SELF-DRIVING CARS AND THEIR IMPACTS ON AMERICANS' LONG-DISTANCE DOMESTIC TRAVEL PATTERNS

Fatemeh Fakhrmoosavi, Ph.D.
Postdoctoral Fellow
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
moosavi@austin.utexas.edu

Priyanka Paithankar
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
priyanka.paithankar@utexas.edu

Kara M. Kockelman, Ph.D., P.E.
(Corresponding Author)
Dewitt Greer Professor in Engineering
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
301 E. Dean Keeton St, Stop C1761, Austin, TX, 78712
kkockelm@mail.utexas.edu
Tel: 512-471-0210

Yantao Huang, Ph.D.
Postdoctoral Researcher
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
yantao.huang@austin.utexas.edu

Jason Hawkins, Ph.D.
Assistant Professor
Department of Civil and Environmental Engineering
University of Nebraska Lincoln
jason.hawkins@unl.edu

Published in *Transportation Planning and Technology in*
November 2023

ABSTRACT

This research estimated models for long-distance domestic passenger trips before and after the introduction of autonomous vehicles (AVs) and their application to a 10% synthetic US population. The authors synthesized 12.1M households and 28.1M individuals across 73,056 US census tracts. To generate disaggregated passenger trips, travel demand models, including trip frequency, season, purpose, party size, mode choice; destination choice models; and vehicle ownership models were estimated. Different datasets, including a 2021 long-distance AV survey, 2016/17 National Household Travel Survey (NHTS) survey, EPA Smart Location data, FHWA rJourney dataset, and a 2017 AV fleet survey, were used for model estimation. The model applications indicated 0.85 vehicles per capita for 2019, which is consistent with the vehicle per capita of 0.83 in 2020 based on US census data. AV ownership is likely to be 0.33 per capita after the introduction of AVs within the marketplace with a $3500 AV technology cost premium in the year 2040. Assuming a $3500 technology cost premium (e.g., in 2040), total person-miles traveled per capita in long-distance trips is estimated to rise 35% (from 280 to 379 miles per month). The results of this study provide insights into how future long-distance travel patterns will change after AVs are in the market on a large scale.

Keywords: Long-distance Travel, Self-driving Vehicles, Travel Demand Modeling, Mode Shift, Destination Choice

MOTIVATION

Long-distance (LD) trips constitute an important part of Americans' intercity travels, with over 7 billion long-distance person-trips (weighted) over 75 miles in 2017 (FHWA 2017). Although these trips are a small portion of all trip counts by Americans (i.e., short- and long-distance), which was reported to be 371 billion person-trips in 2017, they represent almost 50% of the person-miles traveled (PMT) in the US (McGuckin 2018). Thus, LD trips play an important role in traffic congestion and economic growth and it is critical to understand their patterns to control congestion (LaMondia et al. 2016a). Most LD trips are on the ground, especially trips shorter than 750 miles. Autonomous vehicles (AV) and shared autonomous vehicles (SAVs) can raise this share, as these vehicles make driving easier and provide the option of traveling by car for some non-drivers. Most LD efforts have been dedicated to intraregional models (Childress et al. 2015, Harper et al. 2016) and no interregional study has considered the impacts of AVs on users' destination and mode choice, which is the focus of this study.

Most travel demand forecasting models are developed based on models developed for urban intraregional trips, which are different from LD intercity travels in that intraregional trips are differentiated based on homebased vs non-homebased trips, while LD trips are more likely categorized based on trip purpose, such as personal or recreational versus business trips (Schiffer 2012). LD travel patterns contain many other factors, including travel time and cost (Childress et al. 2015), travel party size (Li et al. 2020), and trip duration and schedule (Li et al. 2020, LaMondia et al. 2016b). Household and person-level demographics, such as household income (Sandow and Westin 2010), traveler's age (Collia et al. 2003), education (Holz-Rau et al. 2014), and number of children (LaMondia 2016b), are other important factors reported in previous studies.

The advent of AVs in the market is likely to boost LD passenger travels across the US in the coming years. As mentioned earlier, most studies related to AVs focus on intracity trips. Using an LD survey in Michigan, LaMondia et al. (2016b) investigated the impacts of AVs on LD trip generation and mode choice by assuming lower VOTT for AV users relative to conventional non-autonomous vehicles and higher travel costs for AVs. They predicted the air mode to be the dominating mode for trips longer than 500 miles with a 43.6% share (70.9% of trips greater than 1000 miles). AVs were also anticipated to reduce the share of personal conventional vehicles and airplanes for LD trips shorter than 500 miles. Huang et al. (2019) studied passenger and freight mode splits before and after the introduction of AVs across the Texas Triangle megaregion. They estimated airline passenger travel to fall 82% in that region, as travelers switch to using AVs and SAVs instead. They also estimated that people will choose more distant locations, increasing the average Texas person-trip distance from 14 to 16 miles (using Year 2040 land use forecasts). Childress et al. (2015) investigated the impacts of AVs on travel patterns using an activity-based model for the Seattle, WA region. They made different assumptions about AVs' value of travel time (VOTT) and cost changes relative to conventional passenger vehicles to modify the travel demand model, which was developed for currently available modes. They predicted a rise in VMT considering roadway capacity improvements due to AVs. All aforementioned studies either focused on LD non-AV trips or intraregional AV trips. Harper et al. (2016) estimated a 14% increase in annual light-duty vehicle-miles travelled (VMT) for the US population of 19 years old and older when AVs are an option in the future. They only considered increases in VMT due to the driving option for non-drivers, the elderly, and people with medical conditions that restrict their driving, and did not estimate travel demand models to investigate the impact of AVs on users' mode and destination choices. Gurumurthy and Kockelman's (2020) stated preference survey results among 2588 Americans suggested that over 50% of the US passenger trips between 50 and 500 miles (one way) will be made in an AV or SAV in the future (when AV technology is ubiquitous, but human driving is still permitted). They also estimated a tripling in SAV mode share for such trips if the respondent's annual household income is between \$75,000 and \$120,000 (versus higher or lower income levels), and a 67% increase when it is a business trip (versus personal trip). Their study mostly focused on the willingness of respondents to ride-share after AVs are introduced to the market. Perrine et al. (2020) added AVs as a new mode to the FHWA rJourney mode and destination choice models. In a scenario with AV operating costs equal to 118% of those of traditional cars, they predicted a shift in destination choice by AVs towards longer-distance trips with personal cars (including AVs). They also estimated that the AV mode share would lead to a 53% loss in airline revenue. While adding AVs as a mode to estimate travelers' mode and destination choice in the presence of these vehicles is a useful strategy, the FHWA rJourney data used in their study for mode and destination choices was gathered in 2010, which is rather outdated for this purpose.

Recognizing the potential for such dramatic shifts in travel choices, this study forecasts the impacts of AVs on the destination and mode choices of long-distance passenger trips (over 75 miles one-way) within the US. The travel demand model of this study is composed of several sub-models for vehicle ownership, trip season, trip frequency, trip purpose, travel party size, mode choice, and destination choice. Each model is addressed separately herein. Different datasets were used to estimate these models, including an LD-AV survey (Huang et al. 2022), the 2016/17 National Household Travel Survey (NHTS) dataset, FHWA rJourney travel skim data, and EPA Smart Location dataset. The models were applied to a synthetic population consisting of 12.1M households and 28.1M individuals across 73,056 census tracts throughout the nation to estimate the shifts in travel caused by AVs in the market.

The remainder of this paper is organized as follows. The next section elaborates on the datasets used in this study. The third section explains the framework and methods used to estimate different travel demand models and the application of these models to the synthetic population to generate disaggregate LD trips. Then, the effective parameters in different models will be explained, and the projected impacts of AVs on Americans' LD domestic travel will be summarized, followed by conclusions and limitations of this study.

DATA DESCRIPTION

This study leveraged data from different sources to estimate travel demand models before and after AVs are introduced into the market. The main data source capturing the presence of AVs is an LD-AV survey conducted in 2021 to anticipate Americans' long-distance travel preferences when access to AVs is common. The survey contains responses from 1,004 U.S. respondents (45% residing in Texas and 55% in other US states) to revealed and stated preference questions about recent trips and future trip scenarios (Huang et al. 2022). Sample weights were generated using an iterative proportional fitting (IPF) method (Roth et al., 2017) to match the most recent five years of data from the American Community Survey (ACS).

The 2016/17 NHTS data, containing 924,000 trip observations (~15,000 long-distance over 75 miles) made by almost 130,000 households and 264,000 persons, was used to estimate trip frequencies, trip season, trip purposes, destination choice, and travel party size models. Trips longer than 75 miles (15,100 trips) were filtered from this dataset to estimate trip season, trip purpose, and destination choice models. This dataset contains trip, person, and household tables. The vehicle ownership model leverages the household table and trip models are estimated using the trip table. The person table was matched with the trip table to include individuals' demographics in different models. Sample weights reported in the NHTS data were used to match the sample with the entire US population. NHTS uses ACS data to create expansion factors to scale up survey data to 301 million persons (or 118 million households), making 371 billion person-trips (7 billion long-distance person-trips) every year (FHWA 2017).

Ground and air travel time and cost skims are required to estimate users' mode and destination choices. For this purpose, the FHWA's rJourney dataset was used, which contains a synthetic set of 1.17B long-distance tours by US households, estimated for the year 2010. Travel time and cost estimates in this dataset are across 4,477 National User Model Areas (NUMAs) in the US. NUMAs were generated by FHWA using counties and Census Bureau Public Use Microdata Areas (PUMAs) across the US. FHWA overlays PUMAs and counties and selects the smaller zones as NUMAs. In the United States, there are a total of 3,243 counties and 2,351 PUMAs. The EPA's Smart Location dataset was used for land use details at all 73,056 tract zones, including population density and counts and job counts at each tract zone.

To simulate US travel patterns, the research team synthesized 10% of the US population at the census tract level (73,056 census tracts across the US). The synthesized population is based on marginals from 5-year ACS data in 2019 (for the period between 2015 and 2019), using PopGen 2.0 software developed by Pendyala et al. (2011) and Ye et al. (2009). The household and person data were synthesized across 2,351 PUMAs, to mimic the population distribution of the US (including 50 states and the District of Columbia), consistent with census datasets and geographic-correspondence files. Note that PUMAs are statistical geographic areas defined by the US Census Bureau. These areas are designed to partition each state or equivalent entity into non-overlapping zones, each containing no fewer than 100,000 people. They are primarily utilized for the tabulation and dissemination of decennial census and American Community Survey (ACS) Public Use Microdata Sample (PUMS) data. The authors used the datasets described in this section for estimating travel demand models and model applications. Table 1 summarizes the statistics of the synthesized population (used for model applications) and the 2016/17 NHTS data (used for travel demand model estimations before AVs).

Table 1. Summary Statistics of Synthesized Population (10% Sample of 2019 US Population with 28.1M Persons and 12.1M Households) and 2016/17 NHTS Data (264,000 Persons and 130,000 Households)

Variable	Category	2019 Synthetic Population	2016/17 NHTS
PERSON			
Sex	Male	47.43%	49.07%
	Female	52.56%	50.93%
Race	White	73.54%	72.49%
	Black or African American	12.23%	12.71%
	Asian	5.38%	5.33%
	American Indian or Alaska Native	0.76%	0.86%
	Native Hawaiian/Pacific Islander	0.16%	0.28%
	Multiple responses selected	3.19%	3.96%
	Some other race	4.73%	4.37%
Education	High school graduate or GED	52.71%	33.51%
	Some college or associate degree	23.91%	28.56%
	Bachelor’s degree	14.72%	21.02%
	Graduate or professional degree	8.66%	16.90%
Age	Younger than 10 years old	11.99%	8.37%
	11–17 years old	10.10%	9.69%
	18–24 years old	8.49%	10.37%
	25–34 years old	13.79%	14.07%
	35–44 years old	12.80%	14.03%
	45–54 years old	13.34%	13.43%
	55–64 years old	13.29%	14.45%
	65–74 years old	9.44%	10.05%
	75 years or older	6.75%	5.12%
HOUSEHOLD			
Household Size	1-person HH	27.86%	27.88%
	2 persons in HH	33.93%	33.88%
	3 persons in HH	15.59%	15.67%
	4 persons in HH	12.90%	14.33%
	5 persons in HH	5.97%	5.42%
	6 persons in HH	2.30%	1.93%
	7 or more persons in HH	1.44%	0.89%
	Less than $10,000	5.87%	7.51%
	$10,000–$14,999	4.33%	6.02%

Variable	Category	2019 Synthetic Population	2016/17 NHTS
Annual Household Income	\$15,000–\$24,999	8.95%	9.78%
	\$25,000–\$34,999	8.97%	10.01%
	\$35,000–\$49,999	12.30%	12.37%
	\$50,000–\$74,999	17.26%	16.54%
	\$75,000–\$99,999	12.77%	12.30%
	\$100,000–\$124,999	9.17%	9.38%
	\$125,000–\$149,999	6.07%	5.35%
	\$150,000–\$199,999	6.84%	5.22%
	\$200,000 or more	7.49%	5.50%
# Children	0 children	70.60%	69.92%
	1 child	9.69%	12.13%
	2 children	11.96%	12.29%
	3 children	5.18%	3.94%
	4 children	1.84%	1.22%
	5 or more children	0.72%	4.93%

MODELING FRAMEWORK AND METHODS

To investigate the impacts of AVs on travelers' long-distance trips, this study generates disaggregate trips for the 10% synthetic population before and after AVs are available in the market. Figure 1 illustrates the datasets and steps to generate trips using a synthetic population. This figure shows the sequence of models required for generating trips and distributing them among different destinations and modes. Pre-trip models include the decision to participate in long-distance travel and departure time season, purpose, and frequency over the course of a year. Then, destination and mode choice models should be estimated, with mode choice conditioned on household vehicle ownership decisions and destination choice conditioned on the accessibility term (i.e., mode choice logsum). Party size for each tour should also be estimated before mode choice estimations.

The number of long-distance trips per day was estimated at the individual level using a zero-inflated negative binomial (ZINB) model and the 2016/17 NHTS data. Based on the 2020 AV survey results and prior studies (Huang et al., 2020), it is assumed that trip frequency will increase by 15% after AVs are in the market. Population weights are applied to all models to ensure that parameter estimates better reflect the US household- and person-level populations. A multinomial logit model was used to estimate trip purpose and season models. Purposes include regular home-to-work "commute" trips (9%), as well as work-related business trips (7%), shopping excursions (18%), personal business (11%), religious/community trips (1%), school/daycare trips (1%), medical/dental trips (4%), trips made to visit friends and family (19%), social leisure trips (28%), and other purposes (1%). The party size model also uses a negative binomial specification with the 2016/17 NHTS data set to predict the number of individuals in a trip "party," including non-household members.

Mode choice relies on a joint revealed and stated preference multinomial logit model and the 2021 LD-AV survey data, for all available modes (which vary by household, due to vehicle ownership decisions, and individual preferences). Survey respondents were asked to recall a recent long-distance trip and whether they would be willing to replace the mode used for that trip with AVs, provided the AVs are available with the same travel time as human-driven vehicles. The mode choices before and after the introduction of AVs were used in a joint revealed and stated preference mode choice model. A Poisson model was employed to estimate auto ownership before the advent of AVs. Quarles et al.'s (2021) AV ownership simulation approach was used to predict AV ownership in the future. Their approach estimates households' willingness-to-pay (WTP) for AVs, where all capabilities found in today's human-driven vehicles are maintained in all fully autonomous vehicles, and compares WTP to the technology price in each target year.

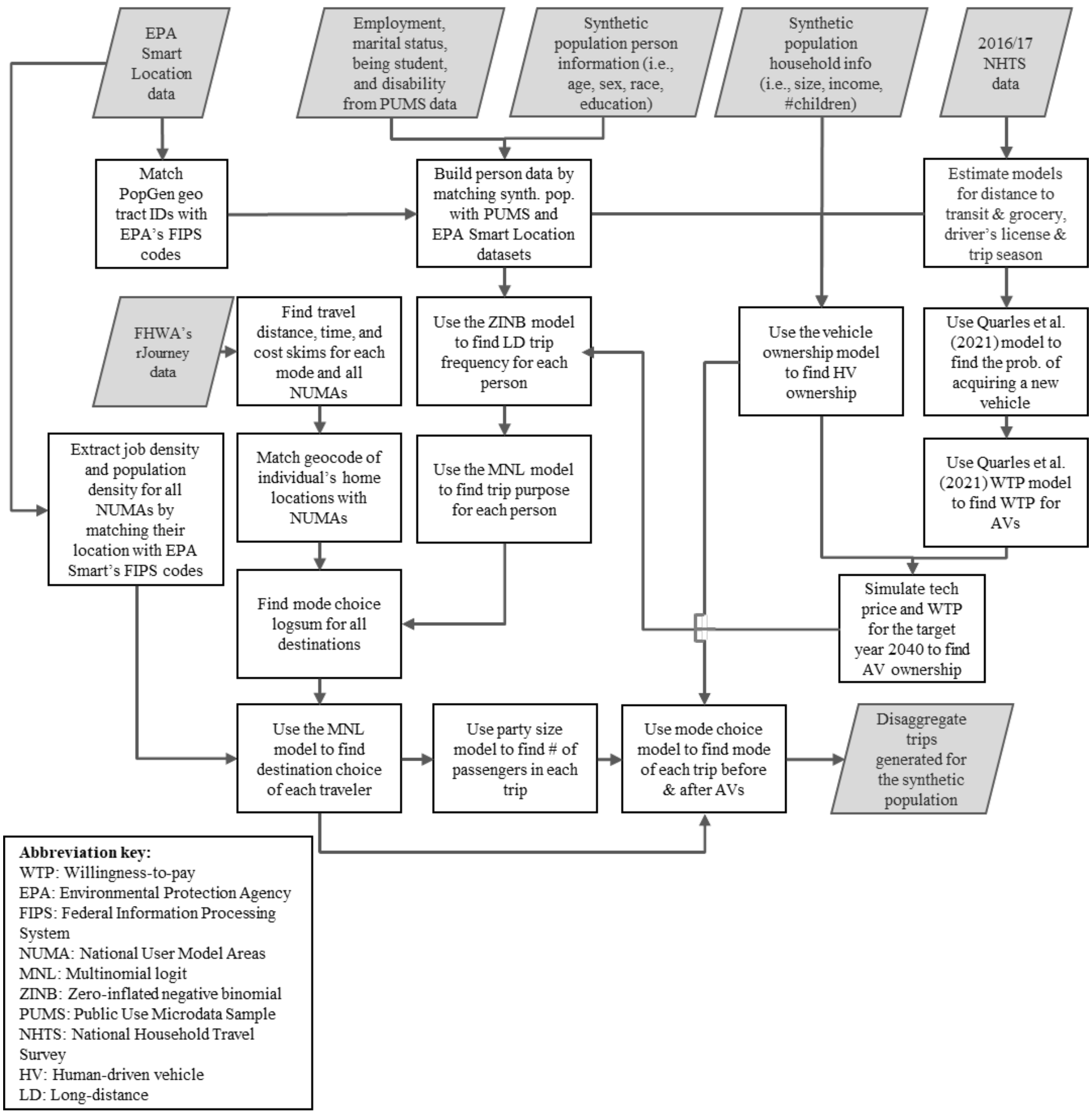


Figure 1. Steps for Applying Travel Demand Models to the Synthetic Population to Generate Disaggregate Trips Before and After AVs

Multinomial logit models were also used to predict destination choice of domestic trips considering NUMAs as different destination zones, which are finer than US counties in the nation's heavily populated big-county regions (like southern California). The destination choice model calibration process tested controls for attraction details (i.e., the logarithm of different job-type counts summed over each destination tract, logsum over mode choice utilities, and population density), in two distinct model equations for business and personal trips. Land use data were extracted from the EPA Smart Location data by mapping NUMA zones to US tracts' Federal Information Processing System (FIPS) codes. The FHWA rJourney travel time and cost skims were used for the mode choice logsum estimates, which are the accessibility terms. Given the very large destination-choice set, Lemp and Kockelman's (2012) strategic sampling approach was used for tractability and reasonable computing time. Strategic sampling for large-set estimation relies on a simple upstream choice logit model with 299 destination alternatives chosen randomly (out of 4477 NUMA zones), alongside the actual chosen zone. A special probability-adjusted logit model is used to draw the 300 alternatives in proportion to initial choice-probability estimates. The sensitivity of this strategic sampling approach to the number of sample alternatives is investigated by Lemp and Kockelman (2012). The travel demand models were run in sequence for the synthesized household and person data and the models were validated by comparing the estimated long-distance trip frequency, purpose, season, party size, modes, and destinations before AVs with those of the NHTS dataset. Then, these specifications were estimated for future scenarios when AVs are readily available.

RESULTS AND DISCUSSIONS

This section summarizes travel demand models and the model application results before and after AVs are readily available. Table 2 illustrates the practically and statistically significant variables in the long-distance trip frequency model, along with the impacts for a one standard deviation increase in each covariate (as a measure of practical significance). Results of the ZINB model for long-distance trip frequencies suggest that shifting the population-weighted sample toward the male gender by 1 SD increased the sample's average long-distance trip frequency by 21.6%. A 1 SD increase in households' vehicles increased long-distance trip-making rates by 21.6%. Shifting the sample toward having at least an associate degree by 1 SD also increased trip frequency by 24%.

Table 3 presents the coefficient estimates of the multinomial logit model to estimate trip season with summertime travel as the base alternative. The probability of taking a long-distance journey in the spring rises with age. Fall is a less popular time for long-distance trips among adults than summer. Table 4 presents the coefficient estimates of the multinomial logit trip purpose models with 10 alternatives, keeping the commute trips as the base. The trip purposes considered in the model include commute (9%), business (7%), shopping (18%), personal business (11%), school (1%), medical/dental (4%), religious or community (1%), visits to friends and relatives (19%), social leisure (28%), and other purposes (1%). The purpose model predicted that as household income increases, the probability of making long-distance business trips and personal trips (except medical/dental trips) increases as compared to daily long-distance work (commute) trips. There is a high probability of making business trips in the spring and fall seasons. With an increase in age, individuals tend to make more medical/dental, business, shopping, religious, and other social leisure trips than commute and school trips.

Table 2. ZINB Model for Long-distance Trip Frequency Using 2016/17 NHTS Household Data

Negative binomial (NB) model coefficients				
Variable	**Estimate**	**t-stat**	**P-value**	**Pract. Sign.**
(Intercept)	0.799	3.62	0.000	-
Male	0.172	7.85	0.000	0.216
Age	-0.002	-3.52	0.000	-0.099
Ln (HH income) ($)	-0.079	-2.72	0.006	0.507
Education associate degree or higher	0.191	6.84	0.000	0.216
#Adults	-0.228	-14.71	0.000	-0.460
Worker	-0.080	-3.95	0.000	-0.077
HH vehicle count	0.141	12.40	0.000	0.657
ln(θ)	15.45	6.44	0.017	-
Zero-inflation (ZI) model coefficients				
Variable	**Estimate**	**t-stat**	**P-value**	**Pract. Sign.**
(Intercept)	7.125	31.49	0.000	-
Ln (HH income) ($)	-0.043	-4.04	0.000	0.507
HH vehicle count	-0.410	-19.80	0.000	0.657

n = 201,820, Pseudo-R^2 = 0.015

Table 3. Coefficient Estimates of the MNL Model for Trip Season (Base Season: Summer)

	Fall Trip			Winter Trip			Spring Trip		
	Estimate	**t-Stat**	**P-value**	**Estimate**	**t-Stat**	**P-value**	**Estimate**	**t-Stat**	**P-value**
(Intercept)	0.034	0.341	0.733	-0.630	-6.92	0.000	-0.828	-6.55	0.000
Male	0.270	6.16	0.000	0.270	6.16	0.000	0.270	6.16	0.000
Age	-	-	-	-	-	-	0.010	7.55	0.000
College Educated or Higher	0.167	2.49	0.013	0.217	3.07	0.002	0.117	1.775	0.076
Income ($1000)	0.001	1.45	0.147	-	-	-	-	-	-
HH Size	-0.097	-5.03	0.000	-0.097	-5.03	0.000	-0.097	-5.03	0.000
#Vehicle Owned	0.091	4.88	0.000	0.091	4.88	0.000	0.091	4.88	0.000
Employed?	-0.250	-5.56	0.000	-	-	-	-0.250	-5.56	0.000
#Adults	-0.113	-3.54	0.000	-	-	-	0.084	2.73	0.006

n = 10,455, Adj. Rho^2: 0.0013

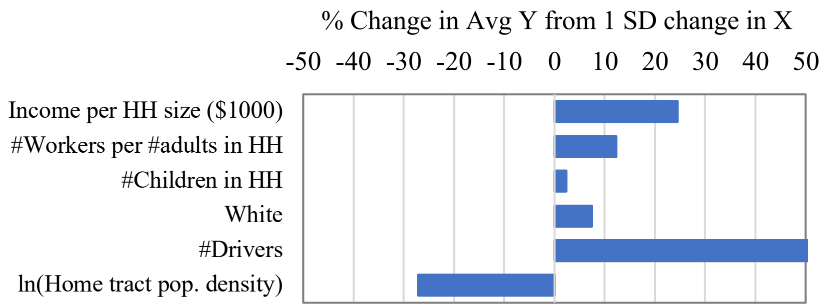
Table 4. Coefficient Estimates of the MNL Model for Trip Purposes (Base Purpose: Commute)

	Business	**Shop**	**Other Personal**	**School**	**Medical & Dental**	**Religious**	**Visit friends/ relatives**	**Social leisure**	**Other**
Intercept	−0.543*	2.916***	2.498***	2.051***	−0.156	−1.665***	2.806***	3.237***	−11.123***
Worker?	-	−2.178***	−1.870***	−3.997***	−3.244***	−2.013***	−2.131***	−2.392***	-
Age	0.012***	0.007***	0.013***	−0.130***	0.041***	0.020***	0.005***	-	0.093***
Male?	-	−0.499***	−0.658***	-	−0.197	-	−0.731***	−0.622***	-
Fall Trip?	0.738***	-	−0.247**	1.018***	0.202	-	0.337***	-	-
Winter trip?	-	−0.602***	−0.556***	−0.567***	-	-	-	−0.616***	-
Spring trip?	0.683***	−0.374***	−0.679***	-	-	-	-	−0.663***	3.172***
Associate degree or higher?	0.422***	0.279***	-	1.980***	-	-	0.358***	0.391***	-
HH size	−0.074*	−0.126	−0.103**	-	-	−0.106	−0.205***	-	-
#Adults	−0.858***	−0.436***	−0.188**	-	-	-	-	−0.419***	-
HH income ($1000)	0.014***	0.007***	0.006***	0.016***	−0.018***	0.009***	0.007***	0.009***	0.022***
White?	-	0.273***	-	−0.548*	-	-	-	0.396***	-
#Vehicle	−0.101**	-	−0.115***	−0.255**	-	-	−0.202***	-	−0.990***

n = 11,414 & Pseudo R^2 = 0.2501, *0.01 to 0.1, **0.001 to 0.01, ***0.000

Figure 2 illustrates the practical significance of all statistically significant variables in the vehicle ownership (Figure 2a) and party size (Figure 2b) models. Figure 2a indicates that a 1-SD change in each household's income or the number of workers per adult in the household increase predicted vehicle ownership counts by 24% and 12%, respectively. A 1-SD rise in the population density (logged) of the census tract of the household home location reduces this ownership by about 27%. Increasing the number of drivers in a household by 1-SD increases vehicle ownership by more than 80%. The average model-predicted number of passengers in a long-distance travel party falls by 25% when the commute-purpose variable rises by 1 SD, and by 19% when the business trip indicator rises by 1 SD. A 1 SD increase in the female gender indicator increases party size by 11%.

a)



b)

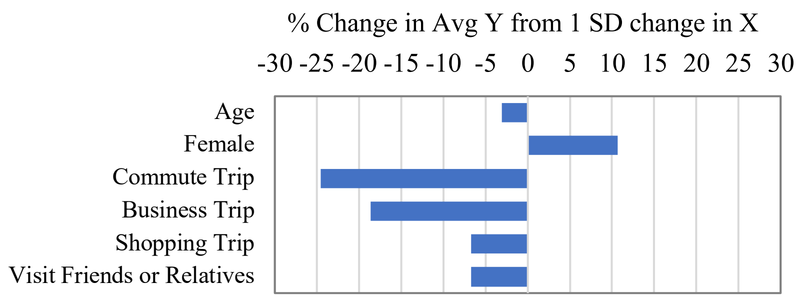


Figure 2. Impacts of Statistically Significant Covariates on a) Vehicle Ownership, b) Trip Party Size (% Average Change in Predicted Y Following a 1 SD Increase in the X Covariate)

Mode and destination choice models for long-distance domestic trips were estimated for business and non-business trips in a joint model before and after AVs become available using the 2021 long-distance AV survey.Uncommon existing long-distance modes (including bus, rail, and boats) were not included, so only air, rental car, personal car, and AVs were permitted. To consider chain trips, we summed the time and costs of all legs of trips. The specifications of the joint revealed and stated preference logit models for non-business trips with AVs are presented in Table 5. The operational cost of AVs was considered $0.70 per mile. The operational cost of human-driven personal vehicles was assumed $0.50 per mile, while the cost of a rental car was $50 per driving day (minimum 1 day) in addition to $0.10 per mile. To avoid the correlation between travel costs and times, the residuals of travel costs from travel times were considered in the mode choice models. The specifications of the mode choice model for non-business trips show that users are more willing to use airplanes for trips longer than 500 miles. In addition, AVs have an inverse relationship with age and a direct relation with having at least a college degree. Rental cars have a higher utility for trips with higher party sizes. Due to the low number of observations in the survey for business trips, the non-business model was adjusted by lowering the impact of cost in these trips' mode choice.

Table 5. Specifications of the Logit Mode Choice Model After AVs Using Joint Stated Preference and Revealed Preference LD-AV Survey Data, EPA Smart, and RSG rJourney Data

	Estimate	**t-ratio**	**P-value**

ASC car	0	-	-
ASC air	-1.187	-7.464	0.000
ASC rental car	-0.710	-10.803	0.000
ASC AV	-0.090	-0.291	0.385
Travel time x car	-0.281	-5.469	0.000
Travel time x air	-0.270	-2.282	0.011
Travel time x rental car	-0.103	-3.618	0.000
Travel time x AV	-0.113	-4.815	0.000
Access/egress distance x air	-0.028	-3.666	0.000
Residual of cost from travel time	-0.002	-3.777	0.000
Long-distance>500 mi x air	1.914	4.120	0.000
Party size rental x car	0.129	2.591	0.005
Female x car	-0.207	-1.336	0.091
Age x AV	-0.023	-3.472	0.000
Associate degree x AV	0.725	2.459	0.007
μ revealed preference	1.000	-	-
μ stated preference	0. 752	11.398	0.000

n = 584, R-squared: 0.3513

The destination choice models with the strategic sampling of 300 alternatives are presented in Table 6. The results of the destination choice model suggest that the number of retail, industrial, service, public administration, and medical jobs at the destination tract are important contributors to business and non-business trips. The utility of destination rises when the accessibility term and/or the population density increases at the destination's tract.

Table 6. Destination Choice Model Specifications Using 2016/17 NHTS, EPA Smart Location, and rJourney Data

	Non-business Trips			Business Trips		
	Estimate	**t-Stat**	**P-Value**	**Estimate**	**t-Stat**	**P-Value**
Mode choice logsum	0.017	122.96	0.000	0.011	50.49	0.000
Destination population density at the tract level (persons/sq mi)	0.002	1.60	0.109	0.005	2.61	0.009
#Retail jobs in tract	-0.068	-8.62	0.000	-0.049	-2.38	0.017
#Industrial jobs in tract	0.027	3.20	0.001	0.021	1.04	0.297
#Service jobs in tract	0.019	2.17	0.030	0.057	2.56	0.010
#Public administration jobs in tract	-0.019	-3.90	0.000	-	-	-
#Medical jobs in tract	-	-	-	-0.044	-2.81	0.005
	n = 9,325, Pseudo-R^2: 0.060			*n = 1802*, Pseudo-R^2: 0.060		

The application of the models presented in Tables 2-6 and Figure 2 to the 10% synthetic US population indicated 0.85 vehicles per capita in 2019, which is consistent with the vehicle per capita of 0.83 in 2020 based on the US census data. After AVs are in the market in the future (e.g., in the year 2040) with an AV technology premium of $3,500, 61% of households are estimated to have AVs. The model applications suggest that Americans conducted 2.00 long-distance trips per person per month in 2019. We also validated the outcomes of all models with their relevant datasets after the application to evaluate their performance. For instance, the application of the synthetic population to the trip season model resulted in the following trip distribution: 30% in summer, 28% in fall, 22% in winter trips, and 19% in spring. In comparison, the distribution of trips in the NHTS data is 31% summer, 25% fall, 20% winter, and 24% spring.

Based on the results of the 2020 AV survey and previous studies, such as Huang et al. (2020), it is assumed that AVs will increase trip counts by 15%. As shown in Figure 3, mode splits for long-distance, domestic trips prior to AV access were estimated as 64.1% by private automobile, 30.4% by rental car, and 5.5% by air. After AVs become available for purchase (with a premium cost of $3,500) and SAVs are available with $0.70/mile operation cost, mode splits shift to 31.7% by conventional human-driven vehicle, 23.0% by conventional rental car, 23.5% by AV, 18.2% by SAVs, and 3.5% by air. Figure 4 summarizes the results of the destination choice model for the synthetic population. Assuming a $3,500 AV technology cost premium in today's dollars in 2040, total person-miles traveled (PMT) per capita in long-distance trips is estimated to rise 35% (from 280 to 379 miles per month). For the same AV technology cost premium scenario, vehicle-miles traveled (VMT) in long-distance trips increases from 121 to 152 miles per capita per month as many travelers shift from air to cars and shorter trips.

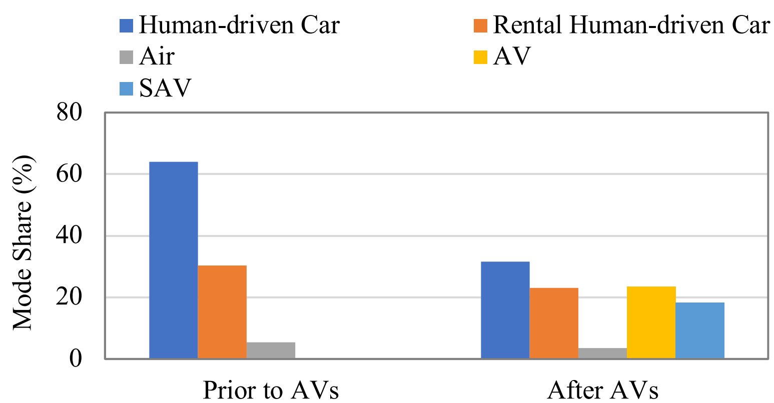


Figure 3. Mode Share Shift Before and After AVs Are in the Market with Technology Cost of $3500

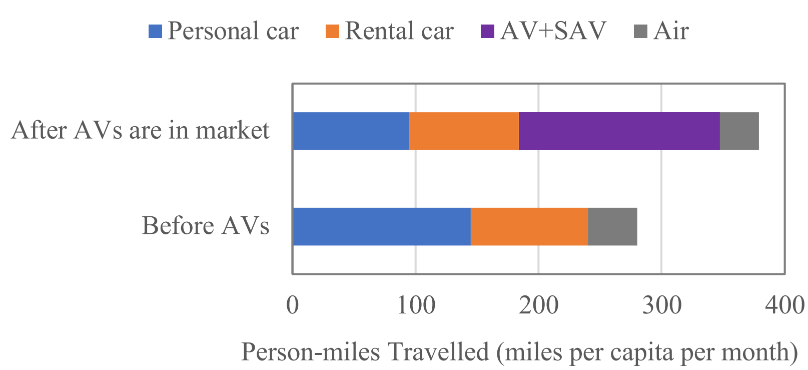


Figure 4. Shift in Person-miles Traveled (PMT) of Long-distance Trips after AVs are in the Market

SUMMARY AND CONCLUSIONS

This research forecasted the effects of automated cars on long-distance (over 75 miles one-way) domestic passenger travel frequency, destination, mode, party size, and scheduling inside the US. Different datasets were used to derive equations for such choices with and without AVs using Poisson, negative binomial, zero-inflated negative binomial distributions, and multinomial logit models. These estimations relied on the nation's PUMS with 2015-19 data (as released in 2019), a survey of 1,004 US respondents (45% residing in Texas and 55% in other US states) in 2021 to revealed and stated preference questions about recent trips and future trip scenarios (Huang et al. 2022), the 2016/17 NHTS, the EPA Smart Location data (for land use attributes at the tract level), and FHWA's rJourney data for long-distance passenger trips in 2010 (to extract travel time and cost skims). To simulate US long-distance domestic passenger travel, this study used synthesized household and person data and the set of estimated travel demand models for trip frequency, trip season, travel purpose, vehicle ownership, party size, mode choice, and destination choice models. The synthetic population is comprised of 28.1M persons in 12.1M households across 2,351 PUMAs, to mimic the nation's population distribution (across 50 states and the District of Columbia). The synthetic population is consistent with census datasets using the nation's 73,056 census tracts.

Model applications with the 10% US synthetic population suggested an average party size of 2.04 persons for long-distance trips, which is assumed to remain stable after the introduction of AVs. Vehicle ownership model application estimated 0.85 vehicles per capita for 2019, which is consistent with the vehicle per capita of 0.83 in 2020 based on the US census data. 2.00 LD trips over 75 miles per month per capita were estimated for the 10% synthetic population, which matches the NHTS data. Assuming a $3,500 technology cost premium (e.g., in the year 2040), total person-miles traveled per capita for existing long-distance trips are estimated to rise 35% (from 280 to 379 miles per month). The increase in person-miles traveled can be attributed to both an uptick in trip frequency and longer trip distances for ground trips, facilitated by the convenience of driving with AVs. It is important to note that as a result of the mode shift from air trips (decreasing from 5.5% to 3.5%) to ground trips following the adoption of AVs, the person-miles traveled for very long-distance trips have been reduced. The results of this study provide insights on how future long-distance travel patterns will change after AVs are in the market on a large scale.

This study utilized comprehensive data sources, including the 2016/17 NHTS dataset and a survey across the US, and synthesized a subset of the US population based on the ACS data to implement the models. Sampling standards were rigorously followed in both the survey and population synthesis. The inclusion of relevant variables, guided by both practical and statistical significance, further enhanced the predictability of the presented models. All models and results in this study are reproducible upon access to the datasets, which were obtained from various US agencies or gathered through a project that funded this study. For future research, the shift in trip destinations could be investigated using a stated preference survey, like the mode choice of this study. In addition, the potential impacts of AVs on international trips, especially to Canada and Mexico, should be investigated.

ACKNOWLEDGMENTS

The authors thank the Texas Department of Transportation (TxDOT) for financially supporting this research, under research project 0-7081, "Understanding the Impact of Autonomous Vehicles on Long Distance Travel Mode and Destination Choice in Texas". The authors also thank Kenneth Perrine for zone geocoding and Jade (Maizy) Jeong for editorial and submission support.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: Fakhrmoosavi, F., Kockelman, K.M., Huang, Y., Hawkins, J.; data collection: Huang, Y.; analysis and interpretation of results: Fakhrmoosavi, F., Paithankar, P., and Kockelman, K.M., Hawkins, J., Huang, Y.; draft manuscript preparation: Fakhrmoosavi, F., Paithankar, P., Kockelman, K.M., Hawkins, J; All authors reviewed the results and approved the final version of the manuscript.

DISCLOSURE STATEMENT

The authors report that there are no competing interests to declare in this paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the Federal Highway Administration (FHWA). Restrictions apply to the availability of these data, which were used under license for this study. Data are available from Dr. Kara M. Kockelman (kkockelm@mail.utexas.edu), Dr. Fatemeh Fakhrmoosavi, and Priyanka Paithankar after the permission of FHWA.

REFERENCES

Bureau of Transportation Statistics (BTS). 2017. *Long Distance Transportation Patterns: Mode Choice*. https://www.bts.gov/archive/publications/america_on_the_go/long_distance_transportation_patterns/entire.

Childress, S., B. Nichols, B. Charlton, and S. Coe. 2015. "Using an activity-based model to explore the potential impacts of automated vehicles." *Transportation Research Record*, *2493*(1), 99-106. DOI: 10.3141/2493-11.

Collia, D. V., J. Sharp, and L. Giesbrecht. 2003. "The 2001 national household travel survey: A look into the travel patterns of older Americans." *Journal of safety research*, *34*(4), 461-470. DOI: 10.1016/j.jsr.2003.10.001

Fagnant, D. J., and K. M. Kockelman. 2014. "The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios." *Transportation Research Part C: Emerging Technologies*, *40*, 1-13. DOI: 10.1016/j.trc.2013.12.001

FHWA. 2017. *National Household Travel Survey*, Federal Highway Administration. https://nhts.ornl.gov/

Gurumurthy, K.M., and K. M. Kockelman. 2020. "Modeling Americans' autonomous vehicle preferences: A focus on dynamic ride-sharing, privacy and long-distance mode choices." *Technological Forecasting and Social Change*, 150. DOI: 10.1016/j.techfore.2019.119792

Harper, C. D., C. T. Hendrickson, S. Mangones, and C. Samaras. 2016. "Estimating potential increases in travel with autonomous vehicles for the non-driving, elderly and people with travel-restrictive medical conditions." *Transportation research part C: emerging technologies*, *72*, 1-9. DOI: 10.1016/j.trc.2016.09.003

Holz-Rau, C., J. Scheiner, and K. Sicks. 2014. "Travel distances in daily travel and long-distance travel: what role is played by urban form?" *Environment and Planning A*, 46(2), 488-507. DOI: 10.1068/a4640

Huang, Y., and K. M. Kockelman. 2019. "What will autonomous trucking do to US trade flows? Application of the random-utility-based multi-regional input–output model." *Transportation*, *47*(5), 2529-2556. https://link.springer.com/content/pdf/10.1007/s11116-019-10027-5.pdf

Huang, Y., and K. M. Kockelman. 2022. "Synthetic US Population with PopGen2." https://sites.google.com/view/yhpopgen

Huang, Y., K. M. Kockelman, and N. Quarles. 2020. "How will self-driving vehicles affect US megaregion traffic? The case of the Texas triangle." *Research in Transportation Economics* 84: 101003. DOI: 10.1016/j.retrec.2020.101003

Huang, Y., N. Zuniga-Garcia, and K. M. Kockelman. 2022. "Long-distance travel impacts of automated vehicles: a survey of American households." Under review at for publication in *Transportation Research part A: Policy and Practice.* https://www.caee.utexas.edu/prof/kockelman/public_html/TRB22LD-AVUSSURVEY.pdf

LaMondia, J. J., M. Moore, and L. Aultman-Hall. 2016a. "Matching Household Life-Cycle Characteristics to Clustered Annual Schedules of Long-Distance and Overnight Travel." *Transportation Research Record*, *2594*(1), 11-17. DOI: 10.3141/2594-02

LaMondia, J. J., D. Fagnant, H. Qu, J. Barrett, and K. M. Kockelman. 2016b. "Shifts in long-distance travel mode due to automated vehicles: Statewide mode-shift simulation experiment and travel survey analysis." *Transportation Research Record*, 2566(1), 1-11. DOI: 10.3141/2566-0

Lemp, J.D. and K. M. Kockelman. 2012. "Strategic sampling for large choice sets in estimation and application." *Transportation Research Part A: Policy and Practice*, 46(3): 602-613. DOI: 10.1016/j.tra.2011.11.004

Li, R., K. M. Kockelman, and J. Lee. 2022. "Reducing Greenhouse Gas Emissions from Long-Distance Travel Business: How Far Can We Go?". *Transportation Research Record*, *2676*(1), 472-486. https://www.caee.utexas.edu/prof/kockelman/public_html/TRB20LDworktravel.pdf

McGuckin, Nancy. 2018. *Analysis brief: Can We Use the NHTS to Estimate Long-Distance Travel?* DOI: 10.13140/RG.2.2.23563.13607. URL: https://www.researchgate.net/profile/Nancy_Mcguckin/publication/329223174_ANALYSIS_BRIEF_Can_We_Use_the_NHTS_to_Estimate_Long-Distance_Travel/links/5bfdd8fa299bf1c2329e7b8d/ANALYSISBRIEF-Can-We-Use-the-NHTS-to-Estimate-Long-Distance-Travel.pdf

Perrine, K. A., K. M. Kockelman, and Y. Huang. 2020. "Anticipating long-distance travel shifts due to self-driving vehicles." *Journal of Transport Geography*, *82*, 102547. DOI: 10.1016/j.jtrangeo.2019.102547

Quarles, N., K. M. Kockelman, and J. Lee. 2021. "America's fleet evolution in an automated future." *Research in Transportation Economics* 90: 101107. DOI: 10.1016/j.retrec.2021.101107

Roth, S., Y. Dai, and J. DeMatteis. 2017. *2017 NHTS Weighting Report.* URL: https://nhts.ornl.gov/assets/2017%20NHTS%20Weighting%20Report.pdf

Sandow, E., and K. Westin. 2010. "The persevering commuter–Duration of long-distance commuting." *Transportation Research Part A: Policy and Practice*, 44(6), 433-445. DOI: 10.1016/j.tra.2010.03.017

Schiffer, R. G. 2012. *Long-distance and rural travel transferable parameters for statewide travel forecasting models (Vol. 735)*. Transportation Research Board. https://link.springer.com/content/pdf/10.1007/978-3-319-05990-7.pdf