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

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


ABSTRACT

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, as well as changes in land use around long-distance transportation facilities. 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 on the United States transportation network, a nationwide long-distance travel demand model was leveraged in attempts to add a new AV mode. With an initial scenario assuming an ownership and operating cost of AVs that is 167% of traditional cars, two outcomes are observed. First, the attractiveness of AVs as a travel mode alternative digs into the airline travel market, reducing revenues to 71% of that experienced before AV introduction. Second, the introduction of AVs results in a shift of destination choice, increasing travel in shorter distances. While this preliminary research has revealed a viable perspective on how existing models can support AVs, the increasing availability of data as AVs emerge onto the market will guide the refinement of long-distance modeling.

BACKGROUND

While there have been several simulations of AVs’ and shared AVs’ effects on intra-regional travel (e.g., Fagnant and Kockelman (2014), Chen and Kockelman (2015), 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) looked at mode choices of those in Michigan making 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.
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. Thanks to easier “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 $10 or more per hour, for many travelers.

This preliminary research investigates 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. These regions are represented by 4,566 National Use Microdata Area zones, or NUMAs. As part of the rJourney effort (Federal Highway Administration, 2015), NUMAs are derived from both Census Bureau Public Use Microdata Areas (PUMAs) and county boundaries. Since aircraft will still travel much faster than AVs between long-distance city pairs (e.g., New York to Los Angeles), those markets may 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 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.

Other research projects have extended existing models with AVs as a new mode. 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 VOTTs, reduced parking costs, and increased car-sharing). They estimated that higher income households are more likely to choose the AV mode. This is expected because of 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. One simulated scenario looked at SAVs with the implication that travelers are no longer associated with the fixed cost (and round-trip restrictions) of personal vehicle ownership and storage.

Fagnant and Kockelman (2014) worked with an agent-based model for SAVs that simulated environmental benefits of such a fleet as compared to that of traditional, personally-owned vehicles, focusing on a dense urban core area. Simulation results indicated that each SAV may replace 11 conventional private owned vehicles while increasing travel distances by up to 10%. When the simulation was extended to a case study of low market penetration (1.3% of trips) in Austin, Texas, each SAV was able to replace 9 conventional vehicles and generated 8% more VMT on average due to empty, unoccupied travel (Fagnant et al. 2015).

Loeb, Liu and Kockelman (2017) expanded upon prior work with SAVs by specifically looking at how an electric-powered shared automated electric vehicle (SAEV) fleet differs from a gasoline-fueled SAV fleet. Within an agent-based simulation, trips under 75 km in the Austin, Texas 6-county region were modeled in scenarios where charging station densities, operating ranges, and charge times were varied. One plausible scenario incurred a 19.8% unoccupied VMT, largely because of the necessity for repositioning to charging stations frequently. A fleet was found where each SAV serves 7 travelers sequentially, where 91% of travelers were served within 10 minutes.
of making a request. Fleet size and number of charging stations both significantly impacted the response time.

In Chen and Kockelman (2017), a mode choice model was added to a previous agent-based framework to anticipate SAEV market shares in direct competition with other modes. In this work, a fleet of 80-mile-range SAEVs were analyzed along with a Level II charging infrastructure to look at possible operations and pricing strategies that coincide with shifting mode shares.

Finally, Zhao and Kockelman (2017) extended the Austin, Texas 6-county region local municipal planning organization’s conventional travel demand model with new connected and autonomous vehicle (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. Also, the mode choice model was 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 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.

The rJourney data and analysis heavily use US customary units for distance measurements. These are the metric equivalents for frequently discussed distances:

- 50 miles = 80.5 km
- 500 miles = 805 km
- 150 miles = 241 km
- 3000 miles = 4,828 km

**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 NUMAs (see Figure 1). The rJourney trips are generated from a synthesized household population. 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 for the “personal” trip type are distributed as shown in the Figure 2 log-chart. Distances for all modes are measured as car distance. Important aspects to note about this distribution are that no trips shorter than 100 miles are expressed in the rJourney trips data set since rJourney only looks at longer-distance trips that involve originating in one NUMA and arriving in another NUMA. As expected, car usage largely dominates shorter trips (less than or equal to 500 miles), while air travel dominates for longer ranges. Bus and rail consistently account for a small portion of all trips.

Figure 2: Log-distribution of rJourney trips for the “personal” trip type

The rJourney set also provides a skim file that includes mode statistics of traveling between most possible pairs of NUMAs. This includes 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.
Summary statistics of the skim files are shown in Table 1. A subset of these are used in this research for evaluating the effects of adding a new AV mode.

Table 1: Summary statistics for the rjourney skim file

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Air File, N = 18,424,925</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>218.73</td>
<td>97.94</td>
<td>25.00</td>
<td>812.00</td>
</tr>
<tr>
<td>Transfers</td>
<td>82.37</td>
<td>50.19</td>
<td>0.00</td>
<td>200.00</td>
</tr>
<tr>
<td>FreqDirect</td>
<td>10.58</td>
<td>24.07</td>
<td>0.00</td>
<td>339.00</td>
</tr>
<tr>
<td>Freq1Stop</td>
<td>145.41</td>
<td>258.47</td>
<td>0.00</td>
<td>2,286.00</td>
</tr>
<tr>
<td>Freq2Stop</td>
<td>348.81</td>
<td>932.69</td>
<td>0.00</td>
<td>10,968.00</td>
</tr>
<tr>
<td>OnTime</td>
<td>88.79</td>
<td>4.00</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>EconomyFare</td>
<td>519.13</td>
<td>327.69</td>
<td>0.00</td>
<td>50,776.00</td>
</tr>
<tr>
<td>BusinessFare</td>
<td>1,199.61</td>
<td>955.59</td>
<td>0.00</td>
<td>152,328.00</td>
</tr>
<tr>
<td>AccessDistance</td>
<td>38.15</td>
<td>25.99</td>
<td>0.00</td>
<td>101.00</td>
</tr>
<tr>
<td>EgressDistance</td>
<td>38.22</td>
<td>26.34</td>
<td>0.00</td>
<td>102.00</td>
</tr>
<tr>
<td><strong>Rail File, N = 8,010,759</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>2,167.24</td>
<td>1,269.59</td>
<td>4.00</td>
<td>6,270.00</td>
</tr>
<tr>
<td>Transfers</td>
<td>134.57</td>
<td>111.05</td>
<td>0.00</td>
<td>800.00</td>
</tr>
<tr>
<td>Frequency</td>
<td>7.77</td>
<td>10.41</td>
<td>3.00</td>
<td>93.00</td>
</tr>
<tr>
<td>EconomyFare</td>
<td>131.75</td>
<td>39.51</td>
<td>9.00</td>
<td>181.00</td>
</tr>
<tr>
<td>BusinessFare</td>
<td>340.56</td>
<td>132.40</td>
<td>18.00</td>
<td>605.00</td>
</tr>
<tr>
<td>AccessDistance</td>
<td>22.82</td>
<td>14.65</td>
<td>0.00</td>
<td>50.00</td>
</tr>
<tr>
<td>EgressDistance</td>
<td>22.16</td>
<td>15.14</td>
<td>0.00</td>
<td>50.00</td>
</tr>
<tr>
<td><strong>Road File, N = 19,727,179</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CarTime</td>
<td>1,161.72</td>
<td>668.14</td>
<td>1.00</td>
<td>3,613.00</td>
</tr>
<tr>
<td>Distance</td>
<td>1,185.41</td>
<td>706.49</td>
<td>1.00</td>
<td>3,582.00</td>
</tr>
<tr>
<td>Toll</td>
<td>67.15</td>
<td>137.85</td>
<td>0.00</td>
<td>1,344.00</td>
</tr>
<tr>
<td>BusTime</td>
<td>1,313.12</td>
<td>1,249.89</td>
<td>0.00</td>
<td>5,617.00</td>
</tr>
<tr>
<td>BusFare</td>
<td>94.71</td>
<td>85.72</td>
<td>0.00</td>
<td>383.00</td>
</tr>
</tbody>
</table>
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, parameters that are difficult to obtain and believed to not be significantly influential in mode choice are not represented in utility functions as they are in the rJourney model, including household size, party size, and number of nights staying.

As a result, this model does not produce an exact replication of the rJourney model results. Indeed, the attempted addition of the AV mode inherently lacks supporting data, already necessitating the use of a subset of existing parameters. A corresponding implication is that all travelers were assumed to own a car in this research. However, in the rJourney simulation, some travelers do not own a car (such as in areas like New York City) and thus are more likely to use alternative forms of transportation. 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 tables containing the coefficient values and T-stats for each of the trip purposes:

\[ V_{Car\_i\_j\_p} = (\beta_{\_\_p} + c_{\_\_p} + \mu_{\_\_p} + T_{\_\_p} + T_{\_\_p} + O_{\_\_p} + D_{\_\_p} + S_{\_\_p} + E_{\_\_p} + P_{\_\_p} + D_{\_\_p} + I_{\_\_p} + J_{\_\_p} + K_{\_\_p} + L_{\_\_p} + M_{\_\_p} + N_{\_\_p} + O_{\_\_p} + P_{\_\_p} + Q_{\_\_p} + R_{\_\_p} + S_{\_\_p} + T_{\_\_p} + U_{\_\_p} + V_{\_\_p} + W_{\_\_p} + X_{\_\_p} + Y_{\_\_p} + Z_{\_\_p}) \]

\[ V_{Bus\_i\_j\_p} = (\beta_{\_\_p} + c_{\_\_p} + \mu_{\_\_p} + T_{\_\_p} + T_{\_\_p} + O_{\_\_p} + D_{\_\_p} + S_{\_\_p} + E_{\_\_p} + P_{\_\_p} + D_{\_\_p} + I_{\_\_p} + J_{\_\_p} + K_{\_\_p} + L_{\_\_p} + M_{\_\_p} + N_{\_\_p} + O_{\_\_p} + P_{\_\_p} + Q_{\_\_p} + R_{\_\_p} + S_{\_\_p} + T_{\_\_p} + U_{\_\_p} + V_{\_\_p} + W_{\_\_p} + X_{\_\_p} + Y_{\_\_p} + Z_{\_\_p}) \]

\[ V_{Rail\_i\_j\_p} = (\beta_{\_\_p} + c_{\_\_p} + \mu_{\_\_p} + T_{\_\_p} + T_{\_\_p} + O_{\_\_p} + D_{\_\_p} + S_{\_\_p} + E_{\_\_p} + P_{\_\_p} + D_{\_\_p} + I_{\_\_p} + J_{\_\_p} + K_{\_\_p} + L_{\_\_p} + M_{\_\_p} + N_{\_\_p} + O_{\_\_p} + P_{\_\_p} + Q_{\_\_p} + R_{\_\_p} + S_{\_\_p} + T_{\_\_p} + U_{\_\_p} + V_{\_\_p} + W_{\_\_p} + X_{\_\_p} + Y_{\_\_p} + Z_{\_\_p}) \]

\[ V_{Air\_i\_j\_p} = (\beta_{\_\_p} + c_{\_\_p} + \mu_{\_\_p} + T_{\_\_p} + T_{\_\_p} + O_{\_\_p} + D_{\_\_p} + S_{\_\_p} + E_{\_\_p} + P_{\_\_p} + D_{\_\_p} + I_{\_\_p} + J_{\_\_p} + K_{\_\_p} + L_{\_\_p} + M_{\_\_p} + N_{\_\_p} + O_{\_\_p} + P_{\_\_p} + Q_{\_\_p} + R_{\_\_p} + S_{\_\_p} + T_{\_\_p} + U_{\_\_p} + V_{\_\_p} + W_{\_\_p} + X_{\_\_p} + Y_{\_\_p} + Z_{\_\_p}) \]
\[ \beta_{\text{air\_access}_p}(X_{\text{Air\_Acc\_i\_j}} + X_{\text{Air\_Egr\_i\_j}}) + \beta_{\text{mgc\_p}} + \text{ASC}_{\text{Air\_p}} + \]
\[ \beta_{\text{air\_oLogDens\_p}}\lambda_{\text{logDens\_i}} + \beta_{\text{air\_dLogDens\_p}}\lambda_{\text{logDens\_j}} + \gamma_{\text{air\_ood\_150\_p}}d_{\text{dist\_150\_i\_j}} \]

Where

1. \(X_{\text{Car\_Time\_i\_j}}\) = Car travel time from NUMA i to j,
2. \(X_{\text{Distance\_i\_j}}\) = Distance in miles from NUMA i to j,
3. \(X_{\text{Toll\_i\_j}}\) = Tolls incurred from NUMA i to j,
4. \(X_{\text{Bus\_Time\_i\_j}}\) = Bus travel time from NUMA i to j,
5. \(X_{\text{Bus\_Fare\_i\_j}}\) = Bus fare from NUMA i to j,
6. \(X_{\text{Rail\_Time\_i\_j}}\) = Rail travel time from NUMA i to j,
7. \(X_{\text{Rail\_Fare\_i\_j\_p}}\) = Rail fare for NUMA i to j, business fare if “employer” purpose,
8. \(X_{\text{Rail\_Xfers\_i\_j}}\) = Rail transfers incurred from NUMA i to j,
9. \(X_{\text{Rail\_Freq\_i\_j}}\) = Rail frequency for traveling from NUMA i to j,
10. \(X_{\text{Rail\_Egr\_i\_j}}\) = Egress time for getting to the rail station for NUMA i to j,
11. \(X_{\text{Air\_Time\_i\_j}}\) = Air travel time from NUMA i to j,
12. \(X_{\text{Air\_Fare\_i\_j\_p}}\) = Airfare for NUMA i to j, business fare if “employer” purpose
13. \(X_{\text{Air\_Xfers\_i\_j}}\) = Air transfers incurred from NUMA i to j,
14. \(X_{\text{Air\_Freq\_Direct\_i\_j}}\) = Air service frequency for direct flights from NUMA i to j,
15. \(X_{\text{Air\_On\_Time\_i\_j}}\) = Air on-time rate for flights from NUMA i to j,
16. \(X_{\text{Air\_Acc\_i\_j}}\) = Access time for getting to the airport for NUMA i to j,
17. \(X_{\text{Air\_Egr\_i\_j}}\) = Egress time for departing from the airport for NUMA i to j,
18. \(\lambda_{\text{log\_Dens\_i}}\) = Log of density for NUMA i,
19. \(C_{\text{car\_ocost}}\) = Car operational cost in dollars per mile, $0.60/mile in initial analysis,
20. \(d_{\text{dist\_150\_i\_j}}\) = Indicator for one-way distance 50 mi. to 150 mi. from NUMA i to j,
21. \(d_{\text{dist\_500\_i\_j}}\) = Indicator for one-way distance > 500 mi. for NUMAs i and j,
22. \(\beta_{\text{mgc\_p}}\) = Coefficient on mode generalized cost for purpose p,
23. \(\beta_{\text{car\_time\_p}}\) = Coefficient for car travel time for purpose p,
24. \(\beta_{\text{cost\_p}}\) = Cost coefficient for direct costs,
25. \(\beta_{\text{bus\_oLogDens\_p}}\) = Coefficient for log of origin NUMA density for bus for purpose p,
26. \(\beta_{\text{bus\_dLogDens\_p}}\) = Coefficient for log of destination NUMA density for bus for purpose p,
27. \(\beta_{\text{rail\_time\_p}}\) = Coefficient on rail travel time for purpose p,
28. \(\beta_{\text{rail\_xfers\_p}}\) = Coefficient on rail transfers for purpose p,
29. \(\beta_{\text{rail\_freq\_p}}\) = Coefficient on rail service frequency for purpose p,
30. \(\beta_{\text{rail\_acc\_p}}\) = Coefficient on rail station accessibility for purpose p,
31. \(\beta_{\text{rail\_oLogDens\_p}}\) = Coefficient for rail concerning the log density of origin for purpose p,
32. \(\beta_{\text{rail\_dLogDens\_p}}\) = Coefficient for rail concerning the log density of dest. for purpose p,
33. \(\beta_{\text{air\_time\_p}}\) = Coefficient on flight travel time for purpose p,
34. \(\beta_{\text{air\_xfers\_p}}\) = Coefficient on airport transfers for purpose p,
35. \(\beta_{\text{air\_freq\_p}}\) = Coefficient on flight service frequency for purpose p,
36. \(\beta_{\text{air\_ot\_p}}\) = Coefficient on air service on-time departure rate for purpose p,
\[ V_{AV,i,j,p} = (\beta_{air\_time,p}X_{CarTime,i,j} + C_{AV\_ocost}X_{Distance,i,j} + \beta_{cost,p}X_{Toll,i,j})\beta_{mgc,p} + Y_{car\_ood500,p}d_{dist500,i,j} \] (5)

Where variables remain as defined earlier, and

\[ C_{AV\_ocost} = \text{AV operational cost in dollars per mile (assumed to be $1 per mile for the initial analysis).} \]

Probability splits for mode choice given each origin, destination, and purpose are then found:
\[ P_{m|i,j,p} = \frac{e^{V_{m,i,j,p}}}{\sum_{m \in M} e^{V_{m,i,j,p}}} \] (6)

Where
\[ V_{m,i,j,p} = \text{Utility function for mode } m, \text{ from NUMA } i \text{ to } j \text{ for purpose } p, \]
\[ M \quad = \text{Set of all modes being analyzed.} \]

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. The corresponding portion in the rJourney model also incorporates a series of size ranges and distance ranges expressed as indicators with corresponding coefficients. Again, for simplicity as well as lack of data, parameters that were believed to be difficult to acquire and not strongly influential in mode choice and destination choice were omitted, including number of nights staying. 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 be helpful for future efforts in better representing destination attractiveness.

The following represents the destination choice model that utilizes the mode choice logsum. In particular, Equation 7 is a simplified representation of attractiveness for a given destination. Future research efforts will evaluate how more of the rJourney destination-choice model can be leveraged for arriving at an improved representation of attractiveness.

\[ A_j = \log(\alpha_1 L_{\text{totHH},j} + \alpha_2 L_{\text{totEmp},j}) \] (7)

\[ \text{LOGSUM}_{i,j,p} = \log \sum_{m \in M} e^{V_{m,i,j,p}} \] (8)

\[ P_{j|i,p} = \frac{e^{A_j + \theta_p \text{LOGSUM}_{i,j,p}}}{\sum_k e^{A_k + \theta_p \text{LOGSUM}_{i,k,p}}} \] (9)

Where variables remain as defined earlier, and
\[ A_j \quad = \text{Attractiveness function for a given NUMA } j, \]
\[ \alpha_1 \text{ and } \alpha_2 \quad = \text{Attractiveness coefficients (set to } 1 \text{ in the initial model),} \]
\[ L_{\text{totEmp},j} \quad = \text{Total employment for NUMA } j, \]
\[ \theta_p \quad = \text{Coefficient for mode choice logsum for purpose } p, \]
\[ N \quad = \text{Number of NUMAs being analyzed.} \]

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

\[ P_{m,j|i,p} = P_{j|i,p} P_{m|i,j,p} \] (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 trips to be redistributed according to the post-AV introduction model.

\[ T_{i,j,m,p} = P_{m,j|i,p} \sum_{k}^N R_{i,k,p} \]  

(11)

Where variables remain as defined earlier, and

\[ R_{i,k,p} = \text{Number of trips in the rJourney trips dataset from origin NUMA } i \text{ to destination } j \text{ for purpose } p. \]

Computation of this model can be classified as a big data problem. For 38 million trips over 2 million NUMA pairs, the rJourney data files constitute 4 GB of data, and files that represent the intermediate and final computational results for all trip purposes amount to 54 GB. The Python Pandas library is used to perform the computations. With a number of considerations made for vectorized matrix operations, the entire set takes on the order of 3 hours to run on a modern, high-end computer. Operations that read and write files from disk account for about half of the run time.

RESULTS

Figure 3 shows the resulting number of trips before and after the AV mode is added to the initial model as described in the methodology. A notable observation is that the distribution of AV trips closely tracks the distribution of traditional vehicles, which can be attributed to the use of several parameters that are represented in the traditional vehicles. The key differences with AVs are the increased in ownership and operating cost, and reduced value of time driving. With some similarity in parameters, this mode split may be inevitably influenced by the independence from irrelevant alternatives property (or, noted many times in the literature as the “red bus/blue bus paradox”) inherent in multinomial logit models.

Figure 3: Number of trips from the mode choice/destination choice analysis, all purposes

Regardless of this split, there are two notable outcomes that offer insight on the possible effects of introduction of AVs to the market, as well as a shift in destination choice. First, results show that the introduction of AVs cuts into the number of trips that had formerly been air trips. See Table 2 for results in terms of shorter and longer trips (e.g. < 500 miles versus ≥ 500 miles). As largely
influenced by the $\beta_{\text{car,ood500}}$ coefficient, 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 the same for AVs as it would be for
traditional cars. Note that in Table 2, “Car+AV” is shown as a means to represent respective totals
of personally owned vehicles.

Table 2: Trip mode choice impact of AV introduction for “personal” trip purpose

<table>
<thead>
<tr>
<th></th>
<th>Car+AV &lt; 500 mi.</th>
<th>Car+AV ≥ 500 mi.</th>
<th>Air &lt; 500 mi.</th>
<th>Air ≥ 500 mi.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before AV</td>
<td>3,655,082</td>
<td>1,840,595</td>
<td>127,331</td>
<td>171,909</td>
</tr>
<tr>
<td>After AV</td>
<td>4,194,300</td>
<td>1,435,313</td>
<td>91,244</td>
<td>122,010</td>
</tr>
<tr>
<td>Percentage</td>
<td>114.8%</td>
<td>78.0%</td>
<td>71.7%</td>
<td>71.0%</td>
</tr>
</tbody>
</table>

Second, among traditional cars and new AVs, more destinations are chosen post-introduction of
AVs that are closer in distance from origins than that of pre-AV introduction. Even though for
both pre- and post-AV introduction the model uses the same number of trip generations per NUMA
per trip purpose, the higher cost of ownership and operation cause the model to result in a shift
toward shorter distances. 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 shorter and longer long-distance trips.

These results inform the next areas to investigate within the model. The first is to alter the cost of
driving from that of vehicle ownership plus operation to purely vehicle operation. The rationale
for this is that drivers may more appropriately be modeled to choose modes and destinations using
criteria that do not include cost of vehicle ownership, as though the car ownership is a given
condition. Currently, the inclusion of ownership cost dominates that portion of the model. It is
anticipated that this will allow for less penalty on long-distance trips. Second, the destination-
choice utility function is rudimentary in this initial analysis, and further refinement and completion
of the utility function would likely better model the split between shorter and longer long-distance
trips.

Airline Revenue

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. Table 3 shows estimated airline revenue before and after
the addition of AVs for all modeled trips of “personal” purpose.

Table 3: Airline revenue for “personal” trip purpose, in millions of USD

<table>
<thead>
<tr>
<th></th>
<th>Trips &lt; 500 miles</th>
<th>Trips ≥ 500 miles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before AV</td>
<td>$38.99</td>
<td>$62.28</td>
</tr>
<tr>
<td>After AV</td>
<td>$27.59</td>
<td>$44.13</td>
</tr>
</tbody>
</table>
In this result, the percentages between revenues between shorter and longer long-distance trips are similar. However, with refinements in the mode choice and destination choice models, the split may evolve to show a greater difference.

### 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 to allow 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 results were produced from this research. First, in the initial scenario where the cost of ownership and operation for an AV is assumed to be $1.00 per mile and VOTT is that of air travel, air travel demand for shorter and longer long-distance trips is changed to be 71.7% of the original value, largely replaced with an increased demand for AVs. It follows that commercial passenger air carriers would do well to more fully predict and understand the implications of AV introduction and target their services and marketing accordingly. Second, with the introduction of AVs, trips are generally distributed to closer distances than before. This can be attributed to the higher cost of AV ownership and operation that is modeled. Currently, vehicle ownership cost significantly influences the model. This prompts further work in determining how to best capture the cost of vehicle operation, as it may be more appropriate to assume that because travelers already own their vehicles, they do not consider the cost of ownership in their travel decision-making process.

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. These are all factors that can help to establish a more accurate 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 the 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, and Will Schievelbein’s efforts in coding to initially read in and calculate data.
REFERENCES


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 Compendium of Papers. Transportation Research Board of the National Academies.

Zhao, Yong, Kara M. Kockelman. Anticipating the Regional Impacts of Connected and Automated Vehicle Travel in Austin, Texas Under review for publication in Transportation Research Record (2017).