Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

Equilibrium Trade in Automobiles: Online Appendix

Kenneth Gillingham *
Yale University

Fedor Iskhakov †
Australian National University

Anders Munk-Nielsen ‡
University of Copenhagen

John Rust §
Georgetown University

Bertel Schjerning ¶
University of Copenhagen

February 2022

*Yale University, School of the Environment, Department of Economics, School of Management, kenneth.gillingham@yale.edu

†Australian National University, Research School of Economics, fedor.iskhakov@anu.edu.au

‡University of Copenhagen, Department of Economics and Centre for Computational Economics (CCE), amn@econ.ku.dk

§**Corresponding Author:** Georgetown University, Department of Economics, Georgetown University, jr1393@georgetown.edu

¶University of Copenhagen, Department of Economics and Centre for Computational Economics (CCE), bertel.schjerning@econ.ku.dk

A Proofs

Lemma L1 (Jacobian Matrix of the Smoothed Bellman Operator)**.** *Let EV be the unique fixed point of the smoothed Bellman operator Γ in (9), and let $\nabla_{EV}\Gamma(EV,P)$ be the Jacobian matrix of Γ with respect to EV. The following holds:*

$$\nabla_{EV}\Gamma(EV,P) = \beta\Omega(P)Q, \tag{A1}$$

where the matrices $\Omega(P)$ and Q are defined in equations (24) and (27), respectively. The norm of this matrix is $\|\nabla_{EV}\Gamma(EV,P)\| = \beta \in (0,1)$.

Proof. This Lemma can be proved via direct calculation, though the algebra involved is extremely tedious. We sketch a more conceptual proof of the result here. First, note that the operator Γ is a smooth nonlinear and recursively nested function of log-sum functions (11). The log-sum functions are in turn functions of the choice-specific value functions $v(i,a,j,d)$ given in equations (3), (6), and (8) of Section 3. These value functions, call them v, are in turn functions of the expected values EV a result we emphasize by writing $\Gamma(EV,P) = \Gamma(v(EP),P)$. Using the chain rule to compute the Jacobian matrix $\nabla_{EV}\Gamma(EV,P)$ with respect to the $J\bar{a}+1\times 1$ vector EV, we have

$$\nabla_{EV}\Gamma(EV,P) = \nabla_v\Gamma(v(EV),P)\nabla_{EV}v(EV). \tag{A2}$$

The Lemma follows by showing that $\nabla_v\Gamma(v(EV),P)$ equals $\Omega(P)$ and $\nabla_{EV}v(EV)$ equals βQ. The former result follows from the Williams-Daly-Zachary Theorem (see McFadden (1981)), and the fact that the Γ operator can be expressed as the expected maximum of the v functions with the additive GEV errors as shown via the representation of $\Gamma(EV,P)$ as the value functions V given in equations (2) and (5). The Williams-Daly-Zachary Theorem implies that the derivatives of the expected maximum of $v+\epsilon$ with respect to v equals the choice probabilities Π such as in equation (13). When the matrix of values v is arrayed in the same order that EV is arrayed (with first $J\bar{a}$ elements of EV equalling $EV(i,a)$ for $i=1,\dots,J$ and $a=1,\dots,\bar{a}$ and $J\bar{a}+1$-st element equal to $EV(\emptyset)$ and we account for the fact that some elements of EV appear in multiple different elements in any given row of the matrix v), it is not difficult to see that $\nabla_v\Gamma(v(EV),P) = Q(P)$, where the latter is a $(J\bar{a}+1)\times(J\bar{a}+1)$ Markov transition probability matrix given in equation (24). Further using the formulas for $v(EV)$ in equations (3), (6), and (8) of Section 3, it is not hard to see that $\nabla_{EV}v(EV) = \beta Q$, where Q is the accident/aging transition probability given in equation (27). Since Q and $\Omega(P)$ are both Markov transition probability matrices, so is their product, $M = \Omega(P)Q$. Recall the notion of a matrix norm, $\|M\| = \sup_{x\neq 0}\|Mx\|/\|x\|$ where $\|x\|$ is a norm of the vector x (e.g. Euclidean norm or sup-norm). When M is a transition probability matrix, it is easy to see that $\|Mx\| \leq \|x\|$, which implies that $\|M\| \leq 1$. Let e be a vector all of whose elements equal 1. Then $Me = e$ which implies $\|M\| \geq 1$, or $\|M\| = 1$. Also it is easy to see from the definition of a matrix-norm, $\|\beta M\| = \beta\|M\|$. It follows that $\|\beta\Omega(P)Q\| = \beta$. □

Lemma L2 (Differentiability)**.** *The unique fixed point EV of the smoothed Bellman operator in (9), the choice-specific value functions $v(\cdot)$ in (2), (5) and (7), the choice probabilities $\Pi(j,d,s|i,a)$ in (13), the trade transition probability matrix $\Omega(P)$ and excess demand function $ED(P)$ exist and are continuously differentiable functions of market prices P. The Jacobian matrix of the fixed point EV with respect to the market prices is given by*

$$\nabla_P EV(P) = -[I - \nabla_{EV}\Gamma(EV,P)]^{-1}\nabla_P\Gamma(EV,P), \tag{A3}$$

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022. "Equilibrium Trade in Automobiles." Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

where $\nabla_P \Gamma(EV, P)$ is the $J\bar{a}+1 \times J(\bar{a}-1)$ Jacobian matrix of Γ with respect to market prices P.

Proof. Existence of a unique fixed point $EV = \Gamma(EV, P)$ for any P follows because the operator Γ can be shown to be a *quasi-linear, monotone mapping* (see Rust, Traub and Wozniakowski (2002)) and thus is a contraction mapping with a unique fixed point EV. By Lemma L1, Γ is a continuously differentiable function of EV with gradient $\nabla_{EV}\Gamma(EV, P)$. By the Implicit Function Theorem we can express EV as a zero, $F(EV, P) = 0 = EV - \Gamma(EV, P)$ and the solution EV will be a continuously differentiable implicit function of P provided that $\nabla_{EV} F(EV, P)$ is invertible. However from Lemma L1 $\nabla_{EV} F(EV, P) = I - \beta\Omega(P)Q$, and this is invertible with the geometric series representation for its inverse

$$[I - \beta\Omega(P)Q]^{-1} = \sum_{t=0}^{\infty} \beta^t [\Omega(P)Q]^t. \tag{A4}$$

Then since $EV(P)$ is a continuously differentiable function of P, it is easy to see from the formulas for v, the conditional choice probabilities Π and the transition probability matrix $\Omega(P)$ are continuously differentiable since they are explicit smooth functions of $EV(P)$. The formula for $\nabla_P EV(P)$ in equation (A3) is a consequence of total differentiation of the identity $\Gamma(EV(P), P) = 0$ with respect to P and solving for $\nabla_P EV(P)$.

□

Lemma L3 (Gradient of invariant distribution)**.** *Consider a $n \times n$ Markov transition probability matrix $P(\theta)$ that depends on a parameter $\theta \in \mathbb{R}^k$ in a continuously differentiable fashion.*[1] *Let $h(\theta)$ be the unique invariant distribution of $P(\theta)$ satisfying the equation*

$$h(\theta) = h(\theta)P(\theta). \tag{A5}$$

Then the $n \times 1$ transpose of $h(\theta)$, $h(\theta)'$, is the unique solution to the expanded $(n+1) \times (n+1)$ linear system given by

$$\begin{bmatrix} I - P(\theta)' & e \\ e' & 1 \end{bmatrix} \begin{bmatrix} h(\theta)' \\ 1 \end{bmatrix} = \begin{bmatrix} e \\ 2 \end{bmatrix} \tag{A6}$$

where e is an $n \times 1$ vector all of whose elements equal 1, and I is a $n \times n$ identity matrix. Moreover, $h(\theta)$ is a continuously differentiable function of θ, and the simple expression for the Jacobian matrix $\nabla_\theta q(\theta)$ are readily available.

Proof. The stationarity condition (A5) can be recast as $h(\theta)$ being a left zero of the matrix $I - P(\theta)$. The usual application of the Implicit Function Theorem to $F(h, \theta) = 0$ where $F(h, \theta) = h(\theta)[I - P(\theta)]$, would provide the result. Unfortunately the prerequisite condition that $\nabla_h F(h, \theta)$ is non-singular at a zero of F fails as $\nabla_h F(h, \theta) = I - P(\theta)$, and this matrix is singular.

Let $A(\theta)$ be the $(n+1) \times (n+1)$ matrix on the left hand side of equation (A6). It is not hard to show that $A(\theta)$ is invertible and we can write

$$\begin{bmatrix} h(\theta)' \\ 1 \end{bmatrix} = \begin{bmatrix} I - P(\theta)' & e \\ e' & 1 \end{bmatrix}^{-1} \begin{bmatrix} e \\ 2 \end{bmatrix}. \tag{A7}$$

[1] Thus, we assume that the mapping $\nabla_\theta P(\theta)$ from R^k to R^{k*n*n} (where the latter can be interpreted as the space of k-tuples of $n \times n$ matrices) exists and is a continuous function of θ. To make things easier to understand, assume initially that $k = 1$ so we are considering $P(\theta)$ and $h(\theta)$ as functions of a single parameter θ. If θ has k components (i.e. $\theta \in R^k$) we simply "stack" the formulas we provide below in the univariate case into a k-tuple.

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

Then we have that $\nabla_\theta h(\theta)'$ is the upper left $n \times n$ submatrix of the product of $\nabla_\theta A^{-1}(\theta)$ times the vector $\begin{bmatrix} e \\ 2 \end{bmatrix}$. Further, the standard formula holds for the gradient of $A^{-1}(\theta)$ with respect to θ

$$\nabla_\theta A^{-1}(\theta) = -A^{-1}(\theta) \left[\nabla_\theta A(\theta)\right] A^{-1}(\theta), \tag{A8}$$

and we have

$$\begin{bmatrix} \nabla_\theta h(\theta)' \\ 0 \end{bmatrix} = \begin{bmatrix} I - P(\theta)' & e \\ e' & 1 \end{bmatrix}^{-1} \begin{bmatrix} -\nabla_\theta P(\theta)' & 0 \\ 0' & 0 \end{bmatrix} \begin{bmatrix} I - P(\theta)' & e \\ e' & 1 \end{bmatrix}^{-1} \begin{bmatrix} e \\ 2 \end{bmatrix}. \tag{A9}$$

□

Results of Lemma L3 are useful in both the differentiating the invariant ownership distribution with respect to the market prices for the equilibrium computations, and the structural parameters for the estimation of the model. In both cases, $P(\theta)$ is given by $\Omega_\tau(P)Q_\tau$ as per equation (34). Computing the required $\nabla_{P_{ia}}\Omega_\tau(P)Q_\tau$ for each price P_{ia} on the secondary market is straightforward. The same applies to all parameters of the utility function. Yet, if θ includes accident probabilities $\alpha(i,a)$ that enter in Q_τ, product rule has to be invoked leading to a slightly more involved expression $\nabla_{\alpha(i,a)}\Omega(\alpha)Q(\alpha) + \Omega(\alpha)\nabla_{\alpha(i,a)}Q(\alpha)$.

A.1 Proof of Theorem 1 (page 23)

Proof. When scale parameters of GEV distribution of random components ϵ are positive, $\sigma \geq \sigma_r \geq \sigma_j \geq \sigma_s > 0$, the choice probabilities are bounded away from zero for all choices and for any price vector P. Thus, the transition probability matrix $\Omega(P)Q$ is irreducible and aperiodic. Uniqueness of the stationary distribution q that satisfies $q = q\Omega(P)Q$ follows from the fundamental theorem of Markov chains.

To show continuous differentiability of the stationary distribution q given as an implicit function of P by $q = q\Omega(P)Q$, it would be enough to apply the Implicit Function Theorem, but unfortunately the prerequisite invertibility condition fails in our case. Indeed, q can be treated as a left zero of the matrix $I - \Omega(P)Q$, where I is the identity matrix of the appropriate size. In other words, $q(P)$ can be written as a zero the non-linear mapping $F(q,P) = q\big(I - \Omega(P)Q\big) = 0$. However when q is ergodic, Lemma L3 provides an explicit solution for $q(P)$ as the inverse of a bordered matrix for which $I - \Omega(P)Q$ is an upper left $(J\bar{a}+1) \times (J\bar{a}+1)$ submatrix. It also provides a formula for $\nabla_P q(P)$ in terms of gradients of the matrix $\Omega(P)$ with respect to P. This proves that $q(P)$ is uniquely defined and continuously differentiable function of P.

□

A.2 Proof of Theorem 2 (page 24)

Proof. The proof of existence of an equilibrium follows from Brouwer's Fixed Point Theorem by defining a mapping $\Psi(P) : R^{J(\bar{a}-1)} \to R^{J(\bar{a}-1)}$ where $\Psi(P) = P + ED(P)$, $J(\bar{a}-1)$ is the dimension of the price vector P and number of used cars traded in secondary markets with J types of cars sold in the primary market and $\bar{a}$ the oldest tradeable car age in each market, and $ED(P)$ is defined in (23). From Lemma L2 it follows that ED and thus Ψ is a continuous mapping from $R^{J(\bar{a}-1)} \to R^{J(\bar{a}-1)}$. Note also that for any P the components of $ED(P)$ lie in the interval $[-1, 1]$. Thus, when prices are sufficiently high, a vanishing number of consumers will wish to buy any new car but nearly all consumers will want to sell their cars, so $ED(P)$ will be close to a vector with all its components equal to -1. Similarly, for a sufficiently low set of prices (possibly negative), nearly all consumers will wish to buy used cars and very few will want to sell their vehicles at such low prices. So for such prices $ED(P)$ will be close to a vector with all of its components equal to $+1$. It follows that we can define a compact box B in $R^{J(\bar{a}-1)}$ where Ψ satisfies an "inward pointing" property on the boundaries of this box, so it follows that $\Psi : B \to B$. Since Ψ is a continuous mapping and B is a compact, convex set, the Brouwer fixed point theorem implies that a fixed point of Ψ exists, and it is clear that any such fixed point satisfies $ED(P) = 0$.

□

A.3 Proof of Theorem 3 (page 25)

Proof. The proof is based on the observation that the physical transition probability matrix Q in equation (27) is block diagonal, implying that the stationarity condition (29) in Theorem 2 can be written separately for each car make/model j as

$$\left(q \begin{bmatrix} \Delta_{1j}(P) \\ \vdots \\ \Delta_{Jj}(P) \\ \Delta_{\o j}(P) \end{bmatrix} + \left[q_{j1}, \ldots, q_{j\bar{a}}\right] \Lambda_j(P) \right) Q_j = \left[q_{j1}, \ldots, q_{j\bar{a}}\right] = q_j, \tag{A10}$$

where $\Delta_{ij}(P)$, $\Lambda_j(P)$, and q_j are defined in equations (25), (26) and (20) respectively. The matrix Q_j is stochastic, therefore we can effectively get rid of it by right-multiplying both side of the equation with the column of ones $e = [1, \ldots, 1]' \in \mathbb{R}^{\bar{a}}$. This is equivalent to taking the sum over all ages between 1 and $\bar{a}$. It also follows from the structure of matrices $\Delta_{ij}(P)$ that the first component in the LHS of the equation above is nothing but the vector of demand for j-type cars from age 1 to $\bar{a} - 1$ with the last element equal to the demand for the new cars $D_{j0}(P, q)$ given in equation (21). We then have

$$\Big(\Big[D_{j1}(P, q), \ldots, D_{j,\bar{a}-1}(P, q), D_{j0}(P, q) \Big] + q_j \Lambda_j(P) \Big) e = q_j e. \tag{A11}$$

Using the market clearing condition (30) in Theorem 2 we can replace the first $\bar{a}-1$ components in the vector of demands with the corresponding supply. To maintain the matrix notation we express supply given in equation (22) with the help of the keeping choice matrix $\Lambda_i(P)$ defined in equation (26), and the similarly constructed diagonal $\bar{a} \times \bar{a}$ matrix $\Lambda^s_j(P)$ of probabilities to scrap $\Pi(1_s | j, a, P)$. Naturally, $\Lambda^s_j(P)$ has to have 1 in the lower right corner corresponding to the choice probability of scrapping a car of the terminal age $\bar{a}$. The vector of supply for all ages between 1 and $\bar{a}$ is then given by $q_j\big(I - \Lambda_j(P)\big)\big(I - \Lambda^s_j(P)\big)$, with the appropriate zero as the last element. Continuing the derivation we have

$$\Big(q_j - q_j \big(I - \Lambda_j(P) \big) \Lambda^s_j(P) + \Big[0, \ldots, 0, D_{j0}(P, q) \Big] \Big) e = q_j e.$$

It then immediately follows that

$$q_j \big(I - \Lambda_j(P) \big) \Lambda^s_j(P) e = D_{j0}(P, q),$$

which is the matrix form of the stationary flow condition (32). □

A.4 Proof of Theorem 6 (page 40)

Proof. By the inversion theorem (Lemma 2 Arcidiacono and Miller, 2011) the differences between any two choice specific value functions given in Section 3 are identified. At this point we do not rely on the location normalization of the utilities, but assume that the scale of the extreme value shocks is fixed. There are many ways to show identification of the monetary part of the utility function, but one simple way is the following. All the derivations in the proof can be carried out for each consumer type, we drop the subscript τ for clarity.

Consider the value differences for the choices to purge an existing car given in equation (3) and the choice to remain in the no car state given in equation (8). We have

$$v(i, a, \emptyset, 1_s) - v(\emptyset, \emptyset) = \mu \underline{P}_i, \tag{A12}$$

and thus the marginal utility of money μ is point identified (up to a scale).

Next, consider the pairwise differences between the choice specific values $v(i, a, j, d, 1_s)$, $v(i, a, j, d, 0_s)$ and $v(i, a, \kappa)$ given in equation (8). Under the assumption [c] seller transaction costs $T_s(i, a) = 0$, and we have the following linear system of two equations with two unknowns

$$\begin{aligned} v(i, a, j, d, 0_s) - v(i, a, j, d, 1_s) &= -\mu[-P_{ia} + \underline{P}_i], \\ v(i, a, i, a, 0_s) - v(i, a, \kappa) &= -\mu[T_b(i, a)]. \end{aligned}$$

Given μ, this system yields identification of the prices and buyer-side transaction costs for all tradable cars $i \in \{1, \ldots, J\}$ and $a \in \{1, \ldots, \bar{a} - 1\}$. □

B Solving the Homogeneous Consumer Economy

The limiting case of our model when $\sigma \to 0$ constitutes the *discrete product market* version of Rust (1985). In this appendix we lay out a simple and efficient numerical solution algorithm for this limiting case, which constitutes the source of precise starting values for the main numerical algorithm in Section 3. We have proven that the following results from Rust (1985) continue to hold in our discrete setting. Proofs for these results are available on request.

Theorem 1 (Equilibrium in homogeneous consumer economy)**.** *Consider the primary and secondary market for automobiles with one car make/model and homogenous consumers ($\sigma = 0$). Assume infinitely elastic supply of new cars at price $\overline{P}$ and infinitely elastic demand for scrapped cars at price $\underline{P}$. The unique stationary equilibrium $\{q, P, a^\star\}$ on this market exists, and is composed of:*

1. *Ownership distribution q, which is the unique invariant distribution corresponding to the physical transition probability matrix Q defined in (27);*
2. *Common scrappage age $a^\star$ equal to the optimal replacement age in the social planner's problem of optimal car replacement (in the absence of secondary market, or equivalently within the class of "buy and hold" strategies);*
3. *Non-increasing price function $P(a)$ defined by*

$$P(a) = \begin{cases} \overline{P} - \frac{1}{\mu}\big(W(0) - W(a)\big), & a \in \{1, \ldots, a^\star - 1\}, \\ \underline{P}, & a \geq a^\star, \end{cases} \tag{A13}$$

where $W(a)$ is the unique fixed point of the Bellman operator in the forementioned social planner's replacement problem.

Corollary C1. *In the equilibrium of the homogeneous consumer economy with no transaction costs defined in Theorem 1, consumers are indifferent between replacing their existing car with the car of any age available in the economy. This holds for owners of all ages of cars with positive shares in the stationary fleet age distribution.*

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

Corollary C2. *The equilibrium in the homogeneous consumer economy with no transaction costs defined in Theorem 1 is welfare maximizing, in particular the discounted expected utility of all consumers is equal to maximum attainable welfare, $V(a) = W(a)$, for all cars with positive shares in the stationary fleet age distribution.*

The main idea of the fast solution algorithm for the homogeneous consumer economy is to express the indifference condition from Corollary C1 as the system of $a^\star - 1$ linear equations to determine the unrestricted prices $P(a)$, $a \in \{1, \dots, a^\star - 1\}$. Because by Corollary C1 consumers are effectively indifferent between any dynamic trading strategies, the strategy of perpetual replacing an existing car of age a results in the maximum attainable expected discounted utility $V(a)$. We have for $a \in \{1, \dots, a^\star - 1\}$

$$V(a) = \tfrac{1}{1-\beta}\Big(u(a) - \beta\mu\big[P(a) - \big(1-\alpha(a)\big)P(a+1) - \alpha(a)\underline{P}\big]\Big). \tag{A14}$$

Let $V(0)$ denote the value of having a new car which is measured right after trading instead of the beginning of the period. Then (A14) also holds for $a = 0$, and $V(0) = W(0)$. Thus, using $V(a) - V(a+1) = W(a) - W(a+1)$ (Corollary C2) and the definition of the price function (A13), we have for $a \in \{0, \dots, a^\star - 2\}$

$$V(a) - \mu P(a) = V(a+1) - \mu P(a+1), \tag{A15}$$

which leads to the following linear equation in prices $(P(a), P(a+1), P(a+2))$:

$$\begin{aligned}\mu P(a) + \mu\big(\alpha(a)\beta - \beta - 1\big)P(a+1) + \mu\beta\big(1 - \alpha(a+1)\big)P(a+2) &= \\ = u(a) - u(a+1) + \beta\mu\underline{P}\big(\alpha(a) - \alpha(a+1)\big).\end{aligned} \tag{1}$$

The collection of equations (1) for $a \in \{0, \dots, a^\star - 2\}$ forms the system of $a^\star - 1$ equations with $a^\star - 1$ unknowns, and can be easily solved numerically under our assumption that $u(a+1) \le u(a)$ for $a \ge 0$. The system be written in matrix form as

$$X \cdot P = Y, \tag{A16}$$

where P is the column vector of prices $\{P(a)\}_{a \in \{1, \dots, a^\star - 1\}}$,

$$X = \begin{bmatrix} \alpha(0)\beta - \beta - 1 & \beta - \alpha(1)\beta & 0 & 0 & \cdots & 0 & 0 \\ 1 & \alpha(1)\beta - \beta - 1 & \beta - \alpha(2)\beta & 0 & \cdots & 0 & 0 \\ 0 & 1 & \alpha(2)\beta - \beta - 1 & \beta - \alpha(3)\beta & \cdots & 0 & 0 \\ 0 & 0 & 1 & \alpha(3)\beta - \beta - 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \alpha(a^\star-3)\beta - \beta - 1 & \beta - \alpha(a^\star-2)\beta \\ 0 & 0 & 0 & 0 & \cdots & 1 & \alpha(a^\star-2)\beta - \beta - 1 \end{bmatrix},$$

$$Y = \begin{bmatrix} \big(u(0) - u(1)\big)/\mu + \beta\underline{P}\big(\alpha(0) - \alpha(1)\big) - \overline{P} \\ \big(u(1) - u(2)\big)/\mu + \beta\underline{P}\big(\alpha(1) - \alpha(2)\big) \\ \big(u(2) - u(3)\big)/\mu + \beta\underline{P}\big(\alpha(2) - \alpha(3)\big) \\ \vdots \\ \big(u(a^\star-3) - u(a^\star-2)\big)/\mu + \beta\underline{P}\big(\alpha(a^\star-3) - \alpha(a^\star-2)\big) \\ \big(u(a^\star-2) - u(a^\star-1)\big)/\mu + \beta\underline{P}\big(\alpha(a^\star-2) - 1\big) \end{bmatrix}.$$

Note that system (A16) is well defined for any $a^\star \ge 2$, and consequently a solution to the system denoted $P(a, a^\star)$ can be computed for any value of $a^\star$. We show that the solution of the social planner's problem $W(a)$ corresponds one-to-one to the solution of the linear system (A16) under the condition on the price vector $\forall a \in \{1,\dots,a^\star\} : \underline{P} \le P(a,a^\star) \le \overline{P}$ and $P(a^\star, a^\star+1) < \underline{P}$. Because solving a system of linear equations implies much lower computational cost than finding a fixed point $W(a)$, and in particular because the dimensionality of the social planner's problem is larger than that of the linear system, an iterative algorithm that solves (A16) for various $a^\star$ to ensure the conditions above hold, is the most efficient numerically.

C Likelihood when accidents are unobserved

Consider the transition probability in a partial information situation where we do not directly observe whether a given car is involved in an accident that leads to its scrappage. As a result we do not fully observe the state x_t representing car holdings of a given household, since for any car state $x_t = (i_t, a_t)$ where $a_t = \bar{a}_{i_t}$ (i.e. cars that have reached the mandatory scrappage state/age $\bar{a}_{i_t}$) we cannot distinguish which of those car are in that state due of an accident (since accidents are not directly observed) and which are in that state not due to an accident, but due to being age $\bar{a}_{i_t} - 1$ in period $t-1$. However we do observe whether the previously chosen car (j_{t-1}, d_{t-1}) was scrapped (exogenously due to accident or endogenously due to voluntary scrappage decision) or not. This implies we fully observe the post-trade ownership state δ_t, and so our approach to inference in the partial information case will be based on the transition probability $P(\delta_{t+1}|\delta_t, \tau, \theta)$.

In the case where $\delta_t = (\emptyset, \zeta_t)$ (either a decision to enter the no car state via a purge decision with an associated scrappage outcome ζ_t, or a decision to remain in the no car state if $x_t = \emptyset$ in which case ζ_t is not relevant), $P(\delta_{t+1}|\delta_t, \tau, \theta)$ is just the conditional choice probability given in equation (13).

If the household did choose to either trade for a car (i_t, a_t) or keep their car (i_t, a_t) at time t, since we do not observe whether that car was involved in an accident between t and $t+1$, the relevant transition probability $P(\delta_{t+1}|\delta_t, \tau, \theta)$ is a mixture of two choice probabilities, depending on whether the car was involved in an accident or not. Thus, when a scrappage event is observed for the car the household chose at t, then $s_{t+1} = 1_s$, (here we let the scrappage indicator, s_{t+1}, equal 1_s if the car the household chose at t was scrapped by period $t+1$, and 0_s otherwise), the transition probability $P(\delta_{t+1}|\delta_t, x_t, \tau, \theta)$ is given by

$$P(j_{t+1}, d_{t+1}, \zeta_{t+1} = 1_s | i_t, a_t, \zeta_t, \tau, \theta) = \Pi(j_{t+1}, d_{t+1}| i_t, \bar{a}_{i_t}, \tau, \theta)\alpha(i_t, a_t) + \\ \Pi(j_{t+1}, d_{t+1}| i_t, a_t+1, \tau, \theta)\Pi(s_{t+1} = 1_s | j_{t+1}, d_{t+1}, i_t, a_t+1, \tau, \theta)(1-\alpha(i_t, a_t)),$$

where $\Pi(j_{t+1}, d_{t+1}| i_t, \bar{a}_{i_t}, \tau, \theta)$ is the conditional choice probability for a household holding a "clunker" which requires a forced scrappage of the vehicle by our modeling assumptions. Thus, in the event of an accident of the car the houseold chose at time t, we represent it as a transition to the clunker state $\bar{a}$ and in this choice there is no choice over whether or not to sell or scrap the car, so $\Pi(s_{t+1} = 1_s | j_{t+1}, d_{t+1}, i_t, \bar{a}, \tau, \theta) = 1$. If there is no accident, then the household does have a choice of whether to scrap or sell the car, so in this case we do include the conditional probability of the scrappage choice, $\Pi(s_{t+1} = 1_s | j_{t+1}, d_{t+1}, i_t, a_t+1, \tau, \theta)$ to calculate the overall transition probability in (2) in the event of an observed scrappage of the chosen car between period t and $t+1$.

If there was no scrappage of the car the household chose at time t between periods t and $t+1$, which we denote by $\zeta_{t+1} = 0_s$, then $P(\delta_{t+1}|\delta_t, x_t, \tau, \theta)$ is given by

$$P(j_{t+1}, d_{t+1}, \zeta_{t+1} = 0_s | i_t, a_t, \zeta_t, \tau, \theta) = \\ \Pi(j_{t+1}, d_{t+1}| i_t, a_t+1, \tau, \theta)(1 - \Pi(s_{t+1} = 1_s | j_{t+1}, d_{t+1}, i_t, a_t+1, \tau, \theta))(1-\alpha(i_t, a_t)),$$

i.e. it is the probability of choosing to trade the existing car but also choosing not to scrap at $t+1$ conditioning on the event no accident occurred between t and $t+1$ either.

In the event of a choice to keep the current car at time $t+1$, which we have denoted earlier in the paper via the special symbol $(j_{t+1}, d_{t+1}) = \kappa$ to distinguish it from a decision to trade the current car (i_t, a_t) for another car with the same type and age, i.e. $j_{t+1} = i_t$ and $d_{t+1} = a_t$, we can conclude that no accident and no voluntary scrappage of the previously chosen vehicle could have occurred. Since the keep decision obviates any choice about selling or scrapping the existing car, the transition probability for the keep decision is given by

$$P(j_{t+1} = i_t, d_{t+1} = a_t + 1, \zeta_{t+1}|i_t, a_t, \zeta_t, \tau, \theta) = \Pi(\kappa|i_t, a_t + 1, \tau, \theta)(1 - \alpha(i_t, a_t)). \tag{A17}$$

We include the previous car ownership state x_t as a conditioning variable in the transition probability $P(\delta_{t+1}|\delta_t, x_t, \tau, \theta)$ just as we did in the full information case above since when a household decides to keep their current car, we need the information in the incoming car state $x_t = (i_t, a_t)$ to determine the coditional probabilities of a scrappage decision and whether the car will have an accident.

The Kullback-Leibler distance in the case of unobserved accidents is given below

$$D(\theta) \equiv$$
$$\sum_\tau \sum_x \sum_\delta \sum_{\delta'} \left[\log\left(P(\delta'|\delta, x, \tau)\right) - \log\left(P(\delta'|\delta, x, \tau, \theta)\right)\right] P(\delta'|\delta, x, \tau) P(\delta|x, \tau) q_\tau(x) f(\tau).$$

We assume that via direct observation post-trading ownership states, δ_t and δ_{t+1}, it is possible to non-parametrically estimate the transition probability $P(\delta'|\delta, x, \tau)$ for each observed household type τ. Note that though x is not always fully observed, it is observed when a household chooses to keep their car. It is not observed for all decisions involving trading the previous car, however the identity of the new car captured in δ is a sufficient statistic for determining the probability distribution of δ' and in the case of a trade, $P(\delta'|\delta, x, \tau)$ is independent of x given δ. It follows that $P(\delta'|\delta, x, \tau)$ can be non-parametrically estimated in the case where accidents are unobserved and thus the state x is not fully observed.

It is also possible to non-parametrically estimate the cross-sectional distribution of car ownership choices $q_\tau(\delta)$ for each household type τ. This is the stationary distribution of car ownership states, prior to accounting for accidents, whereas $q_\tau(x)$ is the actual incoming stationary distribution of car ownership states, accounting for accidents. It is not hard to show that if $q_\tau(x)$ is the invariant distribution of car ownership states at the start of any period t for household type τ, then $q_\tau(\delta) = \sum_x P(\delta|x, \tau) q_\tau(x)$ is an invariant distribution of household car choices, where $P(\delta|x, \tau)$ is the conditional probability that a household of type τ makes a car ownership decision δ at time $t+1$ conditional on their car state being x at time t, and where we have added the probability of keeping the current car $x = (i, a)$, denoted by $\delta = (\kappa, \kappa)$, to the conditional probability of trading for a car of type/age $x = (i, a)$. That is, when $\delta = x$ we define $q_\tau(\delta)$ as $\sum_{x'} P(\delta|x', \tau) q_\tau(x') + P(1|x, \tau) q_\tau(x)$. With this redefined δ variable it is no longer necessary to condition on x when writing the conditional choice probability of choosing a car (and scrapping the existing one) at time $t+1$, we can now simply express it as $P(\delta'|\delta, \tau)$. Then using the invariant distribution over car ownership decisions, $q_\tau(\delta)$, we can write the Kullback-Leibler distance in the case where accidents are not observed as follows

$$D(\theta) \equiv \sum_\tau \sum_\delta \sum_{\delta'} \left[\log\left(P(\delta'|\delta, \tau)\right) - \log\left(P(\delta'|\delta, \tau, \theta)\right)\right] P(\delta'|\delta, \tau) q_\tau(\delta) f(\tau). \tag{A18}$$

In summary, in the full information case the relevant transition probability that we use as

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

a basis for estimation is $P(x_{t+1}|x_t, \tau, \theta)$, where x_t is the fully observed car ownership state of the household, expressed as a product of the conditional choice probability for the household's decision over next period car state, and an accident probability that gives the final realized car state x_{t+1} at the start of the next period $t+1$. When accidents are unobserved, the relevant transition probability is $P(\delta_{t+1}|\delta_t, \tau, \theta)$, using the post-decision state δ_t, which is necessitated by the fact that we do not fully observe the car state x_t at each time period due to the fact that accidents are unobserved.

D Notes on the identification of the model with driving

This appendix is a short note on the identification of the model when we allow for driving, with a linear specification for the predicted optimal driving that can depend on the age of the car. Consider a utility function of the form

$$u(a, x) = \psi_0 + \psi_1 a + \psi_2 a^2 + (\gamma_0 + \gamma_1 a)x - \mu p x + \frac{\phi}{2} x^2 \tag{A19}$$

where a is the age of the car, x is vehicle kilometers driven each period, and the parameter vector is $\theta = (\psi_0, \psi_1, \psi_2, \gamma_0, \gamma_1, \mu, \phi)$. We can consider the sum of the first three terms on the right hand side of equation (A19) to be the utility of ownership *per se,* i.e. the utility a consumer gets from the pure ownership (and option value to drive) even if the household does not do any driving, $x = 0$. The remaining components are the net utility from driving, after deducting the cost of driving (translated into utility terms by multiplying by the marginal utility of money, μ). There are seven parameters to be identified for each consumer type/car type in the model, and for notational simplicity we have suppressed the dependence of all θ parameters except for μ (the marginal utility of money or price-responsiveness coefficient) since we assume that μ depends on consumer type τ but not on the car type j.

We want to consider the identification of these parameters from observations on: 1) consumer trading of automobiles, and 2) observed driving. The symbol p is the per kilometer cost of fuel plus any taxes, and in our model there is *no variation in this price, either over time or across consumers.* However there is variation in p across car types due to different fuel efficiency of different types and ages of cars. However we really cannot use variation in p as a source of identification of the model parameters if we are being fully faithful to the model, which currently does not allow any time series or cross sectional variation in p except over car types as noted above. Fortunately, we now show that we can identify the parameters using the information on car trading, and in a way that is "just identified" so we don't face a trade-off in terms of fitting the model of car driving or the model of car trading: each can be estimated separately and the structural parameters that imply the best-fitting "reduced-form" specifications for driving and car trading are derived below.

We derive the indirect utility function first, by using the utility function above to calculate the optimal level of driving,

$$x^*(p, a) = -\frac{1}{\phi}\left[\gamma_0 + \gamma_1 a - \mu p\right]. \tag{A20}$$

Equation (A20) is the "structural driving equation" implied by the direct utility function $u(x, a)$ in equation (A19). But there is a corresponding "reduced form" or unrestricted driving equation which we denote by

$$x = d_0 + d_1 a + d_2 p. \tag{A21}$$

Though there is no identifying variation in p we do have identifying variation in a, so we can

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022. "Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

separately identify the constant term in the driving equation in (A20) and the age coefficient. So we can treat the age coefficient γ_1/ϕ as known given knowledge of observed driving $x^*(p,a)$ and also the constant term $[\gamma_0 - \mu p]/\phi$, though at this point we cannot separately identify all the parameters from only these two identified coefficients from the "driving equation." Note that optimal driving must be positive, so this implies an additional inequality restriction on the parameters,

$$\gamma_0 + \gamma_1 a \geq \mu p, \quad a = \{0, 1, \ldots, \bar{a} - 1\} \tag{A22}$$

and if, as we expect, $\gamma_1 < 0$, then the set of restrictions below can be reduced to this single inequality restriction at the last age consumers are allowed to own cars

$$\gamma_0 + \gamma_1(\bar{a} - 1) \geq \mu p \tag{A23}$$

but if this inequality is violated, then it is easy to see that $x^*(p,a) = 0$ and the consumer gets zero utility from owning a car (and hence absent extreme value shocks would not buy one since the utility from the outside good is normalized to zero, and there are purchase price and transactions costs involved in buying a car).

Now plugging the optimal driving back into utility, we can derive the indirect utility function,

$$\begin{aligned} u(a, x^*(p,a)) = v(a,p,\tau) &= \psi_0 + \psi_1 a + \psi_2 a^2 - \frac{1}{2\phi}\left[\gamma_0 + \gamma_1 a - \mu p\right]^2 \\ &= u_0 + u_1 a + u_2 a^2 \end{aligned} \tag{A24}$$

where

$$u_0 = \psi_0 - \frac{1}{2\phi}\left[\gamma_0 - \mu p\right]^2, \qquad u_1 = \psi_1 - \frac{\gamma_1}{\phi}\left[\gamma_0 - \mu p\right], \qquad u_2 = \psi_2 - \frac{1}{2\phi}\left[\gamma_1^2\right]. \tag{A25}$$

We can consider the coefficients (u_0, u_1, u_2) and the marginal utility of money as identified from an unrestricted or "reduced form" dynamic discrete choice model of car trading. We now argue that the seven parameter specification of utility $u(x, a, \theta)$ given in equation (A19) is just identified from unrestricted estimation of the reduced-form driving equation (A21) and the dynamic discrete choice model of car trading. First, assume we can identify the marginal utility of money, μ, from estimation of the latter model. Then there are only 6 remaining structural parameters to be identified, $(\psi_0, \psi_1, \psi_2, \gamma_0, \gamma_1, \phi)$ and these are determined from the following 6 equations that provide enough flexibility to ensure perfect unrestricted fit of both the reduced-form driving equation (parameters (d_0, d_1, d_2)) and the reduced-form dynamic discrete choice model (parameters (u_0, u_1, u_2) plus the marginal utility parameter μ).

$$d_0 = -\frac{\gamma_0}{\phi}, \qquad d_1 = -\frac{\gamma_1}{\phi}, \qquad d_2 = \frac{\mu}{\phi} \tag{A26}$$

Thus, given the estimate of the marginal utility of money $\hat{\mu}$ from the dynamic discrete choice model, we can back out $\hat{\phi}$ from the last equation of (A26), and thus also $(\hat{\gamma}_0, \hat{\gamma}_1)$. Then given these parameter estimates we can determine the parameters $(\hat{\psi}_0, \hat{\psi}_1, \hat{\psi}_2)$ from equation (A25) in a way that entails no restrictions on the estimation of the coefficients (d_0, d_1, d_2) of the reduced-form driving model (A21) or the coefficients (μ, u_0, u_1, u_2) of the reduced-form dynamic discrete choice model of car trading.

To derive (A24) we assumed that inequality (A23) holds so that the consumer would want to do some positive driving at all possible ages. However if the inequality does not hold at all ages, then $u(a, x^*(p,a))$ is given by (A24) only for ages a satisfying inequality (A22) and

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022. "Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

for all higher ages $a \geq \hat{a}$ (where $\hat{a}$ is the largest age for which inequality (A24) holds), then $x^*(x,p) = u(a, x^*(p,a)) = 0$ for all $a > \hat{a}$. Notice that $u_0 \geq 0$ and $u_2 \geq 0$ and if $\gamma_1 < 0$, then $u_1 \leq 0$. Thus, consumer preferences over cars of different ages are expected to be decreasing and convex in the age of the car. The strict convexity in age is required in order to imply a finite amount of driving: i.e. if $\phi = 0$ (so indirect utility is linear in a) then $x^*(p,a)$ is predicted to be infinite for any a where the inequality restriction (A22) is strict.

Now let's suppose that we could identify the four parameters (u_0, u_1, u_2, μ) by estimating a model of automobile trading that ignored driving, but with a utility function that is quadratic in the age of the car as in equation (A24). We presume that the marginal utility of money parameter μ is identified by the variation in trading cars of different ages using known prices of used cars, $P(a)$, $a = 0, \ldots, P(\bar{a}-1)$ that are treated as data. Even though these prices are the same for all consumers, there is variation over ages of different cars and combined with the scrappage decision, we will assume these coefficients can be identified.

Now given the additional observation on driving are the other coefficients $(\gamma_0, \gamma_1, \phi)$ identified? The answer is yes. As we noted above, the ratio γ_1/ϕ is identified from the driving equation (A20), and the coefficient γ_1^2/ϕ is identified as the quadratic coefficient u_2 from the dynamic discrete choice model of vehicle trading with the quadratic in age utility function given in (A24). Thus γ_1 is identified as the ratio of these two coefficient estimates. Next, given γ_1 we can identify ϕ from either the linear term of the driving equation, $-\gamma_1/\phi$ or the coefficient u_2 on the a^2 term of the utility of the dynamic discrete choice model, $u_2 = -\gamma_1^2/(2\phi)$. Then using the equation for u_1, the coefficient on a in the dynamic discrete choice model, we can identify γ_0. Thus the parameter vector $\theta = (\gamma_0, \gamma_1, \mu, \phi)$ is in fact overidentified since there are additional restrictions implied by the parameters via the constant term in the driving equation, $x^*(0,p)$ and the constant term u_0 in the utility function in the dynamic discrete choice model.

These parameters can be estimated jointly by maximum likelihood, with observed driving affected by measurement error. But doing this would require estimation of an additional parameter for the variance of the measurement error in driving. We would also have to take seriously the inequality restriction (A22) and make sure that we don't ignore it and get strange results from squaring negative predicted driving, rather than carefully obeying the inequality restriction that implies cars provide zero direct utility when predicted driving is negative.

To get started, I would favor an "indirect least squares" sort of approach where we just estimate the coefficients (u_0, u_1, u_2, μ) of the dynamic discrete choice ingoring driving and make sure we can fit the pattern of trading in cars well. Then with these estimated, we can estimate a linear driving equation and "back out" the implied $(\gamma_0, \gamma_1, \phi)$ coefficients via an "indirect least squares" approach and check that they are reasonable. This would be the "second step". Finally for more efficient parameter estimates, we could estimate the model "structurally" (either by ML or by MM) that impose the "cross equation restrictions" from the observations on driving and observations on car trading using the indirect least squares parameters as starting points and then recasting the parameters of the model directly in terms of the "deep structural parameters" that allow for driving, $\theta = (\gamma_0, \gamma_1, \mu, \phi)$.

E Estimation details

E.1 Data and Institutional Details

The data comes from the Danish demographic registers and covers the period 1996 to 2008. The dataset covers the universe of all Danish households and all cars owned by private individuals. Driving information is obtained from odometer readings that occur when cars are taken to

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

mandatory driving inspections biannually starting from a car age of four.[2] Fuel prices come from eof.dk and are a country-level average.

Car types: Cars are aggregated into four discrete types: we first split cars based on whether the car's weight is above or below the median for that car's vintage cohort, and then within each of those two subsamples, we further split cars based on whether they are above or below the median fuel efficiency. This way, we get four classes of cars, that we name "green" or "brown" for high and low fuel efficiency respectively, and "heavy" and "light" for high and low weights, respectively. Making the splits separately by vintage has the benefit that the distribution of car types is roughly constant over time, but has the drawback that car attributes for the four classes are not constant.[3] Therefore, we simply take the average of the car attributes for each of the four types in our model. Finally, we split cars into 25 age groups from brand new to 24 years old, where the last age category captures cars of 24 years or older. Table A2 presents key summary statistics for our four car types, aggregated over all the years of our sample as well as all the different car ages.

Household types: Households are split into 8 different types based on whether the household is a couple or single, has high or low work distance, and whether income is high or low based. Income splits are based on the median income in the demographic cell. Work distance comes from the Danish tax deduction for travel distance to work. This deduction is only applicable for full-time workers living further than 12 km from their work place (each way), and slightly under half of Danes have high work distance by this definition, although it differs quite a bit by cohabitation status. Table A1 presents summary statistics, where we have computed the weighted average over the years of our data for each household type to simplify the exposition. We also present averages for the number of kids in the household.

Prices: Recall that our model takes the new car price and the scrap price as given. Towards this end, we leverage data on the MSRPs to construct new car prices, which we take as the weighted averages of all underlying car types matched to each of our four discretized types, j. We construct an estimate of the scrappage price by leveraging data on suggested annual depreciation rates of 87% that we have from the Danish Automobile Dealer Association.[4] We construct these numbers for all years 1998 to 2008 and take the unweighted average over years.

The registration tax paid upon the purchase of a new car in Denmark is among the highest in the world. It is a linear tax that has a kink, K, with one rate, τ_1, below K and a higher rate, $\tau_2 > \tau_1$, applying to any price above K. Finally, 25% VAT is paid of the price including the tax. So if the price before the registration fee and VAT is given by P, then the registration fee to be paid is given by $T(P) = \tau_1 \min(1.25P, K) + \tau_2 \max(1.25P - K, 0)$. In 2008, K was 81,000 DKK (around 16,000 USD), τ_1 was 105% and τ_2 was 180%. In our counterfactuals in Table 1, we lower τ_1 and τ_2 to half their initial values, i.e. 52.5% and 90% respectively. There are also annual taxes for car ownership as well as mandatory insurance costs, which we abstract from.

We use a social cost of carbon of US$50/ton (290 DKK) and the other external costs per kilometer travelled are valued at 0.6216 DKK/km and they consist of noise, accidents, congestion and local air pollution, as measured by Transport (2010).

We do not observe scrappage in our data *per se*. Instead, we define a vehicle as having been scrapped if an ownership spell ends and no other ownership spell ever begins afterwards. Since our extract of the ownership register comes from September of 2011 while our last sample year is 2008, this means that a car should have been without owner for 3 years, which typically

[2] That is, a car is inspected at ages 4, 6, 8, 10, and so forth.

[3] In reality, technological progress implies that car attributes are improving over time as engines can drive further per liter of fuel. To accomodate our model's stationary nature, we ignore this "attribute inflation".

[4] The rates vary by car type but that variation is negligible, especially compared to the variation in new car prices across car types.

Table A1: Summary Statistics for Households

τ	Name	N	Income	1(Single)	Work distance	Age	1(Urban)	No. kids
1	Low WD, Couple, Poor	6,500,464	311.68	0.00	0.00	55.03	0.22	0.48
2	Low WD, Couple, Rich	6,352,821	777.19	0.00	0.00	46.38	0.21	1.03
3	Low WD, Single, Poor	7,906,100	109.92	1.00	0.00	54.21	0.35	0.11
4	Low WD, Single, Rich	7,666,452	301.15	1.00	0.00	48.21	0.33	0.20
5	High WD, Couple, Poor	4,031,412	494.61	0.00	34.63	40.58	0.12	0.99
6	High WD, Couple, Rich	3,862,441	862.43	0.00	42.13	43.57	0.12	1.21
7	High WD, Single, Poor	1,217,611	215.04	1.00	26.71	33.85	0.25	0.22
8	High WD, Single, Rich	1,171,919	413.24	1.00	32.98	41.14	0.22	0.24

Note: The column "N" denotes the observations of each household type available across all the years, 1996–2008. The remaining variables are all weighted averages of the corresponding variables with the annual observation counts as weights. Household types are defined based on splitting the sample into cells based on single/couple status, whether work distance is zero or positive, and finally splitting households in two depending on income within the cell is above or below the median. Work distance is based on a travel tax deduction, and it is only positive if one of the household members has more than 12 km to work (each way), and so it is naturally zero for unemployed. The urban dummy is equal to one for the six largest cities in Denmark: Copenhagen, Frederiksberg, Aarhus, Aalborg, and Odense.

Table A2: Summary Statistics for Cars

	No car	light, brown	light, green	heavy, brown	heavy, green
Obs.	16895290	4683737	5594897	5351904	6183392
Diesel share		0.00	0.08	0.14	0.21
Depreciation Factor		0.87	0.87	0.87	0.87
Weight (tons)		1.42	1.28	1.96	1.64
Variables used in the model					
Price, new (1000 DKK)		174.90	144.55	299.45	253.40
Price, new excl. tax (1000 DKK)		67.33	56.41	102.91	89.76
Price, scrap (1000 DKK)		6.20	5.26	9.35	8.76
Fuel efficiency (km/l)		12.84	15.06	9.89	12.63

Note: The four car categories are defined by first splitting cars into two groups based on weight, and then on fuel efficiency within each weight sub-group. The splits are made separately for every car vintage, implying that the attributes of, say, a "light, green" car is changing over time. The variable "Depreciation Factor" is a suggested annual depreciation factor set by the Danish Automobile Dealer Association. The rate varies across cars but not over time, implying that the association uses a constant exponential discounting rule.

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

Table A3: First-stage OLS Estimates of the Driving Model

Dependent variable: thousands of kilometers driven per year			
γ_0	Intercept	19.07	(0.58)
$\hat{\gamma}_1^a/\phi_\tau$	Car age	-0.1325	(0.03)
$\hat{\gamma}_2^a/\phi_\tau$	Car age squared	-0.001975	(0.00)
$\hat{\gamma}_\tau/\phi_\tau$	Intercept, Low WD, Couple, Rich	4.43	(0.70)
$\hat{\gamma}_\tau/\phi_\tau$	Intercept, Low WD, Single, Poor	-3.752	(1.24)
$\hat{\gamma}_\tau/\phi_\tau$	Intercept, Low WD, Single, Rich	-0.0325	(0.79)
$\hat{\gamma}_\tau/\phi_\tau$	Intercept, High WD, Couple, Poor	9.825	(0.79)
$\hat{\gamma}_\tau/\phi_\tau$	Intercept, High WD, Couple, Rich	12.33	(0.74)
$\hat{\gamma}_\tau/\phi_\tau$	Intercept, High WD, Single, Poor	6.436	(1.52)
$\hat{\gamma}_\tau/\phi_\tau$	Intercept, High WD, Single, Rich	12.23	(1.27)
$\hat{\gamma}_j/\phi_\tau$	Car dummy: light, green	-1.994	(0.15)
$\hat{\gamma}_j/\phi_\tau$	Car dummy: heavy, brown	4.345	(0.15)
$\hat{\gamma}_j/\phi_\tau$	Car dummy: heavy, green	3.606	(0.14)
μ/ϕ	Price (common)	-7.074	(0.84)
$\hat{\mu}_\tau/\phi_\tau$	Price, Low WD, Couple, Rich	-4.111	(1.02)
$\hat{\mu}_\tau/\phi_\tau$	Price, Low WD, Single, Poor	4.732	(1.84)
$\hat{\mu}_\tau/\phi_\tau$	Price, Low WD, Single, Rich	0.2781	(1.16)
$\hat{\mu}_\tau/\phi_\tau$	Price, High WD, Couple, Poor	-6.41	(1.17)
$\hat{\mu}_\tau/\phi_\tau$	Price, High WD, Couple, Rich	-9.892	(1.09)
$\hat{\mu}_\tau/\phi_\tau$	Price, High WD, Single, Poor	-1.714	(2.29)
$\hat{\mu}_\tau/\phi_\tau$	Price, High WD, Single, Rich	-9.007	(1.91)
N	Driving periods	19,635,940	

means it has been scrapped. Note also that we do not observe whether a car was involved in an accident in the data, although our model will make a distinction between accidents and voluntary scrappage decisions.

E.2 Estimation

As explained in Section 6, estimation is composed of three steps:

1. Estimate the reduced form driving parameters from (A21) using linear regression: the estimates are in Table A3.
2. Estimate the reduced form dynamic discrete choice parameters from (A24) using Maximum Likelihood: the estimates are in Tables 4, 5 and 7 to 10.
3. Back out the "deep structural parameters" $\theta_{\tau,j} = (\psi_{\tau,j,0}, \psi_{\tau,j,1}, \psi_{\tau,j,2}, \gamma_{\tau,j,0}, \gamma_{\tau,j,1}, \phi_{\tau,j})$ for each of the 8 consumer types τ and 4 car types j using equations (A25) and (A26) of Appendix D.

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

Table A4: Estimates: Accidents

	light, brown	light, green	heavy, brown	heavy, green
Intercept	-5.5876	-6.0006	-5.6697	-5.7399
	(0.0122)	(0.0109)	(0.0096)	(0.0090)
Age slope	0.1725	0.2134	0.2007	0.1969
	(0.0016)	(0.0013)	(0.0009)	(0.0009)

Table A5: Estimates: Scrappage Decision

	Estimate
σ_s: Scrap utility error variance	0.3852
	(0.0093)
Intercept: selling (baseline is scrapping)	-1.5999
	(0.2022)
Selling in inspection years	-2.2955
	(0.0200)

Table A6: Estimates: Marginal Utility of Money

	μ_τ: marginal utility of money
Low WD, Couple, Poor	0.1074
	(0.0006)
Low WD, Couple, Rich	0.1059
	(0.0006)
Low WD, Single, Poor	0.0895
	(0.0007)
Low WD, Single, Rich	0.1024
	(0.0006)
High WD, Couple, Poor	0.0980
	(0.0006)
High WD, Couple, Rich	0.1091
	(0.0006)
High WD, Single, Poor	0.0890
	(0.0007)
High WD, Single, Rich	0.1028
	(0.0007)

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

Table A7: Estimates: Utility Intercept

$u_{\tau,j,0}$: intercept in indirect utility for car ownership				
	light, brown	light, green	heavy, brown	heavy, green
Low WD, Couple, Poor	3.5100	2.9687	4.8387	4.5461
	(0.0162)	(0.0152)	(0.0253)	(0.0231)
Low WD, Couple, Rich	3.8870	3.3155	5.4279	5.1089
	(0.0155)	(0.0146)	(0.0242)	(0.0220)
Low WD, Single, Poor	2.3158	2.0491	3.3213	3.0367
	(0.0177)	(0.0157)	(0.0272)	(0.0251)
Low WD, Single, Rich	3.1192	2.6901	4.4031	4.0691
	(0.0160)	(0.0148)	(0.0249)	(0.0227)
High WD, Couple, Poor	3.7463	3.2959	5.0193	4.8326
	(0.0149)	(0.0138)	(0.0231)	(0.0210)
High WD, Couple, Rich	4.6059	4.1597	6.2197	5.9814
	(0.0165)	(0.0154)	(0.0255)	(0.0232)
High WD, Single, Poor	2.6024	2.3494	3.5814	3.3905
	(0.0195)	(0.0169)	(0.0302)	(0.0276)
High WD, Single, Rich	3.4274	3.0413	4.7046	4.4860
	(0.0189)	(0.0169)	(0.0290)	(0.0265)

Table A8: Estimates: Utility Age Slope

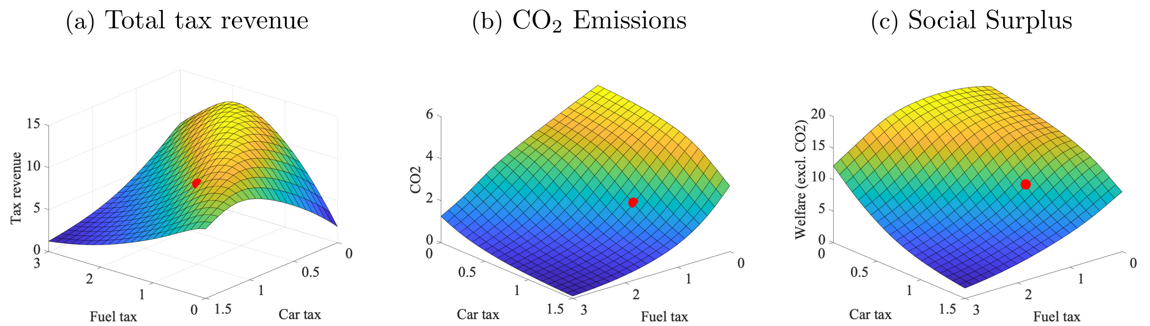
$u_{\tau,j,1}$: coefficient on age in indirect utility for car ownership				
	light, brown	light, green	heavy, brown	heavy, green
Low WD, Couple, Poor	-0.1397	-0.0914	-0.2084	-0.1651
	(0.0008)	(0.0009)	(0.0013)	(0.0011)
Low WD, Couple, Rich	-0.1520	-0.0973	-0.2294	-0.1881
	(0.0008)	(0.0009)	(0.0012)	(0.0010)
Low WD, Single, Poor	-0.0929	-0.0607	-0.1507	-0.1076
	(0.0009)	(0.0009)	(0.0013)	(0.0011)
Low WD, Single, Rich	-0.1253	-0.0836	-0.1955	-0.1487
	(0.0008)	(0.0009)	(0.0012)	(0.0010)
High WD, Couple, Poor	-0.1329	-0.0814	-0.2023	-0.1629
	(0.0008)	(0.0009)	(0.0011)	(0.0010)
High WD, Couple, Rich	-0.1570	-0.1055	-0.2423	-0.1995
	(0.0008)	(0.0009)	(0.0013)	(0.0011)
High WD, Single, Poor	-0.1082	-0.0718	-0.1675	-0.1249
	(0.0010)	(0.0009)	(0.0015)	(0.0012)
High WD, Single, Rich	-0.1436	-0.1001	-0.2105	-0.1673
	(0.0010)	(0.0010)	(0.0015)	(0.0012)

Supplemental Material (not copyedited or formatted) for: Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, Bertel Schjerning. 2022.
"Equilibrium Trade in Automobiles."
Journal of Political Economy 130(10). DOI: https://doi.org/10.1086/720463.

Table A9: Estimates: Transaction Costs (in utility terms)

Utility cost of transacting		
	Intercept	No car
Low WD, Couple, Poor	6.8509	1.7893
	(0.0221)	(0.0029)
Low WD, Couple, Rich	6.6965	1.0742
	(0.0221)	(0.0030)
Low WD, Single, Poor	6.7670	3.0771
	(0.0185)	(0.0045)
Low WD, Single, Rich	6.8514	2.5729
	(0.0209)	(0.0031)
High WD, Couple, Poor	6.4970	0.7784
	(0.0203)	(0.0036)
High WD, Couple, Rich	6.6843	0.1720
	(0.0227)	(0.0045)
High WD, Single, Poor	6.2479	2.3367
	(0.0192)	(0.0066)
High WD, Single, Rich	6.5250	1.7912
	(0.0216)	(0.0064)

Figure 1: The Effects of Varying the Fuel and Registration Tax Rates



Note: All three panels have the same x and y axes, namely the tax rate for fuel and car registrations respectively, normalized by the sample values so that the baseline outcomes occur at (1,1). The panels differ in terms of the rotation and which outcome is on the z axis: tax revenue, CO_2 emissions, and social surplus respectively. Social surplus is the sum of consumer surplus and tax revenue, subtracting the external costs of driving (accidents, congestion, etc.) including CO_2 valued at \$50/ton.

References

Arcidiacono, P. and R. A. Miller (2011): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 79(6), 1823–1867, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA7743.

McFadden, D. (1981): "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data*, ed. by C. Manski and D. McFadden, pp. 198–272. MIT Press, Cambridge, Massachusetts.

Rust, J. (1985): "Stationary Equilibrium in a Market for Durable Assets," *Econometrica*, 53(4), 783–806.

Rust, J., J. Traub and H. Wozniakowski (2002): "Is There a Curse of Dimensionality for Contraction Fixed Points in the Worst Case?," *Econometrica*, 70(1), 285–329.

Transport, D. (2010): "Transportøkonomiske Enhedspriser," Discussion paper, Technical University of Denmark, version 1.1.