This memo compares the vehicle availability model estimated by the Boston Region Metropolitan Planning Organization in its travel demand model (TDM23) with those that can be estimated using national data from the 2017 National Household Travel Survey (NHTS). I estimate two models, one adapting the code provided to attempt to replicate the model estimated in TDM23 as closely as possible and one that adds additional covariates in an attempt to improve the model’s fit.

The code for this analysis can be found in the Github repository at: <https://github.com/wpcs50/examples-wpowell/blob/main/P2/P2_powell.qmd>

**Data and Models**

There are several key factors that result in differences between the NHTS and TDM23 models. First, there are differences in the time period and geographic scope of the data. TDM23 is based on the 2011 Massachusetts Travel Survey while the NHTS is more recent (2017) and national in scope. Second, there are differences in the types of variables available in the data and how they are coded. For example, the ratio of transit accessibility to highway accessibility is available in TDM23, but is not available in the NHTS. Third, the coefficients estimated for TDM23 are calibrated to match observed data.

Vehicle availability is predicted using multinomial logistic regression with three possible outcomes. Households can be vehicle sufficient (at least as many vehicles as drivers) vehicle insufficient (at least one vehicle but fewer vehicles than drivers) or have zero vehicles.

To match TDM23 as closely as possible, the first model includes several household variables, including the number of workers, children, seniors, and drivers (beyond 2), the presence of a third or non-worker driver, income (low is less than 200 percent of the Federal poverty level, high is household income above $100,000 (in 2010 dollars). It also includes a categorical measure of density (less than 7,000 residents per square mile, 7,000-10,000 residents per square mile, or more than 10,000 residents per square mile). The household variables are able to match TDM23 fairly well, but TDM23 has different measures of density and includes the ratio of transit to highway accessibility (measured by the number of jobs within a 30-minute commute boundary).

The second model adds an indicator for whether a household is renting, which may be negatively correlated with the ease of parking, and whether the household lives within a metropolitan statistical area (MSA) with heavy rail and within an urban Census block group, as a proxy for transit access. More difficult parking (expected for renters) and better transit access would both be expected to raise the utility of having zero or insufficient vehicles.

**Results**

The coefficients from the two NHTS models, which correspond to a household’s utility from having a given level of vehicle availability, and their standard errors, are included in the table below. Coefficients from the analogous TDM23 variables are included where available for reference.



All variables are statistically significant for at least one outcome in the NHTS models. In most cases, the NHTS coefficients are directionally similar to those in the TDM23 models. For example, having more workers or children decreases the utility of having zero vehicles. And having more seniors, or extra drivers increases the utility associated with being vehicle insufficient in all models. This makes sense, as these households are likely to require some vehicles, but they may not have enough for all of the extra drivers. Similarly, with most commutes taking place by car, having more workers in a household decreases the utility of having no cars. Low-income households have higher utility of having zero or insufficient vehicles, potentially reflecting their inability to afford vehicles for all drivers.

In all three models, utility of insufficient or zero vehicles is increasing with density. This is expected given the likely positive relationship between density and public transit access and parking difficulty.

Households that rent and which are in an urban block group in an MSA with heavy rail both have higher utility from insufficient or zero vehicles. This is consistent with the intuition outlined earlier. Adding these variables has a limited impact on the other variables in the model, with the primary effect to reduce the magnitude of the estimated effect of density. Given the correlation between renting and density, this is not surprising. Public transit access is another important mechanism by which density would lead to reduced vehicle availability.

In general, the estimated NHTS models do little to improve on the no information rate, with accuracy rates of 0.903 and 0.904, respectively, against the no information rate of 0.885 (which results from the high base rate of sufficient vehicle availability). They particularly struggle with correctly identifying households with insufficient vehicles (and, to a lesser extent, zero vehicles). Adding rental status and heavy rail access marginally improves model fit (as measured by log likelihood) and sensitivity for both of these outcomes.