

**University of Illinois at Urbana-Champaign**

---

**From the Selected Works of Don Fullerton**

---

August, 2013

# Vehicle Choices, Miles Driven, and Pollution Policies

Ye Feng, *University of Texas at Austin*

Don Fullerton, *University of Illinois at Urbana-Champaign*

Li Gan, *Texas A&M University*



Available at: [https://works.bepress.com/don\\_fullerton/9/](https://works.bepress.com/don_fullerton/9/)

# Vehicle choices, miles driven, and pollution policies

Ye Feng · Don Fullerton · Li Gan

Published online: 5 May 2013

© Springer Science+Business Media New York 2013

**Abstract** Mobile sources contribute large percentages of each pollutant, but technology is not yet available to measure and tax emissions from each vehicle. We build a behavioral model of household choices about vehicles and miles traveled. The ideal-but-unavailable emissions tax would encourage drivers to abate emissions through many behaviors, some of which involve market transactions that can be observed for feasible market incentives (such as a gas tax, subsidy to new cars, or tax by vehicle type). Our model can calculate behavioral effects of each such price and thus calculate car choices, miles, and emissions. A nested logit structure is used to model discrete choices among different vehicle bundles. We also consider continuous choices of miles driven and the age of each vehicle. We propose a consistent estimation method for both discrete and continuous demands in one step, to capture the interactive effects of simultaneous decisions. Results are compared with those of the traditional sequential estimation procedure.

**Keywords** Discrete choices · Continuous demands · Environmental incentives

**JEL Classification** D12 · H23 · Q58

---

Y. Feng

Department of Economics, University of Texas, Austin, TX 78712, USA

D. Fullerton (✉)

Department of Finance, University of Illinois, Champaign, IL 61820, USA

e-mail: dfullert@illinois.edu

D. Fullerton · L. Gan

NBER, Cambridge, MA, USA

L. Gan

Department of Economics, Texas A&M University, College Station,  
TX 77843, USA

The standard case for market-based incentives requires a tax or price on each unit of emissions. Each form of abatement is then pursued until the marginal cost of reducing pollution matches the tax per unit of pollution, and the resulting combination of abatement technologies minimizes social costs (Pigou 1920). For vehicles, a tax on emissions could induce drivers to: (1) buy a newer, cleaner car, (2) buy a smaller, more fuel-efficient car, (3) fix their broken pollution control equipment, (4) buy cleaner gasoline, (5) drive less, (6) drive less aggressively, and (7) avoid cold start-ups.<sup>1</sup> Moreover, economic efficiency requires different combinations of these methods for different consumers.

Yet the technology is not available to measure each car's emissions in a reliable and cost-effective manner. On-board diagnostic equipment is imperfect, and it is costly to retrofit millions of vehicles (Harrington and McConell 2003). Remote sensing is less expensive and has been used to identify high-polluting vehicles, but it cannot measure emissions clearly enough to tax each car.<sup>2</sup> Moreover, vehicle emissions are important. In 2001, vehicles in the U.S. contributed 27 percent of volatile organic compounds (VOC), 37 percent of nitrogen oxides ( $\text{NO}_x$ ), and 66 percent of carbon monoxide (CO) emissions.<sup>3</sup>

For these reasons, vehicle emission policies have relied almost solely on mandates: refineries must make clean gasoline, and new cars must meet required emission standards.<sup>4</sup> Yet the mandated combination of abatement methods is unlikely to match the cost-minimizing combination that households would choose if faced with a tax on emissions. In fact, the cost of abatement using such mandates can be several times the minimum cost achieved by using an emissions tax (Newell and Stavins 2003).

While the inability to measure emissions may preclude a vehicle emissions tax, it does not preclude any use of incentives. Those who sell new or used cars or light-trucks can collect tax on vehicle characteristics that are associated with emissions, or provide subsidy for vehicles with low emissions. Most states charge annual registration fees that can be made to depend on vehicle characteristics. Such policies might reduce emission rates, while changes in the gasoline tax can reduce miles driven.

What vehicle characteristics or behaviors should be targeted by a tax or subsidy? How would consumers react to those new incentive instruments? How much would each tax reduce emissions? To address these questions, we build a general purpose model of discrete choices by households about how many cars to own and what types

<sup>1</sup> Heeb et al. (2003) find that cold start emissions rates (in g/km traveled) exceed stabilized emissions rates by a factor of two to five, depending on the pollutant. Sierra Research (1994) finds that a car driven aggressively has carbon monoxide emissions that are almost 20 times higher than when driven normally.

<sup>2</sup> See Sierra Research (1994). Remote sensing in Texas (<http://www.tnrc.state.tx.us/air/ms/vim.html#im3>) and Albuquerque NM (<http://www.cabq.gov/aircare/rst.html>) is used in 2005 to identify polluting vehicles.

<sup>3</sup> See [http://www.bts.gov/publications/transportation\\_statistics\\_annual\\_report/2004/](http://www.bts.gov/publications/transportation_statistics_annual_report/2004/). We focus on local pollutants, where emission rates depend on car characteristics. In contrast,  $\text{CO}_2$  is linked directly to gas use.

<sup>4</sup> In the U.S., new cars face emission standards of .254 g/km of HC's, 2.11 g/km of CO, and .248 g/km of  $\text{NO}_x$ . Light trucks face a variety of weaker standards, but all are scheduled to become more stringent. These figures pertain to a test in the U.S. with a cold start-up phase, a transient phase at different speeds, and a hot start phase, for a total distance of 18 km at an average speed of 34 km/h.

of cars to own, plus continuous choices about how far to drive. In our model, we embrace individual heterogeneity. We estimate all decisions simultaneously, and we use the estimated parameters to predict the effects of certain price changes on choices and on emissions.

Several existing papers explore market incentives that could be used in place of a tax on emissions.<sup>5</sup> In addition, several papers estimate models of the discrete choice among vehicle bundles (including number, size, and age categories).<sup>6</sup> Some models estimate the demand for gasoline or for vehicle miles traveled (VMT) as functions of price and income (as reviewed in [Harrington and McConell 2003](#)). As well, we note that other models predict emissions.<sup>7</sup> A major contribution of our research, then, is to include all such choices simultaneously. In general, we capture the effect of any price change on each household's choices about the number of vehicles to buy, the type and age of each, the consequent emissions rates, miles driven, and the consequent total emissions.

In a two-step procedure, [Dubin and McFadden \(1984\)](#) estimate a discrete choice model (for household appliances) and use the predicted shares to correct for endogeneity in the estimation of a continuous choice (usage hours). Others extend this model to the discrete choice among vehicle bundles and a continuous choice of miles (e.g. [Goldberg 1998](#); [West 2004](#)). A single set of parameters appear both in the indirect utility function used to estimate discrete choices and in continuous demands. Using this sequential procedure, however, the estimated parameters of the continuous demand are not constrained to match the same parameters in the estimated discrete choice model.

[Hanemann \(1984\)](#) proposes a method to estimate discrete and continuous demands at the same time. [Bhat \(2005\)](#) allows consumers to choose multiple discrete alternatives, while [Bhat et al. \(2009\)](#) use this approach to analyze the effects of gasoline price on the choice of vehicle type, vintage, and use. [Bento et al. \(2009\)](#) estimate car choices and miles driven with one set of parameters; they combine those demands with a model of new and used car supplies to simulate efficiency and distributional effects of an increase in the gasoline tax. These papers estimate discrete and continuous demands in one step with one set of parameters, but without the same individual error term in both choices. This unobserved individual heterogeneity is a key factor in the Dubin-McFadden model.

Relative to this literature, we make a number of contributions. First, we capture the simultaneity of these decisions by proposing a method for consistent estimation of both discrete and continuous choices in one step with the same individual error term, yielding a single set of parameters. In other words, whereas the Dubin-McFadden method corrects for selection of vehicle on the choice of miles, our procedure allows for individual unobserved heterogeneity to affect both the fuel demand and the choice

<sup>5</sup> For examples, see [Eskeland and Devarajan \(1996\)](#), [Innes \(1996\)](#), [Kohn \(1996\)](#), [Train et al. \(1997\)](#), [Plaut \(1998\)](#), [Sevigny \(1998\)](#), and [Fullerton and West \(2002, 2010\)](#). For a review article, see [Parry et al. \(2007\)](#).

<sup>6</sup> See [McFadden \(1979\)](#), [Manning and Clifford \(1985\)](#), [Train \(1986\)](#), [Brownstone et al. \(1996\)](#), [Goldberg \(1998\)](#), [Brownstone and Train \(1999\)](#), [West \(2004\)](#), and other papers reviewed in [McFadden \(2001\)](#).

<sup>7</sup> For example, the U.S. Environmental Protection Agency ([U.S. Environmental Protection Agency 1998](#), p.3–68) discusses the use of EPA's MOBILE5a model or California's EMFAC7F model.

of vehicles simultaneously. Second, we allow for two continuous choices of miles – in each vehicle of a two-vehicle household. These choices are bundle-specific.<sup>8</sup> Third, we allow for an additional continuous choice of the age of each vehicle. Fourth, we use the estimated parameters not only to predict changes in choices about vehicles and miles, but also how those choices affect emissions.<sup>9</sup>

As it turns out, results for all continuous demands are broadly similar for the sequential and simultaneous models. For discrete choices, however, our simultaneous model finds substantially larger effects from a change in the gas price per mile, income, or vehicle-specific costs. Signs of some elasticities are reversed. In other words, household-specific heterogeneity does affect discrete choices.

We find that a higher price of gasoline would shift households out of the Car-SUV pair and into the bundle with two cars. It also would reduce miles driven. Both of these changes reduce emissions. A tax on vehicle age would induce shifts to newer vehicles and would also shift families out of bundles with an SUV. Both of these changes also reduce emissions. Similarly, a tax on SUV's would shift families into cars and reduce emissions. The size of such a shift is important for environmental policy. Rather than pin down the exact size of key parameters, however, this paper points to important problems with existing methods and suggests an alternative approach with more internal consistency.

The next section describes a behavioral choice model for one-vehicle households and then extends it to consider two-vehicle bundles. It also presents a new method designed for jointly estimating all discrete and continuous choices. Section 2 describes data sources, while Sect. 3 provides estimation results for both discrete and continuous demands. Section 4 compares elasticities, and Sect. 4 concludes.

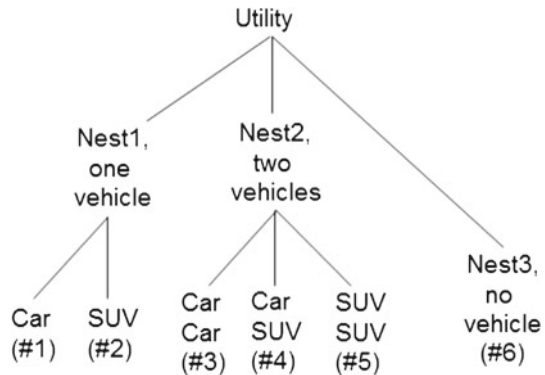
## 1 The model and estimation

All cars in a given year are made to a single emission rate standard, while a weaker standard has applied to “sports utility vehicles” (SUV, for short, but defined here to include all light trucks and vans). For each vehicle, we therefore model the discrete choice between a car and SUV. Each household also faces a discrete choice among a finite number of vehicle bundles. The nesting structure is shown in Fig. 1. One choice is the number of vehicles (0, 1, or 2), and another choice for each vehicle is the type of vehicle (a car or an SUV). Other structures are possible, of course, including all options on the same level. For our final model, we use the structure shown in the figure.

<sup>8</sup> With a higher price of gas, some households might drive fewer miles in their SUV and *more* in their car. We do not estimate separately the miles in each vehicle, but we do estimate a change for the (Car, SUV) bundle that can differ from the (Car, Car) bundle. Other papers have estimated substitution between vehicles within the family, but they treat the vehicles as given rather than chosen. [Greene and Hu \(1985\)](#) find that this kind of substitution occurs to a large extent in some households, while [Sevigny \(1998\)](#) finds small effects.

<sup>9</sup> Our household responses represent market outcomes only if supply curves were horizontal. The simulation of a change in the price of getting a car that is one year newer can be interpreted as a new local tax or subsidy in a small open jurisdiction that can import more of those newer cars at a constant price. However, our demand system could be combined with some other estimates of supply to calculate equilibrium outcomes.

**Fig. 1** Nesting structure for choice among vehicle bundles



The choice of vehicle age is also important for emissions, so we also model this decision.<sup>10</sup> A household that chooses two vehicles has four continuous choices: age of each vehicle and miles to drive each vehicle.<sup>11</sup> Our model requires a price that does not depend on the amount demanded, but the price per year of age is not linear (because owning a new car costs more depreciation per year than owning an old car). We therefore construct a continuous variable called “*Wear*”, the fraction of a vehicle that has depreciated (between 0 and 1). A constant rate of depreciation means that *Wear* is a nonlinear function of age, but then the price per unit of *Wear* does not depend on its amount. This price is estimated for each vehicle type using hedonic regressions below. The household makes a discrete choice about a brand-new “concept vehicle”, and then gets reimbursed by the price of *Wear* for accepting an older car. In other words, in our model, a household makes simultaneous decisions about which concept vehicles, how old, and miles to drive. Specifically, “*Wear*” is calculated for each car in our sample by assuming 20% depreciation per year, so  $Wear = 1 - (1 - 0.2)^{age}$ . Thus, a new car has *Wear* = 0.

Then, since choice of age is considered separately, each discrete vehicle bundle must be defined in a way that is independent of age. For this reason, we define each “concept” vehicle as a bundle of attributes of a brand-new vehicle (car or SUV). The household must pay the price of that brand-new vehicle (the “capital cost”), but then it gets back some money for accepting *Wear* on that vehicle (the “reimbursement” price of *Wear*).

Since the discrete choice in Dubin and McFadden (1984) involves only two alternatives, that paper can use a simple logit model. Our model has six choices, however, and so we require a more general logit structure. We use the nested logit. The next subsection describes the simple case for households with only one vehicle, and the second

<sup>10</sup> Older vehicles have higher emissions both because older vintages were produced to weaker standards and because pollution control equipment deteriorates with age. Panel data would be required to distinguish these.

<sup>11</sup> Fullerton and West (2010) also simulate effects of incentives in a model of heterogeneous households’ continuous choices of car size, car age, and VMT, but they use calibrated rather than estimated parameters. That model avoids discrete choices, but it considers only one car per agent. In our model, we estimate discrete choices to consider the household’s number of vehicles.

subsection considers multi-vehicle households. In the third and fourth sub-sections, we discuss the estimation procedure and elasticity calculations.

### 1.1 Our model of car choice and miles driven

This description starts with the choices of *VMT* and *Wear*, assuming that a one-car household has already chosen vehicle number-and-type bundle  $i$ . Given bundle  $i$ , an agent's direct utility is a function of *VMT*, *Wear*, and another consumption good  $c$ . That is,  $U = U(VMT_i, Wear_i, c_i)$ . Given income  $y$ , the budget constraint is given by:

$$\frac{p_g}{MPG_i} VMT_i - q_i Wear_i + c_i = y - r_i, \quad (1)$$

where  $p_g$  is the price of gasoline (in dollars per gallon), and  $MPG_i$  is fuel efficiency (in miles per gallon), so that  $p_i \equiv p_g/MPG_i$  is the marginal price per mile in the  $i$ th vehicle bundle. The “reimbursement” price of *Wear* for vehicle type  $i$  is denoted as  $q_i$ . The price of the other consumption good is normalized to be 1. The annualized capital cost of the concept-vehicle bundle is  $r_i$ . Thus, gasoline is the only cost per mile, whereas capital cost is a fixed cost of each bundle.<sup>12</sup> The indirect utility for bundle  $i$  is a function of household income and prices, denoted as  $V(y - r_i, p_i, q_i)$ .

One common way to obtain the indirect utility function is to use parametric demand and then solve a system of partial differential equations using Roy's identity (Hausman 1981). For comparability with other studies, we want *VMT* demand as a log-linear function of the price per mile  $p_i$ , available income  $y - r_i$ , and a vector of observed socio-demographic variables  $x$ . We then add the reimbursement price  $q_i$  to that equation to get:

$$\ln(VMT_i) = \alpha_V^i + \alpha_p^i p_i - \alpha_q q_i - \beta(y - r_i) + x' \gamma + \eta, \quad (2)$$

where  $\eta$  represents an agent-specific unobserved factor (see below). Also, we assume

$$r_i = (\delta + \rho) k_i, \quad (3)$$

where  $k_i$  is the total capital value of bundle  $i$  (depreciated or market value),  $\delta$  is the annual rate of further depreciation in value, and  $\rho$  represents the interest and maintenance cost. When we plug (3) into (2) and integrate, the implied indirect utility is:

$$V_i = \frac{1}{\beta} \exp(-\alpha_0^i + \beta y - \beta_1 k_i - x' \gamma - \eta) - \frac{1}{\alpha_p^i} \exp(\alpha_p^i p_i - \alpha_q q_i) + \varepsilon_i, \quad (4)$$

<sup>12</sup> Time variation in gasoline prices may cause time variation in used vehicle prices. Our use of cross-section data helps avoid this problem.

where  $\beta_1 = \beta(\delta + \rho)$ .<sup>13</sup> This equation includes an extra additive error  $\varepsilon_i$  that is bundle-specific. As in the usual discrete choice model, this error term represents the difference between true individual utility at choice  $i$  and the calculated utility level.<sup>14</sup> For households who choose the no-vehicle bundle #6, continuous variables such as  $p_i$ ,  $q_i$ , and  $VMT_i$  are unobservable. Implicitly, we assume that these households may purchase a bicycle or a fare card for public transportation with a fixed fee, similar to the capital cost  $k_i$ . With no cost per mile or of *Wear*, their second exponential term in (4) is 1.0. Their capital cost  $k_i$  is unobserved, so  $\beta_1 k_i$  and  $\alpha_0^6$  are not separately identifiable. Since we allow for a choice-specific intercept, however, we combine both terms into one constant,  $\alpha_0^6$ .

Note that the simple addition of  $\alpha_q q_i$  to Eq. (2) dictates the form of indirect utility in (4). This indirect utility then implies specific forms for both demands:<sup>15</sup>

$$\ln(VMT_i) = \alpha_V^i + \alpha_p^i p_i - \alpha_q q_i - \beta y + \beta_1 k_i + x' \gamma + \eta \quad (5a)$$

$$\ln(Wear_i) = \alpha_W^i + \ln(\alpha_q / \alpha_p^i) + \alpha_p^i p_i - \alpha_q q_i - \beta y + \beta_1 k_i + x' \gamma + \eta \quad (5b)$$

This specification has pros and cons. One limitation is the use of specific functional forms, but these log-linear forms are comparable to prior literature and allow for two different demand functions (5a, 5b) that are consistent with a single indirect utility function (4). An advantage of this specification is that it allows the price of *Wear* ( $q_i$ ) to enter the *VMT* demand, and price of *VMT* ( $p_i$ ) to enter the *Wear* demand, but a limitation is that the expression  $\alpha_p^i p_i - \alpha_q q_i$  enters both demands the same way.<sup>16</sup> Also, both continuous demands have the same income effect,  $\beta$ . A more general model could not be estimated. Note, however, that we have added generality where it matters most. In particular, the price per mile has a bundle-specific coefficient ( $\alpha_p^i$ ), to allow for different effects on the demand for miles in each type of vehicle. Thus a gas tax might decrease miles in an SUV more than in a car, in a way that depends on fuel efficiency, and the change in miles for each household can depend on their vehicle types.

<sup>13</sup> Our model provides estimates of  $\beta$  and  $\beta_1$ , and these can be used to calculate  $(\delta + \rho)$ , but we do not provide separate estimates of  $\delta$  and  $\rho$ . Some of our steps below require an assumption about  $\delta$ , and we use 20 % for this purpose. Estimates of the depreciation rate for automobiles range from 33 % (Jorgenson 1996) or 30 % (Hulten and Wykoff 1996) to 15 %, the rate implicit in the vehicle depreciation schedule currently used by the Bureau of Economic Analysis. We use 20 % because it falls between these bounds.

<sup>14</sup> Also, because of this integration, note that the intercept in (4) may be different from the intercept in (2).

<sup>15</sup> More general demand functions such as translog demand or the almost ideal demand system imply much more complicated indirect utility functions that could not be estimated. Also, note that no-vehicle households have zero marginal prices, so they have constant miles traveled (conditioned on observed socio-demographic variables and total income). Thus, no continuous demand equations are needed for these households.

<sup>16</sup> Thus, a change in  $p_i$  must have the same effect on *Wear* that it has on miles. We tried other models, including one where indirect utility has separate terms  $\exp(\alpha_p^i p_i)$  and  $\exp(\alpha_q q_i)$ , so that  $p_i$  would have no effect on *Wear*, and  $q_i$  would have no effect on *VMT*. That model would not converge, and anyway it is restrictive by assuming no cross-price effects. We also tried models with more coefficients, to relax these restrictions, and we tried many starting points, but only the model in (4) and (5a, 5b) could be estimated simultaneously for discrete and continuous choices (especially for two-vehicle bundles considered below).

## 1.2 Two-vehicle households

So far, the model above considers only one vehicle, but many households have two vehicles and thus two continuous choices of miles and two continuous choices of *Wear*. We have the observed *VMT* and *Wear* for each vehicle, so we can incorporate all four continuous choices.<sup>17</sup> The direct utility for a two-vehicle household choosing bundle  $i$  is  $U(VMT_{i1}, VMT_{i2}, Wear_{i1}, Wear_{i2}, c_i)$ . The budget constraint is given by:

$$\frac{P_g}{MPG_{i1}} VMT_{i1} + \frac{P_g}{MPG_{i2}} VMT_{i2} - q_{i1} (Wear_{i1}) - q_{i2} (Wear_{i2}) + c_i = y - r_i, \quad (6)$$

where  $q_{ij}$  are reimbursement prices for *Wear* in the two vehicles of bundle  $i$  ( $j = 1, 2$ ). Also,  $p_{ij} \equiv p_g / MP \equiv G_{ij}$  is the price per mile using the  $j^{th}$  car of bundle  $i$ . We consider the indirect utility function as follows:

$$V_i = \frac{1}{\beta} \exp \left( -\alpha_0^i + \beta y - \beta_1 k_i - x' \gamma - \eta \right) - \frac{1}{\alpha_{p1}^i} \exp \left( \alpha_{p1}^i p_{i1} + \alpha_{p2}^i p_{i2} - \alpha_{q1} q_{i1} - \alpha_{q2} q_{i2} \right) + \varepsilon_i \quad (7)$$

The indirect utility in (7) is similar to (4) except for two extra terms related to the second vehicle's gasoline price  $p_{i2}$  and reimbursement price  $q_{i2}$ . By Roy's identity, given that the household has chosen bundle  $i$  in (7), the four continuous demands are:

$$\ln(VMT_{i1}) = \alpha_{V1}^i + \alpha_{p1}^i p_{i1} + \alpha_{p2}^i p_{i2} - \alpha_{q1} q_{i1} - \alpha_{q2} q_{i2} - \beta y + \beta_1 k_i + x' \gamma + \eta \quad (8a)$$

$$\ln(VMT_{i2}) = \alpha_{V2}^i + \ln(\alpha_{p2}^i / \alpha_{p1}^i) + \alpha_{p1}^i p_{i1} + \alpha_{p2}^i p_{i2} - \alpha_{q1} q_{i1} - \alpha_{q2} q_{i2} - \beta y + \beta_1 k_i + x' \gamma + \eta \quad (8b)$$

$$\ln(Wear_{i1}) = \alpha_{W1}^i + \ln(\alpha_{q1}^i / \alpha_{p1}^i) + \alpha_{p1}^i p_{i1} + \alpha_{p2}^i p_{i2} - \alpha_{q1} q_{i1} - \alpha_{q2} q_{i2} - \beta y + \beta_1 k_i + x' \gamma + \eta \quad (8c)$$

$$\ln(Wear_{i2}) = \alpha_{W2}^i + \ln(\alpha_{q2}^i / \alpha_{p1}^i) + \alpha_{p1}^i p_{i1} + \alpha_{p2}^i p_{i2} - \alpha_{q1} q_{i1} - \alpha_{q2} q_{i2} - \beta y + \beta_1 k_i + x' \gamma + \eta \quad (8d)$$

These demands generalize those of a one-vehicle household in (5a, 5b) by including terms for  $p_{i2}$  and  $q_{i2}$  (and so we refer to (8a–8d) for “all” demands). The demand for  $VMT_{i2}$  is symmetric to  $VMT_{i1}$  in explanatory variables, but it is non-linear in parameters of both  $p_{i1}$  and  $p_{i2}$ . The demands for  $Wear_{ij}$  ( $j = 1, 2$ ) are similarly defined.

<sup>17</sup> Another interesting question is about each household member's choice of miles driven (in either car), but we have no such data. As described below, we have only data on miles driven in each vehicle.

### 1.3 A procedure to estimate discrete and continuous demands simultaneously

Note that the same parameters appear in both discrete and continuous choice functions, yet previous literature has estimated these choice models separately. Often the estimates for the same parameters are different not only in magnitude but also in sign. In this sub-section, we propose a procedure for simultaneous estimation of bundle choice, vehicle age, and miles driven. We start with separate discussion of car choice and miles driven, and then how we combine them in a single estimation procedure.

Following McFadden's random utility hypothesis, vehicle bundle  $i$  is chosen if and only if:  $V_i \geq V_j$  for all  $j \neq i$ . The unconditional expected share for bundle  $i$  then is:

$$S_i = \int \Pr(V_i > V_j, \forall j \neq i | \eta) f(\eta) d\eta \quad (9)$$

where  $S_i$  is the share choosing bundle  $i$ , and  $f(\eta)$  is the probability density function of the agent-specific error  $\eta$ . We are now in a position to describe the importance of  $\eta$ . On the one hand, individual heterogeneity represented by  $\eta$  could directly affect the choice of bundle. On the other hand, observed demands for *VMT* and *Wear* are conditional on that choice. Since the choice of vehicle bundle is endogenous, the estimated demands for *VMT* and *Wear* could be biased if the influence of  $\eta$  in (9) is ignored. In the model of Dubin and McFadden (1984), the error term  $\eta$  can be cancelled out from the inequality  $\{V_i > V_j, \forall j \neq i\}$ , which simplifies the calculation of probabilities (that is, the integration over  $\eta$  in Eq. (9) is not necessary). In such a model,  $\eta$  appears only in the continuous demands, so this individual heterogeneity does not affect the choice of vehicle bundle directly. They can estimate the discrete model with error  $\varepsilon_i$  for each bundle, and then, given predicted bundle shares, they estimate the continuous choices with errors  $\eta$ .

Yet, our purpose here is to *retain* individual-specific heterogeneity  $\eta$  and its effect on bundle choice. Thus, the evaluation of probabilities in our model involves integration over all error components  $(\varepsilon, \eta)$ , where  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_J)$ , and where  $J$  is the number of possible vehicle bundles. In our model, the  $\varepsilon_i$  are assumed to be distributed with a generalized extreme value (GEV) distribution, and  $\eta$  follows an unknown distribution with a zero mean across individuals. Conditional on  $\eta$ , we integrate over the GEV distribution to obtain conditional choice probabilities as a general nested logit model:

$$\Pr(V_{ni} > V_{lm}, \forall m \neq i, \forall n, l | \eta) = \frac{\exp(V_i/\lambda_n) \left( \sum_{j \in B_k} \exp(V_j/\lambda_n) \right)^{\lambda_n - 1}}{\sum_{l=1}^K \left( \sum_{j \in B_l} \exp(V_j/\lambda_l) \right)^{\lambda_l}} \quad (10)$$

where  $n$  and  $l$  represent nests,  $i$  is an alternative within nest  $n$ ,  $m$  is an alternative within nest  $l$ ,  $K$  is the total number of nests, and  $B_l$  ( $l = 1, \dots, K$ ) represents a nested subset of alternatives. Our nesting structure is illustrated in Fig. 1.

We also integrate over the distribution of  $\eta$  to obtain unconditional probabilities. The literature offers no guidance on the distribution of the  $\eta$ .<sup>18</sup> To reduce the numerical difficulty in estimation, we let  $\eta$  be uniformly distributed in the interval  $[-\xi, \xi]$ . We search for the  $\xi$  that yields a likelihood function with the largest value.<sup>19</sup>

As pointed out by [Dubin and McFadden \(1984\)](#), the random error  $\eta$  does not have a zero mean conditional on each chosen bundle, due to the endogeneity of bundle choice. This can be seen clearly if we rewrite equations (8a–8d) into a more convenient form for estimation (using just equation 8a, as an example):

$$\begin{aligned} \ln(VMT_{i1}) = & \sum_j \alpha_{V1}^j d_{ij} + \sum_j \alpha_{p1}^j p_{j1} d_{ij} + \sum_j \alpha_{p2}^j p_{j2} d_{ij} - \alpha_{q1} \sum_j q_{j1} d_{ij} \\ & - \alpha_{q2} \sum_j q_{j2} d_{ij} - \beta y + \beta_1 \sum_j k_j d_{ij} + x' \gamma + \eta \end{aligned} \quad (11a)$$

where  $d_{ij}$  is a choice indicator variable equal to one when  $i = j$ , and where equations (11b–d) (not shown here) are analogous. The random error  $\eta$  is correlated with the choice indicators  $d_{ij}$ . [Dubin and McFadden \(1984\)](#) suggest sequential estimation to solve this endogeneity problem (a procedure later adopted by [Goldberg \(1998\)](#) and [West \(2004\)](#)). First, the discrete choice model is estimated and the predicted probabilities are calculated. They then suggest three alternative methods that yield consistent estimates of parameters for continuous demands: the instrumental variable method (IV), the reduced form method (RF), and the conditional expectation correction method (CE). They derive the correction terms in terms of probabilities for the CE method based on the assumption of an *i.i.d.* extreme value distribution of  $\varepsilon_i$ . However, since we assume a GEV distribution of  $\varepsilon_i$ , these correction terms cannot be used in our model. We want a method that can be used both for sequential estimation and for our simultaneous estimation, in order to compare them, and so we employ the RF method. Taking expectation of (11a) over  $\eta$ , we have:

$$\begin{aligned} \ln(VMT_{n1}) = & \sum_j \alpha_{V1}^j S_{nj} + \sum_j \alpha_{p1}^j p_{j1} S_{nj} + \sum_j \alpha_{p2}^j p_{j2} S_{nj} - \alpha_{q1} \sum_j q_{j1} S_{nj} \\ & - \alpha_{q2} \sum_j q_{j2} S_{nj} - \beta y + \beta_1 \sum_j k_j S_{nj} + x' \gamma + u_{n1}, \end{aligned} \quad (12a)$$

where  $S_{nj}$  is the probability of individual  $n$  choosing vehicle bundle  $j$  from (9),  $u_{n1}$  is an additional error to represent the difference between observed  $VMT$  and predicted  $VMT$ , and where (12b–d) are analogous (not shown here). The sequential RF method applies least squares to (12a–d), except that the shares  $S_{nj}$  are replaced by estimated shares  $\hat{S}_{nj}$  from the discrete choice model. In contrast, we estimate (9) and (12a–d) simultaneously.

<sup>18</sup> [Dubin and McFadden \(1984\)](#) assume  $\eta$  has a particular form of mean and variance, in order to derive an explicit conditional expectation.

<sup>19</sup> This search yields  $\xi$  equal to 0.65. Since the estimation of the logit model requires integration over the individual heterogeneity term  $\eta$ , our model is a mixed logit model ([McFadden and Train 2000](#)).

Since the same parameters appear in both discrete and continuous choice functions, we propose a joint estimation method to capture this simultaneity. In particular, we obtain a set of parameters that maximize the following objective function:

$$F(\Theta | y, p_1, p_2, q_1, q_2, k, x) = - \sum_n (\ln(VMT_1) - f_1)^2 - \sum_n (\ln(VMT_2) - f_2)^2 \\ - \sum_n (\ln(Wear_1) - g_1)^2 - \sum_n (\ln(Wear_2) - g_2)^2 + \sum_n \ln L, \quad (13)$$

where  $f_1$ ,  $f_2$ ,  $g_1$ , and  $g_2$  represent the right hand sides (without the random error  $u_{n1}$ ) of the four Eqs. (12a-d),  $\ln L$  is the log likelihood function of the nested logit, and  $\Theta$  represents the set of parameters to be estimated by maximizing Eq. (13).

As is consistent with [Dubin and McFadden \(1984\)](#) and other papers in this literature, the maintained hypotheses are that the utility functional form is correct and that consumers maximize it. Under these hypotheses, our procedure produces consistent estimates of parameters. The reasoning is as follows: if the components of (13) were maximized separately, and if some single set of parameters were the solution to all those separate maximizations, then this set of parameters would also maximize the combined objective function. To compare the results, we estimate our model by both the sequential method and the simultaneous estimation method.

#### 1.4 Elasticities

Once we obtain the parameter estimates, we are ready to calculate elasticities. To see the marginal effects of prices on indirect utility, and therefore on bundle choice, we use Eq. (7) to obtain explicit formulas for those derivatives. First, define  $\exp(\cdot) \equiv \exp(\alpha_{p1}^i p_{i1} + \alpha_{p2}^i p_{i2} - \alpha_{q1} q_{i1} - \alpha_{q2} q_{i2})$ . Then:

$$\frac{\partial V_i}{\partial p_{i1}} = -\exp(\cdot), \quad \frac{\partial V_i}{\partial p_{i2}} = -\frac{\alpha_{p2}^i}{\alpha_{p1}^i} \exp(\cdot) \quad (14a)$$

$$\frac{\partial V_i}{\partial q_{i1}} = \frac{\alpha_{q1}}{\alpha_{p1}^i} \exp(\cdot), \quad \frac{\partial V_i}{\partial q_{i2}} = \frac{\alpha_{q2}}{\alpha_{p1}^i} \exp(\cdot) \quad (14b)$$

and the marginal effects of income or capital cost on utility take similar forms:

$$\frac{\partial V_i}{\partial y} = \exp(-\alpha_0^i + \beta y - \beta_1 k_i - x' \gamma - \eta) \quad (15a)$$

$$\frac{\partial V_i}{\partial k_i} = -\frac{\beta_1}{\beta} \exp(-\alpha_0^i + \beta y - \beta_1 k_i - x' \gamma - \eta) \quad (15b)$$

Then we derive the elasticity of choice  $i$  with respect to a change in variable  $z_j$  (where  $z_j$  may be any of the price variables, income  $y$ , or capital cost  $k_j$ ):

$$\frac{\partial S_i}{\partial z_j} \cdot \frac{z_j}{S_i} = \frac{\partial S_i}{\partial V_j} \cdot \frac{\partial V_j}{\partial z_j} \cdot \frac{z_j}{S_i} \quad (16)$$

Since these formulas involve the unconditional probability of vehicle bundle  $i$ , calculating each bundle elasticity requires integration over  $\eta$ . In contrast, calculations of *VMT* elasticities do not involve integration over  $\eta$ . For bundle  $i$  ( $i = 1, \dots, 5$ ), the own- and cross-price elasticities of *VMT* demand are calculated by:

$$\begin{aligned} e_{V1p1}^i &= \frac{\partial \ln(VMT_{i1})}{\partial \ln p_{i1}} = \alpha_{p1}^i p_{i1} = e_{V2p1}^i, \\ e_{V2p2}^i &= \frac{\partial \ln(VMT_{i2})}{\partial \ln p_{i2}} = \alpha_{p2}^i p_{i2} = e_{V1p2}^i \end{aligned} \quad (17)$$

The elasticities of demand for *Wear* with respect to its price have a similar form:

$$\begin{aligned} e_{W1q1}^i &= \frac{\partial \ln(Wear_{i1})}{\partial \ln q_{i1}} = -\alpha_{q1}^i q_{i1} = e_{W2q1}^i, \\ e_{W2q2}^i &= \frac{\partial \ln(Wear_{i2})}{\partial \ln q_{i2}} = -\alpha_{q2}^i q_{i2} = e_{W1q2}^i \end{aligned} \quad (18)$$

We can also calculate the income elasticity, given by:

$$e_{Vy}^i = \frac{\partial \ln(VMT_{i1})}{\partial \ln y} = \frac{\partial \ln(VMT_{i2})}{\partial \ln y} = -\beta y \quad (19)$$

and the total capital cost elasticity, given by:

$$e_{Vk}^i = \frac{\partial \ln(VMT_{i1})}{\partial \ln k_i} = \frac{\partial \ln(VMT_{i2})}{\partial \ln k_i} = \beta_1 k_i. \quad (20)$$

In Eqs. (16–20), elasticities are typically evaluated at each bundle's mean values of  $y$  and  $k$ , the bundle average of gas prices per mile ( $p_1$  and  $p_2$ ) and the bundle average of reimbursement prices ( $q_1$  and  $q_2$ ).

## 2 Data and summary statistics

In order to analyze household choice of vehicles, miles driven, and vehicle *Wear*, we need micro-data on household characteristics, household income or expenditures, and detailed information about household-owned vehicles such as the number of vehicles, miles driven in each, and vehicle characteristics (including miles per gallon, MPG, and emissions per mile, EPM). No single data set contains all such information.

The Consumer Expenditure Survey (CEX) provides data on household income, characteristics, and household-owned vehicles.<sup>20</sup> For each household, we aggregate

<sup>20</sup> The CEX data are collected by the Bureau of Labor Statistics of the U.S. Department of Labor through quarterly interviews of selected households throughout the U.S. Each household is interviewed over five

expenditures over four quarters, taking demographic data and vehicle information from their last quarter in the survey. We use the CEX from 1996–2000, supplemented with the corresponding OVB file (Owned Vehicles Part B Detailed questions). This OVB file includes data on each vehicle type, make, year, number of cylinders, purchase expenses and financing, time since purchase, mileage, gasoline expenditure, and other information. We keep only households that satisfy several criteria. First, expenditures must be reported consecutively for four quarters of 1996–2000. Second, the household must possess the same number of vehicles during these four quarters. Third, we remove households that own more than two vehicles.<sup>21</sup> We also remove households that have vehicles other than automobiles or SUV's (defined to include light trucks or vans). Finally, we are left with 9,027 households, of which 2,077 own no vehicles, 4,211 own one vehicle, and 2,739 own two vehicles. We use yearly total expenditure as a proxy for yearly income of each household. Table 7 in Appendix defines all the variables used in estimations.

Summary statistics are shown in Table 1 for major household characteristics by vehicle bundle. This table shows significant variations in household characteristics across the number of vehicles and bundles. For example, larger households especially with more kids have more vehicles and prefer SUVs. Wealthier households (as measured by total yearly expenditures) possess more vehicles. Households with more workers or income earners have more vehicles. Households with male heads are inclined to have SUVs.

Next, fuel price data are obtained from the ACCRA cost-of-living index for 1996–2000. This index compiles quarterly data for approximately 300 cities in the United States. It also lists average gasoline price for each city for each survey quarter. Since the CEX reports region and state of residence instead of city for each household, we average the city gas prices to obtain a state price for each calendar quarter. For those states reported in the CEX, but not reported in the ACCRA index, we use the average region price as a substitute. Then we assign a gas price to each CEX household based on the state of residence, CEX quarter, and year.

Some of the variables in our model require calculations or additional sources of data. We now describe these extra calculations.

- (1) *Wear*: The vehicle's age is derived by taking the year of the survey minus the year the vehicle was made. We then assume 20 % annual depreciation, and calculate *Wear* as the percentage of the vehicle's value that has wasted away (given all the vehicle characteristics unchanged except vehicle age). *Wear* ranges from zero for a new car, to  $Wear = 1$  for a very old car. Specifically,  $Wear = 1 - (1 - 0.2)^{age}$ .
- (2) *Capital value of the vehicle*: The vehicle's year of purchase and reported purchase price (*pp*) are available in the OVB file, but we want an estimate of current market

---

Footnote 20 continued

consecutive quarters. Each quarter, 20 % of households complete their last interview and are replaced by new households. For CEX data, see <http://elsa.berkeley.edu> or <http://www.icpsr.umich.edu/>.

<sup>21</sup> In the CEX of 1996–2000, 18.4 % of households own more than two vehicles. Some of these households may be wealthy or may include teenagers with their own vehicle, so selection bias is a potential problem. But a two-person household with a third car may not drive it very much. Also, some households may have a third vehicle for business, whereas our model of household choice assumes utility maximization.

**Table 1** Summary of household statistics by vehicle bundle

Characteristics	Number of Vehicles					
	1		2			0
	#1 Car	#2 SUV	#3 Car, Car	#4 Car, SUV	#5 SUV, SUV	#6 none
Number (#) of households (hh)	3,469	742	1,181	1,305	253	2,077
MPG of first vehicle	21.37	16.76	21.88	21.51	17.04	–
MPG of second vehicle	–	–	21.55	16.53	16.50	–
Average hh size	1.92	2.30	2.65	2.94	3.44	1.98
Percentage (%) of hh with kids	23.87	33.56	33.62	43.98	62.45	26.05
Average # of kids per hh	0.44	0.73	0.56	0.89	1.42	0.55
Average hh # over 15 years old	1.52	1.63	2.13	2.12	2.13	1.48
Average hh # of workers	0.85	1.08	1.43	1.49	1.58	0.70
% of hh with male heads	40.10	63.07	65.54	71.80	77.47	33.22
Average age of head	55.24	48.22	51.84	49.45	45.24	55.66
% of hh with white heads	82.07	87.60	83.32	89.04	92.89	67.89
% with head's education more than high school	52.15	52.29	66.05	57.01	57.31	34.33
% in area with population over 4 million	28.37	19.41	30.48	22.68	18.58	38.61
Total expenditures	22,754	24,574	35,472	33,812	34,246	17,795
Gasoline expenditures	648	920	1,103	1,279	1,398	–

value ( $cmv$ ). We calculate the number of “years since purchase” ( $y_{sp}$ ), and we subtract depreciation for each year, again using 20 % as the annual rate of depreciation. The formula is  $cmv = pp \times (1 - 0.2)^{y_{sp}}$ . We then estimate a simple hedonic price regression:

$$cmv = a_0 + a_1cyl + a_2im + b_0(1 - Wear) + b_1(Wear \times cyl) + b_2(Wear \times im) \quad (21)$$

where  $a_0$  through  $a_2$ , and  $b_0$  through  $b_2$  are parameters. The variable  $cyl$  denotes the number of cylinders, while  $im$  is a dummy variable indicating if the vehicle is imported.  $Wear$  is included in the regression to capture the effects of vehicle age on market value. Using a sub-sample of the CEX that has all necessary variables, we run separate regressions for cars and SUV's and report the results in Table 8 of the Appendix.<sup>22</sup> Then, for the value of each brand new “concept” vehicle (with  $Wear = 0$ ), we use:

$$\hat{k} = \hat{a}_0 + \hat{a}_1cyl + \hat{a}_2im + \hat{b}_0. \quad (22)$$

where  $\hat{a}_0$  through  $\hat{a}_2$  and  $\hat{b}_0$  are estimates of parameters in (21).

<sup>22</sup> Some estimates in Table 8 Appendix have large standard errors, implying large confidence intervals, but these are the best estimates possible, given data availability.

- (3) *The price of Wear*: First, we calculate the extra amount paid for a car with no wear on it ( $Wear = 0$ ) compared to a very old car with the same characteristics ( $Wear = 1$ ). From (21), that difference is  $(\hat{b}_0 - \hat{b}_{1cyl} - \hat{b}_{2im})$ . Then,  $q$  is the annual reimbursement price of *Wear*, that is, the amount saved during a year by an owner who accepts one whole unit of *Wear* (an old car instead of a new car). Since a very old car does not depreciate any further, the amount saved is the depreciation during the year from holding a new car. Again assuming 20% depreciation, we have:  $q = 0.2 (\hat{b}_0 - \hat{b}_{1cyl} - \hat{b}_{2im})$ .
- (4) *Fuel Efficiency*: The EPA reports miles per gallon (MPG) of new vehicles, but we need it for vehicles of all ages. The CEX does not contain this information, so we estimate MPG using data of the California Air Resources Board (California Air Resources Board 1997; California Air Resources Board 2000).<sup>23</sup> Their first sub-sample is “series 13”, from November 1995 to March 1997, in which the CARB tested a total of 345 cars, light-duty trucks, and medium-duty vans. The second sub-sample is “series 14”, from November 1997 to August 1999, which includes 332 vehicles (but which reports only 327 vehicles). In total, we use 672 vehicles. We regress MPG against vehicle characteristics in the CARB and then use those estimated coefficients to predict MPG for each vehicle in the CEX. The estimation results are shown in Table 2, where a 4-cylinder SUV is the omitted category. This table shows that fuel efficiency decreases with vehicle age and engine size, both for cars and for SUV’s. Given the same vehicle age and engine size, MPG is higher for cars than for SUV’s. This table will be important later for discussion of the effects of policies on gasoline use and emissions.
- (5) *Emissions per mile (EPM)*: For the same sample of 672 used vehicles, the CARB tests for several pollutants. Following Fullerton and West (2010), we weight each pollutant by estimates of its damages, with the highest weight on nitrous oxides ( $NO_x$ , 0.495), followed by hydrocarbons (HC, 0.405), and carbon monoxide (CO, 0.10). Results appear in Table 2. Cars pollute less than SUV’s because they were produced under stricter standards. Older vehicles pollute more, both because newer vintages faced stricter standards and because pollution control equipment deteriorates over time.<sup>24</sup>
- (6) *Vehicle Miles Traveled (VMT)*: The OVB file provides cumulative miles on each vehicle, but we need yearly miles driven. Odometer data are imperfect. Some later odometer readings are less than the earlier ones, and many readings are missing. For a one-car household, we take observed annual expenditure on gasoline, divide by the price per gallon to get number of gallons, and then multiply by MPG to get miles. For a two-vehicle household, we only know the total gasoline expenditure,

<sup>23</sup> For MPG of new cars, <http://www.fueleconomy.gov/feg/index.htm> is a website of the US Environmental Protection Agency (EPA) and the Department of Energy. The EPA also provides the historical fuel economy of new vehicles at <http://www.epa.gov/otaq/mpg.htm> or at <http://www.epa.gov/otaq/tcldata.htm>.

<sup>24</sup> For vehicles in our sample, the calculated *EPM* is 1.89 g/mile for the average car and 3.56 for the average SUV. It also increases to 6.94 g/mile for a very old vehicle (with  $Wear = 1$ ).

**Table 2** Estimation of miles per gallon (MPG) and emissions per mile (EPM)

Independent variable	Dependent variable			
	MPG		EPM	
	Coefficient	Standard error	Coefficient	Standard error
Constant	24.021	0.496	−0.597	0.663
cyl6	−4.395	0.483	1.103	0.645
cyl8	−7.948	0.581	3.548	0.777
Age	−0.419	0.049	0.285	0.065
Age <sup>2</sup>	0.006	0.002	0.003	0.002
Car	4.262	0.410	−0.589	0.548
cyl6 × car	−1.439	0.560	−0.661	0.749
cyl8 × car	−1.149	0.655	−2.819	0.875
R <sup>2</sup>	0.7598		0.4095	
F value	299.997		65.775	
# of obs.	672		672	

so we need to allocate it between the two vehicles. For this allocation we use the difference in odometer readings.<sup>25</sup>

- (7) *Vehicle bundles*: As listed in Table 1, vehicle choices are classified into six categories according to the number and type of vehicles. For bundle 4, with one car and one SUV, the car is always identified as the first vehicle. For bundles 3 and 5, the first vehicle is identified as the one with higher yearly *VMT*. If two vehicles have the same yearly *VMT*, the identification is random. If *VMT* is missing, then the vehicle with an earlier purchase year is taken as the first vehicle. If the purchase year and miles-driven are both missing, the identification is random.

### 3 Estimation results

The model described in Sect. 1 is estimated by both the sequential and the simultaneous estimation methods. The mean values of key variables are reported by bundle in Table 3. We average the values within each bundle for each bundle-specific variable except gas price per mile. Gas price per mile is calculated by dividing gas price per gallon by a bundle-specific MPG listed in Table 1. Thus, gas prices per mile vary both within and between bundles. The presence of collinearity between the fixed effects  $\alpha_0^i$  ( $i = 1, \dots, 6$ ) and the bundle-specific variables such as  $k_i$  ( $i = 1, \dots, 5$ ) forces us to normalize the fixed effect of bundle one ( $\alpha_0^1$ ) to zero. To facilitate the estimation, we

<sup>25</sup> If the difference in odometer readings is positive for both vehicles, then we divide it by MPG to obtain an estimate of each vehicle's gas consumption. Each gasoline amount divided by their sum gives *shares*, used to allocate the observed total gas consumption. Each vehicle's gallons divided by MPG yields *VMT*. If the difference in odometer readings is positive only for one vehicle, we use this figure as *VMT*<sub>1</sub> and calculate gasoline used in this vehicle. Then total gasoline minus gas used in this vehicle is residual gas, allocated to the other vehicle. Dividing this residual gas by MPG yields *VMT*<sub>2</sub>. If the difference in odometer readings is positive for neither vehicle, then we do imputations based on households with similar characteristics.

**Table 3** Mean values of key variables involved in estimation

Variable	Bundle					
	1 (Car)	2 (SUV)	3 (C, C)	4 (C, S)	5 (S, S)	6 (none)
% of households	38.43	8.22	13.08	14.46	2.80	23.01
$VMT_1$	11,799	12,977	15,283	10,513	16,151	–
$VMT_2$	–	–	5,554	10,771	5,358	–
Price of gas 1 ( $p_1$ )	0.058	0.074	0.056	0.057	0.072	–
Price of gas 2 ( $p_2$ )	–	–	0.057	0.075	0.075	–
Vintage1	8.62	8.24	7.63	7.89	6.87	–
Vintage2	–	–	9.02	8.50	8.78	–
$Wear_1$	0.76	0.73	0.72	0.73	0.68	–
$Wear_2$	–	–	0.77	0.73	0.75	–
Price of $Wear_1$ ( $q_1$ )	15,572	18,010	15,363	15,686	18,052	–
Price of $Wear_2$ ( $q_2$ )	–	–	15,301	18,133	18,105	–
Expenditure ( $y$ )	22,754	24,574	35,472	33,812	34,246	17,795
Capital cost ( $k$ )	17,224	20,187	34,157	37,684	40,551	–
Capital cost 1	17,224	20,187	17,125	17,337	20,232	–
Capital cost 2	–	–	17,032	20,348	20,319	–

also normalize  $y$  in units of 10,000 dollars,  $k_i$  in units of 1,000, and  $q_1$  and  $q_2$  in units of 100 dollars. Accordingly, we multiply  $Wear_1$  and  $Wear_2$  by 100 to keep the total amount of reimbursement unchanged in the budget constraint.

Notice that bundle 3 and bundle 5 each contains two vehicles of the same type, while bundle 4 consists of one car and one SUV. When the retail gas price increases, all gas prices per mile are affected in bundle-specific ways because MPG depends both on vehicle age and type (car or SUV). As revealed by Table 1, MPG is more type-specific than bundle-specific. Thus, we expect that the gas price parameters of car bundles 1 and 3 are quite close to one another, as are those of SUV bundles 2 and 5. For a household with one car and one SUV, however, we wish to allow more substitution. In our estimation, we assign one parameter  $\alpha_{C1}$  to the gas price of the only car in bundle 1 and first car in bundle 3 (and  $\alpha_{C2}$  to the second car). We assign one parameter  $\alpha_{S1}$  to the only SUV in bundle 2 and first SUV of bundle 5 (and  $\alpha_{S2}$  to the second SUV). Then we assign two gas price parameters to bundle 4:  $\alpha_{p1}^4 (= \alpha_{CAR}^4)$  for the car and  $\alpha_{p2}^4 (= \alpha_{SUV}^4)$  for the SUV.

Results from the sequential estimation are discussed first. We follow the procedure suggested by Dubin and McFadden (1984), but at the first stage we estimate a nested logit structure instead of a multinomial logit model. The traditional ML method is employed. The RF method is adopted at the second stage because the correction terms derived by Dubin and McFadden are inappropriate for the GEV error structure. In the second stage we estimate four continuous demand equations jointly (only two equations for the one-vehicle bundles), using an objective function similar to Eq. (13) except that the last term is removed. We constrain parameters to be constant across bundles except those for gas prices and constant terms. The estimation results are reported in the first two columns of Table 4, under “sequential estimation”.

For the discrete choice model in the first column of Table 4, the estimates of  $\alpha_{C1}$  and  $\alpha_{S1}$  are significant at the 1% level, while those of  $\alpha_{C2}$  and  $\alpha_{S2}$  are not statistically

**Table 4** Estimation results

Parameters	Sequential estimation		Simultaneous estimation
	Nested logit	Continuous demands	
$p_{11}, p_{31}(\alpha_{C1})$	-0.246** (0.025)	-0.460** (0.070)	-0.433** (0.073)
$p_{32}(\alpha_{C2})$	-0.045 (0.033)	-0.238* (0.143)	-0.045** (0.008)
$p_{21}, p_{51}(\alpha_{S1})$	-0.237** (0.028)	-0.927** (0.054)	-0.526** (0.105)
$p_{52}(\alpha_{S2})$	-0.011 (0.049)	-0.453 (0.380)	-0.013 (0.080)
$p_{41}(\alpha_{CAR}^4)$	-0.240** (0.024)	-0.374** (0.143)	-0.399** (0.062)
$p_{42}(\alpha_{SUV}^4)$	-0.084** (0.022)	-1.331 (1.582)	-0.662** (0.103)
$q_1(\alpha_{q1})$	-0.012** (0.003)	-0.370E-03 (0.002)	-0.004** (0.001)
$q_2(\alpha_{q2})$	0.010** (0.001)	-0.010** (0.002)	-0.219E-36 (0.936E-36)
$y(\beta)$	-1.408** (0.086)	1.134** (0.134E-03)	-0.420** (0.001)
$k(\beta_1)$	-0.671** (0.108)	-0.456** (0.034)	-0.405** (0.023)
Choice specific			
Constant 2 ( $\alpha_0^2$ )	-1.403** (0.278)		0.645** (0.035)
Constant 3 ( $\alpha_0^3$ )	4.219** (0.516)		1.860** (0.031)
Constant 4 ( $\alpha_0^4$ )	5.057** (0.650)		2.063** (0.051)
Constant 5 ( $\alpha_0^5$ )	2.401** (0.685)		2.320** (0.062)
Constant 6 ( $\alpha_0^6$ )	-2.045** (0.383)		-0.948** (0.132)
Demand-specific			
Constant 1 ( $\alpha_{V1}$ )		9.578** (0.179)	0.302** (0.087)
Constant 2 ( $\alpha_{V2}$ )		7.361** (0.187)	0.805** (0.088)
Constant 3 ( $\alpha_{W1}$ )		9.346* (5.007)	2.580** (0.298)
Constant 4 ( $\alpha_{W2}$ )		5.147** (0.176)	5.114** (1.259)
Famsize	0.332 (0.542)	0.072** (0.002)	0.058** (0.001)
Earnr	0.270** (0.067)	0.067** (0.001)	0.032** (0.183E-03)
Kids	0.510 (0.527)	0.081** (0.002)	-0.031** (0.001)
Drivers	0.190 (0.535)	0.060** (0.001)	-0.041** (0.001)
Metro	-0.552** (0.123)	-0.012** (0.002)	0.012** (0.474E-03)
Pop4	-0.340** (0.085)	-0.013** (0.001)	0.012** (0.290E-03)
Urban	-0.441** (0.161)	-0.058** (0.002)	0.105** (0.001)
Age	0.046** (0.003)	-0.007** (0.290E-04)	0.004** (0.128E-04)
White	0.056 (0.091)	0.136** (0.001)	0.097** (0.386E-03)
Male	0.057 (0.085)	0.109** (0.001)	0.004** (0.240E-03)
Educ	0.020 (0.072)	0.058** (0.001)	0.036** (0.263E-03)
Northwest	0.244 (0.179)	0.042** (0.001)	0.046** (0.386E-03)
Midwest	0.401** (0.173)	0.064** (0.001)	0.059** (0.380E-03)
South	-0.726** (0.121)	-0.150** (0.001)	0.072** (0.374E-03)
$\lambda_1$	0.814** (0.053)		0.138** (0.006)
$\lambda_2$	0.066** (0.003)		0.103** (0.005)
Log Likelihood	-28,917.8	-786,857	-0.310E+07

\* 0.10 significance level; \*\* 0.05 significance level

significant. The estimates of  $\alpha_{p1}^4 (= \alpha_{CAR}^4)$  and  $\alpha_{p2}^4 (= \alpha_{SU}^4)$  are both significant at the 0.01 level. All of them are negative as expected. The *Wear* coefficients  $\alpha_{q1}$  and  $\alpha_{q2}$  are also different from zero at the 0.01 level. The parameter  $\lambda_n (n = 1, 2)$  measures the degree of independence of the errors of alternatives in nest  $n$ . In our model, the estimates of  $\lambda_1$  and  $\lambda_2$  are 0.814 and 0.066, respectively, both significant at the 0.01 level.<sup>26</sup>

Since all the estimates of  $\alpha_{p1}$  and  $\alpha_{p2}$  are negative, Eq. (14a, 14b) indicate that the marginal effects of gas prices per mile are negative. As consistent with expectation, an increase in gas price reduces household utility. Since the coefficient on the reimbursement price  $q_1$  is negative, the marginal effect on utility is positive as expected. A higher reimbursement price means more money back to the household for accepting a given vehicle age or level of *Wear*. However, the coefficient on  $q_2$  has unexpected sign. Since estimates of  $\beta$  and  $\beta_1$  are both negative and significant, Eq. (15a, 15b) indicate that the marginal effect of capital cost is negative while that of income is positive.

We then use those discrete choices from the first column to estimate the continuous demands shown in the second column. A glance down the second column indicates that most of estimated coefficients are quite different from the corresponding estimates in the first column. Yet the parameters in the second column are the same parameters as in the first column, even from the same model, as the continuous demands are supposed to be consistent with a particular indirect utility function. For example, the estimated coefficient on income is  $-1.408$  in the first column and  $+1.134$  in the second column. Both have small errors, and so they are significantly different from each other, even though they are the same parameter of the same model. Many price coefficients also differ significantly in magnitude (and the two estimates of  $\alpha_{q2}$  differ in sign).

Next, the model is estimated by the simultaneous estimation procedure proposed in Section I.C. The point of this procedure is to capture household-specific heterogeneity in both discrete and continuous choices. The two types of choices are connected by the same parameters and the same random error term  $\eta$  appearing in both.<sup>27</sup> In contrast, in the sequential procedure, the bundle choice affects continuous demands (and not vice versa). The simultaneous estimates are reported in the last column of Table 4.

All ten estimates of coefficients on key variables have the expected signs, and all but two are significantly different from zero. Yet, for many parameters, the estimate differs from *both* estimates obtained by sequential estimation. For example, the capital cost coefficient ( $\beta_1$ ) from the simultaneous model ( $-0.405$ ) is smaller in magnitude than either that of the logit model ( $-0.671$ ) or the continuous demand model ( $-0.456$ ). The estimates of coefficients on demographic variables vary with the estimation method, not only in magnitude but also in sign. For most price variables, however, the estimate from the simultaneous model is *between* the two estimates from sequential estimation, which suggests that the simultaneous model might provide more “reasonable” coeffi-

<sup>26</sup> If  $\lambda_n \forall n$  are within the range of zero to one, then “the model is consistent with utility maximization for all possible values of the explanatory variables” (Train 2003, p.85). Since our  $\lambda$  are significantly less than one, the errors within each nest are correlated, evidence in favor of nesting rather than MNL.

<sup>27</sup> The standard deviation for  $x'\gamma$  is about 0.086 within a bundle, and for  $\beta y$  is about 0.78 within a bundle, so the finding that  $\eta$  has a range ( $-0.65, 0.65$ ) reflects a significant amount of individual heterogeneity. Therefore, introducing individual heterogeneity is expected to make a difference in parameter estimates.

**Table 5** Elasticities of discrete choices for each variable

Variable	Bundle					
	#1 Car	#2 SUV	#3 Car, Car	#4 Car, SUV	#5 SUV, SUV	#6 none
Sequential <sup>a</sup>						
$p$	0.015	-0.106	0.006	-0.177E-03	0.034	-
$q$	-0.207	3.618	-0.116	-0.033	-6.077	-
$q_{car}$	1.530	-6.318	0.139	0.127	-3.470	-
$q_{suv}$	-1.737	9.937	-0.255	-0.160	-2.603	-
$y$	-0.106	0.591	-0.042	-0.006	-0.011	-0.006
$k$	0.086	-0.427	0.061	0.008	-0.303	-
$k_{car}$	-0.008	0.127	0.056	-0.944	4.336	-
$k_{suv}$	0.110	-0.413	0.134	-1.099	4.703	-
Simultaneous <sup>b</sup>						
$p$	0.009	-0.073	0.695	-0.793	0.020	-
$q$	0.025	0.193	0.066	0.283	-0.001	-
$q_{car}$	0.177	-0.966	0.151	0.352	-0.147	-
$q_{suv}$	-0.153	1.159	-0.085	-0.069	0.146	-
$y$	0.341	-1.203	-0.818	0.634	0.010	-0.074
$k$	-0.321	0.390	1.655	-6.319	-0.377	-
$k_{car}$	-1.229	7.315	-13.021	7.345	1.263	-
$k_{suv}$	0.908	-6.925	14.676	-13.665	-1.640	-

<sup>a</sup> Calculation based on estimates in column 1 of Table 4<sup>b</sup> Calculation based on estimates in column 3 of Table 4

cients. These coefficients cannot really be compared directly, however, and so we turn to elasticities.

#### 4 Elasticity comparisons

Bundle choice elasticities are presented in Table 5. Each entry in the table is not an elasticity with respect to each price in the model, as it might be difficult to interpret an elasticity such as the change in the probability of holding bundle 3 (two cars) for a change in the price  $p_1$  for gas in the first car only. Instead, we calculate the effect on all choices for a change in the price of gasoline. Also, given vehicle age, a higher reimbursement price  $q$  for *Wear* of a particular bundle means more money back to the household and thus higher probability of choosing that bundle. Again, however, it is difficult to interpret a change in the price  $q_1$  for the first car with no change in  $q_2$  for the household's second car. Instead, we show effects of a change in  $q$  for all vehicles (or for all of one type). Rather than raising  $q$ , however, policymakers can reduce emissions by taxing old vehicles or subsidizing new vehicles (*i.e.* reducing  $q$ ). Table 2 above shows that emissions per mile (EPM) are higher for SUV's than for cars, and rise with either type of vehicle's age.

The upper panel of Table 5 shows elasticities from the sequentially-estimated model, where gas prices have little effect on any bundle choice, and no effect on bundles for two cars or one car and one SUV. Also, a tax on age that reduces  $q$  has large effects that are sometimes the wrong sign (encouraging the purchase of two SUV's).

We therefore focus more on simultaneously estimated elasticities. In the lower part of Table 5, the first row shows that a 1 % increase in the price of gas would decrease most the probability of holding bundle 4 with a car and an SUV (by 0.793 %) while increasing the share holding bundle 3 with two cars (by 0.695 %). In other words, these households sell the SUV for a second car instead. This change is driven by the high price of driving an SUV with low fuel efficiency.<sup>28</sup>

The second row of the lower panel shows that a 1 % tax on *Wear* (lower  $q$  for all vehicles) would decrease the probabilities of holding all bundles except bundle 5 (SUV, SUV). In the next row, a tax on the age of cars would decrease the reimbursement for wear on cars,  $q_{car}$ , and switch households out of cars and into bundle 2 with an SUV and bundle 5 with two SUV's. Conversely, a tax on the age of SUV's that lowers  $q_{suv}$  would induce a switch out of bundles with just SUV's, and into bundles with cars. This tax might actually cut emissions in two ways: by inducing a switch from SUV's to cars (Table 5), and by inducing a switch from older SUV's to newer SUV's (Table 6).

The choice elasticities with respect to  $y$  indicate that households with more income switch from holding no car (bundle 6) to one car (bundle 1), and those with a single SUV (bundle 2) seem to add a car (bundle 4). Additional income reduces the share with two cars (bundle 3). These results are inconsistent with the discrete-choice model in the upper panel, where the only bundle with a positive income elasticity is bundle 2 with one SUV.

We next look at an increase in capital cost in the lower panel of Table 5. Since this change effectively reduces available income, we see that each capital cost elasticity has the opposite sign as that bundle's income elasticity. With higher capital costs, households seem to shift primarily out of two-vehicle bundles with at least one SUV (4 and 5) into bundles with two cars (bundle 3) or only one SUV (bundle 2). While it does not make sense to increase the capital cost only for the first car of a two-car household, it might make sense to increase the capital cost only of cars relative to SUV's or vice versa (to represent a vehicle-type tax). The next row of Table 5 shows that if the increase in capital cost pertains only to cars, then it decreases the shares of the two bundles that have only cars. If it pertains only to SUV's, however, then it has large effects that decrease the shares of all three bundles with SUV's. Such a policy could clearly reduce emissions (given the EPM in Table 2). The 1 % higher cost of an SUV means 13.7% less of bundle 4, which seems too large, but it means that the share falls two percentage points (from 14.5 % of all households in Table 3 to 12.5 % of all households).

Table 6 shows demand elasticities for continuous choices like *Wear* and vehicle miles traveled (VMT). These are "short run" elasticities, in the sense that car choices

<sup>28</sup> This reasoning is confirmed by the choice elasticities with respect to  $p_1$  and  $p_2$  separately. For bundle 4, a 1 % higher price per mile in the car reduces the probability of choosing that bundle by 0.37 %, while a 1 % higher price per mile in the SUV ( $p_2$ ) reduces the probability of choosing that bundle by 0.81 %. Thus, the gas consumption of the SUV has twice as much impact as that of the car.

**Table 6** Short-run elasticities of continuous demands

Variable	Bundle					Total
	#1 Car	#2 SUV	#3 Car, Car	#4 Car, SUV	#5 SUV, SUV	Emissions <sup>c</sup>
Sequential <sup>a</sup>						
$p$	−0.026	−0.066	−0.038	−0.117	−0.098	−0.211
$q$	0.012	0.013	0.306	0.360	0.362	0.631
$q_{car}$	0.012	−	0.306	0.012	−	0.368
$q_{suv}$	−	0.013	−	0.349	0.362	0.263
$y$	−2.581	−2.788	−4.024	−3.836	−3.885	−11.472
$k$	−1.570	−1.840	−3.113	−3.434	−3.695	−8.746
Simultaneous <sup>b</sup>						
$p$	−0.024	−0.037	−0.026	−0.070	−0.038	−0.136
$q$	0.122	0.141	0.120	0.123	0.141	0.434
$q_{car}$	0.122	−	0.120	0.123	−	0.293
$q_{suv}$	−	0.141	−	7.933E−36	0.141	0.141
$y$	0.956	1.032	1.490	1.420	1.438	4.246
$k$	−1.397	−1.637	−2.770	−3.056	−3.288	−7.783

Each entry is the elasticity of  $VMT$  or  $Wear$ , in the first or second vehicle, with respect to each variable

<sup>a</sup> Calculation based on estimates in column 2 of Table 4

<sup>b</sup> Calculation based on estimates in column 3 of Table 4

<sup>c</sup> The last column is the percent change in total emissions,  $E = \sum EPM \times \text{miles}$ , adding over all vehicles in all bundles, for a one percent change in each variable

are fixed, and only continuous choices like driving distances may change (Goldberg 1998).<sup>29</sup> In the top panel, the sequential model uses predictions of discrete choices from Table 5 to estimate continuous demands in Table 6. Most of the results from the sequential procedure have the correct signs, except in the fifth row where a small increase in income ( $y$ ) reduces the demand for miles driven (for every vehicle bundle combination).

Again, we focus primarily on simultaneously estimated elasticities in the bottom panel. In the first row, all elasticities for  $VMT_1$  with respect to gasoline price are negative, as expected, for all bundles. The next row of Table 6 shows the effects of a 1 % increase in the reimbursement price,  $q$ , on  $Wear$ . These elasticities are all positive, as expected: households choose older vehicles when they get higher reimbursement for holding an old vehicle. Conversely, a tax on vehicle age that reduces  $q$  by 10 % would reduce desired  $Wear$  by about 1.2–1.4 % (assuming the desired cars were available).<sup>30</sup> The table also shows similar effects of changing  $q$  just for cars, or just for SUV's.

<sup>29</sup> Panel data would be required to distinguish the effects of lags from contemporaneous price changes.

<sup>30</sup> In Table 3, the average  $Wear$  of 0.75 corresponds to 6.2 years of age, so a 1.2 % decrease in  $Wear$  means a decrease of about one month of age. In the sequential model, the same 10 % lower  $q$  affects desired age of one-vehicle bundles by one-tenth as much, and desired ages of two-vehicle bundles by three times as much.

Next, consider income and capital cost elasticities. Due to symmetric specification of demand functions, a 1 % change in  $y$  or  $k$  has the same effect on both *VMT* and *Wear* (whether for the first vehicle or the second). In the simultaneous model, income elasticities are positive as expected. One percent more income would increase driving distances by about 1–1.5 % for all bundles. The capital cost elasticities are negative as expected.<sup>31</sup>

Finally, the last column in Table 6 reports the percentage change in total emissions when each variable increases by 1 %. In the simultaneous model, for example, a 1 % increase in all gasoline prices would reduce total emissions by 0.136 %, while a tax on age that reduces  $q$  by 1 % would reduce total emissions by 0.434 %.<sup>32</sup> The largest elasticities are from income and capital cost: 1 % higher income raises total emissions as expected, by 4.246 % (but in the sequential model would reduce emissions by 11.47 %) A 1 % increase in capital cost reduces total emissions by about 8 % in either model.

In the simultaneously estimated model, the coefficients are affected by all discrete and continuous choices. The model imposes more constraints on the estimates. Thus, if those constrained estimates are plugged into the likelihood function for either part of the sequential procedure, then the likelihood is not as high as for that *portion* of the sequential procedure. However, the sequentially estimated model yields two sets of estimates for the same parameters. The finding that these estimates are not consistent with each other raises questions about whether the behavioral model is correctly specified.

## 5 Conclusion

This paper focuses on incentive effects of price changes that might be associated with policies to reduce vehicle emissions. We provide a model of household behavior that incorporates both the discrete choice of vehicle type, with different fuel efficiencies and emission rates, and continuous demands for miles driven. Because emission rates depend directly on vehicle age, we also model vehicle age as a continuous choice. To model the effect of prices on the choice of vehicle age, we establish a choice of “concept vehicle” that is separate from the choice of “*Wear*”. Using hedonic price regressions, we quantify the price of *Wear*. Then, after the discrete choice among concept vehicles, both *VMT* and *Wear* become continuous variables that enter utility.

Yearly household data are obtained from the CEX of 1996–2000, supplemented with fuel efficiency estimates from the CARB, and gas prices from the ACCRA cost

<sup>31</sup> The specific form for utility in Eq. (4) means a specific form for demands in Eq. (5a, 5b), where  $\ln(VMT)$  and  $\ln(Wear)$  both depend on  $\alpha_p^i p_i - \alpha_q q_i$ . In other words, the parameter that determines the important effect of gas price on miles ( $\alpha_p^i$ ) also necessarily drives the less-important effect of the gas price on choice of *Wear*. Similarly, the own-price effect of  $q$  on *Wear* also drives the cross-price effect of  $q$  on *VMT*. We note this fact, but we do not mean to emphasize these cross-price elasticities.

<sup>32</sup> These are also short run elasticities, with no change in the number or type of vehicles. Notice that the percentage change in emissions from a change in  $p$  is more than twice the change in driving distance, because the higher  $p$  also reduces demand for *Wear* (which also reduces emissions). The change in  $q$  also affects both *VMT* and *Wear* in the same direction, enlarging the effect on emissions.

of living indexes. First, like many others, we follow the sequential procedure suggested by [Dubin and McFadden \(1984\)](#). This procedure generates two different sets of estimates for the same set of parameters, which we argue is inconsistent with maintained hypotheses about the utility function and utility maximization. We then propose and implement a simultaneous method for consistent estimation of both discrete and continuous choices in one step. Results from the simultaneous estimation differ significantly both in signs and magnitude from both sets of estimates obtained by sequential estimation.

In general, we find that short run price elasticities for continuous variables like VMT are smaller than long-run elasticities for discrete choices. Thus, for example, a tax on age of SUV's might cut emissions both by inducing a switch from SUV's to cars and by inducing a switch from older SUV's to newer SUV's. We emphasize, however, that more work is needed, perhaps using the simultaneous method proposed in this paper.

**Acknowledgments** We are grateful for financial support from Japan's Economic and Social Research Institute (ESRI), and for suggestions from Chunrong Ai, Frank Convery, Stephen Donald, Hilary Sigman, Dan Slesnick, Ken Train, Sarah West, and Cliff Winston. This paper is part of the NBER's research programs in Public Economics (PE) and Energy and Environmental Economics (EEE). Any opinions expressed are those of the authors and not those of the ESRI or the National Bureau of Economic Research.

## 6 Appendix

See Tables 7 and 8.

**Table 7** Variable definitions

Variable	Definition
$y$	Household's yearly expenditure
$k$	Total capital cost of a vehicle bundle
$p_1$	Gas price per mile of the first vehicle
$p_2$	Gas price per mile of the second vehicle
$q_1$	Unit price of <i>Wear</i> of the first vehicle
$q_2$	Unit price of <i>Wear</i> of the second vehicle
$VMT_1$	Miles driven in the first vehicle
$VMT_2$	Miles driven in the second vehicle
$Wear_1$	Continuous variable to measure the wear of the first vehicle
$Wear_2$	Continuous variable to measure the wear of the second vehicle
Famsize	Number of members in a household
Earnr	Number of income earners in a household
Kids	Number of children less than 18 in a household
Drivers	Number of household members 16 years old and over
Metro	=1 if household resides inside a Metropolitan Statistical Area (or else zero)
Pop4	=1 if the household lives in area with population > 4 million (or else zero)
Urban	=1 if the household lives in an urban area (or else zero)

**Table 7** continued

Variable	Definition
Age	Age of household head
White	=1 if the household head is white, and zero otherwise (or else zero)
Male	=1 if the head is male, zero otherwise (or else zero)
Educ	=1 if the head has education higher than high school (or else zero)
Northwest	=1 if in the Northwest (or else zero)
Midwest	=1 if in the Midwest (or else zero)
South	=1 if in the South (or else zero)
West	=1 if in the West (or else zero)

**Table 8** Hedonic price regressions

Dependent variable: <i>cmv</i>	Cars		SUVs	
	Coefficient	Standard error	Coefficient	Standard error
Constant ( $a_0$ )	1,444.64	1,806.08	-1,220.52	2,702.42
cyl ( $a_1$ )	3,150.55	288.44	1,993.56	411.23
Import ( $a_2$ )	2,371.11	894.32	1,417.36	1,584.27
1-Wear ( $b_0$ )	-2,179.03	3,272.66	8,973.32	4,996.71
Wear $\times$ cyl ( $b_1$ )	-3,184.92	546.49	-1,459.66	763.85
Wear $\times$ import ( $b_2$ )	-998.07	1,719.28	-658.35	2,800.80
R <sup>2</sup>	0.4880		0.5147	
F-value	150.004		106.900	
# of obs.	793		510	

## References

- Bento, A., Goulder, L., Jacobsen, M., & Von-Haefen, R. (2009). Distributional and efficiency impacts of increased U.S. gasoline taxes. *American Economic Review*, 99(3), 667–699.
- Bhat, C. R. (2005). A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research: Part B: Methodological*, 39(8), 679–707.
- Bhat, C. R., Sudeshna, S., & Naveen, E. (January 2009). The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research: Part B: Methodological*, 43(1), 1–18.
- Brownstone, D., & Train, K. (1999). Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89, 109–129.
- Brownstone, D., Bunch, D. S., Golob, T. F., & Ren, W. (1996). A vehicle transactions choice model for use in forecasting demand for alternative-fuel vehicles. *Research in Transportation Economics*, 4, 87–129.
- California Air Resources Board. (1997). Test report of the light-duty vehicle surveillance program, Series 13, Project Number 2S95C1.
- California Air Resources Board. (March 2000). Report of the results of the vehicle surveillance program 14, Project Number 2S97C1.
- Dubin, Jeffrey, & McFadden, Daniel. (1984). An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, 52(2), 345–62.

- Eskeland, G., & Devarajan, S. (1996). *Taxing bads by taxing goods: Pollution control with presumptive charges*. Washington, DC: The World Bank.
- Fullerton D., & West S. (2010) Tax and Subsidy Combinations for the Control of Car Pollution. The BE Journal of Economic Analysis & Policy (Advances) 10:1
- Fullerton, D., & West, S. (2002). Can taxes on vehicles and on gasoline mimic an unavailable tax on emissions? *Journal of Environmental Economics and Management*, 43, 135–157.
- Goldberg, P. (1998). The regulation of fuel economy and the demand for 'light trucks'. *Journal of Industrial Economics*, 46(1), 1–33.
- Greene, D. L., & Hu, P. (1985). The influence of the price of gasoline on vehicle use in multi-vehicle households. *Transportation Research Record*, 988, 19–24.
- Hanemann, W. M. (1984). Discrete/continuous models of consumer demand. *Econometrica*, 52(3), 54–62.
- Harrington, W., & McConnell, V. (2003). Motor vehicles and the environment. In H. Folmer & T. Tietenberg (Eds.), *The international yearbook of environmental and resource economics 2003/2004*. Northampton, MA: Edward Elgar.
- Hausman, J. A. (1981). Exact consumer's surplus and deadweight loss. *American Economic Review*, 71(4), 662–676.
- Heeb, N., Forss, A.-M., Saxer, C., & Wilhelm, P. (2003). Methane, benzene and alkyl benzene cold start emission data of gasoline-driven passenger cars representing the vehicle technology of the last two decades. *Atmospheric Environment*, 37, 5185–5195.
- Hulten, C. R., & Wykoff, F. C. (1996). Issues in the measurement of economic depreciation. *Economic Inquiry*, 34, 10–23.
- Innes, R. (1996). Regulating automobile pollution under certainty, competition, and imperfect information. *Journal of Environmental Economics and Management*, 31, 219–239.
- Jorgenson, D. W. (1996). Empirical studies of depreciation. *Economic Inquiry*, 34, 24–42.
- Kohn, R. E. (1996). An additive tax and subsidy for controlling automobile pollution. *Applied Economics Letters*, 3, 459–462.
- Manning, F., & Clifford, W. (1985). A dynamic empirical analysis of household vehicle ownership and utilization. *Rand Journal of Economics*, 16(2), 215–236.
- McFadden Daniel. (1979) Quantitative Methods for Analyzing Travel Behavior of Individuals Some Recent Developments. In: Hensher David., Stopher P. (ed) Behavioral Travel Modeling. Croom Helm, London, 279–318.
- McFadden, D. (2001). Disaggregate behavioural travel demand's RUM side: A 30-year retrospective. In Hensher (Ed.), *Travel behavior research: The leading edge*. London: Pergamon.
- McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5), 447–470.
- Newell, R. G., & Stavins, R. N. (2003). Cost heterogeneity and the potential savings from market-based policies. *Journal of Regulatory Economics*, 23(1), 43–59.
- Parry, I. W. H., Walls, M., & Harrington, W. (2007). Automobile externalities and policies. *Journal of Economic Literature*, 45, 374–400.
- Pigou, A. C. (1920). *The economics of welfare*. London: MacMillan.
- Plaut, P. (1998). The comparison and ranking of policies for abating mobile-source emissions. *Transportation Research D*, 3, 193–205.
- Seigny, M. (1998). *Taxing automobile emissions for pollution control*. Cheltenham, UK: Edward Elgar.
- Sierra Research. (1994). "Analysis of the Effectiveness and Cost-Effectiveness of Remote Sensing Devices". Report SR94-05-05, prepared for the U.S. Environmental Protection Agency, Sacramento, CA: Sierra Research.
- Train, K. (1986). *Qualitative choice analysis: Theory, econometrics, and an application to automobile demand*. Cambridge, MA: MIT.
- Train, K. (2003). *Discrete choice methods with simulation*. Cambridge, UK: Cambridge University Press.
- Train, K., Davis, W. B., & Levine, M. D. (1997). Fees and rebates on new vehicles: Impacts on fuel efficiency, carbon dioxide emissions, and consumer surplus. *Transportation Research E*, 33, 1–13.
- U.S. Environmental Protection Agency. (1998). *Technical Methods for Analyzing Pricing Measures to Reduce Transportation Emissions*. DC: Washington.
- West, S. (2004). Distributional effects of alternative vehicle pollution control policies. *Journal of Public Economics*, 88(3–4), 735–757.