ANTICIPATING LONG-DISTANCE TRAVEL SHIFTS DUE TO SELF-DRIVING VEHICLES

Kenneth A. Perrine
Research Associate
Center for Transportation Research
The University of Texas at Austin
kperrine@utexas.edu
Phone: 512-232-3123

Kara M. Kockelman
(Corresponding author)
E.P. Schoch Professor of Engineering
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
kkockelm@mail.utexas.edu
Phone: 512-471-0210

Yantao Huang
Graduate Student Researcher
Center for Transportation Research
The University of Texas at Austin
yantao.h@utexas.edu

Presented at the 97th Annual Meeting of the TRB

ABSTRACT

Increasing population and travel demand has prompted new efforts to model travel demand across the United States. One such model is rJourney that estimates travel demand among thousands of regions and models mode and destination choice. rJourney includes records representing 1.17 billion long-distance trips throughout the year 2010. Although inter-regional impacts caused by an increase of automated vehicles (AVs) has been investigated, there is little research on inter-regional travel and how longer distance destination and mode choices will change. Because of conveniences offered by AVs, the value of travel time of drivers is expected to fall, thus reducing the generalized cost of AV travel. To initially analyze the impacts of AVs in the United States, a new AV mode was added to a subset of the rJourney mode and destination choice models. With an initial scenario assuming an operating cost of AVs that is 118% of traditional cars, two outcomes are observed that are solely based on model results. First, the avaiability of AVs severely digs into the airline travel market, reducing airline revenues to 53%. Second, the introduction of AVs results in a shift of destination choice, increasing travel in further distances for personal vehicles by 9.6%, but favoring closer distances across all modes with a 6.7% overall trip-miles reduction. While this preliminary research has revealed an initial perspective on how an existing model can support AVs, the increasing availability of data as AVs emerge will refine nationwide long-distance modeling.

Keywords: autonomous vehicle; long distance; travel demand modeling; national scale

INTRODUCTION

As the United States population grows, it is expected that the demand for inter-city travel will rise, running up against the limited capacity of existing infrastructure. The Federal government and states continuously seek to improve long-distance mobility; however, national-scale passenger travel demand modeling is still an emerging area of research. In efforts to enable proactive planning, the Federal Highway Administration (FHWA) commissioned several studies. One of the studies produced a passenger travel demand model called rJourney that models all long-distance travel in the entire United States for the duration of the year 2010 (Federal Highway Administration, 2015).

While the rJourney model surpasses the limitations of traditional travel demand forecasting methods by rigorously incorporating several forms of travel behavior, the prospect of applying the model to an increasingly automated future is challenged by the fact that automated vehicles (AVs) were not a mode of choice in 2010, and therefore are not represented in the model. While traveler behavior may gradually change as the future emerges and AVs continue to enter the marketplace, the most feasible and best-validated future-looking models at hand are inevitably based upon today's knowledge.

This preliminary research leverages the rJourney model to investigate how long-distance travel between pairs of regions across the continental United States may be affected by the option of having vehicles self-drive travelers to their destinations. Possible effects that arise include a general shift in destination choice that promotes a change in overall person-miles traveled (PMT), and a significant change in overall mode choice between personal vehicles and commercial air carriers.

BACKGROUND

AVs and Long-Distance Travel

While there have been several simulations of AVs' and shared AVs' effects on intra-regional travel (e.g., Fagnant and Kockelman (2014) and Childress et al. (2015)), there is little research on inter-regional travel and how longer-distance destination and mode choices will change. LaMondia et al. (2016) explored mode choices in Michigan for trips over 50 miles in length, and forecasted that over 25 percent of airline trips under 500 miles will shift to AVs. Such changes will have important impacts on airlines, infrastructure planning and future land use (especially around long-distance transportation facilities), highway congestion, and the travel industry more generally.

Long-distance travel is common in many countries and regions. Mercedes-Benz responded to the Google challenge in August 2013 with the S500 Intelligent Drive Autonomous Car long-distance test drive between Mannheim and Pforzheim without any driver input. Automated public vehicles may provide much of the long-distance travel between European countries (Heinrichs, 2016). 19% of Americans with disabilities report leaving their homes relatively infrequently, and are less likely to take long-distance trips (BTS, 2003). However, Meyer and Deix (2014) noted that if AVs allow disabled individuals to make the same length and number of car trips, their vehicle-miles traveled (VMT) would probably increase by more than 50 percent.

AVs reduce the burden of travel for drivers and may improve the quality of travel for passengers, who can now focus on more meaningful interactions with those previously focused on driving.

The value of travel time (VOTT) of the driver (or his/her willingness to pay to save travel time) is expected to fall, by 20 to 50 percent or more, so the generalized cost of travel can fall by several dollars per hour to $6 or more per hour, for many travelers. Auld et al. (2017) applied an integrated transportation system model to analyze the impact of hypothesized connected and autonomous vehicle (CAV) scenarios, varying the market penetration, capacity changes and travel time valuations, on performance of the transportation network and changes in mobility patterns for Chicago region. The results show that an increase in capacity of 80% can be achieved with only 4% induced additional VMT. Changes in travel time cost, or VOTT savings, have a significant impact, especially at very low levels of VOTT, increasing VMT by up to 59%.

Extensions of Prior Models

With the impending introduction of AVs as a viable mode choice in the near future, it is necessary for today's future-looking travel demand forecasting models to incorporate them. Childress et al. (2015) used a Seattle, Washington activity-based travel model (including short-term travel choices and long term work-location and auto-ownership choices) to anticipate the impacts of AV technology introduction on regional travel (attributed to higher roadway capacities, lowered value of travel time (VOTT), reduced parking costs, and increased car-sharing). They estimated that higher income households are more likely to choose the AV mode, as costly technology and VOTT reductions for higher-VOTT travelers are likely to be more significant. When shared automated vehicles (SAVs) are modeled to cost $1.65 per mile (similar to costs of current ride-sharing taxi services, like Lyft and Uber), drive-alone trips were estimated to be reduced by one-third and transit shares increased by 140%, as modeled households did away with traditional vehicles and bought AVs, or shifted to SAVs as well as other travel options.

Other existing projects introduced AVs as a new mode in mode choice or destination choice models. Gucwa (2014) used an activity-based model approach to simulate the travel decisions of individuals in the 9-county San Francisco Bay Area. The autonomous vehicle scenarios are modeled under different values of travel time and road capacity, using the Bay Area's Travel Model One. The mode choice confirms to a random utility model. The result showed that the automation can expect a short-run increase of 4-8% in daily VMT. Zhao and Kockelman (2017) extended the Austin, Texas 6-county region local municipal planning organization's conventional travel demand model with new CAV and SAV modes. The gravity model for trip distribution was replaced with a multinomial logit (MNL) model to allow destination choice to be influenced by the new modes. The mode choice model was also simplified and extended to support the new modes. Simulations varied the assumed operating and parking costs. Results suggested that by the year 2020, the introduction of these modes would add 20% demand to the region's current VMT. An added consequence is a reduction of transit system usage. Both of these were attributed to the relative value of time of CAV and SAV travelers as well as an anticipated competitive SAV pricing scheme. Results of this paper suggest that without full realization of other anticipated benefits of CAVs and SAVs (e.g. smaller headways, shared rides), overall congestion would worsen from that of today.

This research investigates a possible use of rJourney to forecast traditional personal car, commercial air carrier, and personal AV mode and destination choice offers insight on future United States inter-city travel. Since aircraft will still travel much faster than AVs between long-distance city pairs (e.g., New York City to Los Angeles), it is intuitive that those markets could be largely immune to this new mode alternative. However, looking at what routes will be significantly changed lacks research and is important for airline and infrastructure planning. If for example the 240-mile (385 km) route between Houston and Dallas is largely dominated by AVs, interstate planners should expect higher traffic on Interstate 45 and the airport managers should expect less short distance travel between the two cities.

This remainder of this paper is organized as follows. First, the rJourney data set that is used in this research is introduced, followed by the preliminary methodology. Then, results of the research model are identified, as well as an exploration of how the model can be used to estimate how the introduction of AVs may affect overall airline industry revenue. Finally, this paper concludes and offers future research directions.

DATA SET

The rJourney data that is leveraged in this research is part of an extensive, nationwide tour-based long-distance travel model created by RSG for the United States Department of Transportation Federal Highway Administration. The motivation for the creation of rJourney is to study intercity travel and to enhance interstate, long-distance modeling efforts. As noted earlier, long distance travel is modeled among almost all pairwise combinations of 4,486 National Use Microdata Area (NUMA) zones as shown in Figure 1. As part of the rJourney effort, NUMAs are derived from both Census Bureau Public Use Microdata Areas (PUMAs) and county boundaries. The 1.17 billion rJourney tours are generated from a synthesized household population of 31.5 million, representing all long-distance travel in the year 2010. Destination and mode choice are modeled with cross-nested logit (CNL), supporting four modes: automobile, bus, rail and airlines. Trip models are organized among five purposes: business travel, commuting, personal business for shopping and relaxation, visiting friends and family, and leisure travel (Outwater et al. 2014).

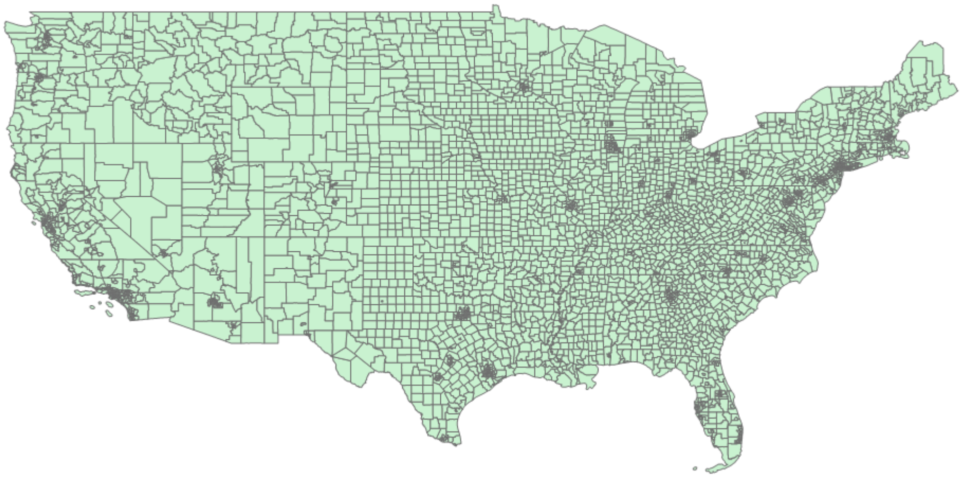

Figure 1: NUMA boundaries within the continental United States

The generated tours provided in the rJourney set across all trip types are distributed as shown in Figure 2. Distances for all modes are measured as round-trip driving distance. All tours consist of one outbound and one return trip over the same path. Important aspects to note about this distribution are that no round-trips shorter than 100 miles (161 km) are expressed in the rJourney tours data set since rJourney only looks at longer-distance trips that involve originating in one NUMA and arriving at a distant NUMA. The longer-distance car trips amount to 1.2 trillion

VMT, which is 40% of the total 3.0 trillion highway VMT reported for 2010 (Bureau of Transportation Statistics, 2011). As expected, car usage largely dominates shorter trips (less than or equal to 500 miles, or 805 km), while air travel dominates for longer ranges. Bus and rail consistently account for a small portion of all trips. The average party size in a tour is 2.15 people.

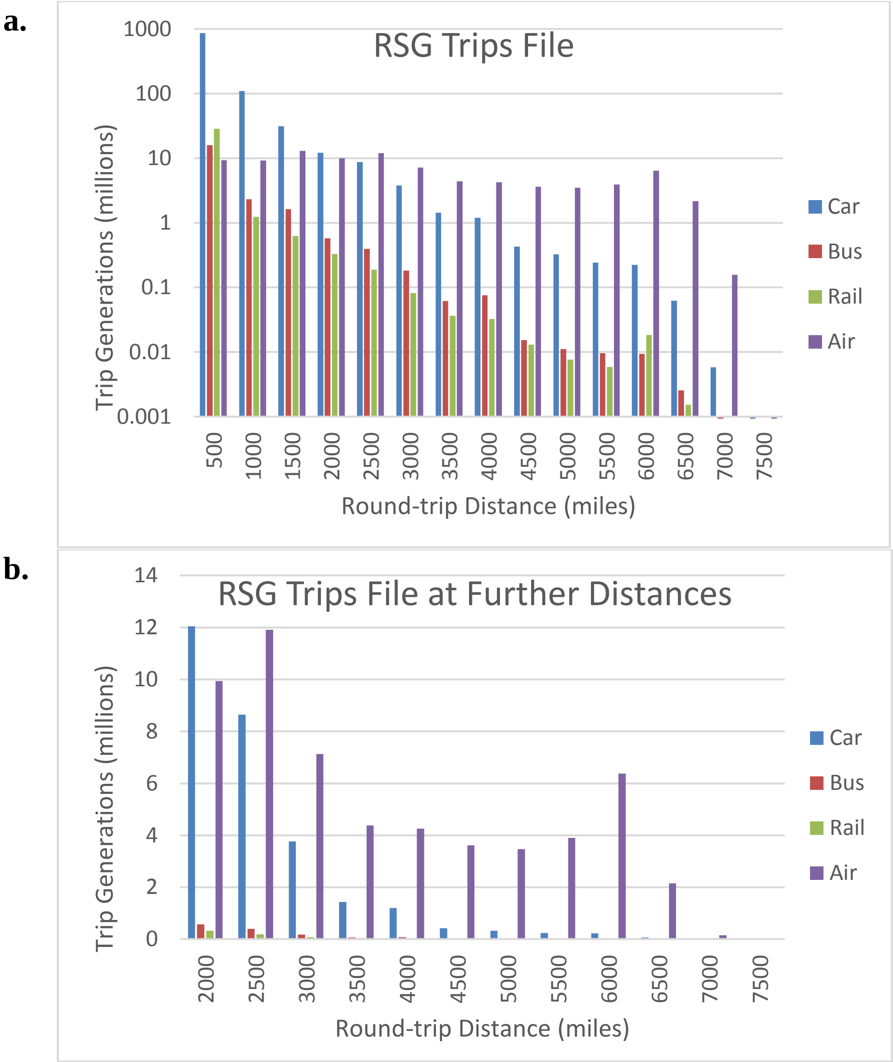

Figure 2: Distribution of rJourney trips for all trip types for a. all distances (shown logarithmically), and b. further distances

The rJourney set also provides a skim file that includes mode statistics of traveling between most possible pairs of NUMAs. These include estimated travel time by car or air, access and egress times, traveling toll or cost, and other factors that would influence a traveler's choice of transportation mode. Corresponding to these are mode choice and destination choice coefficients. In these coefficients, value of travel time for car drivers is $12/hour (in 2010 dollars). These skims and data are used in this research for evaluating the effects of adding a new AV mode.

METHODOLOGY

This analysis leverages a subset of rJourney data and models, and uses pre-existing parameters as a means to quickly characterize the trip distributions for each mode, while leaving the opportunity to add a new mode such as AVs. The subset of data and coefficients were used to closely reproduce the rJourney mode choice results, and then a new AV mode was added. For this analysis, the model was set up as a nested logit model, where mode choice was a nest within an overarching destination choice model.

For finding mode choice from each origin to each destination, parameters include direct costs (value of time, tolls, and fares), NUMA household density, service frequency, transfer frequency, and rail station/airport access and egress penalty. For simplicity, data that are not available to the authors, and parameters not significantly influential in mode choice (e.g. with low T-stats) are not represented in utility functions as they are in the rJourney model, including household size, party size, and number of nights staying. Party size is currently assumed to be 1, and reporting below focuses upon VMT and trip-miles, not person-miles traveled.

As a result, the model subset does not produce an exact replication of the rJourney tours data set. Furthermore, the attempted addition of the AV mode inherently lacks supporting data, already necessitating the use of a subset of existing parameters. Although model subset results show a similar distribution to that of the rJourney tours data set, air travel in particular was underrepresented, showing a correlation of 0.71 overall. To establish a closer representation, a strategy for adjusting (or “pivoting”) the results off of the rJourney tours data set is described further below in Equations 12 and 13.

While future work related to this research will continue to improve upon the rJourney model usage, the preliminary exercise discussed in this paper illustrates the kinds of analyses that are possible with such a model. These are the mode choice utilities, functions of NUMA *i*, destination NUMA *j*, and trip purpose *p*. Refer to (Federal Highway Administration, 2015) for Table 40 that contains the coefficient values and T-stats for each of the trip purposes identified by coefficient subscript number.

$$V_{\text{car\_}i,j,p} = \left(\beta_{11\_p} X_{\text{CT\_}i,j} + C_{\text{OC}} X_{\text{D\_}i,j} + \beta_{10\_p} X_{\text{T\_}i,j}\right)\beta_{1\_p} + \gamma_{112\_p} d_{500\_i,j} + \gamma_{103\_p} \cdot 1 \tag{1}$$

where

$X_{\text{CT\_}i,j}$ = Car travel time from NUMA *i* to *j*,

$X_{\text{D\_}i,j}$ = Distance in miles from NUMA *i* to *j*,

$X_{\text{T\_}i,j}$ = Tolls incurred from NUMA *i* to *j*,

C_{OC} = Car operational cost in dollars per mile, \$0.17/mile in initial analysis,

$d_{500\_i,j}$ = Indicator for one-way distance > 500 mi. (805 km) for NUMAs *i* and *j*, and

$\beta_{1\_p}, \beta_{10\_p}, \beta_{11\_p}, \gamma_{103\_p}, \gamma_{112\_p}$ = Coefficients

$$\begin{aligned} V_{\text{bus\_}i,j,p} = {} & \left(\beta_{21\_p} X_{\text{BT\_}i,j} + \beta_{10\_p} X_{\text{BF\_}i,j}\right)\beta_{1\_p} + ASC_{200\_p} + \beta_{209\_p} Z_{\text{LD\_}i} + \beta_{210\_p} Z_{\text{LD\_}i} + \\ & \gamma_{215\_p} d_{150\_i,j} + \gamma_{207\_p} \cdot 1 \end{aligned} \tag{2}$$

where variables remain as defined earlier, and

$X_{\text{BT\_}i,j}$ = Bus travel time from NUMA *i* to *j*,

$X_{\text{BF\_}i,j}$ = Bus fare from NUMA *i* to *j*,

Z_{LD_i} = NUMA *i* log density (density is the sum of NUMA *i* total households and total employment divided by NUMA *i* square miles),

$d_{150\_i,j}$ = Indicator for one-way distance 50 mi. (81 km) to 150 mi. (242 km) from NUMA *i* to *j*, and

$\beta_{21\_p}, ASC_{200\_p}, \gamma_{207\_p}, \beta_{209\_p}, \beta_{210\_p}, \gamma_{215\_p}$ = Coefficients.

$$V_{\text{rail}\_i,j,p} = \left(\beta_{31\_p} X_{\text{RT}\_i,j} + \beta_{10\_p}\, X_{\text{RF}\_\, i,j,p} + \beta_{32\_p} + X_{\text{RX}\_i,j} + \beta_{33\_p}\, X_{\text{RQ}\_i,j} + \beta_{34\_p}\left(X_{\text{RA}\_i,j} + X_{\text{RE}\_i,j}\right) + \beta_{35\_p}\frac{X_{\text{RA}\_i,j} + X_{\text{RE}\_i,j}}{X_{\text{D}\_i,j}}\right)\beta_{1\_p} + ASC_{300\_p} + \beta_{309\_p}\, Z_{\text{LD}\_i} + \beta_{310\_p} Z_{\text{LD}\_j} + \gamma_{315\_p} d_{150\_i,j} + \gamma_{307\_p} \cdot 1 \tag{3}$$

where variables remain as defined earlier, and

$X_{\text{RT}\_i,j}$ = Rail travel time from NUMA *i* to *j*,

$X_{\text{RF}\_i,j,p}$ = Rail fare for NUMA *i* to *j*, business fare if “employer” purpose,

$X_{\text{RX}\_i,j}$ = Rail transfers incurred from NUMA *i* to *j*,

$X_{\text{RQ}\_i,j}$ = Rail frequency for traveling from NUMA *i* to *j*,

$X_{\text{RA}\_i,j}$ = Access time for getting to the rail station for NUMA *i* to *j*,

$X_{\text{RE}\_i,j}$ = Egress time for departing from the rail station for NUMA *i* to *j*, and

$\beta_{31\_p}, \beta_{32\_p}, \beta_{33\_p}, \beta_{34\_p}, \beta_{35\_p}, ASC_{300\_p}, \gamma_{307\_p}, \beta_{309\_p}, \beta_{310\_p}, \gamma_{315\_p}$ = Coefficients.

$$V_{\text{air}\_i,j,p} = \left(\beta_{41\_p}\, X_{\text{AT}\_i,j} + \beta_{10\ \_p}\, X_{\text{AF}\_i,j,p} + \beta_{42\_p}\, X_{\text{AX}\_i,j} + \beta_{43\_p}\sqrt{X_{\text{AQ0}\_i,j} + \frac{X_{\text{AQ1}\_i,j}}{2} + \frac{X_{\text{AQ2\_i,j}}}{10}} + \beta_{46\_p}\, X_{\text{OT}\_i,j} + \beta_{44\_p}\left(X_{\text{AA}\_i,j} + X_{\text{AE}\_i,j}\right) + \beta_{45\_p}\frac{X_{\text{AA}\_i,j} + X_{\text{AE}\_i,j}}{X_{\text{D}\_i,j}}\right)\beta_{1\_p} + ASC_{400\_p} + \beta_{409\_p} Z_{\text{LD}\_i} + \beta_{410\_p} Z_{\text{LD}\_i} + \gamma_{415\_p} d_{150\_i,j} + \gamma_{407\_p} \cdot 1 \tag{4}$$

where variables remain as defined earlier, and

$X_{\text{AT}\_i,j}$ = Air travel time from NUMA *i* to *j*,

$X_{\text{AF}\_i,j,p}$ = Airfare for NUMA *i* to *j*, business fare if “employer” purpose

$X_{\text{AX}\_i,j}$ = Air transfers incurred from NUMA *i* to *j*,

$X_{\text{AQ0}\_i,j}, X_{\text{AQ1}\_i,j}, X_{\text{AQ2}\_i,j}$ = Air service frequency for direct flights from NUMA *i* to *j* for direct flights, flights with one transfer, and flights with two transfers

$X_{\text{OT}\_i,j}$ = Air on-time rate for flights from NUMA *i* to *j*,

$X_{\text{AA}\_i,j}$ = Access time for getting to the airport for NUMA *i* to *j*,

$X_{\text{AE}\_i,j}$ = Egress time for departing from the airport for NUMA *i* to *j*, and

$\beta_{41\_p}, \beta_{42\_p}, \beta_{43\_p}, \beta_{46\_p}, \beta_{44\_p}, ASC_{400\_p}, \gamma_{407\_p}, \beta_{409\_p}, \beta_{410\_p}, \gamma_{415\_p}$ = Coefficients.

Coefficients are drawn from the rJourney model. In this analysis, the data series pertaining to cost of traditional vehicle operation was drawn using the estimated value of \$0.17/mile. Because this model focuses on mode choice at the time of travel, the ownership cost is not incorporated as in (AAA, 2015). While this serves as a rough estimate, it would be possible with further research to better quantify operation costs as a function of each trip-maker's annual driving distance. The results of the initial analysis shall inform how this function can be evaluated in the future.

The rJourney data includes 285,579 NUMA pairs that lack car mode statistics. These NUMA pairs and corresponding trips are omitted from this analysis because of lack of car-distance data, which is needed in estimating the distance of all modes of travel.

The introduction of AVs into the model presents challenges in implementation, mainly in that the rJourney models and results obviously do not consider the presence of AVs, and little data currently exist to specifically justify model parameters. For AVs to be considered as a new modal alternative, existing data and coefficients are leveraged to arrive at a "best-guess" parameter set. In initially designing how the new modal alternative is integrated, the following assumptions are made: a) a future time is modeled where AVs cost on average \$0.20 per mile to operate; b) the \$6.00 value of time to the occupant is half of that of traditional car; and c) all other parameters are that of traditional cars. The utility function for the AV mode choice is then:

$$V_{\mathrm{AV}\_i,j,p} = \left(\frac{\beta_{11\_p}}{2} X_{\mathrm{CT}\_i,j} + C_{\mathrm{AC}} X_{\mathrm{D}\_i,j} + \beta_{10\_p} X_{\mathrm{T}\_i,j}\right) \beta_{1\_p} + \gamma_{112\_p} d_{500\_i,j} + \gamma_{103\_p} \cdot 1 \qquad (5)$$

where variables remain as defined earlier, and

C_{AC} = AV operational cost in dollars per mile (\$0.20 per mile for the initial analysis).

Probability splits for mode choice given each origin, destination, and purpose are then found:

$$P_{m|i,j,p} = \frac{e^{V_{m\_i,j,p}}}{\sum_{m \in M} e^{V_{m\_i,j,p}}} \qquad (6)$$

where

$V_{m\_i,j,p}$ = Utility function for mode *m*, from NUMA *i* to *j* for purpose *p*,

M = Set of all modes being analyzed.

The destination choice portion of the model incorporates the logsum of the mode choice utility functions along with indicators pertaining to distance ranges, as well as household and employment counts that come from the NUMA zone data set. Again, for simplicity as well as lack of access to data, parameters that are not strongly influential in mode choice and destination choice were omitted. However, as noted later, preliminary results are helpful in identifying investigations of the model in future work. As an observation, the rJourney model does not include gross domestic product per NUMA zone, which could possibly be helpful for future efforts in better representing destination attractiveness.

The following represents the destination choice model, using coefficients drawn from (Federal Highway Administration, 2015) Table 39. Future research efforts will evaluate how more of the rJourney destination-choice model can be leveraged for arriving at an improved representation of attractiveness.

$$LOGSUM_{i,j,p} = \log \textstyle\sum_{m \in M} e^{V_{m\_i,j,p}} \tag{7}$$

$$D_{i,j,p} = \alpha_{1\_p} LOGSUM_{i,j,p} + \alpha_{2\_p} \log\left(X_{\mathrm{D}\_i,j}\right) + \left(a_{3\_p} + \alpha_{6\_p}\right)\left(\frac{X_{\mathrm{D}\_i,j}}{100}\right)^2 + \alpha_{8\_p} d_{100\_i,j} + \alpha_{9\_p} d_{150\_i,j} + \alpha_{10\_p} d_{250\_i,j} + \alpha_{11\_p} d_{500\_i,j} + \alpha_{12\_p} d_{1000\_i,j} + \alpha_{13\_p} d_{1500\_i,j} + \alpha_{14\_p} d_{2000\_i,j} + \alpha_{15\_p} d_{\mathrm{DU}\_j} + \alpha_{16\_p} d_{\mathrm{DR}\_j} + \alpha_{17\_p} d_{\mathrm{ODU}\_i,j} + \alpha_{18\_p}\, d_{\mathrm{ODR}\_i,j} \tag{8}$$

$$P_{j|i,p} = \frac{e^{D_{i,j,p} + \theta_{19\_p} \log S_{j,p}}}{\sum_k^N e^{D_{i,k,p} + \theta_{19\_p} \log S_{k,p}}} \tag{9}$$

where variables remain as defined earlier, and

$S_{j,p}$ = Size for NUMA *j*, purpose *p*. This leverages employment, land use, and education data as identified here, selected by purpose, multiplied with given log-coefficients and summed together:

Coeff.	**1-Personal**	**2-Visit**	**3-Leisure**	**4-Commute**	**5-Employer**
1	Medical	Accommodation	Accommodation	Other service	Accommodation
log α_{20_p}	Entertainment	Entertainment	Entertainment	Entertainment	Entertainment
log α_{21_p}	Other service	Medical	Other service	Retail + wholesale	Retail + wholesale
log α_{22_p}	All other empl	All other empl	All other empl	All other empl	All other empl
log α_{23_p}	University enrollment	Households	Park area (sqm)	University enrollment	University enrollment

$d_{100\_i,j}, d_{150\_i,j}, d_{250\_i,j}, d_{500\_i,j}, d_{1000\_i,j}, d_{1500\_i,j}, d_{2000\_i,j}$ = Indicators for respective distance ranges from NUMA *i* to NUMA *j*,

$d_{\mathrm{DU}\_i}, d_{\mathrm{DR}\_j}$ = Indicators for destination NUMA *j* urban and rural indications, respectively. Urban is defined to have a density of ≥ 1000, and rural is defined to have a density of ≤ 25,

$d_{\mathrm{ODU}\_i,j}, d_{\mathrm{ODR}\_i,j}$ = Indicators for both origin and destination NUMAs *i* and *j* having both urban or rural indications, respectively,

θ_{19_p} = Size multiplier for purpose *p*,

N = Number of NUMAs being analyzed, and

$\alpha_{1\_p}, \alpha_{2\_p}, \alpha_{3\_p}, \alpha_{6\_p}, \alpha_{8\_p}, \alpha_{9\_p}, \alpha_{10\_p}, \alpha_{11\_p}, \alpha_{12\_p}, \alpha_{13\_p}, \alpha_{14\_p}, \alpha_{15\_p}, \alpha_{16\_p}, \alpha_{17\_p}, \alpha_{18\_p}$ = Coefficients.

From this, joint mode/destination choice probabilities are found by combining the mode choice and destination choice conditional probabilities for each origin/destination pair:

$$P_{m,j|i,p} = P_{j|i,p} P_{m|i,j,p} \tag{10}$$

The last step is to use the joint probabilities to distribute trips that are generated from each origin across all modes and destinations. For this analysis, the number of generated trips are obtained from the rJourney tours data that was simulated from generated households across the United States. Because the idea is to study how mode choice and destination choice changes with the introduction of AVs, the mode choices represented in the rJourney tours dataset are ignored to allow the same number of generated tours to be redistributed according to the post-AV introduction model. The modeled tours are defined as:

$$T_{i,j,m,p} = P_{m,j|i,p} \textstyle\sum_k^N R_{i,k,p} \tag{11}$$

where variables remain as defined earlier, and

$R_{i,k,p}$ = Number of trips in the rJourney trips dataset from origin NUMA *i* to destination *k* for purpose *p*.

As mentioned earlier, the model subset does not produce an exact replication of the rJourney tours data set. The authors therefore “pivoted” modeled tours $T_{i,j,m}$ with the rJourney tours data set $R_{i,j,m}$ to arrive at $T*_{i,j,m}$ as follows:

$$T^{*}_{i,j,m\neq\mathrm{AV}} = \frac{T^{\mathrm{w/\ AV}}_{i,j,m}}{T^{\mathrm{w/o\ AV}}_{i,j,m}} \times R^{\mathrm{w/o\ AV}}_{i,j,m},\ m \in \{\mathrm{car, bus, rail, air}\} \qquad (12)$$

$$T^{*}_{i,j,m=\mathrm{AV}} = \frac{T^{\mathrm{w/\ AV}}_{i,j,m=\mathrm{AV}}}{T^{\mathrm{w/o\ AV}}_{i,j,m=\mathrm{car}}} \times R^{\mathrm{w/o\ AV}}_{i,j,m=\mathrm{car}} \qquad (13)$$

Computation of this model can be classified as a big data problem. In representing the expanded 1.17 billion trips, 38 million rJourney trip records over 2 million NUMA pairs constitute 4 GB of data, and files representing the intermediate and final computational results for all trip purposes amount to gigabytes of additional storage requirements. The Python Pandas library is used to perform the computations along with HDF5 file format support. With a number of considerations made for vectorized matrix operations, the entire set takes on the order of 30 minutes to run on a modern, high-end computer. Operations that read and write files from flash storage account for over half of the run time.

RESULTS

Figure 3 shows the resulting number of trips after the AV mode is added to the initial model as described in the methodology. This can be compared with the tours data set distribution in Figure 2. A notable observation is that the distribution of AV trips tracks the distribution of traditional vehicles with an increase in mode share at further distances. This can be attributed to high correlation of several parameters that are represented in the traditional vehicles. The key differences with AVs are the increase in operating cost, and reduced value of time driving. With similarity in parameters, this mode split is influenced by the independence from irrelevant alternatives (IIA) property (or, noted many times in the literature as the “red bus/blue bus paradox”) inherent in multinomial logit models. This property causes highly correlated inputs to be treated as independent, which creates an artificial demand that may not necessarily happen in reality. The high degree of correlation and presence of IIA can best be addressed by creating a nest (e.g. “personal vehicles”) that contains both of the AV and car results, an area for future work.

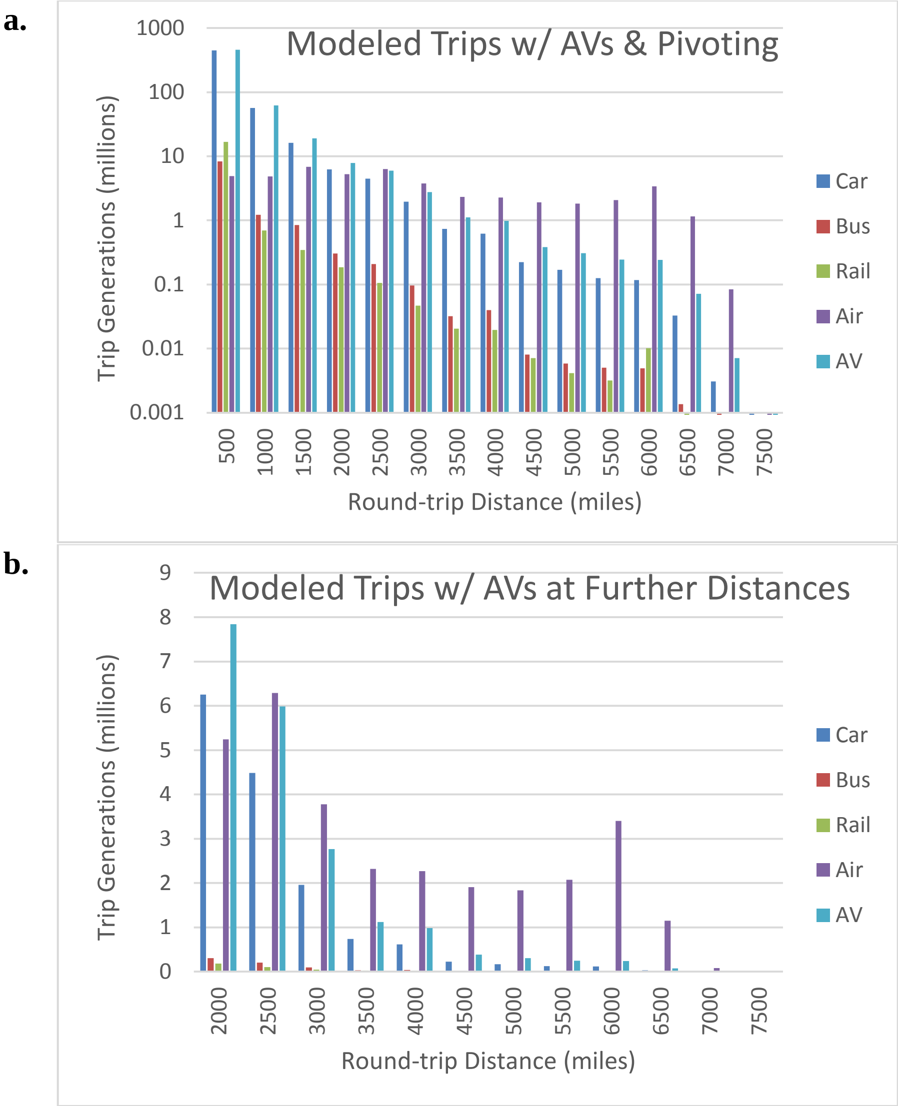

Figure 3: Number of trips from the mode choice/destination choice analysis, all purposes, at a. all distances (shown logarithmically), and b. further distances

There are two notable outcomes that offer insight on the possible effects of AV introduction to the market, as well as a shift in destination choice. First, results show that the introduction of AVs deeply cuts into the number of trips that had formerly been air trips. See the first two sets in Table 1 for results in terms of shorter and longer trips (e.g. < 500 miles (805 km) versus ≥ 500 miles). As largely influenced by the γ_{112_p} coefficient as well as travel time, trips over 500 miles in length are penalized because of the negative "captivity factor" of remaining in a car for a long period of time possibly over several days. It is assumed in this model that this disutility would be similar for AVs as it would be for traditional cars. Note that in Table 1, "Car+AV" is shown as a means to represent respective totals of personally owned vehicles.

Second, among traditional cars and new AVs, more destinations are chosen after introduction of AVs that are further in distance from origins. When looking at the long-distance personal vehicle travel represented in the model, the 3.0 trillion highway VMT reported for 2010 (Bureau of Transportation Statistics, 2011) increases to 3.1 trillion. However, if all modes are considered, the trend is reversed, possibly because of the severe reduction of air trips that dominate the longer-distance trips. The third set in Table 1 shows a change in distribution across overall trip distances. For both pre- and post-AV introduction the model uses the same number of trip generations per NUMA per trip purpose. The significant decrease of air travel may be a consequence of the aforementioned IIA property. In addition to treating cars and AVs as a single nest, further work on characterizing VOTT and operating cost, as well as specifying additional factors in the destination-choice portion of the model may have the outcome of evolving how trip distances are biased among closer and further long-distance trips.

Table 1: Trip mode choice impact of AV introduction for all trip purposes

TOURS BY MODE	AV Market Penetration	Car+AV < 500 mi. round trip	Car+AV ≥ 500 mi. round trip	Air < 500 mi. round trip	Air ≥ 500 mi. round trip
Before AV	0%	860.5 M	168.8 M	9.3 M	79.5 M
After AV	51%	906.9 M	189.0 M	4.9 M	42.0 M
% change	-	105.4%	112.0%	52.9%	52.8%

VEHILCE-MILES	Car+AV < 500 mi. round trip	Car+AV ≥ 500 mi. round trip	Car+AV Total	Air < 500 mi. round trip	Air ≥ 500 mi. round trip	Air Total
Before AV	400.8 B	821.0 B	1,221 B	6.4 B	437.9 B	444.3 B
After AV	425.2 B	913.7 B	1,339 B	3.4 B	232.3 B	235.7 B
% change	106.1%	111.3%	109.6%	52.9%	53.0%	53.0%

TOURS FOR ALL MODES	Tours < 500 mi. one way	VMT for tours < 500 mi.	Tours ≥ 500 mi. one way	VMT for tours ≥ 500 mi.
Before AV	914.1 M	422.4 B	256.1 M	1,294 B
After AV	937.0 M	437.1 B	235.2 M	1,165 B
% change	102.5%	103.5%	91.8%	90.0%

AIRLINE REVENUE	Tours < 500 mi. round trip	Tours ≥ 500 mi. round trip	Total revenue
Before AV	$16.0 B	$159.1 B	$175.1 B
After AV	$8.4 B	$83.9 B	$92.3 B
% change	52.7%	52.7%	52.7%

Market Penetration

The degree that AVs penetrate the market varies according to trip distance. Figure 4 shows market penetration both for AVs among the personal vehicle modes (e.g. car and AV), and also AVs among all mode choices. With respect to personal vehicles, the market penetration increases as distance increases because of the significance of lower VOTT. However, air travel continues to be preferred for longer distances and results in the AV mode share diminishing at further distances. The deviation in penetration for the 7000-mile bin is likely a result of fewer trip samples for that furthest distance.

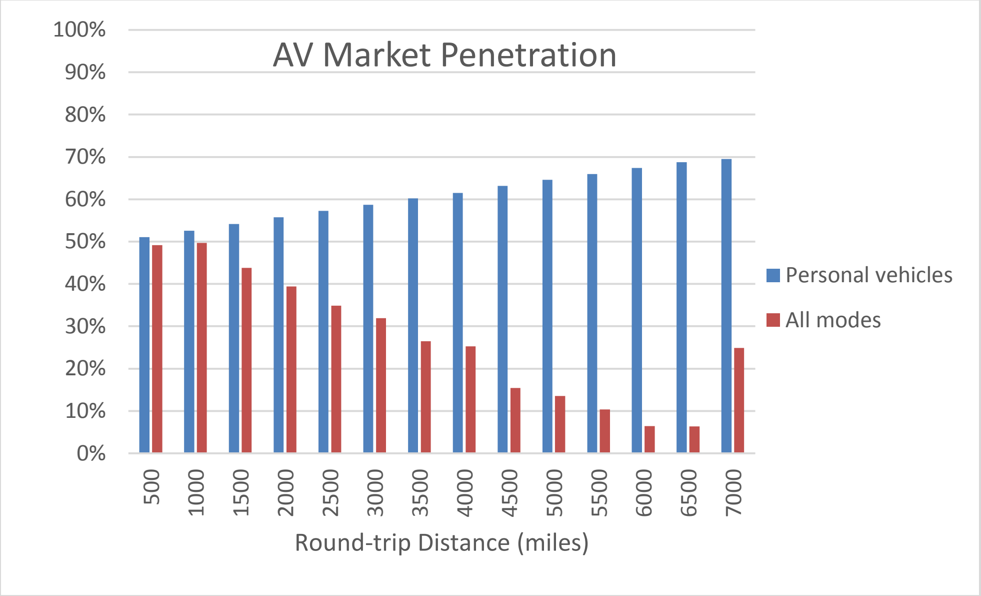

Figure 4: Penetration of AVs among personal vehicles (car+AV) and all modes

Passenger Airline Sales

Given that large-scale introduction of AVs has not yet happened and that no data can be collected directly from AV usage today, a model such as this rJourney subset with AVs added as a new mode can be helpful in roughly estimating market effects that could result from the widespread introduction of AVs. One question that can be addressed with this model is how much revenue the airline industry can possibly lose due to more travelers choosing AVs over air travel. The rJourney data set gives airfare estimates in USD for all NUMA pairs that have suitable access to airports served by commercial passenger carriers. The fourth set in Table 1 shows estimated airline sales before and after the addition of AVs for all modeled trips. Note that because these are based upon cost to the traveler, these sales figures include airport taxes.

In this result, the percent changes between sales between shorter and longer long-distance trips are similar. This is counterintuitive because of the idea that AVs should have a more significant attractiveness for shorter trips and thus cut more into the shorter distance market. It may be here that the model is dominated by the IIA property in adding AVs as a separate mode rather than as a car+AV "personal vehicle" nest. Additionally, with refinements in the mode choice and destination choice models the split may improve in accuracy.

AV Parameter Sensitivity

As mentioned earlier, the parameters and assumptions given to AVs are largely unknown and must be estimated. Two notable parameters include cost of operating the vehicle, as well as personal VOTT. (Another parameter that is relevant but not yet analyzed includes a more pronounced representation of the 500-mile captivity factor, which may be different for car drivers than it is for AV passengers.) A thorough analysis should offer a set of scenarios that span a range of expected operational costs and personal VOTT, given the targeted years, expected AV market penetration, and socioeconomic classes of trip-makers that are being analyzed.

To further understand the sensitivity of these variables on the resulting mode split and destination choice, six new scenarios are created for the "leisure" trip purpose. Scenarios are:

Scenario	**Oper. Cost ($/mile)**	**VOTT ($/hr)**	**Notes**
A	$0.20	$6.00	Base case
B	$0.10	$6.00	Operating cost is cheaper
C	$0.50	$6.00	Operating cost is more expensive
D	$0.20	$3.00	VOTT is decreased
E	$0.20	$9.00	VOTT is increased
F	$1.65	$6.00	AVs are modeled as shared vehicles

Recall that dollar amounts are expressed in year 2010 dollars. The scenario of AVs having the same operating cost and VOTT of cars has been omitted because there would be no distinction between the car and AV modes. Scenario F in particular has been included as a hypothetical scenario to roughly model all AVs on the roadways as shared autonomous vehicles (SAVs). With SAVs, passengers do not own their vehicles, but rather pay per mile for travel in a borrowed vehicle that others can use for other trips, in this case $1.65 per mile. As more data emerges, an improved model would likely offer SAVs as a mode choice that is separate from personally-owned AVs. Table 2 shows the results of each of these scenarios.

Table 2: Trip generations with varied AV parameters, for "leisure" trip purpose

Mode	**Dist.**	**Scenario**	**A**	**B**	**C**	**D**	**E**	**F**
Car+AV	Trips < 500 mi.	Before AV	253.5 M	253.5 M	253.5 M	253.5 M	253.5 M	253.5 M
		After AV	271.5 M	267.2 M	279.4 M	268.1 M	274.3 M	280.3 M
		% change	107.1%	105.4%	110.2%	105.7%	108.2%	110.6%
	Trips ≥ 500 mi.	Before AV	55.7 M	55.7 M	55.7 M	55.7 M	55.7 M	55.7 M
		After AV	63.4 M	65.7 M	57.8 M	65.3 M	61.7 M	46.7 M
		% change	113.8%	118.0%	103.7%	117.2%	110.7%	83.9%
Air	Trips < 500 mi.	Before AV	2.30 M	2.30 M	2.30 M	2.30 M	2.30 M	2.30 M
		After AV	1.23 M	1.20 M	1.28 M	1.21 M	1.24 M	1.40 M
		% change	53.2%	52.1%	55.7%	52.3%	54.0%	60.9%
	Trips ≥ 500 mi.	Before AV	18.11 M	18.11 M	18.11 M	18.11 M	18.11 M	18.11 M
		After AV	9.65 M	9.43 M	10.16 M	9.47 M	9.81 M	11.17 M
		% change	53.3%	52.1%	56.1%	52.3%	54.2%	61.7%

In observing Scenarios B, A, and C in order of increasing operation cost, it can be seen that closer trip generations increase and longer trips decrease because of the significance of operating cost on longer trips. Meanwhile, the cut into the air market decreases as the operation cost increases. In the rough SAV Scenario F, the results coincide with a similar trend, where longer distance trips are more significantly curtailed. For Scenarios D, A, and E in order of increasing VOTT, a similar phenomenon occurs. The reduction of air trips decreases as VOTT increases.

In all cases, the variations that are evaluated do not show an extreme difference in outcomes. In considering travelers' expenses and VOTT, it is possible to reason that the results should be more distinct. Two factors may be dominating the models as these inputs are varied. First, the addition of the AV mode as an independent choice may be an inaccurate model structure that is highly correlated and represented too significantly in the results. As mentioned earlier, it may be more appropriate to treat cars and AVs as a "personal vehicle" nest and estimate the correlation that is to be expected among the mode choices of hypothetical travelers. Second, the representation of AVs in the model is somewhat indistinct from cars, as few parameters exist to offer better differentiation. The addition of new parameters to the car and AV modes can help with this and reduce the correlation between the two modes.

CONCLUSIONS

This preliminary research has leveraged the nationwide, inter-regional rJourney travel demand model for estimating impacts of future introduction of AVs. While models such as rJourney had been created in efforts to better understand intercity travel and offer enhanced capabilities for planning, little research today addresses the introduction of AVs in such models. This effort therefore is intended as an early investigation in allowing AVs to be treated as a viable mode within the same class of modeling framework.

A subset of the rJourney model was implemented to predict mode and destination choice of long-distance travelers with AVs fully considered as a viable mode alternative. The integration of AVs into the model includes some of the preexisting car-specific parameters while employing higher cost of vehicle operation and reduced VOTT that are expected of AVs within the oncoming years.

These preliminary results are solely based upon the rJourney results after adding AVs as a distinct mode. First, in the initial scenario where the cost of ownership and operation for an AV is assumed to be $0.20 per mile and VOTT is half of that of car travel, air travel trip generation for shorter and further long-distance trips is cut to 53% of the original value, largely replaced with an increased demand for AVs. It follows that commercial passenger air carriers may benefit from understanding the implications of AV introduction, perform research on the problem, and target their services and marketing accordingly. Second, with the introduction of AVs, trips among cars and AVs favor further distances for trips; but trips appear to favor closer distances when considering all modes. Here, the total number of car and AV trips increases by 5% for shorter-distance trips and 12% for longer-distance trips; however, among all modes there is a 6.7% reduction in trip-miles. It can be surmised that federal and state DOTs should further investigate possible needs for upgrading interregional infrastructure in preparation for specific levels of AV market penetration.

For further future research, it will be prudent to find and analyze data that is collected in the field as AVs emerge, including willingness to pay, technology cost, travel time savings, and socioeconomic aspects of AV usage. Along the way, it would be helpful to have data on public resistance and acceptance to aid in estimating future AV market penetration. The model would also possibly benefit from nesting together the car and AV modes to account for correlation among the two modes. These are all factors that can help to establish a more accurate, nationwide AV mode and destination-choice model that reflects current and future trends.

ACKNOWLEDGEMENTS

The authors thank the Texas Department of Transportation (TxDOT) for financially supporting this research (under research project 0-6838, "Bringing Smart Transport to Texans: Ensuring the Benefits of a Connected and Autonomous Transport System in Texas") and FHWA's Tianjia

Chang and RSG Inc.'s Maren Outwater and Nazneen Ferdous for sharing the dataset used in this project.

The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing computing and data storage resources that have contributed to the research results reported within this paper. URL: http://www.tacc.utexas.edu

The authors appreciate Sindhu Maiyya's preliminary research on this topic, Will Schievelbein's efforts in coding to initially read in and calculate data, and Scott Schauer-West's assistance in manuscript preparation.

REFERENCES

AAA (2015) Your Driving Costs. Retrieved from: http://exchange.aaa.com/wp-content/uploads/2015/04/Your-Driving-Costs-2015.pdf

Auld, J., Sokolov, V. and Stephens, T., 2017. Analysis of the Impacts of CAV Technologies on Travel Demand. In *Transportation Research Board 96th Annual Meeting* (No. 17-06453). http://vsokolov.org/pdf/17-006453_RV_CP01_11152016110928.pdf

Bureau of Transportation Statistics (2003). Freedom to Travel. Publication #BTS03-08. US Department of Transportation Research and Innovative Technology Administration. http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/freedom_to_travel/index.html.

Bureau of Transportation Statistics (2011). State Transportation Statistics. US Department of Transportation. URL: https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/state_transportation_statistics/state_transportation_statistics_2011/pdf/entire.pdf

Childress, S., Nichols, B., Charlton, B., and Coe, S. (2015) "Using an Activity-Based Model to Explore Possible Impacts of Automated Vehicles," presented at the 94th Annual Meeting of the Transportation Research Board. Washington, DC, January 2015.

Fagnant, D. and Kockelman, K.M. (2014). The Travel and Environmental Implications of Shared Autonomous Vehicles, Using Agent-Based Model Scenarios. Transportation Research Part C Vol (40): 1-13.

Federal Highway Administration. 2015. "Foundation Knowledge to Support a Long-Distance Passenger Travel Demand Modeling Framework: Implementation Report."

Gucwa, M. (2014). Mobility and Energy Impacts of Automated Cars. Presented at 2014 Automated Vehicle Symposium. San Francisco, CA. URL: http://higherlogicdownload.s3.amazonaws.com/AUVSI/3a47c2f1-97a8-4fb7-8a39-56cba0733145/UploadedImages/documents/pdfs/7-16-14%20AVS%20presentations/Michael%20Gucwa.pdf

Heinrichs, D., 2016. Autonomous Driving and Urban Land Use. *Autonomous Driving: Technical, Legal and Social Aspects*. Springer Publishing Company: Berlin.

LaMondia, Jeffrey, Daniel Fagnant, Hongyang Qu, Jackson Barrett, and Kara Kockelman. 2016. “Long-Distance Travel Mode Shifts Due to Automated Vehicles: A Statewide Mode-Shift Simulation Experiment and Travel Survey Analysis.” *Transportation Research Record* 2566. http://www.caee.utexas.edu/prof/kockelman/public_html/TRB16LDModeShifts.pdf.

Meyer, G. and Deix, S., 2014. Research and innovation for automated Driving in Germany and Europe. Part II: Industrial Research and Innovation in *Road Vehicle Automation*, pp. 71-81. Springer: Berlin.

Outwater, Maren, Mark Bradley, Nazneen Ferdous, Chandra Bhat, Ram Pendyala, Stephane Hess, Andrew Daly, and Jeff LaMondia. 2015. “A Tour-Based National Model System to Forecast Long-Distance Passenger Travel in the United States.” TRB 94th Annual Meeting Compedium of Papers. Transportation Research Board of the National Academies.

Transportation Research Part A: Policy and Practice, 86, pp.1-18.Zhao, Yong, Kara M. Kockelman. Anticipating the Regional Impacts of Connected and Automated Vehicle Travel in Austin, Texas. Under review for publication in the *Journal of Urban Planning and Development*.