Analysing Household Vehicle Holdings and Usage in California using a Two-Stage Budgeting-Based Multiple Discrete-Continuous Model

Shobhit Saxena
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA
Email: shobhit.saxena@austin.utexas.edu

Chandra R. Bhat (corresponding author)
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Tel: 1-512-471-4535; Email: bhat@mail.utexas.edu

Abdul Rawoof Pinjari
Department of Civil Engineering
Centre for Infrastructure, Sustainable Transportation, and Urban Planning (C*i*STUP)
Indian Institute of Science (IISc), Bangalore, India, 560012
Email: abdul@iisc.ac.in

ABSTRACT
We analyse households' vehicle ownership and usage decisions in California using the 2019 California Vehicle Survey data. Different from earlier studies, we consider a vehicle type classification based on body type and fuel type, comprising gasoline vehicles, plugin hybrids, battery electric vehicles, and fuel cell electric vehicles. Further, we employ a two-stage multiple discrete-continuous choice framework that endogenously models the total annual vehicle mileage in the first stage and the allocation of the annual mileage to different vehicle types in a utility theoretic manner. In doing so, we extend the prior formulation of the two-stage model to relax the unit scale assumption on model stochasticity. The empirical analysis provides insights into the determinants of households' proclivity for owning and using traditional (gasoline-based) and alterative fuelled vehicles. Policy simulations with the empirical model indicate that regional employment land-use densification and the provision of non-personal commute options can help reduce vehicle ownership levels and their usage, thereby potentially reducing total on-road emissions.

Keywords: Vehicle ownership, vehicle usage, two-stage budgeting model, utility theory, alternative fuelled vehicles.

INTRODUCTION

Background and Motivation

Passenger mobility in the United States heavily relies on private motorized vehicle ownership (in the rest of this paper, we will use the label "vehicle" to refer to "motorized vehicle"). The dependence on personal vehicles in the U.S. may be attributed to several factors, including (a) inadequate public transport infrastructure *(1)*, (b) the ease of accessibility to jobs and other opportunities that a personal vehicle provides in the non-compact geographic footprint of urban areas *(2)*, (c) relatively low vehicle purchase costs compared to the rest of the world *(3)*, and (d) an affective sense of pride and social status attached with vehicle ownership *(4)*. But, at the same time, high vehicle ownership and use are also associated with high traffic congestion, high energy consumption, and high tailpipe emissions. In fact, household vehicle usage contributes to more than half the greenhouse (GHG) emissions from the transportation sector in the U.S. *(5),* thanks to the continued high share of conventional gasoline-powered vehicles *(6)*. In this regard, policies aimed at mitigating GHG emissions and pollutants from household vehicles focus on accelerating the penetration (and use) of cleaner, alternative fuel-powered vehicles, such as plug-in hybrid electric vehicles (PHEVs), battery electric vehicles (BEVs), and fuel cell electric vehicles (FCEVs) (for instance, incentives under the California Clean Vehicle Rebate project (2016) and California's Advanced Clean Cars II (2023) rule to phase out gasoline-powered vehicles by 2035). To aid such policy formulation, vehicle fleet ownership and usage models are often employed to investigate the influence of various factors on households' vehicle fleet mix and miles travelled *(7)*. Beyond policy considerations, such models are embedded within regional travel demand models to predict the evolution of regional vehicle fleets over time and their influence on regional travel demand *(8)*.

From a methodological standpoint, studies have investigated households' vehicle fleet holding and use decisions using discrete-continuous frameworks *(9, 10)*. However, most such studies focus on decisions regarding the most recently purchased vehicle, ignoring the fleet mix held by many households in the U.S. that own multiple vehicles. Bhat and Sen *(11)*, for the first time, used the multiple discrete-continuous (MDC) choice framework to model the simultaneous holding of multiple vehicles by the household and the mileage accrued on each vehicle type owned by the household. The MDC choice framework has garnered substantial attention since then to model household vehicle holdings and usage *(2, 12, 13)*. However, a limitation of most former applications of MDC choice models is that they require information on the total mileage on vehicles as an exogenous input. But this is antithetical to the notion of predicting mileage on individual vehicles as part of the vehicle fleet model, constraining total mileage to be fixed across all vehicles owned. Such models do not accommodate the effect of changes in individual vehicle attributes, such as fuel economy, on total vehicle mileage. To overcome this shortcoming, some other studies include a non-zero mileage for non-motorized modes of travel as an essential outside good (an alternative that is always allocated a non-zero mileage) in the model, which helps accommodate an increase or decrease in the total mileage on motorized vehicles. In another study, Augustin et al. *(14)* developed a stochastic frontier approach to estimate an upper bound on the total mileage (i.e., a maximum possible mileage) that a household can accrue. Such an upper bound is utilized as the total mileage budget in an MDC choice model that is allocated to different vehicle choice alternatives and an outside good that represents the unconsumed portion of the budget. However, the stochastic frontier budget equation in the formulation is not necessarily consistent with the theory of utility maximization. Further, in most earlier MDC applications, the possibility that some households do not own a vehicle was not considered.

Bhat *(15)* recently proposed a two-stage MDC choice model that allowed endogenous estimation of the budget with the allocation to the elementary alternatives, all within a utility maximization framework. This two-step formulation is based on the two-stage budget allocation framework *(16, 17)*. Further, his formulation accommodates situations where none of the elementary alternatives are chosen for consumption (situations corresponding to zero budget) by employing statistical censoring through a Tobit-type *(18)* first stage budget equation. Such situations are common when analysing households' vehicle holdings where it is likely that a household may not own any motorized vehicle.

The Current Paper in Context

In this paper, we present an application of Bhat's *(15)* two-stage MDC model to analyse households' vehicle holding and usage decisions in California using the 2019 California Vehicle Survey data *(19)*. Different from the Bhat's *(15)* application (also a vehicle fleet holding and usage application), we analyse households' vehicle fleet mix comprising a more diverse choice set, with vehicle ownership classification based on body-type and fuel type, including traditional gasoline-powered vehicles and electric and hybrid vehicles. Such an analysis sheds light on the factors that influence the ownership and usage decisions of traditional gasoline-powered vehicles as well as new, alternative-fuelled vehicles such as PHEVs, BEVs, and FCEVs. Such insights from our empirical analysis can potentially help in formulating programs that go beyond monetary incentives for reducing vehicle ownership and accelerating the adoption and usage of alternative-fuelled vehicles. Finally, from a methodological standpoint, we also show how the two-stage MDC model is easily extended to relax the unit scale assumption in its prior formulation.

The rest of the paper is structured as follows. Section 2 discusses the empirical data used in this analysis. Section 3 presents the structure of the two-stage MDC model and discusses the nuances in relaxing the unit scale assumption. Section 4 presents empirical results. Section 5 concludes the paper with a summary and some future research directions.

DATA

The data for the analysis is drawn from the 2019 California Vehicle Survey undertaken by the California Energy Commission *(19)* to understand shifts in both residential and commercial vehicle mix. For this study, we focus on the residential vehicle fleet mix, which comprises vehicle holdings of households spread across the 58 counties of California. Of the surveyed households, 759 households owned PHEVs, BEVs, or FCEVs, based on an intentional oversampling of such households. Thus, different from previous studies (that typically define vehicle type choice alternatives as a combination of vintage and body type), we consider the motorized vehicles owned by the households as a combination of body type and fuel type, leading to the following 7 alternative vehicle categories: a) Gasoline car (includes subcompact, compact, midsize, large, or sports cars that use conventional gasoline/diesel/flex fuel), b) Hybrid car (includes the above car categories, except with PHEV or hybrid-gas fuel options), c) Electric car (includes the above car categories with fully electric engines, or hydrogen cell (FCEV) or a plug-in with hydrogen cell (PFCEV) options), d) Gasoline SUV (includes sub-compact, compact, midsize or large cross-over sports utility vehicles (SUVs) that use gasoline/diesel/flex fuel options), e) 'Hybrid+' SUV (includes the SUV subcategories with PHEV/hybrid-gas fuel options, along with a miniscule proportion (<0.5%) of fully electric SUVs), f) Van (includes small and large vans with gas/diesel as the fuel options), and g) Pickup truck (includes small and large pickup trucks with gas/diesel as fuel options). In addition to the vehicle ownership and annual mileage on each vehicle in the household, a host of person, household and regional attributes were also collected in the survey.

After cleaning the data for missing/incorrect records, the final sample for analysis comprised vehicle ownership and annual mileage information of 2862 households. This sample of 2862 households was further partitioned into an estimation sample comprising randomly selected 2000 households and a holdout validation sample of the remaining 862 households. The details of the vehicle mix and the descriptive statistics of the sociodemographic attributes of the estimation sample are provided in Table 1.

The percentage of households with zero-car ownership is a mere 3.85%, which is not surprising since the U.S. has a high personal vehicle ownership, and California has the highest total number of personal vehicle registrations among all states in the U.S. *(20)*. As expected, gasoline car continues to be the largest owned vehicle category, with 56.25% of the households owning a gasoline car. This is followed by gasoline-powered SUVs, with 40.5% of the households owning a gas SUV. At the same time, the proportion of households owning a hybrid/electric car is small but not negligible, with close to 15% of the households owning an electric or hybrid car. This figure goes down to just around 4% in the case of hybrid SUVs. In sum, across all households, nearly 20% of the vehicle fleet is either hybrid or electric. Notably, this number is higher than the average penetration of electric vehicles (EVs) and hybrid cars in California (around 7% as per *(20)*. This could be because the survey data oversampled households with at least one PEV or FCEV (leading to nearly 13% of the vehicle fleet being either electric or hybrid). Pickup trucks are owned by nearly 13% of the households in our sample, which is close to the proportion of pickup trucks in California (around 12%) *(21)*. Only a small proportion of households in our sample (around 7%) own a passenger van. Interestingly, the average mileage on hybrid and electric cars and SUVs is higher than that on gasoline-powered vehicles. This is presumably because hybrid and electric vehicles are more fuel efficient. They are also likely to have been purchased more recently and therefore driven more. Overall, across all vehicle types, cars are the most driven, which is representative of the trends in vehicle usage in California.

In addition to the vehicle ownership and usage information, the fuel efficiency of the vehicles was also available in the data. The fuel efficiency was reported as miles per equivalent gallon (MPeG), which is the same as miles travelled per gallon of gasoline for gasoline fuelled vehicles. For other alternative fuelled vehicles, MPeG was determined using the gallon equivalent of the alternative fuel, which is the amount of alternative fuel that gives the same energy as one gallon of gasoline (see (*22)* for details). The fuel efficiency is the highest for the electric car (with MPeG of 81.7 miles per equivalent gallon of gasoline) and lowest for pickup trucks (with MPeG of 18.5 miles per gallon of gasoline). These values are consistent with the market standards, with electric cars having an efficiency of 2.5-3 miles per kWh (or 78-100 miles per equivalent gallon).

In the context of sociodemographic attributes, the sample reasonably reflects the makeup of the population in California. 12.6% of the households in the sample have an income of more than $200,000, which is quite close to that in the California population, where 15% of the households belong to this income category *(23)*. In the context of employment characteristics, nearly 70% of the households have at least one full-time or part-time worker, which is nearly 78% in the population *(24)*. The average household size in our sample is 2.2, which is lower than the population's average household size of 2.9. The average number of driving license holders per household in our sample is around 1.9 which is close to the value of 2.1 in the California population *(24)*. Nearly 61% of the housing in California are detached single houses (California housing statistics), which is well reflected in our sample, with 61.4% households living in detached houses.

TABLE 1 Descriptive Statistics of the Estimation Sample

Sample size: 2000 households	**Vehicle types**						
	Gasoline car	**Hybrid car**	**Electric car**	**Gasoline SUV**	**"Hybrid+" SUV**	**Van**	**Pickup truck**
Percentage of households with zero vehicle	3.85%						
Percentage of households with the vehicle category	56.25	15.60	14.75	40.55	4.30	7.15	13.85
Average annual mileage (miles)	10744	11444	11219	10271	10197	9972	9041
Average total miles driven in a year (miles)	16052						
Average fuel efficiency (in miles per equivalent gallons) of the vehicle category	25.8	49.30	81.71	22.83	46.40	21.77	18.51

Exogenous variables	**Sample shares/Average values**	**Exogenous variables**	**Sample shares/Average values**
Household income		Average number of driving licenses in the household	1.9
Income less than $25K	8.8%	Single individual households	32.8%
Income between $25K to $50K	16.6%	***Household type***	
Income between $50K to $100K	31.9%	Detached	61.4%
Income between $100K to $200K	30.1%	Apartments	26.1%
Income more than $200K	12.6%	Other	12.5%
Presence of workers		***Parking facilities***	
at least 1 full time/part time worker	62.9%	Absence of a covered parking in the vicinity of the household	18.8%
		Absence of a parking facility with electrical charging (in office/frequently visited places)	18.3%
Household composition		***Regional attributes***	
Average household size	2.2	Average population density (person/ sq. miles)	2020.9
Average number of children (with age less than 15 years)	0.3	Proportion of households in areas with population density > 2000 per sq. mile	41.2%
Household with children	15%	Average employment density (person/sq. mile)	1048.4
Average number of adults	1.9	Proportion of households in areas with employment density > 500 per sq. mile	57.3%

In the context of regional attributes, our sample households live in locations with an average population and employment densities of 2020.9 and 1048.4 persons per square mile, respectively. In contrast, the corresponding values for the state of California (averaged across all 58 counties) are 702.9 persons per sq. mile and 397.12 persons per sq. mile *(24)*, indicating that our sample over-represents the urban population.

MODEL STRUCTURE

This section presents the model formulation for each of the two stages and the linking between the two stages.

Fractional Split MDC Model for the Second Stage Allocation to Elemental Alternatives in the Product Group

Consider the following constrained utility maximization problem for the second stage of the two-stage formulation for the allocation among the inside goods for a product group:

$$U(\tilde{\mathbf{f}}) = \sum_{k=1}^{K} \gamma_k \psi_k \ln\left\{\left(\frac{\tilde{f}_k}{\gamma_k}+1\right)\right\} \qquad (1)$$

$$\text{subject to } \sum_{k=1}^{K} \tilde{f}_k = 1 \text{ and } \tilde{f}_k \geq 0$$

Here, $U(\tilde{\mathbf{f}})$ is an increasing, continuously differentiable, and additively separable utility function, with $\tilde{f}_k$ being the fraction of the product group level budget allocated to an inside alternative *k*. In this utility function, γ_k are the satiation parameters and ψ_k are the baseline preference parameters. γ_k are specified as $\exp(\boldsymbol{\omega}'\mathbf{w}_k)$ to accommodate heterogeneity in satiation due to alternative attributes and household characteristics in the vector $\mathbf{w}_k$. The baseline preference is expressed as:

$$\psi_k = \exp\left(\boldsymbol{\beta}'\mathbf{z}_k - \frac{1}{\sigma}\ln p_k + \varepsilon_k\right) \qquad (2)$$

where $\mathbf{z}_k$ is a set of attributes that characterize alternative *k* and the decision maker (including a constant), p_k is the unit price for good *k*, the inverse of $\sigma (\sigma > 0)$ is the coefficient on $\ln p_k$ (called the price coefficient from now on), and ε_k is an error term to recognize the unobserved factors influencing the baseline preference. $\boldsymbol{\beta}$ is the vector of parameters corresponding to the set of attributes $\mathbf{z}_k$. The identification considerations allow estimating only *K*–1 alternative specific coefficients of the variables in $\mathbf{z}_k$ that do not vary across alternatives. Note that in the above specification, Bhat *(15)* assumed the ε_k terms to be of type I extreme value distribution based on minimum (i.e., reverse Gumbel distribution). He also assumed the distributions to be of unit scale. However, we relax that assumption and show that one can freely estimate the scale parameter. To do so, we denote the scale parameter by κ. Note that this scale of the error terms is different from that of $1/\sigma$, the coefficient on the logarithm of p_k. That is, the σ parameter is not the scale of the error terms. In fact, it is possible to estimate the price coefficient separately from the scale parameter if price variation exists across individuals for at least two alternatives.

One can set up the Karush-Kuhn-Tucker (KKT) conditions of optimality for the utility maximization problem in Equation (1) and derive the following likelihood function for the optimal fractional allocations ($\tilde{f}_k^{op}\ \forall k$) among the elemental alternatives in the product group (see *15* and *25* for the relevant derivations):

$$L\left(\tilde{f}_1^{op},\ldots,\tilde{f}_M^{op},0,0,\ldots,0\right)$$

$$=\frac{|J|(M-1)!}{\kappa^{(M-1)}}\left[\frac{\exp\left(\sum_{i=1}^{M}\overline{V}_k\right)}{\left(\sum_{k=1}^{M}\exp(\overline{V}_k)\right)^M}+\sum_{D\subset\{M+1,M+2,\ldots,K\},|D|\geq 1}(-1)^{|D|}\frac{\exp\left(\sum_{i=1}^{M}\overline{V}_k\right)}{\left(\sum_{k=1}^{M}\exp(\overline{V}_k)+\sum_{k\in D}\exp(\overline{V}_{k0})\right)^M}\right] \tag{3}$$

where $|J| = \left(\prod_{i=1}^{M} c_i\right)\left(\sum_{i=1}^{M}\frac{1}{c_i}\right)$, where $c_i=\left(\frac{1}{\tilde{f}_i^{op}+\gamma_i}\right)$, with $\overline{V}_k=\frac{V_k}{\kappa}$, and $\overline{V}_{k0}=\frac{V_{k0}}{\kappa}$,

$V_k=-\boldsymbol{\beta}'\mathbf{z}_k+\ln\left(\frac{\tilde{f}_k^{op}}{\gamma_k}+1\right)$ (k = 1, 2, 3,…, K), and $V_{k0}=-\boldsymbol{\beta}'\mathbf{z}_k$ (k = 1, 2, 3,…, K).

D in the above equation represents a specific combination of the vehicle segments, and $|D|$ is the cardinality of the specific combination D. The above expression is the same as that of an MDCEV model with reverse Gumbel error terms and fractional allocations *(25)*.

Linking Function to Link the Second-Stage Model (for Fractional Allocations) with the First-Stage Model (for Total Budget to the Product Group)

Define the following quality-adjusted, scalar composite price index for each elementary inside good alternative:

$$-\ln\tilde{\pi}_k=\frac{1}{1/\sigma\kappa}\left[\ln\left(\overline{\psi}_k\gamma_k+\gamma_k\sum_{\substack{j=1\\j\neq k}}^{M}\gamma_j\left(\overline{\psi}_k-\overline{\psi}_j\right)\right)\right],\ \overline{\psi}_k=\left(\psi_k\right)^{1/\kappa}$$

$$\text{or }\tilde{\pi}_k^{-1/\sigma\kappa}=\overline{\psi}_k\gamma_k+\gamma_k\sum_{\substack{j=1\\j\neq k}}^{M}\gamma_j\left(\overline{\psi}_k-\overline{\psi}_j\right). \tag{4}$$

In the above expression, the presence of $\overline{\psi}_k$ allows for quality and scale adjustment. Also, the negative sign in "$-\ln\tilde{\pi}_k$" ensures that as the good becomes more desirable (through an increase in baseline preference or a decrease in satiation), its quality-adjusted price index reduces. Notably, an increase in the unit price p_k of an alternative increases its quality-adjusted price, *ceteris paribus*, since an increase in p_k leads to a reduction in baseline preference. $1/\sigma$ represents the marginal utility of income.

Note that the above price-index, in its logarithmic form, is a quality- and scale-adjusted price-index. Specifically, the normalization by the scale parameter recognizes the need to scale the "signal" with the "noise". In Equation (4), we write the normalized baseline preference parameter as:

$$\overline{\psi}_k=\left(\psi_k\right)^{1/\kappa}=\mu_k^{1/\kappa}\exp(\overline{\varepsilon}_k),\ \mu_k=\exp(\boldsymbol{\beta}'\mathbf{z}_k)\text{ and }\overline{\varepsilon}_k=\frac{\varepsilon_k}{\kappa}. \tag{5}$$

with $\overline{\varepsilon}_k$ now standardized to unit scale.

Using the price-indices of all elemental alternatives within a product group, the fractional allocation to the alternatives in that product group can be derived, conditional on positive budget allocation to the product group, as below (see *15* for details):

$$\tilde{f}_k^{op} = \left(\frac{\tilde{\pi}_k^{-1/\delta}}{\sum_{m=1}^{M} \tilde{\pi}_m^{-1/\delta}} \right) = \frac{\bar{\psi}_k \gamma_k + \gamma_k \sum_{\substack{j=1 \\ j \neq k}}^{M} \gamma_j \left(\bar{\psi}_k - \bar{\psi}_j \right)}{\sum_{j=1}^{M} \bar{\psi}_j \gamma_j}, \; k = 1, 2, ..., M \text{ (for consumed goods)} \tag{6}$$

$$\tilde{f}_k^{op} = 0, \; k = M+1, M+2, ..., K \text{ (for non-consumed goods)}$$

In the above expression, for notation convenience, $\sigma\kappa$ is denoted by δ.

Next, the price indices ($\tilde{\pi}_k$) of all inside goods within a product group can be aggregated using the following constant elasticity of substitution (CES) form (which is homogenous of degree 1) to define a group-level price index *(15)*:

$$b(\tilde{\boldsymbol{\pi}}) = \left[\sum_k \left(\left(\tilde{\pi}_k \right)^{-1/\delta} \right) \right]^{-\delta} = \left[\sum_{j=1}^{M} \bar{\psi}_j \gamma_j \right]^{-\delta} = \left[\sum_{j=1}^{M} \left(\mu_k^{1/\kappa} \exp(\bar{\varepsilon}_k) \right) \gamma_j \right]^{-\delta} \tag{7}$$

The product group-level price index defined above is used to link the second-stage fractional allocation to the first-stage budget allocation (i.e., to enable the second-stage allocation to influence the first-stage). However, to keep the fractional allocation to each good within the product group exogenous to the group-level budget (a requirement of the Gorman polar form), Bhat *(15)* used a different set of error terms (i.e., τ_k instead of $\bar{\varepsilon}_k$) for the baseline preference parameters in the group-level price index. This helps avoid any correlation or influence of the group-level budget on the second-stage fractional allocation.

The linking function, in logarithmic form, can then be rewritten as:

$$\ln b(\tilde{\boldsymbol{\pi}}) = -\delta \ln \left[\sum_k \left(\mu_k^{1/\kappa} \exp\left(\tau_k \right) \right) \gamma_k \right] = -\delta \ln \left[\sum_k \bar{a}_k \exp\left(\tau_k \right) \right], \text{ where } \bar{a}_k = \mu_k^{1/\kappa} \gamma_k. \tag{8}$$

with τ_k following a standardized reverse Gumbel distribution, $\mu_k^{1/\kappa}$ representing the scale-normalized deterministic part of the baseline preference parameter, and the scale of the error terms (κ) also appearing in the term δ.

Structure of the First-Stage, Product Group-Level Budget Equation

The group-level budget amount needs to be non-negative, with a positive probability of zero budget to the group. Consider the following Tobit-like equation:

$$\begin{aligned} y^* &= \boldsymbol{\theta}'\mathbf{s} - \lambda \ln b(\tilde{\boldsymbol{\pi}}) - \lambda\zeta = \boldsymbol{\theta}'\mathbf{s} + \lambda\delta \left[\ln \sum_k \bar{a}_k \exp(\tau_k) \right] - \lambda\zeta \\ &= \boldsymbol{\theta}'\mathbf{s} - \lambda\eta, \text{ with } \eta = \left(\zeta - \delta \left[\ln \sum_k \bar{a}_k \exp(\tau_k) \right] \right) \end{aligned} \tag{9}$$

$$y = \begin{cases} 0 & \text{if } y^* \leq 0 \\ y^* & \text{if } y^* > 0 \end{cases}$$

In the above expression, $\mathbf{s}$ is an exogenous variable vector, $\boldsymbol{\theta}$ is a corresponding coefficient vector, λ is a scalar link parameter (λ>0), and ζ is a random variable capturing the effects of unobserved variables in the budget equation. The linking parameter appendage (i.e., λ) to the error term ζ in the first line of Equation (9) is innocuous and is only for presentation ease in the characterization of the error term η. As the price of any inside good k (p_k) decreases, or as the non-cost systematic (log) baseline utility element for any inside good k ($\boldsymbol{\beta}'\mathbf{z}_k$) increases, the value of η falls, and the value of the budget allocated to group g, (i.e., y) increases.

As in *(15)*, we assume ζ to be reverse Gumbel distributed, however with scale δ, to derive the distribution of η. This new distribution, apparently not seen in the earlier statistical literature, was characterized by Bhat *(15)* as the minLogistic distribution, who derived its properties. The resulting cumulative distribution and probability density functions are closed-form expressions, pawing way to a simple likelihood expression, which is extendable to the case where the scale of the error terms of the MDC model are explicitly estimated.

Model Estimation

From the first and second stages of the two-stage model, collect the parameters to be estimated in a vector $\mathbf{r} = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \kappa, \sigma, \boldsymbol{\theta}, \lambda)'$. The likelihood of zero allocation to the product group (i.e., none of the inside goods from the product group are chosen for consumption) in the first stage is the expression below based on the properties of the minLogistic distribution:

$$L(y=0) = F_{y^*}(0) = \frac{1}{\prod_{k=1}^{K}\left(1+\bar{a}_k e^{(\boldsymbol{\theta}'\mathbf{s})/\vartheta}\right)} \text{, with } \vartheta = \lambda\delta = \lambda\kappa\sigma. \tag{10}$$

The likelihood for allocation of a positive budget amount b to the product group, along with that of the fractional allocations to the inside alternatives in product group is given by:

$$L\left\{\left(y=b\right),\left((\tilde{f}_2^*,...,\tilde{f}_M^*,0,0,...,0)\right)\right\} = f_{y^*}(b) \times L\left(\tilde{f}_2^*,...,\tilde{f}_M^*,0,0,...,0\right) \tag{11}$$

where, $f_{y^*}(\cdot)$ is the density of y^*, given by:

$$f_{y^*}(t) = \frac{1}{\vartheta}\left(e^{[(\boldsymbol{\theta}'\mathbf{s}-t)/\vartheta]}\right) \times F_{y^*}(t) \times \sum_k \frac{\bar{a}_k}{\left(1+\bar{a}_k e^{[(\boldsymbol{\theta}'\mathbf{s}-t)/\vartheta]}\right)},\ F_{y^*}(t) = \frac{1}{\prod_{k=1}^{K}\left(1+\bar{a}_k e^{[(\boldsymbol{\theta}'\mathbf{s}-t)/\vartheta]}\right)} \tag{12}$$

and $L\left(\tilde{f}_2^*,...,\tilde{f}_M^*,0,0,...,0\right)$ is the likelihood for the observed fractional allocation in the second stage, given by Equation (3).

Once the parameters are estimated, the total budget and budget allocations to individual inside goods may be obtained using the procedures laid down in Bhat *(15)*.

EMPIRICAL ANALYSIS

Estimation Results

The two-stage MDC model is useful for analysing expenditures when data on unit prices are available. In our empirical case, information on fuel efficiency (in MPeG) for each vehicle type (that was owned by a household) was available as a proxy for price. In our estimation trials, we first started by using expenditure across each vehicle category and the total expenditure as the dependent variable. However, the estimated log-price coefficient turned out to be insignificant, possibly because of using aggregate price measures across vehicle categories that were not owned by the households (since information on fuel efficiency was available only for vehicles that were owned by households). As a result, we resorted to using mileage accrued on each vehicle type and the total annual mileage as the dependent variable in our analysis. The empirical specification of the model was systematically built by considering each of the relevant attributes in the budget equation (i.e., the first stage) and the fractional split model (i.e., the second stage) and retaining those variables that were statistically significant with t-statistic of more than 1. The use of a low confidence interval is to guide future investigations of these variables in influencing vehicle fleet ownership and usage analysis with large national-level datasets.

MDCEV Fractional Split Model Results
The parameter estimates are provided in Table 2. The baseline preference constants (presented toward the end of the first row panel of Table 2) do not offer any substantive interpretation once other attributes are included in the specification. Household demographics significantly impact the vehicle type choice. As expected, high-income households (income > $100K) are more likely to own electric and hybrid variants of cars and SUVs. These results reaffirm findings from the literature that income is one of the most consequential determinants of (or a barrier to) EV ownership. This finding warrants the need for policies to address EV adoption equity concerns *(26)*. Interestingly, low-income households (income < $50,000) have a lower preference for SUVs, probably owing to their higher cost. On the other hand, households with income more than $50,000 prefer pickup trucks (see *27* for a similar finding). In the context of household composition, individuals living alone (i.e., single-person households) are more likely to own gasoline and electric cars, and less likely to own vans. This may be because owning smaller cars satisfies their travel needs without any hassles of owning and maintaining large-sized vehicles. Similar findings can be observed for women living alone, with a higher preference for gasoline and hybrid cars, and lower preference for pickup trucks. The presence of children (<15 years age) in the household results in a higher preference for vans and SUVs, presumably because such spacious vehicles are convenient for carpooling arrangements. Interestingly, the presence of an employed adult increases the preference for owning an electric vehicle (see *28* for a similar finding for German households). Surprisingly, the presence of employed individuals in the household also increases the preference for vans. This could be because of a seemingly increasing attraction towards nomadic lifestyle, where individuals undertake weekend getaways using larger sized vehicles (such as camper vans) *(29)*. Additional evidence is needed to corroborate this hypothesis.

Households living in detached housing units have a higher preference for larger vehicles and a lower preference for gasoline cars. Interestingly, the preference for owning larger vehicles (vans or pickup trucks) reduces if the household does not have access to covered parking. As expected, the absence of a charging facility in home-parking (or workplace parking or other frequently visited places) reduces the preference for owning an electric car. Finally, the regional attributes (population density and employment density) of household's location also influence car ownership and usage behaviour. Further, households located in regions of higher employment density (> 500 persons per sq. mile) are less likely to own pickup trucks and SUVs. This result is not surprising since households in regions of lower employment density are likely to be self-employed in activities such as farming and construction, which require hauling material and manoeuvring in rugged terrains.

The parameters influencing the satiation functions (γ_k) are reported in the lower pane of Table 2. The parameter estimates reveal that as household income decreases, more mileage is put on gasoline-powered vehicles. Households with employed individuals tend to put less mileage on SUVs and vans, conditional on owning these vehicle types. Interestingly, recall from the baseline

TABLE 2 Estimation Summary: Parameter Estimates of the Fractional Split Model

	Parameter estimates (*t-stats*)						
Explanatory variables	**Gasoline car**	**Hybrid car**	**Electric car**	**Gasoline SUV**	**"Hybrid+" SUV**	**Van**	**Pickup truck**
Baseline preference function							
Household income characteristics (Base: Income more than $200K)							
Income less than $25K	--	-1.26 *(-3.10)*	-1.77 *(-3.09)*	-1.02 *(-3.02)**	-1.02 *(-3.02)**	--	--
Income less than $25K to $50K	--	-0.46 *(-2.73)*	-0.88 *(-3.13)*	-0.36 *(-2.50)*	-0.50 *(-2.28)*	--	--
Income less than $50K to $100K	--	-0.07 *(-1.10)*	-0.41 *(-2.35)*	--	-0.19 *(-1.83)*	--	0.22 *(1.74)*
Income between $100K to $200K	--	--	--	--	--	--	0.19 *(1.60)*
Household composition							
Single person household	0.23 *(2.18)*	--	0.23 *(2.11)*	--	--	-0.21 *(-1.50)*	--
Single female household	0.13 *(1.39)*	0.23 *(2.13)*	--	--	--	--	-0.49 *(-2.18)*
Presence of children (age less than 15 years)	--	--	--	0.19 *(2.04)*	0.10 (1.29)	0.15 *(2.23)*	--
Presence of employed individuals	--	--	0.19 (1.94)	--	--	0.10 *(1.40)*	--
Household structure (Base: Apartment/attached and other)							
Detached household	-0.21 *(-2.58)*	--	--	0.09 *(1.03)*	0.25 *(2.24)*	0.22 *(2.14)*	0.34 *(2.82)*
Parking facilities in the household							
Absence of a covered parking	--	--	--	-0.19 *(-2.22)*	-0.19 *(-2.22)*	-0.20 *(-1.91)*	-0.39 *(-2.37)*
Absence of electric charging facility in the parking of home/workplace	--	--	-0.18 *(-1.89)*	N.A.	--	N.A.	NA
Household's regional attributes							
Employment density more than 500 person/ sq. mile	--	--	--	-0.08 *(-1.51)*	--	--	-0.44 *(-3.14)*
Constant	--	-0.88 *(-3.42)*	-0.94 *(-3.28)*	-0.25 *(-2.37)*	-1.44 *(-3.46)*	-1.52 *(-3.45)*	-1.01 *(-3.35)*
Satiation function							
Household income characteristics ((Base: Income more than $200K))							
Income less than $25K	0.68 *(1.70)*	--	--	1.84 *(1.22)*	--	--	--
Income less than $25K to $50K	0.73 *(2.92)*	1.69 *(1.93)*	--	1.65 *(3.51)*	--	--	--
Income less than $50K to $100K	0.57 *(3.71)*	1.63 *(2.89)*	--	0.62 *(3.94)*	--	--	--
Income between $100K to $200K	--	1.53 *(2.77)*	--	--	--	--	--
Household composition							
Presence employed individuals	--	--	--	-0.47 *(-2.81)*	--	-1.17 *(-2.01)*	--
Regional attributes							
Employment density		--	--	--	--	--	0.31 *(1.61)*
Constant	*-0.02 (-0.98)*	0.44 *(1.06)*	1.33 *(2.30)*	0.24 *(0.60)*	1.97 *(3.72)*	1.81 *(2.82)*	0.14 *(0.37)*

--: Insignificant/same as the base category. *Same parameter across the alternatives preference parameter estimates that households with employed individuals had a higher preference for owning a van. Despite the high preference, such households are less likely to use them than households that do not have any employed individuals. This result reaffirms our earlier hypothesis that vans serve more of a leisurely purpose in households with employed individuals rather than providing day-to-day functionality. Finally, despite the lower preference for owning pickup trucks by households in regions of high employment densities, mileage accrued to these vehicles, if they are owned, is higher than that in regions of low employment densities. This is potentially because these vehicles typically serve the purpose of hauling goods and materials and therefore, would travel longer distances in urban regions with high employment density.

Budget Equation Results

The following Tobit budget equation (based on Equation (9)) was estimated:

$$\begin{aligned} y_q^* = &-10.22 + 4.83(\text{no. of driving licenses})_q + 3.92(\text{presence of workers})_q - 4.24(I(HInc_q < \$25\text{K})) \\ &-5.90(I(\$25\text{K} < HInc_q < \$50\text{K})) - 3.16(I(\text{Household uses work shuttle})_q) \\ &-1.94(I(\text{Population density} > 2000/\text{sq. mile})) - \underbrace{10.42}_{\lambda} \times (\zeta - \underbrace{0.76}_{\kappa}\left[\ln \sum_k \bar{a}_k \exp(\tau_k)\right], \end{aligned} \quad (13)$$

where index q represents a household, $HInc_q$ represents the qth household's annual income and, $I(\cdot)$ is an indicator function.

Notably, the scale of the error terms in the fractional split MDCEV model (κ) is estimated with a value of 0.76 and a corresponding t-statistic 2.25 (against 1), indicating that the scale parameter is significantly different from 1 (with more than 95% confidence level). Also, the linking parameter is significantly different from 0 (with a value of 10.42 and a t-statistic of 37.36), indicating that the desirability of a vehicle type increases the total mileage driven by the household. These effects can potentially provide better insights from a policy standpoint. For instance, the densification of a region by increasing employment density may likely reduce the total mileage driven by households in that region. However, this effect could not be captured directly in the budget equation, with the corresponding parameter turning out insignificant. On the other hand, this effect is indirectly captured through the linking function, where an increase in employment density results in lower preference for gasoline SUVs and pickup trucks, thereby reducing the total price index.

The parameters offer intuitive insights into the annual mileage. As the number of valid driving license holders increases in the household, total annual mileage is also likely to increase. The same trend follows with the presence of workers in the household as well. In the context of annual household income, the annual mileage by high-income households (with income more than $50,000) is likely to be higher than low-income households. Intuitively, households where individuals use work shuttle to commute accrue lesser mileage Finally, households in regions with high population density (more than 2000 persons per sq. mile) are likely to drive less.

Model Fit Statistics

Likelihood-Based Goodness of Fit Statistics

The log-likelihood of the final model (log-likelihood of -12429.02 with 72 parameters) is superior to the constant-only model (log-likelihood of -13150.50 with 16 parameters), indicating the importance of our empirical specification. Further, allowing free estimation of the scale parameter in the fractional split model generally improves the statistical fit of the model (final log-likelihood of -12450.10 with unit scale assumption and a total of 71 parameters). This result is demonstrated by the superior fit of the model with free scale parameter when compared with the model with restricted scale parameter (corresponding likelihood ratio test with a test statistic of 42.16 is greater than the critical squared value for a single degree of freedom at any reasonable significance level). The same trend is also observed in the holdout sample (with an LRT test statistic of 4.48 which is greater than the critical chi-squared value for a single degree of freedom at 5% significance level), indicating the importance of freely estimating the scale parameter. Overall, freely estimating the scale parameter will generally result in improvement in likelihood-based model-fit measures.

Non-Likelihood-Based Goodness of Fit Statistics
To further assess the importance of freely estimating the scale parameter, we undertake a comparative assessment of the prediction performance of the estimated model on the holdout sample. The level of accuracy in the predicted vehicle holding pattern is computed using weighted mean absolute percentage error (MAPE), which is computed as:

$$\text{Weighted MAPE } = \frac{\sum_i \left(\frac{O_i - P_i}{O_i} \right) O_i}{\sum_i O_i} \times 100 \, , \tag{14}$$

where O_i is the observed share (or continuous mileage) and P_i is the predicted share (or continuous mileage) for alternative *i*.

The predicted vehicle holding and usage pattern is presented in Table 3 (under the heading non-likelihood-based data fit). As indicated from the results, the discrete shares of vehicle holdings are predicted close to the observed vehicle holding patterns with the weighted MAPE of 5.19%. However, the predictions of mileage accrued to each of the vehicle category has a slightly magnified error of 8.67%. Despite this slight error, the total average mileage is predicted with great accuracy (a weighted percentage error of 1.57%). However, the number of households predicted with zero vehicle ownership was way off from the observed value (7.95% of households predicted with zero vehicle holdings as opposed to the observed value of 2.55%). This high error could be attributed to an inherent small proportion of households with zero car ownership. Overall, these predicted vehicle holding and usage patterns are reasonably close to the observed values in the holdout sample, with an overall weighted MAPE of 6.28%. In comparison, errors in predictions from the model with unit scale were slightly higher (a weighted MAPE of 7.03%). This result underscores the importance of freely estimating the scale parameter in the model.

Policy Evaluation
The two-stage MDC allows better evaluation of policies since it allows the possibility of both complementarity and/or substitution effects (through income effects). In this study, we undertake analysis of two specific policies that demonstrate the advantages of the two-stage MDC model. Specifically, we focus on understanding the effects of: (a) regional densification by increasing employment density, and (b) providing non-private commute opportunities to every household with employed individuals. In the context of quantifying changes in vehicle holdings due to densification of employment opportunities, households' vehicle holdings and mileage decisions were compared between two scenarios: i) all households living in regions with employment density less than 500 persons per sq. mile, and ii) all households living in regions of employment density more than 500 persons per sq. mile. Similarly, in the context of providing non-private commute options to workers, vehicle holdings were forecasted for two scenarios: i) none of the households

TABLE 3 Non-Likelihood-Based Data Measures (Prediction Performance) and Policy Evaluations

Non-likelihood-based data fit (holdout sample)	Gasoline car	Hybrid car	Electric car	Gasoline SUV	"Hybrid+" SUV	Van	Pickup truck
Observed in the holdout sample							
Percentage of zero-car households				2.55			
Percentage of households owning the vehicle type	56.72	14.50	13.10	41.29	3.71	7.07	16.24
Average mileage on the vehicle type (in 1000 miles)	9.84	10.83	10.87	10.97	10.74	9.07	8.49
Total average mileage driven (in 1000 miles)				15.53			
Predictions from the two-stage MDCEV model with free scale parameter							
Percentage of zero-car households				7.95			
Percentage of households owning the vehicle type	56.02	14.86	13.48	38.41	3.96	6.68	13.3
Average mileage on the vehicle type (in 1000 miles)	11.43	10.05	10.13	10.43	9.42	8.17	8.22
Weighted MAPE	Discrete WMAPE: 5.19%; Continuous WMAPE: 8.67%; Overall WMAPE: 6.28%						
Average total miles driven by a household (in 1000 miles)				15.29			
Policy evaluations							
Policy I: Employment densification of regions (Base case: Employment density <500 persons/sq. mile; Policy case: Employment density >500 persons/sq. mile							
% change in H.H.s with zero vehicle ownership				6.34			
% change in vehicle ownership	4.01	10.62	10.45	-3.79	-5.44	-0.12	-52.46
% change in vehicle mileage accrued on each vehicle type#	6.51	21.15	11.91	-6.11	-8.63	1.51	-55.78
% change in total average miles driven				-1.60			
Policy II: Provision of non-private commute services (Base case: No household using work shuttle services; Policy case: All households with employed individuals using work shuttle							
% change in H.H.s with zero vehicle ownership				25.15			
% change in vehicle ownership	-2.81	-5.23	-3.17	-1.71	-1.20	0.58	-1.06
% change in vehicle mileage accrued on each vehicle type#	-12.34	-14.74	-14.59	-10.87	-14.46	-7.61	-9.67
% change in total average miles driven				-12.11			

#Averaged across all households

with employed individuals (full-time or part-time) using work shuttle, and ii) all households with employed individuals (full-time or part-time) using work-shuttle services. Considering the benefits of freely estimating the scale parameter, the policy evaluation is undertaken using the model that freely estimates the scale parameter. The results for the above two policy evaluations are provided in the second set of rows in Table 3 (under the heading Policy evaluations).

Policy I: Effect of Employment Densification

The effects of employment densification manifest as a reduction in ownership of gasoline SUVs and pickup trucks (see the set of rows in Table 3 under the heading Policy I). Interestingly, densification of employment opportunities also results in a reduction in ownership of hybrid SUVs, and vans as well (akin to a complementarity effect). At the same time, ownership of cars (across all three fuel categories) increases (akin to a substitution effect). Despite this increase in car ownership, the total vehicle ownership (indicated by an increase in the percentage of households with zero-car ownership) as well as the total average miles driven, reduces. This reduction in total vehicle ownership levels and miles driven is an indirect effect of employment densification (through the linking effect). From a policy standpoint, these findings are indicative that employment densification can be an important lever to control total miles driven and reduce total vehicle ownership. However, it can still result in an increase in ownership of small cars (both gasoline-powered and electric). Overall, with a decrease in ownership of SUVs, vans and trucks, the total on-road emissions in this scenario can potentially decrease since light-duty trucks are the biggest source of GHG emissions among all vehicle classes in the U.S. *(30)*.

Policy II: Effect of Provision of Non-Private Commute Options

The effects of the provision of non-private commute options were evaluated through the usage of work shuttle services by households. The results of this policy evaluation are provided in the second set of rows under the heading Policy II in Table 3. The total average miles driven by households show a significant 12.1% decrease as households start using work shuttles (see the last row of Table 3). Interestingly, the use of work shuttles by workers also leads to a reduction in vehicle ownership, with the number of households with zero vehicles showing a significant 25% increase. This reduction in vehicle ownership also translates into a small reduction in vehicle ownership of each vehicle type. Therefore, from a policy standpoint, providing work shuttle services can potentially result in reduced total miles driven and vehicle ownership levels, thereby potentially reducing overall on-road GHG emissions. This result again highlights the importance of endogenous modelling of budget within the MDC model framework since the optimal allocations derived through the traditional MDC models are not sensitive to income effects.

CONCLUSIONS

This paper analyses households' vehicle holdings and usage decisions in California using a two-stage MDC choice model framework that allows endogenous estimation of the budget as well as the allocation of the budget to the elementary alternatives, all within a utility theoretic framework. In doing so, we revisit the prior formulation of the two-stage MDC model and relax the unit scale requirement in the model. With this extended formulation, the empirical assessment of households' vehicle holding and usage in California is undertaken using the 2019 California Vehicle Survey data. The empirical results indicate that freely estimating the scale parameter improves model fit. Policy simulations using the empirical model indicated that regional densification measures, such as increasing employment density reduces ownership of SUVs, vans, and pickup trucks, while increasing the ownership of passenger cars (both gasoline and electric). Despite this increase in passenger cars, the total vehicle ownership reduces as a result of employment densification. Further, provision of alternative, non-personal commute options (such as work shuttles) can significantly reduce personal vehicle ownership and miles travelled. Overall, both policies can potentially lead to reduction in total on-road emissions.

ACKNOWLEDGMENTS

This research was partially supported by the U.S. Department of Transportation through the Center for Understanding Future Travel Behavior and Demand (TBD) (Grant No. 69A3552344815 and No. 69A3552348320). All authors acknowledge the support from the Indian Ministry of Education through SPARC for encouraging international collaborations. The authors are grateful to Lisa Macias for her help in formatting this document.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: S. Saxena, C.R. Bhat, A.R. Pinjari; analysis and interpretation of results: S. Saxena, C.R. Bhat, A.R. Pinjari; draft manuscript preparation: S. Saxena, C.R. Bhat, A.R. Pinjari. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Mallett, W. J. (2018). Trends in public transportation ridership: Implications for federal policy. *Washington, DC: Congressional Research Service*.
2. Morris, E. A., Blumenberg, E., and Guerra, E. (2020). Does lacking a car put the brakes on activity participation? Private vehicle access and access to opportunities among low-income adults. *Transportation Research Part A*, 136, 375-397.
3. Andor, M. A., Gerster, A., Gillingham, K. T., and Horvath, M. (2020). Running a car costs much more than people think—stalling the uptake of green travel. *Nature*, 580(7804), 453-455.
4. Moody, J., and Zhao, J. (2020). Travel behavior as a driver of attitude: Car use and car pride in U.S. cities. *Transportation Research Part F*, *74*, 225-236.
5. Hula, A., Maguire, A., Bunker, A., Rojeck, T., and Harrison, S. (2022). *The 2022 EPA Automotive Trends Report: Greenhouse Gas Emissions, Fuel Economy, and Technology Since 1975* (No. EPA-420-R-22-029).
6. Bleviss, D. L. (2021). Transportation is critical to reducing greenhouse gas emissions in the United States. *Wiley Interdisciplinary Reviews: Energy and Environment*, *10*(2), e390.
7. Zhang, W., Guhathakurta, S., and Khalil, E. B. (2018). The impact of private autonomous vehicles on vehicle ownership and unoccupied VMT generation. *Transportation Research Part C*, *90*, 156-165.
8. Kim, S. H., Mokhtarian, P. L., and Circella, G. (2020). Will autonomous vehicles change residential location and vehicle ownership? Glimpses from Georgia. *Transportation Research Part D*, *82*, 102291.
9. Fang, H. A. (2008). A discrete–continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B*, *42*(9), 736-758.
10. Nguyen, N. T., Miwa, T., and Morikawa, T. (2017). Vehicle type choice, usage, and CO2 emissions in Ho Chi Minh city: analysis and simulation using a discrete-continuous model. *Asian Transport Studies*, *4*(3), 499-517.
11. Bhat, C. R., and Sen, S. (2006). Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model. *Transportation Research Part B*, *40*(1), 35-53
12. You, D., Garikapati, V. M., Pendyala, R. M., Bhat, C. R., Dubey, S., Jeon, K., and Livshits, V. (2014). Development of vehicle fleet composition model system for implementation in activity-based travel model. *Transportation Research Record*, *2430*(1), 145-154.
13. Vyas, G., Paleti, R., Bhat, C. R., Goulias, K. G., Pendyala, R. M., Hu, H. H., and Bahreinian, A. (2012). Joint vehicle holdings, by type and vintage, and primary driver assignment model with application for California. *Transportation Research Record*, *2302*(1), 74-83.
14. Augustin, B., Pinjari, A. R., Eluru, N., and Pendyala, R. M. (2015). Estimation of annual mileage budgets for a multiple discrete-continuous choice model of household vehicle ownership and utilization. *Transportation Research Record*, *2493*(1), 126-135.
15. Bhat, C.R., (2022). A new closed-form two-stage budgeting-based multiple discrete-continuous model. *Transportation Research Part B*, *164*, 162–92.

16. Rouwendal, J., and Boter, J. (2009). Assessing the value of museums with a combined discrete choice/count data model. *Applied Economics*, *41*(11), 1417-1436.
17. Hausman, J. A., Leonard, G. K., and McFadden, D. (1995). A utility-consistent, combined discrete choice and count data model assessing recreational use losses due to natural resource damage. *Journal of Public Economics*, *56*(1), 1-30.
18. Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: Journal of the Econometric Society*, 24-36.
19. Transportation Secure Data Center. (2019). National Renewable Energy Laboratory. Accessed May 2023, *https://nrel.gov/tsdc*
20. U.S. Department of Energy (2021). Vehicle Registration Counts by State. https://afdc.energy.gov/vehicle-registration Accessed July 2023.
21. Blackley, J (2023). Which states drive the most pickup trucks? iSeeCars. https://www.iseecars.com/which-states-drive-the-most-pickup-trucks-studycks Accessed July 2023.
22. California Energy Commission (2018). 2015-2017 California Vehicle Survey. RSG. https://www.nrel.gov/transportation/secure-transportation-data/assets/pdfs/cec_2015-2017_california_vehicle_survey_report.pdf Accessed July 2023.
23. U.S. Census Bureau (2017-21). https://www.census.gov/quickfacts/fact/table/CA/INC110221. Accessed July 2023.
24. Population and Census Data (2020). Employment Development Department. State of California. https://labormarketinfo.edd.ca.gov/Population_and_Census.html#POP. Accessed July 2023.
25. Mondal, A., and Bhat, C. R. (2021). A new closed form multiple discrete-continuous extreme value (MDCEV) choice model with multiple linear constraints. *Transportation Research Part B*, *147*, 42-66.
26. Brecht, P. (2020). 2020-2023 Investment Plan Update for the Clean Transportation Program: Lead Commissioner Report. California Energy Commission.
27. Brownstone, D., and Fang, H. (2014). A vehicle ownership and utilization choice model with endogenous residential density. *Journal of Transport and Land Use*, *7*(2), 135-151.
28. Plötz, P., Schneider, U., Globisch, J., and Dütschke, E. (2014). Who will buy electric vehicles? Identifying early adopters in Germany. *Transportation Research Part A*, *67*, 96-109.
29. Woofter, P. (2021). Van life on the rise – a camper van California road trip turned permanent. https://themilsource.com/2021/03/03/van-life-on-the-rise-a-camper-van-california-road-trip-turned-permanent/ Accessed July 2023.
30. Statista (2021). Greenhouse gas emissions from on-road vehicles in the United States. https://www.statista.com/statistics/1120499/us-road-vehicle-ghg-emissions-by-vehicle-type/#:~:text=Light%2Dduty%20trucks%20in%20the,some%20374.2%20MtCO%E2%82%82e%20that%20year. Accessed July 2023.