Investigating Public Transportation Use in the United States

Mikita Zhylinski
Lake Forest College, zhylinskim@lakeforest.edu

Follow this and additional works at: https://publications.lakeforest.edu/seniortheses

Part of the Econometrics Commons, and the Public Economics Commons

Recommended Citation
Investigating Public Transportation Use in the United States

Abstract
Public transportation is considered to be more cost-efficient and ecological than private transportation. Americans, however, use public transportation much less than citizens of European countries. How can public policy change American preferences, to increase usage of public transportation? I examine the relationship between personal characteristics (e.g., education, income, race) and public transportation usage while considering accessibility and geographic factors. In order to investigate these issues, I use the National Household Travel Survey 2009 (NHTS). Two samples of people with different access to public transportation are compared to better understand the relationship. The results show that African Americans and people in high population density areas use public transportation more often than are white Americans and people in low population areas respectively, but income has no significant effect on public transportation use. These results match most previous studies on public transportation. In conclusion, the policy should improve public transportation network systemwide, but simple-aimed gasoline taxes are predicted to be ineffective.

Document Type
Thesis

Degree Name
Bachelor of Arts (BA)

Department or Program
Economics, Business, and Finance

First Advisor
Amanda J. Felkey

Second Advisor
Robert J. Lemke

Third Advisor
Enrique Treviño

Subject Categories
Econometrics | Public Economics
LAKE FOREST COLLEGE

Senior Thesis

INVESTIGATING PUBLIC TRANSPORTATION USE IN THE UNITED STATES

by

Mikita Zhylinski

April 16, 2018

The report of the investigation undertaken as a Senior Thesis, to carry two courses of credit in the Department of Economics, Business, and Finance

_______________________  _______________________
Davis Schneiderman       Amanda J. Felkey, Chairperson
Krebs Provost and Dean of the Faculty

_____________________
Robert J. Lemke

_____________________
Enrique Treviño
ABSTRACT

Public transportation is considered to be more cost-efficient and ecological than private transportation. Americans, however, use public transportation much less than citizens of European countries. How can public policy change American preferences, to increase usage of public transportation? I examine the relationship between personal characteristics (e.g., education, income, race) and public transportation usage while considering accessibility and geographic factors. In order to investigate these issues, I use the National Household Travel Survey 2009 (NHTS). Two samples of people with different access to public transportation are compared to better understand the relationship. The results show that African Americans and people in high population density areas use public transportation more often than are white Americans and people in low population areas respectively, but income has no significant effect on public transportation use. These results match most previous studies on public transportation. In conclusion, the policy should improve public transportation network systemwide, but simple-aimed gasoline taxes are predicted to be ineffective.
DEDICATION

To my teachers
I. Introduction

Pollution and organization concerns threaten modern US cities. Both concerns are caused by a variety of urban issues that will be discussed later. I am going to address these problems through the lens of urban transportation. Over 90% of US households own a private vehicle (Federal Highway Administration, 2009) – a convenient means of individual transportation, and a sign of status. For the last century a private vehicle was a part of the American dream, a representation of the freedom to move. However, all freedoms come with certain responsibilities: cars produce pollution and, and the prevalence of cars transporting people to and from work can cause congestion, leading to longer commute times.

Both problems are most prevalent in urban areas, as rural areas have lower population density and fewer cars. Cities, on the other hand, are exposed to the effect commonly known as “smog” – an intense city pollution in the form of a gas-dust cloud. Personal vehicles are among the top contributors to smog and other types of pollutions. The amount of smog varies across cities largely due to geographical and climate differences (Sierra Club, 2001). The most dangerous result of smog, of course, is the negative effects on health and nature (Environmental Protection Agency, 1999). I will focus only on the health issues caused by cars because the impact of greenhouse gases on ecology is unquantifiable. In their study, McCubbin and Delucchi estimate the lower bound for the total costs of health issues caused by vehicles to be twenty-four billion dollars for the year 1990 (McCubbin & Delucchi, 1996, pp. 212, 275-80). Due to a 40% increase in highway vehicles quantity since 1990 (Bureau of Transportation Statistics, 2017), the annual amount of money spent on health issues caused by vehicles is much bigger now than it was 30 years ago.
In addition to health and environmental concerns, city-planning and urbanization over the last 100 years has been set largely to accommodate car usage. Roads and highways take city space away from the people, as the car-centered logistic of a city creates a conveyor-like environment where everybody must move in one direction. This creates passengers and vehicle congestion that increase the commute time. During rush hours cars fill the street of cities with pollutant emissions, and people have to breathe it in order to get to their destination. Imagine the streets of New York, Chicago or Los Angeles with only one car lane in each direction. Suddenly there will be space for trees and plants, as streets have the potential to become a place of socialization, instead of a noisy, smelly place you have to walk through. Unfortunately, such an idealistic scenario would cause massive traffic problems. Even now, when most space of a street is covered by automobile roads, the traffic problem is not extinct. In the period of 1977-1990, the US enlarged the high-way infrastructure, hoping to solve the traffic problem. The outcome was an increased amount of traffic (Hansen, 1995). Adding more roads does not solve the traffic problem, as new roads can belogistically inefficient and lower the traffic flow even further (Virtual Cell Program n.d). This happens as construction of more roads increases the demand for more road activity, which is why more roads do not always lead to less congestion. That said, building more roads or regulating old ones can help to reduce traffic congestion if done properly.

Rather than building more roads, another way to solve the traffic problem is to remove cars from the roads. What if instead of 10,000 vehicles on the street, there were 1,000? It would reduce the pollution ten times and would require far less space for vehicle infrastructure. This solution is public transportation.

In this thesis, I am going to discuss the demand for transportation. With demand as a given, I will further assume that there are two transportation options that “compete” for users – public and
private modes of transportation. Let me define the terms to avoid misinterpretations. Public transportation is a publicly provided service of transportation; it includes buses, trains, planes, etc. Private transportation is a privately provided service of transportation; it includes private vehicles, as well as cabs and various taxi services. Despite the fact that many modern taxi services like Uber provide a communal transportation service, I do not consider it a public transportation because it is a privately-owned service. As will be discussed later, private transportation poses a larger environmental concern and less social benefit compared to public transportation.

Extensive public transportation systems can, in theory, be comprehensive enough to provide most transportation needs in modern cities. Buses and trolleybuses can transport dozens or hundreds of passengers at once. In European countries one can live in a city without relying on private transportation. The national transportation system (e.g., trains and buses) can allow transportation to almost every town. According to national data, passenger-miles per gallon of gasoline for public transportation (buses and trains) is higher than for personal vehicles (cars and trucks) (Alternative Fuels Data Center, 2018). This means public transportation is more fuel efficient, and, on average, occupies less road space compared to private vehicles, while transporting more passengers. Bus routes inside the city can cover small streets, and unpopular corners, while heavy passenger flow is better processed by the subway and trains. London developed the first electric traction train system in 1890, and it remains a very energy efficient system of transportation. Electric locomotives produce no pollution in the city, and, in many cases like in London, the tracks run underground. In the US, the Chicago subway system uses a mixed system of tracks with some tracks on the ground, some underground, and some elevated above ground. This mixed system of tracks results in the most cost-effective use of the space available. The subway removes a large amount of traffic from the city roads, allowing the
citizens to leave their vehicle at home, and relieving parking, gasoline use, and engine repair issues.

Despite the various benefits of public transportation, Americans can be reluctant to use public transportation, even when it is available, for a variety of reasons. The most common reasons not to use public transportation are often caused by passengers, not the vehicle itself. Indeed, a closed space with a lot of people inside a bus can be potentially more dangerous than a car, however, the statistics reports public transportation vehicles to be much safer compared to private vehicles (American Public Transportation Association, 2018). Another reason why public transportation is not used more often is due to a common belief that buses are for the poor. However, people of all income levels tend to use public transportation more when it is easily accessible and policies are made to increase availability of public transportation for US citizens (Mallett, 2018).

The US cities’ policies aimed at public transportation development do not increase public transportation use at an appropriate scale. Inflation-adjusted amount of investments in public transportation increased by 50% between 1990 and 2009, but the total number of public transportation trips grew only by 20% over the same period of time (Hughes-Cromwick & Dickens, 2018, pp. 7, 24). The diminishing return to investments is expected; however, Taylor emphasizes that it happens due to an inefficient distribution of resources among public transportation projects in US cities. Instead of concentrating on expenditure and service improvements on single-line projects, the public transportation authorities should have prioritized a systemwide improvement of service frequency and innovation in transit pricing for the whole public transportation system (Taylor, Miller, Hiroyuki, & Camille, 2008, p. 16).
II. Literature Review

1. Economic Influence of Transport

The physical mobility of people strongly depends on the availability of transportation services. People of low economic status can rarely afford to buy a car. In the absence of public transportation, it makes them physically immovable. Physical immobility of population is one of the causes of poverty in low-income communities (Wachs, 2010). This issue is present in the US, as despite the top-chart economy of the world, 10% to 15% of the US population is in poverty (McCarthy, 2019). There are many ways to address this issue and the physical mobility problem is the one that causes the perpetuating effect of poverty.

Immobility of a population contributes to unemployment. Physical mobility of people who do not own a car and do not have access to public transportation is limited to their walking distance. Such individuals are unable to search for employment beyond their household area. Welfare recipients significantly improve their chances of finding a job if they own a car, as it significantly improves their physical mobility (Blumenberg & Ong, 1997, pp. 10-15). Public transportation provides a similar service to a car, as it extends the physical mobility reach of people from their household area to all areas that are reachable by the transportation network.

That is why people who are able to access a public transportation service by walking are better able to reach job opportunities in areas well-served by public transportation (Cervero, Sandoval, & Landis, 2002, p. 61). As a result, an increase in physical mobility among the low-income population improves their economic stability by providing them with job opportunities that were beyond their physical reach.

Western countries address the issue of physical mobility in two main ways. Central nations of the European Union (France, Germany) use public transportation to equate the income
distribution among the population. In these countries, a fraction of income spent on transportation increases with the individual income (Institute for Transportation and Development Policy, 2019). Low-income populations have an easier access to a cheap public transportation service. In comparison, the US responded by increasing the car ownership, and loosening access to personal vehicles. The number of privately-owned cars in the US has been steadily increasing every five years since 1960; currently there are 800 privately owned cars per 1000 population (Bureau of Transportation Statistics, 2017). This happened as an average passenger per mile cost, that includes the cost of buying, insuring and repairing a private vehicle, is comparable to the average passenger per mile cost of a public transportation (O’Toole, 2018). These statistics account only for the national average, while the cost of yearly public transportation pass in the largest US cities is actually less ($4,000 to $6,000)\(^1\) than the average annual cost of vehicle maintenance ($8,000).\(^2\) However, low-income people spend a bigger fraction of their income to buy a private vehicle compared to higher income groups. People in poverty or welfare-recipients have to save a larger fraction of their income in order to buy a car and increase their mobility. Low-income individuals who have a car spend up to 30% of their income as a transportation cost, while people with higher annual income spend 15% of their income as a transportation cost (Institute for Transportation and Development Policy, 2019). Such distribution of transportation spending contributes to a lack of savings among a low-income population in the US. This perpetuates as one of the main factors of poverty, which contributes to physical immobility.

\(^1\) The numbers above are an average cost for a 12 monthly passes for public transportation in New-York (MTA, 2019), Chicago, Los Angeles (Reynolds, 2019), summarized with a lower and an upper bounds for a 12 monthly passes for a commuter rail in Chicago (Metra, n.d.) and Boston (Massachusetts Bay Transportation Authority, n.d.).

\(^2\) The number above is an average annual cost of maintenance for all vehicle’s categories (Hybrid, Electric, SUV, Sedan, Pickup) estimated by September 2019 (American Automobile Association, 2019).
In order to understand the problem better, let’s consider a possible example with real numbers. People with the annual income of $20,000 to $30,000 can rarely afford the annual cost of a private vehicle – $8,000. The only choice they have is to buy a cheaper vehicle that will require more frequent maintenance, in the hope to find a job that will pay these expenses; or to use public transportation. Given their city has a public transportation network and the potential employment locations do not exceed the city limits, annual public transportation costs should not exceed $1,500 per individual (Reynolds, 2019). Such decrease in transportation spending, compared to a national average of 30% for a low-income population, will allow for an increase in economic stability of the low-income population.

Private vehicles carry an additional social cost in the form of negative externalities: reduced city space, a worse environment, and increased medical expenses. While the cost of city space that is occupied because of the excess number of vehicles is hard to evaluate, it suffices to say that some US cities have as much as five parking stalls per household (Scharnhorst, 2018). The second type of externality caused by private vehicles, which directly affects our lives, is medical expenses that are caused by pollution and accidents. According to McCubbin and Delucci (1996), the lower bound for medical expenses caused by private vehicles was $24 billion dollars in 1990. Whatever the actual amount was in 1990, it is unlikely to decrease due to a 40% increase in vehicle quantity (Bureau of Transportation Statistics, 2017) and an 8% increase in total road mileage due to new roads built (Federal Highway Administration, 2018). Ischemic heart diseases, respiratory infections, strokes, and lung cancer are on the top of the list of diseases caused by vehicles emissions. For the year 2010, five of the top ten causes of death or disability were associated with injuries and air pollution caused by motorized vehicles (World Bank, 2014, pp. 22-31). For the same year in North America, the death rate from air pollution due to motorized vehicles is estimated to be 5 deaths per 100,000 population. Apart from the long-
term diseases, vehicles also cause injuries in a form of traffic accidents. The US road injury death rate was 14 deaths per 100,000 population in 2010 (World Bank, 2014, pp. 22-31). These negative externalities caused by private vehicles are distributed among the entire US population, and not just vehicle owners. The solutions would be to increase vehicles’ safety or to reduce the number of vehicles on the roads.

Since vehicle safety regulations are beyond the scope of this thesis, I will concentrate on the positive benefits that could accrue to society from reducing the number of privately owned vehicles in the US. A 10% decrease of vehicle usage would presumably reduce the annual medical costs associated with vehicle usage by at least $2.4 billion dollars according to the 1990 estimations (McCubbin & Delucchi, 1996, pp. 212, 275-285). Due to an increase of vehicle usage since 1990, the medical costs associated with vehicle usage are likely to be higher now, and a decrease of vehicle usage now would probably decrease the medical costs in 2020 more than in 1990.

The way to achieve reduced dependence on privately-owned vehicles, of course, is by expanding the use of public transportation. This will lead to a decrease of overall medical expenditures as public transportation travel is ten times safer per mile of travel than a private vehicle (American Public Transportation Association, 2018). In addition, buses and trains occupy less road space, consume less gasoline, and do not require parking lots inside the city as compared to private transportation vehicles. However, the reduction of negative externalities is not the only reason why government should invest in public transportation.

Investments in public transportation generate a larger positive effect on the economy than investments in private transportation do. According to 2011 estimations, public transportation projects generate 15% to 31% more jobs per dollar spent compared to the construction and maintenance of roads and bridges (Smart Growth America, 2011, p. 2). According to the same
study, new public transportation routes in low employment communities produce 2.5 times more jobs compared to new public transportation lanes in the high employment communities. This happens because low employment is often caused by low physical mobility of population in an area. Once the mobility increases (due to a new public transportation lane), employment also increases, as people have an access to more job opportunities. This supports the earlier claim that the public transportation access yields high employment opportunities for the low-income population group and unemployed people. In addition to the accessibility issue, public transportation also yields direct economic opportunities to the population. Every $1 billion invested in public transportation generates $4 billion in economic returns and creates over 50,000 jobs (American Public Transportation Association, 2018). Consequently, public transportation is a much better solution to the physical immobility problem than private transportation. In order for people to understand that, they should be given an appropriate economic and social reason through various transportation policies.

2. Transportation Policies Research

According to the latest reports, if we were to exclude the New York area, the national ridership in public transportation has declined by 7% over the last decade (Mallett, 2018, p. 2). Despite a lot of research, there is no comprehensive explanation for the decline in ridership. One of the issues that prevents concrete answers is that national trends are not always reflected on a local level. Different local areas have different factors that affect the public transportation ridership. The two national factors that are persistent seem to be the drop of the price of the substitute good – cars, and an increase in the public transportation service supply (Mallett, 2018, p. 5). The price of a private vehicle is correlated with two factors: gasoline price and vehicle ownership. The price of transportation per vehicle mile follows the trends of gasoline prices. A
drop in gasoline prices leads to more frequent private vehicle rides due to a cost decrease. This causes a decrease in public transportation ride share, as a private vehicle’s cost quickly adjusts to a change in gasoline prices. Another factor that was found important in explaining transit ridership is the availability of personal vehicles (Taylor, Miller, Hiroyuki, & Camille, 2008). The argument is that the availability of personal vehicles is the most significant factor in declining public transportation ridership. The increase in the number of vehicles in local households is strongly correlated with a decrease of public transportation use in the area (Manville, Taylor, & Blumenberg, 2018, pp. 9-11). One of the control methods proposed in 2018 by Mallett – a specialist in transportation policy – was to raise users’ fees on automobiles. For example, the average federal and state tax per gallon of gasoline in the US was $0.43 in 2014, compared with $4.19 in Germany (Federal Highway Administration, 2016, pp. Table IN-1). The increase in users’ fees can be spent on eliminating the externalities of driving such as air pollution and traffic incidents (Litman 2019b, 20-23). At the same time, raised users’ fees on private transportation will make public transportation comparatively more attractive.

According to Alam, Nixon and Zhang (2015, 34) “the greater the supply, the greater the demand for transit.” There are two ways to increase the public transit supply: larger network, and more frequent rides. By extending the public transportation network, the number of people who can access the public transportation increases. Since transit supply drives the demand, it will increase the use of the public transportation network. The second method – increased frequency of the rides will improve the transit quality and increase the supply for people within the reach of the network. Public transit fares are also a significant factor, as high-income people will evaluate against the price of a private vehicle trip, while low-income people may choose not to travel, or consider slower but cheaper mode of transportation (bus compared to a rail), if the fare is too high for an individual (Reinhold, 2008). These instruments of control over the public
transit supply should be carefully used to regulate the number of travelers such that no line is under-used or overcrowded.

Since the stable growth of population in US cities, the number of potential passengers constantly increases. Therefore, the under-use is a rare concern for a city developer. However, an increasing number of passengers is a concern. For example, the New York (Ćosic, Šimunovic, & Šojat, 2017) subway system experiences significant delays due to a capacity constraint of stations and carts. This happens because despite an 80% increase of passengers since 1990, the New York subway did not increase the number of carts nor the tracks mileage (Fitzsimmons, Fessenden, & Lai, 2017). The new riders are packed into a system that has the same size as decades ago. The obvious solution would be to restructure the subway, by increasing frequency of rides or building additional lanes that would reduce a passenger flow on the most crowded routes. However, the other way to reduce a passenger flow in a subway is a more efficient bus system.

The data of 2016 shows a decrease in public transit ridership in almost every major city, except Houston and Seattle. Houston and Seattle reported the highest increase of 2.3% and 4.1% in their transit ridership (Schmitt, 2017). The two cities have in common a systematic redesign of a bus network in accordance with the public demand. Changes done to Houston transportation network did not require big monetary investments (Vock, 2017). Thorough research allowed to determine most and least popular routes in the city, as well as to determine the main directions of a passenger flow. This allowed to reconfigure bus routes to better serve demands of passengers. Unpopular routes were remodeled, cut or blended with some other unpopular routes which allowed to increase frequency of rides on the main passenger traffic directions (Vock, 2017).
Summarizing the information above, in order to increase public transportation ridership a government should introduce a coordinated package of mutually supportive policies that will cover all bases of public transportation issues: service quality, fare pricing, convenient ticketing, multimodal and regional integration, increased car taxes, etc. (Buehler & Pucher, 2012, p. 541)

In order to develop such policies, a thorough research is required. The research needs to determine the directions and routes of passengers at each time of day (Čosić, Šimunovic, & Šojat, 2017), explore the social and economic characteristics of passengers, introduce new ticketing options and explore passengers attitude towards different ticket types – by-rides or by-time, and so on. My research concentrates on exploring individual characteristics that influence public transportation use in the US cities.

3. Public Transportation Studies

Previous works in the field of travel mode choice were done by Chiou, Jou, and Yang (2015), Scott, George and Prybutok (2016), Haque et al. (2019), and Chen (2015). They concentrated their efforts on household decisions or statistics of a geographical area, I will focus on the individual decisions. This paper will formulate a model for the individual’s choice to use public transportation and empirically test the model. My findings support (Chiou, Jou, & Yang, 2015, p. 176) conclusion about factors having varying impacts on public transportation use in different geographical areas.

According to Scott, George and Prybutok (2016) and Mantouka et al. (2019) perceptions of service quality and beliefs about public transportation have great influence on the individual’s transportation mode choice. If a person who has never, or rarely uses public transportation happened to use a bus or a train, there are no other significant factors except his satisfaction with the ride, that will further affect his decisions to use public transportation (Scott, George, &
Prybutok, 2016, pp. 1062-1063). On the other hand, for people who use public transportation on a regular basis, all factors documented in 2016 by Scott, George and Prybutok had a significant impact on the choice of a transportation mode. Satisfaction with the ride is therefore an important factor of a travel utility. The travel utility is largely affected by a travel mode choice (Mantouka, et al., 2019, p. 8). It means that no other variable (weather, day of the week, trip purpose) has a significant impact on travel utility; individual’s choice between buses and cars is the only significant factor that determines the individual’s utility of a trip. When an individual uses public or private transportation, he updates the expectation about this transportation mode, if the expectation does not drop significantly, such individual will keep his travel mode choice behavior, given there is no other factor to change the behavior.

Heath and Gifford (2002, 2177-80) support the findings of the works above, regarding the economic dependency between the individual’s expectation of public transportation and its use. Heath and Gifford do it through the psychology case-study on a group of students in British Columbia. They find that the group of students who had never or rarely used public transportation had lower expectations of their initial rides in public transportation. During the study, students received monthly passes to the local bus system to encourage the use of public transportation. In the end of the month students, who did not like public transportation but used buses, changed their perception of public transportation and became more prone to use. This case study shows that there are people who do not have sufficient information about public transportation to form an opinion, and their low expectation of public transportation can be changed through exposure. However, the expectation factor impact can be different for different social and economic groups of people.

The preference factors of travel modes are evaluated based on perception of alternatives. The perception of alternative is influenced by measured variables and latent variables, including a
social position of an individual and his perception of socially appropriate behavior (Koppelman & Pas, 1980, pp. 27-28). The impact of these factors can be different for people in different social groups. The same result is suggested and confirmed by Chiou, Jou and Yang (2015). They show that there are variables that impact travel mode choice differently in different geographical regions of Taiwan. For example, some variables, like percentage of minors in an area, can influence public transportation mode choice positively or negatively in areas with different urbanicity characteristic. Public transportation usage increases with the percentage of minor population in rural areas and decreases in areas with high population density. Such result can be explained by the change of behavior of people once the characteristics of their residence change. This result is supported by Haque et al.’s (2019) research on British households that changed their place of residence. They find that people tend to change their transportation behavior when moving in an area with different transportation infrastructure. Households that move into areas with high public transportation availability tend to sell their private vehicles and use public transportation more often.

As you can see, availability of services, personal, regional and other latent variables are essential in understanding individual travel choice behavior. Another important variable is the service quality of public transportation. Service quality improves customers’ perception and expectation about the use of public transportation; added with the positive beliefs about public transportation, it makes people more likely to use it in the future (Chen, 2015). Chen and Li (2015) use rational utility theory, which considers people who use public transportation as consumers. The consumers pay for the transportation service, and therefore their decision-making behavior can be analyzed as a consumer choice behavior. Chen and Li (2015) use discrete choice model (DCM) to analyze consumer behavior regarding transportation mode choice. DCM explains consumer behavior as a consequence of individual’s preferences, with the
assumption that the consumer chooses the most preferred option. Chen and Li (2015) use structural equation model (SEM) to incorporate preferences and other latent variables as factors of a decision-making process. The resulting integrated SEM-DCM model fits the data better than the traditional logit model, and shows more significant impact of latent variables (expectations and beliefs) on the travel mode choice behavior. The works above show that any model of the transportation choice should account for a latent variables effect.

Despite a common belief about public transportation being a poor man’s alternative to a car, there are different social groups that use it more often than others. Women, young people, and singles use public transportation more often than do men, older people and families. Aside from these, people who are concerned about issues caused by private vehicles use public transportation more frequently (Steg, 2003, pp. 33-34). The household urbanicity factor seems to be the most definitive indicator of public transportation use (Anderson, 2016), as public transportation is rarely available in the rural areas of the US. Without accounting for any other factors, income and car ownership factors explain the decline in public transportation demand (Paulley, et al., 2006). However, Holmgren (2013) suggests that there is no evidence that the increase in income causes a decline in public transportation demand. According to Holmgren, a confusion regarding income variables effect comes from the omission of car ownership effect on the public transportation demand. Contradicting results regarding the influence of different factors may be explained by different data origins. It supports the result of (Chiou, Jou, & Yang, 2015) about opposite effect of some variables in different geographical regions. In order to understand how individual variables affect public transportation use, it makes sense to account for the local effects of each geographical area separately.
III. Theoretical Model

1. The Choice Problem

This paper models an individual’s transportation mode choice when traveling intermediate distances. Intermediate distance limits the trip radius to fifty miles from the start location of the trip. The model assumes an individual has already chosen to travel and has determined a destination location. Let us assume that there are no other choices that an individual can make except the choice for type of vehicle that will deliver him to the destination location. The individual can choose between two modes of transportation: public transportation and private transportation (personal vehicle). Each mode of transportation has its own costs and benefits that are associated with it. Public transportation is typically associated with a lower monetary cost, higher time cost and usually bears lower comfort benefit compared to a private vehicle. Most of these differences are due to a dedicated route of travel and a communal nature of services provided by public transportation. Personal vehicles usually bear higher monetary cost, but may provide faster and more convenient service, as personal vehicles go directly from the start to the destination locations and can pick up and drop passengers on a street. There are situations, when the time cost of a private vehicle would be comparable to a public transportation; this happens if a parking space is hard to access, or due to a traffic congestion. Such situations are typical for high density areas. The goal of a rational individual is to maximize his utility by choosing the vehicle type for the trip, given a certain budget constraint. In order to focus on the individual choice of travel mode this model will make the following simplifying assumptions.
4. Assumptions

1. Travelers are rational in making transportation mode choice and they will choose the mode yielding them the highest utility.

2. Individuals can choose between two modes of transportation: public and private (both modes of transportation are available).

5. Budget Constraint

A budget constraint of an individual is the combination of all possible goods and services for a given price that the individual may purchase given his savings and credit in the current period. Formula (1) demonstrates an equation of the budget constraint for all goods. On the left side of the equation is the individual’s income $Y$, on the right side of the equation is the price $P_t$ multiplied by the quantity $X_t$ of transportation services added with the price $P_{other}$ multiplied by the quantity $X_{other}$ of other goods. Increased price or amount of transportation services leads to a decrease in the amount of income dedicated for other goods and services.

$$Y = P_t * X_t + P_{other} * X_{other}$$ (1)

To simplify the transportation choice, the individual considers two modes of transportation: private or public. The two types of transportation have different prices, therefore, an individual is limited to different number of rides in each type of transportation. An individual with an income $Y$ can afford more trips on public transportation due to its lower cost compared to private vehicle trips. The impact of the transportation mode choice of an individual should be included in the formula of the budget constraint (1). Formula (2) illustrates the budget constraint of an individual with the income $Y$, who can choose between two modes of transportation. On the right side of the equation $P_{pt}$ and $X_{pt}$ represent the price and the
quantity of public transportation rides, $P_{pt}$ and $X_{pt}$ represent the price and the quantity of private vehicle rides.

$$Y = P_{pt} * X_{pt} + P_{pv} * X_{pv} + P_{other} * X_{other}$$ (2)

By choosing the number of public transportation ($pt$) and private vehicle ($pv$) trips for the current period, the individual chooses $X_{pt}$ and $X_{pv}$, thus limiting the amount of spending for other goods $X_{other}$.

A. Monetary costs

Price of public transportation is constant for all individuals. Every individual, regardless of their income or other factors, pays the same price for the same public transportation trip. This is not the case for the private vehicle trips.

The price of the private vehicle trip depends on the type of a vehicle. People that own a car receive a smaller price for choosing a private mode of transportation. If an individual does not own a car, his private transportation options are to use a rented car or a taxi service.

Consequently, the price of a private trip is a function of gasoline price, vehicle maintenance, rent cost, and driver’s salary, see formula (3).

$$P_{pv} = f_{full} \begin{cases} f_v (\text{gasoline price, maintenance}) & \text{if car is owned} \\ f_r (\text{gasoline price, rent}) & \text{if car is rented} \\ f_t (\text{gasoline price, driver’s salary}) & \text{if taxi} \end{cases}$$ (3)

Gasoline price is a function of the length of a trip. Cost of gasoline for the same trip is the same, regardless of whether a vehicle is owned, rented or a taxi. Vehicle maintenance includes insurance and engine repair. These depend on the vehicle cost, which is significantly correlated with the car owner’s income. On average the American household spends 15%-20% of its annual income on transportation costs associated with private transportation (Litman, Evaluating
Transportation Equity, Guidance For Incorporating Distributional Impacts in Transportation Planning, 2019, p. 24). High-income households use more expensive cars that require more expensive associated services, while low-income individuals rent and buy cheaper cars, with low associated costs. The driver’s salary variable is equal to zero if the individual is the driver on the trip. However, in case of a taxi service use, driver’s salary and gasoline price are the same for everyone, given the fixed trip locations. As a result, in this model, the price of a private trip is a function of a single variable – income, as it is the only variable that determines the cost of a vehicle purchase, rent, and maintenance. See formula (4).

\[ P_{pv} = f_{full} = \begin{cases} f_o(\text{maintenance}) & \text{if car is owned} \\ f_r(\text{maintenance, rent}) & \text{if car is rented} \\ f_t(\text{maintenance, driver's salary}) & \text{if taxi} \end{cases} = P_{pv}(Y) \]  

(4)

Let’s use formula (4) to update the budget constraint formula (2):

\[ Y = P_{pt} * X_{pt} + P_{pv}(Y) * X_{pv} + P_{other} * X_{other} \]  

(5)

B. Time Cost

Time cost of a transportation mode q is the individual perception of monetary losses associated with the longer commute time of one mode of transportation compared to the alternative mode of transportation. Time cost of a transportation mode can be evaluated as a function of an individual’s income that he could get if he had chosen an alternative mode of transportation.

Formula (6) explains the time cost of a transportation mode \( TQ_{a} \). On the right side of the equation \( TQ_{a} \) is the trip commute time in minutes via the intended mode of transportation, and \( TQ^{*} \) is the trip commute time in minutes via the alternative mode of transportation. The individual’s opportunity cost over time \( f_{oc}(Y) \) is the amount of money that he loses per minute of extra time spent on a trip, which is dependent on an individual’s income.
\[ TC_q = \begin{cases} (T_q - T_q^*) * f_{oc}(Y) & \text{if } T_q > T_q^* \\ 0 & \text{if } T_q < T_q^* \end{cases} \] (6)

An individual can select between two modes of transportation: public transportation \( pt \), and private vehicle \( pv \). Formula (6) reflects a monetary amount of time lost due to selecting a slower mode of transportation. In the case when public transportation is the intended mode of transportation, that is \( T_q = T_{pt} \); and it is faster than a private vehicle: \( T_{pt} < T_{pv} \), then the time cost of the trip is zero. In the case when public transportation is the intended mode of transportation, that is \( T_q = T_{pt} \); but it is slower than a private vehicle: \( T_{pt} > T_{pv} \), then the time cost of the trip will be the difference between both trip’s times \( T_{pt} - T_{pv} \) multiplied by the opportunity cost of time \( f_{oc}(Y) \) of an individual with income \( Y \). Using the new information let’s update the budget constraint formula (5) to include the time cost:

\[ Y - TC_q = P_{pt} * X_{pt} + P_{pv}(Y) * X_{pv} + P_{other} * X_{other} \] (7)

6. Trip Utility Function

The problem for an individual is to maximize his utility, by choosing between modes of transportation, given the budget constraint. Expected trip utility \( U \) of an individual using transportation mode \( q \) is a function of three groups of variables: individual \( \Gamma \), household \( \Delta \), and social \( \Lambda \). See formula (8).

\[ U_q = f(\Gamma, \Delta, \Lambda) \] (8)

A. Individual Variables

i. Trip Comfort

Comfort positively affects utility, which means more comfortable trips yield higher utility.

Individual perception of comfort of public and private vehicles can vary. Trip safety is one of the
factors of trip comfort. Public transportation is more secure compared to a private vehicle, as buses and trains are 90% less likely to get into an accident than cars (American Public Transportation Association, 2018). But private vehicles are on average more comfortable compared to public transportation vehicles due to a higher control over one’s surroundings. However, there are other issues that influence the expected trip comfort variable. Both modes of transportation have an accessibility issue which reduces trip comfort. Public transportation stops and parking lots can be located far from the trip start or destination locations. This creates an additional cost in the form of an accessibility issue. The accessibility cost of a private vehicle is expected to be lower on average due to a large amount of parking spaces in urban and rural areas. Consequently, in this case, private transportation yields lower comfort costs as compared to public transportation.

ii. Income

Diminishing marginal utility of income suggests that as income increases, an individual’s utility from income grows at a decreasing rate. It means, utility difference between prices of $5 and $25 is smaller for an individual with an annual income of $100,000 than for an individual with an annual income of $50,000. As a result, a high-income individual faces a smaller utilitarian cost from a more expensive mode of travel, as a result receiving higher trip utility from private transportation as compared to a low-income individual.

The comfort of a trip is a factor of individual income. High-income individuals drive more expensive cars as compared to low-income people; on average, car quality increases with individual income. Higher quality of the vehicle contributes to higher comfort, which yields higher utility of a trip. In addition, high-income people may receive lower utility from public transportation trips due to a shifted perception of comfort as compared to low-income people. This may happen as high-income people usually enjoy a privilege of controlling their
environment, thus the inability to do so in public transportation is an additional comfort cost. Summarizing the above, high-income people face higher utility cost when using public transportation compared to low-income people.

The opposite effect of income can be observed by considering the municipal funding of an individual’s household area. Wealthy households are more likely to appear in wealthy neighborhoods, and low-income people are likely to buy a house in a low-income neighborhood (Fry & Taylor, 2012). The result of the income segregation of districts is an uneven distribution of municipal services. High-income districts, and people who live there, enjoy a better quality of municipal services due to higher tax base. Municipal services include an access to public transportation. Consequently, the individual who lives in a high-income area has a better supply of public transportation, and therefore faces lower accessibility costs as compared to low-income people in other districts.

iii. Employment Status

Employment contributes to the opportunity cost of time function $f_{op}(Y)$ in formula (6). If an individual is employed, his opportunity cost function produces higher value compared to an unemployed individual. Therefore, an employed individual is expected to face higher utility cost from public transportation trips.

An employed individual often has a stable home-work-home route of travel, which means that such an individual can plan the time frame of his home to work and work to home trips. Scheduled trips are easier to take using an appropriate public transportation route. An employed individual is expected to face lower comfort cost from scheduled public transportation trips as compared to other public transportation trips. As a result, the individual may receive higher utility from public transportation on a home-work-home route.
iv. **Education**

A higher education level may imply higher awareness of the transportation issues, public transportation benefits, and social costs associated with personal vehicle use. High education therefore increases the utility of public transportation rides and decreases the utility of private vehicle trips.

Education has strong correlations with income and employment status. I expect the correlation between education and income to diminish the significance of awareness factors.

v. **Driver status**

An individual with the driver’s license faces lower costs by choosing a private mode of transportation. A licensed driver can always be a driver on a trip, while a car can still be rented or owned. Such individual uses taxi services less often compared to non-drivers. Consequently, in most cases the private mode of transportation is cheaper for drivers and therefore bears higher utility for them.

vi. **Vehicle ownership**

Private vehicle owners are expected to receive higher utility from private transportation. Private vehicle users receive higher utility from a car trip as compared to people who hardly or never travel by car (Steg, 2003, pp. 30-34). It makes sense to suggest that car owners receive higher utility from a private vehicle trip, as compared to people who do not own a car.

People who often or always travel by car have a more negative perception of public transportation as compared to people who hardly ever or never travel by car (Steg, 2003, pp. 33-34). The same bias is likely to be spread among car owners. They are expected to receive lower utility from a public transportation trip, as compared to people who do not own a car.
vii. Age

Older individuals are expected to value comfort over monetary and time costs. Due to the higher comfort of private vehicles, older people are expected to receive higher utility from private transportation trips. However, age also has a diminishing marginal return. Older drivers are less cognitively coherent, and – as such – are worse drivers. Consequently, people who use a car because they like the act of driving are likely to receive less utility from private vehicles as they become older.

B. Household Variables

i. Size of a Household

The size of a household directly influences individual’s income and savings. A new dependent member of a household would decrease the average individual income, while a new employed member of a household can increase average individual income. For example, if the household size increases due to a child, the individual income will decrease, as a child cannot earn income; if the household size increases due to a marriage, the individual income is likely to increase, as a new working member of a household is likely to bring more money to the household. Therefore, a change of a household size effect can be similar to an increase or decrease of the individual income.

Apart from the direct influence, a change in the household size can also change travel behavior of an individual by changing his attitudes and beliefs. For example, a new child would provoke a protective behavior from the parents. Even if a parent is alone on a trip, the security concern would be higher than for a childless individual. As a result, such people are expected to face higher comfort costs from public transportation due to an uncontrollable environment. On the other side, if a new household member holds strong beliefs against a certain mode of
transportation, it may change the attitude of all household members. In case of a multiple member household, it would make sense to take into account the beliefs of all members of a household when considering a public transportation choice. Consequently, the effect of a household size on travel utility is hard to quantitatively evaluate.

ii. **Age of the Youngest Child in a Household**

The direct influence of the age of the youngest child on an individual travel utility is ambiguous. According to the 2015 data, monetary cost of raising a child increases with the child’s age only for households with high income (Lino, Kuczynski, Rodriguez, & Schap, 2017, pp. 10-12). Households with annual income less than $100,000 rarely experience an increase in family expenditures from a child between ages zero to seventeen. It follows, that the age of the youngest child is not expected to influence an individual travel utility through the monetary means.

However, the age of the youngest child influences its parents’ utilitarian costs. Parents of an infant are much more concerned about safety issues than parents of a teenager. Parents of an infant are expected to be more concerned about safety issues even when traveling alone, as they have a higher parental responsibility compared to a teenager’s parents. The safety concern is a part of a comfort variable. It would make sense to expect people who are more concerned about safety to receive higher utility form public transportation, as you are 90% less likely to get into an accident in a public transportation vehicle than in a car (American Public Transportation Association, 2018). Real life implication of this principle may contradict the theoretical model, as some people believe that cars are safer than buses; or they are more concerned about children’s environment which is easier to control in private transportation vehicles.
C. **Social variables**

i. **Social Stigma**

Social stigma is a factor that prevents people from deviating from socially expected behavior. As such, high-income people are expected to prioritize time and comfort over monetary cost, and low-income people—vice versa. A deviation from socially expected behavior may cause interruptions in social relations. For example, in a society where most wealthy people use cars, it is a part of a wealthy individual’s social image to drive a private vehicle. In this society, a wealthy individual who uses a bus may receive a negative social consequence in a form of various assumptions, such as that the individual pretends to be wealthy while he cannot afford a car. These assumptions are a utilitarian cost associated with the public transportation use. The same applies to other variables and their combinations, as there always is a social cost associated with the choice that contradicts social expectations. As a result, people receive lower utility from travel mode choice that contradicts their social image.

ii. **Population Density and Urbanicity Status of a Household**

Due to the fact that the household location is the most popular trip start location, the population density and urbanicity status of a household should have a significant impact on a trip utilitarian cost. Population density is strongly correlated with the public transportation accessibility, which is why people in high density areas are expected to face lower utilitarian costs from public transportation trips compared to people in low density areas, due to higher rides frequency and easier accessibility (Haque, Choudhury, Hess, & Crastes dit Sourd, 2019). The population density variable interacts with the age of a youngest child variable. A childless individual is more tolerant of lower ride frequency of public transportation, which is typical for low-density areas, compared to an individual who has children. Parents and guardians are
expected to value time with kids more than regular leisure. It means that time spent waiting for a bus bears higher utility cost for guardians than for childless people. Summarizing, guardians in high-density areas are expected to receive higher utility from public transportation travel compared to guardians in low-density areas.

iii. Heavy Rail System

A heavy rail system increases the utility of public transportation trips as it increases the network effect of the system by attaching additional locations. Rails offer faster transportation to distant locations compared to bus travel. A trip that may require two or more bus transfers may be covered in one train ride. Heavy rail decreases time costs of public transportation and therefore increases the utility of a public transportation trip. People who live in an area with a heavy rail system are expected to receive higher utility from public transportation compared to people with no access to the rail system.

D. Trip Utility Equation

Utility of a trip is a factor of three groups of variables that are described above. Some variables above are linearly codependent. Therefore, the utility $U$ of an individual given transportation mode $q$ is a linear utility function $f_{iu}$ of all variables from these groups: individual $\Gamma$, household $\Delta$, and social $\Lambda$, see formula (9).

$$U_q = f_{iu}(\Gamma, \Delta, \Lambda) \quad (9)$$

7. Decision Model

The individual chooses a mode of transportation that will maximize his utility subject to his budget constraint. The theoretical optimization of this choice yields the individual demand equations for three goods: public transportation, private transportation, and other goods. Each
demand equation in reduced form is a function of individual, household, and social variables. An individual demand $D_g$, see formula (10), of an individual for the type of good $g$ is a function $f_g$ subject to variables from the three groups mentioned above

$$D_g = f_g(I', \Delta, A)$$  \hspace{1cm} (10)

In the formula (10), type of goods $g$ are public transportation, private transportation, and other goods.

The combination of demands for public and private transportation yields to an individual a number of public and private transportation trips that the individual makes in the current period. Using these numbers, a fraction of public transportation trips from all trips can be calculated.

IV. Empirical Model

1. Data Description

A. The NHTS 2009 Dataset

The National Household Travel Survey (U.S. Department of Transportation, Federal Highway Administration, 2009) is a repeated phone-survey. The NHTS collects data from the civilian, non-institutionalized population of the US. The design of the survey is the following. Each selected phone number is assigned a specific date as a “travel day.” The respondent receives a call, and completes a recruitment interview. Follow that, the respondent is assigned a travel day on which he or she records information about all travel taken on the travel day. The respondent then receives follow-up call and provides the information to the NHTS. The data collecting
period for the 2009 dataset was March 28, 2008 to April 30, 2009. Travel day dates were assigned in those date boundaries to all seven days of the week.

The NHTS data is provided in three different data files – a personal file, a household file, and a travel/trip file. The first dataset provides information about respondents (travelers). It contains data on age, education, whether the person is a licensed driver, and employment status. The second dataset contains household data: income, household size, race of a respondent, vehicle count, age of the youngest child, state of origin, characteristics of geographic household location, etc. To analyze this information, it needs to be extracted and combined into a single file. The research unit is a respondent (traveler), so each variable will be a value for a single person. The third dataset provides information about trips that a member of a household took on the travel day, for which they had to complete their travel day diary. The trips dataset has trip information from the travel day diary about the specifics of the trips. It includes the type of a vehicle used, the purpose of the trip, miles traveled, time of the trip, day of the week, etc.

Due to the theoretical assumptions, the empirical research will only consider respondents who have access to two modes of transportation. Therefore, the population sample is limited to respondents whose age is 16 years or older, who live in areas where at least one respondent has reported to use public transportation for the past month. This is done to exclude respondents who live in areas with no access to public transportation.

B. Variables

Variables that hold personal information of respondents are self-reported. Variables that hold characteristics of geographic household location are derived by the survey designers. For the summary statistics of all available data, see Table 1.
i. Dependent Variable

The dependent variable of the research is the fraction of public transportation trips to all trips taken on the travel day by a respondent. The variable is measured in a percentage ranging from 0 to 100. For example, a respondent who traveled five times during the travel day, and used public transportation on one of the trips, is coded as a 20 meaning that 20 percent of the respondent’s trips included public transportation.

ii. Independent Variables

The independent variables of the research represent individual, household, and regional characteristics of a respondent.

1. Individual Variables - $I'$

The respondents age is a continuous variable that reports the full number of years of the respondent’s age. It holds integer values between 16 and 92. According to the theoretical model, an increase in a respondent’s age is expected to raise the utility of private transportation and decrease the utility of public transportation. Consequently, an increase in a respondent’s age is expected to be negatively correlated with the dependent variable.

The square of the respondent’s age is a continuous variable measured in integer values and is equal to the arithmetic square of the respondent’s age. This variable is used to determine any non-linear effects associated with age. Since the effect of age is not expected to be linear, the square of the respondent’s age variable checks for a quadratic dependence between age and the dependent variable.

The driver status of a respondent is a binary variable which equals 1 if a respondent reports to be a licensed driver, and equals 0 if the respondent is not a licensed driver. Intuitively, licensed drivers are predicted to be less likely to use public transportation compared to non-drivers.
The employment status of a respondent is a binary variable which equals 1 if a respondent reports to be currently employed, including self-employment. The variable equals 0 if a respondent reports to be unemployed. Due to an equivocal theoretical impact of employment on trip’s utility, an impact of the employment status on the dependent variable is expected to be low or not significant.

The respondent’s education is a group of binary variables that show the highest education grade completed by a respondent, including no high school education, high school degree (and no college degree), associate degree, bachelor degree, and graduate or professional degree.

According to the theoretical estimation, higher education level of a respondent is expected to raise his social and ecological awareness, on the other side higher education level is correlated with higher standards of living, which include private vehicle usage. Therefore, the impact of education is expected to be equivocally related to the dependent variable.

The respondent’s income is a group of binary variables that show a derived individual income of a respondent. According to the theoretical model, the transportation decision is made on an individual level, therefore the research uses the individual income of a respondent instead of a household income. The respondent’s income $Y$ is derived from the household income $Y_{HH}$ in order to account for possible differences in decision making of people from a single member household of a certain income, and people from a multiple member household with the same income. The respondent’s income $Y$ is derived using a modified OECD equivalence scale (OECD) – an instrument that assigns income to household members using a certain proportion. I use the modified scale with the following proportions: a household head has a value of 1, each additional member of a household has a value of 0.5. For example, in a household of three, with total annual household income of $Y_{HH}$, each member will be assigned an individual income of:

$$
Y = \frac{Y_{HH}}{1 + 0.5 + 0.5} \tag{11}
$$
In formula (11) the denominator is a sum of values assigned to each member of a household. The household head has the value of 1, while two other members have the value 0.5. The individual income is split into three groups by the annual income amount: low – $0 to $34,999, middle – $35,000 to $79,999, high – $80,000 and more. Since I use the equivalence scale to calculate the respondent’s income from a household income, all effects of other variables on a household income will be presented in the respondent’s income variable effects. For example, a monetary effect of an additional household member (positive or negative) will be presented in the respondent’s income variable effects. According to the theoretical estimations on income and other variables that affect it (education, size of a household, age of a youngest child, etc.), income is expected to be negatively related with the dependent variable.

2. Household Variables - $\Delta$

Household size is a continuous integer variable, which represents the number of household members as reported by the respondent. According to the theoretical estimations, household size is expected to have a low or insignificant impact on the dependent variable.

Status of the youngest member of a household is a group of binary variables that reflect the status of underage members in the respondent’s household. The group includes the following variable: no underage members, some underage members of age zero to five, and some number of underage members of age six to twenty-one. According to the theoretical estimations, households with no children are expected to take more public transportation than households with children, due to a lower utility cost of a waiting time. The fraction of public transportation trips in households with an infant is expected to be the same as in households with an older child due to contradicting influences of security (better for public transport) and control over the surrounding (better for private transport).
The race of a respondent is a group of binary variables that best describe the race of the respondent, according to the respondent. The possible responses are coded as: white, black, Hispanic (or Mexican origin), or other races and ethnicities. Race is not expected to have a significant impact on the dependent variable.

Relation of the number of vehicles to the number of drivers in a household is a group of binary variables. Only one of these variables can be true per respondent. The variables are: no vehicles in a household, the number of vehicles is smaller than the number of drivers, and the number of vehicles is equal to or larger than the number of drivers. According to the theoretical estimations, individuals who do not own a vehicle are expected to have significantly more public transportation trips than individuals who own a vehicle, due to a significantly lower monetary cost of public transportation. Individuals who live in households with a high number of vehicles are expected to use public transportation significantly less often than everybody else, as they typically get more utility from a private vehicle trip due to a lower cost and higher comfort of private vehicle trips.

3. Regional Variables - \( A \)

Population density in a household's area is a group of binary variables that represent the number of people per square mile that live in the household's block group. Only one variable of this group holds a true value. The variables are: low population density – 0 to 3,999 people per square mile, medium population density – 4,000 to 9,999 people per square mile, and high population density – 10,000 to 999,999 people per square mile. According to the theoretical estimation, population density is expected to be positively and significantly correlated with the dependent variable, due to a higher accessibility of public transportation in higher density areas.

The heavy rail status of a household is a binary variable that equals 1 if an area where a household is situated has a heavy rail transportation system. The variable equals 0 otherwise.
According to the theoretical estimation, the heavy rail variable is expected to be positively and significantly related with the dependent variable.

The home address in urbanized area variable. It is a binary variable which equals 1 if a respondent’s house is in an urbanized area. The variable equals 0 if a respondent’s house is in a rural or a sub-urban area. According to the theoretical estimation, the urbanized status of a household is expected to be positively and significantly correlated with the dependent variable.

The states variable is a group of binary variables that equals 1 if a household is located in a specific state. All other variables in this group for a unit of research equal 0. Among all states I select seven with subjectively better public transportation systems. See Table 2 for the description of the seven states with good public transportation. According to the theoretical estimation, individuals who live in states with a good public transportation system are expected to use public transportation more often compared to individuals in other states. The group of state variables controls for unobservable effects of a state geographic location, local transportation policies, and regional preferences on the fraction of trips taken on public transportation over the total number of trips.

C. Data Issues

The NHTS 2009 contains data from both complete and incomplete diaries. Out of 308,000 participants, only 223,000 were associated with complete data for the travel day. However, of the 85,000 observations with missing data, some might contain all of the data required in this analysis. Therefore, this study uses observations with all data completion status. Regardless of this selection, the total number of observations that fit to the regression parameters is 218,776. This number is smaller than the number of observations associated with complete data due to omitting the observations with missing values. This presents a couple of selection biases.
The complete or incomplete diary selection bias is not important if not completing the diary is determined randomly, because the survey designers created survey weights to adjustment for each respondent’s representation in the general population. Thus, as long as the omitted observations are randomly determined, the regression results, although less efficient due to a lower sample size, will remain unbiased.

The second selection bias is the omittance of research units that have missing variable’s values. People with complete and incomplete diaries are omitted from the regression if they selected an “appropriate skip”, “don’t know”, “not ascertained,” or “refused” options as an answer for the independent variables of the research. This is a selection bias as it means that the research will analyze data from only those people who shared the required information. It means, that the research is likely to miss a group of people who do not want to share their information for various reasons. At the same time, the research collect data from people who own a dedicated phone-number. Further, the reader should take this information into account.

D. Survey Data Weights

The unit of the research is a respondent (traveler). The primary component of the survey’s base weights is the inverse probability of selecting a household telephone number from the sample group frame. The survey interviewed each of the eligible individuals within each of the selected households. In terms of sampling, it means that every individual in the household had a probability of selection equal to the probability of selecting the household. Then, each individual’s base weight is equal to the base weight of the household (Rizzo, et al., 2011). In a single member household, an individual adjusted weight is equal to the weight of the household. In a household with multiple members, a different method was used to calculate adjusted weights of respondents from the same household.
Multiple member households were defined as useable if at least half of the eligible adults completed the survey. Within useable households were members who did not answer all questions in the survey. The weights of the responding adults needed to be adjusted to account for the nonresponding adults. The survey designers used the following approach: to generate cells across useable households, so that a particular nonresponding adult is adjusted for in the weights of a number of responding adults across a number of useable households. A non-responded adult was adjusted within a pool of responded adults who have many of the same characteristics (sex, age, driver status, etc.) (Rizzo, et al., 2011).

As a result of the survey’s design, it follows that the individual adjusted weight of a respondent is designed to be the inverse of the probability that the respondent is included in the dataset because of the sampling design. It happened because the survey designers selected certain telephone numbers from each state and assigned each member in these households an individual weight. The individual weight is roughly equal to the inverse probability of the person’s chance to take the survey. In Stata program, which I use to analyze the data, the weight option for the regression of such data is probability-weights – “pweight.”

8. Model Description

The following empirical model is designed to explore the effects of observable individual variables on the fraction of public transportation trips to all trips (the dependent variable). Observable individual variables are grouped into individual, household, and regional categories that are described above.

Interdependent relations between observable and unobservable variables, such as the dependence between car ownership and of public transportation perception (Steg, 2003, pp. 33-
may impact the value of a variable’s coefficient. The coefficient estimator of a variable may reflect the impact of the variables itself combined with a related unobservable variable. Despite the model containing personal, household, and trip characteristics, unobserved heterogeneity may still be a concern. Given the comprehensive nature of the dataset, the most likely source of this heterogeneity is geography. Large cities vary in climate, government spending at both the state and local level, industrial production, etc. Therefore, in order to control for unobserved heterogeneity at the state level, a complete set of state fixed effects is included in the model.

Due to the theoretical estimation on the demand for public transportation, private transportation, and other goods, the dependent variable $D_{pt}$ is a function $f_{pt}$ of the following variables:

$$D_{pt} = f_{pt}(Age, Square \; Age, Driver \; status, Employment \; status, Highest \; education \; level, Income, Household \; size, Age \; of \; the \; youngest \; child, Race, Number \; of \; vehicles, Population \; Density, Heavy \; Rail \; status, Urbanicity \; status, State \; effects)$$

According to the theoretical model, the travel mode choice process might be different for people with different possible choices available to them. For example, people in cities with low public transportation supply are likely to act differently from people in areas where public transportation supply is high. In order to explore the difference in travel mode choice behavior of the two population groups, I select two population samples. The first sample includes people living in areas with high public transportation supply, while the second sample includes people living in areas with low public transportation supply. Unfortunately, the NHTS does not report which city people live in, nor does it collect direct information on the supply of public
transportation. However, there are two methods of selecting such samples. One method would be to differentiate population based on the respondent’s household area urbanicity status and population density characteristics; such that one sample includes only people in high-density urban areas and the other sample includes people who live in lower density urban or rural areas. This method is good to explore the difference between urban and rural population. Indeed, people who live in suburbs act differently from people in commuter towns who live in flats. However, there are plenty cities, towns and suburbs in the US that have good, bad or absent public transportation systems. Therefore, this method is inapplicable for the country-level research. Instead, I am going to select population samples based on states public transportation characteristics. To address this problem, I assign anyone who lives in one of the seven states listed in Table 2 as living in a state with a high supply of public transportation as these seven states are known for their public transportation infrastructure and investments. Anyone not living in these seven states is conversely grouped as being in a low public transportation supply state.

A. States with a High Supply of Public Transportation Options (First Sample)

This sample includes people who live in states with good public transportation systems. This will allow the research to explore the travel mode choice behavior of people, who are most likely to have access to both transportation modes. A respondent’s data from this sample will be used to regress the fraction of public transportation trips over the total number of trips against individual characteristics of a respondent. See Table 3 for the summary statistics of the first sample data.
B. States with a Low Supply of Public Transportation Options (Second Sample)

The travel mode choice behavior of people who live in absence of easily accessible public transportation is expected to be different from people who have easy access to public transportation. The obvious differences are the increased accessibility cost and reduced network effect. However, there are good reasons to suspect that a low supply of public transportation changes the evaluation criteria of the travel mode choice. Comparing the coefficient estimates from the two datasets will also allow for insight into which regional variables have an important impact on travel mode choice behavior. See Table 3 for the summary statistics of the second sample data.

C. Total Population Sample

When using the entire NTHS sample, the regression results apply “on average” to the entire US population, given the selection criteria described above. This will allow to better understand the relationship between the explanatory variables and public transportation usage based on access to public transportation. The first sample and the second sample are separately used to estimate the model. A Chow test will be used to check for structural differences in travel mode choice behavior between people who live in states with high public transportation access and the rest of the US. Comparing the coefficient estimates from the two datasets will also allow for insight into how different supply of public transportation impacts travel mode choice behavior. See Table 3 for the summary statistics of all three samples.
V. Results

Table 4 reports the estimated coefficients from the linear regression model for all three samples. transportation

1. Regression Results from States with a High Supply of Public Transportation

Age and squared age variables are statistically significant predictors of using public transportation. Public transportation ridership is decreasing at a decreasing rate up to the age of about 60-year-old after which ridership on public transportation starts increasing. This means that the group of 55- to 65-year-old people uses public transportation less than any other age group. This is consistent with (Steg, 2003, p. 33) research which states that younger people use public transportation more often than older people.

Driver status, employment status, and education variables do not have a significant impact on public transportation usage.

While theoretically important, income is not a statistically significant predictor of using public transportation. Therefore, there is no evidence of a significant impact of higher income on the dependent variable, holding other variables constant. Since income and education measure similar characteristics of an individual, it might be that education picked up all the effects of income for the regression. However, education variables are also not statistically significant in states with a high supply of public transportation.

Household size has a statistically significant impact on the public transportation usage. For each additional member in the household, an individual’s fraction of trips taken on public transportation over the total number of trips decreases by 0.96 percentage points. According to
the theoretical estimations, it may imply that people in multiple member households value the control over the surrounding (which is better in a private vehicle) more than people in single member households.

Race was not expected to have a significant impact on using public transportation. The regression results, however, do show a certain degree of racial bias. A black respondent is predicted to be 5.1 percentage points more likely to use public transportation compared to a white respondent. This result is significant at 1% level. Hispanic respondents and people who report another race do not show a significant difference in the fraction of trips taken on public transportation over the total number of trips compared to white respondents.

The relative number of vehicles to drivers variables is also statistically related to the dependent variable. A person in a household with at least one vehicle, but with more drivers than cars, is 27.6 percentage points less likely to take public transportation than is a person in a household with no cars. Similarly, a person in a household with at least as many cars as drivers, is 35.2 percentage points less likely to take public transportation than is a person in a household with no cars. The result proves the theoretical estimation on the impact of personal vehicle availability on the public transportation demand.

Population density is a strong predictor of public transportation usage. People who live in areas of medium population density are 1.1 percentage point more likely to use public transportation than people who live low a population density area. This result is significant at 5% level. People who live in areas of high population density are 4.9 percentage points more likely to use public transportation than people who live low a population density area. This result is significant at 1% level. This may be a consequence of a positive correlation between population density and public transportation supply; people who live in areas with higher public transportation supply
use it more often. It also may be caused by an increased cost of private transportation usage in high density areas, such as a more expensive and less accessible parking.

The heavy rail availability variable is positively and significantly (at 1% level) correlated with the dependent variable. Availability of heavy rail system in a respondent’s household area increases the usage of public transportation by 2 percentage points compared to a respondent who lives in an area without heavy rail transportation system. This result was expected and can be explained by an increased availability of public transportation for people in the area. The second explanation is through the decreased time cost of distant trips. When considering mode of transportation for a distant trip people in areas without a rail people choose between a bus and a private vehicle; people in areas with a rail people choose between a train, a bus and a private vehicle. Consequently, some people are likely to choose rail over car, which raises the fraction of trips taken on public transportation over the total number of trips.

The urbanicity status of a household variable has a positive non-significant effect on the dependent variable. It shows that the supply and demand for public transportation does not depend on a de jure status of a household area. Population density, and heavy rail availability are much better characteristics to explain why people use more or less public transportation.

States variables in the regression account for various unobservable regional variables: weather, local social norms, and transportation policies. The estimated coefficients and their significance indicate that there is an impact of state level unobservable variables on the travel mode choice behavior. But evaluation of these factors is not a goal of this research, that is why Table 4 does not show the coefficient estimators for states fixed effects.
2. Regression Results from States with Low Public Transportation Supply

Age variable is a statistically significant (at 5% level) predictor of using public transportation. The estimator on the square of age is not significant which indicates a linear dependence between age and public transportation usage. Public transportation usage by people in the states with low public transportation supply is decreasing at a constant rate of 0.1 percentage point per year of age. Still it is important to notice that the negative coefficient of age in the second sample regression holds twice lower absolute value than in the first sample regression. The absolute effect of age on the dependent variable in the second sample data compared to the first sample data is lower for people that are younger than 65 years old.

Driver status is statistically significant at 1% level. Drivers are 7 percentage points more likely to use public transportation than non-drivers. The regression on the first sample data showed this variable to be not significant. It indicates that drivers among the first sample population have similar travel mode choice behavior to not drivers, while in the second sample driver status makes a significant impact on the travel mode choice behavior. It may happen due to a significantly higher accessibility and time cost of public transportation in the US outside the states with good and developed public transportation systems.

All coefficient estimators of education variables are significant at 1% level. Compared to people with no high school education: high school graduates are 3 percentage points less likely to use public transportation, college or associate degree graduates are 3 percentage points less likely to use public transportation, bachelor degree graduates are 3.3 percentage points less likely to use public transportation, and graduate or professional degree graduates are 2.8 percentage points less likely to use public transportation. It follows that people with the graduate or
professional degree use public transportation more than those who only attended high school, college, or got a bachelor degree. It may happen due to a strong income-education correlation, or it may mean that people who completed graduate education live in areas with good public transportation networks compared to people who got an associate degree.

Household size has barely significant (at 10% level) correlation with the dependent variable. Every additional member of a household is expected to decrease the public transportation usage by 0.26 percentage points.

The regression results show a certain degree of racial bias. A black respondent is predicted to be 2 percentage points more likely to use public transportation compared to a white respondent. This result is significant at 1% level. Hispanic respondents and people who report another race do not show a significant difference in the fraction of trips taken on public transportation over the total number of trips compared to white respondents.

The relative number of vehicles to drivers variables is statistically (at 1% level) related to the dependent variable. A person in a household with at least one vehicle, but with more drivers than cars, is 26.2 percentage points less likely to take public transportation than is a person in a household with no cars. Similarly, a person in a household with at least as many cars as drivers, is 28.7 percentage points less likely to take public transportation than is a person in a household with no cars. These results are similar to the first sample data.

Population density variable is a less significant predictor of public transportation usage in states with lower public transportation supply. There is no significant difference in public transportation usage of people who live in low population density areas compared to people who live in medium population density areas. People who live in areas of high population density are 5.6 percentage points more likely to use public transportation than people who live low a population density area. This result is significant at 1% level.
The heavy rail availability variable is positively and significantly (at 5% level) correlated with the dependent variable. Availability of heavy rail system in a respondent’s household area increases the usage of public transportation by 0.9 percentage points compared to a respondent who lives in an area without heavy rail transportation system.

3. The Chow Test

As expected, usage of public transportation is greater in high supply states compared to low supply states. This is most readily seen in Table 3, where respondents in high supply states use public transportation on 3.5% of trips whereas respondents in low supply states use public transportation on 1.4% of public transportation trips.

Given this difference, it follows that there can be a difference in the travel mode choice making process. In order to check for structural differences in the decision-making process for the population in the two samples, a Chow test will demonstrate if the independent variables have different impacts on the two subgroups of the US population. The null hypothesis for the Chow test is that the coefficients in the two sub-sample regressions are equal. The Chow test statistic follows the F distribution with \( k \) degrees of freedom in the numerator and \( n - 2k \) degrees of freedom in the denominator, where \( k \) is the total number of parameters (in this case, 73) and \( n \) is the number of observations in the total sample. Let \( RSS_t \) is the residual sum of squares of the linear regression for the total population sample, and let \( RSS_1 \) and \( RSS_2 \) be the residual sums of squares of the linear regression for the two sub-samples. Then, the Chow statistic is calculated as:

\[
Chow = \frac{RSS_t - RSS_1 - RSS_2}{RSS_1 + RSS_2} \frac{1}{n-2k} \tag{13}
\]
The calculation of the Chow test statistics yields the $F(73, 218776) = 16.7$, and is associated with a $p$-value of under 1% meaning that there is significant evidence for the rejection of the null hypothesis. Consequently, there is strong evidence for a structural difference in transportation mode choice between the two samples of data. Because there is a structural difference in the decision-making process between high supply and low supply states, relevant policy proscriptions may also differ or vary in their effectiveness across the high supply and low supply states.

VI. Policy Implications

The following section presents interpretations of the results of the empirical analysis and uses the results to identify possible policy implications. Since public transportation appears to be socially better than private transportation, the policies will aim to increase the use of public mode of transportation. As mentioned earlier, public transportation is more efficient in the long run because it carries lower social costs (health, space, fuel) compared to private transportation. That is why it is beneficial for a society to encourage its members to use public transportation. Because the potential policy impact will be greater in areas with relatively high public transportation supply, the discussed policies stem from the results of the high supply states.

1. Targeting Increased Ridership in High Supply States

Driver status, employment, education, income variables have no significant impact on the fraction of public transportation rides in high supply states. It is important to notice that the impact these variables might be different if to select a more specific population set, either by restricting it to a certain age range, social status, or geographic area. That said, the evidence suggests that policies focussed at these characteristics will be ineffective.
However, population density variable is significant and has a positive impact on public transportation demand. This means that, on average, people who live in higher population density areas tend to use more public transportation. There are two ways to interpret this result.

First, a high population density area is somehow related to an increase in public transportation demand. For example, high density areas are typically associated with high cost of parking, which increases private transportation cost. This makes public transportation comparatively more attractive in high density areas. The area with high demand of public transportation is likely to have a typical look of a high-demand market – a deficit of goods – insufficient supply. If an area has overcrowded bus or train lines, it is likely that economic reasons to take public transportation outweigh possible issues of comfort. Such area needs a restructuring or an expansion of its public transportation system.

Second, a high population density area has developed public transportation systems, and people who live there enjoy a sufficient supply of public transportation. This area type is less likely to have overcrowded buses, which means demand meets an appropriate supply level. I suggest no further policies in such areas. However, one of the major qualities of public transportation system is its network coverage. If there is a single bus-route in a city, it will only transfer passengers locally. In this case, an underuse of public transportation can occur when public transportation coverage is so small that there is no point in using it due to a low number of possible destinations; or when the rest of the public transportation system is overcrowded.

Thus, the matter of public transportation should be planned out on a largest possible scale. Local bus routes should cover small districts and neighborhoods, leading to a bigger bus, train or subway lines that will transfer people to various places in a city. These are general ideas; the main question of such policies often lies in sources of funding.
The income variable shows a statistically insignificant effect in the regression, which means that an additional private vehicles gas tax is unlikely to change the travel mode choice behavior of a general consumer – traveler. There is not enough evidence to suggest that an increase of monetary costs of private transportation would cause a substitution effect, and people would use more public transportation. However, the money can be allocated to policies which would produce the desired substitution effect. These policies include fixing old or constructing new public transportation systems, thus increasing the supply. Suggestion of Alam, Nixon and Zhang (2015, 34) “the greater the supply, the greater the demand for transit” is likely to work in areas with low or absent public transportation, such as neighborhoods and towns without or with only local public transportation system. People in such areas rely on their cars to get to the state capital or other social or economic hubs. If an additional public transportation lane is introduced from such areas to a commuter rail or to a larger city, people are likely to start using it. At the same time, a new public transportation lane creates additional customers for local businesses, due to an increase in the number of local tourists.

Racial characteristics are statistically significant predictor of public transportation rides. Being African American increases the fraction of public transportation rides. Due to a degree of racial segregation of housing districts in the US, it is impossible to unequivocally interpret this result. There can be various reasons for people, who identify as African-American, tend to use public transportation significantly more often than people who identify as White/Caucasian. Possible reasons can be higher supply of public transportation in districts that are predominantly inhabited by people who associate themselves with Black race; or there is some other shared characteristic that makes African Americans use more public transportation. This property is unlikely to relate to economic reasons or household size since these characteristics are accounted in the regression. Therefore, further research is required.
Based on the previous results and their interpretations I designed an incentive program that is likely to increase public transportation usage and reduce vehicle occupancy of city’s centers. The program is intended to reduce the number of vehicles in a city. In order to achieve this, I propose to create a system of encouragements in a form of cheaper public transportation tickets for people who park their cars outside the city. The possible implementation can be through using a parking ticket as a transportation ticket in buses and subways. Without such a system, people are unlikely to leave a car far from their destination location and then use public transportation to get there. This feels as a double price: parking cost and fare cost. Dedicated parking, where people can leave their car and further use public transportation services without additional monetary costs, will encourage people to use more public transportation, and additionally reduce vehicle occupancy in cities. Such program does not require high initial investments, rather a certain degree of coordination between different city authorities and parking firms. For example, Chicago Transit Authority provides a “Park & ride” service, which allows drivers to leave their cars in dedicated parking lots near subway stations. Additional encouragement in a form of reduced subway fee would certainly increase an amount of their customers. This can be accompanied by a parking lot tax, which would increase costs of in-city parking lots. As a result, people will be more likely to leave their vehicles further from the center due to higher cost of parking in the center, and more likely to use public transportation due to a reduced cost and lower impact of the accessibility issue.

The most important element to any transportation policy would be to inform population about benefits that this policy provides. Providing information to population is vital since the unawareness of population about certain policy can totally negate its effect.
### VII. Appendix A

Table 1. Descriptive Statistics of the Survey Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of P.T. trips</td>
<td>2.036</td>
<td>12.600</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Respondent's Age</td>
<td>53.770</td>
<td>17.277</td>
<td>16</td>
<td>92</td>
</tr>
<tr>
<td>Square of Age</td>
<td>3189.752</td>
<td>1814.110</td>
<td>256</td>
<td>8464</td>
</tr>
<tr>
<td>Driver Status</td>
<td>0.952</td>
<td>0.214</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Employment Status</td>
<td>0.568</td>
<td>0.495</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No High School</td>
<td>0.059</td>
<td>0.235</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>High School</td>
<td>0.252</td>
<td>0.434</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>College or Associate Degree</td>
<td>0.277</td>
<td>0.447</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Bachelor's degree</td>
<td>0.216</td>
<td>0.412</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Graduate or Professional Degree</td>
<td>0.161</td>
<td>0.368</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Low Income</td>
<td>0.339</td>
<td>0.473</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Middle Income</td>
<td>0.579</td>
<td>0.494</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>High Income</td>
<td>0.081</td>
<td>0.273</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Household Size</td>
<td>2.684</td>
<td>1.305</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>No Children</td>
<td>0.654</td>
<td>0.476</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A Child of Age 0 to 5</td>
<td>0.102</td>
<td>0.302</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A Child of Age 6 to 21</td>
<td>0.244</td>
<td>0.429</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>White</td>
<td>0.866</td>
<td>0.340</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>0.052</td>
<td>0.222</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.025</td>
<td>0.157</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>0.045</td>
<td>0.208</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No Vehicles</td>
<td>0.021</td>
<td>0.142</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Fewer Vehicles Than Drivers</td>
<td>0.106</td>
<td>0.308</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>More Vehicles Than Drivers</td>
<td>0.882</td>
<td>0.323</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Heavy Rail</td>
<td>0.176</td>
<td>0.381</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Low Population Density</td>
<td>0.738</td>
<td>0.440</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Medium Population Density</td>
<td>0.207</td>
<td>0.405</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>High Population Density</td>
<td>0.055</td>
<td>0.228</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Urbanicity Status</td>
<td>0.708</td>
<td>0.455</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: number of observations is 218,776
### Table 2. High P.T. Supply States

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York</td>
<td>23,444</td>
</tr>
<tr>
<td>California</td>
<td>31,771</td>
</tr>
<tr>
<td>District of Columbia</td>
<td>331</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>636</td>
</tr>
<tr>
<td>Illinois</td>
<td>1,218</td>
</tr>
<tr>
<td>Washington</td>
<td>608</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>1,224</td>
</tr>
<tr>
<td>Seven Public Transportation States</td>
<td>59,232</td>
</tr>
</tbody>
</table>
Table 3. Descriptive Statistics of the Regression Data

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>High P.T. Supply States</th>
<th>Low P.T. Supply States</th>
<th>All States Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 59,232</td>
<td>N = 159,544</td>
<td>N = 218,776</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>Fraction of P.T. trips</td>
<td>3.412</td>
<td>16.211</td>
<td>1.525</td>
</tr>
<tr>
<td>Respondent's Age</td>
<td>52.719</td>
<td>17.417</td>
<td>54.161</td>
</tr>
<tr>
<td>Square of Age</td>
<td>3082.630</td>
<td>1812.930</td>
<td>3229.520</td>
</tr>
<tr>
<td>Driver Status</td>
<td>0.941</td>
<td>0.236</td>
<td>0.956</td>
</tr>
<tr>
<td>Employment Status</td>
<td>0.583</td>
<td>0.493</td>
<td>0.563</td>
</tr>
<tr>
<td>No High School</td>
<td>0.050</td>
<td>0.219</td>
<td>0.062</td>
</tr>
<tr>
<td>High School</td>
<td>0.226</td>
<td>0.419</td>
<td>0.261</td>
</tr>
<tr>
<td>College or Associate Degree</td>
<td>0.279</td>
<td>0.449</td>
<td>0.276</td>
</tr>
<tr>
<td>Bachelor's Degree</td>
<td>0.218</td>
<td>0.413</td>
<td>0.216</td>
</tr>
<tr>
<td>Graduate or Professional Degree</td>
<td>0.189</td>
<td>0.391</td>
<td>0.151</td>
</tr>
<tr>
<td>Low Income</td>
<td>0.320</td>
<td>0.466</td>
<td>0.346</td>
</tr>
<tr>
<td>Middle Income</td>
<td>0.593</td>
<td>0.491</td>
<td>0.574</td>
</tr>
<tr>
<td>High Income</td>
<td>0.087</td>
<td>0.282</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td>2.782</td>
<td>1.364</td>
<td>2.648</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>Household Size</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No Children</strong></td>
<td>0.625</td>
<td>0.484</td>
<td>0.665</td>
</tr>
<tr>
<td><strong>A Child of Age 0 to 5</strong></td>
<td>0.106</td>
<td>0.308</td>
<td>0.100</td>
</tr>
<tr>
<td><strong>A Child of Age 6 to 21</strong></td>
<td>0.269</td>
<td>0.444</td>
<td>0.234</td>
</tr>
<tr>
<td><strong>White</strong></td>
<td>0.830</td>
<td>0.376</td>
<td>0.880</td>
</tr>
<tr>
<td><strong>Black</strong></td>
<td>0.034</td>
<td>0.180</td>
<td>0.059</td>
</tr>
<tr>
<td><strong>Hispanic</strong></td>
<td>0.044</td>
<td>0.204</td>
<td>0.019</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td>0.075</td>
<td>0.264</td>
<td>0.034</td>
</tr>
<tr>
<td><strong>No Vehicles</strong></td>
<td>0.032</td>
<td>0.177</td>
<td>0.016</td>
</tr>
<tr>
<td><strong>Fewer Vehicles Than Drivers</strong></td>
<td>0.123</td>
<td>0.328</td>
<td>0.100</td>
</tr>
<tr>
<td><strong>More Vehicles Than Drivers</strong></td>
<td>0.856</td>
<td>0.351</td>
<td>0.891</td>
</tr>
<tr>
<td><strong>Heavy Rail</strong></td>
<td>0.455</td>
<td>0.498</td>
<td>0.073</td>
</tr>
<tr>
<td><strong>Low Population Density</strong></td>
<td>0.550</td>
<td>0.498</td>
<td>0.808</td>
</tr>
<tr>
<td><strong>Medium Population Density</strong></td>
<td>0.293</td>
<td>0.455</td>
<td>0.175</td>
</tr>
<tr>
<td><strong>High Population Density</strong></td>
<td>0.157</td>
<td>0.364</td>
<td>0.017</td>
</tr>
<tr>
<td><strong>Urbanicity Status</strong></td>
<td>0.796</td>
<td>0.403</td>
<td>0.676</td>
</tr>
<tr>
<td>Variable</td>
<td>High P.T. Supply States</td>
<td>Low P.T. Supply States</td>
<td>All States</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-------------------------</td>
<td>------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Age</td>
<td>-0.242&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.111&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.165&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.047)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>Square of Age</td>
<td>0.002&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.001</td>
<td>0.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Driver Status</td>
<td>-2.230</td>
<td>-7.007&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-5.063&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(1.471)</td>
<td>(1.159)</td>
<td>(0.910)</td>
</tr>
<tr>
<td>Employment Status</td>
<td>0.395</td>
<td>-0.215</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.660)</td>
<td>(0.339)</td>
<td>(0.313)</td>
</tr>
<tr>
<td>High School</td>
<td>-0.751</td>
<td>-3.097&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-2.271&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(1.074)</td>
<td>(0.793)</td>
<td>(0.642)</td>
</tr>
<tr>
<td>College or Associate Degree</td>
<td>-1.370</td>
<td>-3.078&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-2.419&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(1.016)</td>
<td>(0.757)</td>
<td>(0.611)</td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td>-0.994</td>
<td>-3.265&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-2.370&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(1.027)</td>
<td>(0.778)</td>
<td>(0.623)</td>
</tr>
<tr>
<td>Graduate or Professional Degree</td>
<td>-0.524</td>
<td>-2.785&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-1.869&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(1.145)</td>
<td>(0.820)</td>
<td>(0.671)</td>
</tr>
<tr>
<td>Middle Income</td>
<td>-0.740</td>
<td>0.029</td>
<td>-0.176</td>
</tr>
<tr>
<td></td>
<td>(0.597)</td>
<td>(0.306)</td>
<td>(0.281)</td>
</tr>
<tr>
<td>High Income</td>
<td>0.305</td>
<td>0.739</td>
<td>0.616</td>
</tr>
<tr>
<td></td>
<td>(1.154)</td>
<td>(0.829)</td>
<td>(0.673)</td>
</tr>
<tr>
<td>Household Size</td>
<td>-0.964&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.256&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-0.485&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.307)</td>
<td>(0.155)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>A Child of Age 0 to 5</td>
<td>-0.782</td>
<td>-0.570</td>
<td>-0.718</td>
</tr>
<tr>
<td></td>
<td>(0.915)</td>
<td>(0.564)</td>
<td>(0.486)</td>
</tr>
<tr>
<td>A Child of Age 6 to 21</td>
<td>0.490&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.191</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>(0.820)</td>
<td>(0.419)</td>
<td>(0.394)</td>
</tr>
<tr>
<td>Black</td>
<td>5.135&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.005&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.717&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(1.578)</td>
<td>(0.620)</td>
<td>(0.627)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.813&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.230</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.901)</td>
<td>(0.880)</td>
<td>(0.619)</td>
</tr>
<tr>
<td>Other</td>
<td>1.299&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.051</td>
<td>0.630</td>
</tr>
<tr>
<td></td>
<td>(1.012)</td>
<td>(0.525)</td>
<td>(0.529)</td>
</tr>
<tr>
<td>Fewer Vehicles Than Drivers</td>
<td>-27.649&lt;sup&gt;a&lt;/sup&gt;</td>
<td>26.184&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-27.819&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(2.586)</td>
<td>(3.237)</td>
<td>(2.011)</td>
</tr>
<tr>
<td>More Vehicles Than Drivers</td>
<td>-35.235&lt;sup&gt;a&lt;/sup&gt;</td>
<td>28.691&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-32.134&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(2.479)</td>
<td>(3.211)</td>
<td>(1.971)</td>
</tr>
<tr>
<td>Medium Population Density</td>
<td>1.134&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.238</td>
<td>0.577&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.527)</td>
<td>(0.403)</td>
<td>(0.325)</td>
</tr>
<tr>
<td>High Population Density</td>
<td>4.906&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.574&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.636&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.748)</td>
<td>(1.606)</td>
<td>(0.727)</td>
</tr>
<tr>
<td>Heavy Rail</td>
<td>1.982&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.912&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.763&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.375)</td>
<td>(0.343)</td>
</tr>
<tr>
<td>Urbanicity Status</td>
<td>0.356</td>
<td>-0.020</td>
<td>-0.092</td>
</tr>
<tr>
<td></td>
<td>(0.524)</td>
<td>(0.236)</td>
<td>(0.220)</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>Constant</td>
<td>46.345a</td>
<td>43.480a</td>
<td>44.948a</td>
</tr>
<tr>
<td></td>
<td>(3.219)</td>
<td>(3.504)</td>
<td>(2.215)</td>
</tr>
<tr>
<td>Observations</td>
<td>59,232</td>
<td>159,544</td>
<td>218,776</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.229</td>
<td>0.132</td>
<td>0.188</td>
</tr>
</tbody>
</table>

Notes:  
Each regression includes states fixed effects. 
Robust standard errors are reported in parentheses below the coefficients estimates.  
a, b, c indicate statistical significance at 1%, 5%, and 10% levels respectively.
References


https://www.bts.gov/content/number-us-aircraft-vehicles-vessels-and-other-conveyances

https://journals.sagepub.com/doi/pdf/10.1177/0739456X0202200105


https://www.pewsocialtrends.org/2012/08/01/the-rise-of-residential-segregation-by-income/


https://nhts.ornl.gov/download.shtml#citation


http://vcp.med.harvard.edu/braess-paradox.html


http://documents.worldbank.org/curated/en/984261468327002120/pdf/863040IHME0T4H0ORLD0BANK0compressed.pdf