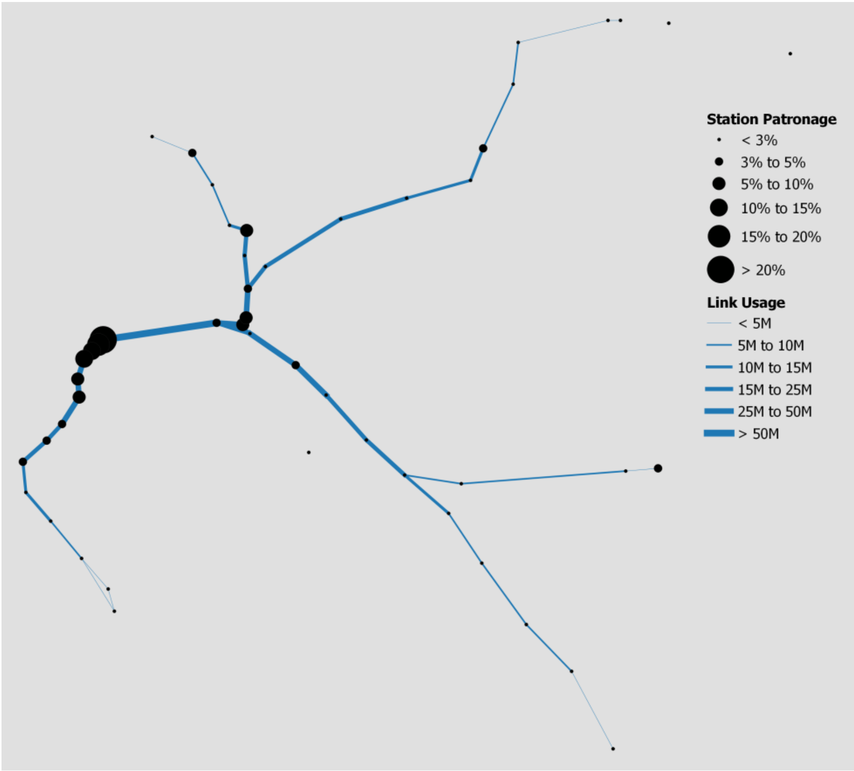
**As reported in 2019 Agency Profiles of National Transit Database*

Table 1: Agency Profiles

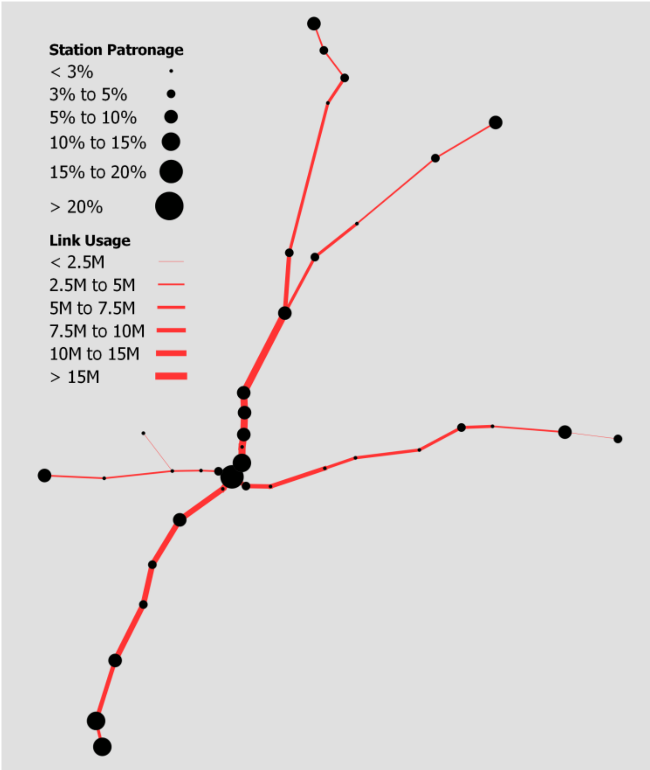
During FY19, the MARTA system had 38 stations, resulting in 703 two-way OD pairs while BART had 48 stations and 1,128 two-way OD pairs. However, BART's network includes three rail service types — standard BART trackage (mainline), a diesel multiple-unit segment of service (eBART), and the Oakland Airport Connector (OAC), which operates as a cable-powered people mover. As in the cost allocation research (Mallett, 2022), while I include all stations, I include only mainline portions of the BART track network to ensure consistency in analysis.

The spatial pattern of travel in the two systems is also distinct, which can influence how trip length and OD pairs' centralization around the urban core influences spatial cost recovery patterns. BART's OD ridership patterns revolve around a monocentric center, downtown San Francisco, whereas MARTA's OD ridership patterns are polycentric if not broadly dispersed. About two-thirds of all trips taken in the BART system begin or end at its four busiest stations in downtown San Francisco. Further, every other BART station's ridership has its highest destination relationship with one of the four busiest stations, ranging from 10% to 30%, with an average of 19%. And when the busiest stations are analyzed as a group, between 29% and 73% (with a 50% average) of trips from every other BART station is destined for downtown San Francisco. By comparison, just more than half of all trips taken on MARTA begin or end at its four busiest stations, these stations are associated with two epicenters of ridership, downtown Atlanta and the airport, and just 79% of other stations have their strongest destination relationship with one of these. Just 51% of trips that begin or end at a downtown Atlanta station — North Avenue, Civic Center, Peachtree Center, Five Points, Garnett, Georgia State, Dome/CNN Center, and Vine City stations. In both systems, the four busiest stations are the only stations with a double-digit share of passenger patronage.

While station patronage is relatively monocentric for BART and dispersed for MARTA, the bidirectional flow of ridership along links is centrally concentrated in both networks. This is reflected in . **Figures 2** and makes sense, as even in a polycentric or dispersed travel environment, trip paths can be centrally concentrated along links of a network even if origins and destination are not. This is especially likely if the network has a defined central node that many OD trips must pass through, as is the case for MARTA. I investigate these observations more thoroughly in the Results section.



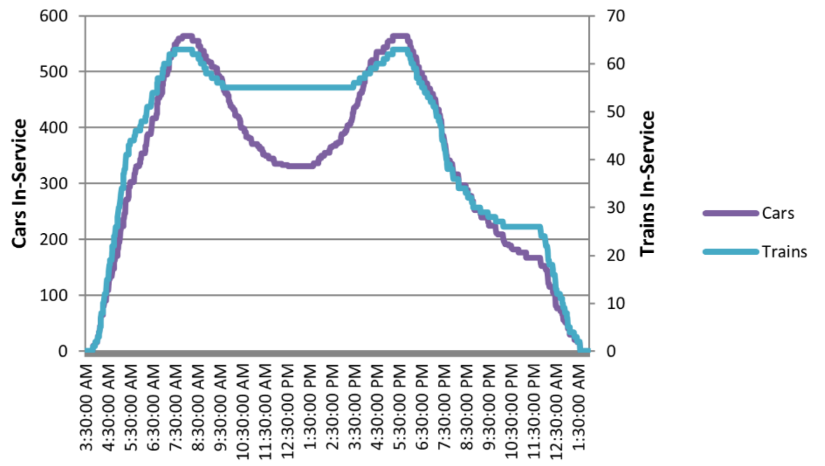
(A) BART



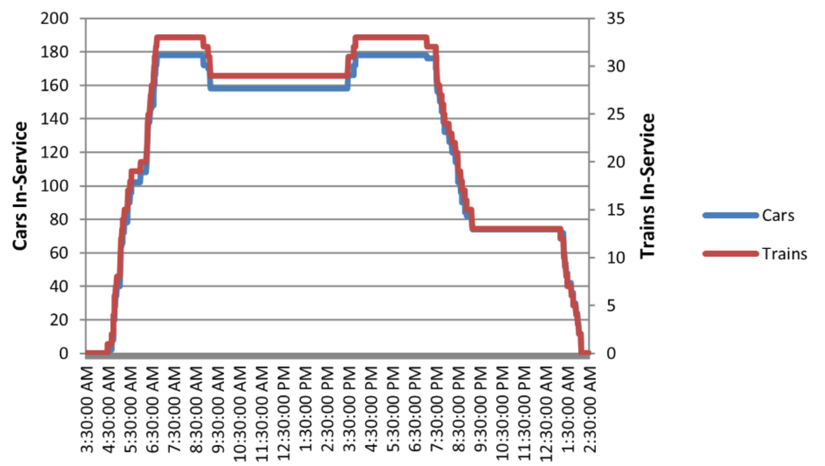
(B) MARTA

Figure 2: Spatial Pattern of Station Patronage and Link Usage

Finally, BART's and MARTA's temporal scaling of service output is not nearly as "peaked" as many other systems. **Figure 3**, from Mallett (2022), shows the number of trains and railcars in service during every minute of a weekday in the two networks. Conceptually, trains correlate with labor (i.e., train operators), and railcars correlate with assets and their maintenance. Notably, while BART scales the number of railcars in service by "making" and "breaking" train consists, unlike MARTA, both systems have relatively consistent train frequency from opening to evening. This may have relevance in the temporal analysis and its transferability to other systems or modes.



(A) BART (effective July 1, 2018 to February 10, 2019)



(B) MARTA

Figure 3: Weekday Temporal Distribution of Trains and Railcars in Service

4. Data and Methods

My objective in this research is to estimate if travelers at different times and in different locations pay a different share of their travel costs through fares in any statistically significant way such that transit operations are not only inefficient, but inequitably so on spatial and temporal dimensions. "Times" include the various time periods that a transit operator scales its output to, while "locations" are links and stations of the railroad. In addition, I test whether OD trip subsidies are more explained by trips' lengths or orientation around the transit system's urban core, and whether the subsidy patterns correlate with the socioeconomic makeup of riders. Finally, I infer fare structure implication by contrasting the difference in findings between BART and MARTA, two transit systems with similar operating environments but different fare structures.

I do not control for directionality in my analysis; a trip from one station to another is analyzed in tandem with a trip in the reverse direction. Again, for consistency in analysis, I include only mainline tracks of the BART network (i.e., where traditional BART trains/technology operate), so exclude any trips that solely use non-mainline portions of track. This results in 1,124 two-way OD pairs. Finally, for any trips that partially used a non-mainline portion of track, I assign trip costs using only the mainline links of the networks — that is, a trip's cost is equal to the summation of the costs per rider of the origin station, destination station, and each mainline link used to fulfill the trip.

For cost recovery analysis, I use FY19 budgetary, operating, and OD ridership and fare data provided by BART and MARTA. To analyze socioeconomic impacts, I use rider survey data from BART's 2015 Station Profile Study and the Atlanta Regional Commission's (ARC) 2019 Transit On-Board Survey . While the survey data provided by BART are already weighted to trip count, I manually weighted the survey data provided by ARC. **Table 2** lists and defines model terms I use.

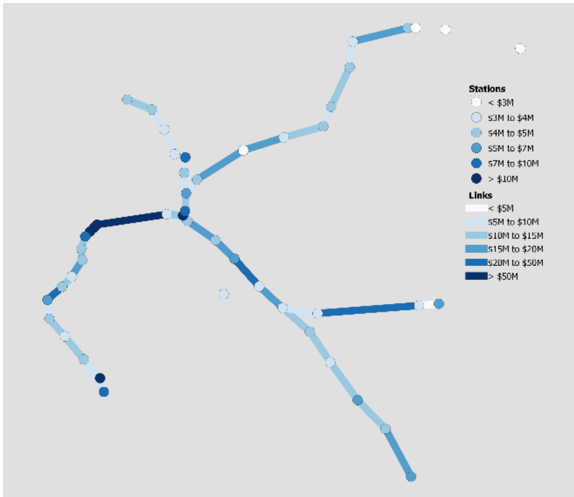
Variable	Description
$cost_{od}$	The cost of serving a particular OD trip, *od*.
$cost_t$	The cost expended serving a particular time period, *t*.
$cost_{tot}$	The cost expended during the study period.
costppm	The cost per passenger-mile of a particular link; that is, the allocated cost of a particular link divided by the sum product of the length (in track-miles) and number of annual trips that traverse the link.
costppx $costppx_l$ $costppx_o$ $costppx_d$	The cost per passenger of a particular station or link; that is, the allocated cost of a particular station or link divided by the number of trips into or out of the particular station or that traverse the particular link. Subscripts "l" "o," and "d" correspond to a unique link, origin station, and destination station, respectively.
$fare_{od}$	The average fare paid for consuming a particular OD trip, *od*.
$fares_t$	The total fare revenue generated during a particular time period, *t*.
$fares_{tot}$	The total fare revenue generated during the study period.
$percentpaid_{od}$	The average fare paid for a particular OD trip divided by the cost of serving the OD trip.
$percentpaid_l$ $percentpaid_s$	The average OD cost recovery ($percentpaid_{od}$) for all OD trips associated with a particular link, *l*, or station, *s*.
$percentpaid_t$	The costs recovered through fares generated in a particular time period, *t*.
pxmiles	The total passenger-miles generated — that is, the sum product of the count of trips associated with each OD trip and the trip length of each OD trip.
triplength	The length of a particular OD trip, in track-miles; limited to mainline track only.
$triplength_{average}$	The average length, in track-miles, of all OD trips made that used a particular station or link of the network or were taken during a particular time period; limited to mainline track only.
$distancecore_l$ $distancecore_s$ $distancecore_{average}$	The Euclidean (straight-line) distance that a particular station, *s*, link, *l*, or the average of the origin and destination stations of an OD pair, *average*, is from a defined core station of the network — West Oakland Station for BART, Five Points Station for MARTA. For links, the distance to the station farthest from the core station defines this value.

trips$_{od}$ *trips*$_{l}$ *trips*$_{s}$ *trips*$_{t}$ *trips*$_{tot}$	The number of trips generated across the study period. Subscripts "od," "l," "s," "t," and "tot" correspond to unique OD pairs, links, stations, time periods, and the total for the entire study period, respectively.
race_ai *race_api* *race_black* *race_hisp* *race_other* *race_white*	The percentage of a station's ridership that is of a particular racial group and/or Hispanic — American Indian (*ai*), Asian and Pacific Islander (*api*), Black or African American (*black*), Hispanic (*hisp*), Other or Mixed-Race (*other*), or White (*white*). For BART, respondents are either Hispanic of any race (*hisp*) or non-Hispanic with a race defined (*ai*, *api*, *black*, *other*, or *white*). For MARTA, Hispanic is a separate identification; respondents are a race and either Hispanic or non-Hispanic.
*income_0to20** *income_0to25*** *income_20to30** *income_25to35*** *income_30to40** *income_35to40*** *income_40to50* *income_50to60* *income_60to75* *income_75to100* *income_100to120** *income_100to150*** *income_120plus** *income_150plus***	The percentage of a station's ridership whose annual household income is in a particular income band — $0 to $20,000 (*0to20*), $20,000 to $30,000 (*20to30*), and so forth. ** Unique MARTA band* *** Unique BART band*

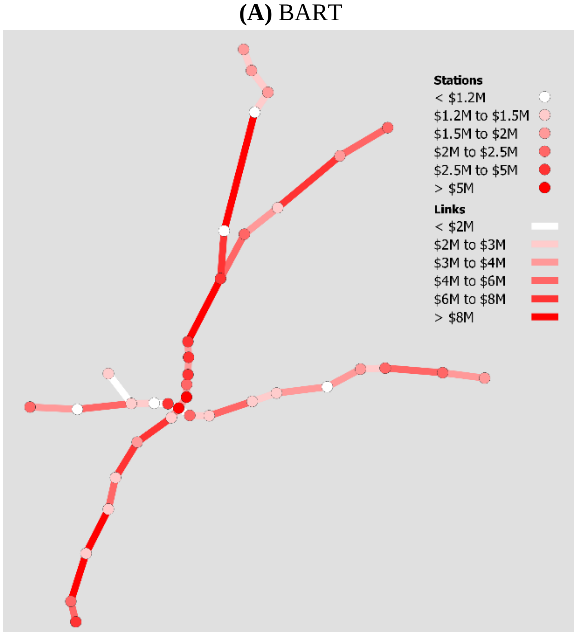
Table 2: Table of Variables

Measuring cost recovery variability requires analyzing how both costs and fare receipts vary. Accordingly, I first allocate costs to time periods for the temporal analysis and to links and stations for the spatial analysis. Then, I divide the allocated costs by the number of riders who travel during each time period, traverse each link, and use each station. This results in costs per rider for time periods, links, and stations (for more details, see Mallett, 2022).

In brief, I find that costs decrease but costs per rider increase with distance from the core of the BART system, that there is no clear spatial pattern in the MARTA system, and that the weekday peak period — out of eight operating time periods for BART and five for MARTA — has the lowest cost per rider in both networks but the highest cost for BART and barely lower costs than the weekday base period for MARTA. **Figures 4** and **5** show the spatial cost and cost per rider findings from Mallett (2022), respectively; **Figure 6**, the temporal findings.

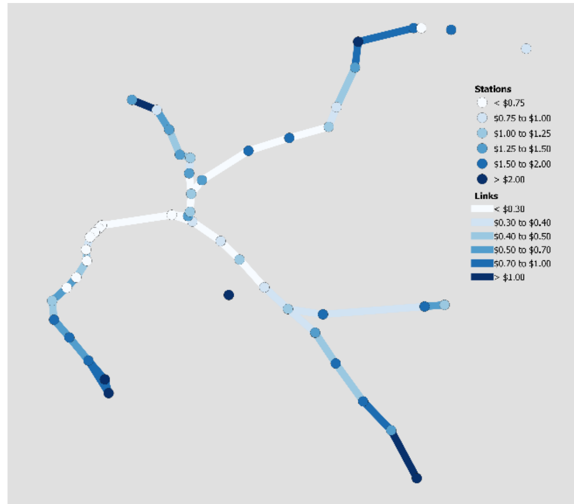


(A) BART

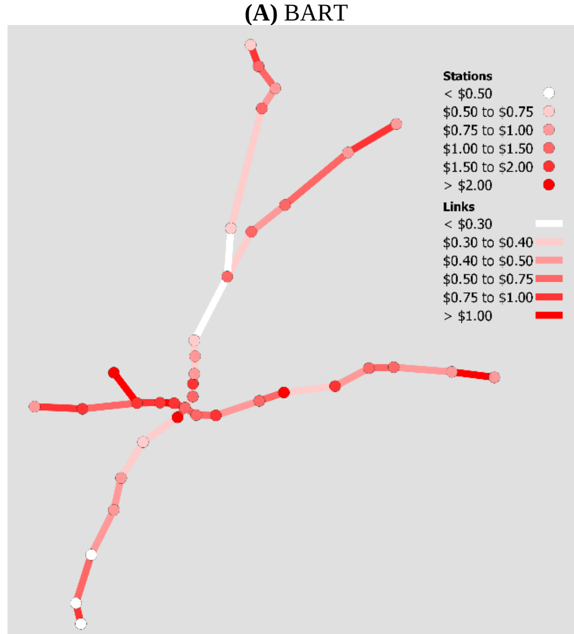


(B) MARTA

Figure 4: Spatial Cost Variability

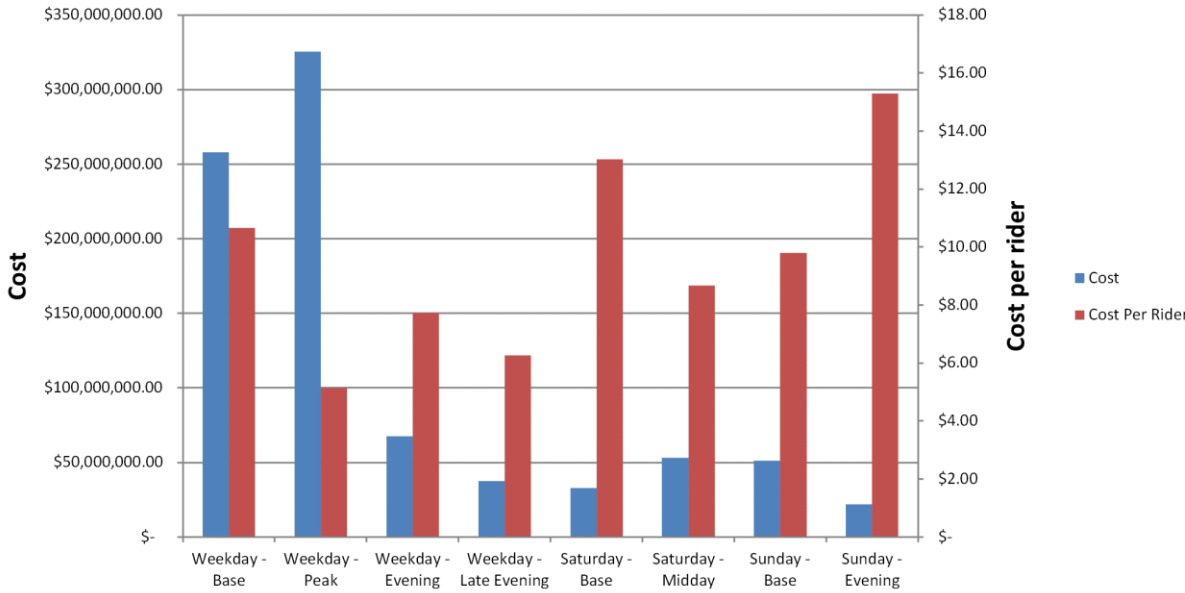


(A) BART

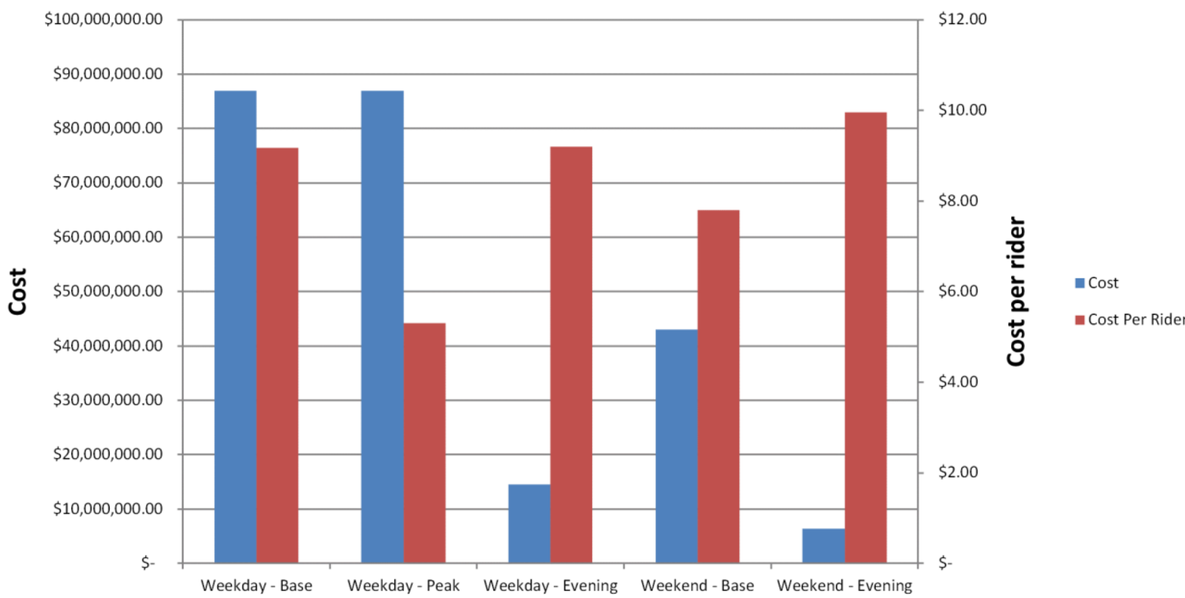


(B) MARTA

Figure 5: Spatial Cost per Rider Variability



(A) BART



(B) MARTA

Figure 6: Temporal Cost and Cost per Rider Variability

To analyze temporal cost recovery variability, I calculate the ratio between total fare revenues and total costs by time period (**Equation 1**). I also run pairwise correlations to assess how costs, fare revenues, cost recoveries, passenger-miles, trip counts, and the spatial variables explained below interact across time periods. Because ARC and BART did not survey weekend riders, socioeconomic data are insufficient to investigate how temporal subsidies interact with the makeup of riders of different time periods.

$$percentpaid_t = \frac{fares_t}{cost_t} \tag{1}$$

Analyzing spatial cost recovery variability is more complex, as there are both OD and location (i.e., stations and links) considerations. Furthermore, unlike with time periods, which are one-dimensional, there is no practical way to divvy up fares into link and station parts. Thus, links and stations do not have a cost recovery per se. Instead, I calculate base cost recoveries at the OD level, then scale to stations and links using averages to understand the spatial incidence of subsidies.

I first assign every OD pair a cost equal to the sum of the costs per rider of the entry station, exit station, and each link used to complete the trip (**Equation 2**). I then calculate the OD cost recovery rate, which is the ratio of the average fare paid for an OD trip to the OD trip cost (**Equation 3**).

$$cost_{od} = costppx_o + costppx_d + \sum_{l=1}^{n} costppx_l \tag{2}$$

$$percentpaid_{od} = \frac{fare_{od}}{cost_{od}} \tag{3}$$

To test how much trip subsidies are explained by trip length relative to their orientation around the urban core, I run ordinary least squares regressions. I regress the log of OD cost recovery onto the log of both trip length and the average straight-line distance that the origin and destination stations are from a defined core station of the network — West Oakland for BART and Five Points for MARTA. This is defined in **Equation 4**.

$$\ln(percentpaid_{od}) = \beta_0 + \beta_1 \ln(triplength) + \beta_2 \ln(distancecore_{average}) \tag{4}$$

In essence, the second term proxies for an OD pair's monocentricity. An OD pair whose origin and destination are distant from the core will have a large value for this term; one whose origin and destination are both near the core, a small value; and one with a distant origin or destination and the other near the core, a moderate value. Notably, this term is not inherently collinear with trip length, as OD pairs can have origins and destinations distant from the core and be either short or long in length.

To understand the geographic incidence of subsidies, I devise cost recovery profiles of stations and links equal to the weighted average of what users of each station and link pay as a share of their trip costs. Hence, subject to an OD trip being associated with a station or link, the average cost recovery of all OD trips consumed defines the station or link cost recovery profile. I define these calculations in **Equations 5** and **6**, respectively. Referring to Figure 1(a), a trip from MacArthur to Fruitvale and a trip from Dublin/Pleasanton to Montgomery Street will both be included in the weighted average cost recovery of the Lake Merritt–Fruitvale link, but only the former trip will be included in the weighted average cost recovery of Fruitvale Station.

$$percentpaid_s = \frac{\sum_{s=o,d}(trips_{od} * percentpaid_{od})}{trips_s} \tag{5}$$

$$percentpaid_l = \frac{\sum(trips_{od} * percentpaid_{od}) \in l}{trips_l} \tag{6}$$

Finally, as with the temporal analysis, I evaluate the covariance of costs, fares, cost recovery, average trip lengths, and more across stations and links of the railroads. For the station-level analysis, I include socioeconomic variables of race and income, as the data provided by ARC and BART are scaled for station profile purposes. Admittedly, both regression analysis and use of OD data would be preferred. However, they are infeasible because there are insufficient sample sizes of the socioeconomic data at the OD level, and with 38 and 48 stations for MARTA and BART, respectively, there are too few observations for multivariate regression across stations.

5. Descriptive Statistics

Tables 3 and **4** show descriptive statistics of variables in the OD cost recovery and regression analysis for BART and MARTA, respectively. In these tables, I show both unweighted distributions and distributions weighted to OD trip count. While a particular OD pair may have a cost recovery higher than another, if the former has more consumption (i.e., trip counts) than the latter, the mean weighted to trip count will be higher than the unweighted mean. Similarly, the average, median, and standard deviation of variables for all actualized trips will be different from the same statistics without weighting. I use unweighted values for the regression analysis, as these explain the *occasions* of subsidies, but show both unweighted and weighted statistics to demonstrate how OD pair patterns differ from aggregate trip consumption patterns — which is pertinent for the station and link analyses that describe the *incidence* of subsidies.

	Minimum	**Mean**		**Median**		**Standard Deviation**		**Maximum**
		Unweighted	**Weighted**	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**	
$percentpaid_{od}$	10.16%	44.24%	63.31%	39.35%	57.11%	19.24%	25.6%	157.46%
$distancecore_{average}$	0.64	11.64	8.76	11.03	8.02	5.3	4.16	29.81
triplength	0.35	21.82	14.98	20.97	12.63	12.5	10.06	56.44
trips	354	104,514.5	—	31,702.5	—	184,272.8	—	1,372,116

N (two-way OD pairs) = 1,124

Table 3: Descriptive Statistics of OD Pairs — BART

	Minimum	**Mean**		**Median**		**Standard Deviation**		**Maximum**
		Unweighted	**Weighted**	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**	
$percentpaid_{od}$	10.2%	30.10%	36.05%	27%	31.44%	13.38%	17.51%	105.25%
$distancecore_{average}$	0.18	4.99	5.06	4.78	4.83	2.75	2.68	12.88
triplength	0.38	9.62	9.13	9.05	8.45	5.96	6.06	26.39
trips	725	58,682.16	—	30,181	—	77,612.73	—	687,245

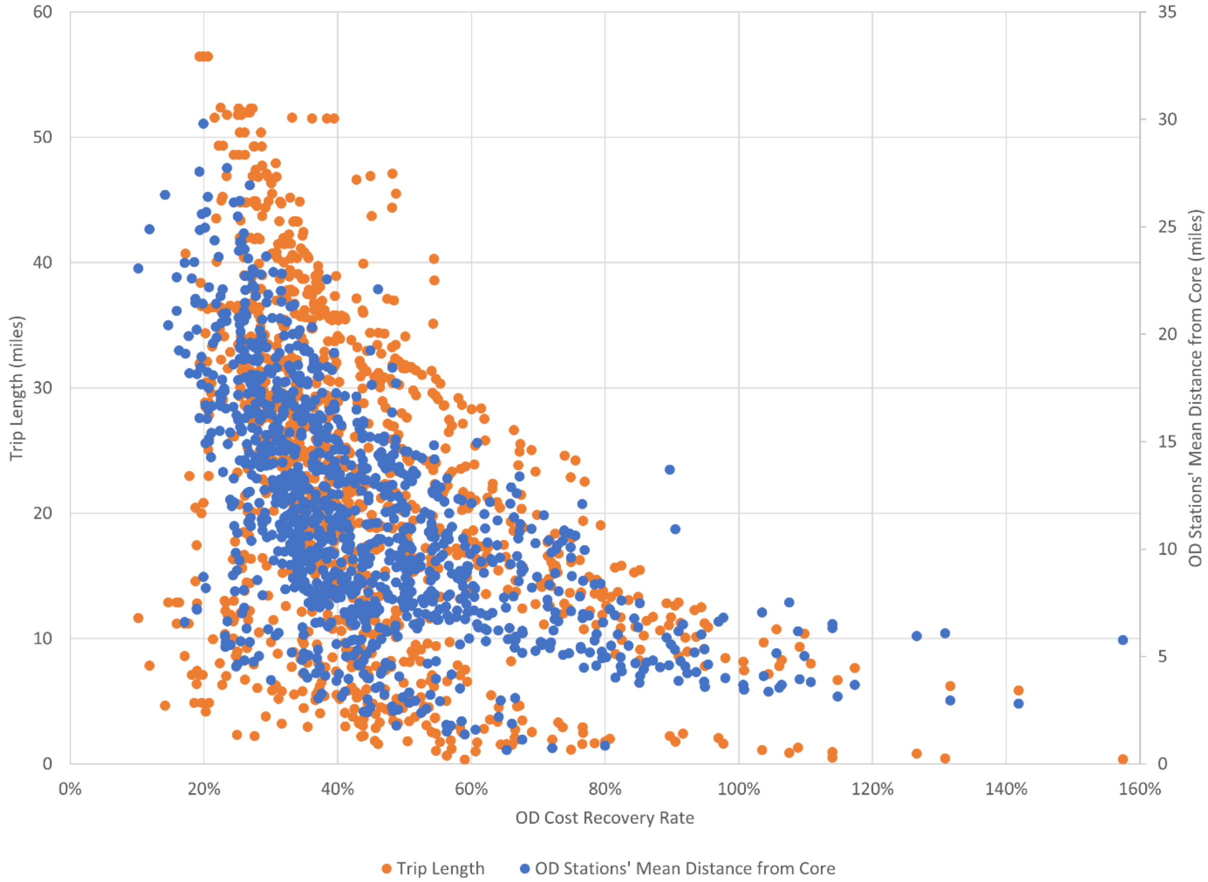
N (two-way OD pairs) = 703

Table 4: Descriptive Statistics of OD Pairs — MARTA

The difference between the two networks in how unweighted (i.e., OD pairs, irrespective of trip count) and weighted (i.e., total OD trips) statistics vary is noteworthy. BART's weighted data are markedly different from the unweighted data, implying that trip consumption patterns vary significantly across the network; trip consumption is not normally distributed. Specifically, the average cost recovery of trips consumed (weighted) is 1.4 times *more*, average trip length 31% *less*, and core orientation of travel 25% *more*, than if trips were equally spread across all OD pairs (unweighted). By comparison, MARTA's trip consumption patterns are similar to the network's unweighted spread of OD pairs. In addition, while BART trips are longer and more "suburbanized" in whole number terms, as a percentage of the maximum extents of the networks, MARTA trips are longer and more suburban. Trips taken on BART (weighted data), on average, have a station pair whose average straight-line distance from the core station is 8.8 miles, or 29% of the maximum possible. The same statistic for MARTA is 5.0 miles, equal to 39% of the maximum possible. Similarly, the average length of consumed trips taken on BART is about 15 miles, or 27% of the maximum possible trip length; on MARTA, 9.1 miles, or 35% of the longest possible trip.

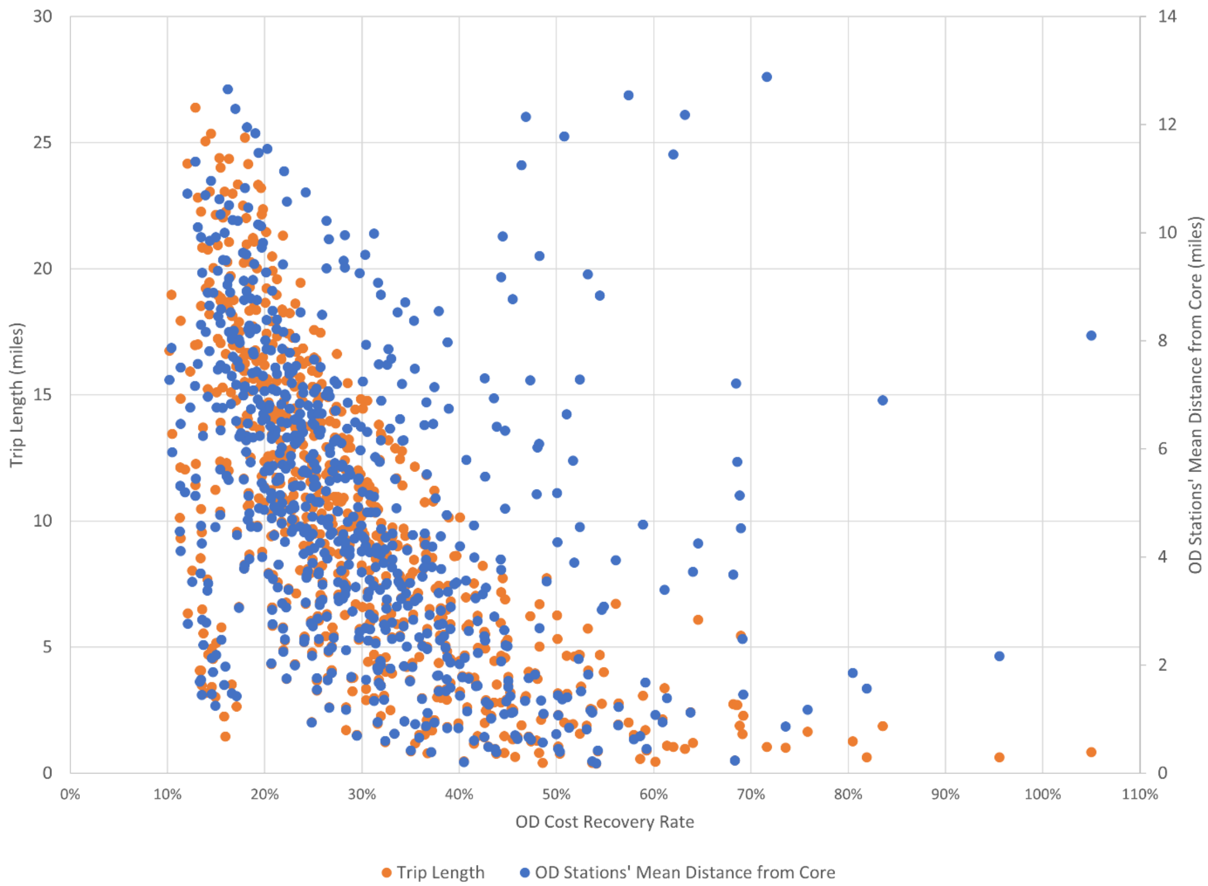
One last point on Tables 3 and 4. It may go against intuition that the systemwide cost recoveries of 53% for BART and 25% for MARTA defined in Table 1 do not align with the unweighted or weighted mean cost recoveries in Tables 3 and 4. This is because OD pair and trip cost recoveries are derived from a subset of costs and fare revenues, whereas the systemwide cost recovery is the quotient of aggregate fare revenue divided by aggregate costs. Therefore, OD pair and trip cost recoveries will not revolve around the systemwide cost recovery value since their source of costs and fares are different. This underscores an objective of this research: the aggregate pattern of a network does not necessarily reflect patterns across its parts.

Figure 7 graphically shows the two independent variables' relationship with the dependent variable. Consistent with my general hypotheses, both independent variables have an increasingly negative effect with the independent variable; cost recovery decreases with trip length and average distance of the entry and exit stations from the urban core. However, particularly for MARTA, there are outliers for the second independent variable.



60
50
40
30
20
10
0
Trip Length (miles)
0%
20%
40%
60%
80%
100%
120%
140%
160%
OD Cost Recovery Rate
35
30
25
20
15
10
5
0
OD Stations' Mean Distance from Core (miles)
Trip Length
OD Stations' Mean Distance from Core

(A) BART



(B) MARTA

Figure 7: Cost Recovery vs. Trip Length and Average OD Station Distance from Core

Tables 5 and **6** show the descriptive statistics for the station-level analysis for BART and MARTA, respectively; **Tables 7** and **8**, the same for the link-level analysis. As previously discussed, these data are station and link averages of the above-summarized OD data, weighted to OD tip count and conditional on a station being an origin or destination station, or a link being one that is traversed in completing an OD trip.

Variable	Minimum	Mean	Median	Standard Deviation	Maximum
$percentpaid_s$	24.86%	54.47%	54.57%	14.36%	99.79%
$distancecore_s$	0	11.68	9.89	7.73	31.05
$distancecore_{average}$	3.44	9.85	8.71	3.79	19.51
$triplength_{average}$	7.67	17.36	15.25	7.8	42.25
$trips_s$	617,004	4,897,543	3,473,553	5,053,904	24,571,444
race_ai	0%	0.32%	0.28%	0.2%	0.79%
race_api	9.91%	22.9%	21.9%	9.38%	50.89%
race_black	4%	12.16%	9.76%	7%	36.02%
race_hisp	7.88%	17.62%	16.04%	5.89%	37.35%
race_other	1.27%	3.39%	3.46%	0.88%	5.2%
race_white	23.58%	43.61%	45.28%	11.83%	69.51%
income_0to25	2.25%	8.31%	7.68%	3.09%	15.17%
income_25to35	1.95%	5.23%	5.12%	1.79%	8.9%

income_35to40	1.48%	4.87%	4.61%	1.87%	9.33%
income_40to50	2.29%	8.47%	7.89%	2.68%	13.71%
income_50to60	5.18%	12.36%	12.08%	3.41%	21.34%
income_60to75	8.81%	15.24%	15.36%	2.56%	21.11%
income_75to100	8.06%	15.21%	15.19%	2.45%	20.08%
income_100to150	6.28%	16.08%	16.12%	4.1%	23.23%
income_150plus	2.59%	14.22%	12.98%	7.29%	40.82%

N (stations) = 48 (45 for socioeconomic variables)

Table 5: Descriptive Statistics of Station Profiles — BART

Variable	Minimum	Mean	Median	Standard Deviation	Maximum
$percentpaid_s$	13.38%	29.37%	29.93%	6.27%	40.71%
$distancecore_s$	0	5.66	4.73	4.67	15.95
$distancecore_{average}$	2.58	4.97	4.34	1.87	8.56
$triplength_{average}$	5.02	8.81	7.47	3.21	16.36
$trips_s$	430,987	2,171,240	1,728,397	1,529,423	7,356,599
race_ai	0%	0.96%	0.9%	0.53%	2.45%
race_api	0.53%	3.41%	2.4%	2.51%	9.92%
race_black	39.2%	66.49%	65.44%	14.59%	89.38%
race_hisp	2.98%	5.72%	5.22%	2.28%	13.38%
race_other	0.38%	4.4%	4.19%	2.33%	11.13%
race_white	6.64%	24.74%	25.67%	11.22%	43.86%
income_0to20	4.96%	15.24%	15.2%	5.26%	26.32%
income_20to30	4.96%	11.52%	11.43%	3.3%	17.99%
income_30to40	7.75%	14.83%	14.72%	3.87%	25.57%
income_40to50	8.68%	15.42%	15.25%	2.95%	19.87%
income_50to60	8.19%	13.65%	13.9%	2.57%	19.9%
income_60to75	3.79%	11.5%	11.37%	2.66%	16.27%
income_75to100	2.83%	8.94%	8.23%	4.27%	20.99%
income_100to120	0.32%	4.23%	3.63%	2.86%	11.45%
income_120plus	0.33%	4.67%	3.91%	3.37%	13.02%

N (stations) = 38

Table 6: Descriptive Statistics of Station Profiles — MARTA

Variable	Minimum	Mean	Median	Standard Deviation	Maximum
$percentpaid_l$	24.86%	47.61%	48%	9.29%	64.58%
$distancecore_l$	1.59	13.61	12.16	8.21	30.02
$distancecore_{average}$	6.72	11.46	10.9	3.24	18.92
$triplength_{average}$	1.82	22.14	19.88	7.36	39.78
$trips_l$	154,130	18,199,306	13,028,305	15,014,428	64,656,418

N (links) = 47

Table 7: Descriptive Statistics of Link Profiles — BART

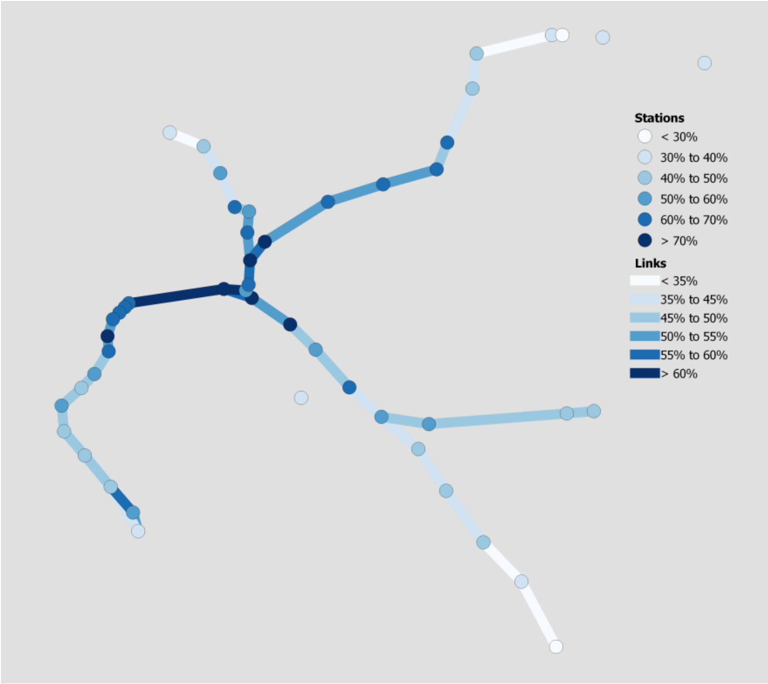
Variable	Minimum	Mean	Median	Standard Deviation	Maximum
$percentpaid_l$	13.38%	23.61%	23.83%	2.85%	27.35%
$distancecore_l$	0.38	5.82	4.8	4.64	15.95
$distancecore_{average}$	3	5.95	5.76	1.43	8.56
$triplength_{average}$	6.37	12.28	12.42	2.45	16.57
$trips_l$	430,987	8,194,683	6,654,606	4,931,187	17,505,807

N (links) = 37

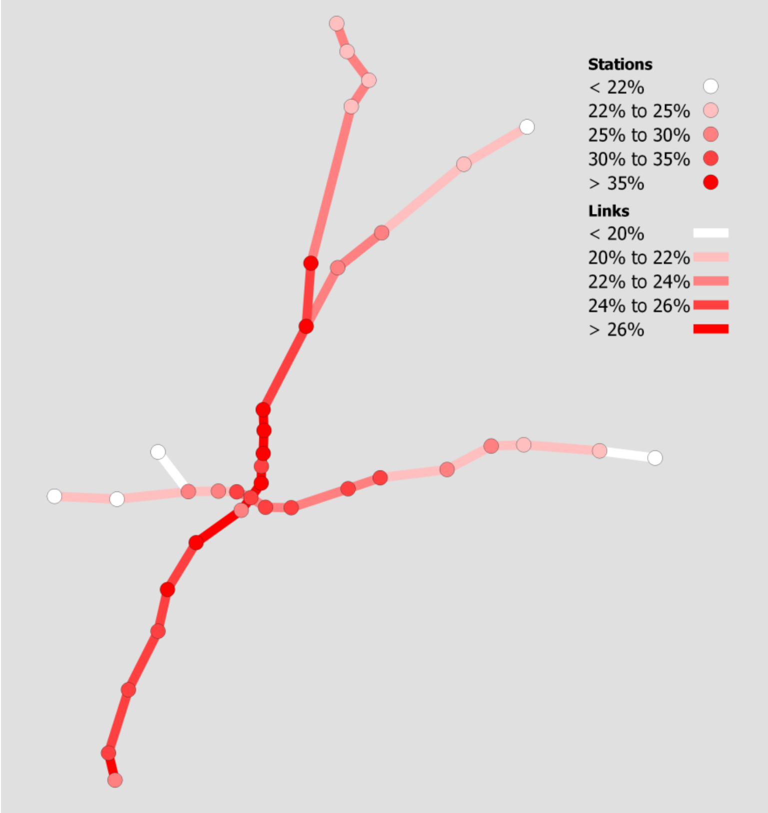
Table 8: Descriptive Statistics of Link Profiles — MARTA

Comparing these statistics with the weighted OD trip statistics in Tables 3 and 4, station and link mean cost recovery profiles are noticeably less than the mean weighted OD trip cost recovery. With smaller standard deviations, the spread is also much less than weighted OD trip cost recoveries. This is principally explained by these being different measurements. The latter is a measurement of trip cost recoveries, while the former are measurements of trip cost recovery *averages* for trips that are associated with a link or station. As a result, although high cost recovery OD trips "crowd out" low cost recovery OD trips in the OD trip cost recovery statistics, those same OD trips share links and stations with low cost recovery OD trips. Thus, even the links or stations with the highest cost recovery profile will inherently have a value less than the highest cost recovery OD trip. Furthermore, there are far fewer links and stations than there are OD pairs, and high cost recovery trips may be concentrated across few links and stations, leaving the remainder of links and stations to be associated with lower cost recovery OD trips. I evaluate this concept in the Results section.

Figure 8 shows the geographic distribution of cost recovery across stations and links in the BART and MARTA networks. To illustrate interpretation of these maps, the "average rider" who travels to or from MARTA's Five Points Station pays 34.1% of the costs of their trip, while the "average rider" who traverses BART's link between West Oakland and Embarcadero stations (the Transbay Tube) pays 63.9% of their total trip cost. With few exceptions, this figure suggests that the average cost recovery of riders of different links and stations generally *declines* with distance from each system's core — though, as suggested by the different scale on the maps, the magnitude of variance is significantly less for MARTA relative to BART. In addition, consistent with my cost vs. cost per rider analysis (Mallett, 2022), BART's most expensive stations and links to operate — and likely to build due to their being tunneled — also have the lowest cost per rider and highest cost recovery profile.



(A) BART



(B) MARTA

Figure 8: Average Cost Recovery Across Links and Stations

Lastly, for the temporal analysis, **Tables 9** and **10** show descriptive statistics for different time periods for BART and MARTA, respectively. For five of the variables — cost, fare revenue, cost recovery, trip count, and passenger-miles — I show the aggregate value for the time period. For the other variables — average origin and destination station straight-line distance from the core station, and trip length — I show the statistics corresponding to OD trips made during each time period. Hence, the variation in these latter values reflects how trip consumption patterns vary across time periods.

Time Period		Weekday				Saturday		Sunday/Holiday		Total
		Base	Peak	Evening	Late Evening	Base	Midday	Base	Evening	
$cost_t$		$258,037,518	$325,430,391	$67,444,710	$37,682,645	$32,792,752	$53,279,185	$51,104,798	$22,027,128	$847,799,127
$fares_t$		$89,697,710	$243,146,564	$35,114,074	$23,259,610	$9,622,910	$22,504,842	$19,464,467	$5,878,558	$448,688,735
$percentpaid_t$		34.76%	74.72%	52.06%	61.72%	29.34%	42.24%	38.09%	26.69%	52.92%
$trips_t$		24,177,808	63,242,768	8,732,436	6,017,829	2,514,618	6,141,670	5,208,938	1,438,202	117,474,269
pxmiles		343,857,182	970,670,140	138,290,468	89,440,215	36,825,580	84,332,354	72,345,257	22,255,671	1,758,016,867
triplength	**Minimum**	0.35				0.35		0.35		0.35
	Mean	14.22	15.35	15.84	14.86	14.64	13.73	13.89	15.47	14.97
	Standard Deviation	10.16	9.96	10.45	10.15	10.01	9.84	9.86	10.6	10.07
	Maximum	56.92				56.92		56.92		56.92
$distancecore_{average}$	**Minimum**	0.64				0.64		0.64		0.64
	Mean	8.61	8.82	8.92	8.72	8.73	8.53	8.66	9.06	8.76
	Standard Deviation	4.18	4.1	4.3	4.3	4.24	4.1	4.16	4.39	4.15
	Maximum	29.81				29.81		29.81		29.81

Table 9: Temporal Descriptive Statistics — BART

Time Period		Weekday			Weekend/Holiday		Total
		Base	Peak	Evening	Base	Evening	
$cost_t$		$86,991,391	$86,950,887	$14,582,218	$43,032,908.02	$6,435,315	$237,992,718
$fares_t$		$15,839,966	$29,559,350	$2,626,714	$9,450,864	$1,099,602	$58,576,496
$percentpaid_t$		18.21%	34%	18.01%	21.96%	17.09%	24.61%
$trips_t$		12,391,117	19,898,647	2,057,750	6,985,699	841,506	42,174,719
pxmiles		109,206,799	188,547,757	17,270,565	61,508,060	6,977,054	383,510,235
triplength	**Minimum**	0.38			0.38		0.38
	Mean	8.81	9.48	8.39	8.8	8.29	9.09
	Standard Deviation	6.22	5.94	5.86	6.04	5.79	6.04
	Maximum	26.39			26.39		26.39
$distancecore_{average}$	**Minimum**	0.18			0.18		0.18
	Mean	4.95	5.16	5.12	4.91	4.99	5.05
	Standard Deviation	2.74	2.66	2.56	2.65	2.54	2.68
	Maximum	12.88			12.88		12.88

Table 10: Temporal Descriptive Statistics — MARTA

As is evident from these data, the peak period is the most expensive to operate by far for BART and almost as expensive to operate as the weekday base period for MARTA. But in both cases, the weekday peak period serves so many more riders and passenger-miles that the cost per rider and per passenger-mile is lowest during this period. Similarly, by serving so many more riders and passenger-miles, much more fare revenue is generated during the weekday peak period — so much that it offsets any additional costs of serving the period, even when semi-fixed costs are accounted for. As a result, the weekday peak period recovers the greatest share of costs of any time period.

6. Results — Spatial Analysis

Tables 11 and **12** show the results of the OD cost recovery model for BART and MARTA, respectively. I show both base units of measurement (e.g., miles) as well as standard normal units, and use robust standard errors.

Variable	**Coefficient**		**Standard Error (Robust)**	**95% Confidence Interval**
Base / Standard Normal (SN)	**Base**	**SN**		
ln(triplength)	-0.04*	-0.082*	0.021	(-0.081, 0.002)
ln($distancecore_{average}$)	-0.409***	-0.556***	0.031	(-0.47, -0.349)
constant	4.777***		0.058	(4.663, 4.89)
N	1,124			
R-squared	0.3729			

Statistical Significance: *** p ≤ 0.001, ** p ≤ 0.01, * p ≤ 0.1

Table 11: Origin-Destination Cost Recovery Model Results — BART

Variable	**Coefficient**		**Standard Error (Robust)**	**95% Confidence Interval**
Base / Standard Normal (SN)	**Base**	**SN**		
ln(triplength)	-0.44***	-0.898***	0.021	(-0.481, -0.399)
ln($distancecore_{average}$)	0.136***	0.243***	0.023	(0.091, 0.182)
constant	4.003***		0.031	(3.943, 4.063)
N	703			
R-squared	0.5183			

Statistical Significance: *** p ≤ 0.001, ** p ≤ 0.01, * p ≤ 0.1

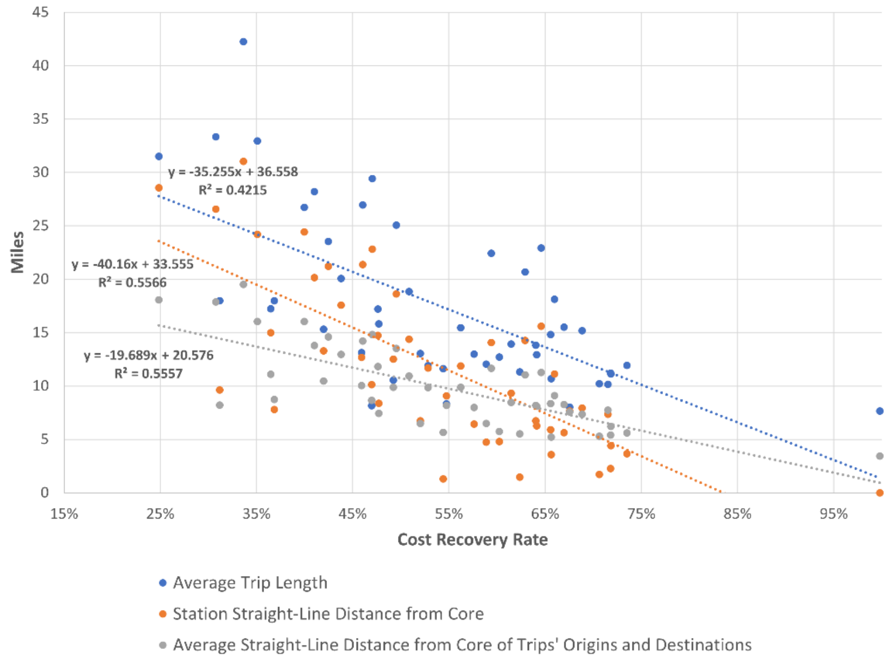
Table 12: Origin-Destination Cost Recovery Model Results — MARTA

In both systems, trip length is negatively associated with cost recovery, though the magnitude is markedly greater for MARTA. Whereas a 1% increase in trip length is associated with a 0.04% *decrease* in cost recovery at BART, the impact is eleven-fold (at a value of -0.44%) for MARTA. By contrast, the association between cost recovery and average Euclidean distance from the core of the origin and destination stations is opposing between the two networks. A 1% increase in this average distance is associated with a 0.41% *decrease* in OD cost recovery for BART but a 0.14% *increase* for MARTA. Apart from the opposing directions and magnitudes, the core orientation of travel more statistically significantly explains cost recovery for BART than does trip length — perhaps because trip length is incorporated in the fares.

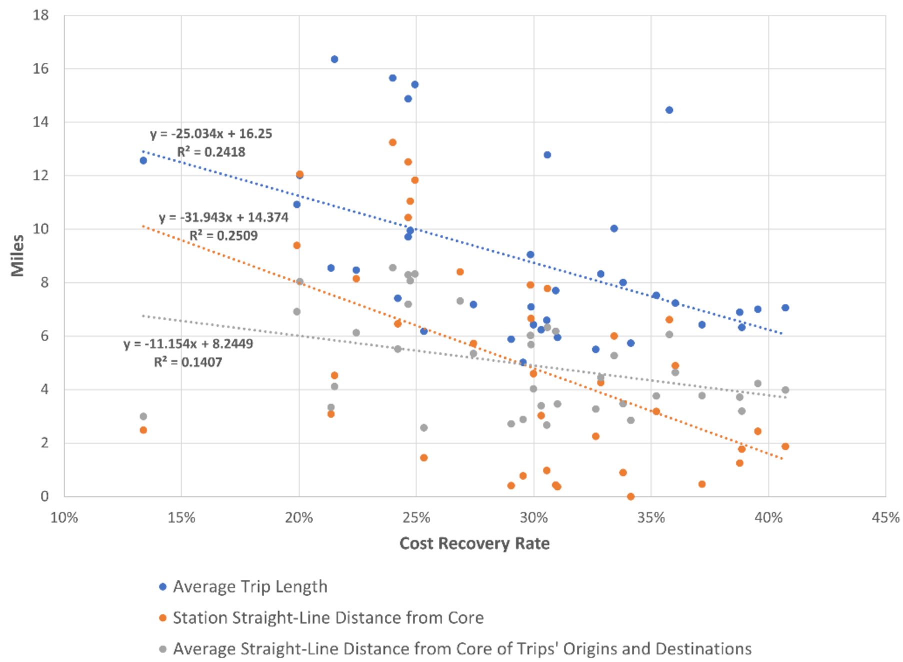
On their face, these results imply that, holding all else equal, "suburban" travelers are less subsidized than urban or monocentric travelers in the MARTA system; to the extent they are more subsidized, it is through trip length. Yet, in reviewing Figure 8, there is clearly a logarithmic negative relationship between travel monocentricity and cost recovery; indeed, the one-to-one logarithmic correlation is -0.607 for BART and -0.47 for MARTA, and both are statistically significant at a $p \leq 0.001$ level. Upon further review, this result in the MARTA regression is explained by the outliers in Figure 8; several short, but distant from the core, OD pairs have high OD cost recoveries. Furthermore, these OD pairs are concentrated along the northern portion of the network. With local travelers sharing the cost of these stations and links with longer distance travelers but paying the same rate, there is a higher cost recovery for local OD pairs concentrated in this area. And because trip length and travel monocentricity have similar relationships with cost recovery apart from these observations, the regression accounts for these outliers by making the travel monocentricity variable positive and amplifying the magnitude of trip length to compensate. An expanded model that controls for additional centers of activity — for example, multiple of the travel monocentricity terms, one each for each activity center — may overcome this limitation. BART also serves additional activity centers, most notably downtown Oakland, so its results may also benefit from an expanded model. In fact, West Oakland Station was selected as the core station because it is the point of ridership crush load and is the one stop that separates downtown Oakland and downtown San Francisco, so is central to the activity centers. Affording each activity center its own monocentric proxy variable may be an improvement.

In addition, spatial patterns of travel in the MARTA network are markedly more dispersed than in the BART network. The more evenly dispersed travel in a network is, the less influence a node's distance from the core of the network will have on the cost recovery, all else being equal. Finally, in a flat-rate fare structure environment, like MARTA, *all else being equal*, every additional mile traveled on a network will be associated with a lower cost recovery. That is, the consumer is consuming more wear, labor-hours, etc., and not paying anything more for it, unlike in the BART network. Taken together, trip length will reasonably have a greater influence on cost recovery outcomes in the MARTA network than in the BART network.

Scaled to stations and links — represented in **Figures 9** and **10**, respectively — average trip length, station or link straight-line distance from the core station, and the average origin and destination straight-line distance from the core station for trips associated with a station or link, negatively correlate with cost recovery. The one exception is links of the MARTA network, in which there is a positive correlation between cost recovery and trips' average length and core-orientation. However, as suggested by the R-squared value, these relationships for MARTA links are also weak. Lastly, while BART is more efficient on aggregate than MARTA, when accounting for the scale difference in these charts, they reinforce an observation in Figure 8: MARTA's inefficiency is more geographically equitable. That is, there is far less link and station average cost recovery variability and spread for MARTA than for BART.

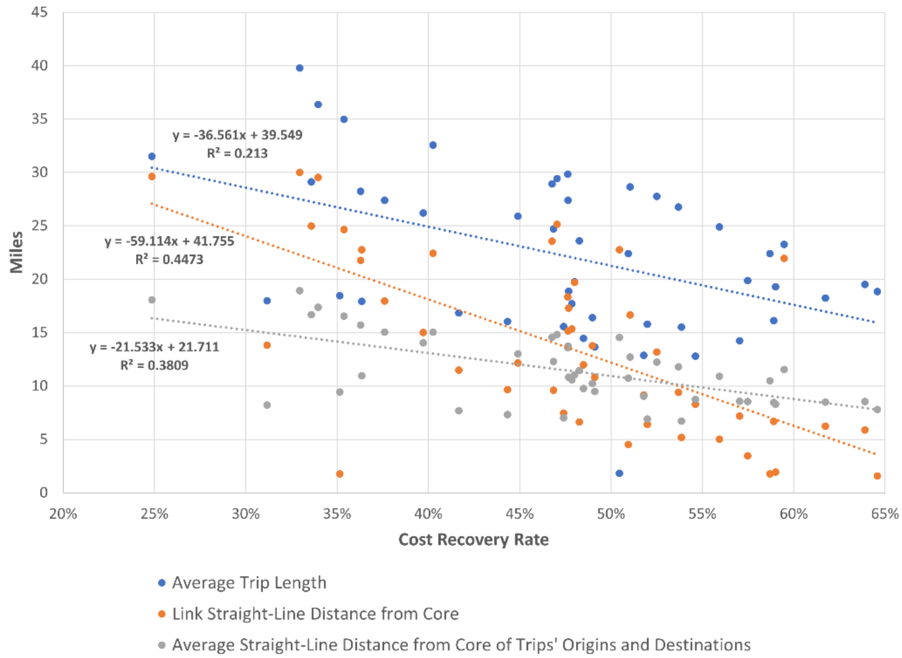


(A) BART

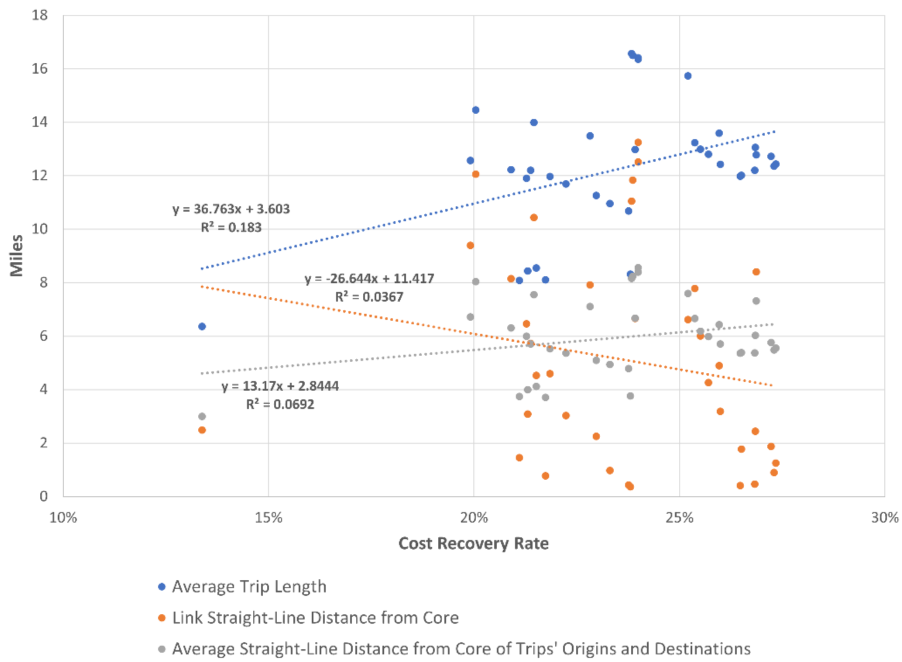


(B) MARTA

Figure 9: Station Cost Recoveries and Spatial Variables



(A) BART



(B) MARTA

Figure 10: Link Cost Recoveries and Spatial Variables

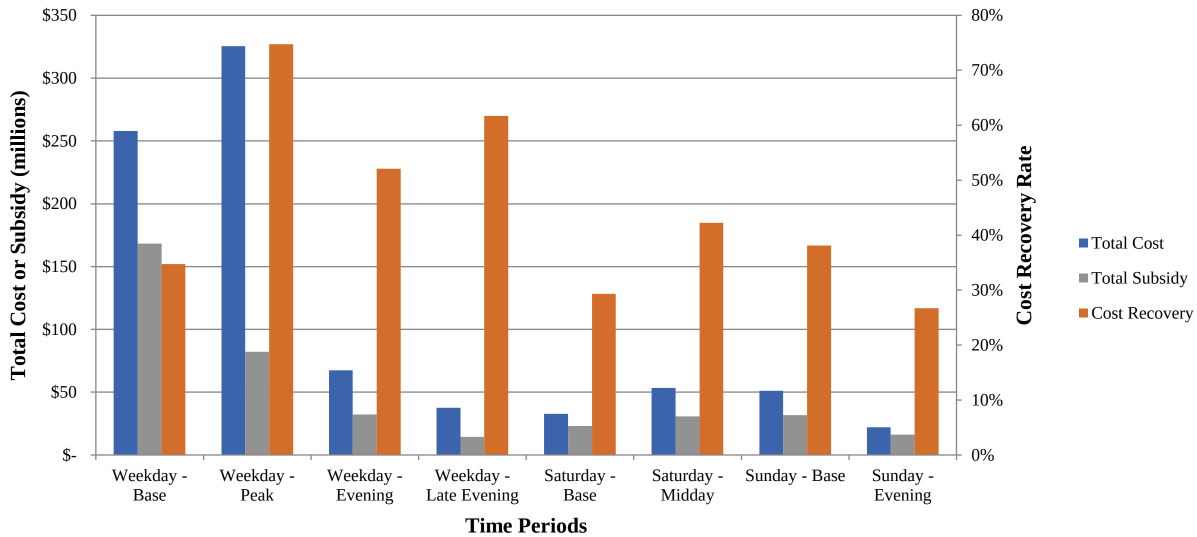
To understand the relationship between the socioeconomic makeup of ridership and spatial incidence of subsidies, I ran pairwise correlations at the station level — a matrix for which is in the **Appendix A**, wherein cells shaded blue correspond to BART and those shaded red correspond to MARTA. BART riders who identify as Asian and Pacific Islander or Hispanic are, as a group, negatively associated with station cost recoveries in a statistically significant way. The only racial group associated with positive station cost recoveries in a statistically significant way identify as White. In other words, Asian and Pacific Islander and Hispanic populations consume more travel through stations whose trips receive more subsidy than average, while the opposite relationship is the case for White riders. Therefore, contrary to past research, subsidy patterns are progressive with respect to race. In the MARTA network, there is no statistically significant racial disparity in station cost recovery variability.

In both networks, only persons with a reported household income between $60,000 and $75,000 have a statistically significant relationship with station cost recoveries. However, the direction of influence is opposing; at BART it is negative, while at MARTA it is positive. Other interactions between socioeconomic and spatial variables are abound, including that in both networks, stations with higher shares of Asian and Pacific Islander riders and higher-income riders, are associated with less core-centric travel. At MARTA, the same pattern is true for stations with high shares of riders who identify as Hispanic, White, or Other or Mixed, and the same racial groups have a statistically significant positive relationship with average trip lengths when aggregated to stations. Only stations with a high share of Black riders have a statistically significant relationship with these variables in the opposing direction; the stations they use are associated with travel that is more urban core-centric and of a shorter distance.

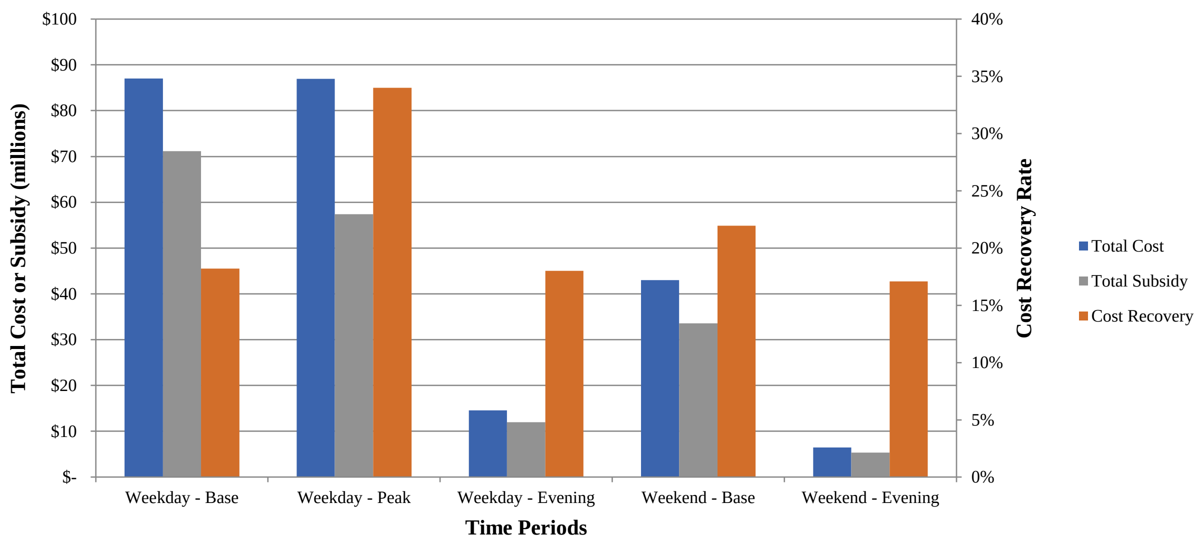
Importantly, these socioeconomic findings are based on aggregate patterns at the station level, so they may not depict true travel patterns of different groups (i.e., the ecological fallacy). That is, different demographic groups may consume OD travel differently across stations, making what is true about the whole or average population of a station not necessarily reflective of its different demographics. Nevertheless, these cumulative results clearly show a spatial aspect of subsidy patterns with outlying stations and links receiving more subsidy, and the socioeconomic correlation with station subsidy patterns suggests that subsidies are partially progressive.

7. Results — Temporal Analysis

Although the weekday peak period is the most expensive to serve for BART and about as costly as the weekday base period for MARTA, it recovers the highest proportion of its costs through fare payment in both systems — 75% at BART and 34% at MARTA. In fact, the weekday peak period recovers twice as much of its costs compared to the weekday base period in both systems — so much that even the monetary subsidy expended on weekday peak service is less than weekday base service. This is even after accounting for semi-fixed asset costs, including, for example, BART’s peak period being solely responsible for the annualized purchase price of 37% of the agency’s railcars and responsible for 57% of the annualized purchase of railcars overall through its share in use of all railcars (Mallett, 2022). Other time periods’ total and marginal costs — that is, the additional costs of serving a time period relative to the next-highest-cost time period — are much smaller compared to weekday base and peak periods. Even so, it is noteworthy that the weekend base period at MARTA and Sunday base period at BART also have higher cost recoveries than the weekday base period. These findings are reflected in **Figure 11**, which shows the total costs that are allocated (left axis), cost recovery (right axis), and resulting monetary subsidy (left axis) of different time periods in the BART and MARTA systems.



(A) BART



(B) MARTA

Figure 11: Time Period Costs, Subsidies, and Cost Recoveries

Finally, in running a pairwise correlation of the relationship between cost recovery and select travel pattern variables *across time periods* I found that cost recoveries increase in a statistically significant way with time periods that have higher trip and passenger-mile generation in both networks. For MARTA, cost recovery also has a positive and statistically significant relationship with average trip length — meaning, time periods with longer trip lengths, on average, generate higher cost recovery. However, travel centrality around the urban core does not appear to have a strong relationship with temporal cost recovery patterns. Thus, to the extent cost recovery is influenced by how "suburban" a trip is, the influence does not vary across time. The matrix in **Appendix B** shows the findings — again, with BART highlighted in blue; MARTA, in red.

8. Discussion and Conclusion

In the preceding analysis, I employed a long-run partially allocated cost allocation model (Mallett, 2022) and OD trip and fare data to measure the spatial and temporal distribution of travel subsidies in the BART and MARTA networks. I find that travelers with trips less concentrated around the urban core and those who travel during off-peak travel periods pay a lower share of their costs relative to travelers whose travel is concentrated in the urban core and during peak travel periods — meaning the former are subsidized more. In addition, when OD cost recoveries are weighted to trip count and averaged at the station level, I find that subsidy patterns are moderately progressive.

My spatial analysis included unweighted and weighted OD trip cost recoveries, the latter of which informs the spatial incidence of subsidies at the link and station level. Unweighted OD trip cost recoveries explain the *occasions* of subsidies. They take the distribution of costs and travel in the network as fixed and calculate what the corresponding cost recovery is of a one-off OD trip. Weighted OD trip cost recoveries explain the OD *incidence* of subsidies. They describe the pattern of subsidies based on the aggregate consumption of OD trips.

My regression analysis (unweighted OD trips) reveals that trips being less concentrated around the core explains subsidy patterns more than trip length in the BART network, both in terms of magnitude and statistical significance. In MARTA, both variables have a statistically significant negative one-to-one relationship with cost recovery, but the core-centrality of an OD pair has a positive association in the regression. I posit that this is explained both by my model not accounting for the existence of multiple activity centers in the MARTA network and fare structure. On the latter point, MARTA's flat fare structure means longer distance trips that use more miles of service irrespective of cost will cover less of their cost, while BART's structure being distance-based mitigates the influence of distance but does not absolve it since outlying areas cost more per-rider and distance-based fares do not account for this. When aggregate patterns (weighted OD trips) are averaged at the link and station level, I find that trips associated with outlying links and stations receive more subsidy than trips associated with inner links and stations, on average, in both networks.

My station-level pairwise correlation shows that stations with higher shares of Asian and Pacific Islander or Hispanic riders have a lower cost recovery profile, and stations with higher shares of White riders have a higher cost recovery profile, than average. Only stations with higher shares of an income between $60,000 and $75,000 have a statistically significant association with cost recovery; it is negative for BART and positive for MARTA. These findings have mixed consistencies with past research. Many scholars find that flat fares are regressive on income and racial bases; lower-income persons and marginalized minorities pay disproportionately more due to traveling shorter distances and during off-peak periods that are less costly, but they pay the same for such travel (e.g., Brown, 2018; Cervero, 1981; Taylor et al., 2000). My findings are consistent with this in that I find that BART, an agency with distance-based fares, has cost recovery patterns that are progressive; some marginalized racial groups consume travel through stations with low cost recoveries, while persons who identify as White consume more travel through stations with high cost recoveries, on average. With MARTA, I find no cost recovery disparity across racial groups, despite finding that travel patterns like trip lengths have racial variability at the station level. Thus, despite racial minorities traveling shorter distances for the same base price, I do not find cost recovery disparity. As acknowledged previously, the ecological fallacy caveat applies; what is true about stations' populations may not be true about segments of that population. Additional inconsistencies in my findings include that, whereas I measure equity based on cost recovery, Brown (2018) bases equity on the fare paid per mile of travel, and Taylor et al. (2000) focus on expenditure patterns. My measurement of equity is closest to Cervero (1981), though he used bus yard cost centers as his spatial unit of analysis rather than unique stations or stops and focused on bus transit rather than rail transit. Accordingly, my spatial units of analyses are more granular than his, and I focus on a different mode.

My temporal analysis involved calculating the cost recovery of different operating time periods (aggregate fares divided by allocated costs). I find that the weekday peak period, while the costliest to operate for BART and about as costly as the weekday base period for MARTA, is the most efficient in both networks. This runs counter to findings from past research. Cervero (1981), Parody et al. (1990), and others show that the peak period is the costliest to serve in net terms — meaning even after cost recovery is considered. This is partly because these other studies include fixed asset costs in their cost allocations, whereas I do not. However, as shown in the cost allocation research (Mallett, 2022), this divergence in findings is also explained by the fixed headway schedule that BART and MARTA operate throughout the day. Because of this practice, the weekday peak period has a low marginal cost of operating compared to what is typical of traditional commuter rail systems and bus networks. At BART, the additional cost of serving the weekday peak period is almost exclusively driven by capital because the agency resizes the length of its trains without compromising frequency of service. By comparison, MARTA has a nominal number of additional trains in service during the weekday peak period relative to weekday base period and does not resize its trains, leading to an even smaller marginal gross cost difference — so small that the additional operating hours associated with base service make it slightly more expensive than the weekday peak period. A growing equity argument for more off-peak service is that there is no marginal cost to providing it because assets sit unused otherwise. However, the BART and MARTA experience show this is false because there is a net increase in off-peak marginal cost in such case through it reducing the marginal cost of the weekday peak period.

The correlations between time period cost recoveries and travel pattern variables associated with different time periods indicate that the number of trips and passenger-miles generated across time periods explain much of the difference in time period cost recoveries. Given the relatively fixed headway schedules of both agencies, this makes sense; holding all else equal, more trips and trip-miles will result in more revenues, which will increase cost recovery. This is especially true for BART, given that its fare structure is distance based. By contrast, the core-centrality of travel does not have any statistically significant influence on temporal cost recovery variation. Indeed, these terms vary little across time periods. Taken together with the spatial finding that cost recovery is negatively associated with decentralization of travel in both networks, this suggests that outlying areas are subsidized regardless of the time of travel. Thus, while there is geographic incidence of subsidies, it is not temporally variable. Even so, research that explicitly interacts space and time could more decisively test this.

While this research shows that core stations and links and the weekday peak period are less subsidized than suburban stations and links and other time periods, there may be other purposes for charging premiums to travel in these areas and during these times. My cost allocation study (Mallett, 2022) does not account for externalities, including the costs of congestion beyond production costs of serving it. Core areas of these networks and peak times of travel may generate so much crowding and inconvenience costs for passengers — for example, BART passengers departing downtown San Francisco during the evening commute often backtrack to secure a seat — that a premium is warranted for use in these areas or during these times to internalize delay time costs or manage demand. In addition, minimum fares or trip length restrictions are often used to manage capacity or product differentiate transit services, such as a regional service relative to a local service. Among other examples, BART's minimum fare assumes a trip of at least six miles to distinguish its regional focus from peer agencies' local focus, and New York's Metro-North Railroad does not sell fares for travel between Harlem/125th Street and Grand Central Terminal stations to distinguish its commuter rail orientation from New York City Transit's local travel focus. In this analysis, I do not control for these policies or objectives, which can inflate the cost recovery level found for peak-period travel and travel in core areas of each network.

In addition, as is true with most fare equity studies, I do not test how demand would respond to a fare structure that corrects for spatial and temporal inequities found. It is conceivable that changing fares to achieve cost recovery equity in this cross-section would result in a change in travel patterns that merely redistributes cost recovery inequities. Furthermore, core segments of networks benefit from ridership associated with both core-centric and less core-centric travel; riders who travel within the core "overpay" partly because they are sharing links and stations with riders who "underpay" in terms of cost recovery equity. Changes to fares may change this pattern as well.

Nonetheless, I have shown that, in the BART and MARTA networks, travel to and from outlying areas and during off-peak times receive more subsidy than travel in core areas and during the weekday peak travel period. Further, the spatial concentration of subsidies is moderately progressive. This suggests that suburban and exurban location choices may be enabled through transport subsidies in these networks, that the socioeconomic beneficiaries are disadvantaged populations, and that prioritizing constant service frequency throughout the day mitigates the net marginal cost of serving peak period travel. Future research should explore whether other modes of transport have spatial and temporal dimensions of subsidies to help explain if transport subsidies more broadly effect location choice and times of travel.

9. Acknowledgements

My doctoral dissertation and education, which this research is a product of, were generously funded through merit-based awards provided by the American Public Transportation Foundation, California Transportation Foundation, Railway Association of Southern California, United States Department of Transportation Dwight David Eisenhower Transportation Fellowship Program, and the University of Southern California. I extend sincerest appreciation to these organizations and their members, donors, and funding sponsors. Special thanks are also owed to many colleagues and members of my dissertation committee who provided valuable feedback, including Marlon Boarnet, Nancy Brooks, Genevieve Giuliano, Nicholas Klein, Clemens Pilgram, and Brian Taylor; and to staff of BART and MARTA.

Bibliography

Bandegani, Milad, and Meisam Akbarzadeh. "Evaluation of horizontal equity under a distance-based transit fare structure." *Journal of Public transportation* 19.3 (2016): 10.

Brown, Anne E. "Fair fares? How flat and variable fares affect transit equity in Los Angeles." Case Studies on Transport Policy 6.4 (2018): 765-773.

Cervero, Robert. "Flat versus differentiated transit pricing: what's a fair fare?." Transportation 10.3 (1981): 211-232.

Cervero, Robert, and Martin Wachs. "An answer to the transit crisis: the case for distance-based fares." *Journal of Contemporary Studies* 5.2 (1982): 59-70.

Cherwony, W., and S. R. Mundle. "Peak-Base Cost Allocation Models." *Transportation Research Record 663* (1978): 52–56.

Cherwony, Walter, and Subhash R. Mundle. "Transit cost allocation model development." *Transportation Engineering Journal of ASCE* 106.1 (1980): 31-42.

Farber, Steven, et al. "Assessing social equity in distance based transit fares using a model of travel behavior." *Transportation Research Part A: Policy and Practice* 67 (2014): 291-303.

Frisk, Mikael, et al. "Cost allocation in collaborative forest transportation." European Journal of Operational Research 205.2 (2010): 448-458.

Hodge, David C. "Fiscal equity in urban mass transit systems: A geographic analysis." Annals of the Association of American Geographers 78.2 (1988): 288-306.

Iseki, Hiroyuki. "Equity in regional public transit finance: Tradeoffs between social and geographic equity." Journal of Urban Planning and Development 142.4 (2016)

Jones Jr., David W. "Urban transit policy: An economic and political history." (1985).

Guajardo, Mario, and Mikael Rönnqvist. "A review on cost allocation methods in collaborative transportation." International transactions in operational research 23.3 (2016): 371-392.

Mallett, Zakhary. "Pricing Equity, Expenditure Equity, and Pricing-Expenditure Equity of Transit." (2023a) [Manuscript submitted for publication.]

Mallett, Zakhary. "Bridging equity and efficiency in transit subsidy policy research: A literature review." (2023b) [Manuscript submitted for publication.]

Mallett, Zakhary. "Spatial and Temporal Variability of Rail Transit Costs and Cost Effectiveness." Transportation Research Record (2022): 03611981221104807.

Morales Sarriera, Javier, and Frederick P. Salvucci. "Rising costs of transit and Baumol's cost disease." Transportation Research Record 2541.1 (2016): 1-9.

Nuworsoo, Cornelius, Aaron Golub, and Elizabeth Deakin. "Analyzing equity impacts of transit fare changes: Case study of Alameda–Contra Costa Transit, California." *Evaluation and program planning* 32.4 (2009): 360-368.

Parody, Thomas E., Mary E. Lovely, and Poh Ser Hsu. "Net Costs of Peak and Offpeak Transit Trips Taken Nationwide by Mode." Transportation Research Record.1266 (1990): 139-145

Pickrell, Don H. "Rising deficits and the uses of transit subsidies in the United States." Journal of Transport Economics and Policy (1985): 281-298.

Reilly, J. M. Transit Costs During Peak and Off-Peak Hours. Transportation Research Record 625 (1977): 22–26.

Rosenthal, Edward C. "A cooperative game approach to cost allocation in a rapid-transit network." Transportation Research Part B: Methodological 97 (2017): 64-77.

Rubensson, Isak, Yusak Susilo, and Oded Cats. "Is flat fare fair? Equity impact of fare scheme change." Transport Policy 91 (2020): 48-58.

Savage, Ian. "The analysis of bus costs and revenues by time period. II. Methodology review." Transport Reviews 9.1 (1989): 1-17.

Serebrisky, Tomás, et al. "Affordability and subsidies in public urban transport: What do we mean, what can be done?." Transport reviews 29.6 (2009): 715-739.

Taylor, Brian D., Mark Garrett, and Hiroyuki Iseki. "Measuring cost variability in provision of transit service." Transportation Research Record 1735.1 (2000): 101-112.

Taylor, Brian D., and Alexandra Tassiello Norton. "Paying for transportation: What's a fair price?." Journal of Planning Literature 24.1 (2009): 22-36.

United States Department of Transportation, Federal Transit Administration. "2018 National Transit Summaries and Trends: Appendix." *National Transit Database*, December 2019, https://www.transit.dot.gov/sites/fta.dot.gov/files/2020-06/2018-ntst-appendix.pdf.

United States Department of Transportation, Federal Highway Administration. "Table HF-1 — Highway Statistics 2019" *Office of Highway Policy Information*, 2019, https://www.fhwa.dot.gov/policyinformation/statistics/2019/.

Wachs, Martin. "US transit subsidy policy: In need of reform." Science 244.4912 (1989): 1545-1549.

Zhao, Pengjun, and Yixue Zhang. "The effects of metro fare increase on transport equity: New evidence from Beijing." Transport Policy 74 (2019): 73-83.

	percentpaid_s	trips	triplength_average	distancecore_s	distancecore_average	r_ai_p	r_api_p	r_black_p	r_hisp_p	r_other_p	r_white_p	i_0to25_p / i_0to20_p*	i_25to35_p / i_20to30_p*	i_35to40_p / i_30to40_p*	i_40to50_p	i_50to60_p	i_60to75_p	i_75to100_p	i_100to150_p / i_100to120_p*	i_150plus_p / i_120plus_p*
percentpaid_s	1																			
	1																			
trips	0.36**	1																		
	0.33*	1																		
triplength_average	-0.65***	-0.28*	1																	
	-0.44**	1.44	1																	
distancecore_s	-0.75***	-0.38**	0.9***	1																
	-0.5**	0.26	0.93***	1																
distancecore_average	-0.75***	-0.28*	0.92***	0.99***	1															
	-0.38*	0.13	0.94***	0.93***	1															
r_ai_p	0.07	0.06	-0.01	-0.07	-0.09	1														
	-0.01	0.19	0.18	0.15	0.26	1														
r_api_p	-0.41**	0.07	0.16	0.43**	0.46**	-0.09	1													
	0.09	0.25	0.5**	0.4*	0.52***	0.37*	1													
r_black_p	0.12	-0.09	-0.08	-0.22	-0.21	0.06	-0.4**	1												
	-0.19	-0.13	-0.5**	-0.39*	-0.48**	-0.3*	-0.77***	1												
r_hisp_p	-0.34*	-0.1	-0.05	0.09	0.08	0.16	-0.14	0.45**	1											
	-0.18	0.14	0.45**	0.39*	0.46**	0.3*	0.62***	-0.56***	1											
r_other_p	0.21	0.09	-0.21	-0.33*	-0.31*	-0.12	-0.41**	0.24	0.17	1										
	0.01	0.17	0.5**	0.36*	0.46**	0.15	0.69***	-0.69***	0.87***	1										
r_white_p	0.4**	0.04	-0.04	-0.23	-0.26*	-0.06	-0.45**	-0.52***	-0.67***	0.03	1									
	0.23	0.07	0.43**	0.33*	0.4*	0.23	0.62***	-0.97***	0.39*	0.53***	1									
i_0to25_p	0.02	-0.09	-0.28*	-0.29*	-0.31*	0.16	-0.17	0.45**	0.43**	0.31*	-0.37*	1								
i_0to20_p*	-0.15	-0.31*	-0.6***	-0.5**	-0.51**	-0.12	-0.49**	0.71***	-0.28*	0.49**	-0.7***	1								
i_25to35_p	0.07	-0.14	-0.29*	-0.25*	-0.27*	0.22	-0.29*	0.46**	0.62***	0.24	-0.37*	0.76***	1							
i_20to30_p*	-0.2	-0.04	-0.34*	-0.2	-0.32*	-0.16	-0.53***	0.66***	-0.33*	-0.45**	-0.64***	0.49**	1							
i_35to40_p	-0.03	-0.15	-0.28*	-0.13	-0.15	0.03	-0.07	0.57***	0.62***	-0.01	-0.59***	0.5***	0.6***	1						
i_30to40_p*	-0.26	-0.12	-0.24	-0.12	-0.19	-0.07	-0.58***	0.77***	-0.34*	-0.47**	-0.77***	0.54***	0.75***	1						
i_40to50_p	0.01	-0.07	-0.28*	-0.18	-0.18	-0.02	0.02	0.56***	0.57***	0.24	-0.65***	0.54***	0.51***	0.7***	1					
	0.26	0.06	-0.33*	-0.25	-0.25	-0.05	-0.38*	0.45**	-0.38*	-0.36*	-0.42**	0.16	0.26	0.42**	1					
i_50to60_p	-0.03	0.04	-0.29*	-0.17	-0.18	-0.01	-0.02	0.57***	0.64***	0.08	-0.65***	0.4**	0.45**	0.68***	0.75***	1				
	0.14	0.06	-0.13	-0.24	-0.21	-0.34*	0.02	0	0.14	0.19	-0.03	-0.2	-0.33*	-0.27	-0.09	1				
i_60to75_p	-0.4**	-0.15	0.09	0.23	0.22	-0.22	0.3*	0.17	0.44**	-0.11	-0.54***	-0.01	0.03	0.32*	0.33*	0.55***	1			
	0.36*	0.22	0.18	0.05	0.12	0.04	0.29*	-0.57***	0.19	0.35*	0.6***	-0.6***	-0.58***	-0.69***	-0.27	0.35*	1			
i_75to100_p	-0.08	0.1	0.25	0.32*	0.31*	0.03	0.22	-0.31*	-0.21	-0.16	0.12	-0.58***	-0.45**	-0.43**	-0.42**	-0.34*	0.14	1		
	0.14	0.24	0.54***	0.41**	0.48**	0.28*	0.66***	-0.79***	0.39*	0.53***	0.76***	-0.74***	-0.74***	-0.78***	-0.47**	0.07	0.55***	1		
i_100to150_p	-0.08	0.11	0.4**	0.3*	0.33*	-0.05	0.2	-0.58***	-0.65***	-0.18	0.52***	-0.65***	-0.71***	-0.77***	-0.74***	-0.81***	-0.4**	0.4**	1	
i_100to120_p*	-0.04	0.14	0.65***	0.56***	0.6***	0.21	0.72***	-0.85***	0.45**	0.55***	0.82***	-0.67***	-0.65***	0.78***	-0.57***	-0.04	0.42**	0.84***	1	
i_150plus_p	0.2	0.08	0.16	0.01	0.01	-0.02	-0.13	-0.56***	-0.72***	-0.12	0.8***	-0.56***	-0.58***	-0.72***	-0.81***	-0.82***	-0.64***	0.17	0.74***	1
i_120plus_p*	-0.82	-0.04	0.56***	0.47**	0.5**	0.18	0.58***	-0.85***	0.37*	0.5**	0.87***	-0.67***	-0.65***	-0.69***	-0.51**	-0.07	0.42**	0.72***	0.85***	1

Appendix A: Matrix of Spatial and Socioeconomic Variables Across Stations

percentpaid_t				
0.39*	trips_t			
0.84*				
0.7*	1***	pxmiles		
0.86*	1***			
0.33	0.22	0.24	triplength_average	
0.92*	0.97***	0.98**		
0.01	0.01	0.03	0.89**	distancecore_average
0.56	0.31	0.34	0.34	

Appendix B: Matrix of Spatial Variables Across Time Periods