

Full analysis code can be found here: <https://github.com/trzh/examples-tirzahkhan>

In constructing an alternative to the Boston TDM23 model, five variables were added to the original model. These variables are:

- PRICE: whether price of gasoline affects travel
- PLACE: whether travel is a financial burden for the household
- reduce_burden_action: if travel is a financial burden, does the household use an alternative mode of transportation to travel
- med_cond_impact: whether someone in the household has a medical condition that impacts how they travel
- use_med_device: whether someone in the household uses a medical device

The variables PLACE and PRICE were taken as-is from the 2017 NHTS data, but the other three variables were constructed manually. It's also important to note that I compared the Sufficient and Insufficient vehicle categories to the reference category of Zero vehicles, because I was interested in looking at how drastically utility of vehicle availability would change if the state of having zero cars was used as a baseline.

In both models, lower income households are negatively associated with vehicle availability. However, in the alternative model, low income shows a stronger negative effect on vehicle availability, with a coefficient of -1.296 for the Insufficient category and -1.889 for the Sufficient category, compared to the Boston model. This suggests that, according to the alternative model, lower income households are even less likely to have access to sufficient vehicles, and more likely to be impacted by not having enough vehicles. The number of workers in the household variable is a much stronger predictor in the alternative model than in the Boston model. The coefficients for Insufficient and Sufficient vehicle availability are 3.18 and 2.72, respectively, indicating that households with more workers are much more likely to have sufficient vehicles compared to the zero vehicle reference group.

The variables med_cond_impact and use_med_device in the alternative model both show negative associations with vehicle availability, meaning that households with medical conditions or devices are less likely to have enough vehicles, which is something that the Boston model does not account for. The PLACE variable, which shows whether the household feels travel is a financial burden, shows a positive correlation with vehicle availability. This suggests that households that

One interesting variable was the price of gasoline: Each response category under the PRICE variable had a coefficient of over 10. This indicates that all households, no matter whether they

strongly disagree or strongly agree (or somewhere in between) with the idea that the price of gasoline affects their travel, have a higher likelihood of reporting either insufficient or sufficient vehicles compared to the reference group of having no vehicles. This makes sense, as households without vehicles are less likely to be impacted by fluctuating gas prices, as they're less likely to be relevant to the household's activities and travel behavior. However, the standard error was extremely large, meaning that the coefficients are unreliable, and that the model does not indicate that there is strong evidence for a relationship between being burdened by gas price and vehicle availability.

I used three different measures to determine which model is a better predictor of household vehicle availability. One measure was the Akaike Information Criterion: for the Boston model, the AIC score was 39937.11, while for the alternative model, the AIC score was 37825.44. The lower AIC score for the alternative indicates that that model is better, because it has a better fit to the data. The second measure I used was the McFadden R-squared, which is used as a comparative indicator of how well a model fits data. The Boston model's pseudo-R-squared is 0.27, and the alternative model's is 0.31. As the alternative's R-squared is higher, it fits the data better than the Boston model. The third measure used was a less objective measure, and more of an explanation of why certain variables were included over others: The alternative model simply answers more of the questions that I am interested in than the original model does, meaning that it's more personally relevant to my interests and gives me a better understanding of what factors influence household vehicle availability.

Despite the alternative model answering more of my questions about travel behavior and vehicle availability than the Boston region model, there are other questions I am interested in that neither model can answer. For example, the impact of more specific mental health indicators (i.e. diagnoses of depression, anxiety, etc) on vehicle availability, as well as household proximity to nature and person-level ratings of social connectedness. I'm also interested in whether someone in the household being a first- or second-generation immigrant affects vehicle availability. From my own personal experience living in a high-immigration community, I know that first-generation immigrant households, particularly in the first few years following their immigration, often have a lot of difficulty getting access to vehicles to be able to go to work or school. Given that Boston has a foreign-born population of 28.2%¹, it would be interesting to know how being an immigrant in the Boston region affects vehicle availability.

¹ "Immigrant Demographics | Boston.Gov," April 6, 2021. <https://www.boston.gov/departments/immigrant-advancement/immigrant-demographics>.

```
Call:
mlogit(formula = choice ~ 0 | WRKCOUNT + n_child + n_seniors +
      n_extra_drivers + three_drivers + non_work_driver + income +
      density | 0, data = veh_dfidx_train, relevel = "Zero", method = "nr")
```

Frequencies of alternatives:choice

```
Zero  Insuff.  Suff.
0.047038 0.069064 0.883898
```

nr method

9 iterations, 0h:0m:7s

$g'(-H)^{-1}g = 4.37E-05$

successive function values within tolerance limits

Coefficients :

	Estimate	Std. Error	z-value	Pr(> z)	
(Intercept):Insuff.	-4.482944	0.112126	-39.9814	< 2.2e-16	***
(Intercept):Suff.	-0.128626	0.087512	-1.4698	0.141613	
WRKCOUNT:Insuff.	3.568485	0.073608	48.4798	< 2.2e-16	***
WRKCOUNT:Suff.	3.120896	0.067026	46.5623	< 2.2e-16	***
n_child:Insuff.	0.324458	0.043831	7.4025	1.337e-13	***
n_child:Suff.	0.115785	0.040717	2.8436	0.004460	**
n_seniors:Insuff.	0.873691	0.052509	16.6389	< 2.2e-16	***
n_seniors:Suff.	0.546476	0.047558	11.4908	< 2.2e-16	***
n_extra_drivers:Insuff.	-1.048223	0.425564	-2.4631	0.013773	*
n_extra_drivers:Suff.	-1.231308	0.423496	-2.9075	0.003643	**
three_driversTRUE:Insuff.	1.154920	0.598076	1.9311	0.053476	.
three_driversTRUE:Suff.	0.398942	0.595002	0.6705	0.502547	
non_work_driver:Insuff.	5.389559	0.089709	60.0785	< 2.2e-16	***
non_work_driver:Suff.	4.100179	0.072355	56.6672	< 2.2e-16	***
incomelow:Insuff.	-1.307347	0.067646	-19.3264	< 2.2e-16	***
incomelow:Suff.	-1.956106	0.057752	-33.8708	< 2.2e-16	***
incomehigh:Insuff.	-0.362192	0.128222	-2.8247	0.004732	**
incomehigh:Suff.	-0.083561	0.118924	-0.7026	0.482276	
densityLow:Insuff.	0.200666	0.069994	2.8669	0.004145	**
densityLow:Suff.	0.522840	0.059280	8.8199	< 2.2e-16	***
densityMedium:Insuff.	-0.605220	0.089094	-6.7931	1.098e-11	***
densityMedium:Suff.	-1.483996	0.071231	-20.8335	< 2.2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -19947

McFadden R²: 0.27173

Likelihood ratio test : chisq = 14885 (p.value = < 2.22e-16)

```
Call:
mlogit(formula = choice ~ 0 | WRKCOUNT + n_child + n_seniors +
  n_extra_drivers + three_drivers + non_work_driver + income +
  PLACE + PRICE + use_med_device + reduce_burden_action + med_cond_impact +
  density | 0, data = veh_dfidx_train, reflevel = "Zero", method = "nr")
```

Frequencies of alternatives:choice

```
Zero Insuff. Suff.
0.047038 0.069064 0.883898
```

```
nr method
14 iterations, 0h:0m:23s
g'(-H)^-1g = 5.26E-07
gradient close to zero
```

Coefficients :

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept):Insuff.	-16.261324	3215.095741	-0.0051	0.9959645
(Intercept):Suff.	-14.761875	3214.332802	-0.0046	0.9963357
WRKCOUNT:Insuff. **	3.176038	0.076194	41.6837	< 2.2e-16 *
WRKCOUNT:Suff. **	2.721191	0.070059	38.8414	< 2.2e-16 *
n_child:Insuff. **	0.261776	0.046113	5.6768	1.372e-08 *
n_child:Suff.	0.048389	0.043267	1.1184	0.2634017
n_seniors:Insuff. **	0.890924	0.056647	15.7277	< 2.2e-16 *
n_seniors:Suff. **	0.556001	0.052041	10.6839	< 2.2e-16 *
n_extra_drivers:Insuff.	-0.826524	0.411673	-2.0077	0.0446731 *
n_extra_drivers:Suff.	-1.002201	0.410039	-2.4442	0.0145189 *
three_driversTRUE:Insuff.	0.833674	0.580835	1.4353	0.1512009
three_driversTRUE:Suff.	0.068430	0.577956	0.1184	0.9057506
non_work_driver:Insuff. **	5.005028	0.093621	53.4603	< 2.2e-16 *
non_work_driver:Suff. **	3.710387	0.077414	47.9292	< 2.2e-16 *
incomelow:Insuff. **	-1.296087	0.074372	-17.4271	< 2.2e-16 *
incomelow:Suff. **	-1.889069	0.064746	-29.1766	< 2.2e-16 *
incomehigh:Insuff.	-0.136301	0.135488	-1.0060	0.3144134
incomehigh:Suff.	0.153065	0.126554	1.2095	0.2264773
PLACE01:Insuff.	0.473469	0.198795	2.3817	0.0172333 *
PLACE01:Suff.	0.369471	0.156406	2.3623	0.0181643 *
PLACE02:Insuff. **	0.858558	0.188452	4.5559	5.217e-06 *
PLACE02:Suff. **	0.683690	0.148092	4.6166	3.900e-06 *
PLACE03:Insuff. **	1.119382	0.190965	5.8617	4.581e-09 *
PLACE03:Suff. **	0.990777	0.150639	6.5772	4.795e-11 *
PLACE04:Insuff. **	1.232716	0.200552	6.1466	7.916e-10 *
PLACE04:Suff. **	0.932711	0.159963	5.8308	5.516e-09 *
PLACE05:Insuff. **	1.624777	0.235508	6.8990	5.236e-12 *
PLACE05:Suff. **	1.373737	0.191171	7.1859	6.677e-13 *
PRICE-9:Insuff.	10.360308	3215.095753	0.0032	0.9974289
PRICE-9:Suff.	13.999284	3214.332803	0.0044	0.9965250

PRICE01:Insuff.	13.085911	3215.095737	0.0041	0.9967525
PRICE01:Suff.	16.291245	3214.332799	0.0051	0.9959561
PRICE02:Insuff.	12.792736	3215.095736	0.0040	0.9968253
PRICE02:Suff.	16.099539	3214.332799	0.0050	0.9960037
PRICE03:Insuff.	11.468677	3215.095736	0.0036	0.9971538
PRICE03:Suff.	14.763337	3214.332798	0.0046	0.9963354
PRICE04:Insuff.	11.900511	3215.095736	0.0037	0.9970467
PRICE04:Suff.	15.142892	3214.332798	0.0047	0.9962411
PRICE05:Insuff.	10.998861	3215.095737	0.0034	0.9972704
PRICE05:Suff.	14.122436	3214.332799	0.0044	0.9964944
use_med_device:Insuff.	-0.398674	0.106802	-3.7328	0.0001893 *
**				
use_med_device:Suff.	-0.437150	0.089975	-4.8586	1.182e-06 *
**				
reduce_burden_actionTRUE:Insuff.	-1.430178	0.069812	-20.4860	< 2.2e-16 *
**				
reduce_burden_actionTRUE:Suff.	-1.987903	0.059610	-33.3485	< 2.2e-16 *
**				
med_cond_impact:Insuff.	-0.318835	0.095316	-3.3450	0.0008227 *
**				
med_cond_impact:Suff.	-0.551724	0.082503	-6.6873	2.274e-11 *
**				
densityLow:Insuff.	0.043420	0.074505	0.5828	0.5600432
densityLow:Suff.	0.308885	0.064577	4.7832	1.725e-06 *
**				
densityMedium:Insuff.	-0.372833	0.094158	-3.9597	7.505e-05 *
**				
densityMedium:Suff.	-1.170995	0.077536	-15.1026	< 2.2e-16 *
**				

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -18863

McFadden R²: 0.3113

Likelihood ratio test : chisq = 17052 (p.value = < 2.22e-16)