Transit dependents, choice riders, and service criticality: an analysis of the determinants of bus ridership in the Philadelphia Region

Erick Guerra
https://orcid.org/
0000-0002-7769-2581

FINAL RESEARCH REPORT

Contract # 69A3551747111

Research Assistants

Han Qiao
Department of City and Regional Planning
University of Pennsylvania, Philadelphia, Pennsylvania, 19104
Email: h.qiao@mail.utoronto.ca

Xuezhu Zhao
Department of City and Regional Planning
University of Pennsylvania, Philadelphia, Pennsylvania, 19104
Email: gxzhao@alumni.upenn.edu

Andrew Tzuyuan Wang
Department of City and Regional Planning
University of Pennsylvania, Philadelphia, Pennsylvania, 19104
Email: tzuyuan@alumni.upenn.edu

Qi Si
Department of City and Regional Planning
University of Pennsylvania, Philadelphia, Pennsylvania, 19104
Email: qisi@alumni.upenn.edu

Overview

This report presents the results of two interrelated projects on transit ridership in the SEPTA region. The first focuses on developing and testing an empirically based theory of transit-dependency using a predictive model of transit mode choice in the Philadelphia region. The second examines station-level shifts in transit use in response to the Covid-19 outbreak in the Philadelphia region and how these correspond with the distribution of transit dependents throughout the region.

Part I

As local, state, and federal agencies began investing substantial resources into subsidizing transit in the 1960s and 70s, public documents argued that transit agencies should focus on attracting choice riders instead of dependent riders, who have no alternatives and use transit regardless of service quality. After six decades, the definitions, uses, and implications of the terms choice and dependent rider have remained consistent in the academic and professional literature. These definitions, however, lack a strong theoretical grounding or empirical evidence to support them. Using travel diary data from the Philadelphia region, I estimate discrete choice models to identify choice riders, who I define as those who have close to a 50% probability of choosing between a car or transit for a given trip. The Philadelphia region, which has a diverse range of transit users and transit services, is an ideal place to develop and fit an empirical model of choice ridership. Attributes assumed to be associated with dependent riders, such as a lack of a car, low income, and being a racial or ethnic minority, are much more prevalent among choice riders than the general metropolitan population. Choice riders are also diverse, with a mix of racial backgrounds, income levels, educational attainment, and access to private cars. Transit dependency, by contrast, is rare. The lowest and highest income residents generally only choose transit when service quality is high, and transit is cost- and time-competitive with the car.

Part II

The Covid-19 pandemic outbreak led to a long-lasting shock that jeopardizes health, economy, education, cultural and social activities around the world. People adjusted their travel behaviors according to new lifestyles as social distancing and travel restrictions were being implemented to prevent the spread of the virus. Huge declines in all modes of transportation were seen across the world and buses were among the most impacted modes of transportation. In this paper, we try to

understand how different types of places lost bus ridership at the beginning of the pandemic by unpacking how the impact of demographics, socioeconomics and land use factors on bus ridership changed during the initial few weeks of the pandemic outbreak. We adopt Philadelphia as a case context and utilize a mixed-effect multilevel linear regression model to reveal the underlying correlations. Results show that factors negatively correlated with bus ridership (i.e. income, precipitation) became stronger in driving bus ridership and factors positively correlated with bus ridership (i.e. job accessibility, population density, parking cost, transfer station, weekday) became weaker after the pandemic outbreak. The result emphasizes the challenges that transit agencies face, especially during the immediate period after a society-wide change, and sheds light on future transit network planning and policies in providing more resilient and equitable travel mode choices in challenging times.

Part I. What the heck is a choice rider? A theoretical framework and empirical model

1.1 Introduction

The term "choice rider" enters the English lexicon in the early 1960s (Google, 2021). Choice riders first appear in technical transportation-planning documents alongside captive, necessity, or dependent riders to categorize existing and future transit users. For example, the 1961 Pittsburgh Area Transportation Study groups metropolitan transit users into captive and choice riders based on private vehicle availability and describes a decline in captive ridership as more households acquire cars. While many documents provide circular definitions—a choice rider is someone who can choose to use transit—several keywords frequently coincide with each term. Choice riders have cars, licenses, and suburban homes in wealthy neighborhoods. They are white, white-collar, male workers, who take trains to downtown jobs. Captive riders are poor, racial minorities, housewives, the old, the young, the carless, and the handicapped. They rely on urban bus services to accomplish their daily tasks regardless of service quality.

The introduction of the term choice rider coincides with a broad shift in the provision of US transit. Prior to World War 2, a combination of private for-profit companies and city agencies provided transit services throughout the US. Although vehicle registrations were already rapidly increasing in the early 20th century, the combination of increased automobility and suburbanization in the post-war era contributed to a substantial reduction in transit ridership, the closure of multiple transit lines, the public takeover of many private transit companies, and an increasing need for public subsidies to maintain remaining transit services. In greater Philadelphia, the city of Philadelphia and State of Pennsylvania began to subsidize transit services as early as 1960 (Hepp, 2018). In 1963, the State formed the Southeastern Pennsylvania Transportation Authority (SEPTA), which began to take over transit operations from the patchwork of transit companies throughout the region. In 1964, the Federal Government created the Urban Mass Transportation Administration (now, Federal Transit Administration) and began to pass a series of bills to support and subsidize urban transit systems throughout the country (Federal Transit Administration, n.d.). As public agencies continued to take over and subsidize transit operations throughout the 1960s and 1970s, the number reports referencing choice and dependent riders increases and peaks in 1977 (Google, 2021).

Within these technical reports, the differentiation between choice and captive riders has implicit and explicit connections to the economic and environmental justifications for subsidizing transit. One common line of argument is that, "[t]he captive rider has no choice but to wait, regardless of the headway between buses or trains, but the choice rider can get back in his car and drive (Bates, 1981, p. 13)." The captive rider market, "…will always exist…" but the choice rider market "…will exist only as long as transit service is attractive (Keefer et al., 1963, p. 58)." If transit ridership is to increase or draw passengers away from cars and thus reduce associated pollution and congestion, transit agencies should ignore captive riders and focus on choice riders. In an early article on the economics and political economy of transit subsidies, Haines (1978, pp. 64–66) argues there is no economic justification for subsidizing transit for captive riders and little reason to do so, since "…in the nature of things, captive riders are not a particularly potent political force." The direct implication of these early uses of the terms choice and captive rider is that agencies should generally focus investments and service improvements on suburban rail services to downtown job centers in wealthier, whiter suburban communities. Urban bus services in low-income and minority neighborhoods can be safely ignored.

After six decades, the definitions, uses, and implications of the terms choice and captive rider have persisted, though the term dependent rider has largely supplanted the term captive rider. These definitions, however, are theoretically weak and empirically inaccurate. For example, just 18% of US households earning below $25,000 per year do not have a car. The adults in these low-income carless households take 25% of trips by transit compared to 27% by car (U.S. Department of Transportation, 2017). The uses and implications of the terms may have also contributed to racist public policies. For example, the Los Angeles Bus Riders Union's sued the Metropolitan Transit Authority in 1994. The plaintiffs argued that the agency's focus on rail investments at the expense of bus investments violated the 14th amendment and 1964 Civil Rights Act by discriminating against the racial and ethnic minority groups that disproportionately used buses. The lawsuit led to a court injunction and reforms to improve bus services and stabilize transit fares (Elkind, 2014; Grengs, 2002). Additionally, mischaracterizations of choice ridership may encourage transit investments that not only attract fewer transit riders per dollar invested but also fail to draw as many transit users out of cars.

The purpose of this paper is to develop a theoretically robust and measurable definition of choice transit riders, estimate models of choice ridership, and describe the factors associated with choice transit travel. In the proceeding section, I summarize academic definitions and uses of the terms choice and dependent riders. The academic literature has a different focus than technical planning reports, but generally accepts and frequently expands on early definitions of choice and dependent riders. Next, I describe my methodological approach, definition of choice ridership, case context, data, and model specification. Relying on travel survey diaries from the Philadelphia region, I describe choice transit users as people who have close to a 50% estimated probability of choosing transit instead of a car for a given trip. Philadelphia, which has a high number of transit users that match existing definitions of dependent and choice transit users, is an ideal place to estimate and describe a model of choice transit use. Next, I describe Philadelphia's choice riders and compare them to the general population. Many of the people traditionally associated with transit dependency, such as low-income urban residents, minorities, and those without cars, are most likely to be on the fence about choosing to take transit or a car. The strongest associations with choice ridership relate to high-frequency bus and rail services near residents' trip origins and destinations.

Last, I conclude with takeaways for researchers and policymakers. Existing characterizations of choice riders are almost certainly inaccurate. If attracting people out of cars is a key transit objective, then agencies would do well to focus service improvements in dense urban areas with high concentrations of low-income residents without cars. These are the kinds of places where residents are likeliest to respond to service improvements by riding transit more. Moreover, researchers and policymakers should stop referring to dependent or captive riders altogether. Even in a large city with relatively good transit, the people most likely to be characterized as transit-dependents only take transit consistently when service quality is high enough to make it a reasonable choice.

1.2 Academic references to choice and dependent riders

The academic literature generally follows and expands upon early planning documents' definitions of choice and dependent riders. Specifically, the term transit dependency is associated with keywords, such as carless, low-income, bus, racial minority, age, disability, and travel to places outside of the downtown. Grengs (2002, p. 170) even makes the explicit argument that transit operators have at least some justification in ignoring transit-dependents to focus on luring choice riders out of cars:

> The dilemma of serving either "choice" or "captive" riders gets even more complicated. To lure people out of their cars requires highly attractive service. And attractive service means higher costs for cash-strapped agencies, especially for distant, low-density suburbs. Keeping transit-dependent customers, by contrast, does not require good service because these riders have no other choice.

In terms of overarching research topics, academic papers that reference transit dependency and choice ridership generally either focus on defining transit-user markets or showing unfairness in the transportation system. Many of these studies also reveal that those defined as transit-dependent exercise a substantial amount of choice and frequently rely on cars. Although I focus on findings from the US and Canada below, the terms transit-dependency and choice ridership are also used in a variety of international contexts, including China (Cai et al., 2020; Xiaoshu Cao et al., 2018; Sun & Fan, 2018), India (Cheranchery & Maitra, 2018), Korea (Sohn & Yun, 2009), Australia (Chia et al., 2016), and Colombia (Márquez et al., 2018).

1.2.1 Defining transit markets

Researchers frequently define and group choice and dependent riders as inputs into empirical models or for comparisons of travel behavior across groups. For example, Polzin et al. (2000), divide the US population into choice and dependent riders based on age, driver's license, and household vehicles to compare travel behavior across these groups. Lachapelle et al. (2016), who define transit-dependency by car availability, find that transit dependents participate in more physically active travel than choice riders or car users. Beimborn et al. (2003) add a third category of auto captives and use the three categories (transit dependents, choice riders, and auto dependents) as inputs to improve predictive models of transit ridership in metropolitan Portland. The authors define choice and captivity based on car availability, transit quality, and proximity to a transit stop. Similarly, van Lierop and El-Geneidy (2016) add another category of captive-by-choice riders, who are wealthy enough to own a car but do not, and use these categorizations to develop models predicting customer satisfaction with transit.

Several researchers apply clustering algorithms, such as factor analysis or K-means clustering, to travel diary and other survey data to group and describe various transit markets and submarkets. For example, Krizek and El-Geneidy (2007) use factor analysis to group residents of the Twin

City metropolitan area into four groups, which they define as transit captives, choice riders, potential riders, and auto captives based on the covariance of survey data about travel preferences, views on transit quality, and available transportation modes. Although Krizek and El-Geneidy (2007) further distinguish these four groups as regular and irregular commuters, the authors argue that transit users fall neatly into two main categories with 46% of the sample being captive riders and the remaining 54% being choice riders. Similarly, Zhao, Webb, and Shah (2014) group transit users from customer survey data in Chicago using factor analysis and structural equation models. The authors differentiate between choice and captive riders primarily based on whether they are likely to continue to use transit when they perceive service quality as poor.

Further distinctions within categories are also common. For example, Chia et al. (2016) distinguish between true and nontrue transit captives—similar to Lierop and El-Geneidy's (2016) captive-by-choice riders—based on access to alternative modes of transportation. Jacques et al. (2013) cluster transit users from a travel survey of students, faculty, and staff at McGill University in Montreal into four market segments that they describe as captivity, utilitarianism, dedication, and convenience. Captivity relates to transit users who are dissatisfied with transit and whose transit service is relatively poor, while the other three groups have higher quality transit or higher satisfaction with transit.

1.2.2 Unfairness in transit systems

The terms choice and dependent rider also frequently occur in studies that test or discuss unfairness in transportation systems or policy. Cervero (1981), using data from three Californian transit operators, shows that flat fare systems are less fair to transit dependents who tend to travel shorter distances outside of peak hours, than those based on distance and time of day. Grengs (2001) finds that poorer, minority residents of Syracuse, New York, have worse accessibility to supermarkets by transit than wealthier, whiter residents, who are less dependent on transit. Using similar definitions of transit dependency, Jiao and Dillivan (2013) define transit deserts as places with relatively high shares of transit-dependent individuals but relatively poor transit service. This definition has since been used to identity transit deserts in major cities in Texas (Jiao, 2017) and China (Cai et al., 2020). Comparing spatial relationships between shared-mobility services and transit deserts in New York City, Jiao and Wang (2020) conclude that shared mobility services are mostly located in wealthier neighborhoods that already have good access to transit.

In addition to investigating unfairness, several studies highlight the gap between transit investment priorities and transit's existing customer base. Grengs (2002) examines how Los Angeles' Bus Riders Union pursued a lawsuit claiming that Los Angeles' investments in suburban rail were at the expensive of investments in bus services and discriminated against poor and minority urban bus riders characterized as transit-dependents. Taylor and Morris (2015) expand on this theme using data on transit operations, travel surveys, and a survey of 50 transit agencies. Only a small share of agency representatives view serving the needs transit-dependent populations as an important goal for public transit. As a result, agencies tend to prioritize commuter-oriented rail investments that appeal to wealthier residents with more political capital instead of urban bus services on which transit-dependents rely. These biases may also exist within modes. For example, Wells and Thill (2012) examine whether transit dependent neighborhoods—defined as non-white, poorer, elderly, students, low car—get worse bus service than other neighborhoods in Asheville, North Carolina, Charlotte, North Carolina; Mobile, Alabama, and Richmond, Virginia. While the authors find better transit service in low-car-ownership neighborhoods, they find worse bus service in minority neighborhoods when controlling for other factors, such as car ownership and income.

Daily experiences with transit may also reveal biases in the delivery of transit services. For example, Lubitow et al. (2017) use focus groups to examine transit-dependent riders' experiences in Portland, Oregon, and conclude that public transit investments are biased toward the experiences and the benefit of white, relatively wealthy, able-bodied, male commuters. These differences in experiences and services may also have important implications for poorer residents' overall life-satisfaction and quality of life. Comparing life-satisfaction with available transportation alternatives and residential location, Makarewicz and Németh (2018) find that only low-income transit-dependent residents of Denver have substantial differences in subjective wellbeing based on whether they live in the urban core or other parts of the region. The authors argue that access to transit service is particularly important for the overall wellbeing of poorer residents.

1.2.3 Evidence of choice

Finally, the existing literature provides substantial evidence that so-called transit-dependents exercise a great deal of choice. For example, in an analysis of the travel-behavior of choice and dependent riders, Polzin et al. (2000) find that transit dependents, defined by auto-availability, age, and license-status, use transit for just 16% of trips. As Giuliano (2005) observes, most poor households are car-dependent rather than transit-dependent and only use transit when service quality is high enough to meet daily travel needs. Policymakers should therefore focus high-capacity investments in high-density and high poverty areas, instead of suburban rail services that are unlikely to attract substantial numbers of new transit riders (Giuliano, 2005). Thomspon et al. (2012) and Brown et al. (2014) find that transit-dependents are highly responsive to service quality, price, travel time, and how well transit serves job centers outside of downtown locations in Broward County, Florida, and Atlanta, Georgia.

The overall observation that dependents exercise a substantial amount of choice and will only choose transit when it suits their needs is also consistent with research on income, car-availability, and other keywords associated with transit dependency. King et al. (2019) argue that the US built environment is so auto-oriented that, outside of older, denser urban centers, poor households need a private vehicle to participate in basic economic activities. In order to afford a car, people frequently drive without collision insurance (Clark & Wang, 2010) and even turn to crowd-funding to pay to replace a car lost due to unexpected circumstances (Klein et al., 2019). Many low-income residents without cars borrow them or carpool to get to work (Blumenberg & Smart, 2014; Lovejoy & Handy, 2011; Rogalsky, 2010).

1.3 Research approach

I use a discrete-choice random utility modeling framework to define and generate estimates of choice riders. Discrete choice models are commonly applied to estimate the probability that an individual chooses one available alternative, such as transit, over others, such as a car and other modes (Ben-Akiva & Lerman, 1985; Train, 2009). Estimating transit ridership has been particularly important to the early development of discrete choice models. In his Nobel lecture, McFadden (2001) details how the success of early applications to predict the ridership of a new rapid transit system in the San Francisco Bay Area was particularly important to the popularization of random utility models.

I define choice riders simply as those travelers who have close to a 50% probability of choosing transit based on estimates generated using a random utility model. For conceptual clarity and to emphasize the existing literature's focus on drawing transit riders out of cars, I discuss and estimate models of travelers choosing between transit and a private car. Discrete choice models make a clear and direct connection between the probability of choosing transit and the relative attractiveness of transit. When transit is substantially less attractive than a car, a traveler not only has a low probability of choosing transit but is generally unresponsive to changes in the attractiveness of either cars or transit. Similarly, when the attractiveness of transit is high, changes in the attractiveness of cars or transit will only have a small effect on the probability of choosing transit. Choice transit riders, by contrast, are highly responsive to changes in the attractiveness of cars or transit and have a much higher likelihood of adjusting their travel behavior as transit agencies improve or reduce service quality. While this design is conceptually clear for choice riders, it is likely less relevant for examining transit dependency. While a choice rider is making the choice between transit and driving, a transit dependent may be choosing between transit, walking, or not taking a trip at all.

1.3.1 Case context

Greater Philadelphia, which has a diverse range of transit users and transit services, is an ideal place to develop and fit an empirical model of choice ridership. SEPTA and New Jersey Transit provide bus, subway, commuter rail, and trolley services throughout the region. The centrally located cities of Philadelphia, PA, and Camden, NJ, have substantial numbers of low-income, minority residents who use the cities' urban bus systems. These residents are characteristic of the literature's general definitions of transit dependents. The region also has a large network of commuter rail lines, many of which extend into wealthy, low-density, majority-white towns and neighborhoods. The term the Main Line refers to the original operator of several of SEPTA's commuter rail lines and has become shorthand for Philadelphia's wealthy western suburbs.

1.3.2 Data summary

Table 1.1 presents the predictor variables used to estimate whether an individual chooses to use transit or a car on a weekday trip in the Philadelphia region. Predictor variables include socioeconomic information about the individual making a trip, characteristics about the trip, and environmental characteristics around the trip's origin and destination. The existing literature commonly includes these predictor variables, many of which also feature in descriptions of choice and dependent transit ridership. I pay special attention to including variables that feature in descriptions of choice ridership, such as income, race, car ownership, gender, urban location, and service quality.

Table 1.1 Socioeconomic, trip, and environmental characteristics of trips and trip makers

Variable	Share/Mean	Std Dev	Min	Max
Socioeconomic characteristics				
Age				
18 - 24	0.051			
25 - 44	0.216			
45 - 64	0.459			
64+	0.270			
Unreported	0.004			
Race/Ethnicity				
White/Caucasian	0.854			
Black/African American	0.072			
Other/Unreported	0.078			
Educational Attainment				
High school or lower	0.176			
Associate or some college	0.195			
Bachelor	0.306			
Graduate	0.317			
Unreported	0.005			
Gender				
Female	0.553			
Male and other	0.447			
Child(ren) under 5 in household	0.084			
Occupation				
Manufacturing/production/agriculture	0.046			
Non-office services	0.278			
Office/other/unreported/not employed	0.676			
Income below $10,000	0.016			
Household motor vehicles				
0	0.050			
1	0.295			
2	0.478			
3+	0.177			

Trip characteristics				
Trip tour				
Home-based work	0.420			
Home-based other	0.518			
Non-home-based	0.062			
Travel time (transit minus car)	34	31	-43	199
Travel cost (transit minus car)	-1.36	3.53	-31.96	24.45
Environmental characteristics				
Land use mix at origin	0.55	0.24	0.00	1.00
Kilometers to Philadelphia City Hall	24.8	14.9	0.3	71.4
Hourly parking price at destination	0.17	0.44	0.00	2.00
Average bus frequency at bus stops	3.3	6.0	0.0	69.6
High-capacity rail station presence				
None	0.552			
Origin and destination	0.149			
Origin or destination	0.300			
Number of observations	20895			

Socioeconomic and travel data are from the Delaware Valley Regional Planning Commission's (DVRPC) (2012) household travel survey. This survey provides data on 20,216 individuals and 81,940 trips undertaken by the members of 9,235 households in Philadelphia and the surrounding suburban counties of Pennsylvania, New Jersey, and Delaware between July 2012 and September 2013. The DVRPC Office of Modeling and Analysis also provided land use data, average parking prices, and travel times, costs, and trip distances by mode during four time periods (6AM-10AM, 10AM-3PM, 3PM-7PM, and 7PM-6AM) drawn from the 2010 TIM2.1 Travel Model, which was run in VISUM 12.5 and validated for a 2010 base year. Estimated tolls, fares, and parking charges are in 2010 dollars with an additional $0.575 operating cost assigned to each mile of car travel. The land use entropy index includes commercial, residential, and institutional land uses and varies from zero, when there is no land use mix, to one, when there is an equal share of all three land uses.

Bus service frequency is estimated for each transportation analysis zone using 2013 and 2015 SEPTA and NJ Transit's station-level GTFS data at the four time periods presented above. The presence of high-capacity transit is estimated by whether there is a subway, commuter rail, or trolley station within 800 meters of the centroid of a transportation analysis zone. Distances to City Hall are calculated by assigning the shortest road-network path.

The dataset and analysis exclude trips made by modes other than transit or car, trips made by individuals under 18, trips outside of the Philadelphia region, trips for which transit was not a viable alternative due to a lack of service, and trips within the same transportation analysis zone. Travel time and cost estimates by mode are not available for trips outside of the service area and within the same analysis zones. The final dataset includes 26,033 trips, of which 5,138 are selected randomly by household and set aside for model validation and testing.

1.3.3 Model specification

The reported model fits the data using a binomial logit model predicting the probability that an individual chooses transit or a private vehicle as a function of socioeconomic information about the trip-maker, characteristics of the trip, and land use and transportation characteristics near a trip's origin and destination. To account for unobserved correlations in the mode choice of individuals making multiple trips and members of the same household, I estimate and report cluster bootstrapped standard errors by household. The reported parameter estimates have not been transformed and correspond to an estimate of the shift in the systematic utility of transit associated with each predictor variable. Exponentiating individual parameter estimates will provide the odds-ratio for readers who prefer this measure.

To emphasize legibility and parsimony, I drop variables with low statistical significance from the model and group factor variables based on legibility and model fit prior to bootstrapping standard errors. For example, population density and job density are not included in the final models because they are not statistically significantly associated with mode choice when including data on price, travel time, vehicle ownership, and other trip characteristics. The ten income groups provided in the travel survey are grouped into a single category because only those earning less than $10,000 per year have a statistically significantly higher likelihood of choosing transit over a car when including other predictor variables. The relationships between mode choice and socioeconomic and environmental predictors, which are correlated across household members, generally weaken after bootstrapping clustered standard errors. Several are no longer statistically different from zero at the 95th or 90th percent confidence level in the final reported models. In terms of legibility, I group employment categories into three types. Including all 26 employment classifications improves model fit at the margin but takes up substantial space, reduces legibility, and creates some overfitting problems for several of the categories with limited observations.

1.4 Model results

The model predicting whether someone uses transit instead of a car for a given trip fits the data well with a pseudo R-squared of 0.54 and generates statistically significant parameter estimates that are consistent with the existing literature on transit ridership (Table 1.2). Travelers from the poorest households and households without cars are more likely to use transit than those from other households. Women are less likely to use transit when controlling for other predictor variables, as are those from households with young children. These differences may reflect differences in travel patterns and safety concerns by gender and presence of young children. There is only a small and statically insignificant difference in the probability that White or Black travelers choose transit when including other covariates. Other races and ethnicities, predominantly Asian and Latin American, are more likely to choose transit than either White or Black residents. The difference, however, is not statistically significant when bootstrapping standard errors. The probability of choosing transit decreases with age and educational attainment.

Table 1.2 Binary logit model predicting the probability of choosing transit instead of a car for trips within the Philadelphia region

	Coefficient Estimation	Bootstrapped standard errors
Socioeconomic characteristics		
Age		
18 - 24	Reference	
25 - 44	-1.260***	0.222
45 - 64	-1.548***	0.214
64+	-1.762***	0.239
Unreported	-1.134	1.856
Race/Ethnicity		
White/Caucasian	Reference	
Black/African American	0.055	0.175
Other/Unreported	0.238	0.177
Educational Attainment		
High school or lower	Reference	
Associate or some college	-0.291	0.177

Bachelor	-0.370**	0.168
Graduate	-0.372**	0.165
Unreported	0.708	0.939
Female	-0.304***	0.105
Child(ren) under 5 in household	-0.535**	0.213
Occupation		
Manufacturing/production/agriculture	Reference	
Non-office services	-1.061***	0.355
Office/other/unreported/not employed	-0.367***	0.115
Income below $10,000	-0.606	0.364
Household vehicles		
0	Reference	
1	-3.108***	0.196
2	-3.713***	0.201
3+	-4.099***	0.247
Trip characteristics		
Trip tour		
Home-based work	Reference	
Home-based other	0.942***	0.125
Non-home-based	0.771***	0.198
Travel time (transit minus car)	-0.054***	0.005
Travel cost (transit minus car)	-0.168***	0.014
Environmental characteristics		
Land use mix at origin	0.361*	0.204
Kilometers to Philadelphia City Hall	-0.009*	0.005
Hourly parking price at destination	0.647***	0.068
Average bus frequency at bus stops	0.020***	0.005
High-capacity rail station presence		
None	Reference	
Origin and destination	1.629***	0.173
Origin or destination	0.635***	0.159
Constant	2.159***	0.473
Observations		20,895
Log Likelihood		-2,831
McFadden Pseudo R2		0.541

*Notes. (1) Significance codes: *p<0.1; **p<0.05; ***p<0.01. (2) Bootstrapped standard errors are clustered by household.*

As expected, residents are substantially more likely to choose transit when the price and speed of transit improves relative to the car. Every additional minute saved by car reduces the utility of transit by 0.05. For an average commuter, increasing the speed of transit by 10 minutes relative to car increases the odds of choosing transit by 70%. Dividing the travel time by the travel cost parameter estimate suggests that the average person is willing to spend about $20 to save an hour of travel time. In addition to overall estimated time and cost, residents are much more likely to use transit on trips that are connected by rail stations and in areas and times of day with higher bus frequency. Higher parking meter prices, a greater mix of land uses, and closer proximity to downtown Philadelphia are all also statistically associated with a higher probability of choosing transit. Whether a given feature increases or decreases the probability of taking transit, however, does not provide information about whether that feature is associated with choice or dependent ridership.

1.4.1 Understanding choice riders

To better understand choice ridership, I apply the model from Table 1.2 to generate estimates of the probability of choosing transit and summarize the data by low, middle, and high probability of transit choice (Table 1.3). These categorizations best correspond to auto-dependents, choice riders, and transit dependents. I draw four main findings from these groupings.

Table 1.3 Share or mean of socioeconomic, trip, and environmental characteristics by probability of choosing transit

	Estimated probability of choosing transit			
Variable	0% - 100%	<1%	40% - 60%	>95%
Socioeconomic characteristics				
Age				
25 - 44	0.216	0.178	0.345	0.409
45 - 64	0.459	0.472	0.415	0.330
64+	0.270	0.318	0.165	0.056
Unreported	0.004	0.003	0.003	0.000
Race/Ethnicity				
Black/African American	0.072	0.030	0.174	0.242
Other	0.078	0.062	0.104	0.135
Educational Attainment				
Associate degree or some college	0.195	0.198	0.175	0.181
Bachelor	0.306	0.331	0.264	0.298
Graduate	0.317	0.310	0.380	0.256
Unreported	0.005	0.003	0.014	0.000
Female	0.553	0.574	0.491	0.484
Child(ren) under 5 in household	0.084	0.091	0.073	0.042
Occupation				
Manufacturing/production/agriculture	0.046	0.063	0.023	0.000
Non-office services	0.278	0.270	0.266	0.251
Income below $10,000	0.016	0.003	0.042	0.205

Household vehicles				
1	0.295	0.203	0.500	0.028
2	0.478	0.556	0.252	0.033
3+	0.177	0.239	0.049	0.000
No car	0.050	0.002	0.199	0.940
Trip characteristics				
Trip tour type				
Home-based work	0.420	0.332	0.674	0.544
Non-home-based	0.062	0.044	0.082	0.153
Travel time (transit minus car)	34.417	51.258	4.655	-2.760
Travel cost (transit minus car)	-1.362	-0.845	-2.552	-2.340
Environmental characteristics				
Land use mix at origin	0.550	0.479	0.737	0.795
Kilometers to Philadelphia City Hall from residence	24.786	31.072	9.586	5.691
Hourly parking price at destination	0.169	0.031	0.734	1.052
Average bus frequency by station	3.288	1.528	10.498	13.699
High-capacity rail station presence				
Origin and destination	0.149	0.003	0.689	0.930
Origin or destination	0.300	0.184	0.262	0.070
Number of observations	**20895**	**10827**	**576**	**215**
Share of data sample	**1**	**0.518**	**0.028**	**0.010**

First, many of the socioeconomic factors associated with the literature's definitions of transit dependency are much more common for choice riders than for the general population. For example, choice riders are substantially more likely to be non-white, low income, and carless than other residents of the Philadelphia region. For example, 17% of the sample of choice riders are Black compared to 7% of the total sample and 3% of auto-dependents. Although just 4% of choice riders earn less than $10,000 per year, that share is almost 3 times higher than the total sample and 12 times higher than the sample of auto-dependents. Choice riders are 1.6 and 4.0 times likelier to have one car or no car than the metropolitan sample average.

Second, there is substantial diversity within the group of choice transit riders. There are low-income bus users and high-income commuter rail users. There are old, young, male, female, Black, White, Asian, and Hispanic choice riders. Some have graduate degrees while others have not completed high school. Some have multiple cars. Others have none. Some work in downtown office jobs while others have retail jobs outside of Philadelphia. The strongest commonality across choice riders is that transit is generally competitive with the car in terms of cost, travel time, and convenience. Differences in the share of choice riders using transit for trips to and from work may also partially reflect transit service's general orientation toward serving job centers and peak travel hours.

This leads directly to the third main finding that transit service quality is critical to choice ridership. The travel time difference between transit and cars is just 4.5 minutes for choice riders compared to 51 minutes for auto-dependents and 34 minutes for the total sample. On average, an auto-dependent would need to have a value of time of less than a dollar per hour to choose transit. Choice transit riders have rail stations near the origins and destinations of their trips for two-thirds of all trips and have nearby buses arriving every six minutes on average. They are also more likely to be traveling in places with diverse land uses that are close to downtown Philadelphia. In terms of the predictive models, the measures of transit service quality are substantially stronger and more statistically significant predictors of transit use than socioeconomic predictors. Low-income and minority residents, like wealthier white residents, generally only choose transit when service quality is high. The early definitions of choice transit ridership are correct in that choice riders will choose their personal cars over transit when service quality is low. These definitions, however, miss that many choice riders do not have a car but will borrow one or get a ride from friends, family, and coworkers. Across the entire sample, carless travelers used cars for 38% of all trips.

Fourth and finally, transit dependents who use transit regardless of service quality because they are low income or do not own a car appear to be rare. Just 31 trips (0.015% of the sample) had a greater than 99% chance of being by transit. Due to the small number of absolute trips in this category, I expand the column to include trips with over 95% probability of being taken by transit. While these were more likely to be taken by low-income and minority residents without cars, these trips also had the highest quality transit options in the dataset. Wealthy white residents with multiple cars also have a high likelihood of taking transit when buses arrive every 5 minutes, high speed rail connects both trip ends, parking is expensive, and transit is both faster and less expensive than driving. When buses arrive infrequently and there is no high-capacity rail, most of the sample moves by car, regardless of race, income, gender, education, or the number of household vehicles. As a result, the probability of residents choosing transit is substantially higher when they live in the urban core, where service quality is high for a high share daily travel (Figure 1.1). Of the more than five thousand trips with below average transit service, just 36 were made by transit. Of note, white residents with one or more cars took most of these trips.

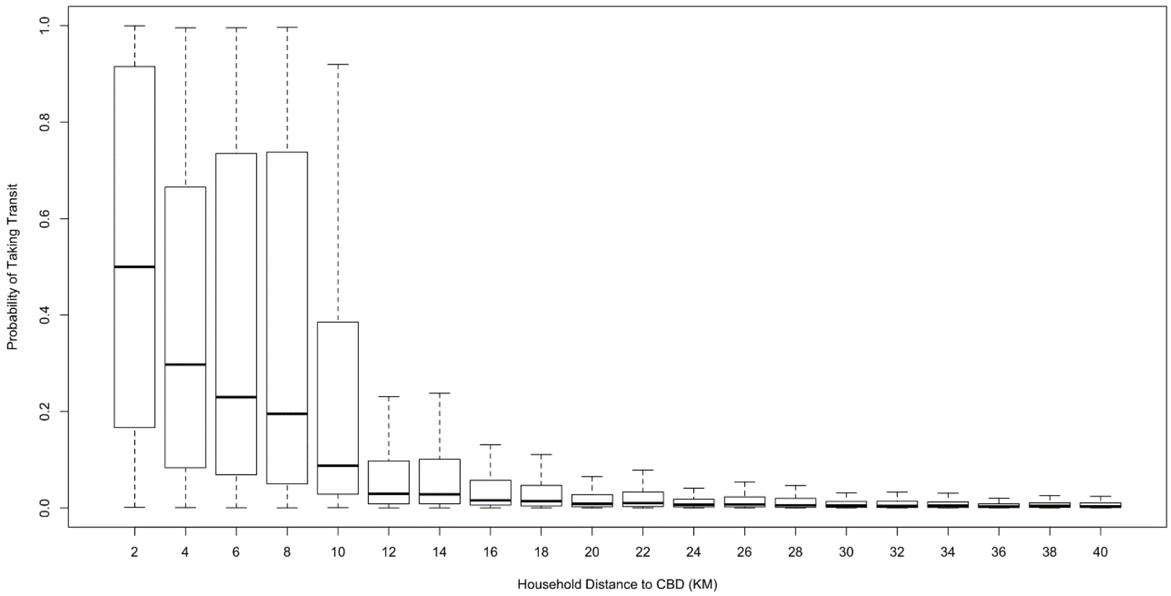

Figure 1.1 Estimated probability of choosing transit by how far a traveler lives from Philadelphia City Hall

Two caveats apply to this broad observation about the rarity of transit dependents. First, just because low-income individuals without cars are not systematically transit-dependent does not mean that there are not individuals who depend partially or entirely on transit to meet their daily and weekly needs. That no one can depend on transit where no service exists does not contradict the fact that low-income carless travelers are more likely to use transit despite worse service than others. For the two-thirds of transit trips taken by carless low-income survey respondents, a private car would have been 12 minutes faster than transit on average. While this is a much smaller differential than the sample average of 34 minutes, it is also much larger than the 4.7-minute average for choice riders. Second, the research design is focused on identifying choice riders rather than transit dependents. Those who most depend on transit despite poor transit service are probably more likely to be choosing between taking transit and deferring a trip than between taking transit and taking a car. Living without a car in an area with poor transit service quality almost certainly reduces travel and constrains access to employment, school, shopping, recreation, and other important destinations.

1.4.2 A note on residential self-selection and vehicle ownership

Due to issues of residential self-selection (Xinyu Cao et al., 2009; Handy et al., 2005), it is difficult to say by how much increasing or reducing transit services into specific neighborhoods would affect transit ridership. For example, residents in auto-dependent neighborhoods may be particularly disinterested in transit and unlikely to choose transit even if service levels improve. These unobserved preferences for and against travel modes might influence the size and significance of the parameter estimates presented in Table 2. Although the model does not include controls for preferences beyond a robust set of socioeconomic predictors, accounting for preferences would likely strengthen the overall finding that: (1) features commonly associated with dependent riders are more prevalent among choice riders than the general population; (2) transit service quality is critical for choice ridership; and (3) people who choose transit regardless of service quality appear to be exceedingly rare. Lower-income households without cars are least able to make travel and housing decisions to match their personal preferences.

Vehicle ownership decisions are also highly related to mode choice decisions. People who do not like to drive, for example, are unlikely to purchase a car. Including vehicle ownership likely attenuates the strength of income and other predictor variables that are associated with both vehicle ownership and mode choice. I include vehicle ownership directly in the model for two primary reasons. First, vehicle ownership is the most common defining characteristic of choice ridership in the literature and thus an essential predictor variable. Second, the research design is focused on predicting transit ridership rather than assessing the causal determinants of transit ridership. Vehicle ownership is a strong predictor of mode choice even after including variables on trip characteristics and travelers. A model focused on causal relationships would require a different modeling approach.

1.5 Conclusion

In this paper, I develop a theoretical model of choice ridership and apply it to data from a travel survey in the Philadelphia region using a random utility model. The Philadelphia region, which has substantial bus services in low-income urban neighborhoods and high-capacity commuter rails services into wealthy suburban neighborhoods, is an ideal place to study dependent and choice ridership. The reported model produces relatively strong predictions of whether individuals choose transit or a private vehicle for trips outside of their neighborhood that start and end within the region. The model also produces individual parameter estimates that are generally consistent with the existing literature on mode choice and willingness to pay for travel time savings. Low-income urban residents with no car and who are commuting to work in areas with high-quality transit service and costly parking prices are particularly likely to travel by transit.

Based on these models of transit ridership, I analyze features that are most common across choice riders. Many of the attributes assumed to be associated with dependent riders, such as a lack of a car, low income, and being a racial or ethnic minority, are much more prevalent among choice riders than the general metropolitan population. Moreover, transit dependency is rare. Residents generally only choose transit when service is high quality and transit is cost- and time-competitive with the car for a given trip. Those without cars frequently borrow a car or get a ride with family, friends, or colleagues.

Based on these findings, researchers and policymakers should avoid undertheorized and under analyzed descriptions of choice and dependent riders. The prevailing descriptions of choice and dependent riders are inaccurate and may divert investments and service improvements away from riders who are most likely to choose transit for more trips as service improves. The analysis of choice riders, moreover, suggests that there may be opportunities to differentiate across types of choice riders. For example, some may be particularly sensitive to travel time and service frequency, while others may be more sensitive to the price of parking or the ease of access to commuter and heavy rail stations. Latent-class choice modeling may offer opportunities to better understand whether there are important and systematic differences in choice riders, by location, income, and other features.

The findings also suggest that the debate between using transit investments to reduce automobile use or increase accessibility for low-income transit users is largely misplaced. Improving urban bus service into low-income neighborhoods almost certainly attracts people out of cars. Future analysis could help shed light the relative costs and benefits of attracting specific types of trips from specific locations. Suburban commuter rail trips, for example, may be more expensive to attract, but they are also more likely to replace longer-distance car trips. Urban bus trips may replace shorter car trips, but these trips may occur on relatively congested local urban streets. These shorter car trips may also be relatively harmful if they are likelier to occur in older vehicles that produce more local pollution and crash with higher frequency and severity. In any case, focusing transit policy on expanding its existing customer base, who tend to live in the urban core where service quality is already relatively high, is probably the most cost-effective way to draw new riders.

Finally, even in a transit-friendly region like Philadelphia, most residents live and travel in areas where transit is simply not a reasonable option. Across our sample, taking a trip by car saves an average of 34 minutes per trip. For the 51% of the sample with a lower than 1% chance of taking transit, the difference is nearly an hour per trip. Attracting auto-dependents to transit through either transit investments or land use policy therefore is likely to be prohibitively expensive and time-consuming. Unless new technologies or business models can make transit competitive with cars on these types of trips, focusing investments into urban areas with high concentrations of high-probability transit riders will not only improve service for existing users, but likely do the most to draw riders out cars.

1.6 Acknowledgments

Funding support was provided by the Mobility 21 National University Transportation Center (USDOT Award No. 69A3551747111). Matthew Gates (Delaware Regional Planning Commission) provided data and support related to the regional travel model outputs and inputs. Han Qiao and Tzu (Andrew) Wang from the University of Pennsylvania helped process the data, estimate models, and generate figures. Grant Engel (SEPTA), Chris Puchalsky (City of Philadelphia), and Andrew Simpson (City of Philadelphia) provided data support and feedback on takeaways for public policymakers.

1.7 References

Bates, J. (1981). *Case Study VI: Marketing in Marta* (DOT-I-81-1 Final Rpt.; Mass Transit Management: Case Studies of the Metropolitan Atlanta Rapid Transit Authority). University of Georgia. https://rosap.ntl.bts.gov/view/dot/527

Beimborn, E. A., Greenwald, M. J., & Jin, X. (2003). Accessibility, Connectivity, and Captivity: Impacts on Transit Choice. *Transportation Research Record*, *1835*(1), 1–9. https://doi.org/10.3141/1835-01

Ben-Akiva, M., & Lerman, S. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand* (1st ed.). The MIT Press.

Blumenberg, E., & Smart, M. (2014). Brother can you spare a ride? Carpooling in immigrant neighbourhoods. *Urban Studies*, *51*(9), 1871–1890.

Brown, J., Thompson, G., Bhattacharya, T., & Jaroszynski, M. (2014). Understanding Transit Ridership Demand for the Multidestination, Multimodal Transit Network in Atlanta, Georgia: Lessons for Increasing Rail Transit Choice Ridership while Maintaining Transit Dependent Bus Ridership. *Urban Studies*, *51*(5), 938–958. https://doi.org/10.1177/0042098013493021

Cai, M., Jiao, J., Luo, M., & Liu, Y. (2020). Identifying transit deserts for low-income commuters in Wuhan Metropolitan Area, China. *Transportation Research Part D: Transport and Environment*, *82*, 102292. https://doi.org/10.1016/j.trd.2020.102292

Cao, Xiaoshu, Chen, H., Liang, F., & Wang, W. (2018). Measurement and Spatial Differentiation Characteristics of Transit Equity: A Case Study of Guangzhou, China. *Sustainability*, *10*(4), 1069. https://doi.org/10.3390/su10041069

Cao, Xinyu, Mokhtarian, P. L., & Handy, S. (2009). Examining the Impacts of Residential Self-Selection on Travel Behaviour: A Focus on Empirical Findings. *Transport Reviews*, *29*(3), 359–395. https://doi.org/10.1080/01441640802539195

Cervero, R. (1981). Flat versus differentiated transit pricing: What's a fair fare? *Transportation*, *10*(3), 211–232. https://doi.org/10.1007/BF00148459

Cheranchery, M. F., & Maitra, B. (2018). Investigating perception of captive and choice riders for formulating service standards of ordinary and premium buses in Indian cities. *Transport Policy*, *72*, 89–96. https://doi.org/10.1016/j.tranpol.2018.10.002

Chia, J., Lee, J. (Brian), & Kamruzzaman, M. (2016). Walking to public transit: Exploring variations by socioeconomic status. *International Journal of Sustainable Transportation*, *10*(9), 805–814. https://doi.org/10.1080/15568318.2016.1156792

Clark, W. A., & Wang, W. W. (2010). The automobile, immigrants, and poverty: Implications for immigrant earnings and job access. *Urban Geography*, *31*(4), 523–540.

Delaware Valley Regional Planning Commission. (2012). *2012-2013 Household Travel Survey for the Delaware Valley Region*. https://www.dvrpc.org/Products/14033/

Elkind, E. (2014). *Railtown: The Fight for the Los Angeles Metro Rail and the Future of the City*. University of California Press. https://www.ucpress.edu/book/9780520278271/railtown

Federal Transit Administration. (n.d.). *About FTA*. Retrieved January 29, 2021, from https://www.transit.dot.gov/about-fta

Giuliano, G. (2005). Low Income, Public Transit, and Mobility. *Transportation Research Record: Journal of the Transportation Research Board*, *1927*(1), 63–70. https://doi.org/10.3141/1927-08

Google. (2021). *Google Books Ngram Viewer*. https://books.google.com/ngrams#

Grengs, J. (2001). Does Public Transit Counteract the Segregation of Carless Households? Measuring Spatial Patterns of Accessibility. *Transportation Research Record*, *1753*(1), 3–10. https://doi.org/10.3141/1753-01

Grengs, J. (2002). Community-Based Planning as a Source of Political Change: The Transit Equity Movement of Los Angeles' Bus Riders Union. *Journal of the American Planning Association*, *68*(2), 165–178. https://doi.org/10.1080/01944360208976263

Haines, R. L. (1978). The Politics of Economics in Transit Planning. *Urban Affairs Quarterly*, *14*(1), 55–77. https://doi.org/10.1177/107808747801400103

Handy, S., Cao, X., & Mokhtarian, P. L. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D: Transport and Environment*, *10*(6), 427–444. https://doi.org/10.1016/j.trd.2005.05.002

Hepp, J. (2018). SEPTA. In *Encyclopedia of Greater Philadelphia*. Mid-Atlantic Regional Center for the Humanities. https://philadelphiaencyclopedia.org/archive/septa/

Jacques, C., Manaugh, K., & El-Geneidy, A. M. (2013). Rescuing the captive [mode] user: An alternative approach to transport market segmentation. *Transportation*, *40*(3), 625–645. https://doi.org/10.1007/s11116-012-9437-2

Jiao, J. (2017). Identifying transit deserts in major Texas cities where the supplies missed the demands. *Journal of Transport and Land Use*, *10*(1), 529–540.

Jiao, J., & Dillivan, M. (2013). Transit Deserts: The Gap between Demand and Supply. *Journal of Public Transportation*, *16*(3). https://doi.org/10.5038/2375-0901.16.3.2

Jiao, J., & Wang, F. (2020). Shared mobility and transit-dependent population: A new equity opportunity or issue? *International Journal of Sustainable Transportation*, *0*(0), 1–12. https://doi.org/10.1080/15568318.2020.1747578

Keefer, L., Witheford, D., Thomas, S. J., Schwartz, A., Kolesar, E. R., Cominsky, R., Sullivan, S., Irwin, R., Luster, G., Blake, G. W., Fox, W., Latta, J., Dorgan, J., Lunardini, R., Jorgenson, D., Hamburg, J., Jones, G., Schneider, M., & Carroll, J. D. (1963). *Pittsburgh Area Transportation Study: Final Report in Two Parts. Volume 2 Forecasts and Plans*. Commonwealth of Pennsylvania, County of Allegheny, City of Pittsburgh, and U.S. Department of Commerce: Bureau of Public Roads.

King, D. A., Smart, M. J., & Manville, M. (2019). The Poverty of the Carless: Toward Universal Auto Access. *Journal of Planning Education and Research*, 0739456X18823252. https://doi.org/10.1177/0739456X18823252

Klein, N. J., Tran, M., & Riley, S. (2019). *"Desperately in Need of Car": Analyzing Crowdfunding Campaigns for Car Purchases and Repairs on Gofundme.com* [Preprint]. SocArXiv. https://doi.org/10.31235/osf.io/8x7d2

Krizek, K., & El-Geneidy, A. (2007). Segmenting Preferences and Habits of Transit Users and Non-Users. *Journal of Public Transportation*, *10*(3). https://doi.org/10.5038/2375-0901.10.3.5

Lachapelle, U., Frank, L. D., Sallis, J. F., Saelens, B. E., & Conway, T. L. (2016). Active Transportation by Transit-Dependent and Choice Riders and Potential Displacement of Leisure Physical Activity. *Journal of Planning Education and Research*, *36*(2), 225–238. https://doi.org/10.1177/0739456X15616253

Lovejoy, K., & Handy, S. (2011). Social networks as a source of private-vehicle transportation: The practice of getting rides and borrowing vehicles among Mexican immigrants in California. *Transportation Research Part A: Policy and Practice*, *45*(4), 248–257.

Lubitow, A., Rainer, J., & Bassett, S. (2017). Exclusion and vulnerability on public transit: Experiences of transit dependent riders in Portland, Oregon. *Mobilities*, *12*(6), 924–937. https://doi.org/10.1080/17450101.2016.1253816

Makarewicz, C., & Németh, J. (2018). Are multimodal travelers more satisfied with their lives? A study of accessibility and wellbeing in the Denver, Colorado metropolitan area. *Cities*, *74*, 179–187. https://doi.org/10.1016/j.cities.2017.12.001

Márquez, L., Pico, R., & Cantillo, V. (2018). Understanding captive user behavior in the competition between BRT and motorcycle taxis. *Transport Policy*, *61*, 1–9. https://doi.org/10.1016/j.tranpol.2017.10.003

McFadden, D. (2001). Economic Choices. *The American Economic Review*, *91*(3), 351–378.

Polzin, S., Chu, X., & Rey, J. (2000). Density and Captivity in Public Transit Success: Observations from the 1995 Nationwide Personal Transportation Study. *Transportation Research Record: Journal of the Transportation Research Board*, *1735*(1), 10–18. https://doi.org/10.3141/1735-02

Rogalsky, J. (2010). Bartering for basics: Using ethnography and travel diaries to understand transportation constraints and social networks among working-poor women. *Urban Geography*, *31*(8), 1018–1038.

Sohn, K., & Yun, J. (2009). Separation of car-dependent commuters from normal-choice riders in mode-choice analysis. *Transportation*, *36*(4), 423–436. https://doi.org/10.1007/s11116-009-9209-9

Sun, T., & Fan, Y. (2018). Inequitable job accessibility across educational and hukou groups in Beijing: An analysis of transit-based accessibility to sectoral jobs. *Journal of Transport and Land Use*, *11*(1), 791–803.

Taylor, B. D., & Morris, E. A. (2015). Public transportation objectives and rider demographics: Are transit's priorities poor public policy? *Transportation*, *42*(2), 347–367. https://doi.org/10.1007/s11116-014-9547-0

Thompson, G., Brown, J., & Bhattacharya, T. (2012). What Really Matters for Increasing Transit Ridership: Understanding the Determinants of Transit Ridership Demand in Broward County, Florida. *Urban Studies*, *49*(15), 3327–3345. https://doi.org/10.1177/0042098012443864

Train, K. (2009). *Discrete Choice Methods with Simulation* (2nd ed.). Cambridge University Press. http://elsa.berkeley.edu/books/choice2.html

U.S. Department of Transportation. (2017). *National Household Travel Survey*. http://nhts.ornl.gov/

van Lierop, D., & El-Geneidy, A. (2016). Enjoying loyalty: The relationship between service quality, customer satisfaction, and behavioral intentions in public transit. *Research in Transportation Economics*, *59*, 50–59. https://doi.org/10.1016/j.retrec.2016.04.001

Wells, K., & Thill, J.-C. (2012). Do Transit-Dependent Neighborhoods Receive Inferior Bus Access? A Neighborhood Analysis in Four U.S. Cities. *Journal of Urban Affairs*, *34*(1), 43–63. https://doi.org/10.1111/j.1467-9906.2011.00575.x

Zhao, J., Webb, V., & Shah, P. (2014). Customer Loyalty Differences between Captive and Choice Transit Riders. *Transportation Research Record*, *2415*(1), 80–88. https://doi.org/10.3141/2415-09

PART II. How Covid-19 Impacted Bus Ridership in Philadelphia: An Analysis of the Initial Shock to Bus Ridership

2.1 Introduction

In December 2019, the COVID-19 pandemic outbreak hit the world and has hugely impacted the way we live since then. As the World Health Organization declared a pandemic in March 2020, countries around the globe imposed air travel restrictions, travel bans and stay-at-home orders. The health and social impact associated with these lockdown measures led to unemployment, business closure and tremendous shock in people's everyday routine including working, traveling, buying groceries and education. Public transportation is among the most impacted activities, given the health concerns of contacting a high number of people in a small and restricted space. At the same time, transit helped thousands of frontline workers get to critical jobs in the healthcare system.

Public transportation modes including interregional air flights and intraregional mass urban transit all experienced drastic ridership decline. In May 2020, half of all the industry's planes were parked in airports and desert airstrips (Chokshi, 2020). US passenger airlines lost more than 46 billion dollars or more than 62.5 percent of the total revenue in 2020 compared to 2019 (Airlines for America, 2021). Mass urban transit, including buses, subways and trains all experienced tremendous decrease in ridership, which was caused by reductions in both travel demand and services supply. Transit ridership in North American cities dropped by 90 percent by the end of March 2020 (DeWeese, 2020). Transit agencies in Philadelphia, Maryland and Los Angeles reported huge revenue loss projections, from $300 million to $1.8 billion, through 2021 (Garza, 2020).

The aggregated loss in transit ridership cannot reveal the differences in the impact of the pandemic across different socioeconomic groups. In fact, not every occupation has the privilege of being able to work from home and not every person can completely avoid taking public transit during the pandemic. Essential workers in healthcare, food, public works and transportation sectors, are among the first groups that return to work and these are also the people who rely on public transit, who accounted for 36 percent of all transit commuters in the U.S. (TransitCenter, 2020). Therefore, it is crucial to reveal the stories behind aggregated transit ridership numbers and to understand the socioeconomic implications of transit ridership change before and after the pandemic outbreak.

In this study, we try to understand how different types of places lost ridership during the immediate phase after the pandemic hit by unpacking how the impact of demographics, socio-economics and land use factors on bus ridership changed before and after the Covid-19 outbreak. Through identifying the key characteristics of places that lost ridership most quickly, we can help show where service is most vital as well as reveal information about which types of riders are least able to stay at home or substitute other modes for transit. The reminder of the paper is organized as follows: Section 2 Literature review provides a summary of how past works looked into Covid-19's impact on trip making and especially transit ridership; Section 3 Methodology goes into the details about the data used and the reasoning behind using a mixed effect linear regression model for the study; Section 4 Result interprets the model estimations and their meanings; Section 5 Conclusions and Discussions presents how the model results from this study can be incorporated into social and policy insights for transit planning in the future.

2.2 Literature Review

2.2.1 Factors of Transit Ridership Generation

The rich collection of past literature on understanding factors influencing transit ridership informed the variable selection and modeling of this study. Generally, factors can be categorized into external factors and internal factors (Taylor & Fink, 2003). External factors are influencing factors outside of the transit system and internal factors are factors related to the provision of service.

A wide range of external factors have been examined, which include socio-economic, demographic, spatial, meteorological as well as financial factors (Boisjoly et al., 2018; Taylor & Fink, 2003; Tao et al., 2018). Demographic and socio-economic factors such as population density, job density, income level, car ownership are widely studied and found to be significantly impacting transit ridership across North America cities (Boisjoly et al., 2018; Boisjoly et al., 2018; Gómez-Ibáñez, 1996; Taylor et al., 2009). Similar correlations were found throughout these studies -- larger population, greater population density, lower income and lower vehicle ownership tend to positively correlate with higher ridership. Spatial factors that have been studied involve land use and urban form, such as types of land use, transportation infrastructures, parking availability and parking price (Chakour & Eluru, 2016; Pasha et al., 2016). Another area of focus is meteorological, or weather, factors. Among a wide range of weather types, those that have been studied and found to be significant include temperature, rainfall, snow, wind speed and fog etc (Guo et al., 2007; Böcker et al., 2013; Koetse & Rietveld, 2009). Some research also indicates that financing policy, such as transit subsidy, also significantly influences transit ridership (Gomez-Ibanez, 1996; Taylor & Fink, 2003).

Internal factors can be categorized into pricing, service quantity as well as service quality factors. Kain and Liu (1995) found that fares, as part of a combination of internal factors and external factors including employment, gas prices, and service quantity, contribute significantly to transit ridership. Other research also found the same negative correlation between fares and transit ridership and special pricing scheme has been noted also as an influencing factor (Chen et al., 2011; Kain & Liu, 1999; McLeod et al., 1991; Taylor et al., 2009). Transit quantity factors include service coverage as well as service frequency. Some metrics being used in modeling include fleet size, number of vehicles operated in maximum service, vehicle revenue hours and vehicle revenue miles etc (Gómez-Ibáñez, 1996; Guerra & Cervero, 2011; McLeod et al, 1991; Kain and Liu, 1999; Taylor et al., 2009). Service quality factors include customer service satisfaction, station safety, reliability, dependability and information availability were investigated by previous researchers (Bates et al., 2001; Currie and Wallis, 2008; Figler et al., 2011). There is a general consensus that higher service quantity and service quality lead to higher transit ridership.

It is noteworthy that the question of what explains transit ridership is complex (Taylor & Fink, 2013). As the different factors intertwine and correlate with each other, it is hard to conclude the relative importance of each of the factors and the relationship between these factors. Therefore, in this study, our model was built with the consideration of a broad range of potential factors in available data and then simplified based on collinearity and parsimoniousness given the specific context of Philadelphia and model testing.

2.2.2 Covid-19 and Trip Making

The pandemic has largely impacted how we navigate our everyday lives. Researchers have also been studying Covid-19's influence on trip making and travel behavior since the pandemic outbreak. The literature mainly consists of two methods of investigating how the pandemic shifted the way we travel -- empirical mobility data from mobile devices and surveys on human perception of travel and the pandemic.

Through using data collected mainly through mobile devices, researchers were able to find patterns of how people changed their travel behaviors during the pandemic outbreak. In an empirical analysis of a comprehensive mobility dataset of Andorra, researchers found that all metrics of mobility including number of trips and people making trips, dropped sharply at the start of the country's lockdown and gradually rose again as the restrictions were gradually lifted (Doorely, 2022). Another study on mobility patterns in NYC found that distance traveled everyday dropped by 70% and number of social contact in places decreased by 93% when comparing weekends in late February and March (Bakker, 2020). The national emergency declaration on March 14th 2020 and school closures in NYC resulted in a surge of trips to groceries, shopping, food and outdoor places as well as a surge in activities in the beaches and the Hamptons (Bakker, 2020). Research in Australia broke down trip making by modes of transportation and found that while trips by all modes decreased after the pandemic outbreak, private vehicle trips recovered faster than public transit trips and among public transportations active modes recovered the fastest (Beck, 2021). In terms of trip purposes, work business trips by public transport have returned to 60% of before COVID-19 levels but for most other trip purposes the recovery is slower (Beck, 2021).

Another group of researchers adopted surveys to investigate reasonings and human perceptions on their trip making decisions. Aaditya and Rahul using surveys conducted across India found that people were more willing to reduce essential and recreational trips compared to work trips (2021). People’s awareness regarding the disease was also a significant factor influencing people’s decision on trip making (Aaditya & Rahul, 2021). Another research project using surveys to understand the relationship between risk perceptions and trip making decisions filtered out many factors including experience with influenza, gender, perceptions on destination that are significant in influencing travel behavior during the pandemic (Hotle, 2020). Results from another worldwide survey conducted both before and during Covid-19 explained that trip purpose, mode choice, distance traveled, and frequency of trips were significantly different before and during the pandemic (Abdullah, 2020). The majority of trip purposes was shopping during the pandemic and mode choice became more influenced by pandemic related concerns (Abdullah, 2020).

2.2.3 COVID-19 and Transit Ridership

Public transit is among the hardest hit industries by the pandemic. Since the outbreak of the COVID-19 pandemic, researchers, news media and governmental institutions all have conducted initial studies on the changes of transit ridership and the corresponding socioeconomic implications. To summarize, the past studies mainly adopted two approaches -- trend analysis on transit ridership data and correlation analysis between socioeconomic factors and ridership change during the pandemic. All past studies agree upon a tremendous decline in public transit ridership and some studies revealed more detailed changes across different times and types of transits. Through correlation analysis, researchers have also found the disparities in transit ridership change among different socioeconomic groups.

Studies from MTA, MBTA and WMATA all reported over 70% ridership decline in the first few months of the COVID-19 pandemic (MTA,2020; MBTA, 2020; WMATA, 2020). A more recent study also looked into the detailed trends in bus and subway ridership changes and found shift in peak periods, increasing average trip distance for subway and local bus routes and decrease in express bus usage (Halvorsen et al, 2021).

Researchers have also started looking into how different socioeconomic groups respond differently in taking transit during the pandemic. Studies found that people with lower income, people of color and essential workers remain more reliant on public transit than their counterparts who can work from home or use private vehicles (Halvorsen et al, 2021; Hu & Chen, 2021; Parker et al., 2021; Qi et al., 2021). Through correlation analysis, researchers found that ridership decline in New York subway from February to April 2020 was negatively associated with the percentage of black population, foreign born population, population under poverty rate and essential workers, and subway ridership decline is positively associated with median household income and being in the CBD area (Halvorsen et al, 2021). News media also used graphics and charts to report how essential workers, low-paid workers and people of color are relying on public transit during the pandemic (Gothehrer-Cohen, 2020; Bliss et al, 2020). Similarly, TransitCenter looked at MBTA's automated passenger counter bus data and found that bus stops in neighborhoods with higher proportion of people of color and low incomes remained high ridership throughout the pandemic, whereas acute declines in ridership are found in business, cultural and university districts (TransitCenter, 2020). A recent study looking into transit demand across 113 county-level transit systems in 63 metro areas and 28 states across the

US again revealed that cities with more essential workers and a more vulnerable population tend to maintain higher transit demand levels during COVID-19 (Liu et al, 2020). Beyond transit for work, Long et al. (2023) also found that male, younger, non-White passengers are more likely to return to public transportation for non-work trips as soon as Covid restrictions are lifted, which revealed inequalities in public transport demand.

Despite the rich research and reports produced by transit agencies and researchers since the pandemic outbreak, there are still knowledge gaps that need to be addressed. First, to our knowledge, previous studies have only focused on understanding the correlations of individual factors with transit ridership during the pandemic or using graphics to show such correlations. We propose using a multilevel mixed-effect regression model to account for the compounding and controlled effects of various meteorological, demographic and land use variables to see how their effects changed before and after the outbreak of the pandemic. Secondly, due to constraints in data availability, most research studies have focused on a few cities where transit data are more easily accessible. We hope to expand the past literature with a case context of Philadelphia, a city with many characteristics that makes it important to study during the pandemic outbreak -- Philadelphia has a high concentration of hospitals and medical workers; it is the poorest large city in the US as well as a city with a high number of population relying on the extensive public transit network. To fill these gaps, this study uses data from APC (automatic passenger counters) on buses in Philadelphia before and after the outbreak of COVID-19 in March 2020 to identify and quantify the various factors associated with bus ridership decline.

2.3 Methodology

2.3.1 Study Context

The Philadelphia County is chosen as the study context due to its vast transit network and the diverse population relying on transit. The city of Philadelphia has a total population of 1.6 million, which consists of 39.3% black or African American population, 36.3% white population and 24.4% other racial groups (United States Decennial Census, 2020). The city is also dubbed as the "poorest large city" in the U.S., with a poverty rate around 23.3% (American Community Survey, 2019). The unemployment rate of Philadelphia reached 9.6% as of August 2021, which is much higher than the 5.2% unemployment rate of the U.S. as a whole (U.S. Bureau of Labor Statistics, 2021).

Public transit is a popular mode of travel for residents in Philadelphia. Around 25.5% of the population commute to work using public transportation compared to only 5% of the U.S. as a whole (American Community Survey, 2019). The city's accessible and bike-friendly transit system is operated by Southeastern Pennsylvania Transportation Authority (SEPTA). SEPTA offers a variety of transit services, including buses, trolleys, subways and Regional Rail. In the 2019 Fiscal Year, SEPTA operated 152 routes of public transits with the total unlinked trips adding up to 292 million that covers the entire Philadelphia County and the adjacent counties (SEPTA, 2019).

2.3.2 Ridership Data Processing

In this study, we obtained ridership data of SEPTA buses from February 23rd 2020 (Sunday) to April 1st 2020 (Wednesday). This period encompassed bus service data before the COVID outbreak, when the service was provided on a regular schedule, as well as the immediate period following the COVID pandemic, when SPETA reduced its bus services on March 22nd, 2020 (Streva, 2020).

Bus ridership estimation was developed from automated passenger count (APC) data and transit schedule based on General Transit Feed Specification (GTFS). SEPTA only has APC data available on a sample of buses, providing information on onboarding and offboarding passenger numbers for each stop. SEPTA partners with two vendors, Infodev and UTA, to collect ridership information. The sampled buses for each vendor were mutually exclusive, so we combined data from both vendors in order to obtain a more accurate estimation. Figure 2.1 shows the total number of sampled bus and scheduled bus by time of day throughout the study period. GTFS provides bus arrival times for each stop, which was used to calculate the scheduled bus frequencies for all stops on each day. GTFS data was obtained from SEPTA github website. Twelve different GTFS schedules were used to cover the entire study period. During the period where SEPTA ran headway based bus service, GTFS did not include arrival times for each stop. Instead, it provided the trip start time and the time it takes for a bus to travel between two stops. This information was used to calculate the number of scheduled buses at each stop during the period.

Figure 2.1 Total number of sampled stops by time of day (left) and total number of scheduled stops by time of day (right). We can see that the sample follows the pattern of the schedule.

Through linking APC data with GTFS by stop identification number and date, we were able to extrapolate the sampled APC bus ridership to an estimated total ridership for each stop on each day. We first calculated a sample rate for each stop on each day by **Equation 1**:

$$APC\ sample\ rate\ = \frac{Number\ of\ APC\ sampled\ \ buses}{Number\ of\ GTFS\ scheduled\ buses}, \qquad (1)$$

By summing up the total onboarding passenger count from APC data, we obtained the total APC sampled ridership for each stop on each day. To estimate the total ridership of the stop, we use **Equation 2**:

$$Total\ ridership\ estimation\ = \frac{Total\ APC\ sampled\ ridership}{APC\ sample\ rate}, \qquad (2)$$

In this way, the APC sampled data was expanded to an estimation of total ridership for each stop on each day during the study period.

2.3.3 Mixed-effect Multilevel Linear Regression Model

In this study, we adopted a multilevel mixed-effect regression model. Using a multilevel model helps examine ridership variation on both spatial and temporal scale, allowing us to incorporate the impacts from individual stops as well as individual days on ridership. While both group effects and group level predictors will contribute to the coefficient, with random intercept, we are allowing flexibility on having different intercepts for each stop each day (University of Bristol, Center for Multilevel modeling).

Both temporal and cross-sectional variables are used in this analysis. Temporal variables are a continuous record of daily precipitation in decimeters and an indicator variable differentiating between weekend and weekdays. Cross-sectional variables allow us to include information of the location of the stops in our model. Data were gathered and calculated for the census tracts where the stops are located. While information on riders of each stop and for smaller geographic units are not available, using census tract as the unit of analysis is a simple yet helpful approximation. Cross-sectional variables used in our model include an indicator for transfer stops, median household income, number of jobs accessible within 45 minutes of transit, population density, daily parking cost, and vehicle ownership rate. These variables were selected for their theoretical importance and statistical significance. More specifically, the variable *jobs accessible within 45 minutes* was first calculated at the census block level, using the travel time to all other census blocks within 60 kilometers for each departure time at 1-minute intervals between 7 and 9 a.m. based on the U.S. Census Longitudinal Employer-Household Dynamics (LEHD) and Origin-Destination Employment Statistics (LODES). The total number of accessible jobs was calculated for each block and departure time using a 45-minute threshold. Then the average number of accessible jobs was calculated for each block between 7 and 9 a.m. Finally, we calculated the average number of accessible jobs for each census tract by weighing the number of workers in each block.

Our study has a sample size of 398,546, which includes 7,877 stops in 63 days and not every stop has data on every day. Although we had options of including 13,697 stops in the larger SEPTA service area as part of our analysis, we opted to focus on stops that are in the City of Philadelphia's boundary for more accurate predictions. Places outside of the city are more suburban and have different travel behaviors. We tested them to conclude that those stops serve more as noises than meaningful data since they did not follow the same regression distribution pattern as the city. During the data cleaning process, we also eliminated duplicated stations, trolley stations, train stations, and temporary stations in the bus dataset to avoid unnecessary errors. We have applied log transformation to the dependent variable, daily ridership per stop, to help normalize its distribution. Figure 2.2 graphically shows that log transformation had led to a more normalized distribution pattern. We have also examined the effects of log transformation on the independent variables. However, we have decided to only use the log on ridership to keep the independent variables simple to interpret.

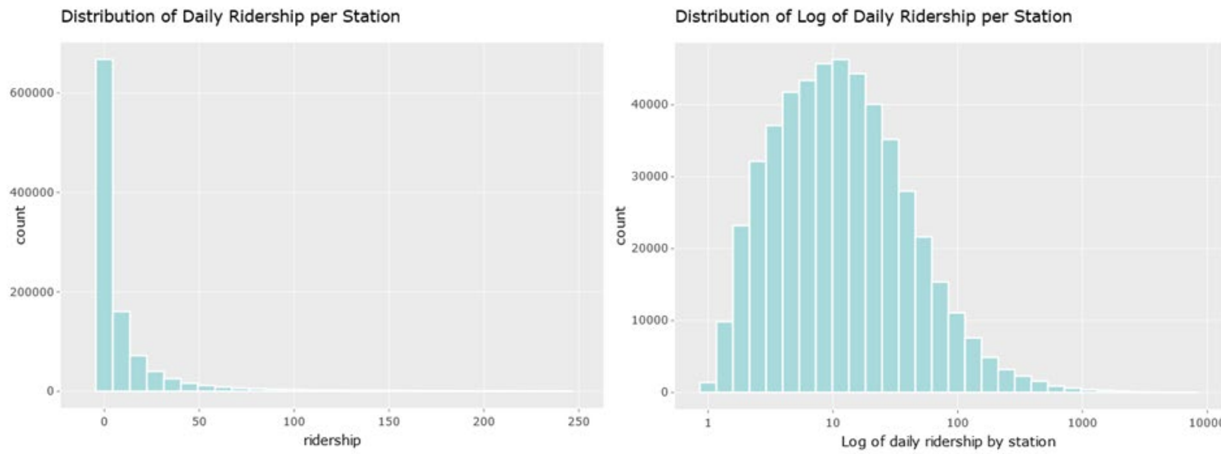

Figure 2.2 Comparison of distribution of daily bus ridership per station before and after log transformation.

This study incorporated three categories of independent variables: demographic and socioeconomic variables, transportation infrastructure variables, and temporal indicators. Table 2.1 presents the distribution for the selected variables in the final model. Demographic and socioeconomic variables, including median household income and population density selected in the final model, were collected from the 2015-2019 ACS 5-Year Estimates. Temporal variables such as precipitation were obtained from Visual Crossing using the Weather Data API. Other than precipitation, this research also set the lockdown threshold on March 23, 2020, based on SEPTA's announcement of service reduction. Transportation infrastructure variables were collected from several sources. The number of jobs accessible within 45 minutes was collected from Access Across America: Transit 2019 Data that retrieved the Data Repository for the University of Minnesota; the parking cost was obtained from DVRPC provided TAZ zonal data, and the transfer station indicator variable is derived from the GTFS data.

Table 2.1 Distribution for selected variables in the final model

Continuous Variables							
Variable Name	Min	1st Qu.	Median	Mean	3rd Qu	Max.	Count of 0
Extrapolated ridership	0	0	3.333	22.513	16.32	20882.33	110635

Precipitation (Decimeter)	0	0	0	0.02937	0.01270	0.42926	260815
Median Household Income (100k)	0	0.2878	0.4320	0.4772	0.6207	1.4365	14499
Transit Job Accessibility in 45 mins (100k)	0	2.369	4.734	3.996	5.355	7.082	63
Population Density (1000/km^2)	0	3.686	6.348	6.773	9.170	33.681	12619
Daily Parking Cost	0	0	0	1.761	3	21.5	261763
Indicator Variables							
Variable Name	False Count			True Count			
Is Weekend	285391			113155			
Is Transfer Center	397547			999			
Is During Lockdown	217137			181409			

2.4 Results

We constructed a multi-level regression model using the log transformation of ridership data at each stop as the dependent variable. The regression modeling was conducted in R using the mle4 package. We started refining the model by including all variables mentioned in the methodology section and eliminating insignificant variables. The final model regression result is shown in Table 2.2**,** with variables and their interaction terms with the Covid lockdown indicator variable being separated into the left and the right column.

Table 2.2 Regression Model Result

Variable (Units)	**Estimate**	**Standard Error**	**Variable * Lockdown**	**Estimate**	**Standard Error**
Covid lockdown order After lockdown: 1 Before lockdown: 0	-0.193*	0.114			
Weekday Weekday: 1 Weekend: 0	0.856***	0.097	Weekday * Lockdown	-0.630***	0.134
Rain (Decimeter)	-0.803**	0.391			
Median household income (100k dollars)	-0.699***	0.056	Median household income * Lockdown	-0.071***	0.011
Job accessible within	0.189***	0.009	Job accessible within 45	-0.022***	0.002

45 mins of transit (100k jobs)			mins of transit * Lockdown		
Population density (100k/km^2)	3.742***	0.378	Population density * Lockdown	-1.938***	0.077
Daily parking cost (10 dollar)	0.365***	0.049	Daily parking cost * Lockdown	-0.150***	0.010
Transfer station Transfer: 1 Non-transfer: 0	3.429***	0.289	Transfer station * Lockdown	-0.133**	0.055
Constant	1.078***	0.093			
Observations: 398,546 Log Likelihood: -523,287.800 Akaike Inf. Crit.: 1,046,612.000 Bayesian Inf. Crit.: 1,046,8080.000					
*p<0.1; **p<0.05; ***p<0.01					

The estimated coefficients of the independent variables show us the correlation between each variable and bus ridership. The estimated coefficients of the interaction terms between a variable and the Covid-19 indicator variable can be interpreted as the marginal effect of the variable after the outbreak of Covid-19. Detailed interpretation of the estimated coefficients will be discussed in the next three subsubsections.

2.4.1 Demographic and Socioeconomic variables

The coefficient of Median Household Income variable is -0.699, which indicates that before the Covid-19 lockdown order, with a 100k dollars increase in median household income of a census tract, ridership will decrease by around 70%. The coefficient of the interaction term between Median Household Income and the Covid Lockdown Order indicator variable is -0.071, indicating that the effect of Median Household Income on lowering bus ridership increased after the Covid-19 lockdown. With a 100k dollars increase in median household income, ridership will decrease by around 77% ((-0.699)+(-0.071)=-0.77) after the Covid-19 lockdown order.

The variable we used to indicate job density and accessibility, Job Accessible within 45 mins of Transit, is indicated as positively correlated with bus ridership with a coefficient of 0.189, meaning that with an increase of 100k jobs, transit ridership will increase by 19% before Covid-19. However, after the Covid-19 lockdown, the marginal effect decreased with a coefficient of -0.022 for the interaction term between Job Accessible within 45 mins of Transit and the Covid Lockdown Order, indicating that with an increase of 100k jobs, transit ridership will only experience around 17% increase (0.189+(-0.022)=0.167).

Similarly, population density shows the same decrease in marginal effect after the Covid-19 lockdown order. The variable, Population Density (100k per square kilometers) has a coefficient of 3.742 and the interaction of the variable with the Covid Lockdown Order variable has a coefficient of -1.938.
This means that before the Covid-19 lockdown, with all other variables being constant, with an increase of 100k people in a square kilometers, bus ridership will increase by 374%, but after the lockdown, with the same number of increase in population density, bus ridership will only increase by around 180% (3.742+(-1.938) = 1.804). We see a more exacerbated negative effect (median household income) and a milder positive effect (population density and job accessibility) in bus ridership after the Covid-19 lockdown order being implemented in Philadelphia.

2.4.2 Transportation Infrastructure Variables

Transportation infrastructure related variables include Daily Parking Cost, which indicates the average parking cost in a census tract in cents and Transfer Station, which is an indicator variable of whether the bus stop is a transfer station for other lines of services. The coefficient of Daily Parking Cost is positively associated with bus ridership (0.365), indicating that the higher the parking cost, the more ridership for buses in the area. After the Covid-19 lockdown order, the marginal effect decreased, indicated by the negative coefficient (-0.150) of the interaction term. In other words, with a 10 dollar increase in the average parking cost, bus ridership will increase by 21.5% after Covid-19 compared to the 36.5% before the pandemic.
Transfer Station has a coefficient of 3.429, which means that a bus stop that is a transfer station will have 343% more ridership than its counterpart without any transfer lines. The interaction term between Transfer Station and Covid Lockdown Order has a coefficient of -0.113, which indicates that after lockdown, the marginal effect is negative and the effect of a bus station being a transfer station was 332% than its counterpart. We see a decrease in transit ridership across both types of stations, transfer and non-transfer. Although the difference between these two types of stations became smaller after the Covid-19 outbreak, there is still a much higher number of transit riders at transfer stations.

2.4.3 Temporal and Meteorological Variables

The final model includes three temporal and meteorological variables. Temporal variables include an indicator variable named Covid Lockdown Order indicating before or after the Philadelphia Covid-19 lockdown order on March 23rd 2020 and an indicator variable Weekday indicating whether the date of recorded ridership is a weekday or a weekend. The only meteorological variable included in the model is Rain measuring precipitation in decimeters.

The negative coefficient of Covid Lockdown Order (-0.193) indicates ridership decrease after the implementation of lockdown. The negative coefficient of Rain (-0.803) aligns with our intuition and previous literature that bus ridership decreases during rainy weather (Guo et al., 2007; Böcker et al., 2013; Koetse & Rietveld, 2009).

Lastly, Weekday's positive coefficient (0.856) shows that bus ridership is generally higher on weekdays compared to weekends. Before the Covid-lockdown, ridership during the weekdays is 85.6% higher than ridership on weekends. After the Covid-lockdown, there is a marginal decrease in the effect, reflected through the negative coefficient of interaction term between Weekday and Covid Lockdown Order (-0.630). This means that although weekdays still have more ridership compared to weekends, after the Covid-lockdown, the difference dropped by 63% and became only 22.6%.

2.5 Conclusion

This study examines how the effect of various demographic, socioeconomic, infrastructural factors on bus ridership changed before the pandemic and immediately after the Covid-19 outbreak. We situate the study in the city of Philadelphia and build a mixed-effect linear regression model using extrapolated ridership data from SEPTA's automatic passenger count. Results from the model show us how different factors correlated to bus ridership before and during the initial days of the pandemic. Factors that are positively correlated with bus ridership include: weekdays, job accessibility, population density, daily parking cost and transfer station; factors that are negatively correlated with bus ridership includes: precipitation and income. Covid-19 lockdown order is also negatively correlated with bus ridership, so as all the interaction terms between all the variables and the Covid-19 lockdown indicator variable. The negative coefficient of all the interaction terms, indicates a marginal increase in the effect of the factors negatively correlated with bus ridership and a marginal decrease in the effect of the factors positively correlated with bus ridership. This overall negative trend shown in the model coefficients reflects the drastic decrease of bus ridership after the Covid-19 outbreak and the challenges of driving ridership up faced by transit agencies. More precisely, the extrapolated APC ridership data shows that the average daily ridership decreased from around 270k to around 76k during the study period.

To visualize the effect of a single variable, we create simulative datasets where all other variables stay constant but the one chosen variable fluctuates up and down by a fixed amount. The first variable chosen is median household income. In Figure 2.3, we show how the average daily ridership would change if all data entries' household median income decrease by \$10k, \$20k, \$30k and increase by \$10k, \$20k, \$30k. We see that besides the overall drastic decrease in bus ridership after lockdown, the average daily ridership fluctuates more before Covid-19 than after the Covid-19 outbreak. From Figure 2.4, we also see a similar pattern when data is presented as daily aggregates. Figure 2.4 also reflects how the weekday and weekend ridership pattern being flattened after the Covid-19 outbreak which aligns with the model result of having a negative marginal effect for the Weekday factor. The weakening of the higher weekday and lower weekend ridership trend can also be explained by a decrease in commute trips as people started to work from home.

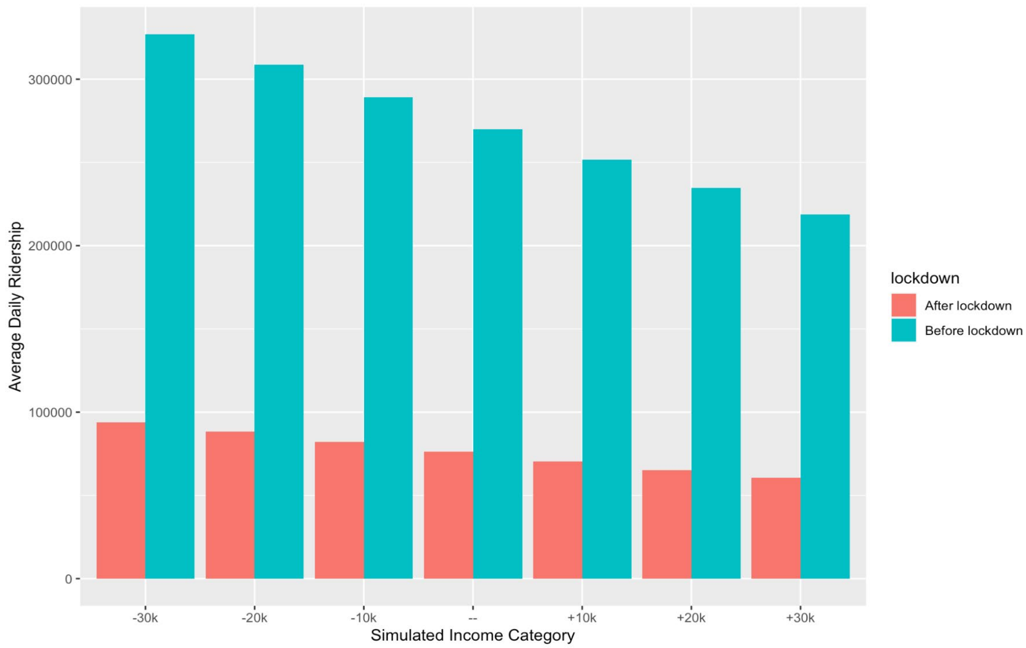

Figure 2.3 Average daily bus ridership prediction before and after the Covid-19 lockdown using simulated income data ranging from \$30k less than to \$30k more than the actual recorded average household income.

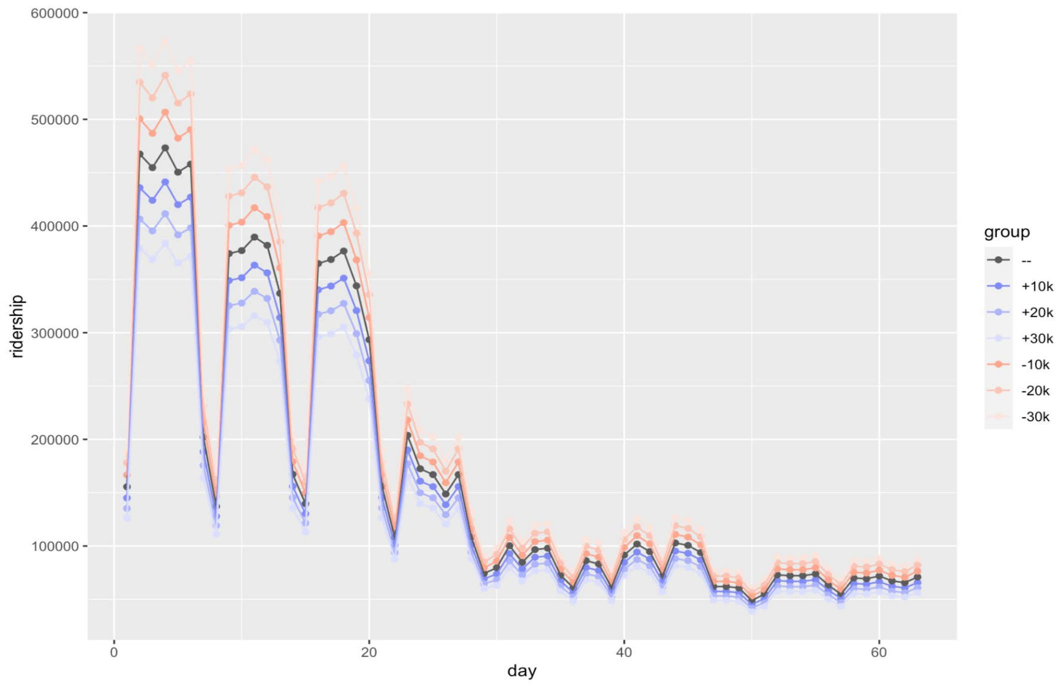

Figure 2.4 Bus ridership prediction by day using simulated income data ranging from $30k less than to $30k more than the actual recorded average household income.

We also run the same procedure on the population density variable by creating simulative datasets where all other variables stay constant and the population density variable gets fluctuated by 3, 6 and 9, which represents 300k, 600k and 900k people per square kilometer. In Figure 2.5, we show how the average daily ridership would change if all data entries' population density decreases by 3, 6 and 9 and increases by 3, 6 and 9 units. Similarly to the income variable, we see that the average daily ridership fluctuates more before Covid-19 than after the Covid-19 outbreak when we change the population density. From Figure 2.6, we present predicted bus ridership aggregated by day. We again see the same trend, where ridership decreased drastically after the Covid-19 lockdown and with the same amount of increase in population density, ridership before the Covid-19 pandemic can be increased more than ridership after the pandemic outbreak.

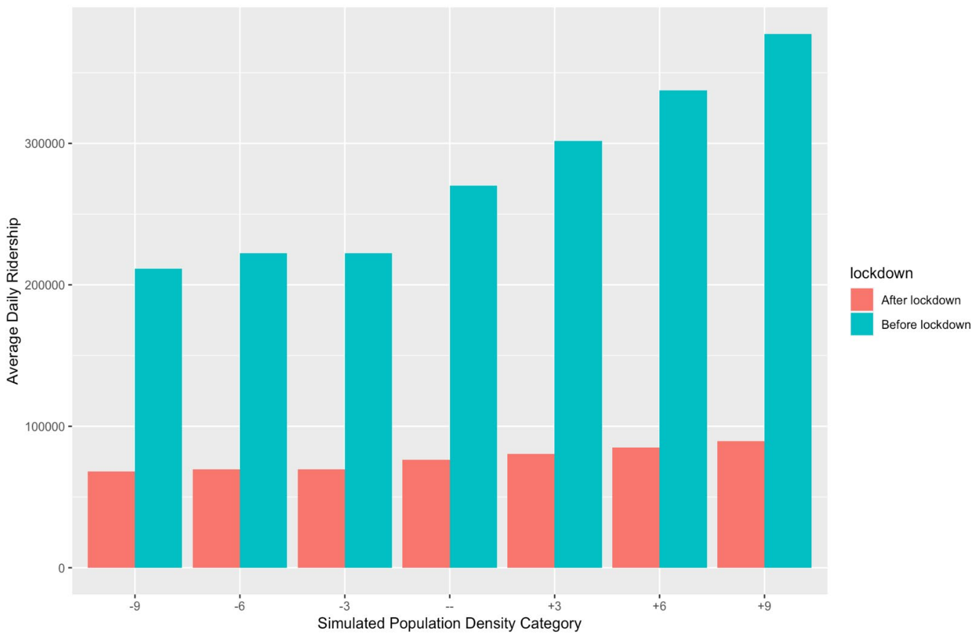

Figure 2.5 Average daily bus ridership prediction before and after the Covid-19 lockdown using simulated population density data from 900k people/km^2 less than to 900k people/km^2 more than the actual recorded population density.

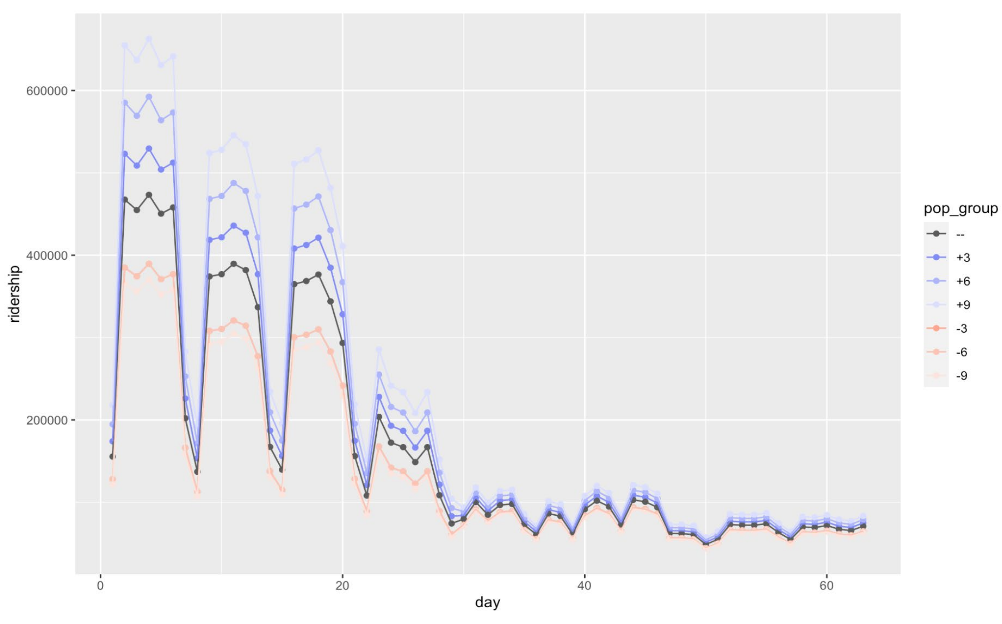

Figure 2.6 Bus ridership prediction by day using simulated population density data from 900k people/km^2 less than to 900k people/km^2 more than the actual recorded population density.

Through unpacking the effects from various socioeconomic factors on ridership, our research brings out a few highlights for the service practitioners. As all interaction terms between a variable and the Covid lockdown indicator variable have negative coefficients, this result confirms that the lockdown order has led to the reduction in ridership across all factors. The amplification of negatively correlated factors and the reduction of positively correlated factors means that during the lockdown, the elements used to generate ridership are less effective and the characters usually making ridership less attractive are more powerful. Additionally, since the drastic ridership decrease is experienced in all areas, it is hard for transit agencies to reroute during an unpredictable period of shock like the Covid-19 pandemic. It is challenging to strategically maintain or increase service through focusing on just a few elements or fewer areas in the short term.

Although it is challenging for transit agencies to continue operating during the immediate time period of a societal shock, the transit services still remain crucial to support many people and communities. Thus it is especially important for transit agencies to focus on matching cost effectiveness and be more targeted on where services go. Stations in areas with more jobs and areas that are more dense could bounce back faster. Services should be maintained and improved especially in lower income communities. Transit agencies could also keep operations going through transfer stations. Although their effect on ridership is weaker during the lockdown, the combined positive coefficient indicates the essential need of bus service in those service areas. Although public transit has gone through a tough time during the pandemic, it remains to play a crucial role of providing equitable mobility options. Better preparation for unpredictable ridership shock, solutions for faster recovery and network designs for maintaining resilience should be a continuous topic for transportation planners and researchers to focus on.

2.5.1 Limitations

We acknowledge that there are three key limitations of our paper, two of which are centered around limited access to quality public transit data and the third is around accounting for user perception of transit. Firstly, the estimated ridership data was extrapolated from a sample dataset consisting of automatic passenger count data from two vendors working with SEPTA. The variation in sampling rate and the changes in bus schedules throughout the study period led to inevitable costs of accuracy in the extrapolated data. Although we did our best to present an optimal estimation of the ridership information throughout the period, there is room for improvements in collecting ridership data to ensure higher accuracy. We also suggest future researchers and public transit agencies partner to adopt better processes of data collection and curation. Secondly, the period of data that we had access to was limited to approximately one month prior and after the Covid-19 pandemic outbreak. As the pandemic unfolds over the past two years and people's behavior changed gradually as some adopted the norm of remote living while others slowly going back to the in-person lifestyle, we encourage future studies to look into data from a longer period of time to gain insights into how a shock event impacts bus ridership in the long term and how would bus ridership recovers throughout time. Lastly, even when focusing on the immediate period of the pandemic outbreak, our study only focused on external factors and did not include user perceptions and sentiments which could have an important role in determining one's travel mode. Just as Talyor and Fink (2003) argue that a large variety of factors play a role in affecting transit ridership and it is hard to tease out the relative importance and interactions among them and even more broadly as Rittel and Webber (1973) argues that planning problems are essentially wicked problems, our study contributes as a highlight for the challenges that transit agencies face right after a societal shock and an example of how one can potentially draw patterns and insights from that initial period.

2.6 References

"2019 American Community Survey 5-Year Estimates." Accessed July 30, 2022. https://data.census.gov/cedsci/table?y=2019&d=ACS%205-Year%20Estimates%20Detailed%20Tables.

Abdullah, Muhammad, Charitha Dias, Deepti Muley, and Md. Shahin. "Exploring the Impacts of COVID-19 on Travel Behavior and Mode Preferences." *Transportation Research Interdisciplinary Perspectives* 8 (November 1, 2020): 100255. https://doi.org/10.1016/j.trip.2020.100255.

Bates, John, John Polak, Peter Jones, and Andrew Cook. "The Valuation of Reliability for Personal Travel." *Transportation Research Part E: Logistics and Transportation Review*, Advances in the Valuation of Travel Time Savings, 37, no. 2 (April 1, 2001): 191–229. https://doi.org/10.1016/S1366-5545(00)00011-9.

Beck, Matthew J., David A. Hensher, and John D. Nelson. "Public Transport Trends in Australia during the COVID-19 Pandemic: An Investigation of the Influence of Bio-Security Concerns on Trip Behaviour." *Journal of Transport Geography* 96 (October 1, 2021): 103167. https://doi.org/10.1016/j.jtrangeo.2021.103167.

Bh., Aaditya, and Rahul T.M. "A Comprehensive Analysis of the Trip Frequency Behavior in COVID Scenario." *Transportation Letters* 13, no. 5–6 (May 28, 2021): 395–403. https://doi.org/10.1080/19427867.2021.1906478.

Bliss, Laura, Jeremy C. F. Lin, and Marie Patino. "Pandemic Travel Patterns Hint at Our Urban Future." *Bloomberg.Com*. Accessed February 3, 2022. https://www.bloomberg.com/graphics/2020-coronavirus-transportation-data-cities-traffic-mobility/.

Böcker, Lars, Martin Dijst, and Jan Prillwitz. "Impact of Everyday Weather on Individual Daily Travel Behaviours in Perspective: A Literature Review." *Transport Reviews* 33, no. 1 (January 1, 2013): 71–91. https://doi.org/10.1080/01441647.2012.747114.

Boisjoly, Geneviève, Emily Grisé, Meadhbh Maguire, Marie-Pier Veillette, Robbin Deboosere, Emma Berrebi, and Ahmed El-Geneidy. "Invest in the Ride: A 14 year Longitudinal Analysis of the Determinants of Public Transport Ridership in 25 North American Cities." *Transportation Research Part A: Policy and Practice* 116 (October 1, 2018): 434–45. https://doi.org/10.1016/j.tra.2018.07.005.

Bristol, University of. "Random Intercept Models." University of Bristol. Accessed July 30, 2022. http://www.bristol.ac.uk/cmm/learning/videos/random-intercepts.html.

Bureau, US Census. "Decennial Census of Population and Housing Datasets." Census.gov. Accessed July 30, 2022. https://www.census.gov/programs-surveys/decennial-census/data/datasets.html.

Chakour, Vincent, and Naveen Eluru. "Examining the Influence of Stop Level Infrastructure and Built Environment on Bus Ridership in Montreal." *Journal of Transport Geography* 51 (February 1, 2016): 205–17. https://doi.org/10.1016/j.jtrangeo.2016.01.007.

Chen, Cynthia, Don Varley, and Jason Chen. "What Affects Transit Ridership? A Dynamic Analysis Involving Multiple Factors, Lags and Asymmetric Behaviour." *Urban Studies* 48, no. 9 (July 1, 2011): 1893–1908. https://doi.org/10.1177/0042098010379280.

Currie, Graham, and Ian Wallis. "Effective Ways to Grow Urban Bus Markets – a Synthesis of Evidence." *Journal of Transport Geography*, Growing Public Transport Patronage, 16, no. 6 (November 1, 2008): 419–29. https://doi.org/10.1016/j.jtrangeo.2008.04.007.

Doorley, Ronan, Alex Berke, Ariel Noyman, Luis Alonso, Josep Ribó, Vanesa Arroyo, Marc Pons, and Kent Larson. "Mobility and COVID-19 in Andorra: Country-Scale Analysis of High-Resolution Mobility Patterns and Infection Spread." *IEEE Journal of Biomedical and Health Informatics* 26, no. 1 (January 2022): 183–93. https://doi.org/10.1109/JBHI.2021.3121165.

Figler, Scott A., P. S. Sriraj, Eric W. Welch, and Nilay Yavuz. "Customer Loyalty and Chicago, Illinois, Transit Authority Buses: Results from 2008 Customer Satisfaction Survey." *Transportation Research Record* 2216, no. 1 (January 1, 2011): 148–56. https://doi.org/10.3141/2216-16.

Gomez-Ibanez, Jose A. "Big-City Transit Rider Snip, Deficits, and Politics: Avoiding Reality in Boston." *Journal of the American Planning Association* 62, no. 1 (March 31, 1996): 30–50. https://doi.org/10.1080/01944369608975669.

Gottehrer-Cohen, Zach. "Interactive Map Shows Which Parts Of NYC Are Relying On The Subway Most During Coronavirus." Gothamist, May 5, 2020. https://gothamist.com.

Guerra, Erick, and Robert Cervero. "Cost of a Ride." *Journal of the American Planning Association* 77, no. 3 (July 1, 2011): 267–90. https://doi.org/10.1080/01944363.2011.589767.

Guo, Zhan, Nigel H. M. Wilson, and Adam Rahbee. "Impact of Weather on Transit Ridership in Chicago, Illinois." *Transportation Research Record* 2034, no. 1 (January 1, 2007): 3–10. https://doi.org/10.3141/2034-01.

Halvorsen, Anne, Daniel Wood, Darian Jefferson, Timon Stasko, Jack Hui, and Alla Reddy. "Examination of New York City Transit's Bus and Subway Ridership Trends During the COVID-19 Pandemic." *Transportation Research Record*, September 8, 2021, 03611981211028860. https://doi.org/10.1177/03611981211028860.

Hu, Songhua, and Peng Chen. "Who Left Riding Transit? Examining Socioeconomic Disparities in the Impact of COVID-19 on Ridership." *Transportation Research Part D: Transport and Environment* 90 (January 1, 2021): 102654. https://doi.org/10.1016/j.trd.2020.102654.

Kain, John F, and Zvi Liu. "Secrets of Success: Assessing the Large Increases in Transit Ridership Achieved by Houston and San Diego Transit Providers." *Transportation Research Part A: Policy and Practice* 33, no. 7 (September 1, 1999): 601–24. https://doi.org/10.1016/S0965-8564(99)00009-9.

Kim, Karl. "Impacts of COVID-19 on Transportation: Summary and Synthesis of Interdisciplinary Research." *Transportation Research Interdisciplinary Perspectives* 9 (March 2021): 100305. https://doi.org/10.1016/j.trip.2021.100305.

Kim, Suji, Sujin Lee, Eunjeong Ko, Kitae Jang, and Jiho Yeo. "Changes in Car and Bus Usage amid the COVID-19 Pandemic: Relationship with Land Use and Land Price." *Journal of Transport Geography* 96 (October 1, 2021): 103168. https://doi.org/10.1016/j.jtrangeo.2021.103168.

Koetse, Mark J., and Piet Rietveld. "The Impact of Climate Change and Weather on Transport: An Overview of Empirical Findings." *Transportation Research Part D: Transport and Environment* 14, no. 3 (May 1, 2009): 205–21. https://doi.org/10.1016/j.trd.2008.12.004.

Liu, Luyu, Harvey J. Miller, and Jonathan Scheff. "The Impacts of COVID-19 Pandemic on Public Transit Demand in the United States." *PLOS ONE* 15, no. 11 (November 18, 2020): e0242476. https://doi.org/10.1371/journal.pone.0242476.

———. "The Impacts of COVID-19 Pandemic on Public Transit Demand in the United States." *PLOS ONE* 15, no. 11 (November 18, 2020): e0242476. https://doi.org/10.1371/journal.pone.0242476.

Long, Alfie, Ffion Carney, and Jens Kandt. "Who Is Returning to Public Transport for Non-Work Trips after COVID-19? Evidence from Older Citizens' Smart Cards in the UK's

Second Largest City Region." *Journal of Transport Geography* 107 (February 1, 2023): 103529. https://doi.org/10.1016/j.jtrangeo.2023.103529.

"Mapping Route-By-Route Who's Riding Transit During the Pandemic," July 15, 2020. https://transitcenter.org/mapping-route-by-route-whos-riding-transit-during-the-pandemic/.

McLeod Jr, Malcolm S., Kevin J. Flannelly, Laura Flannelly, and Robert W. Behnke. "MULTIVARIATE TIME-SERIES MODEL OF TRANSIT RIDERSHIP BASED ON HISTORICAL, AGGREGATE DATA: THE PAST, PRESENT AND FUTURE OF HONOLULU (WITH DISCUSSION AND CLOSURE)." *Transportation Research Record*, no. 1297 (1991). https://trid.trb.org/view/359621.

Murphy, Brendan, and Andrew Owen. "Access Across America: Transit 2019 Data," January 26, 2021. https://doi.org/10.13020/xgr6-pt58.

Osorio, Jesus, Yining Liu, and Yanfeng Ouyang. "Executive Orders or Public Fear: What Caused Transit Ridership to Drop in Chicago during COVID-19?" *Transportation Research Part D: Transport and Environment* 105 (April 1, 2022): 103226. https://doi.org/10.1016/j.trd.2022.103226.

Parker, Madeleine E. G., Meiqing Li, Mohamed Amine Bouzaghrane, Hassan Obeid, Drake Hayes, Karen Trapenberg Frick, Daniel A. Rodríguez, Raja Sengupta, Joan Walker, and Daniel G. Chatman. "Public Transit Use in the United States in the Era of COVID-19: Transit Riders' Travel Behavior in the COVID-19 Impact and Recovery Period." *Transport Policy* 111 (September 2021): 53–62. https://doi.org/10.1016/j.tranpol.2021.07.005.

Pasha, Mosabbir, Shakil Mohammad Rifaat, Richard Tay, and Alex De Barros. "Effects of Street Pattern, Traffic, Road Infrastructure, Socioeconomic and Demographic Characteristics on Public Transit Ridership." *KSCE Journal of Civil Engineering* 20, no. 3 (April 1, 2016): 1017–22. https://doi.org/10.1007/s12205-016-0693-6.

Polzin, Steven, Tony Choi, and United States. Department of Transportation. Office of the Assistant Secretary for Research and Technology. "COVID-19's Effects on The Future of Transportation," January 14, 2021. https://doi.org/10.21949/1520705.

Qi, Yi, Jinli Liu, Tao Tao, and Qun Zhao. "Impacts of COVID-19 on Public Transit Ridership." *International Journal of Transportation Science and Technology*, November 1, 2021. https://doi.org/10.1016/j.ijtst.2021.11.003.

Rittel, Horst W. J., and Melvin M. Webber. "Dilemmas in a General Theory of Planning." *Policy Sciences* 4, no. 2 (1973): 155–69.

Shields, Mike. "Economy League - The Changing Distribution of Poverty in Philadelphia." Economy League of Greater Philadelphia, 2020. http://http://economyleague.org//providing-insight/leadingindicators/2020/12/16/phlpov19.

Streva, Virginia. "SEPTA Reduces Transit Services, All Lines to Operate on Saturday Schedule Because of COVID-19 | PhillyVoice," 2020. https://www.phillyvoice.com/septa-saturday-schedule-coronavirus-covid-19-regional-rail-market-frankford-broad-street-subway/.

Subbarao, Saladi S. V., and Raghuram Kadali. "Impact of COVID-19 Pandemic Lockdown on the Public Transportation System and Strategic Plans to Improve PT Ridership: A Review." *Innovative Infrastructure Solutions* 7, no. 1 (November 19, 2021): 97. https://doi.org/10.1007/s41062-021-00693-9.

Tao, Sui, Jonathan Corcoran, Francisco Rowe, and Mark Hickman. "To Travel or Not to Travel: 'Weather' Is the Question. Modelling the Effect of Local Weather Conditions on Bus Ridership." *Transportation Research Part C: Emerging Technologies* 86 (January 1, 2018): 147–67. https://doi.org/10.1016/j.trc.2017.11.005.

Taylor, B D, and C N Y Fink. "Explaining Transit Ridership: What Has the Evidence Shown?" *Transportation Letters* 5, no. 1 (January 1, 2013): 15–26. https://doi.org/10.1179/1942786712Z.0000000003.

Taylor, Brian D., and Camille N. Y. Fink. "The Factors Influencing Transit Ridership: A Review and Analysis of the Ridership Literature," September 1, 2003. https://escholarship.org/uc/item/3xk9j8m2.

Taylor, Brian D., Douglas Miller, Hiroyuki Iseki, and Camille Fink. "Nature and/or Nurture? Analyzing the Determinants of Transit Ridership across US Urbanized Areas." *Transportation Research Part A: Policy and Practice* 43, no. 1 (January 1, 2009): 60–77. https://doi.org/10.1016/j.tra.2008.06.007.

Tiikkaja, Hanne, and Riku Viri. "The Effects of COVID-19 Epidemic on Public Transport Ridership and Frequencies. A Case Study from Tampere, Finland." *Transportation Research Interdisciplinary Perspectives* 10 (June 1, 2021): 100348. https://doi.org/10.1016/j.trip.2021.100348.

"Transit Is Essential: 2.8 Million U.S. Essential Workers Ride Transit to Their Jobs," March 24, 2020. https://transitcenter.org/2-8-million-u-s-essential-workers-ride-transit-to-their-jobs/.

Wang, Haoyun, and Robert B. Noland. "Bikeshare and Subway Ridership Changes during the COVID-19 Pandemic in New York City." *Transport Policy* 106 (June 1, 2021): 262–70. https://doi.org/10.1016/j.tranpol.2021.04.004.