International Journal of Sustainable Transportation

ISSN: 1556-8318 (Print) 1556-8334 (Online) Journal homepage: https://www.tandfonline.com/loi/ujst20

Factors affecting Atlanta commuters' high occupancy toll lane and carpool choices

Randall Guensler, Joonho Ko, Daejin Kim, Sara Khoeini, Adnan Sheikh & Yanzhi Xu

To cite this article: Randall Guensler, Joonho Ko, Daejin Kim, Sara Khoeini, Adnan Sheikh & Yanzhi Xu (2020) Factors affecting Atlanta commuters' high occupancy toll lane and carpool choices, International Journal of Sustainable Transportation, 14:12, 932-943, DOI: 10.1080/15568318.2019.1663961

To link to this article: https://doi.org/10.1080/15568318.2019.1663961

INTERNATIONAL JOURNAL OF SUSTAINABLE TRANSPORTATION
2020, VOL. 14, NO. 12, 932–943
https://doi.org/10.1080/15568318.2019.1663961

Factors affecting Atlanta commuters' high occupancy toll lane and carpool choices

Randall Guensler[a], Joonho Ko[b], Daejin Kim[a], Sara Khoeini[a], Adnan Sheikh[a], and Yanzhi Xu[a]

[a]School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA; [b]Graduate School of Urban Studies, Hanyang University, Seoul, South Korea

ABSTRACT

As more high occupancy toll (HOT) facilities are planned and under development, a comprehensive understanding of HOT operations is required for establishing appropriate HOT policies. To enhance the understanding, this study attempts to investigate the factors affecting drivers' choices on HOT lane use and carpooling in the Atlanta I-85 HOT corridors. The investigation utilizes a questionnaire-based survey addressed to 12,000 households/commuters in the corridors by employing a mail-out/mail-back method. A total of 642 surveys were retrieved and about 300 surveys after data screening were utilized for developing classification tree and logistic regression models that explain commuters' HOT lane and carpool choices. The estimated models indicate that HOT lane and carpool choices can be affected by various factors including age, gender, income, commute distance, education, number of household workers and car ownership. In addition, the models show that respondents' positive perception of commute condition changes after the HOT installation can substantially increase the chance of using the HOT lanes. The results also imply that the HOT installation cannot always boost carpool formation, requesting policy makers to develop appropriate strategies that encourage drivers to carpool.

ARTICLE HISTORY
Received 21 July 2018
Revised 21 June 2019
Accepted 2 September 2019

KEYWORDS
Carpool; classification tree; high occupancy toll lane; high occupancy vehicle lane; mail survey

1. Introduction

Freeways are crucial elements of transportation systems as they provide travelers with a high level of mobility services. However, freeways in urban areas often experience severe traffic congestion due to growing mobility demand that exceeds facility capacity, worsening cities' economic competitiveness and quality of life (American Highway Users Alliance, 2015). Reducing freeway congestion requires implementing strategies designed to increase facility capacity or reduce transportation demand. To this end, transportation agencies have introduced high occupancy toll (HOT) lanes. HOT lanes typically allow 3-person carpools to use the lanes for free, which limits the demand for the use of the lane. Tolls are then instituted to allow non-carpools to use the lane for a fee, which is the notable difference from high occupancy vehicle (HOV) lanes where single occupancy vehicles are restricted from traveling. With proper variable pricing, which adjusts the toll to ensure demand remains lower than capacity, congestion can be prevented and travel speeds in HOT lanes can be maintained (e.g., 45+ mph more than 90% of the time). Hence, traveling on HOT lanes in particular in a form of carpool or transit is an attractive option during peak hours because of saved time and cost. This aspect suggests that an introduction of HOT lanes can be considered as a measure enhancing the sustainability of transportation systems as they nudge travelers to switch from solo driving to carpooling or transit. The reduced amount of traffic by increasing vehicle occupancy can contribute to energy saving and emissions reduction.

Since California's State Route (SR) 91 Express Lanes opened in December 1995 (the first HOT lane in the US), approximately 20 HOT facilities have been installed (Guensler et al., 2013). In some cases, such as on Atlanta's I-85 corridor, HOT lanes are created from the conversion of HOV lanes. Still, numerous HOT facilities are under development across the county. To improve operations of the current facilities and design of future facilities, transportation agencies need to understand the various impacts and travelers' behavior changes resulting from the introduction of HOT lanes.

In order to enhance the understanding, this study investigates factors affecting drivers' choices on HOT lane use and carpooling in the Atlanta I-85 HOT corridor. Although this research is not the first effort to assess factors that affect HOT operations in the I-85 corridor, it is unique in that it utilizes data obtained from a survey administered to commuters who traveled the corridors. Indeed, previous studies

CONTACT Joonho Ko jko@hanyang.ac.kr Graduate School of Urban Studies, Hanyang University, 222 Wangsimni-ro Seongdong-gu, Seoul, 04763, South Korea.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ujst.

utilized marketing (credit report) data for identifying the socio-economic characteristics of drivers traveling on HOT or regular general purpose (GP) lanes (Khoeini & Guensler, 2014a, 2014b; Sheikh, Misra, & Guensler, 2015). Their approach is advantageous in obtaining a sizable data set at a low cost, but one issue of uncertainty appears to be unavoidable; driver/household information in the marketing data will not always match those of the observed drivers. This study that employs a survey approach is expected to avoid the issue as the survey data allow for directly relating drivers' responses to their perception, socio-economic characteristics and trip patterns (e.g., commute distance and work start time).

As well as the impacts of drivers' socio-demographic and travel characteristics, this study attempts to capture the impacts of two factors: drivers' perception of HOT performance and whether the driver was a carpooler or HOV lane user before HOT implementation. Their impacts have rarely been examined in the previous studies. It should be also noted that carpool behavior after the installation of a conversion of an HOV lane to a HOT lane has not yet been handled in the previous studies. The findings from this study will add knowledge to the literature on drivers' behavior concerning HOT lane and carpool choices in HOT corridors.

2. Literature review

The existing studies concerning HOT lanes have usually focused on their operational impacts. One of the potential impacts identified is a change in transit ridership. Previous studies demonstrated that not a small portion of new bus riders in HOT lane corridors (for example, 23% in Minneapolis and 53% in Miami) were influenced to take transit by HOT lanes, resulting in an increase in transit ridership (Pessaro, Turnbull, & Zimmerman, 2013). However, the bus ridership increase seems not to be always guaranteed by the introduction of HOT lanes. Castrillon, Roell, Khoeini, and Guensler (2014) reported that person throughput for commuter buses remained stable even with an 18% increase in express bus throughput in the Atlanta I-85 HOT corridor. By comparing the numbers of carpools before and after HOT lane implementations, Burris, Alemazkoor, Benz, and Wood (2014) and Goel and Burris (2012) revealed that HOT lanes tended to decrease carpooling.

It should be noted that the impacts can somewhat vary by the characteristics of the HOT facility (e.g., HOT lane capacity) and exogenous factors such as gas prices, as pointed out by Burris et al. (2014). Transportation demand management (TDM) activities (e.g., carpooling, vanpooling, and transit use) and toll exemption policies can also influence travelers' choices, which in turn affect traffic conditions on both managed lanes and regular GP lanes. Pessaro and Buddenbrock (2015) illustrated how such impacts are expected for the I-95 Express Lanes using scenario-based traffic simulation approaches. This aspect requires ongoing research efforts on HOT operations, so that transportation agencies can gain a comprehensive understanding that helps in establishing appropriate HOT design and operation policies that can enhance the sustainability of the facility.

The review reveals that the decision to travel in HOT lanes can be affected by drivers' socio-economic characteristics, such as gender, age, income, household size, vehicle ownership, and education level (Khoeini & Guensler, 2014a; Sheikh et al., 2015). Khoeini and Guensler (2014b) also explored this issue using vehicle value as a proxy for income. They showed that the value of vehicles using HOT lanes is approximately 23% higher than that of vehicles using regular GP lanes in the case of the Atlanta I-85 HOT corridor. This result implies that high-income people are more likely to use HOT lanes, as was also found in the study of the SR 91 Express Lanes (Li, 2001). The SR 91 study concluded that household income, vehicle occupancy, commute trip, and age are important predictors of HOT lane use. Meanwhile, the study identified that such factors as gender, trip length, trip frequency and household size and type (family with or without children) are not significant. Building on these studies, further research is encouraged to firmly understand drivers' behavior on HOT lane choices.

3. Study corridor: the Atlanta I-85 HOT lanes

The spatial scope of this study is the Atlanta I-85 HOT lanes. The HOT lanes were installed by converting existing HOV lanes over a 15.5-mile length. The HOV-to-HOT conversion is the first to simultaneously introduce tolling while increasing the occupancy requirement (from HOV2+ to HOV3+), but did not add additional lanes (Guensler et al., 2013). The HOV2+ lane still exists just to the south of the HOT lane corridor, extending into Downtown Atlanta. Since the HOT lane opened on October 1, 2011, dynamic tolling varies toll price in response to congestion. Toll exempt vehicles include vehicles carrying three or more persons (HOT3+), transit vehicles, emergency vehicles, motorcycles, and alternative fuel vehicles with proper license plates (hybrid vehicles do not qualify). All vehicles must register for a Peach Pass toll tag, even if they are toll-exempt, so that corridor activity can be monitored.

The field survey data collected over the corridor showed that the number of vehicles traveling as HOV2+ in the HOT lanes decreased after the conversion, from 3,966 to 613 average weekday travelers during the morning peak period and from 3,941 to 697 travelers during the evening peak period. Meanwhile, the number of HOV2+ carpoolers in the regular GP lanes about doubled (Burris et al., 2014; Guensler et al., 2013). That is, carpools shifted out of the managed lane into the GP lane. Concerning this situation, Guensler, Xu, Sheikh, Li, and Khoeini (2015) identified that the 2-person carpoolers considered the toll cost is too high, perceiving that the amount of saved travel time is not worth the cost. In addition, some carpoolers complained that it is too difficult to get into and out of the express lanes. Overall, morning commute carpooling on the corridor after HOV-to-HOT conversion decreased by more than 30% (Guensler

Table 1. Respondents' mode and route choices before and after HOT installation (before data screening).

Mode and lane choice		After: Drive alone GP	Drive alone HOT	Carpool GP	Carpool HOT	Other[a]	Total
Before	Drive alone GP	232	99	13	10	13	367
	Carpool GP	17	5	16	9	2	49
	Carpool HOV	28	14	53	35	4	134
	Other[a]	1	1	0	0	73	75
Total		278	119	82	54	92	625

Note: HOT, high occupancy toll; HOV, high occupancy vehicle; GP, general purpose.
[a]This category includes walk, bicycle, transit, local roads and work at home.

et al., 2013). Identifying those who are likely to carpool in the HOT lanes, will help inform future managed lane operational strategies.

4. Survey data

A questionnaire-based survey was designed to explore the behavioral changes of the travelers along the I-85 HOT lane corridor. The first hurdle of the survey was to identify a sample pool, given that HOT lane users and carpoolers tend to constitute only a small portion of the overall traveling public. Fortunately, the research team had collected about 1.5 million license plates of the vehicles traveling the I-85 corridor one year before and one year after the HOV-to-HOT conversion. The collected data allowed the identification of households that retain frequent HOV/HOT users of the corridor (Khoeini & Guensler, 2014b). Based on the database, 10,000 survey targets were selected and questionnaires were mailed out in the form of an eight-panel folded sheet with a prepaid return envelope. It was later discovered that an issue with the user database resulted in incorrect names printed in the survey address, which might have affected the response rate. Numerous surveys that were completed included notes indicating that the name on the form was incorrect. Given the potential problem, a second stage involved sending out 2,000 additional surveys to households that were not previously targeted in the initial deployment. This second stage survey was designed to check if the previous problem (incorrect name on the survey form) affected the response rate. The research team conducted the mail-out/mail-back survey in November and December 2014 and obtained 642 responses among the 12,000 target households (a retrieval rate of 5.4%). The response rates were roughly equal during the both stages, indicating that the name errors in the first stage probably did not significantly influence the response rates.

4.1. Survey questionnaire

The survey questionnaire was composed of four general sections, asking:

1. Primary routes and modes for morning commute before and after the HOT lane implementation,
2. Perception of the HOT lane effectiveness on their commute traffic conditions,
3. Reasons why the respondent chose to use or not use the HOT lanes, or to carpool, after the HOT lane implementation, and
4. Individual and household socioeconomic/demographic characteristics.

The questions about the lane choices include: use of HOV lanes or regular GP lanes before HOT implementation, and use of HOT lanes or regular GP lanes after HOT implementation. The perception was measured by a question ("Have the HOT lanes improved your own commute conditions?") and respondents were requested to answer based on a 5-point rating scale from "definitely yes" to "definitely no." The socioeconomic and demographic questions include age, gender, household income, number of children, number of workers, car ownership, education, work start time and job locations described by zip code. Based on the job location, home to work distances were estimated by utilizing the time-based shortest path between the zip codes of their home and work place with the help from the function of the Google Maps API. As a unique element of this study, the identified information on whether the respondent is a former HOV user and/or former carpooler, the perception and the socioeconomic characteristics were employed to explain commuters' HOT lane and carpool choices.

4.2. Data selection

Self-administered mail-out/mail-back surveys are susceptible to missing values and inconsistency of answers, requiring a careful data selection procedure. During an initial check, it was found that 17 respondents (2.6% of the retrieved 642 replies) did not properly provide their mode and route/lane choices. After excluding the 17 respondents, a table illustrating changes in mode and lane choices (considering all the combinations of carpool, drive alone, HOT and regular GP lanes) after the HOT lane installation was developed as shown in Table 1. The table, unfortunately, revealed that the travel patterns of 73 respondents are irrelevant to this study as they did not drive on the freeway either before or after HOT implementation.

Further data screening procedures took into account whether the choices are multiple (for example, cases in which respondents marked both HOT and regular GP lanes for their usual travel lanes) and whether respondents answered all the questions related to the explanatory variables discussed in the previous section. The multiple choices can be regarded as normal for some drivers, but it is highly

Table 2. Respondents' mode and route choices before and after HOT installation (after data screening).

Mode and Lane choices		After: Drive alone GP	Drive alone HOT	Carpool GP	Carpool HOT	Other[a]	Total
Before	Drive alone GP	138	59	0	2	3	202
	Carpool GP	1	0	2	2	0	5
	Carpool HOV	12	7	28	13	1	61
	Other[a]	0	0	0	0	3	3
Total		151	66	30	17	7	271

Note: HOT, high occupancy toll; HOV, high occupancy vehicle; GP, general purpose
[a]This category includes walk, bicycle, transit, local roads and work at home.

Table 3. Demographic and opinion responses in the sample.

Variables		HOT lane choice ($n = 313$) Observations	HOT lane choice Proportion	Carpool choice ($n = 332$) Observations	Carpool choice Proportion	Metro Atlanta Proportion[b]
Gender	Male	168	53.7%	178	53.6%	47.8%
	Female	145	46.3%	154	46.4%	52.2%
Age	<40	47	15.0%	48	14.5%	42.9%
	40–50	103	32.9%	111	33.4%	21.3%
	>50	163	52.1%	173	52.1%	35.8%
Annual Household Income	<$60k	31	9.9%	38	11.4%	51.8%
	$60–$100k	87	27.8%	99	29.8%	24.0%
	>$100k	195	62.3%	195	58.7%	24.2%
Number of Children	0	165	52.7%	173	52.1%	61.8%
	1	64	20.4%	64	19.3%	16.2%
	2+	84	26.8%	95	28.6%	22.0%
Number of Workers	1	83	26.5%	91	27.4%	54.9%
	2	180	57.5%	185	55.7%	38.0%
	3+	50	16.0%	56	16.9%	7.2%
Number of vehicles for commute	1	130	41.5%	136	41.0%	35.3%
	2	144	46.0%	154	46.4%	43.0%
	3+	39	12.5%	42	12.7%	21.7%
Education	Less than a bachelor's degree	76	24.3%	80	24.1%	35.4%
	Bachelors	131	41.9%	141	42.5%	64.6%
	Masters/doctorate	106	33.9%	111	33.4%	
Typical Work Start Time	Before 7 a.m.	27	8.6%	28	8.4%	—
	7–9 a.m.	256	81.8%	270	81.3%	—
	After 9 a.m.	30	9.6%	34	10.2%	—
Commute Distance (miles)	<20	47	15.0%	50	15.1%	—
	20–30	107	34.2%	116	34.9%	—
	>30	159	50.8%	166	50.0%	—
Were you a HOV lane user or carpooler?[a]	No	231	73.8%	251	75.6%	—
	Yes	82	26.2%	81	24.4%	—
Have the HOT Lanes improved your own commute conditions?	Definitely no	151	48.2%	159	47.9%	—
	Probably no	32	10.2%	34	10.2%	—
	Not sure	14	4.5%	15	4.5%	—
	Probably yes	50	16.0%	58	17.5%	—
	Definitely yes	66	21.1%	66	19.9%	—

Note: HOT, high occupancy toll; HOV, high occupancy vehicle.
[a]The question of the HOV lane use is applied for the HOT lane choice sample, while the carpooling question is for the carpool choice sample.
[b]*Source*: US Census Bureau, 2010 5-Year American Community Survey.

complex to include such answers in a statistical model. The screening procedure removed cases with multiple choices or missing values, resulting in a significant loss of data from 625 to 271 (Tables 1 and 2). Concerning the missing values, approximately 150 respondents did not provide any personal information. In total, 58% (371 out of 642) of the retrieved surveys were not usable for choice-based analysis, where all explanatory variables associated with choice need to be entered into the statistical models.

An attempt to develop a multinomial logistic regression model to predict post-opening travel for former carpoolers considered four choices (drive alone in the regular GP lanes, drive alone in the HOT lane, carpool in the regular GP lanes, and carpool in the HOT lane), but the results were unsatisfactory. The main reason for the unsatisfactory result was most likely the small sample size. Multinomial regression using a maximum likelihood estimation method usually requires an even larger sample size than ordinal or binary logistic regression (Agresti, 1996). Given this situation, the authors developed binary choice models separately by route (HOT or regular GP lanes) and mode choice (carpool or drive alone) made after the HOT lane began operating. The route and mode choices before the HOT installation were also utilized as independent variables.

To minimize the loss of data, the authors conducted separate data selection procedures for HOT lane and carpool choice models. This is because more samples are likely to be screened out when HOT lane and carpool choices are simultaneously considered. As previously demonstrated, the procedure screened out cases with multiple choices for both HOT and regular GP lanes (likewise carpool and drive alone) and missing values for the explanatory variables,

Table 4. Respondents' behavioral changes in HOT lane and carpool choices.

HOT lane choice				Carpool choice			
	After				After		
Before	HOT lane	GP lanes	Total	Before	Carpool	Drive alone	Total
HOV lane	29	53	82	Carpool	56	25	81
GP lanes	78	153	231	Drive alone	7	244	251
Total	107	206	313	Total	63	269	332

Note: HOT, high occupancy toll; HOV, high occupancy vehicle; GP, general purpose.

resulting in 313 and 332 valid cases for HOT lane and carpool choice models, respectively.

Table 3 summarizes the selected sample characteristics based on 11 factors including socio-demographic characteristics, commute option, and perception on HOT lane, showing that the two data sets are very similar. This is not surprising, given that the two data sets share 285 identical respondents (91% and 86% of the data sets for the HOT lane and carpool choices respectively). The sample is composed of slightly more males than females. More than half of the respondents are older than 50 years (about 52%). More than 60% of the respondents belong to a high-income group, above USD $100,000 per year. The income distribution seems to be reasonable since the sample contains a group of HOT lane users who are likely to have a higher value of time (Khoeini & Guensler, 2014b). More than half of the households have children. Single-worker households comprise less than 30% of the sample, with two-worker households being dominant. About 60% of households own multiple cars for commuting. About 75% of the respondents have a Bachelors' degree or higher. Compared to the general population of Metro Atlanta, the sample has higher proportions of males, older people, high-income households and educated people. The sample also tends to have more children and workers in the household. Interestingly, the number of vehicles for commute is smaller compared to that of the general population.

It appears that most respondents (about 80%) start their work between 7 and 9 a.m. and about half of the respondents commute more than 30 miles. In the HOT lane choice sample, 26.2% of the respondents replied that they usually used HOV lanes before the HOT lanes opened. In the carpool choice sample, former carpoolers occupy 24.4%. With respect to respondent opinion about whether the HOT lanes have improved their commutes, 58% were negative (definitely no and probably no), about 37% were positive (definitely yes, and probably yes), and less than 5% were not sure. Note, however, that these percentages do not control for whether the respondents are or are not regular HOT users. In fact, it turns out that users are generally positive and non-users are generally negative.

The selected respondents' behavioral changes in HOT lane and carpool choices are summarized in Table 4. The table shows that 34% (107 out of 313) of the respondents usually use the HOT lanes, while the remaining 66% of respondents are regular GP lane users. In addition, it indicates that 65% (53 out of 82) of the former HOV lane users switched to the regular GP lanes. Concerning the carpool choice, 19% (63 out of 332) of the respondents usually carpool while the remaining 81% drive alone for commuting. Changes in carpooling behavior are also observed. Responses indicate that carpool breakups outpaced carpool formation. Carpool formation was only 2.8% (7 out of 251) while 31% (25 out of 81) of former carpoolers left their carpools. The data indicate that 89% (56 out of 63) of the remaining carpoolers were former carpoolers.

5. Analytical approaches

Two approaches, classification trees and logistic regressions, were used for developing statistical models designed to explain drivers' behavioral responses in their commute travel. The approaches can be used to estimate the class membership of a categorical dependent variable (Camdeviren, Yazici, Akkus, Bugdayci, & Sungur, 2007). Indeed, this study uses binary dependent variables by assigning an indicator value of one for cases where respondents choose HOT lanes (or carpool lanes), and zero otherwise.

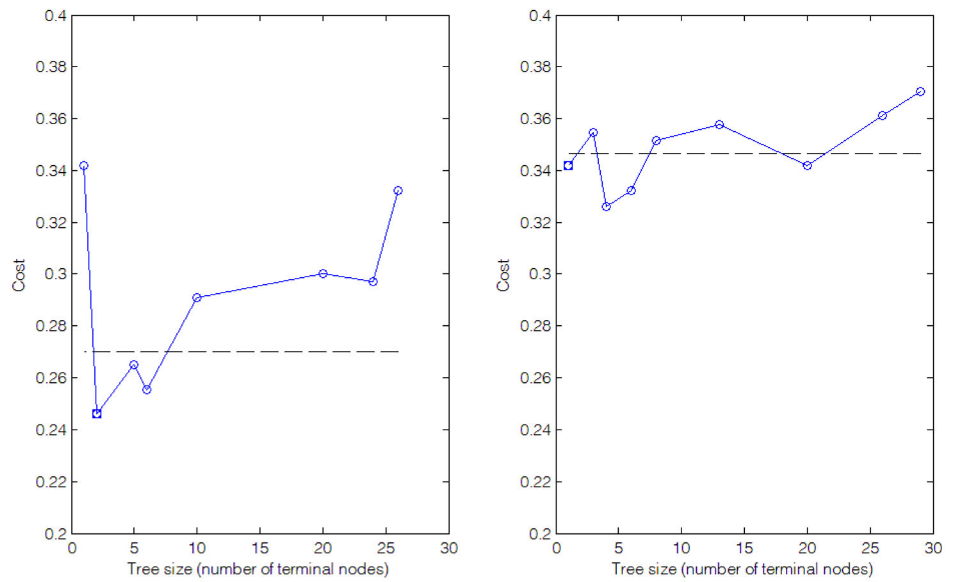
5.1. Classification trees

To obtain a better understanding of commuter characteristics, a multi-dimensional analysis considering interactions between factors was conducted using the tree-based regression and classification technique. This approach is attractive because the resultant trees provide a symbolic representation that lends itself to easy human interpretation (Camdeviren et al., 2007). In particular, this study applies classification trees to discrete dependent choices of HOT lane use (or carpool lane use), with the selected independent variables. The technique splits the data through a binary partition, thus generating two resultant regions. As the partitioning process continues, the tree tends to grow, resulting in over-fitted and complicated models. Meanwhile, a tree that is too small might not capture the important structure of the data. Thus, an optimal tree size should be adaptively chosen from the data.

This study utilizes the cross-validation technique in finding an optimal tree. In the approach, the cost of the tree by tree size is computed based on the 10-fold cross-validation method (Breiman, Friedman, Stone, & Olshen, 1984; Hastie, Tibshirani, & Freidman, 2001). The cost is the sum over all terminal nodes of the estimated probability of that node times the sum of the misclassification errors of the observations in that node. The best tree size, or the number of terminal nodes, is the one that produces the smallest tree that is within one standard error of the minimum-cost subtree.

5.2. Logistic regression

Logistic regression models were also applied to identify the factors affecting drivers' choices of HOT lanes and carpooling in the I-85 corridor. In the model, the response variable has only two possible outcomes, whether the respondent usually uses HOT lanes or whether they do not. When Y_i is an independent Bernoulli random variable for the i^{th} observation with an expected value E$\{Y_i\}$, the logistic regression model with k predictor variables, known constants x and



(a) Models for all the 11 Factors (b) Models without the Perception Variable

Figure 1. Classification tree cost (error) by tree size for HOT lane choice. (a) Models for all the 11 factors; (b) models without the perception variable. Note: The dashed line indicates one standard error of the minimum-cost subtree.

coefficients to be estimated β is expressed as follows (Kutner, Nachtsheim, Neter, & Li, 2005):

$$E\{Y_i\} = \frac{\exp(\beta_{0+}\beta_1 x_{i1} + \ldots + \beta_k x_{ik})}{1 + \exp(\beta_{0+}\beta_1 x_{i1} + \ldots + \beta_k x_{ik})}.$$

The interpretation of the estimated regression coefficients in the fitted logistic response function is not straightforward as in a linear regression model. The reason is that the effect of a unit increase in predictor variables varies depending on the location of the starting point on the predictor variable scale (Kutner et al., 2005). Thus, the odds ratio, which is computed by taking the exponent value of the estimated coefficient, is used for associating the outcome with explanatory variables. Odds ratios above one indicate that the event is more likely to occur while odds ratios smaller than one indicate lower chances of the event to occur.

Kim (2009) showed that logistic regression models can be more efficiently developed by utilizing the results of the tree-based regression and classification technique. This is because classification trees may reveal statistically meaningful interactions between the explanatory variables, helping analysts identify which interaction effects should be entered in the regression models. In particular, the approach is substantially helpful when numerous and complex interaction effects may exist.

6. Results

6.1. HOT lane choice classification trees

Classification tree analyses considering the eleven factors were performed using MATLAB 2015b to assess HOT lane choices. Firstly, the costs of the models by the tree size were estimated based on the 10-fold cross-validation approach to identify the best tree size. The graphs in Figure 1 show the estimated costs for the two models using all the 11 factors and excluding the variable of the perception of the effectiveness of the HOT lanes. The graphs imply that the perception variable has a substantially strong explanatory power, which can be explained in two ways. First, when all the 11 factors are considered, the cost is minimized at two terminal nodes, with the perception being the single variable dividing the choices. Second, the misclassification errors (costs) become much larger when the perception variable is excluded, which can be easily identified by comparing the costs in the two graphs in Figure 1.

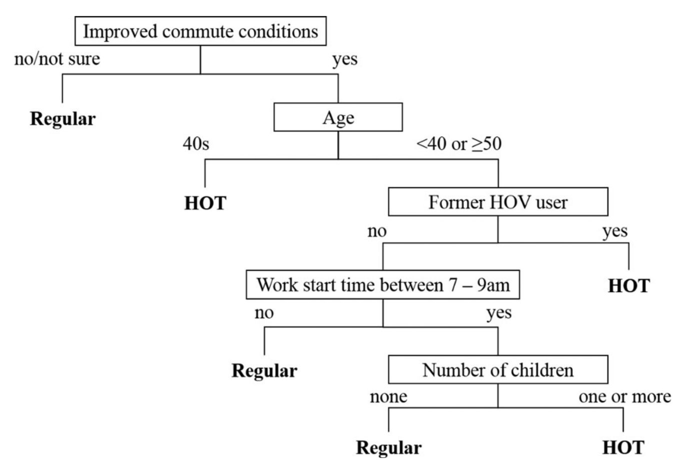


Figure 2. Classification tree with six terminal nodes for the HOT lane choice (including the perception variable).

The estimated costs in Figure 1 indicate that the best tree for the model with the perception variable has only two terminal nodes. The two-node model, however, may fail to capture the important aspects of the choices because of its simplicity. Thus, a classification tree with six terminal nodes, the second best tree in terms of the cost, was developed as an alternative for explaining the lane choices. The developed tree with six terminal nodes is shown in Figure 2,

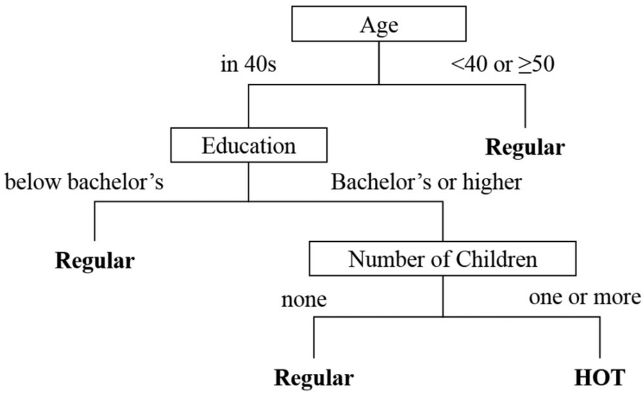


Figure 3. Classification tree for HOT lane choice (excluding the perception variable).

illustrating that five variables are critical factors: perception of benefit, age, former HOV user, work start time and number of children. The tree implies that the respondents who do not perceive the HOT lanes have improved their own commute conditions are more likely to choose regular GP lanes instead of the HOT lanes. Of the respondents who perceive the positive effects of the HOT lanes, the ones in their age of 40s are more likely to choose the HOT lanes. In addition, the model implies that the respondents who usually used HOV lanes are more likely to use the HOT lanes. The respondents who usually start to work between 7 and 9 a.m. and have children, likely time-constrained commuters during morning peak hours, are also found to have a stronger tendency to choose the HOT lanes.

Because of the dominant impact of the perception variable, the influences of other variables may be concealed. Thus, an examination of a classification tree without the perception variable is also of interest. As suggested by the model costs shown in Figure 1, a tree with four terminal nodes was constructed as the best model for the perception-excluded data. Figure 3 illustrates the tree depicted by three variables: age, education, and number of children. Unlike the previous model, the education variable is found to be an important factor; the respondents with a bachelor's degree or higher are more likely to choose the HOT lanes. It is conjectured that the education level may reflect the financial ability of the respondents to pay a toll.

6.2. HOT lane choice logistic regression

It seems that the constructed classification trees successfully identified potential relationships between lane choices and influential factors. However, they do not appear to be sufficient to show the factors' statistical significances in a measurable way. To overcome this situation, further investigations were attempted by developing three logistic regression models using IBM SPSS Statistics 23 with different independent variables: models with main effects only (Model 1), with both main and interaction effects (Model 2), and without the perception variable (Model 3). The results of the classification trees were fully utilized when identifying the appropriate forms of independent variables. More specifically, the main effect variables were re-defined based on the cut points revealed by the classification trees, treating them as categorical variables. As a result, all 11 factors were simply classified into binary cases. Moreover, potentially influential interaction terms were identified in an efficient and effective way based on the resultant classification trees. In other words, without the information provided by the classification trees, potential 55 interaction terms, all combinations of two from 11 main effects, should have been considered for specifying the interaction models. However, the interactions of the factors revealed by the classification trees significantly reduce the number of interaction terms to be entered in the model to a practically implementable level.

The resultant logistic regression models were summarized in Table 5, where only statistically significant variables at a significance level of 0.10 were captured based on a backward stepwise procedure, eliminating variables that do not add explanatory power to the model. This stepwise procedure is beneficial to systematically exclude correlated independent variables (Kutner et al., 2005). The table also shows the Hosmer-Lemeshow goodness-of-fit statistics, of which p-values are at least 0.181, implying that the estimated models properly follow the key property of the logistic response function at a significant level of 0.05. The Nagelkerke R-square statistics suggest that Model 2 considering main and interaction effects together has the strongest explanatory power among the three models although the difference from Model 1 appears to be marginal. This justifies the inclusion of the interaction effects. Moreover, the inclusion can be beneficial in that the interaction terms can present useful perspectives for those aiming at improving HOT lanes use. Meanwhile, the model excluding the perception variable has the least explanatory power, indicating the variable's influential impact on the lane choices as already revealed in the classification tree analyses.

The model considering only main effects captured four statistically significant variables: the perception, former HOV user, commute distance, and age. In particular, the odds ratio for the perception variable indicates that respondents are about 11 times more likely to use the HOT lanes when they positively perceive the effectiveness of the HOT lanes. Age also appears to be influential in the lane choice decision; respondents in their 40s are 2.8 times more likely to choose the HOT lanes than respondents in other age groups. In addition, commute distance, which was not a significant factor in the classification tree analyses, was found to be a critical one although its impact is rather weaker compared to the other three factors. The longer-distance commuters, particularly longer than 30 miles (48 km), have a stronger tendency to use the HOT lanes. This finding may be ascribed to the aspect that the longer travelers can gain more travel time-saving benefits by traveling on the HOT lanes during congested peak hours.

When the interaction effects are considered, five variables, including two main effects (the perception and commute distance) and three interaction terms (combinations of former HOV user, age, and perception) are found to be significant. Interestingly, the odds ratio for the perception variable decreases by about half from 10.880 to 5.246, compared to Model 1, although perception is still significantly meaningful in explaining the choices. It seems that the explanatory power of the perception

Table 5. Logistic regression models for the HOT lane choices.

	Model 1 (main effects only)			Model 2 (main + interaction effects)			Model 3 (without the perception variable)		
Variable	*B*	*p*	Exp(*B*)	*B*	*p*	Exp(*B*)	*B*	*p*	Exp(*B*)
Constant	−2.592	.000	.075	−2.097	.000	.123	−1.749	.000	.174
Main effects									
Improved commute conditions (yes = 1)	2.387	.000	10.880	1.657	.000	5.246			
Former HOV user	.851	.013	2.341						
Commute distance (>30 mile)	.610	.033	1.840	.639	.027	1.894	.787	.002	2.196
Age in 40s	1.026	.001	2.791						
Annual household income (>USD $100k)							.543	.045	1.721
Interaction effects									
Former HOV user and Improved commute conditions				1.589	.053	4.901			
Age in 40s and Improved commute conditions				1.671	.001	5.318			
Age in 40s and Former HOV user				1.180	.013	3.254			
Age in 40s and Bachelor's degree or higher							1.263	.000	3.535
	Nagelkerke R^2 = 0.353 Hosmer Lemeshow = 10.135 (*p* =.181)			Nagelkerke R^2 = 0.364 Hosmer Lemeshow = 1.248 (*p* =.940)			Nagelkerke R^2 = 0.145 Hosmer Lemeshow = 2.552 (*p* =.769)		

Note: HOV, high occupancy vehicle.

variable is dispersed over the two interaction terms combined with former HOV user and age. In fact, the interaction terms, former HOV user by perception, and age by perception, have relatively high odds ratios, 4.901 and 5.318, respectively. Model 2 also shows that two main effects of former HOV user and age in 40s are no longer significant by themselves. Instead, they appear to be significant only when they are combined with other factors, implying a simple consideration of main effects may fail to fully capture the characteristics of the data. A potential benefit of the model with interaction terms seems to be its enhanced capability to predict the choices more specifically.

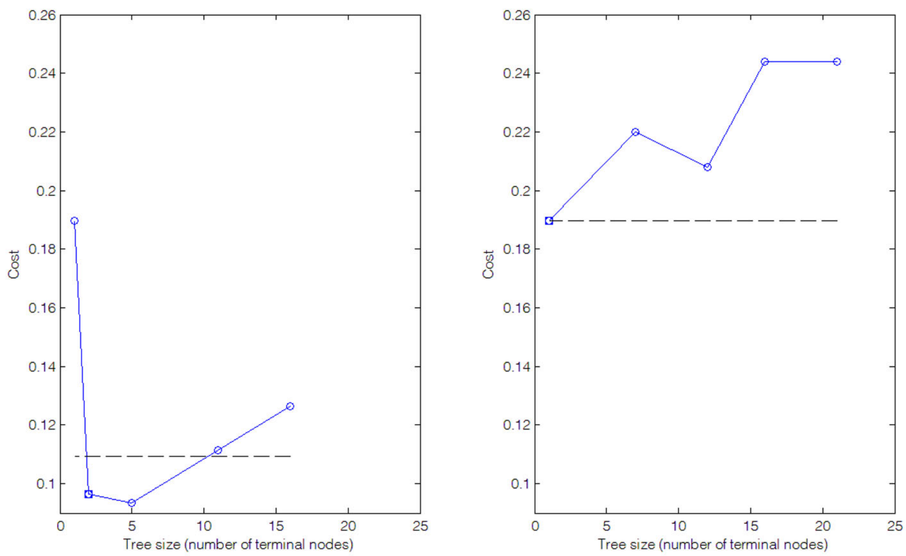
The model excluding the perception variable reveals additional significant variables not shown in the previous models: household income and age by education. HOT lane positive perception may be related to some extent to these variables, perhaps tied to employment in some way. In the later section, a model is presented to show the relationship between the perception and other variables. However, the substantially lowered explanatory power of the model measured by Nagelkerke R^2 (from 0.364 to 0.145) indicates that the variables cannot fully replace the perception variable in explaining the lane choices. This aspect may justify the use of the perception variable for the model development. The estimated model shows that the respondents with a high income and a higher education degree are more likely to choose the HOT lanes. In particular, the respondents in their 40s and with a bachelor's degree or higher are found to have 3.5 times more chances to use the HOT lanes.

6.3. Carpool choice classification trees

Classification trees were developed to analyze the commuters' carpool choices using the selected 332 samples. As illustrated in the HOT lane choice models, the best tree size was first identified using the cost functions of the trees. The cost changes of the trees considering all the 11 factors, shown in Figure 4, indicate that a tree with two or five terminal nodes may be adequate for explaining the carpool choice behavior. When the two-terminal node tree was considered, the former carpooler variable was found to be the single factor predicting carpool choice, given that the majority of prior carpoolers are still carpooling. Indeed, the cost graph in Figure 4 illustrates that the cost becomes much larger when the former carpooler variable was excluded. Although the two-terminal node tree is meaningful, its simplicity may fail to provide sufficient information on the data structure. Thus, a five terminal node tree was selected for analyzing the data. Meanwhile, when constructing a tree without the former carpooler variable, seven terminal nodes were considered as suggested by the tree costs in Figure 4.

Figure 5 illustrates a tree with five terminal nodes depicted by four factors: former carpooler, number of workers, age, and income. The tree strongly supports that respondents are less likely to carpool unless they were already carpoolers before the HOT implementation. In fact, the importance of the former carpooler variable was expected by the sample characteristics; 89% of the carpoolers are the former carpoolers (see Table 4). The figure also shows that even among the former carpoolers, respondents in their 40s whose households have a single worker are more likely to drive alone for their commutes.

It is plausible that under a situation in which one single variable explains almost all of the variability, the effects of other important factors can be obscured. Thus, further analyses were conducted by developing an additional classification tree without the former carpooler variable.



(a) Models for all 11 Factors (b) Models without the Former Carpool User Variable

Figure 4. Classification tree cost (error) by tree size for carpool choice. (a) Models for all 11 factors; (b) models without the former carpool user variable. *Note*: The dashed line indicates one standard error of the minimum-cost subtree.

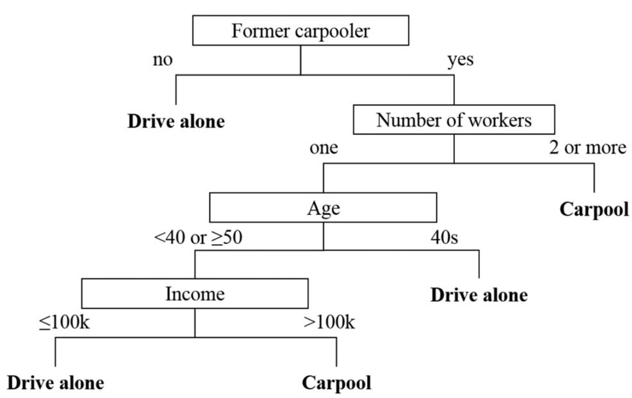


Figure 5. Classification tree for carpool choice.

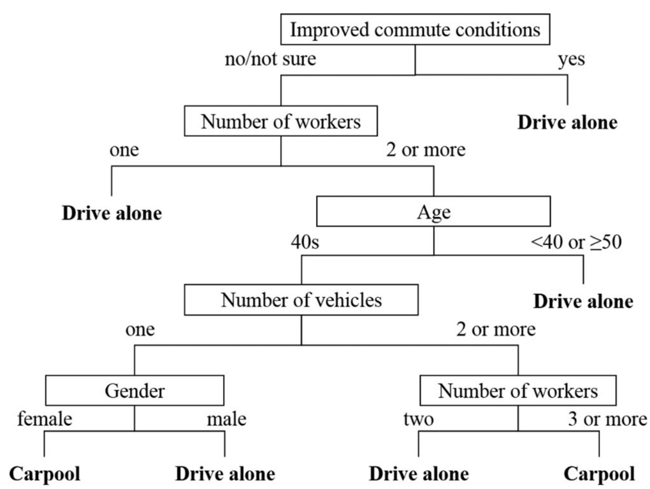


Figure 6. Classification Tree for Carpool Choice (Excluding the Former Carpooler Variable).

Interestingly, the developed tree with seven terminal nodes, shown in Figure 6, illustrates that the perception of the effectiveness of the HOT lanes has the strongest impact on the carpool choice decision, pointing out that the respondents who have a positive perception of the HOT lanes are less likely to carpool. It should be noted that the positive perception of the HOT lanes is also correlated with higher chances of using the HOT lanes, as identified in the HOT lane choice models. Thus, HOT lane use and carpool choices are negatively associated, at least for the I-85 HOT corridor commuters. The tree revealed that the number of vehicles for commuting can play a role in determining the decision. However, the influence of vehicle ownership varies depending on subtree factors such as gender and number of workers. The tree structure also indicates that females are more likely to carpool.

6.4. Carpool choice logistic regression

Further analyses were performed to examine the factors that may influence carpool choice using logistic regression models. The model specifications and procedures were identical to those of the HOT lane choice models, except that the number of worker variable was reclassified to have three groups (1, 2, and 3+), reflecting the cut points suggested in the classification trees. In addition, the number of vehicles available for commuting was not used as an explanatory variable, because the variable was found to be significantly correlated with the number of household workers (Pearson's correlation coefficient = 0.625).

The model considering only main effects revealed that the former carpooler variable is the single dominant factor at a significance level of 0.05 with a Nagelkerke R^2 value of 0.611. (This model is not reported for conserving space.) This may be incurred by the data characteristics; a majority of the carpoolers are the former carpoolers (56 out of 63). This feature became more pronounced in another model which considers both the main and interaction effects. (This model is also not reported for conserving space.) Indeed, the

Table 6. Logistic regression models for carpool choice (excluding the former carpool user variable).

Variable	*B*	*p*	Exp(*B*)
Constant	−2.914	.000	.054
Main effects			
Work start time between 7 and 9 a.m.	.822	.082	2.275
Number of household workers (reference = 1)			
2	1.102	.008	3.010
3+	.978	.040	2.659
Age in 40s	.690	.038	1.993
Interaction effects			
Age in 40s and improved commute conditions	−2.182	.051	.113
Number of household workers = 2, and improved commute conditions	−1.176	.052	.308
Nagelkerke R^2 = 0.165 Hosmer Lemeshow = 5.362 (p = .498)			

Table 7. Ordered Probit Models for the Perception of Improved Commute Conditions.

		Model with all variables		Model with significant variables only	
Variables		*B*	*p*	*B*	*p*
Age (reference = ≥50)	Under 40	−.277	.164		
	40–49	−.055	.733		
Gender (female)		.279	.050	.250	.073
Single worker household		.360	.020	.370	.016
No children		−.124	.391		
Annual household income (>USD $100k)		.320	.034	.412	.003
Bachelor's degree or higher		.163	.324		
Former HOV user		−.718	.000	−.731	.000
Work start time between 7 and 9 a.m.		−.344	.046	−.324	.057
Commute distance (mile)		.015	.014	.014	.026
Threshold $\tau1$		.381	.265	.401	.174
Threshold $\tau2$		.672	.050	.691	.020
Threshold $\tau3$		.805	.019	.824	.006
Threshold $\tau4$		1.336	.000	1.350	.000
Goodness of fit					
−2 log likelihood of null constant only model		836.422		819.551	
−2 log likelihood of full model		783.316		769.475	
		$p < .001$		$p < .001$	
Nagelkerke R^2		0.167		0.159	

Note: τ_j ($j = 1, 2, 3, 4$) is the threshold parameter (cut-off point) for ordered probit models.
HOV, high occupancy vehicle.

estimated parameters in that model seem to be inflated, implying the maximum likelihood estimates are not properly obtained. This situation clearly indicates that the data have a separation problem which occasionally happens in logistic or probit regressions (Heinze & Schemper, 2002). In other words, the former carpooler variable separates the carpool choices almost completely except for seven cases. When separation occurs, two approaches are frequently employed: 1) "mechanical" measures including increasing sample size, combining the category, with similar ones and omitting the category, and 2) statistical measures such as Firth's penalized maximum likelihood method (Gim & Ko, 2017).

This study developed a carpool choice model by omitting the former carpooler variable (one of the common "mechanical measures") to be consistent with the HOT lane choice models. Also, it was conjectured that the use of this influential variable would obscure the impacts of other important factors, in particular for a small sample size data set. Future studies may consider other alternative approaches for this modeling. Table 6 illustrates the result of the estimated model, pointing out three main effects and two interaction effects that are statistically significant. The model suggests that the respondents who are in their 40s, start to work between 7 and 9 a.m., and have two or more workers in their households are more likely to carpool. Combined with the finding that the respondents in their 40s are prone to use the HOT lanes more, this result implies that they are also more likely to use the HOT lanes in carpool mode.

As found in the classification tree, the interaction effects reveal that the participants who have a positive perception of the HOT lanes have a weaker tendency to carpool, which may statistically support that HOT lanes may negatively influence carpooling. The perception variable is found to interact with age (40s) and the number of workers (two-worker households) and their impacts seem to be substantial as suggested by the magnitudes of the estimated parameters (−2.182 and −1.176). The resultant Nagelkerke R^2 value of 0.165 suggests that the model lacks the ability to strongly predict the carpool choices. Future studies are encouraged to incorporate more factors including travelers' perceptions and attitudes into the model for better understanding carpool behavior.

6.5. The perception model

It was suspected that the perception of the HOT lanes might have associations with other factors. To examine this, ordered probit models were developed, considering that the perception was measured by a 5-point Likert scale from one (definitely not improved) to five (definitely improved). The model was developed based on the data set of the HOT lane choice model and the car ownership variable was excluded due to its strong correlation with the number of household workers. Table 7 presents the resultant models, illustrating the six factors that are statistically significant at a level of 0.10: gender, number of household workers, income, former HOV user, work start time, and commute distance. Interestingly, the former HOV users appear to negatively perceive the HOT lanes, implying the HOT implementation might not be preferred by them, and thus may influence the breakup of carpools. The HOT implementation is also negatively perceived by commuters who usually start their work between 7 and 9 a.m., which may be ascribed to decreased travel speeds even in the HOT lanes during morning peak hours. Further studies are encouraged to explore these phenomena in more detail for better interpretations.

Despite the appearance of the significant variables, the overall explanatory power of the perception model seems unsatisfactory as suggested by the low value of Nagelkerke R^2 (0.159), suggesting the lack of capability of the model to predict HOT lane perception using the variables. This situation may justify the inclusion of the perception variable in the choice models together with other variables. More research in this area is definitely warranted.

7. Conclusions

The understanding of commuters' responses to HOT installations is important in that it can help transportation agencies identify operational strategies designed to maximize the sustainable use of HOT facilities. This study explores Atlanta's HOT lane implementation and carpool choices over the I-85 HOT corridors using data collected through a questionnaire-based survey. The self-administered mail-out/mail-back survey asked respondents about their lane choices (HOT or regular GP lanes) and carpool choices, both before and after the HOT lane installation, as well as overall trip patterns and demographic information. This survey is meaningful in that it was designed as the first attempt to assess carpool behavior after the installation of a conversion of an HOV lane to a HOT lane. As expected, the retrieval rate of the survey was low (about 5%), and a significant number of the retrieved surveys were not usable for developing certain statistical models, due to missing values and multiple answers for the same questions. Although low sample size restricted this study from fully utilizing respondents' various behavioral responses before and after the HOT installation, the binary choice models via classification trees and logistic regressions produced interpretable results that help explain the commuters' lane and carpool choices.

The HOT lane choice models showed that the perception of the effectiveness of the HOT lanes, a unique variable rarely treated in HOT behavior studies, exerts the strongest impact on the choices. More specifically, commuters are more likely to choose HOT lanes when they perceive HOT lanes have improved their own commute conditions. This finding implies that HOT operators should maintain an adequate level of HOT lane performance for maximizing the utilization of the lanes. The models also suggested that HOT lane choices can be affected by commuters' socio-economic characteristics. Commuters in their 40s, commuters with higher income, and commuters with higher education levels are more likely to choose the HOT lanes. This suggests that commuters with a high value of time are more likely to use HOT lanes, as expected. The importance of the age variable was also illustrated in the SR91 Express Lanes study, but in the SR91 study the age group in the 50s showed a stronger tendency to use HOT lanes (Li, 2001). Concerning trip patterns, commuters making longer trips were found to more likely choose HOT lanes. This situation appears to be intuitively correct, in that these travelers may have a stronger intention to save their travel times. The impact of trip length was not found significant in Li's (2001) study. The models pointed out that former HOV lane users tended to choose HOT lanes, suggesting many former HOV lane users might opt to use HOT lanes even after the HOT conversion. However, it is not clear how they use HOT lanes: paying a toll or HOV3+. Future studies are encouraged to investigate these choices in detail for a better understanding of commuters' behavior.

Regarding carpool choices, the selected data set showed that most carpoolers after the HOT installation were composed of former carpoolers. Weak carpool formation was noted, even after the HOT conversion, which is in the same vein as the conclusion of Burris et al. (2014). Likewise, the developed models revealed that the former carpooler variable dominated the effect on the carpool choice. This aspect clearly indicates that before-and-after carpooling behavior should be considered together to firmly understand drivers' behavior in particular when the HOT lanes are converted from HOV lanes. Statistical models also showed that commuters' socioeconomic characteristics could affect the carpool choice. Commuters in their 40s, commuters that have two or more workers in their households, and commuters that start work between 7 and 9 a.m. are more likely to carpool. In addition, females have a higher tendency to carpool, which conforms to the findings of studies conducted in France (Delhomme & Gheorghiu, 2016) and in Dallas-Fort Worth and Houston in Texas (Li et al., 2007). However, the models also indicate that commuters who have a positive perception of the HOT lanes are less likely to carpool. In particular, the constructed classification tree revealed that perception was the most important factor when the former carpooler variable was excluded. Based upon the survey data, this HOT project did not enhance carpooling as originally expected by the project proponents, which was also confirmed by vehicle occupancy evaluation in the previous before-after study (Guensler et al., 2013). This may also mean that carpools could continue to break-up as the performance of HOT lanes continues to improve. From the perspective that higher vehicle occupancy can generally heighten the sustainability level of transportation systems, this finding appears frustrating in spite of a potential sample bias that the mail-out/mail-back survey can retain. Policy makers may need to rethink strategies designed to increase carpool formation and retention as they implement HOT projects throughout the region.

Complementing previous studies, this study has enhanced the understanding of HOT lane and carpool choices on HOT corridors in particular by revealing the strong association between perception and mode/lane choices. However, the findings obtained from the binary choice observations seem to still leave numerous unexplained behavioral responses of the commuters, which might have been overcome with larger samples and more complete survey responses. A sufficient sample may be able to provide researchers with more chances to examine their complex decision-making mechanisms. It is also uncertain what actually happened about travelers' departure times and routes after the HOT lane installation. This stresses the importance of a better before-and-after data collection for the next managed lane conversion. In addition, the limited number of factors considered can explain only a small portion of HOT lane or carpool mode decision-making processes. Indeed, the explanatory power of the lane choice model was at most 0.36 in terms of Nagelkerke R^2. Work place TDM options and toll pricing effects, which were not captured in the survey, may deserve to be considered for the additional factors. Future study efforts are encouraged to capture larger samples and explore additional variables for developing improved models.

Acknowledgments

The authors of this paper wish to thank the National Center for Transportation Systems Productivity and Management, a University Transportation Center supported by the U.S. Department of Transportation, for their support and assistance.

References

Agresti, A. (1996). *An introduction to categorical data analysis*. New York, NY: John Wiley & Sons, Inc.

American Highway Users Alliance. (2015). Unclogging America's arteries 2015. Retrieved from http://www.highways.org/wp-content/uploads/2015/11/unclogging-study2015-hi-res.pdf

Breiman, L., Friedman, J., Stone, C., & Olshen, R. (1984). *Classification and regression trees*. Boca Raton, FL: CRC Press.

Burris, M., Alemazkoor, N., Benz, R., & Wood, N. (2014). The impact of HOT lanes on carpools. *Research in Transportation Economics*, *44*, 43–51. doi:10.1016/j.retrec.2014.04.004

Castrillon, F., Roell, M., Khoeini, S., & Guensler, R. (2014). The I-85 HOT lane's impact on Atlanta's commuter bus and vanpool occupancy. *Transportation Research Record: Journal of the Transportation Research Board*, *2470*(1), 169–177. doi:10.3141/2470-18

Camdeviren, H. A., Yazici, A. C., Akkus, Z., Bugdayci, R., & Sungur, M. A. (2007). Comparison of logistic regression model and classification tree: An application to postpartum depression data. *Expert Systems with Applications*, *32*(4), 987–994. doi:10.1016/j.eswa.2006.02.022

Delhomme, P., & Gheorghiu, A. (2016). Comparing French carpoolers and non-carpoolers: Which factors contribute the most to carpooling? *Transportation Research Part D: Transport and Environment*, *42*, 1–15. doi:10.1016/j.trd.2015.10.014

Gim, T. T., & Ko, J. (2017). Maximum likelihood and Firth logistic regression of the pedestrian route choice. *International Regional Science Review*, *40*(6), 616–637. doi:10.1177/0160017615626214

Goel, R., & Burris, M. W. (2012). Hot lane polices and their implications. *Transportation*, *39*(6), 1019–1033. doi:10.1007/s11116-011-9382-5

Guensler, R., Elango, V., Guin, A., Hunter, M., Laval, J., Araque, S., … Zinner, S. (2013). Atlanta, Georgia, I-85 high-occupancy vehicle to high-occupancy toll conversion: Analysis of vehicle and person throughput. Georgia Department of Transportation, Report # FHWA-GA-13-10-03. Georgia Institute of Technology, Atlanta.

Guensler, R., Xu, A., Sheikh, A., Li, H., & Khoeini, S. (2015). *Atlanta I-85 HOT lane carpool survey: Initial findings*. SE UTC Conference. Retrieved from http://www.ncitec.msstate.edu/wp-content/uploads/TS-S5-P2.pdf

Hastie, T., Tibshirani, R., & Freidman, J. (2001). *The elements of statistical learning: Data mining, inference, and prediction*. New York, NY: Springer-Verlag.

Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine*, *21*, 2409–2419. doi:10.1002/sim.1047

Khoeini, S., & Guensler, R. (2014a). Socioeconomic assessment of the I-85 high-occupancy vehicle to high-occupancy toll conversion in Atlanta, Georgia. *Transportation Research Record: Journal of the Transportation Research Board*, *2450*(1), 52–61.

Khoeini, S., & Guensler, R. (2014b). Using vehicle value as a proxy for income: A case study on Atlanta's I-85 HOT lane. *Research in Transportation Economics*, *44*, 33–42. doi:10.1016/j.retrec.2014.04.003

Kim, M. (2009). Two-stage logistic regression model. *Expert Systems with Applications*, *36*(3), 6727–6734. doi:10.1016/j.eswa.2008.08.063

Kutner, M., Nachtsheim, C., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). New York, NY: McGraw-Hill, Inc.

Li, J. (2001). Explaining high-occupancy-toll lane use. *Transportation Research Part D: Transport and Environment*, *6*(1), 61–74. doi:10.1016/S1361-9209(00)00013-4

Li, J., Embry, P., Mattingly, S., Sadabadi, K. F., Rasmidatta, I., & Burris, M. W. (2007). Who chooses to carpool and why? Examination of Texas carpoolers. *Transportation Research Record: Journal of the Transportation Research Board*, *2021*(1), 110–117. doi:10.3141/2021-13

Pessaro, B., & Buddenbrock, P. (2015). Managed lane toll prices: Impact of transportation demand management activities and toll exemptions. *Transportation Research Record: Journal of the Transportation Research Board*, *2484*, 159–164. doi:10.3141/2484-17

Pessaro, B., Turnbull, K., & Zimmerman, C. (2013). Impacts to transit from variably priced toll lanes. *Transportation Research Record: Journal of the Transportation Research Board*, *2396*, 117–123. doi:10.3141/2396-13

Sheikh, A., Misra, A., & Guensler, R. (2015). High-occupancy toll lane decision making: Income effects on I-85 express lanes, Atlanta, Georgia. *Transportation Research Record: Journal of the Transportation Research Board*, *2531*(1), 45–53. No. doi:10.3141/2531-06