smart cities

MDPI

Article

Does Adoption of Ridehailing Result in More Frequent Sustainable Mobility Choices? An Investigation Based on the National Household Travel Survey (NHTS) 2017 Data

Vivekananda Das

Golisano Institute for Sustainability, Rochester Institute of Technology, Rochester, NY 14623, USA; vd1706@rit.edu

Received: 27 March 2020; Accepted: 7 May 2020; Published: 11 May 2020

Abstract: Among many changes potentially induced by the adoption of ridehailing, one key area of interest in transportation and urban planning research is how these services affect sustainable mobility choices, such as usage of public transit, walking, and biking modes and lower ownership of household vehicles. In this study, by using subsamples of the National Household Travel Survey (NHTS) 2017 data, propensity score matching technique is applied to generate matched samples of ridehailing adopters and non-adopters from ten different core-based statistical areas in the U.S. Results from multivariable count data regression models built on the matched samples indicate that, on average, the count of public transit trips is greater for adopters compared against identical non-adopters in all ten areas. Regarding average counts of walking and biking trips, adopters tend to make more trips in most of the places, although a few exceptions are also found. However, the relationship between ridehailing adoption and count of household vehicles appears to be more complicated as adopters, on average, seem to have a lower or higher number of vehicles than identical non-adopters, depending on the area. One major limitation of this study is that, in the statistical analyses, effects of attitudinal and detailed geographic variables are not directly controlled for, which complicates causal interpretations of findings.

Keywords: ridehailing; public transit; walking; biking; vehicle ownership; sustainable mobility; propensity score matching; count data regression

---

1. Introduction

Cities across the globe, faced with challenges posed by rapid urbanization and climate change, are exploring strategies to deal with growing demands in sustainable ways [1]. Sustainably meeting the mobility needs of city dwellers remains to be an intricate issue for urban planners and policymakers, as greenhouse gas emissions continue to rise from the transportation sector [2], while marginalized communities experience widening inequality due to a lack of adequate and affordable mobility [3].

Although sustainability has become a crucial topic in policy dialogues and academic research, there appears to be no universally agreed-upon definition of the terms sustainability, sustainable development, or sustainable mobility [4]. Despite disagreements on the definition, transportation and urban planning practitioners generally consider the promotion of multimodality, which incorporates higher usage of public transit, walking, and biking modes, and lower ownership and usage of private vehicles, as more sustainable in nature [4–8]. In this paper, I refer to trip making by public transit, walking, and biking modes and lower ownership of household vehicles as "sustainable mobility choices."

Over the last decade, cities around the U.S. have witnessed tremendous growth in the availability of shared mobility services [9]. Among the different varieties of shared mobility, ridehailing—services that enable users to get on-demand and short-term mobility access using smartphone apps—have drawn particular attention among urban residents, academics, and policymakers. According to the Pew Research Center, about 36% of U.S. adults mentioned using ridehailing services by 2018 [10].

Quite similar to the case of sustainability, practitioners often disagree on the definitions of different shared mobility services; consequently, services provided by the Transportation Network Companies (TNCs), such as Uber and Lyft, have been referred to as ridesharing, ridehailing, ridesourcing, app-based on-demand rides, etc. [11–13]. In this paper, I use the term "ridehailing" to describe both individual (e.g., UberX) and shared-ride (e.g., UberPool) services provided by the TNCs.

Studies conducted so far to explore the impact of ridehailing services on sustainable mobility choices hint at a wide range of possibilities: some indicate that these services can positively affect sustainable mobility choices, while others suggest that they can have a negative effect. Understandably, some adopters can use ridehailing to solve the first- and last-mile connection issues before and after transit trips, and also use these services when public transit services are unavailable; on the contrary, some other adopters, who can afford a higher fee, can replace public transit trips in situations when ridehailing works as a more convenient alternative. Although findings of previous studies have improved our understanding of the different mechanisms through which ridehailing services can affect other modes and vehicle ownership, limitations of existing studies include: usage of nonprobability samples, usage of samples from a specific urban area, usage of samples containing only adopter data, and application of only univariate statistical analyses. As a result, in terms of assessing the net effect, question remains on both generalizability and depth of many of these findings. As cities across the country are exploring opportunities to promote urban sustainability through reviving transit services, creating more walkable and bikeable spaces, and encouraging people to reduce private vehicle ownership and usage [14], it is crucial to have a deeper understanding of the net effect of ridehailing on these sustainable mobility choices.

In this paper, by building multivariable regression models on matched samples of adopters and non-adopters from ten core-based statistical areas in the U.S., I attempt to advance the existing understanding of the net effect of ridehailing adoption on sustainable mobility choices. As this study used data obtained from probability sampling and applied a statistical method widely used for impact assessment, findings can provide a more nuanced understanding.

The paper begins with a review of the existing literature relevant to the topic (Section 2). Next, in Section 3, the materials and methods used in this research—including a brief description of the National Household Travel Survey (NHTS) 2017 data, attainment of matched samples using propensity score matching, and analyses of the matched samples by building multivariable count data regression models—are discussed. In Section 4, results of the regression models are presented and interpreted. The final section (Section 5) explains the implications of findings, limitations of this research, and directions for future research.

2. Literature Review

So far, a number of studies have tried to investigate how the adoption of ridehailing affects sustainable mobility choices. In terms of reduction of private vehicle ownership, analyzing data from an online panel of respondents living in seven major U.S. cities, Clewlow and Mishra [15] found that, on average, ridehailing adopters did not possess significantly fewer vehicles than non-adopters. However, the authors mentioned that 9% of ridehailing adopters participating in their study decided to give up one or more vehicles. In another study, by collecting data from an intercept survey of ridehailing adopters in San Francisco and comparing findings against other data sources, Rayle et al. [11] found that 43% of ridehailing adopters belonged to zero vehicle households, whereas 35% of regular taxi users and 19% of the overall population of the city lived in similar households. They also found that

90% of vehicle owners had not changed vehicle ownership after beginning to use ridehailing, and suggested that presence of ridehailing had not influenced vehicle ownership decisions.

Regarding the impact of ridehailing on different travel modes (public transit, walking, biking, etc.), existing studies provide divergent results [16]. To assess how the adoption of ridehailing affects other modes, researchers often ask a stated preference question similar to "what modes would you have used if ridehailing were not available?" [16]. Combining findings from five studies (Table 1), it appears that ridehailing trips replace sustainable trips (transit, walking, and biking) to a greater extent than private automobile trips, although results indicating the extent of supposed replacements vary among studies.

Table 1. Findings indicating how the adoption of ridehailing affects different travel modes.

Would Have Chosen If Ridehailing Services were Unavailable *	Article					
	Rayle et al. [11]	Clewlow and Mishra [15]	Alemi [12]		Henao and Marshall [17]	Feigon and Murphy [18]
			Generation X	Millennial		
Would not have made/Fewer Trips	–	22%	7%	9.2%	12.22%	0%
Transit	33%	15%	11.9%	27.4%	22.2%	15%
Walk	8%	17%	–	–	–	6%
Bike	2%	7%	–	–	–	7%
Walk or Bike	–	–	11.9%	24.6%	11.9%	–
Drive private vehicle	6%	21%	38.3%	37.8%	19%	20%
Other	51%	19%	30.9%	1%	34.7%	52%

* Categories may not add up to 100 due to rounding. – indicates that the response option was not provided in the survey. Also, some of the response options have been rephrased (from original studies) in this table to create a generalized structure.

Interestingly, Hampshire at al. [19] investigated the same question, but in a revealed preference situation. After Uber and Lyft temporarily suspended services in 2016 in Austin, Texas, the authors asked previous Uber/Lyft passengers how they made a reference trip given the unavailability of Uber/Lyft. According to their findings, among those who participated in the survey, only 2.9% used public transit, whereas the majority of them either used a private vehicle (45%) or another ridehailing service (41%). Additionally, people who opted for private vehicles also included 8.9% of respondents who purchased a vehicle after Uber and Lyft had stopped operating in the city. These findings indicate that the presence of Uber and Lyft possibly replaces more private automobile trips than public transit trips, and in some instances, stops adopters from purchasing private vehicles. However, the authors mentioned using a convenience sample for the study, which limits the generalizability of the findings.

A number of other studies have attempted to explore the impact of ridehailing on public transit ridership at a macro level. Hall et al. [20] used a difference-in-differences design to estimate the effect of Uber on public transit ridership across major metropolitan cities in the U.S. According to their findings, Uber's entry increases ridership for an average transit agency. Although they found that the presence of Uber can increase or decrease transit ridership depending on the size and location of a transit agency, they concluded that the net effect of Uber on transit ridership was apparently positive. In another study, by using a longitudinal multilevel mixed-effect regression approach, Boisjoly et al. [21] investigated the determining factors of public transit ridership over the period of 2002 to 2015 for 25 transit agencies in North America. According to the authors, characteristics of the metropolitan area (size and population), average fares, and car ownership were the major factors influencing ridership, and ridehailing had an overall positive (but statistically insignificant) impact on the outcome. Contrary to the findings of these two studies, Graehler et al. found negative association between the presence of ridehailing and transit ridership [22]. The authors updated Boisjoly's approach by using more recent data and segmenting different transit services. According to their findings, in every year since the entrance of ridehailing services in a market, heavy rail and bus ridership can decrease by 1.3% and 1.7%, respectively.

Overall, conflicting indications on the net effect of ridehailing adoption on sustainable mobility choices may be due to a number of reasons, such as differences in the geographic area under study, in sampling procedures (e.g., probability sampling versus nonprobability sampling), in the type of survey data (e.g., stated versus revealed), in the time period of the survey, etc. Above all, one major limitation in the existing understanding is that many of the suggestions have been given based on descriptive statistics, which cannot control for self-selection bias. This bias occurs when participants of a study choose their own treatment condition (e.g., whether to adopt ridehailing or not), which in turn makes it difficult to estimate whether the outcome is due to receiving the treatment or due to the inherent differences between those who receive the treatment and those who do not [23]. For estimating treatment effects from observational data, quantitative social science researchers use a number of econometric methods, such as propensity score matching, instrumental variable, and regression discontinuity design, which require stronger theoretical assumptions apart from standard statistical modeling [24]. This study applied the propensity score matching technique as an attempt to explore the effect of ridehailing adoption on the four sustainable mobility choices in the context of ten different areas. Findings can provide deeper insights into the effect of ridehailing adoption in these areas, and limitations pointed out in the final section (Section 5) should facilitate better design of future studies for a more precise estimation of the effect.

3. Materials and Methods

3.1. National Household Travel Survey (NHTS) 2017 Data

The National Household Travel Survey (NHTS), conducted by the Federal Highway Administration (FHWA), is a nationally representative survey that collects data relevant to noncommercial trips made by American households [25]. A sample of 264,234 individuals aged 5 or above from 129,696 households was collected during the NHTS 2017. However, for the purpose of this research, a subset of the NHTS 2017 sample, consisting of people living in the ten leading core-based statistical areas in terms of public transit usage—New York–Newark–Jersey City, Atlanta–Sandy Springs–Roswell, Boston–Cambridge–Newton, Chicago–Naperville–Elgin, Los Angeles–Long Beach–Anaheim, Miami–Fort Lauderdale–West Palm Beach, Philadelphia–Camden–Wilmington, San Francisco–Oakland–Hayward, Seattle–Tacoma–Bellevue, and Washington–Arlington–Alexandria [26]—was chosen. Given the fact that public transit availability, walkability, and bikeability differ widely across the ten areas, it is possible to observe how sustainable mobility choices differ, not only between adopters and non-adopters living in the same area, but also among adopters from different areas. Next, as ridehailing services require users to be 18 or older [27,28], the cases of respondents aged below 18 are eliminated. Although the NHTS 2017 recorded responses of 31,210 respondents aged 18 or above living in the ten areas considered in this research, by eliminating the cases of respondents who did not provide answers to one or more questions relevant to the dependent and independent variables of this research, I initially narrowed down to a sample of 29,695 respondents. Figure 1 presents the ridehailing adoption rate in the ten areas based on the initial sample.

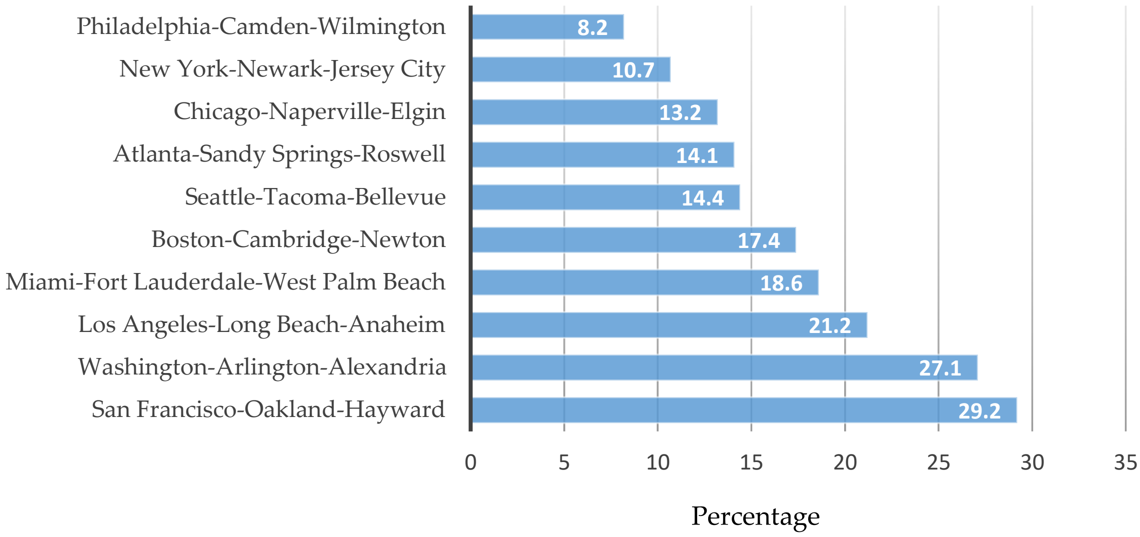

Figure 1. Ridehailing adoption rate in the ten areas.

3.2. Need for a Matched Sample

In order to assess the effect of a treatment (e.g., adoption of ridehailing) on an outcome (e.g., count of public transit trips), ideally, we need to conduct a randomized experiment in which we manipulate the treatment by randomly assigning people to a treatment group (e.g., adopters) and a control group (e.g., non-adopters) [29]. Randomization ensures that the treatment and control groups are, on average, identical to one another considering all predictors except the treatment variable, and therefore, the two groups can be considered as counterfactual of one another [29]. In that case, by subtracting the average outcome for the control group from the average outcome for the treatment group, we get an estimate of the average treatment effect [29]. In simple mathematical terms, for any participant (i) in the experiment, average treatment effect = $\mathrm{E}\,(y_i^1 - y_i^0) = \mathrm{E}\left(y_i^1\right) - \mathrm{E}\left(y_i^0\right)$, where y_i^1 = outcome in the case of receiving the treatment and y_i^0 = outcome in the case of not receiving the treatment [30].

However, in many practical situations, such as for assessing the effect of ridehailing adoption on sustainable mobility choices, conducting a randomized experiment can be extremely difficult, and researchers have to rely on observational data. A fundamental challenge in making causal inferences from observational data is that treatment and control groups often differ from one another, based not only on the treatment variable but also on a number of other predictors [29,30]. One possible way of countering this problem is to build a multivariable regression model which can isolate the effect of the treatment variable by controlling for the effects of other variables; however, if there are large differences between the treatment and control groups, regression coefficients can be biased, which would result in biased estimation of the treatment effect [31]. To solve this problem, by using propensity score matching, we can obtain an observationally equivalent subsample of the treatment and control groups, which provides a stronger case to find an estimate of the treatment effect [29,32]. In a way, this technique can potentially mimic a randomized experiment by creating a treatment and a control group which are, on average, identical to one another, depending on observed predictors [30]. After matched samples are obtained, a regression model, that includes the treatment status indicator (a dummy variable) and propensity scores as predictors, can be built to estimate the average treatment effect. For example, the coefficient (β) in the following model can be interpreted as an estimator of the average treatment effect:

$$y_i = \alpha + \beta\, T_i + \gamma\, P(x_i) + e_i, \tag{1}$$

where T_i = treatment indicator and $\mathrm{P}(x_i)$ = proponsity score [30]

Also, the regression method makes it possible to adjust for the effects of other confounders which are not considered during the matching process [30].

3.3. Getting Matched Samples using Propensity Score Matching

Propensity scores for respondents from the ten areas were separately obtained using logit models that predict the probability of receiving the treatment (ridehailing adoption) based on nine predictors: urbanicity of household location, annual household income, educational attainment, race, sex, generational cohort, life cycle classification for the household, household size, and household worker count. A description of these variables is provided in Table 2. In the logit models built for generating propensity scores, outcome variables of this research were not considered as predictors following recommendations provided in [31,33]. Next, based on the propensity score of each adopter, a similar non-adopter (ratio = 1) was selected using the nearest-neighbor matching technique, and unmatched non-adopters were discarded. Finally, ten matched samples, each having an equal number of adopters and non-adopters from a particular area, were found. All statistical analyses were done separately for the ten matched samples. A description of the dependent and independent variables of the regression models (described in Section 3.6) is provided in Table 3.

Table 2. Description of the predictors used in the matching process.

Variable Name	Variable Description	Levels/Values
HBHUR	Urbanicity of Household Location (as defined by Claritas [34])	Urban Suburban Second City Small Town Rural
INCOME *	Annual Household Income	Less than $50,000 $50,000 to $100,000 $100,000 or more
EDUCATION *	Educational Attainment	Below Bachelor's Bachelor's or above
RACE *	Race	White Black or African American Asian Others (American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, Multiple responses selected, Some other race)
R_SEX	Sex	Male Female
HHSIZE	Count of Household Members	1–13
GENERATION *	Generation (as defined by the Pew Research Center [35])	Post-Millennial (Aged between 18 and 20)Millennial (Aged between 21 and 36) Generation X (Aged between 37 and 52) Baby Boomer (Aged between 53 and 71) Silent and Greatest (Aged 72 or above)
WRKCOUNT	Number of workers in household	0–7
LIF_CYC	Life Cycle classification for the household, derived by attributes pertaining to age, relationship, and work status	One adult, no children 2+ adults, no children One adult, youngest child 0–5 2+ adults, youngest child 0–5 One adult, youngest child 6–15 2+ adults, youngest child 6–15 One adult, youngest child 16–21 2+ adults, youngest child 16–21One adult, retired, no children 2+ adults, retired, no children

* These variables have been created by recategorizing relevant NHTS variables. For example, INCOME has been recreated from HHFAMINC, EDUCATION from EDUC, RACE from R_RACE, and GENERATION from R_AGE.

Table 3. Description of the variables used in the regression models.

Variable Name	Question in the NHTS 2017 [36]	Values/Levels	Data Type
PTUSED	In the past 30 days, about how many days have you used public transportation such as buses, subways, streetcars, or commuter trains?	0–240	Count
NWALKTRIP	In the past 7 days, how many times did you take a walk outside including walks to exercise, go somewhere, or to walk the dog (e.g., walk to a friend's house, walk around the neighborhood, walk to the store, etc.)?	0–200	Count
NBIKETRIP	In the past 7 days, how many times did you ride a bicycle *outside* including bicycling to exercise, or to go somewhere (e.g., bike to a friend's house, bike around the neighborhood, bike to the store, etc.)?	0–99	Count
HHVEHCNT	How many vehicles are owned, leased, or available for regular use by the people who currently live in your household? *Include motorcycles, mopeds, and RVs.*	0–12	Count
CARSHARE	In the past 30 days, how many times did you use a car-sharing service where a car can be rented by the hour (e.g., Zipcar orCar2Go)?	0 = Non-adopter 1 = Adopter	Binary
RIDEHAIL *	In the past 30 days, how many times have you purchased a ride with a smartphone rideshare app (e.g., Uber, Lyft, Sidecar)?	0 = Non-adopter 1 = Adopter	Binary

* Although the NHTS 2017 used "RIDESHARE," I used "RIDEHAIL/Ridehailing" throughout this paper to refer to the same service. Also, the RIDESHARE variable was recorded as count data in the NHTS 2017. In this research, I created the RIDEHAIL variable by coding RIDESHARE as a binary variable to separate out the adopters and non-adopters. The same procedure was followed to recreate the CARSHARE variable.

3.4. Summary Statistics

Table 4 shows the average count of public transit, walking, and biking trips, and the average count of household vehicles for adopters and non-adopters in the matched samples from all ten areas. Sample sizes for each area are also shown. Additionally, for each mobility choice, an independent sample t-test was conducted to identify whether the average count for adopters and non-adopters are significantly different at the 5% significance level. Based on the findings, counts of public transit trips for adopters are significantly greater than the same for non-adopters, regardless of the area. However, public transit trip counts among adopters vary widely depending on the area; for example, on average, ridehailing adopters in New York–Newark–Jersey City make 12.57 trips, whereas adopters in Miami–Fort Lauderdale–West Palm Beach make only 2.02 trips. Possibly, these variations are caused by differences in the quality of public transit across cities.

In terms of walking trip counts, on average, adopters make significantly more trips than non-adopters in all areas except in Atlanta–Sandy Springs–Roswell, Boston–Cambridge–Newton, and Seattle–Tacoma–Bellevue. Similarly, average biking trip count seems to be significantly greater for adopters in all areas except in Boston–Cambridge–Newton, Chicago–Naperville–Elgin, and Miami–Fort Lauderdale–West Palm Beach. Similar to the variation in transit trips, differences in walking and biking trips among ridehailing adopters from different areas are possibly caused by the disparity in walkability and bikeability across places.

Regarding counts of household vehicles, there appears to be no significant difference between adopters and non-adopters living in Chicago–Naperville–Elgin, Miami–Fort Lauderdale–West Palm Beach, and Philadelphia–Camden–Wilmington areas, but adopters have significantly lower household vehicles in all other areas.

Table 4. Average sustainable mobility choices by ridehailing adopters and non-adopters in the matched samples from the ten areas.

Core-Based Statistical Area (CBSA) of the Respondent's Home Address	Public Transit Trips		Walking Trips		Biking Trips		Household Vehicle	
	Adopter	Non-Adopter	Adopter	Non-Adopter	Adopter	Non-Adopter	Adopter	Non-Adopter
New York–Newark–Jersey City, NY–NJ–PA ($N_{adopter} = 1037$, $N_{non-adopter} = 1037$)	12.57 *	6.86 *	12.25 *	8.68 *	0.80 *	0.29 *	1.48 *	1.85 *
Atlanta–Sandy Springs–Roswell, GA ($N_{adopter} = 665$, $N_{non-adopter} = 665$)	2.54 *	0.72 *	6.09	5.67	0.29 *	0.17 *	1.90 *	2.14 *
Boston–Cambridge–Newton, MA–NH ($N_{adopter} = 114$, $N_{non-adopter} = 114$)	9.19 *	5.39 *	11.60	9.50	0.62	0.68	1.26 *	1.68 *
Chicago–Naperville–Elgin, IL–IN–WI ($N_{adopter} = 203$, $N_{non-adopter} = 203$)	7.45 *	4.19 *	8.79 *	6.71 *	0.56	0.48	1.85	2.00
Los Angeles–Long Beach–Anaheim, CA ($N_{adopter} = 1175$, $N_{non-adopter} = 1175$)	2.14 *	1.30 *	7.11 *	5.24 *	0.51 *	0.29 *	2.02 *	2.15 *
Miami–Fort Lauderdale–West Palm Beach, FL ($N_{adopter} = 97$, $N_{non-adopter} = 97$)	2.02 *	0.49 *	7.26 *	4.57 *	0.62	0.31	2.24	2.00
Philadelphia–Camden–Wilmington, PA–NJ–DE–MD ($N_{adopter} = 89$, $N_{non-adopter} = 89$)	7.28 *	2.13 *	10.46 *	7.31 *	0.62 *	0.13 *	1.85	2.10
San Francisco–Oakland–Hayward, CA ($N_{adopter} = 1156$, $N_{non-adopter} = 1156$)	8.01 *	3.91 *	9.46 *	6.16 *	0.92 *	0.45 *	1.69 *	2.16 *
Seattle–Tacoma–Bellevue, WA ($N_{adopter} = 83$, $N_{non-adopter} = 83$)	6.52 *	2.78 *	7.93	6.66	0.93 *	0.08 *	1.71 *	2.10 *
Washington–Arlington–Alexandria, DC–VA–MD–WV ($N_{adopter} = 380$, $N_{non-adopter} = 380$)	9.54 *	4.70 *	11.86 *	7.41 *	0.84 *	0.35 *	1.36 *	1.70 *

* indicates average counts for adopters and non-adopters that are significantly different at the 5% significance level.

3.5. Multivariable Regression Models

As all four dependent variables of this research are count data (non-negative integers), building ordinary Linear regression models may not be the ideal approach, because count data often violate two important assumptions—normality and homoscedasticity—of Linear regression [37], and thus Linear regression models built on count data can produce biased estimates [38]. For such cases, Poisson regression models can be considered; however, Poisson distribution assumes equality of mean and variance [37,38], which seemingly does not hold true for any of the four dependent variables (regardless of the area) considered in this paper. From Table 5, it is evident that, for all ten areas, the variance is greater than the mean (indicating overdispersion) for the count of public transit, walking, and biking trips, whereas the variance is smaller than the mean (indicating underdispersion) for the count of household vehicles.

To model overdispersed count data, researchers often use a Negative Binomial (NB) regression model, which is a generalized version of the Poisson model and capable of dealing with the overdispersion by incorporating an extra parameter α that accounts for unobserved heterogeneity among observations [37,38]. NB regression models have been used in transportation research for finding answers to a wide range of questions. For example, Marshall and Ferenchak (2019) modeled crash counts from twelve large U.S. cities [39]; Zahran et al. (2008) modeled counts of walk commuters [40]; Wang et al. (2014) estimated mixed-mode urban trail traffic [41]; Hu et al. (2012) modeled crash frequency at highway-railroad grade crossings [42]; Cao et al. (2006) estimated frequencies of strolling trips and pedestrian shopping trips [43]; Zhao and Kockelman (2002) modeled household vehicle ownership [44]; Young and Lachapelle (2017) modeled travel time for different travel modes and frequency of trips for different purposes in Canadian cities [45] using NB regression models. On the contrary, underdispersed count data modeling appears to be less explored [46] as the phenomenon is less commonly observed [47]. To model underdispersed count data, Quasi-Poisson (QP) regression models can be useful, as illustrated by Harris et al. [46] and Wilson et al. [48].

Table 5. Differences between Mean and Variance in Dependent Variables.

Core-Based Statistical Area (CBSA) of the Respondent's Home Address	Public Transit Trips		Walking Trips		Biking Trips		Household Vehicle	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
New York–Newark–Jersey City, NY–NJ–PA	9.72	157.56	10.47	154.75	0.54	10.34	1.66	1.58
Atlanta–Sandy Springs–Roswell, GA	1.63	26.66	5.88	65.12	0.23	1.28	2.02	0.98
Boston–Cambridge–Newton, MA–NH	7.29	103.83	10.55	112.84	0.65	6.34	1.47	1.28
Chicago–Naperville–Elgin, IL–IN–WI	5.82	110.44	7.75	80.93	0.52	3.56	1.93	1.42
Los Angeles–Long Beach–Anaheim, CA	1.72	32.86	6.17	64.48	0.40	3.00	2.09	1.32
Miami–Fort Lauderdale–West Palm Beach, FL	1.26	24.37	0.46	2.89	0.46	2.89	2.12	1.01
Philadelphia–Camden–Wilmington, PA–NJ–DE–MD	4.71	95.77	8.89	109.98	0.38	2.28	1.98	1.28
San Francisco–Oakland–Hayward, CA	5.96	91.43	7.81	73.32	0.69	5.06	1.92	1.39
Seattle–Tacoma–Bellevue, WA	4.65	63.91	7.30	49.17	0.51	3.37	1.90	1.21
Washington=Arlington–Alexandria, DC–VA–MD–WV	7.12	103.59	9.72	98.99	0.59	4.68	1.53	1.09

For the QP model, expected count, E(Y) = μ and variance, V(Y) = θ μ, where μ and θ denote conditional mean and dispersion parameter, respectively [49].

Also, for the NB model, expected count, E(Y) = λ and variance, V(Y) = λ (1 + α λ), where λ and α denote conditional mean and dispersion parameter, respectively [49].

Given the possible presence of overdispersion, multivariable NB regression models are built to investigate the extent to which the count of public transit trips, walking trips, and biking trips can get affected by the adoption of ridehailing. Also, due to the possible presence of underdispersion, multivariable QP regression models are built to explore how household vehicle ownership can be impacted by the adoption of ridehailing.

Although adopters and non-adopters were matched based on nine predictors related to sociodemographic and geographic characteristics, there could be some other differences between them that were not been controlled for. Findings from previous studies indicate that people who use public transit are more likely to make more walking and biking trips, and belong to households with no/fewer vehicles [50,51]. As a consequence, it is possible that differences in the count of walking and biking trips are due to differences in the count of public transit trips rather than ridehailing adoption. Furthermore, adoption of carsharing can be related to all four outcome variables of this research [31]. To negate these possibilities, in the regression models predicting the count of trips made by a certain mode, count of trips made by other modes, adoption of carsharing, and count of household vehicles were used as confounding variables. For the same reason, in the models predicting household vehicles, count of trips made by public transit, walking, and biking modes and adoption of carsharing were used as confounding variables.

It is worth mentioning that a number of other variables relevant to the physical characteristics of respondents' household and workplace locations and their attitudes towards different travel modes were not used in the analyses due to either unavailability of relevant variables (physical characteristics) or lack of enough responses to relevant questions (attitudinal) in the NHTS 2017. Regarding questions relevant to attitudes towards walking and biking trips, the majority of the respondents skipped answering. However, the predictors used in the matching process can potentially account for the ones that were not included. For example, Bhat and Pulugurta used descriptors of residential location based on urbanization (Urban and Suburban) as proxy variables for other location-relevant factors, such as opportunities to perform activities by transit, level of service from auto and transit modes, and auto maintenance and insurance costs [52]. Also, Na Chen argued that sociodemographic variables can partially control for the effects of attitudinal variables [53]. Consequently, based on the assumption that predictors used in the analyses could control for the effects of other possible confounders, subsequent analyses were conducted and the results are interpreted.

3.6. Implementation of Regression Models

The four multivariable regression models for each area were designed as follows:

$$\text{Model NB-TRANSIT: } \log(\lambda_{PTUSED}) \sim f\,(\text{RIDEHAIL, Propensity Score, Other Confouding Variables})$$
$$\text{Model NB-WALK: } \log(\lambda_{NWALKTRIP}) \sim f\,(\text{RIDEHAIL, Propensity Score, Other Confouding Variables})$$
$$\text{Model NB-BIKE: } \log(\lambda_{NBIKETRIP}) \sim f\,(\text{RIDEHAIL, Propensity Score, Other Confouding Variables})$$
$$\text{Model QP-VEHICLE: } \log(\lambda_{HHVEHCNT}) \sim f\,(\text{RIDEHAIL, Propensity Score, Other Confouding Variables})$$

where λ denotes expected counts.

3.7. Software and Packages Used for Analyses

All statistical analyses were done in RStudio version 3.5.1. Propensity score matching analysis was done using MatchIt package [54]. The glm.nb function in MASS package [55] was used to fit the Negative Binomial regression models, and the glm function in stats package [56] was used to fit the Quasi-Poisson regression model.

4. Results

4.1. Interpretation of Regression Model Output

The outcomes of both Negative Binomial and Quasi-Poisson regression models take the form of natural logarithm of expected counts (ln λ) [46,57]. For numeric predictors, coefficient β expresses β units change in the outcome with one unit change in the predictor, whereas for categorical predictors with multiple levels, β denotes the outcome changes by β units for the considered level compared against the base level. Another way of expressing model outcomes is in terms of exponentiated coefficients, also known as Incidence Rate Ratios (IRR), in which the outcomes take the form of expected counts (λ). In this transformed scale, IRR denotes that one unit change in the numeric predictor is associated with exp (β) times change in the outcome. For categorical predictors with multiple levels, IRR expresses exp (β) times change in the outcome for the considered level compared against the base level.

In Table 6, regression outputs have been shown in terms of IRR as it is relatively easier to interpret. An IRR value of greater than 1 indicates a positive association between the predictor and the outcome, whereas a value of less than 1 suggests a negative association between them.

4.2. Multicollinearity and Goodness of Fit

The goal of building multivariable regression models in this research was to investigate the effect of ridehailing adoption on sustainable mobility choices after controlling for the effects of as many confounding variables as possible. Although adding many predictors in a regression model can minimize omitted variable bias, it increases the possibility of multicollinearity, a phenomenon that can affect the estimation of both the coefficient and the standard error of the coefficient of a predictor which is correlated with one or more predictors in the model [58]. However, multicollinearity can be ignored in cases where predictors added to the model only for controlling purpose are correlated with each other but not with the predictors of research interest (treatment variable) [59]. Variance Inflation Factor (VIF), a diagnostic tool to detect the presence of multicollinearity in multivariable regression models, was calculated to confirm the absence of multicollinearity. Low VIF values for the RIDEHAIL variable in all four models suggest that multicollinearity should not be a concern in the context of this research. The VIF analyses were conducted using CAR package [60] in RStudio.

To evaluate the goodness of fit for generalized linear models (Logistic, Poisson, Negative Binomial, etc.), several Pseudo R-squared methods have been developed [61–63]. For all the models, I calculated McFadden's Pseudo R-squared values, which can be mathematically expressed as:

$$R^2_{McFadden} = 1 - \frac{\ln \hat{L}\ (Model_{full})}{\ln \hat{L}\ (Model_{intercept\ only})} \quad (2)$$

where $\hat{L}$ = Estimated likelihood, $Model_{full}$ = Full estimated model, $Model_{intercept\ only}$ = Intercept-only model.

The value of $R^2_{McFadden}$ indicates improvement in the full estimated model compared against the intercept-only model. $R^2_{McFadden}$ values were calculated using DescTools package [64] in RStudio.

The complete findings of all forty (four for each area) regression models, including estimates of coefficients, exponentiated coefficients (IRR), $R^2_{McFadden}$ values, and VIF scores, are provided in the Supplementary Materials.

4.3. Discussion on the Association between Ridehailing Adoption and Sustainable Mobility Choices after Controlling for the Confounders

In this section, I limit the discussion to the IRR of the treatment variable (ridehailing adoption); to put it simply, I avoid explaining the IRR of other predictors as they have been added to the models only for controlling purpose. IRR values of the RIDEHAIL variable from all forty models are presented in Table 6. $R^2_{McFadden}$ values for the models mostly range between 0.1 and 0.2, which seems to be satisfactory given the fact that these values tend to be considerably lower than the R^2 values found in Linear regression, and $R^2_{McFadden}$ values between 0.2 and 0.4 represent excellent fit [65].

Table 6. IRR (exponentiated coefficient) of the RIDEHAIL variable from the regression models. (Base case: Non-adopter)

Area Relevant to the Model	Model			
	NB-TRANSIT	NB-WALK	NB-BIKE	QP-VEHICLE
New York–Newark–Jersey City, NY–NJ–PA	1.570 *	1.154 *	1.872 *	0.954
Atlanta–Sandy Springs–Roswell, GA	3.312 *	0.952	1.096	0.930 *
Boston–Cambridge–Newton, MA–NH	2.434 *	1.120	0.765	0.935
Chicago–Naperville–Elgin, IL–IN–WI	1.651 *	1.183	1.394	1.104
Los Angeles–Long Beach–Anaheim, CA	1.552 *	1.274 *	1.508 *	0.976
Miami–Fort Lauderdale–West Palm Beach, FL	2.740	1.511 *	1.260	1.153 *
Philadelphia–Camden–Wilmington, PA–NJ–DE–MD	3.524 *	1.099	6.101 *	1.002
San Francisco–Oakland–Hayward, CA	1.528 *	1.226 *	1.063	0.911 *
Seattle–Tacoma–Bellevue, WA	2.047 *	0.942	7.975 *	0.899
Washington—Arlington–Alexandria, DC–VA–MD–WV	1.639 *	1.221 *	2.222 *	1.007

* indicates IRR that is significant at the 5% significance level.

Findings presented in Table 6 show statistically significant positive association between adoption of ridehailing and count of public transit trips in all areas except in Miami–Fort Lauderdale–West Palm Beach, where the association is positive but insignificant. These findings indicate that, compared against identical non-adopters, on average, adopters make significantly more public transit trips in almost all areas. Also, the magnitude of the positive association varies across urban areas: adopters in Philadelphia–Camden–Wilmington area, on average, make 3.524 times more transit trips than identical non-adopters, whereas adopters in Los Angeles–Long Beach–Anaheim, on average, make 1.552 times more transit trips than similar non-adopters.

In terms of walking trips, on average, ridehailing adopters make significantly more trips than identical non-adopters in five areas: New York–Newark–Jersey City, Los Angeles–Long Beach—Anaheim, Miami–Fort Lauderdale–West Palm Beach, San Francisco–Oakland–Hayward, and Washington–Arlington–Alexandria. In other places, adoption of ridehailing seems to have a mostly positive but statistically insignificant relationship with the count of walking trips.

Count of biking trips seems to have significant positive association with ridehailing adoption in New York–Newark–Jersey City, Los Angeles–Long Beach–Anaheim, Philadelphia–Camden–Wilmington, Seattle–Tacoma–Bellevue, and Washington–Arlington–Alexandria. No significant association is observed in other areas, although the association tends to be positive except in Boston–Cambridge–Newton.

Regarding the count of household vehicles, on average, ridehailing adopters have significantly lower household vehicles compared against similar non-adopters only in Atlanta–Sandy Springs–Roswell and San Francisco–Oakland–Hayward, but the opposite holds true in Miami–Fort Lauderdale–West Palm Beach. In other places, the relationship between ridehailing adoption and the count of household vehicles appears to be insignificant. Nevertheless, it should be noted that some of the associations may have been found to be insignificant due to relatively small sample sizes used in the analyses. Also, given limitations in the data, it is difficult to explain why these associations vary across the ten areas.

In the context of this research, perhaps, the more important question is: do these associations express causal effects of ridehailing adoption? The answer to the question largely depends on the extent to which confounding variables included in the analyses controlled for the effects of all possible confounders. To elaborate on this, it is worth taking a look back at the relationship between ridehailing adoption and count of household vehicles in the Chicago–Naperville–Elgin area. From the summary statistics (Table 4), it appears that average count of household vehicles is 1.08 times higher for non-adopters; however, results of the multivariable regression analysis (Table 6)—that included propensity score, count of public transit trips, count of walking trips, count of biking trips, and adoption of carsharing as confounders—show that, after controlling for the five confounders, adopters, on average, have 1.104 times higher household vehicles. This example illustrates the importance of controlling for the effects of as many confounders as possible for precise estimation of effects. In future research, detailed data on respondents' attitudes, lifestyle preferences, trip purpose and duration, and built environment characteristics of their living and workplace locations should be collected, as directly controlling for these factors should result in more reliable effect estimation.

5. Discussion and Conclusions

In this research, using matched samples of ridehailing adopters and non-adopters from ten core-based statistical areas in the U.S., multivariable regression models were built as an attempt to assess the effect of the service adoption on sustainable mobility choices. Findings indicate that, regardless of the urban area, ridehailing adopters use public transit more often than identical non-adopters. Also, adopters tend to walk and bike more frequently in most areas. However, ridehailing's relationship with the count of household vehicles is apparently more location-dependent and can be either positive or negative. Overall, although not yet conclusive, results suggest that ridehailing adoption enhances multimodality to some extent in all ten areas but may not be related to lower vehicle ownership everywhere.

As this study was focused only on ten core-based statistical areas, results may not be applicable to other areas. From a policy perspective, transit agencies across the country—which are considering the option to incorporate ridehailing into their broader transportation planning—should conduct studies improving upon the limitations articulated in this paper to figure out how the adoption of ridehailing affects sustainable mobility choices in their local contexts. In terms of environmental impact, understandably, ridehailing increases Vehicle Miles Traveled (VMT) and greenhouse gas emissions mostly due to "deadheading," which refers to the extra miles that a ridehailing vehicle travels without a passenger in between two rides. Presumably, the overall negative externalities created by ridehailing vehicles can be reduced if these services play a part in inducing and retaining multimodality and reducing private vehicle ownership. However, as findings of this study suggest, the magnitude of the probable impact of ridehailing varies across urban areas, and so the true potential of ridehailing to enhance multimodality and to reduce dependence on private vehicles at a certain area may not be achieved unless the built environment is transformed for the same purpose. Interestingly, over the last five years, a number of public transit agencies across the U.S., including Dallas Area Rapid Transit (DART) in Dallas, Texas, and Pinellas Suncoast Transit Agency in St. Petersburg, Florida, have partnered with Uber to increase the reach of transit services and to reduce operational costs [66,67]. Detailed case studies on the outcomes of these partnerships can provide further insights on whether such coordinated efforts from the public and private sectors facilitate sustainable mobility choices.

Regarding limitations of this study, it should be re-emphasized that the ignorability assumption, which assumes that the effects of all confounders have been controlled for, required for making causal inferences from regression models, may have been violated in case the predictors used in the models for controlling purpose did not capture the effects of potentially omitted confounders. Additionally, in this research, the possibility of reverse causality was not accounted for. A number of studies have found that people who use public transit and belong to households with a lower number of vehicles are more likely to adopt ridehailing [68–70]. Although no study conducted so far has confirmed the causal direction, failing to account for the possibility of reverse causality further complicates the making of any causal inferences from the findings of this research. In future work, for more precise estimation, researchers should collect more detailed data and keep using econometric methods developed for making causal inferences in the context of nonexperimental studies. However, a key challenge here is that, although academic researchers can gather detailed data by developing better survey instruments, as a method of data collection at a regional or national level, they have to depend on commercial online survey panels which may not result in a sample as representative (due to nonprobability-based recruitment) as the one from the NHTS. Therefore, in future NHTS, collection of information relevant to detailed geographic and attitudinal factors should be considered.

Finally, similar to all other studies using statistical modeling on cross-sectional survey data, this study has some other limitations. Travel decisions can vary over time, and in the case of emerging information-communication-technology-based mobility systems, we would expect even more rapid changes. As the NHTS 2017 asked respondents to mention counts of public transit and ridehailing trips made in the previous 30 days and counts of walking and biking trips made in the previous 7 days, this study fails to capture the long-term trends. Also, respondents might have made mistakes in retrieving information on the count of trips they had made using different travel modes.

Supplementary Materials: The following are available online at http://www.mdpi.com/2624-6511/3/2/20/s1.

Author Contributions: The author affirms sole responsibility for the research conception and design, statistical modeling, interpretation of the results, and overall preparation of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The author expresses gratitude to Eric Williams and Qing Miao for their valuable comments during the preparation of this paper. The author also thanks two anonymous reviewers for their useful suggestions to improve the paper.

Conflicts of Interest: The author declares no conflict of interest.

References

1. The World Bank. Cities Around the World Want to Be Resilient and Sustainable. But What Does This Mean? Available online: https://www.worldbank.org/en/news/feature/2019/10/15/resilient-sustainable-cities-brazil-gpsc-conference. (accessed on 1 February 2020).

2. Wang, S.; Ge, M. Everything You Need to Know About the Fastest-Growing Source of Global Emissions: Transport. Available online: https://www.wri.org/blog/2019/10/everything-you-need-know-about-fastest-growing-source-global-emissions-transport. (accessed on 1 February 2020).
3. White, G.B. Stranded: How America's Failing Public Transportation Increases Inequality. Available online: https://www.theatlantic.com/business/archive/2015/05/stranded-how-americas-failing-public-transportation-increases-inequality/393419/ (accessed on 1 February 2020).
4. Sustainable Transportation and TDM. Available online: https://www.vtpi.org/tdm/tdm67.htm (accessed on 1 February 2020).
5. Olafsson, A.S.; Nielsen, T.S.; Carstensen, T.A. Cycling in multimodal transport behaviours: Exploring modality styles in the Danish population. *JTRG* **2016**, *52*, 123–130. [CrossRef]
6. Nobis, C. Facets and Causes of Sustainable Mobility Behavior. *Transp. Res. Rec. J. Transp. Res. Board* **2007**, *1*, 35–44. [CrossRef]
7. Spickermann, A.; Grienitz, V.; von der Gracht, H.A. Heading towards a multimodal city of the future? Multi-stakeholder scenarios for urban mobility. *Technol. Forecast. Soc. Chang.* **2014**, *89*, 201–221. [CrossRef]
8. Thematic Discussion 8: Multi-?-Modal Sustainable Transport and Transit Solutions: Connecting Rail, Maritime, Road and Air. Available online: https://sustainabledevelopment.un.org/content/documents/11694Thematicdiscussion8conceptnote.pdf (accessed on 1 February 2020).
9. Bricka, S.; Carlson, T.; Geiselbrecht, T.; Miller, K.; Moran, M. Shared Mobility Programs: Guidebook for Agencies. Available online: https://groups.tti.tamu.edu/communications/files/2016/10/Shared-Mobility-Guidebook-0-6818-P1.pdf (accessed on 20 March 2020).
10. Jiang, J. More Americans Are Using Ride-Hailing Apps. Available online: https://www.pewresearch.org/fact-tank/2019/01/04/more-americans-are-using-ride-hailing-apps/ (accessed on 11 March 2020).
11. Rayle, L.; Dai, D.; Chan, N.; Cervero, R.; Shaheen, S. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transp. Policy* **2016**, *45*, 168–178. [CrossRef]
12. Alemi, F. *What Makes Travelers Use Ridehailing? Exploring the Latent Constructs behind the Adoption and Frequency of Use of Ridehailing Services, and Their Impacts on the Use of Other Travel Modes*; University of California: Davis, CA, USA, 2018.
13. Furuhata, M.; Dessouky, M.; Ordóñez, F.; Brunet, M.; Wang, X.; Koenig, S. Ridesharing: The state-of-the-art and future directions. *Transp. Res. Part B* **2013**, *57*, 28–46. [CrossRef]
14. Chu, T. America's Transportation History is Full of Mistakes. Let's Not Make Another One. Available online: https://www.citylab.com/perspective/2019/09/transportation-future-mobility-technology-regulations-data/597748/ (accessed on 20 March 2020).
15. Clewlow, R.R.; Mishra, G.S. *Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States*; UC Davis Institute of Transportation Studies: Davis, CA, USA, 2017.
16. Rodier, C. *The Effects of Ride Hailing Services on Travel and Associated Greenhouse Gas Emissions*; UC Davis Institute of Transportation Studies: Davis, CA, USA, 2018.
17. Henao, A.; Marshall, W.E. The impact of ride-hailing on vehicle miles traveled. *Transportation* **2019**, *46*, 2173–2194. [CrossRef]
18. Feigon, S.; Murphy, C. Shared Mobility and the Transformation of Public Transit. Available online: http://www.trb.org/Publications/Blurbs/174653.aspx (accessed on 10 February 2020).
19. Hampshire, R.C.; Simek, C.; Fabusuyi, T.; Di, X.; Chen, X. Measuring the Impact of an Unanticipated Disruption of Uber/Lyft in Austin, TX. *Soc. Sci. Res. Netw.* **2017**. [CrossRef]
20. Hall, J.D.; Palsson, C.; Price, J. Is Uber a substitute or complement for public transit? *J. Urban. Econ.* **2018**, *108*, 36–50. [CrossRef]
21. Boisjoly, G.; Grisé, E.; Maguire, M.; Veillette, M.; Deboosere, R.; Berrebi, E.; El-Geneidy, A. Invest in the ride: A 14 year longitudinal analysis of the determinants of public transport ridership in 25 North American cities. *Transp. Res. Part A* **2018**, *116*, 434–445. [CrossRef]
22. Graehler, M.J.; Mucci, R.A.; Erhardt, G.D. Understanding the recent transit ridership decline in major US cities: Service cuts or emerging modes? In Proceedings of the Transportation Research Board Annual Meeting, Washington, DC, USA, 13–17 January 2019.
23. APA Dictionary of Psychology. Available online: http://dictionary.apa.org/self-selection-bias (accessed on 10 February 2020).

24. Winship, C.; Morgan, S.L. The estimation of causal effects from observational data. *Annu. Rev. Sociol.* **1999**, *25*, 659–707. [CrossRef]
25. 2017 National Household Travel Survey. Available online: https://nhts.ornl.gov (accessed on 10 February 2020).
26. Anderson, M. Who Relies on Public Transit in the U.S. Available online: https://www.pewresearch.org/fact-tank/2016/04/07/who-relies-on-public-transit-in-the-u-s/ (accessed on 5 February 2020).
27. Uber Help, Requests from Underage Riders. Available online: https://help.uber.com/driving-and-delivering/article/requests-from-underage-riders---?nodeId=43b84de6-758b-489e-b088-7ee69c749ccd (accessed on 24 January 2020).
28. Lyft, Safety Policies. Available online: https://help.lyft.com/hc/en-us/articles/115012923127-Safety-policies (accessed on 10 February 2020).
29. Gelman, A.; Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*; Cambridge University Press: Cambridge, UK, 2006.
30. Piquero, A.R.; Weisburd, D. *Handbook of Quantitative Criminology*; Springer: New York, NY, USA, 2010.
31. Mishra, G.S.; Clewlow, R.R.; Mokhtarian, P.L.; Widaman, K.F. Research in Transportation Economics the effect of carsharing on vehicle holdings and travel behavior: A propensity score and causal mediation analysis of the San Francisco Bay Area. *Res. Transp. Econ.* **2015**, *52*, 46–55. [CrossRef]
32. Rosenbaum, P.R.; Rubin, D.B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* **1983**, *70*, 41–55. [CrossRef]
33. Ho, D.E.; Imai, K.; King, G.; Stuart, E.A. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Polit. Anal.* **2007**, *15*, 199–236. [CrossRef]
34. Claritas, Assessing the Role of Urbanicity. Available online: https://nhts.ornl.gov/assets/Assessing_the_Role_of_Urbanicity.pdf (accessed on 5 February 2020).
35. Center, P.R. The Generations Defined. Available online: https://www.pewresearch.org/fact-tank/2018/04/11/millennials-largest-generation-us-labor-force/ft_18-04-02_generationsdefined2017_working-age/ (accessed on 2 February 2020).
36. NHTS, Main Study Retrieval Questionnaire. Available online: https://nhts.ornl.gov/2016/pub/NHTS_Retrieval_Instrument_20161006.pdf (accessed on 10 February 2020).
37. Coxe, S.; West, S.G.; Aiken, L.S. The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives. *J. Personal. Assess.* **2009**, *91*, 121–136. [CrossRef] [PubMed]
38. Long, J.S.; Freese, J. *Regression Models for Categorical Dependent Variables using Stata*; Stata Press: College Station, TX, USA, 2006.
39. Marshall, W.E.; Ferenchak, N.N. Why cities with high bicycling rates are safer for all road users. *J. Transp. Health* **2018**, *13*, 285–301. [CrossRef]
40. Zahran, S.; Brody, S.D.; Maghelal, P.; Prelog, A.; Lacy, M. Cycling and walking: Explaining the spatial distribution of healthy modes of transportation in the United States. *Transp. Res. Part D* **2008**, *13*, 461–469. [CrossRef]
41. Wang, X.; Lindsey, G.; Hankey, S.; Hoff, K. Estimating Mixed-Mode Urban Trail Traffic Using Negative Binomial Regression Models. *J. Urban Plan. Dev.* **2014**, *140*, 1–9. [CrossRef]
42. Hu, S.; Li, C.; Lee, C. Journal of the Chinese Institute of Engineers Model crash frequency at highway-railroad grade crossings using negative binomial regression. *J. Chin. Inst. Eng.* **2012**, *35*, 841–852. [CrossRef]
43. Cao, X.; Handy, S.L.; Mokhtarian, P.L. The influences of the built environment and residential self-selection on pedestrian behavior: Evidence from Austin, TX. *Transportation* **2006**, *33*, 1–20. [CrossRef]
44. Zhao, Y.; Kockelman, K.M. Household vehicle ownership by vehicle type: Application of a multivariate negative binomial model. In Proceedings of the Transportation Research Board's 81st Annual Meeting, Washington, DC, USA, 13–17 January 2002.
45. Young, M.; Lachapelle, U. Transportation behaviours of the growing Canadian single-person households. *Transp. Policy* **2016**, *57*, 41–50. [CrossRef]
46. Harris, T.; Hardin, J.W.; Yang, Z. Modeling underdispersed count data with generalized Poisson regression. *Stata J.* **2012**, *12*, 736–747. [CrossRef]
47. Giuffrè, O.; Granà, A.; Roberta, M.; Corriere, F. Handling Underdispersion in Calibrating Safety Performance Function at Urban, Four-Leg, Signalized Intersections. *J. Transp. Saf. Secur.* **2011**, *3*, 174–188. [CrossRef]

48. Wilson, S.R.; Leonard, R.D.; Edwards, D.J.; Swieringa, K.A.; Underwood, M. Inference for Under-Dispersed Data: Assessing the Performance of an Airborne Spacing Algorithm. Available online: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20190027465.pdf (accessed on 10 February 2020).
49. Hoef, J.M.V.; Boveng, P.L. Quasi-poisson vs. negative binomial regression: How should we model overdispersed count data? *Publ. Agencies Staff U.S. Dep. Commer.* **2007**, *142*, 2766–2772.
50. Lachapelle, U.; Noland, R.B. Does the commute mode affect the frequency of walking behavior? The public transit link. *Transp. Policy* **2012**, *21*, 26–36. [CrossRef]
51. Lachapelle, U. Walk, Bicycle, and Transit Trips of Transit-Dependent and Choice Riders in the 2009 United States National Household Travel Survey. *J. Phys. Act. Heal.* **2015**, *12*, 1139–1147. [CrossRef] [PubMed]
52. Bhat, C.R.; Pulugurta, V.A. Comparison of Two Alternative Behavioral Choice Mechanisms for Houshold Auto Ownership Decisions. *Transp. Res. Part B Methodol.* **1998**, *32*, 61–75. [CrossRef]
53. Chen, N. *How Do Socio-Demographics and The Built Environment Affect Individual Accessibility Based on Activity Space as A Transport Exclusion Indicator?* Ohio State University: Columbus, OH, USA, 2016.
54. Ho, D.E.; Imai, K.; King, G.; Stuart, E.A. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *J. Stat. Softw.* **2011**, *42*, 1–28. [CrossRef]
55. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S*, 4th ed.; Springer: New York, NY, USA, 2002.
56. R Core Team. R: A Language and Environment for Statistical Computing. In *R Foundation for Statistical Computing*; R Core Team: Vienna, Austria, 2013; ISBN 3-900051-07-0.
57. Liu, L.; Jiang, C.; Zhou, S.; Liu, K.; Du, F. Impact of public bus system on spatial burglary patterns in a Chinese urban context. *Appl. Geogr.* **2017**, *89*, 142–149. [CrossRef]
58. Highly Correlated Predictors. Available online: https://online.stat.psu.edu/stat462/node/179/ (accessed on 10 February 2020).
59. Allison, P. When Can You Safely Ignore Multicolliniearity. Available online: https://statisticalhorizons.com/multicollinearity (accessed on 10 February 2020).
60. Fox, J.; Weisberg, S. *An R Companion to Applied Regression*, 3th ed.; Sage: Thousand Oaks, CA, USA, 2019.
61. Negative Binomial Regression|Stata Annotated Output. Available online: https://stats.idre.ucla.edu/stata/output/negative-binomial-regression/ (accessed on 10 February 2020).
62. What Are the pseudo R-Squareds? Available online: https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-what-are-pseudo-r-squareds/ (accessed on 10 February 2020).
63. Pseudo R-Squared Measures. Available online: https://www.ibm.com/support/knowledgecenter/SSLVMB_24.0.0/spss/tutorials/plum_germcr_rsquare.html (accessed on 10 February 2020).
64. Signorell, A. DescTools: Tools for Descriptive Statistics. Available online: https://cran.r-project.org/package=DescTools (accessed on 10 February 2020).
65. McFadden, D. *Quantitative Methods for Analyzing Travel Behavior of Individuals: Some Recent Developments*; University of California: Berkeley, CA, USA, 1977.
66. Jaffe, E. Uber and Public Transit are Trying to Get Along. Available online: https://www.citylab.com/solutions/2015/08/uber-and-public-transit-are-trying-to-get-along/400283/ (accessed on 2 February 2020).
67. Building the Future of Public Transit Together. Available online: https://www.uber.com/us/en/transit/agency/ (accessed on 2 February 2020).
68. Zhang, Y.; Zhang, Y. Exploring the Relationship between Ridesharing and Public Transit Use in the United States. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1763. [CrossRef]
69. Conway, M.W.; Salon, D.; King, D.A. Trends in Taxi Use and the Advent of Ridehailing, 1995–2017: Evidence from the US National Household Travel Survey. *Urban Sci.* **2018**, *2*, 79. [CrossRef]
70. Sikder, S. Who Uses Ride-Hailing Services in the United States? *Transp. Res. Rec. J. Transp. Res. Board* **2019**. [CrossRef]