Equilibrium Trade in Automobiles

Kenneth Gillingham

Yale University

Fedor Iskhakov

Australian National University

Anders Munk-Nielsen

University of Copenhagen

John Rust

Georgetown University

Bertel Schjerning

University of Copenhagen

We introduce a computationally tractable dynamic equilibrium model of automobile markets with heterogeneous consumers, focused on stationary flow equilibria. We introduce a fast, robust algorithm for computing equilibria and use it to estimate a model using nearly 39 million observations on car ownership transitions from Denmark. The estimated model fits the data well, and counterfactual simulations show that Denmark could raise total tax revenue by reducing the new-car registration tax rate. We show that reducing this tax rate while raising the tax rate on fuel increases aggregate welfare, tax revenue, and car ownership, while reducing car ages, driving, and CO_2 emissions.

Electronically published August 11, 2022

Journal of Political Economy, volume 130, number 10, October 2022.

https://doi.org/10.1086/720463

I. Introduction

Modeling the automobile market is particularly challenging because of the trading and substitution possibilities that exist as a result of the presence of a secondary market where used cars are traded. Not only are there dozens or even hundreds of different makes and models of new cars to choose from in the primary market: consumers have a huge array of used-car options as well. They also decide whether to scrap or sell their car and can respond to an increase in new-car prices by switching to the outside good (i.e., not having a car) or holding their existing used car longer. Endogenous scrappage of cars is also of interest for safety and environmental reasons, since it is well documented that used cars become less safe and pollute more as they age.[1]

We develop a dynamic model of trading in new and used cars that demonstrates how secondary markets lead to significant gains from trade via clear patterns of specialization in the holdings of cars by heterogeneous consumers. The secondary market facilitates a "hand-me-down chain" for cars, where rich consumers buy brand new cars and hold them for a few years before selling to other, slightly less rich consumers, who hold the cars for a few more years before trading them to even poorer consumers, who may hold the car until it is involved in an accident or voluntarily scrapped.

The secondary market creates substitution possibilities that can limit market power and affect pricing decisions by new-car producers in the primary market. High government taxation of new cars can also cause consumers to hold their used cars longer or to substitute to the "outside good," that is, not to own a car. Government sales taxes and regulations on emissions and safety can also interfere with the operation of secondary markets and reduce trade and consumer welfare. Beyond some point, sufficiently high taxation and overly onerous safety/emissions regulations can serve to kill off the secondary market and push consumers into

We dedicate this paper to James A. Berkovec, whose contributions to the development of microfounded equilibrium models of the auto market was far ahead of his time and so inspirational to our own work. His untimely death at age 52 in 2009 remains a huge loss to the economics profession. We also acknowledge helpful comments from the editor and referees and from Charles Manski and Dmitriy Stolyarov, whose contributions to modeling dynamic equilibrium in auto markets are equally inspiring and important. We received additional feedback and suggestions from Aureo de Paula, Nathan Miller, Eduardo Souza-Rodrigues, and Clifford Winston. We are grateful for funding from the IRUC (Intelligent Road User Charging) research project, financed by the Danish Council for Independent Research. Rust acknowledges financial support from the Gallagher Family Chair in Economics at Georgetown University. Iskhakov acknowledges financial support from ARC (Australian Research Council) Future Fellowship grant FT180100632. Replication files are available in a zip file. This paper was edited by Chad Syverson.

[1] See, e.g., NHTSA (2013) and Borken-Kleefeld and Chen (2015). The evidence is less clear on whether a car's fuel efficiency (measured as kilometers per liter) declines with age.

the outside good. Thus, there is a "Laffer curve" and the possibility of increasing total tax revenues by decreasing tax rates.

We use our model to analyze the fiscal and welfare effects of the new-car registration tax in Denmark, one of the highest in the world, which in the sample period amounted to 180% of the new-car price on top of a 25% value-added tax (VAT). The Danish government is highly reliant on this tax, which accounts for approximately 4% of all tax revenues, or about 2% of Danish gross domestic product. Our model predicts that the Danish tax rate is "over the top of the Laffer curve" and that Denmark could raise more tax revenue by reducing the registration tax rate. We show that reducing new-car taxes and raising fuel taxes improves aggregate welfare and significantly increases tax revenue and car ownership while reducing average car age and per-household driving and aggregate CO_2 emissions. However, the new policy is not a Pareto improvement absent offsetting transfers: the change in taxes reduces the welfare of those with long commute distances at the expense of those with shorter ones.

Our results are made possible by the fact that we can rapidly compute equilibria of the model using a fast and robust Newton-based solution algorithm that can be nested within a maximum likelihood estimation algorithm. Using Danish register data, which record the car ownership and trading decisions of all Danish citizens, we empirically estimate a version of our model with eight types of households and four types of cars and show that it provides a good approximation to car holdings and trading in Denmark. Our model provides a simple explanation for a striking zig-zag pattern in scrappage rates of older cars, whereby cars of even ages are scrapped with significantly higher probability than cars whose ages are odd numbers. We show that this is consistent with the rigorous biannual safety inspections in Denmark. Our estimation results reveal that these inspections have high perceived "hassle costs," so that once cars are sufficiently old, most Danes prefer to scrap their vehicles rather than incur the time and expense to repair them to pass the mandatory inspection.

The primary contribution of this paper is to advance the state of the art for computing equilibria in the primary and secondary markets for automobiles and other durable goods. We introduce a computationally tractable dynamic equilibrium model where new and used vehicles of multiple types (e.g., makes and models) are traded by heterogeneous consumers. Prices of used cars equate supply and demand for all car types and traded vintages. The ages at which cars are scrapped are also determined endogenously as part of the equilibrium. The model allows for transaction costs, taxes, and flexible specifications of car characteristics, consumer preferences, and heterogeneity. Our framework can be used to address a wide range of research and policy questions relating to the primary and secondary markets for vehicles. We also show how to incorporate a utility-based model of driving into the model, which is crucial for analyzing environmental policies.

We derive market demand from microaggregation of an individual-level dynamic discrete-choice model of ownership and trade in automobiles. Our specification of consumer heterogeneity includes additive idiosyncratic generalized extreme-value (GEV) preference shocks that can be interpreted as unobserved costs of maintaining an existing car, consumer-specific variations in search/transaction costs, and idiosyncratic variations in transaction prices and other costs involved in trading cars that constitute an important source of gains from trade that explain the existence of secondary markets. By varying the scale of these additive extreme-value preference shocks, we show how reductions in consumer heterogeneity reduce gains from trade and ultimately kill off secondary markets when trade frictions are sufficiently large.

The GEV specification results in logit or nested logit conditional choice probabilities for the decisions to keep or trade different types and ages of vehicles. We show how additional persistent consumer heterogeneity can be added, giving us the flexibility to match rich patterns of trading, including consumers who choose not to own cars (i.e., the “outside good”) and brand loyalty and brand-switching behaviors. We show that the choice probabilities, and thus aggregate demand, are smooth functions of car prices, which allows us to use fast derivative-based methods such as Newton’s method to solve for consumers’ dynamic trading strategies and equilibrium prices.

We formulate our model and define equilibrium in an infinite-horizon stationary environment. We use the machinery of Markov processes to describe trading behavior and characterize the vehicle holdings of different types of consumers as invariant distributions to certain Markov chains. These Markov chains reflect the trading of vehicles, their aging, and the impact of stochastic accidents that result in premature scrappage of some vehicles. Our stationary-equilibrium concept results in a very compact and elegant description of equilibrium that can be extended to nonstationary environments with macroeconomic shocks and overlapping generations of consumers with finite life spans.

Section II reviews the large theoretical and empirical literature on modeling auto markets and other durable goods on which we build. Section III introduces the basic model with multiple car brands and idiosyncratic consumer heterogeneity, and section IV adds persistent consumer heterogeneity that increases trade of cars between different consumer types. Section V describes how the model parameters can be structurally estimated by maximum likelihood using a doubly nested fixed-point algorithm that recomputes equilibrium prices, holdings, and consumer trading strategies each time the likelihood function is evaluated. We also establish the identification of the structural parameters. In section VI, we estimate the model using Danish register data and analyze the welfare and environmental impacts of changes in Danish car tax policies. Section VII concludes with a discussion of various directions in which the model can be extended.

II. Previous Work on Modeling Automobile Markets

A starting point of any discussion of the literature on equilibrium models of automobile markets is the well-known BLP model (Berry, Levinsohn, and Pakes 1995). This influential work focuses on the primary market for new vehicles but ignores the presence of the secondary market and the substitution possibilities it offers consumers. Rust (1985b) and Esteban and Shum (2007) were the first to tackle the challenging problem of solving for a full equilibrium in both the primary and secondary markets for automobiles. Rust studied the simultaneous determination of price and durability by a monopolist new-vehicle producer, while Esteban and Shum studied oligopolistic pricing of competing new-vehicle producers. To make progress, both of these studies assumed stationarity and zero transaction costs, which implies that consumers trade each period for their most preferred vehicle in the entire market.[2]

We build on a substantial literature focused on modeling equilibrium in secondary markets for automobiles, taking the price of new vehicles as given. The earliest work that we are aware of in this literature is a series of papers by Manski (1980, 1983) and Manski and Sherman (1980). These papers introduced theoretical models of equilibrium in secondary markets for cars that could be numerically solved for prices and quantities and used for policy forecasting of a wide range of policies of interest.

The next important early contribution was by Berkovec (1985), who microeconometrically estimated and numerically solved a large-scale equilibrium model of the new- and used-vehicle markets using a nested logit model. He defined "expected excess demand" by summing estimated discrete choice probabilities for cars of each age and class, net of scrappage.[3] Berkovec computed equilibrium prices using Newton's method to find a zero to a system of 131 nonlinear equations representing the excess demand for the vehicles in his model.[4]

[2] Esteban and Shum (2007) also assume quality-ladder preferences, which further simplifies the choice problem.

[3] Berkovec used a probabilistic model of vehicle scrappage from Manski and Goldin (1983), where the probability that a vehicle is scrapped is a decreasing function of the difference between the second-hand price of the vehicle (net of any repair costs) and an exogenously specified scrap value for the vehicle. This implies that, except for random accidents, there is very little chance that new vehicles are scrapped, but the probability that a used vehicle is scrapped increases monotonically with the age of the vehicle.

[4] Berkovec showed that the Jacobian matrix had a special structure he called "identity outer product" that enabled him to invert the Jacobian via inverting a smaller 48×48 matrix and doing some additional matrix vector multiplications.

The contributions of Manski, Sherman, and Berkovec were extremely advanced, given the computing power at the time, and in many respects represent the closest point of departure for our own work. However, their work was based on short-run, static equilibrium holding models of the market. Implicit in the static discrete-choice formulation is the assumption that consumers keep their vehicle for only a single period, so that at the end of each period consumers trade their current vehicle for their most preferred vehicle. Rust (1985a) formulated the first dynamic equilibrium model of automobile trading.[5] He assumed that the state of a vehicle is captured by its odometer reading x_t, which evolves according to an exogenous Markov process representing variable usage of cars with transition probability $\Phi(x_{t+1}|x_t)$ that reflects stochastic usage and deterioration of vehicles.[6]

When there are no transaction costs and the economy is in a stationary equilibrium (i.e., no macroeconomic shocks or other time-varying factors altering the market), the optimal trading strategy involves trading every period for the most preferred age/condition of vehicle $x^*(\tau)$, where τ indexes potentially heterogeneous preferences over "newness" of vehicles. However, the assumption of zero transaction costs is unrealistic, and so is the excessive trading behavior it implies. When there are transaction costs (which are distinct from trading costs, i.e., the difference between the price $P(x)$ of the car x a consumer wishes to buy and the price $P(x')$ of the car x' that the consumer wishes to sell), the optimal trading strategy involves less frequent trading, and consumers will generally keep cars for multiple periods. The optimal strategy then takes the form of an "(S, s) rule" reminiscent of optimal inventory theory: trade is characterized by two thresholds $(\underline{x}^*(\tau), \overline{x}^*(\tau))$, where $\underline{x}^*(\tau) < \overline{x}^*(\tau)$ and $\underline{x}^*(\tau)$ is the state of the optimal replacement vehicle whenever the consumer trades in for a new one. The value $\overline{x}^*(\tau)$ is the *replacement threshold*, or the odometer threshold where it is optimal to trade the current car in condition x for a replacement car in condition $\underline{x}^*(\tau)$. When transaction costs are zero, then $\overline{x}^*(\tau) = \underline{x}^*(\tau) = x^*(\tau)$, and it is optimal to trade for the optimal car $x^*(\tau)$ every period. However, in a homogeneous-agent economy, the slightest transaction costs will completely kill off the secondary market, driving all consumers into an autarkic "buy-and-hold" equilibrium where all consumers buy a brand new vehicle whenever they trade (i.e., $\underline{x}^*(\tau) = 0$) and hold it until it is optimal to scrap the car when the odometer exceeds an optimal replacement threshold $\overline{x}^*(\tau)$.

[5] Other dynamic models of vehicle choice appeared around this time, such as Mannering and Winston's (1985), but their analysis focused on dynamics of utilization and did not consider dynamics of car trading or equilibrium. In subsequent work, Winston and Yan (2021) developed an empirically estimable model of the dynamics of utilization and trading of cars, but in a partial-equilibrium framework.

[6] Since x_t fully captures the state of a car and is observable by both parties in a transaction, Rust's analysis avoided "lemons problem" information asymmetries, of the type analyzed in the seminal work of Akerlof (1970), that can kill off the secondary markets for cars.

There are potential gains from trade in a heterogeneous-agent economy that enable the existence of a secondary market and a wide range of car-trading strategies. However, establishing the existence of a stationary equilibrium in such an economy in the presence of transaction costs is challenging. Consider a consumer of type τ who desires to buy a vehicle with $\underline{x}^*(\tau) > 0$. When there are transaction costs, there is no guarantee that some other consumer τ' is willing to sell their vehicle at $\underline{x}^*(\tau)$. Using advanced methods from functional analysis (e.g., the Fan-Glicksburg fixed-point theorem), Konishi and Sandfort (2002) established the existence of a stationary equilibrium in the presence of transaction costs under certain conditions. Their proof shows that it is possible for the equilibrium price function $P(x)$ to adjust to prevent such coordination failures. However, to our knowledge, there has been no work actually calculating equilibria with transaction costs in this infinite-dimensional setting.

Stolyarov (2002) advanced the literature by assuming that the state of a vehicle can be summarized by its age a, which can take only a finite number of values, $a = 0, 1, 2, \dots, \bar{a}$, where $\bar{a}$ is age when cars are scrapped. Stolyarov introduced a continuous unidimensional parameterization of consumer heterogeneity with quasi-linear preferences and computed equilibria in the presence of stochastic transaction costs, using a fixed-point formulation of the problem. Gavazza, Lizzeri, and Roketskiy (2014) extended Stolyarov's approach by allowing households to own up to two vehicles, using a two-dimensional specification of consumer heterogeneity. They find that transaction costs have a large effect on equilibrium trade.[7]

Our model can be thought of as combining Stolyarov (2002) with the earlier work by Manski and Berkovec by using a multidimensional extreme-value specification to capture idiosyncratic consumer heterogeneity. We use a hierarchical specification of heterogeneity that includes both time-varying idiosyncratic preference shocks (i.e., the extreme-value error terms in the model) and flexible specifications for persistent heterogeneity and

[7] There is a close connection between models of automobile trading that incorporate transaction costs and models that emphasize information asymmetries, such as Akerlof (1970). House and Leahy (2004, 582) show how adjustment costs of the (S, s) variety discussed above "arise endogenously from adverse selection in the secondary market." For example, there are "lemon laws" in many countries that require sellers to compensate buyers for defects or problems in a car that were not disclosed and negotiated on at the time of sale. Dealers typically perform inspections and repair cars before selling and often provide a limited-term warranty, all of which mitigate the informational asymmetries and result in transaction costs that are often borne by the dealer. As a result, it is not clear that informational asymmetries seriously inhibit trade in used vehicles, but they would be expected to show up in transaction costs. Hendel and Lizzeri (1999) study equilibria in auto markets with and without asymmetric information and find that adverse selection does not necessarily kill off the secondary market. They find it difficult to empirically distinguish between predictions of models with asymmetric information and those with transaction costs and argue that, for Fords and Hondas at least, the evidence does not support adverse selection as the primary reason for steeper price declines of Fords as the vehicles age. In light of this, we use transaction costs to capture various trade frictions in auto markets, including informational ones.

fixed consumer types τ. The extreme-value distribution allows for continuous formulas for choice probabilities even in the case where there is no other time-invariant heterogeneity, and this continuity permits us to demonstrate the existence of equilibrium via the Brouwer fixed-point theorem. More importantly, we show that the excess-demand function for used cars in our model, $\mathrm{ED}(P)$, is a continuously differentiable function of P that enables us to rapidly and accurately calculate equilibrium prices by solving the system of nonlinear equations $\mathrm{ED}(P) = 0$ by Newton's method. This makes our approach very attractive for use in empirical work and policy modeling.

III. Equilibrium with Idiosyncratic Consumer Heterogeneity

In this section, we introduce a dynamic model of equilibrium trade in the automobile market. We use the concept of "stationary flow equilibrium" in the market of stochastically deteriorating durable goods from Rust (1985a) but adapt it for the discrete-goods trade in presence of flexible transaction costs. We start by considering equilibrium with J different makes/models of cars and a unit mass of consumers whose preferences for cars, as well as the outside option, are idiosyncratically heterogeneous. We adopt a GEV specification of consumer heterogeneity that results in a nested logit specification for choice probabilities similar to that of Berkovec (1985). In subsequent sections, we extend the framework to persistent heterogeneity in consumer preferences.

A. *Key Assumptions and Restrictions*

We consider a stationary equilibrium in an infinite-horizon economy where cars are initially sold as new in the primary market and then traded in used-car markets called "secondary markets." Consumers make purchase, replacement, trading, and scrapping decisions to maximize expected discounted utility with a common discount factor $\beta \in (0, 1)$. We focus on a stationary environment and do not allow for any "macro shocks" that could lead to time-varying fuel prices or prices of new cars.[8]

Our concept of equilibrium results in endogenous determination of a vector of equilibrium prices P with typical element P_{ja}, where $j \in \{1, \dots, J\}$ indexes makes/models and $a \in \{1, \dots, \bar{a} - 1\}$ indexes the ages of the traded cars. When the cars reach the upper bound $\bar{a}$, they are no longer safe to drive and are not allowed to be kept or traded

[8] Although it is possible to extend our framework to allow for macro shocks, this fundamentally changes the definition of the equilibrium. We defer this extension to future work because of the vastly greater computational challenges that it presents, as noted in the work of Krusell and Smith (1998) and Cao (2020).

and must be scrapped.[9] We treat the model as a "small open economy" where new-car prices are determined in the world market with an infinitely elastic supply of new cars at prices $\bar{P}_j$.[10] We assume that there is an infinitely elastic demand for cars at any age, including $\bar{a}$ for their scrap value $\underline{P}_j$, which normally results in $P_{ja} \geq \underline{P}_j\ \forall\ j, a$, provided that the level of transaction costs is not too high.

In our framework, all persistent differences between the cars are captured by the make/model $j \in \{1, \ldots, J\}$, and all time-varying characteristics of cars are reflected by the car age $a \in \{1, \ldots, \bar{a}\}$. The unit mass of cars in the economy is distributed among $J\bar{a}$ types given by the combination of car make/model and age (j, a). Clearly, used cars of the same age and type have idiosyncratic features, such as odometer reading, that we ignore, making it inconsistent with a single common price P_{ja} for all used cars of age a and make/model j. This is partially accounted for in our framework, via the stochastic GEV shocks that reflect not only idiosyncratic heterogeneity in consumer preferences but also idiosyncratic features of different used cars of the same age and type. Thus, we can interpret P_{ja} as the average price of a car of type j and age a, and components of the idiosyncratic shocks reflect customer- and car-specific deviations in these prices from the market average prices that are determined endogenously in equilibrium.

We assume that consumers' preferences are characterized by a common quasi-linear utility function

$$U(\cdot) = u(j, a) - \mu(\text{operating costs} + \text{trade and transaction costs}),$$

where the first term captures the utility of owning and using a car and the second term accounts for the monetary costs of ownership and trade. We assume that the marginal utility of an additional car is sufficiently small that no consumer would want to own more than a single car.[11] The parameter $\mu > 0$ is a simple way to capture income/wealth effects in the model. High values of μ can be interpreted as "being poor" because the cost of buying a new car will involve a high opportunity cost in terms of forgone consumption of other goods. We expect that the function $u(j, a)$ would be nonincreasing in a for all j, and we show below how it captures the utility of

[9] The same upper bound is assumed to hold for the cars of all makes/models without loss of generality to simplify exposition. It is straightforward to allow the upper bound to be j-specific with more complicated notation.

[10] However, our framework can be used for modeling competition in the primary market for new cars, where $\bar{P}_j$ can be set taking into account the substitution effects not only between different types of new cars but also between new and used cars.

[11] This is a reasonable assumption for a country like Denmark, where most households own only a single car. It also greatly simplifies notation and the presentation of our model. The assumption could be relaxed to extend the model to countries like the United States, where most households own multiple cars.

driving (utilization) as well as the expected nonmonetary cost of maintaining a car of age a.

Trade costs consist of the difference in prices of traded cars, with the addition of transaction costs. Let (i, a) be the make/model and age of the existing car and (j, d) denote the car the household purchases. Transaction costs are given by a function $T(i, a, j, d)$ that depends on both the traded cars and the whole set of prices $\{\bar{P}_j, \underline{P}_j\} \cup \{P_{ja}\}$, $a \in \{1, \dots, \bar{a} - 1\}$, $j \in \{1, \dots, J\}$.[12] We assume that the transaction costs are borne by both buyers and sellers. Even though it would be possible to work with general nonseparable specifications for transaction costs $T(i, a, j, d)$, for simplicity we assume that these costs are additively separable into two components denoted $T_{\mathrm{b}}(j, d)$ and $T_{\mathrm{s}}(i, a)$. The first component is associated with searching for and buying another car and the second with undertaking repairs and improvements to make the car of age a that the consumer is trying to sell acceptable to potential buyers. We do not make further restrictions on the functional form of the transaction costs, so we can allow for both fixed and proportional costs, that is, sales taxes and registration fees.

The total trade and transaction costs associated with selling car (i, a) to buy another car (j, d) are given by $P_{jd} - P_{ia} + T_{\mathrm{b}}(j, d) + T_{\mathrm{s}}(i, a)$. If a consumer chooses to "purge" their car and choose the outside option of not owning a car, the buyer-side components depending on (j, d) disappear from the expression, so the consumer faces only the seller-side transaction cost. Similarly, if a person without a car decides to purchase one, the seller-side components depending on (i, a) disappear from the trade costs. Finally, we assume that the trade-in price $\underline{P}_i$ for the cars being scrapped already includes all the costs associated with deregistering and transporting the clunker to the scrapyard. That is, we normalize the seller-side transaction cost of scrapping to zero. Similarly, we assume that the search cost is negligible when buying a new car and therefore normalize the buyer-side transaction cost for the new car buyers to zero.

B. *Consumer States and Choices*

The *state* of a consumer in any period t is given by the vector (i, a, ϵ), where $i \in \{\varnothing, 1, \dots, J\}$ denotes the make/model of car the consumer owns at the start of the period and $a \in \{\varnothing, 1, \dots, \bar{a}\}$ denotes its age. We use the symbol $\varnothing$ to denote the state of not owning a car. The random component of the state vector ϵ incorporates the (idiosyncratic) heterogeneity in cars and consumers.

[12] All quantities in the consumer choice problem depend on these prices, but for clarity we do not write this explicitly until sec. III.D.

We assume that at the start of each period a consumer who owns a car (i, a) can choose whether to keep it or trade it for another car of make/model (j, d) or choose the outside option of not having a car at all. We assume that trade occurs instantaneously at the start of each period, and thus the cars (j, d) that households hold after trading are utilized until the end of the period. Cars deterministically age from d to $d + 1$ but may be involved in total-loss accidents, which we model by stochastic transition to the terminal age $\bar{a}$ with probability $\alpha(j, d) \in [0, 1)$. The realized state of the car constitutes the car state at the start of period $t + 1$. Then instantaneous trading occurs, and the process repeats this way for the infinite future.

The cars that reach the terminal age $\bar{a}$ by either natural aging or as a result of an accident are exogenously scrapped and removed from the market during the trading stage. In addition, unless the existing car (i, a) is kept, it can be endogenously scrapped instead of being sold on the secondary market. It would seem that all consumers would prefer to sell their existing used car in the market rather than scrap it; however, there are transaction costs that a seller must incur, and the net value that a consumer might receive from selling a sufficiently old used car may be lower than the value from simply scrapping it. Our model allows consumers to choose whether to sell or scrap their existing car, depending on which option they prefer, which can also include unobserved idiosyncratic inspection/repair costs that are incorporated in the GEV shocks we describe below.

Let $C(i, a)$ denote the choice set for the consumer who enters the period with the existing model i car that is a years old. If the consumer has no car $(a = \varnothing)$, they can choose to remain in the no-car state $(d = \varnothing)$ or buy a new car $(d = 0)$ of any type $j \in \{1, \dots, J\}$ or one of the vintages available for sale in the secondary market. If the consumer already owns a car model i of age $a < \bar{a}$, they have an additional option of keeping it, which we denote $d = \kappa$. However, once a car reaches the terminal age $\bar{a}$, it is no longer possible to keep it, according to our assumption of exogenous scrappage.[13] Every time an existing car is traded, the consumer chooses to either sell it on the secondary market, which we denote by $s = 0_s$, or to take it to the scrapyard, which we denote by $s = 1_s$. The set of feasible choices for a consumer in a car state (i, a) is thereby summarized as follows:[14]

[13] The difference between the decision $d = \kappa$ to keep the current car of age a and the decision $d = a$ to trade for another car of the same age and model $j = i$ is in the incurred transaction costs.

[14] To simplify notation here and throughout the paper, we use single ø symbol for the no-car state, i.e., $C(\varnothing)$ instead of $C(\varnothing, \varnothing)$.

$$C(\emptyset) = C(i,\bar{a}) = \{\emptyset\} \cup \{1,\dots,J\} \times \{0,1,\dots,\bar{a}-1\}, \quad \forall\, j,$$
$$C(i,a) = \{(\emptyset,1_s),(\emptyset,0_s),\kappa\} \cup \{1,\dots,J\} \times \{0,1,\dots,\bar{a}-1\} \times \{1_s,0_s\}, \quad \forall\, j, a<\bar{a}. \tag{1}$$

Since cars must be scrapped when they reach the upper bound on age $\bar{a}$, they cannot be traded, and therefore the oldest car that be purchased in the secondary market is $\bar{a}-1$ periods old.

The random component ϵ is a vector whose dimension is equal to the number of elements in the choice set. We assume that ϵ has a multivariate GEV distribution that allows for flexible dependence between its elements but that the vectors are drawn independently between time periods and individuals, capturing the idiosyncratic heterogeneity between them. The elements of the vector ϵ, in turn, capture the differences between the discrete choices available to each individual, such as maintenance expenditures, search costs, and variability in the prices of traded cars, that reflect their idiosyncratic features.[15]

The GEV distribution we apply in our framework generalizes the standard multivariate extreme-value distribution and results in the choice probabilities that take the form that McFadden (1981) called the "nested multinomial logit" (NMNL), rather than the standard multinomial logit (MNL). Under this specification, it is possible to control for correlation patterns in different subsets of the overall choice set $C(i, a)$, and the choice itself can be represented as a sequence of choices from nested subsets (i.e., a filtration) of $C(i, a)$. However, even with this representation, all decisions are made simultaneously and instantaneously at the beginning of each period, as described above. The patterns of interdependence of the GEV distribution can be illustrated by a choice tree, a directed acyclical graph such as the one shown in figure 1.

The choice tree illustrated in figure 1 is one of the many possible ways to introduce dependence patterns into the components of the idiosyncratic shocks ϵ. This particular tree is for a consumer who owns a car, and it has four levels: the top level consists of the choices to purge ($j=\emptyset$), keep ($j=\kappa$), or replace ($j \in \{1,\dots,J\}$) the current car. Conditional on the decision to replace, the second level is the choice of the make/model j of the replacement vehicle. The third level contains the choice of the age d of the car to buy; for a new car, $d=0$. Finally, the fourth level

[15] The components of ϵ should be interpreted in an ex ante sense, as the idiosyncratic utility/disutility the consumer can expect from undertaking a search for a used car of a given type. We do not explicitly model the sequential search process in this paper, nor do we model the "microstructure" of auto dealers and other places consumers go to search for and buy used cars.

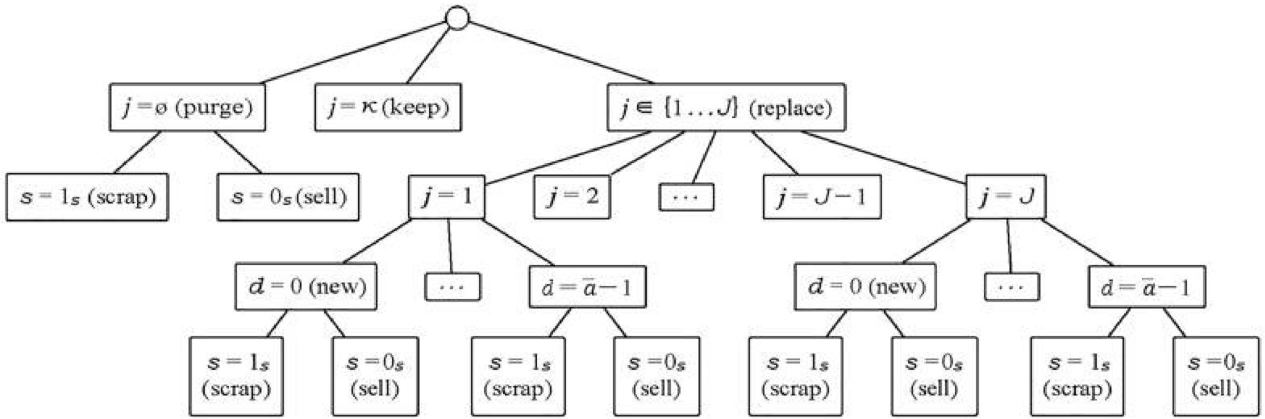

FIG. 1.—Example of the choice tree for a consumer who owns a car under the nested logit specification of the GEV distribution of the idiosyncratic heterogeneity term ϵ. Note the three choice nests at the top level (to purge, keep, or trade the existing car), two intermediate levels of nesting in the case of trading, and an additional scrappage choice of the existing car in the cases when it is traded.

contains the choice of whether to sell or scrap the existing car if it is not kept.[16] For each level choice, apart from top choice $j = \kappa$ in our example, there is a subtree of lower-level choices eventually leading to a distinct leaf of the tree corresponding to a particular alternative in the choice sets defined in equation (1).

Dependence patterns in the distribution of the elements of ϵ are determined by a set of scale parameters that can be defined individually for the alternatives immediately below the top node of each subtree of the overall choice tree. For example, in figure 1, the parameter σ controls the scale of idiosyncratic shocks at the top-level choice set $\{\o, \kappa, \{1, \ldots, J\}\}$ involving the decision to have no car, keep the current car, or trade for some other new or used car. The parameter σ_{r} controls the degree of similarity in the unobserved shocks affecting the choice of one of the J different types of cars, that is, idiosyncratic heterogeneity in "brand effects." The parameters σ_j and $j \in \{1, \ldots, J\}$ control the scale of idiosyncratic shocks affecting the choices of different ages of cars of a given make/model. Finally, the parameter σ_{s} controls the scale of idiosyncratic shocks reflecting unobserved components of transaction, inspection, and repair costs involved in selling the current car versus scrapping it. McFadden (1981) showed that in order for the GEV distribution to be a valid multivariate probability distribution, the similarity parameters must form a nonincreasing sequence along any particular branch. In our example this implies that $\sigma \geq \sigma_{\mathrm{r}} \geq \sigma_j \geq \sigma_{\mathrm{s}}$ for all $j \in \{1, \ldots, J\}$. As any of the similarity parameters approaches zero, the choice in the corresponding choice subset becomes deterministic, as do the choices in the subtrees below. When $\sigma = \sigma_{\mathrm{r}} = \sigma_j = \sigma_{\mathrm{s}}$ the choice tree collapses to a one-level tree with all alternatives in the choice set $C(i, a)$ on the same level. This is the MNL case with the implied independence-from-irrelevant-alternatives property: that is, there is no dependence among the components of ϵ.

C. *Consumer Dynamic Choice Model*

The optimal trading/holding strategy for cars is given by the solution of their infinite-horizon expected utility maximization problem, which constitutes a discrete-choice dynamic programming problem (Rust 1987, 1994). Let $V(i, a, \epsilon)$ be the value function for a consumer in state (i, a, ϵ), $i \in \{\o, 1, \ldots, J\}$, $a \in \{\o, 1, \ldots, \bar{a}\}$. For a consumer who does not own a car, it is given by

[16] The fourth-level scrappage choice appears on level two for the top-level decision to purge, without loss of generality.

$$V(\text{ø}, \epsilon) = \max\left(v(\text{ø}, \text{ø}) + \epsilon(\text{ø}); \max_{\substack{j \in \{1, \dots J\}, \\ d \in \{0, 1, \dots, \bar{a} - 1\}}} (v(\text{ø}, j, d) + \epsilon(j, d)) \right), \quad (2)$$

where the choice-specific value functions of remaining in the no-car state and of leaving the no-car state to buy a car of type j age d are respectively given by

$$\begin{aligned} v(\text{ø}, \text{ø}) &= u(\text{ø}) + \beta \mathrm{EV}(\text{ø}) \quad \text{and} \\ v(\text{ø}, j, d) &= u(j, d) - \mu\left(P_{jd} + T_{\mathrm{b}}(j, d)\right) \\ &\quad + \beta(1 - \alpha(j, d))\mathrm{EV}(j, d + 1) + \beta\alpha(j, d)\mathrm{EV}(j, \bar{a}). \end{aligned} \quad (3)$$

The expected-value functions $\mathrm{EV}(\text{ø})$ and $\mathrm{EV}(i, a)$ provide the conditional expected values of starting the next period without a car and with car of type i and age a, respectively, and are given by

$$\begin{aligned} \mathrm{EV}(\text{ø}) &= \int_{\epsilon} V(\text{ø}, \epsilon) f(\epsilon|\text{ø})\, d\epsilon \quad \text{and} \\ \mathrm{EV}(i, a) &= \int_{\epsilon} V(i, a, \epsilon) f(\epsilon|i, a)\, d\epsilon, \end{aligned} \quad (4)$$

where $f(\epsilon|\cdot)$ is the corresponding probability density function of a GEV distribution of the idiosyncratic shocks ϵ. Implicit in these formulas is the assumption that idiosyncratic shocks affecting the consumer's choice are independent of their past realizations. This implies that the $\mathrm{EV}(\cdot)$ functions depend only on the car that has been driven and aged during the current period and constitutes the car state in the beginning of the next period. Under our assumption of the GEV distribution of idiosyncratic shocks ϵ, the integrals in equation (4) can be expressed in closed form. The formulas depend on the assumed nesting structure of the choices. Later in this section, we provide the analytic formulas for $\mathrm{EV}(\cdot)$ corresponding to the nesting structure in figure 1.

The value function for a consumer who starts the period with a car of terminal age $\bar{a}$ is similarly given by

$$V(i, \bar{a}, \epsilon) = \max\left(v(i, \bar{a}, \text{ø}) + \epsilon(\text{ø}); \max_{\substack{j \in \{1, \dots, J\}, \\ d \in \{0, 1, \dots, \bar{a} - 1\}}} (v(i, \bar{a}, j, d) + \epsilon(j, d)) \right), \quad (5)$$

where again the first component corresponds to the decision to choose the outside option of not owning a car and the second to the purchase of a new car of type j and age d. Because at terminal age $\bar{a}$ the car is scrapped exogenously, the consumer does not have the option of keeping their car and also does not have an additional choice of endogenous scrappage of the existing car $(i,\ \bar{a})$. This is why the Bellman equation (5) looks very similar to that for the consumers who do not own a car (eq. [2]). The difference, however, is in the net scrap value $\underline{P}_i$ the consumer receives from the scrapyard net of towing and other costs of scrapping. Therefore, the relevant choice-specific value functions are given by

$$\begin{aligned} v(i, \bar{a}, \varnothing) &= u(\varnothing) + \mu \underline{P}_i + \beta \mathrm{EV}(\varnothing), \\ v(i, \bar{a}, j, d) &= u(j, d) - \mu\big(P_{jd} - \underline{P}_i + T_\mathrm{b}(j, d)\big) \\ &\quad + \beta(1 - \alpha(j, d))\mathrm{EV}(j, d + 1) + \beta\alpha(j, d)\mathrm{EV}(j, \bar{a}). \end{aligned} \tag{6}$$

Finally, the value function for a consumer who starts the period with a car of type i and age a is given by

$$\begin{aligned} &V(i, a, \epsilon) \\ &= \max\left(v(i, a, \kappa) + \epsilon(\kappa); \max_{s \in \{1_\mathrm{s}, 0_\mathrm{s}\}} (v(i, a, \varnothing, s) + \epsilon(\varnothing, s)); \right. \\ &\qquad \left. \max_{\substack{j \in \{1, \ldots, J\}, \\ d \in \{0, 1, \ldots, \bar{a} - 1\}, \\ s \in \{1_\mathrm{s}, 0_\mathrm{s}\}}} (v(i, a, j, d, s) + \epsilon(j, d, s)) \right). \end{aligned} \tag{7}$$

In this case, the endogenous scrappage choice $s \in \{1_\mathrm{s}, 0_\mathrm{s}\}$ has to be accounted for, so the complete set of choice-specific value functions that correspond to all the alternatives in the choice set $C(i,\ a)$ defined in equation (1) is given by

$$
\begin{aligned}
v(i, a, \varnothing, 1_s) &= u(\varnothing) + \mu \underline{P}_i + \beta \mathrm{EV}(\varnothing), \\
v(i, a, \varnothing, 0_s) &= u(\varnothing) + \mu(P_{ia} - T_s(i, a)) + \beta \mathrm{EV}(\varnothing), \\
v(i, a, \kappa) &= u(i, a) + \beta(1 - \alpha(i, a))\mathrm{EV}(i, a + 1) + \beta\alpha(i, a)\mathrm{EV}(i, \bar{a}), \\
v(i, a, j, d, 1_s) &= u(j, d) - \mu\left(P_{jd} - \underline{P}_i + T_b(j, d)\right) \\
&\quad + \beta(1 - \alpha(j, d))\mathrm{EV}(j, d + 1) + \beta\alpha(j, d)\mathrm{EV}(j, \bar{a}), \\
v(i, a, j, d, 0_s) &= u(j, d) - \mu\left(P_{jd} - P_{ia} + T_s(i, a) + T_b(j, d)\right) \\
&\quad + \beta(1 - \alpha(j, d))\mathrm{EV}(j, d + 1) + \beta\alpha(j, d)\mathrm{EV}(j, \bar{a}).
\end{aligned} \tag{8}
$$

Here, $v(i, a, \varnothing, 0_s)$ is the value of selling the car on the market and not replacing it, $v(i, a, \varnothing, 1_s)$ is the value of scrapping the car and not replacing it—in both of these cases, the customer has no car to drive and ends up in the no-car state in the next period.[17] The value of keeping the existing car is given by $v(i, a, \kappa)$, and the values of trading the existing car (i, a) for a replacement car (j, d) are denoted $v(i, a, j, d, 0_s)$ and $v(i, a, j, d, 1_s)$ depending, respectively, on whether the existing car is sold or scrapped.

Combined, the value functions defined in equations (2), (5), and (7) cover the whole state space of the problem (i, a, ϵ), $i \in \{\varnothing, 1, \ldots, J\}$, $a \in \{\varnothing, 1, \ldots, \bar{a}\}$. With their corresponding choice-specific values and the general formula for the expected-value function (4), we can define the Bellman equation for the consumer choice problem as a mapping of the space of value functions $V(i, a, \epsilon)$ to itself, and standard contraction mapping arguments guarantee that V is the unique fixed point of the "Bellman operator." However, the problem can be solved in a computationally much easier fashion in terms of the "projection" $\mathrm{EV}(i, a)$, which, as we noted above, is a much lower–dimensional object, since it does not depend on the continuously distributed idiosyncratic state variables ϵ.[18] Vector EV is just a finite-dimensional vector, $\mathrm{EV} \in \mathbb{R}^{J\bar{a}+1}$, whose elements are the expected values of starting the next period in car state (i, a) $(i \in \{1, \ldots, J\}$, $a \in \{1, \ldots, \bar{a}\})$ or with no car.[19] Applying equations (4) for each element of vector EV, plugging in the expressions (2), (5), and (7) and

[17] Recall that, by our assumption, the transaction cost of selling to the scrapyard is normalized to zero, i.e., included in $\underline{P}_i$.

[18] Another, though inferior, possibility is to formulate and solve the Bellman equation in the space of choice-specific value functions, which depend on both state and choice variables and thus constitute even higher-dimensional objects than the value functions themselves. All three ways to set up the fixed-point problem are equivalent, in the sense that they lead to the same solution and the corresponding functional mappings are contractions (Ma and Stachurski 2021).

[19] Recall that because of our timing assumption it is not possible to start the period with a new car $(a = 0)$: all new cars purchased during the trading stage become 1-year-old cars by the start of the next period.

their corresponding choice-specific value functions, we derive the system of $J\bar{a}+1$ nonlinear equations

$$\mathrm{EV} = \Gamma(\mathrm{EV}), \tag{9}$$

whose solution enables us to reconstruct V and characterize optimal trading behavior. Here, Γ is the smoothed Bellman operator that constitutes a contraction mapping and hence has a unique fixed point. Further, Γ is a smooth mapping from $\mathbb{R}^{J\bar{a}+1}$ to $\mathbb{R}^{J\bar{a}+1}$, which enables us to use Newton's method in combination with the method of successive approximations to rapidly compute this unique finite-dimensional fixed point and thus solve the consumer choice problem for any set of car prices.[20]

Under the GEV distributional assumption, the integrals in equation (4) have closed-form expressions, further contributing to the computational tractability of the problem. The analytic form of these depends on the assumed nesting of choices in the decision process, as in the example illustrated in figure 1, and the implied dependency structure of the elements of ϵ. Generally, the choice-specific values $v(i, a, \cdot)$ that correspond to the bottom layer of the nodes in the tree are combined following the tree structure with the help of McFadden's (1981) log-sum (smoothed max) function $\mathcal{I}(\cdot)$, defined as

$$\mathcal{I}(\lambda, \chi_1, \ldots, \chi_n) = \lambda \log\left(\exp\frac{\chi_1}{\lambda} + \ldots + \exp\frac{\chi_n}{\lambda}\right), \tag{10}$$

where λ takes the value of the scale parameter in each particular grouping of alternatives (nest). Using the example choice tree in figure 1, we have

$$\begin{aligned}&\mathrm{EV}(i, a) \qquad (11)\\ &= \mathcal{I}(\sigma, \underbrace{\mathcal{I}(\sigma_{\mathrm{s}}, v(i, a, \varnothing, 1_{\mathrm{s}}), v(i, a, \varnothing, 0_{\mathrm{s}}))}_{\text{purge}}, \underbrace{v(i, a, \kappa)}_{\text{keep}}, \underbrace{\mathcal{I}(\sigma_{\mathrm{r}}, I_1, \ldots, I_J)}_{\text{replace}}),\end{aligned}$$

where I_j are the inclusive values of trading to a car make/model j, which are found by recursively applying the log-sum function (10) to the further nests as

$$I_j = \mathcal{I}(\sigma_j, \underbrace{I(i, a, j, 0,)}_{\text{new car}}, \ldots, I(i, a, j, \bar{a}-1)),$$

where $I(i, a, j, d) = \mathcal{I}(\sigma_{\mathrm{s}}, v(i, a, j, d, 1_{\mathrm{s}}), v(i, a, j, d, 0_{\mathrm{s}}))$ is the inclusive value of the nested scrappage decision corresponding to the choice of car j of age $d \in \{0, \ldots, \bar{a}-1\}$, including the new car. When scale parameters are equalized, $\sigma = \sigma_{\mathrm{r}} = \sigma_j = \sigma_{\mathrm{s}}$, the above nested log-sum functions

[20] See lemma L2 for details and proof.

collapse and the expected value involves a single log-sum formula where the choice-specific values correspond to all alternatives in the choice set $C(i, a)$. The expected values $\mathrm{EV}(\text{ø})$ and $\mathrm{EV}(i, \bar{a})$ involve similar closed-form expressions.

Let $\Pi(j, d, s|i, a)$ be the conditional probability of choosing a feasible alternative (j, d, s) from the choice set $C(i, a)$ by a consumer in a given car state (i, a). Under the GEV assumption, these choice probabilities take NMNL closed-form expressions that also depend on the structure of the choice tree (McFadden 1981). For example, under the choice structure in figure 1, the top-level probability of keeping the existing car follows directly from equation (11):

$$\Pi(\kappa|i, a) = \frac{\exp(v(i, a, \kappa)/\sigma)}{\exp(\mathcal{I}(\sigma_{\mathrm{s}}, v(i, a, \text{ø}, 1_{\mathrm{s}}), v(i, a, \text{ø}, 0_{\mathrm{s}}))/\sigma) + \exp(v(i, a, \kappa)/\sigma) + \exp(\mathcal{I}(\sigma_{\mathrm{r}}, I_1, \ldots, I_J)/\sigma)}. \tag{12}$$

For more complicated nested choices, such as the choice of scrapping/selling the existing car (i, a) and replacing it with car (j, d) following the choice tree in figure 1, the choice probability $\Pi(j, d, s|i, a)$ can be decomposed into products of conditional probabilities as

$$\Pi(j, d, s|i, a) = \Pi(\text{replace}|i, a) \cdot \Pi(j|\{1, \ldots, J\}, i, a) \cdot \Pi(d|j, i, a) \cdot \Pi(s|j, d, i, a), \tag{13}$$

where $\Pi(\text{replace}|i, a)$ denotes the probability of trading for some type of car, the choice probability $\Pi(j|\{1, \ldots, J\}, i, a)$ corresponds to choosing a particular make/model j conditional on having decided to replace the existing car (i, a), $\Pi(d|j, i, a)$ is the choice probability for a particular age d, and $\Pi(s|j, d, i, a)$ is the probability of the choice of scrapping the existing car (i, a) or selling it in the secondary market. These probabilities are given by the following expressions:

$$\Pi(\text{replace}|i, a) = \frac{\exp(\mathcal{I}(\sigma_{\mathrm{r}}, I_1, \ldots, I_J)/\sigma)}{\exp(\mathcal{I}(\sigma_{\mathrm{s}}, v(i, a, \text{ø}, 1_{\mathrm{s}}), v(i, a, \text{ø}, 0_{\mathrm{s}}))/\sigma) + \exp(v(i, a, \kappa)/\sigma) + \exp(\mathcal{I}(\sigma_{\mathrm{r}}, I_1, \ldots, I_J)/\sigma)}, \tag{14}$$

$$\Pi(j|\{1, \ldots, J\}, i, a) = \frac{\exp(I_j/\sigma_r)}{\exp(I_1/\sigma_{\mathrm{r}}) + \ldots + \exp(I_J/\sigma_r)}, \tag{15}$$

$$\Pi(d|j, i, a) = \frac{\exp(I(i, a, j, d)/\sigma_j)}{\exp(I(i, a, j, 0)/\sigma_j) + \ldots + \exp(I(i, a, j, \bar{a} - 1)/\sigma_j)}, \tag{16}$$

$$\Pi(s|j, d, i, a) = \frac{\exp(v(i, a, j, d, s)/\sigma_s)}{\exp(v(i, a, j, d, 1_{\mathrm{s}})/\sigma_s) + \exp(v(i, a, j, d, 0_{\mathrm{s}})/\sigma_s)}. \tag{17}$$

The remaining choice probabilities for all feasible alternatives in the choice sets (1) corresponding to all other states in the model (including having no car, $i = \text{ø}$, or having a car of terminal age $a = \bar{a}$) have similar closed-form expressions that combine the corresponding choice-specific value functions implied by the assumed nesting structure of choices.

As a result of additively separable transaction costs, and provided that the scale parameter σ_s is the same in all nests of the choice tree where the decision is relevant, the choice between scrapping and selling the existing car is independent of the choice of the replacement car and has no implications for the future periods. This implies that the scrap/sell decision is static, so $\Pi(s|j, d, i, a) = \Pi(s|i, a)$ and we can generally factor any choice probability as $\Pi(j, d, s|i, a) = \Pi(j, d|i, a)\Pi(s|i, a)$, where we refer to $\Pi(j, d|i, a)$ as the trading choice probability and $\Pi(s|i, a)$, $s \in \{1_s, 0_s\}$, as the endogenous scrappage probability. It is readily verified that the scrappage probability is the same for both purging and replacing the current car (i, a), so for all $i \in \{1, \ldots, J\}$, $a \in \{1, \ldots, \bar{a} - 1\}$, it is given by

$$\Pi(1_s|i, a) = \left(1 + \exp\left(\frac{\mu}{\sigma_s}(P_{ia} - T_s(i, a) - \underline{P}_i)\right)\right)^{-1}. \tag{18}$$

In other words, car owners choose to scrap or sell their existing car on the basis of the difference between the market price net of seller transaction cost and the scrap value of their car, conditional the marginal utility of money μ and the scale parameter σ_s.

D. *Equilibrium with Idiosyncratic Consumer Heterogeneity*

With the consumer dynamic choice problem fully described, in this section we turn to the definition of stationary equilibrium in the secondary market for automobiles. Recall that we assume that the prices of the new cars $\bar{P}_j$, $j \in \{1, \ldots, J\}$, are fixed and that the supply of new cars is infinitely elastic. Similarly, we assume that there is an infinitely elastic demand for cars for their scrap value $\underline{P}_j$, and so we also treat the scrap value of a car of each type j as fixed. We also assume that all cars have to be scrapped at the upper-bound age $\bar{a}$.

The used cars of age from $a = 1$ to $a = \bar{a} - 1$ of each make/model are traded in the secondary market. Supply and demand of these $J(\bar{a} - 1)$ tradable goods are balanced by $J(\bar{a} - 1)$ prices, which we combine into the J-block partitioned price vector

$$P = (P_1, \ldots, P_J) = ((P_{11}, \ldots, P_{1\bar{a}-1}), \ldots, (P_{J1}, \ldots, P_{J\bar{a}-1})) \in \mathbb{R}^{J(\bar{a}-1)}. \tag{19}$$

The value functions and choice probabilities derived in the previous subsection are all implicit functions of P, though we have suppressed their dependence on P so as not to overload the notation.

Let $0 \leq q_{ia} \leq 1$ denote the fraction of the unit mass of households in car state (i, a), namely, those who own a car of make/model i of age a at the start of the period before the trading phase. Let $0 \leq q_{\o} \leq 1$ be the fraction of households without a car. The *ownership distribution* vector q

summarizes the distribution of the unit mass of consumers in the economy over all possible car states:

$$q = (q_1, \dots, q_J, q_\varnothing) = ((q_{11}, \dots, q_{1\bar{a}}), \dots, (q_{J1}, \dots, q_{J\bar{a}}), q_\varnothing) \in \mathbb{R}^{J\bar{a}+1}. \tag{20}$$

The ownership distribution q is a proper probability vector (its elements sum to 1) and thus belongs to the $J\bar{a}$-dimensional unit simplex. The subvectors of the ownership distribution q_i correspond to the particular makes/models of the car and do not represent a proper distribution unless normalized. The conditional distribution (i.e., market shares) of all cars in the economy is a vector with one less element than q and can be constructed by using its first $J\bar{a}$ elements, multiplied by the normalization constant $1/(1 - q_\varnothing)$.

Though we have a continuum of consumers, we are studying an economy with a finite number of goods, so our concept of equilibrium involves the traditional approach of finding a vector P that equates supply and demand for all used cars in the secondary market. However, with a continuum of consumers, we define supply and demand in terms of the fraction of the total population of consumers who wish to sell and to buy a car of a given type and age, (j, d).[21] Let $D_{jd}(P, q)$ be the demand for make/model j cars of age d. Conditional on the ownership distribution of consumers q, for $j \in \{1, \dots, J\}$, $d \in \{1, \dots, \bar{a} - 1\}$ it is given by

$$D_{jd}(P, q) = \Pi(j, d \mid \varnothing, P) q_\varnothing + \sum_{i=1}^{J} \sum_{a=1}^{\bar{a}} \Pi(j, d \mid i, a, P) q_{ia}, \tag{21}$$

where we now include P as an argument of the choice probabilities $\Pi(j, d|\cdot, P)$ to emphasize their dependence on market prices.

The supply of used cars to the secondary market consists of those that are not kept and not scrapped. The corresponding fractions of consumers are given by the complements to the choice probability of keeping, $\Pi(\kappa|i, a, P)$, and scrapping, $\Pi(1_s|i, a, P)$. Let $S_{jd}(P, q)$ be the supply of cars of make/model j of age d, $j \in \{1, \dots, J\}$, $d \in \{1, \dots, \bar{a} - 1\}$. It is given by

$$S_{jd}(P, q) = (1 - \Pi(\kappa|j, d, P))(1 - \Pi(1_s|j, d, P)) q_{jd}. \tag{22}$$

Supply and demand of cars of all makes/models and all ages can be stacked in the same way as in price vector (19) to form the J-block partitioned vectors of demand and supply, $D(P, q)$ and $S(P, q)$. We then define the vector of *excess demand* as

$$\begin{aligned} &\mathrm{ED}(P, q) \\ &\quad = (D_{11}(P, q) - S_{11}(P, q), \dots, D_{J,\bar{a}-1}(P, q) - S_{J,\bar{a}-1}(P, q)) \in \mathbb{R}^{J(\bar{a}-1)}. \end{aligned} \tag{23}$$

[21] Each consumer in the economy chooses the alternative that maximizes their payoff conditional on their independent draw of the random component ϵ; by the law of large numbers, these deterministic individual choices aggregate to population shares that are defined in terms of q and choice probabilities.

In equilibrium, the prices equate supply and demand of used cars, resulting in zero excess demand, so equilibrium prices are a solution to the nonlinear system of $J(\bar{a}-1)$ equations given by $\mathrm{ED}(P,q)=0$ in $J(\bar{a}-1)$ unknown prices P.

Besides the market-clearing condition, which has to hold in each time period, in a stationary flow equilibrium we require the ownership distribution q to be time invariant. The evolution of q can be broken into two stages: (1) an instantaneous trading phase in the beginning of each period and (2) the rest of the period, when car utilization takes place. After the trading phase, ownership of cars changes as a result of trade between households. Also, old cars are scrapped and new cars purchased. Then, between periods t and $t+1$, the state of cars changes as they either become one period older or are involved in an accident.

To describe the two phases of the evolution of the ownership distribution, we rely on the tools from Markov chain theory and describe car ownership and state transitions using two transition probability matrices Q and $\Omega(P)$, defined below. Let $\Omega(P)$ be the $(J\bar{a}+1)\times(J\bar{a}+1)$ *trade transition probability matrix* given by

$$\Omega(P)=\begin{bmatrix} \Delta_{11}(P)+\Lambda_1(P) & \Delta_{12}(P) & \dots & \Delta_{1J}(P) & \Delta_{1ø}(P) \\ \Delta_{21}(P) & \Delta_{22}(P)+\Lambda_2(P) & \dots & \Delta_{2J}(P) & \Delta_{2ø}(P) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Delta_{J1}(P) & \Delta_{J2}(P) & \dots & \Delta_{JJ}(P)+\Lambda_J(P) & \Delta_{Jø}(P) \\ \Delta_{ø1}(P) & \Delta_{ø2}(P) & \dots & \Delta_{øJ}(P) & \Pi(ø|ø,P) \end{bmatrix}, \quad (24)$$

where the typical $\bar{a}\times\bar{a}$ block of replacing choice probabilities $\Delta_{ij}(P)$ is given by

$$\Delta_{ij}(P)=\begin{bmatrix} \Pi(j,1|i,1,P) & \dots & \Pi(j,\bar{a}-1|i,1,P) & \Pi(j,0|i,1,P) \\ \vdots & \ddots & \vdots & \vdots \\ \Pi(j,1|i,\bar{a}-1,P) & \dots & \Pi(j,\bar{a}-1|i,\bar{a}-1,P) & \Pi(j,0|i,\bar{a}-1,P) \\ \Pi(j,1|i,\bar{a},P) & \dots & \Pi(j,\bar{a}-1|i,\bar{a},P) & \Pi(j,0|i,\bar{a},P) \end{bmatrix} \quad (25)$$

and the typical $\bar{a}\times\bar{a}$ block of keeping choice probabilities $\Lambda_i(P)$ is given by

$$\Lambda_i(P)=\begin{bmatrix} \Pi(\kappa|i,1,P) & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \Pi(\kappa|i,\bar{a}-1,P) & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix}. \quad (26)$$

In addition, the bottom row and the rightmost column blocks in equation (24) are given by

$$\Delta_{\o j}(P) = [\Pi(j,1|\o,P),\ldots,\Pi(j,\bar{a}-1|\o,P,),\Pi(j,0|\o,P)],$$
$$\Delta_{i\o}(P) = \begin{bmatrix} \Pi(\o|i,1,P) \\ \vdots \\ \Pi(\o|i,\bar{a},P) \end{bmatrix}.$$

Each $\bar{a} \times \bar{a}$ block in the trade transition probability matrix $\Omega(P)$ refers to the cars of each make/model. The trade probabilities $\Pi(j, d|i, a, P)$ are strictly positive for nondegenerate GEV distributions of ϵ and form the bulk of the interior of $\Omega(P)$, and the probabilities of keeping $\Pi(\kappa|i, a, P)$ appear on the diagonal. The bottom row contains the probability of buying a car $\Pi(j, d|\o, P)$ by households who do not have one. The last column contains the probability $\Pi(\o|i, a, P)$ of choosing the no-car state, and the bottom-left corner element is the probability of remaining in the no-car state $\Pi(\o|\o, P)$.

Note that because the cars of the terminal age $\bar{a}$ cannot be traded, we use the last column of each block $\Delta_{ij}(P)$, the subtransition probabilities for trading car i for car j, to hold the choice probabilities $\Pi(j, 0|i, a, P)$ corresponding to buying a new car of type j. Similarly, the last diagonal element in the each block $\Lambda_i(P)$ is zero because the cars of age $\bar{a}$ cannot be kept and instead have to be scrapped during the trading phase.

The structure of the trade transition probability matrix $\Omega(P)$ corresponds to the block structure of the ownership distribution q in equation (20). The matrix product $q\Omega(P)$ represents the distribution of car ownership in the economy after the trading phase, assuming that all demand is satisfied (which is true in equilibrium). The result is the *posttrade holdings* distribution $q\Omega(P)$, which reflects the distribution of car holdings after the instantaneous trading phase has occurred, where new cars are delivered to households who demand them and used cars that households choose to scrap (both endogenous and exogenous scrappage) are removed from the economy. Because of the special arrangement of the columns in the trade transition probability matrix, the elements in the posttrade holdings distribution are reordered such that the fraction of owners of new cars in $q\Omega(P)$ is the last element in each subvector of length $\bar{a}$.

After the instantaneous trading phase, households own and drive their cars, and the aging and accidents of these cars is governed by the $(J\bar{a} + 1) \times (J\bar{a} + 1)$ block-diagonal stochastic matrix that we refer to as the *physical transition probability matrix* Q:

$$Q = \begin{bmatrix} Q_1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & Q_J & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}, \quad \text{where} \tag{27}$$

$$Q_j = \begin{bmatrix} 0 & 1-\alpha(j,1) & \dots & 0 & \alpha(j,1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1-\alpha(j,\bar{a}-2) & \alpha(j,\bar{a}-2) \\ 0 & 0 & \dots & 0 & 1 \\ 1-\alpha(j,0) & 0 & \dots & 0 & \alpha(j,0) \end{bmatrix}. \tag{28}$$

Each $\bar{a} \times \bar{a}$ block Q_j forms a transition probability matrix that governs the evolution of cars of make/model j. The first $\bar{a} - 1$ rows of the matrix describe the joint effect of deterministic aging and stochastic exogenous scrappage. As described in the previous section, the latter is modeled as a direct transition to the terminal age $\bar{a}$ with probability $\alpha(j, d) \in [0, 1)$, resulting in a compulsory scrappage next period. Cars of age $\bar{a} - 1$ reach the terminal age $\bar{a}$ with certainty, as both aging and exogenous scrappage lead to the same outcome $\bar{a}$. The last row governs aging and accidents of new cars. Finally, the bottom-right corner element of Q denotes the transition by households who choose the no-car state.

It follows immediately that the product of the posttrade ownership distribution and the physical transition matrix, $q\Omega(P)Q$, gives the ownership distribution in the beginning of the next period. It is then clear that the stationary ownership distribution is simply given by an invariant distribution of the matrix $\Omega(P)Q$, which the following theorem shows is unique for any P.

THEOREM 1. Let $\sigma \geq \sigma_r \geq \sigma_j \geq \sigma_s > 0$. Then for any vector of prices P, there is a unique invariant distribution q that satisfies the stationarity condition $q = q\Omega(P)Q$ and is a continuously differentiable function of P.

Proof. See appendix A.1.

Uniqueness of the stationary ownership distribution in theorem 1 simply follows from the fundamental theorem of Markov chains, once we realize that with positive GEV scale parameters the choice probabilities have full support and that therefore transition matrix $\Omega(P)Q$ is irreducible and aperiodic. However, to show differentiability of $q(P)$ with respect to P, we need a more involved argument, given in the appendix.

We are now in position to formally define and prove the existence of equilibrium in the automobile market, which extends the stationary flow equilibrium concept of Rust (1985a) to economies with positive transaction costs and discrete goods.

Definition D1 (Stationary equilibrium in the automobile market). A stationary equilibrium in the economy with a unit mass of consumers and cars of J makes/models and ages bounded above by $\bar{a}$ is given by the price vector and the ownership distribution probability vector $(P, q) \in \mathbb{R}^{J(\bar{a}-1)} \times \mathbb{R}^{J\bar{a}+1}$, such that the following conditions are satisfied:

(a) consumers follow their optimal trading strategies, which arise from the solution of the dynamic problem in equations (2)–(8);
(b) the market-clearing conditions are satisfied: the excess demand is zero;
(c) the ownership distribution q is time invariant;
(d) new cars are supplied at fixed prices $\bar{P}_i$ and scrapped at prices $\underline{P}_i$, $i \in \{1, \ldots, J\}$, infinitely elastically.

Theorem 2. The stationary equilibrium in the economy without persistent consumer heterogeneity defined in definition D1 exists. In equilibrium, the ownership distribution q satisfies the stationarity condition and the equilibrium prices P satisfy the market-clearing condition:

$$q\Omega(P)Q = q, \tag{29}$$

$$\mathrm{ED}(P, q) = 0. \tag{30}$$

Proof. See appendix A.2.

Briefly, the proof uses theorem 1 to guarantee that the equilibrium distribution q given by equation (29) is a smooth function of market prices P. The same is true for all major components of the model, namely, the expected values EV that constitute the fixed point of the Bellman operator Γ in equation (9), all choice probabilities, and the excess demand $\mathrm{ED}(P, q)$ given in equation (23). Then, given that the excess demand $\mathrm{ED}(P, q)$ in our framework is bounded to the $(-1, 1)$ hypercube, we construct a continuous map satisfying the conditions of the Brouwer fixed-point theorem, which establishes the existence of equilibrium. A number of intermediate results that the proof of theorem 2 relies on, and turn out to be very useful for our computational framework, are formulated as separate lemmas in the appendix.

It follows directly from the stationarity of the ownership distribution q that the fraction of the population without cars is also time invariant. Algebraically, it can be seen from comparing the last elements in the left- and right-hand sides of the stationarity condition (eq. [29]). Because the last column in Q has only one nonzero element, it follows that

$$\sum_{i=1}^{J}\sum_{a=1}^{\bar{a}}\Pi(\text{ø}|i, a, P)q_{ia} + \Pi(\text{ø}|\text{ø}, P)q_{\text{ø}} = q_{\text{ø}}. \tag{31}$$

Thus, a consequence of the stationarity of the equilibrium is that the fraction of consumers who demand the outside good—that is, choose not to have a car—in the left-hand side of equation (31) equals the "supply of the outside good."

Another consequence of definition D1 and the two conditions in theorem 2 is that the economy is in stationary flow equilibrium, that is, it exhibits the steady-flow property that the outflow of cars due to endogenous and exogenous scrappage equals the inflow of new cars of each make/model $j \in \{1, \ldots, J\}$. If the economy were not in a stationary flow equilibrium, there would be either a continual increase or a continual decrease in the total stock of cars of each type j in the economy over time.

THEOREM 3. In a stationary equilibrium under the conditions of theorem 2, the steady-flow property is satisfied for each car make/model $j \in \{1, \ldots, J\}$:

$$\begin{aligned}&\underbrace{\sum_{a=1}^{\bar{a}-1}\Pi(1_s|j, a, P)(1 - \Pi(\kappa|j, a, P))q_{ja} + q_{j\bar{a}}}_{\text{outflow of scrapped cars of make/model } j}\\ &= \underbrace{\sum_{i=1}^{J}\sum_{a=1}^{\bar{a}}\Pi(j, 0|i, a, P)q_{ia} + \Pi(j, 0|\text{ø}, P)q_{\text{ø}}.}_{\text{inflow of new cars of make/model } j}\end{aligned} \tag{32}$$

Proof. See appendix A.3.

With the result of theorem 3, it is clear that definition D1 defines a stationary flow equilibrium in the market of new and used cars. Note that theorem 2 establishes only the existence of the stationary flow equilibrium but is silent about its uniqueness. Uniqueness of the equilibrium ownership distribution under any market prices P is established by theorem 1; however, we have been unable to find high-level conditions that guarantee uniqueness of the equilibrium price vector P. Despite this, we have computed many equilibrium solutions and have never encountered an issue of multiplicity of equilibria for a variety of utility function specifications and parameter values. Thus, we conjecture that there are conditions under which a stationary flow equilibrium not only exists but is unique.

E. Numerical Implementation

The key to success for our numerical implementation is the possibility to use efficient Newton-based methods for finding the fixed point of the smooth Bellman operator Γ in the dynamic programming part of model (9) and, when solving the nonlinear system of equations (30), to find the equilibrium price vector P.

As is well known, Newton's method has a quadratic convergence rate when initiated from a sufficiently close starting point in a domain of attraction of the solution. In the dynamic programming part of the algorithm, we rely on the globally convergent method of successive approximations before switching to Newton-Kantorovich iterations, in the same way as it is done in the nested fixed-point (NFXP) estimator in Rust (1994). In the equilibrium price search, we initialize the Newton solver at the equilibrium prices of a similar model without consumer heterogeneity and transaction costs that can be computed as a solution to a system of linear equations, as shown in appendix B.

Appendix A contains several lemmas that establish the prerequisite smoothness properties and specify the analytical formulas for the required gradients. We first show that EV is a smooth function of P implicitly defined by the fixed-point condition, $\mathrm{EV} = \Gamma(\mathrm{EV}, P)$. It follows via the chain rule that all value functions and all choice probabilities are also continuously differentiable in P. Differentiability of excess-demand function $\mathrm{ED}(P, q)$ in P for any q immediately follows.

Uniqueness and differentiability of $q(P)$ as an implicit function of P are established by theorem 1. Using the chain rule, we obtain the Jacobian matrix for $\mathrm{ED}(P, q(P))$ and solve the market-clearing conditions $\mathrm{ED}(P, q(P)) = 0$ as a system of $J(\bar{a} - 1)$ nonlinear equations in prices. We can use Newton's method for this, but also using the chain rule to compute the total derivative of $\mathrm{ED}(P, q(P))$ with respect to P, which for lack of better notation we denote by $\nabla_P \mathrm{ED}(P, q(P))$. The computational algorithm involves the following steps:

1. for a given vector of market prices P, solve the Bellman equation (9) for the fixed point $\mathrm{EV}(P)$ using Newton-Kantorovich iterations;
2. compute choice probabilities and form the trading and physical transition probability matrices $\Omega(P)$ and Q;
3. compute the ownership distribution $q(P)$ as an invariant distribution of $\Omega(P)Q$;
4. calculate excess demand and update prices via Newton's method $P' = P - (\nabla_P \mathrm{ED}(P, q(P)))^{-1}\mathrm{ED}(P, q(P))$;
5. exit if the convergence criterion for $\mathrm{ED}(P', q(P')) = 0$ is satisfied; otherwise, replace P with P' and return to step 1.

Thus, it is possible to use the gradient-based Newton method in all steps of our numerical implementation, resulting in a fast algorithm for computing the stationary flow equilibrium in the automobile market.[22] Given

[22] Code and data are available in a zip file.

how quickly the equilibria can be computed for various parameter values and specifications of the model, it can be nested within other algorithms, such as the maximum likelihood estimator that we develop in section V.

IV. Equilibrium with Persistent Consumer Heterogeneity

The economy with idiosyncratically heterogeneous households analyzed in the previous section is sufficient to support trade in the presence of transaction costs. However, when we allow for more persistent consumer heterogeneity, there are even larger gains from trade that result in equilibrium sorting of different ages of cars by different types of households.

A. Time-Invariant Heterogeneity

We start with the case superficially opposite to the idiosyncratic heterogeneity considered in the previous section, namely, the economy with permanent household types. However, the key to tractability of our framework is that we introduce persistent heterogeneity in addition to the idiosyncratic heterogeneity due to GEV random components. Later, we add an intermediate form of time-varying heterogeneity as well, which in the end gives us the best of both worlds: more realistic, flexible forms of consumer heterogeneity while retaining the elegance and computational tractability from the nested logit GEV specification introduced in the previous section. It is essentially a form of "mixed logit" that has been so useful in empirical work.

We introduce the symbol $\tau \in \{1, \ldots, N_\tau\}$ to denote households of different permanent types and denote as f_τ the fraction of households of type τ. The structure of the household decision problem (eqq. [2]–[5]) is identical for all households but, as a result of arbitrary differences in preferences, the solution becomes type-specific. Demand and supply from all types are aggregated in market equilibrium, and households endogenously specialize in their holdings of different makes/models and ages of automobiles in response to market prices and differences in their preferences. We allow for essentially unlimited flexibility in how the preferences of households of different types τ differ, as long as each type conforms to the general structure introduced in section III.A.

Let $u_\tau(i, a)$ be the utility for owning a car of make/model i and age a by households of type τ. Solving the Bellman equations (2)–(5) N_τ times, we obtain the decision-specific value functions $v_\tau(i, a, j, d, s)$, expected-value functions $\mathrm{EV}_\tau(j, d)$, and choice probabilities $\Pi_\tau(j, d, s|i, a, P)$ for each household type τ for all states $a \in \{1, \ldots, \bar{a}, \varnothing\}$, $i \in \{1, \ldots, J\}$ and all choices in $C(i, a)$ defined in equation (1).

Denote as q_τ the ownership distribution of type-τ households, which we define similarly to equation (20) as a proper stochastic vector in $\mathbb{R}^{J\bar{a}+1}$. The full ownership distribution in the economy can be written as $q = (q_1 f_1, \ldots, q_{N_\tau} f_{N_\tau}) \in \mathbb{R}^{N_\tau(J\bar{a}+1)}$, but it is sufficient to work with its type-specific subvectors. Repeating the definitions of supply and demand from section III.D, we can derive τ type-specific excess-demand functions $\mathrm{ED}_\tau(P, q_\tau)$.

To extend definition D1 for stationary equilibrium to the economy with permanent household types, note that conditions a and d remain unchanged and that we have to modify only the market-clearing and stationarity conditions. Bearing in mind that trade is allowed between household types, $\mathrm{ED}_\tau(P, q_\tau)$ need not be zero for each τ; instead, the integrated demand has to clear, leading to the following condition:

$$\mathrm{ED}(P, q) = \sum_{\tau=1}^{N_\tau} \mathrm{ED}_\tau(P, q_\tau) f_\tau = 0. \tag{33}$$

Further, with multiple types of households we require stationarity of the ownership distribution for each household type,

$$q_\tau = q_\tau \Omega_\tau(P) Q_\tau, \qquad \forall\ \tau, \tag{34}$$

and thus stationarity of the ownership distribution in the whole economy. By making the aging transition probability matrix Q_τ household-type specific in equation (34), we allow scrappage probabilities to vary by household type.

Theorem 4. The stationary equilibrium in the economy with $\tau \in \{1, \ldots, N_\tau\}$ time-invariant household types in addition to idiosyncratic heterogeneity—see definition D1—exists. In equilibrium, the ownership distribution $q \in \mathbb{R}^{N_\tau(J\bar{a}+1)}$ is composed of type shares–weighted subvectors q_τ, each of which satisfies the stationarity condition (34), and equilibrium prices P satisfy the market-clearing condition (33). The steady-flow property of the equilibrium continues to hold.

We omit the proof of theorem 4 because it follows from straightforward modifications of the proofs of theorems 1–3 of section III. Fully detailed proofs are available on request. But to provide a rough idea of how the proof works, first note that we can use theorem 1 to prove that, for each consumer type τ, there is a unique invariant distribution $q_\tau = q_\tau \Omega_\tau(P) Q_\tau$ and that this q_τ is continuously differentiable function of P. Then, it follows from lemma L2 that $\mathrm{ED}(P, q)$, given in equation (33), is a smooth function of P. Then, following the proof of theorem 2, we can appeal to the Brouwer fixed-point theorem to prove the existence of an equilibrium with persistent heterogeneity. Finally, the steady-flow equilibrium condition must also hold (otherwise, the stock of cars would be continually increasing or decreasing over time), and this result can be proven via a proof similar to that for theorem 3.

Overall, the stationary equilibrium is exactly as described in section III: the only additional step is aggregation of τ-specific excess demand. Otherwise, a price vector P sets excess demand to zero per equation (33), subject to the constraint that the ownership distributions for all types τ are stationary per equation (34). Moreover, most of the theoretical results from section III apply directly for each household type one by one, with the key exception that excess demand need not be zero type by type; that is, $\mathrm{ED}_\tau(P, q_\tau)$ may not necessarily equal zero for each type τ even though aggregate excess demand must be zero, $\mathrm{ED}(P, q) = 0$.

The computational approach from section III.E does not change much at all: we compute the equilibrium by first solving equation (34) for $q_\tau(P)$, which is a smooth implicit function of P, and repeat this calculation N_τ times for every τ. Then, the functions $q_\tau(P)$ are jointly substituted into the excess demand and the corresponding nonlinear system of equations in prices is solved, again with Newton's method. Therefore, as one part of the solution algorithm is repeated for each household type and the other does not depend on N_τ, we conclude that the solver is only linearly more computationally costly.

B. *Time-Varying and Hybrid Heterogeneity*

Now consider the case of time-varying types. Consider an exogenous Markov process with state space $\mathcal{Y}$ and transition density $\rho(y_{t+1}|y_t)$ for some time-varying variable y_t that is household specific (e.g., income) and evolves independently for each household. Assuming that y enters the utility for cars, $u(j, a, y)$ or the marginal utility of money, $\mu(y)$, the dynamic problem becomes more complex, since the household now has to account for stochastic variation in the y_t state variable when considering his/her optimal car-trading strategy. An unexpected negative income shock may induce the household to keep their older car and delay replacement, or, conversely, a positive income shock may induce them to trade their existing car and buy a new one or upgrade to a different car make/model.

The Bellman equations (2)–(7), which describe the optimal trading strategy, need to be altered to account for the extra state variable y_t and need to include an extra integration with respect to the transition density $\rho(y_{t+1}|y_t)$.[23] Assume that $\{y_t\}$ is ergodic and has an invariant distribution $\lambda(y)$ satisfying

$$\lambda(y') = \int_{\mathcal{Y}} \rho(y'|y)\lambda(y)\, dy. \tag{35}$$

[23] Since these extensions are straightforward, we omit the Bellman equations to save space.

In the case when $\{y_t\}$ is a finite-state Markov chain instead of a continuous state process, ρ is simply a transition probability matrix and equation (35) can be written as $\lambda = \lambda\rho$, where λ is a probability distribution vector completely analogous to the distribution of permanent types $(f_1, \ldots, f_{N_\tau})$ in the previous section.

With time-varying heterogeneity, the value functions, choice probabilities, and ownership transition probability matrices are indexed by y similarly to the way they were indexed by τ in the time-invariant case. Let q_y denote the ownership distribution conditional on y, which we again define similarly to equation (20) as a proper stochastic vector in $\mathbb{R}^{J\bar{a}+1}$. Its typical element q_{jay} is the share of households who own car type j of age a while in income state y.

Continuing with the analogy, let $\mathrm{ED}_y(P, q_y)$ denote the excess-demand function for a household whose income state is y. Though y changes over time for different households, there is a stationary cross-sectional distribution of y, given by the invariant density $\lambda(y)$ defined above, and there is a stationary joint density of car ownership states and y, given by $q_y(P)$. So, to extend definition D1 for stationary equilibrium to the economy with time-variant household types, we modify the market-clearing condition to

$$\int_{\mathcal{Y}} \mathrm{ED}_y(P, q_y)\lambda(y)\, dy = 0, \tag{36}$$

which is still the system of $J(\bar{a} - 1)$ nonlinear equations in prices and ownership distribution. However, the latter is pinpointed by the modified stationarity condition that takes into account the stochastic evolution of types according to the transition density $\rho(y'|y)$, namely,

$$q_{y'} = \int_{\mathcal{Y}} q_y \Omega_y(P) Q_y \rho(y'|y)\lambda(y)\, dy. \tag{37}$$

THEOREM 5. The stationary equilibrium in the economy with time-varying household types given by an exogenous ergodic Markov process $\{y_t\} \in \mathcal{Y}$ with transition density $\rho(y'|y)$ and stationary distribution $\lambda(y)$, defined in definition D1, exists. In equilibrium, the joint ownership-type distribution is given by the $\lambda(y)$ and q_y that satisfy the stationarity condition (eq. [37]), and the equilibrium prices P satisfy the market-clearing condition (eq. [36]). The steady-flow property of the equilibrium continues to hold.

This proof is also very similar to the proof of theorem 4 and is omitted for brevity, though a full detailed proof is available on request. It is also possible to layer combinations of time-invariant and time-varying heterogeneity, and these cases can be handled by combining theorems 4 and 5. For example, we could have a finite number of types τ with different transition densities $\rho_\tau(y'|y)$. We can extend the equilibrium conditions by integrating excess demand $\mathrm{ED}_{\tau y}(P, q_{\tau y})$ over all y for each type τ and then sum over types. This requires computing stationary ownership distributions $q_{\tau y}$ for each (τ, y) combination using a τ-specific analog of equation (37) and stationary distributions λ_τ for each time-invariant type τ. We can then substitute these invariant distributions (taken as smooth implicit functions of P) into the formula for excess demand $\mathrm{ED}_{\tau y}(P, q_{\tau y})$ and compute the equilibrium prices by searching for a vector P that solves the system of nonlinear equations

$$\mathrm{ED}(P, q) = \sum_{\tau=1}^{N_\tau} f_\tau \int_y \mathrm{ED}_{\tau y}(P, q_{\tau y})\lambda_\tau(y)\, dy = 0 \quad (38)$$

formed by integrating excess demand over all time-varying and time-invariant household types.

C. *Illustrative Example: Sorting in Stationary Flow Equilibrium*

Figure 2 illustrates the stationary flow equilibrium in a heterogeneous-agent economy with two permanent types of households who differ in their marginal utility of money, $\mu_2 = 0.3 > 0.1 = \mu_1$. The households who have a lower marginal utility of money are the rich households in this economy. The utility of the outside good is set to 0 for both households types. We assume that this economy has 50% rich and 50% poor households.

In this example, we collapse the GEV structure of random components to a simple extreme-value EV1 distribution with common scale parameter $\sigma = 1$. Consumers also have the same discount factor $\beta = .95$, and the utility function is $u(a) = 10 - 0.5a$. There is a single car make/model $J = 1$ traded in this economy, with new-car price $\bar{P} = 200$ and scrap value $\underline{P} = 1$.

We also illustrate the effect of transaction costs by computing an equilibrium with buyer-side transaction costs of $T_\mathrm{b}(P, d) = 0$ and $T_\mathrm{s}(P, a) = 0$ (low transaction cost), as well as an equilibrium with high transaction costs, $T_\mathrm{b}(P, d) = 10$.

It seems reasonable to conjecture that in equilibrium a hand-me-down chain will emerge in which the rich are more likely to buy brand new cars whereas the poor households will buy the used cars previously owned by the rich. However, it is not clear a priori what relative fractions of the two types of households will select into what fractions of the new and used cars and which fraction will choose to have no car at all. These questions and the effect of transaction costs on holdings can be answered by numerically computing the equilibrium prices and ownership distributions.

Figures 2 and 3 plot prices and holdings for two equilibria corresponding to the low- and high-transaction-cost cases, respectively. For comparison, we

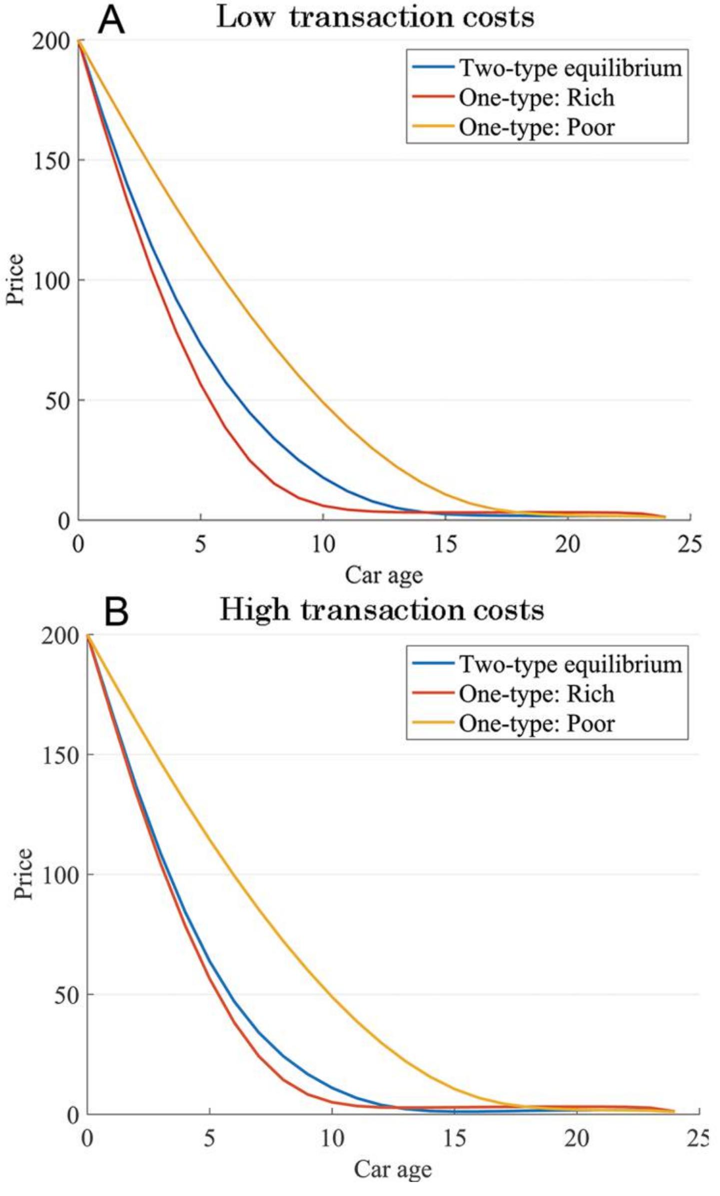

FIG. 2.—Equilibrium price functions in a two-household-type economy. Both panels show the equilibrium price functions of the heterogeneous-agent economy as well as homogeneous economies without transaction costs where all households are of either type.

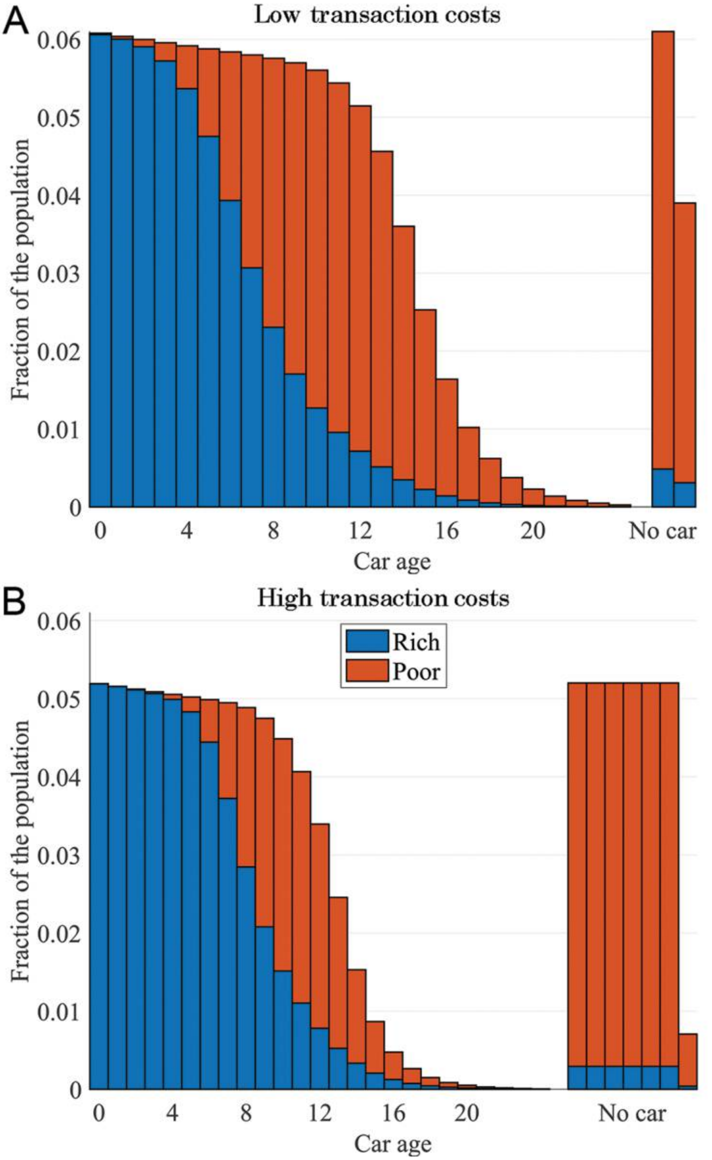

Fig. 3.—Equilibrium ownership in a two-household-type economy. The total area of the bars in each panel adds up to 1, presenting both the distribution of households over car ages and the share of outside good. There are equal numbers of households of each type. Both panels show the posttrade ownership distributions. Since the density is much higher in the no-car state than in each car-by-age category, we have split the no-car bar into several bars placed horizontally next to each other.

also plot the prices that arise in the economies with single household types, one where all households have high marginal utility of money and one where it is low. When transaction costs are high, equilibrium prices are closer to prices in the one-type economy where all consumers are rich, because many poor consumers are now driven out of the market. Moreover, cars are scrapped earlier. Thus, the higher transaction cost limits the gains from trade and partially "kills off" the market for used cars.

There is clear evidence of sorting of households into the ages of the car in figure 3*A*, which shows the posttrade ownership distribution $q\Omega(P)$. The rich households hold the newer cars and in particular are much more likely to buy new cars than the poor households. In addition, the fraction of poor households who do not to own a car is much higher. Overall, we see that poor households are driven out of the market to a much larger extent than rich households when transaction costs increase.

These findings confirm our conjecture that a "hand-me-down chain" arises endogenously in equilibrium, created by type-specific specialization in holdings that facilitates gains from trade between the two types of households. The rich households buy brand new cars, hold them for several years, and then sell them to poor households, who also hold them for several years, trading the cars over a succession of poor owners until the cars are scrapped. Thus, in this example the rich households are net suppliers of older cars, and the poor households are net demanders of older cars. Most of the trade between rich and poor households occurs for cars of roughly middle age: the rich supply their middle-aged cars to the poor households, and the market clears in aggregate, but not for each type.

V. Identification and Structural Estimation

We have introduced a dynamic model of trade in automobiles that allows for rich specifications of observed and unobserved consumer heterogeneity whose equilibrium can be rapidly computed. By embedding the equilibrium solver into other algorithms, we can therefore extend our framework in multiple directions, including using the model in empirical applications. In this section, we develop the doubly nested fixed-point (DNFXP) algorithm to implement the maximum likelihood estimator for our modeling framework.[24] We also provide a discussion on model identification, including the case when the market prices are not observed.

[24] Another natural extension of our framework is to include the equilibrium in the primary market of automobiles. However, given our empirical application in sec. VI to the Danish automobile market, where all new cars are imported, we defer this extension to a future paper.

A. Maximum Likelihood Estimation

Let θ be the vector of parameters characterizing consumer preferences, car transaction costs, and car-type-specific accident rates that we wish to estimate. The solution of the model, including the equilibrium prices, quantities, and choice probabilities, then consists of implicit functions of the parameter vector θ.

Suppose that we observe types, states, choices, and accidents for a random sample of consumers (households) indexed by $h \in \{1, \ldots, N_H\}$, where each consumer is observed over T_h separate time periods, which may or may not form a consecutive sequence. In other words, we may have a balanced or unbalanced panel or even a cross section, if $T_h = 1$ for all h. Denote the total number of observations $N_{HT} = \sum_{h=1}^{H} T_h$. Let τ_h denote the observed type of consumer h.[25] Let x_{ht} denote the observed predecision state of consumer h in period t, so $x_{ht} = (i_{ht}, a_{ht})$ if the consumer owned a car of type i_{ht} and age a_{ht} at the start of period, before making any decision about trading, and similarly $x_{ht} = \o$ if the consumer enters the period with the outside good. Let $(c_{ht}, s_{ht}) \in C(x_{ht})$ denote the consumer's decision from the choice set $C(\cdot)$ given in equation (1), where c_{ht} denotes the choice of whether to keep, trade, or purge the current car and s_{ht} denotes the voluntary scrappage decision, which is relevant only when $x_{nt} \neq \o$ and $c_{nt} \neq \kappa$. Finally, let $z_{ht} \in \{0, 1\}$ denote whether household h experienced a total-loss accident leading to an exogenous scrappage of their car during period t. Note that, given the structure of the state transitions, the next period state x_{ht+1} is fully determined by $(x_{ht}, c_{ht}, s_{ht}, z_{ht})$, and we can treat it as a deterministic function of the current state and decision, $x_{ht+1} = f(x_{ht}, c_{ht}, z_{ht})$. This implies that, for any decision involving owning a car $c_{ht} \neq \o$, knowledge of (x_{ht}, c_{ht}) enables us to predict the probability of x_{ht+1} in terms of the accident probability $\alpha(c_{ht})$, so we can write a conditional density for (x_{ht+1}, s_{ht}), given x_{ht}, as

$$\begin{aligned}\pi(x_{ht+1}|x_{ht}, \tau_h)\pi_s(s_{ht}|x_{ht}, \tau_h) &\equiv \pi(f(x_{ht}, c_{ht}, z_{ht})|x_{ht}, \tau_h)\pi_s(s_{ht}|x_{ht}, \tau_h)\\ &= (I\{z_{ht} = 1\}\alpha(c_{ht}) \qquad\qquad (39)\\ &\quad + I\{z_{ht} = 0\}(1 - \alpha(c_{ht})))\Pi_{\tau_h}(c_{ht}|x_{ht})\Pi_{\tau_h}(s_{ht}|x_{ht}).\end{aligned}$$

The probability $\pi(x_{t+1}|x_t, \tau)$ is simply an element of the matrix $\Omega_\tau(P)Q$ from the equilibrium condition equation (29) of theorem 2, and the accident probabilities $\alpha(\cdot)$ are those used in the definition of the Q matrix in equation (27).[26] The conditional choice probability $\Pi_\tau(c_{ht}|x_{ht})\Pi_\tau(s_{ht}|x_{ht})$

[25] We focus on the case of time-invariant consumer heterogeneity. The cases of time-varying and mixed consumer heterogeneity are straightforward generalizations.

[26] Note that if $c_{ht} = \o$, then the accident probability is not relevant and $\pi(x_{ht+1}|x_{ht})\pi_s(s_{ht}|x_{ht})$ reduces to $\Pi(c_{ht}|x_{ht})\Pi_s(s_{ht}|x_{ht})$, and if $x_{ht} = \o$ or $c_{ht} = \kappa$, then the scrappage indicator is not relevant, so we can define $\pi_s(s_{ht}|x_{ht}) = 1$ in such cases.

on the right-hand side of equation (39) is composed of the type-specific analogs of the choice probabilities given in equations (13) and (18) of section III.C.

The (average) log likelihood collects contributions from state transitions and the choice probabilities similar to Rust (1987):

$$L_H(\theta) = \frac{1}{N_{HT}} \sum_{h=1}^{H} \sum_{t=1}^{T_h} [\log(\pi(x_{ht}|x_{ht-1}, \tau_h, \theta)) + \log(\pi_s(s_{ht}|x_{ht-1}, \tau_h, \theta))], \tag{40}$$

where the first term of $L_H(\theta)$ represents information from trading decisions, while the second term reflects information from scrappage decisions. We show below, in theorem 6, that the information on scrappage decisions is crucial for identification of the key marginal utility of money parameters $\{\mu_\tau\}$ as well as transaction costs $\{T_s(i, a), T_b(i, a)\}$, even when secondary prices P_{ia} are unobserved by the econometrician. Thus, information on scrappage of cars is key to the identification of the model.

In applications with a large sample size, we can speed up the computation of the likelihood by avoiding the summation over h and t in equation (40) and instead using counts of similar observations. Let $N_{x'x\tau} = \sum_{h=1}^{H}\sum_{t=1}^{T_h} I\{x_{ht} = x', x_{ht} = x, \tau_h = \tau\}$ denote the total number of type-τ consumers in state x who transit to state x'. We have $N_{HT} = \sum_\tau \sum_x \sum_{x'} N_{x'x\tau}$. Similarly, let $N_{sx\tau} = \sum_{h=1}^{H}\sum_{t=1}^{T_h} I\{s_{ht} = s, x_{h\tau} = x, \tau_h = \tau\}$ be the total number of observations for scrappage choice $s \in \{0_s, 1_s\}$. In our empirical application to Denmark, H is the total number of households, and using repeated cross sections of the entire population over $T = 12$ years, we end up with $N_{HT} = 39$ million observations. By the ergodic law of large numbers, we have, with probability 1,

$$\begin{aligned} \lim_{N_{HT}\to\infty} \frac{N_{x'x\tau}}{N_{HT}} &= \pi(x'|x, \tau) q_\tau(x) f_\tau, \\ \lim_{N_{HT}\to\infty} \frac{N_{sx\tau}}{N_{HT}} &= \pi_s(s|x, \tau) q_\tau(x) f_\tau. \end{aligned} \tag{41}$$

Thus, the large-sample limit of the cell-based likelihood function $L_H(\theta)$ is

$$\lim_{N_{HT}\to\infty} L_H(\theta)$$

$$= \sum_\tau \sum_x \left[\sum_{x'} \log(\pi(x'|x, \tau, \theta)) \pi(x'|x, \tau) + \sum_s \log(\pi_s(s|x, \tau, \theta)) \pi_s(s|x) \right] q_\tau(x) f_\tau, \tag{42}$$

so our data can be condensed into a much smaller number of empirical transition probabilities and market shares, $\{\pi(x'|x, \tau), \pi_s(s|x, \tau), q_\tau(x), f_\tau\}$, along with the consumer-type distribution f_τ in the population. It is far faster to evaluate the likelihood in equation (42) than to sum over 39 million individual observations, as in equation (40).

In our empirical application in section VI, we observe scrappage outcomes but not voluntary scrappage decisions, because we do not observe accidents, including ones that lead to the forced scrappage of a vehicle. This implies that we cannot always observe the household's predecision state x_{ht} needed to evaluate the full-information likelihood $L_N(\theta)$ given in equation (40) above. Nevertheless, we show in appendix C that we can still identify the accident probabilities $\alpha(i, a)$ for different car types and ages even when accidents are unobserved. We do this via a marginal likelihood function that integrates out with respect to accidents by calculating the probability of a scrappage outcome as the sum of two probabilities: (1) "exogenous scrappages" due to accidents, which occur with probability $\alpha(i, a)$, and (2) "endogenous scrappages" due to a voluntary choice by the owner, which occurs with probability $(1 - \alpha(i, a))\Pi_\tau(1_s|i, a)$. Technically, the likelihood function can be expressed using a transition probability for the postdecision state augmented with the scrappage outcome, which we denote by ζ_t, where $\zeta_t = 1$ if the car owned before trading is scrapped (either voluntarily or exogenously as a result of an accident) and $\zeta_t = 0$ if the car x_t is not scrapped. Define δ_t as the vector consisting of the postdecision state $c_t = (j_t, d_t)$ (i.e., the car the consumer has after the instantaneous trading decision) and the scrappage outcome ζ_t on the previous car $x_t = (i_t, a_t)$ (i.e., the predecision state). Thus, we have $\delta_t = (c_t, \zeta_t)$, which equals $\delta_t = (j_t, d_t, \zeta_t)$ if the consumer traded car $x_t = (i_t, a_t)$ for car $c_t = (j_t, d_t)$ at time t, and in our Danish data, δ_t is always observed.[27] In appendix C, we derive a transition probability $\pi(\delta'|\delta, \tau, \theta)$ and show that a likelihood function using these observed transition probabilities succeeds in identifying all parameters θ. This likelihood can also be written in both standard and cell forms, as in equations (40) and (42).

Note that the formulation of the likelihood function above fully imposes the equilibrium constraints and relies on the equilibrium prices $P(\theta)$ implied by the model rather than on observed prices at the secondary market. As shown below, conditional on the model being well specified, these equilibrium constraints allow us to identify the structural parameters θ even when secondary prices P are unobserved. We can thus consistently estimate the true parameter θ^* using the maximum likelihood estimator $\hat{\theta}$ defined by $\hat{\theta}_H = \text{argmax}_{\theta \in \Theta} L_H(\theta)$.[28]

[27] If the predecision state is $x_t = \emptyset$ or the observed decision $c_t = \kappa$, then the scrappage outcome is not relevant, and δ_t simply equals x_t in such cases.

[28] When prices P are fully observed, we can do maximum likelihood subject to the constraint that $P(\theta) = P$, which results in a more efficient estimator because of the imposition of the constraint. However, even in situations where prices P are actually observed, researchers may not want to estimate subject to this constraint but rather to rely on the model's ability to identify

B. DNFXP Algorithm

To implement $\hat{\theta}$, we adopt a full-solution approach similar to the NFXP algorithm in Rust (1987), but where the additional equilibrium constraints require an additional nested loop to compute equilibrium prices. We refer to this as the DNFXP algorithm: while searching over the parameter space, DNFXP invokes the solution algorithm described in sections III and IV to compute the equilibrium prices, quantities, and choice probabilities necessary to compute the likelihood function. This requires an additional nested loop for calculation of the equilibrium prices P after the expected-value functions $\mathrm{EV}_\tau(P, \theta)$, and the corresponding choice probabilities are found as a result of solving consumers' dynamic programming problems.

While this seems like a daunting computational task, our extensive use of gradient-based methods and principles of back propagation to compute them makes it fast and robust. For example, computing the gradient $\nabla_\theta L(\theta)$ of the likelihood function incorporates, through the chain rule of calculus, the computed gradient of the equilibrium prices $\nabla_\theta P(\theta)$, which in turn incorporates $\nabla_\theta \mathrm{EV}_\tau(P, \theta)$. The gradients needed at the outer loop (likelihood maximization) of the algorithms are produced as by-products of solving the model at the inner loops (equilibrium and Bellman optimality).[29] This enables us to use fast implementations of quasi-Newton algorithms such as the BHHH algorithm of Berndt et al. (1974) with accurate analytic gradients, without the need to compute the Hessian of $L(\theta)$, which is even more tedious. Together, this makes DNFXP feasible even on ordinary laptop computers.

C. Model Identification

In this section, we establish the identification of the structural parameters θ of our equilibrium model under the most unrestricted parametric specification. Following the description in sections III and IV, under the least restriction on the parameters of the model vector θ is composed of β and $\{u_\tau(i, a), \mu_\tau, T_b(i, a), T_s(i, a)\}$, for all $a \in \{0, \dots, \bar{a} - 1\}$, $i \in \{1, \dots, J\}$, and $\tau \in \{1, \dots, N_\tau\}$ and thus has $1 + 2J(\bar{a} - 1) + N_\tau(J\bar{a} + 1)$ elements.[30]

equilibrium prices using only micro data on car ownership transitions. The predicted prices $P(\hat{\theta})$ constitute some of the strong overidentifying restrictions of the model, so a Wald or likelihood ratio test of the hypothesis $H_0 : P(\theta^*) = P$ is likely to be a powerful test of the combined hypotheses of stationarity, individual optimality, and market equilibrium.

[29] Because we extensively use the chain rule of calculus, our framework is compatible with many different specifications of preferences and can easily accommodate additional structural parameters. The computational cost of adding parameters is small—there is practically no additional time spent to compute additional derivatives, and so the run time for a single evaluation of the likelihood function and its gradient hardly changes.

[30] The accident probabilities $\{\alpha(i, a)\}$ can also be treated as structural parameters, as in our empirical application in sec. VI.

In a theoretical analysis of identification, the choice probabilities $\{\pi(x'|x, \tau), \pi_s(s|x, \tau)\}$ for each observed consumer type τ and the implied market shares $q_\tau(x)$, as well as type proportions f_τ, are treated as known. The choice probabilities constitute the reduced-form objects of the model. Subject to an arbitrary location and scale normalization on utilities, the model is identified if only one vector of structural parameters θ^* implies the reduced-form probabilities.

A necessary condition for identification is that we have more observed "moments" (i.e., probabilities) than parameters being estimated. It is straightforward to show that we cannot identify the structural parameters using the aggregate market share data alone. This is simply due to the fact that such data provide only $J\bar{a}$ independent moments, but even after a location and scale normalization, the unrestricted specification for θ has far more parameters. Using data on consumer-type-specific market shares may suffice, depending on the particular application. However, we have far more data than market shares in ownership transitions, which constitute a total of $J\bar{a}(J\bar{a} + 1)$ independent moments for each of the consumer types, and an additional $J(\bar{a} - 1)$ moments for the age-specific scrappage probabilities for each car type. Thus, with microdata or cell count data as described in section V.A, we typically have vastly many more moments than structural parameters in our model.[31]

There are two key reasons why we are able to identify μ_τ even when secondary prices P are unobserved. First, we do observe prices of new cars $\bar{P}_i$ as well as scrap values $\underline{P}_i$ for each make/model. Second, the quasi-linear structure of preferences, together with the assumption that the market is in equilibrium, imposes strong "cross-equation restrictions" on car ownership transitions, holdings, and prices. In particular, $P_{ia}(\theta)$ is a nonlinear function of model parameters, so the price terms $\mu_\tau P_{ia}(\theta)$ and the car utilities $u_\tau(j, a)$ will not be collinear. Intuitively, we can identify μ_τ by observing where consumers endogenously "locate" in the "hand-me-down chain" in the car market—richer consumers (those with lower μ_τ) are more likely to buy newer cars and more expensive brands, whereas poorer consumers are more likely to buy older cars and less expensive brands. Formally, identification of $\{\mu_\tau\}$ (up to a scale normalization) and transaction costs under additional identifying assumptions is provided by the following theorem.

Theorem 6. Provided that the market is in equilibrium and under the assumptions mostly laid out in sections III and IV, namely,

[31] In the empirical analysis in sec. VI, we allow for $N_\tau = 8$ (observed) household types and $J = 4$ car types and assume $\bar{a} = 25$. This results in 1,001 parameters in the unrestricted parameterization, which cannot be estimated using only 96 independent market shares or even 768 type-specific market shares. The total number of "moments" in the ownership state transitions is, however, 80,800. Nevertheless, as described in sec. VI, we actually use a parsimonious restricted specification with only 131 parameters.

(a) random components ϵ of the utilities follow the GEV distribution,
(b) prices of new cars $\bar{P}_i$ and scrap values $\underline{P}_i$ are observed for each car make/model i,
(c) transaction costs of the new car buyer, as well as transaction costs on the seller side, for all $i \in \{1, \ldots, J\}$, $a \in \{1, \ldots, \bar{a}\}$ are included in the prices, $T_{\mathrm{b}}(i, 0) = T_{\mathrm{s}}(i, a) = 0$,

the equilibrium prices P_{ia} and transaction costs $T_{\mathrm{b}}(i, a)$ are point identified for all tradable cars $i \in \{1, \ldots, J\}$, $a \in \{1 \ldots, \bar{a} - 1\}$, and the marginal utilities of money μ_τ are point identified (up to a scale) for all consumer types $\tau \in \{1, \ldots, N_\tau\}$.

Proof. See appendix A.4.

The proof of theorem 6 relies on the inversion theorem of Hotz and Miller (1993), namely, the version of it for the GEV distributed shocks in lemma 2 in Arcidiacono and Miller (2011). We can then treat the choice-specific value function differences as data and, using the quasi-linear structure of the utility function, are able to pick out its different parts by considering value contrasts for various pairs of choices. Under assumption c that the seller-side costs are "passed on" to the buyer, all the monetary components in consumer preferences are identified.[32]

Once all the monetary components in consumer preferences are identified per theorem 6, our model reverts to the standard dynamic discrete-choice setting, where the remaining structural parameters $\{\beta, u_\tau(i, a), \}$, $a \in \{0, \ldots, \bar{a} - 1\}$, $i \in \{1, \ldots, J\}$, can be nonparametrically identified under the standard normalization and exclusion restrictions. In particular, proposition 2 in Magnac and Thesmar (2002) establishes identification of the utilities $u_\tau(i, a)$, assuming that the discount factor β is known and after location normalization by adding constants to all utilities so that $u_\tau(\emptyset) + \beta v_\tau(\emptyset, \emptyset) = 0$.

To identify the discount factor β, we may resort to the exclusion restrictions shown in Abbring and Daljord (2020) being sufficient for identification of the discount factor in the dynamic discrete-choice models. An example of the appropriate exclusion restriction is the flatness of the utility function $u(i, a)$ for older cars, such as $u_\tau(i, \bar{a} - 2) = u_\tau(i, \bar{a} - 1)$ for some i. Results of Abbring and Daljord (2020) then establish identification of the discount factor.[33]

[32] There are other types of identifying restrictions. In our empirical application in sec. VI, we are able to identify transaction costs of both buyers and sellers even when accidents and secondary prices are unobserved under including driving and additional functional form restrictions on utilities.

[33] Theorem 6 does not pretend to provide the weakest possible conditions for identification, and identification can be established using alternative types of restrictions and incorporating other types of data not contemplated in that result. For example, in the model we actually estimate in sec. VI, we also use data on driving to show how to extend the model to allow for driving. Since the marginal utility parameters μ_τ also affect observed driving,

Even though preferences are identified only up an arbitrary location and scale normalization, this can be sufficient to use the model to make counterfactual predictions of a wide number of policy changes, including changes to car taxation policy, as we show in our analysis in section VI.[34] However, we acknowledge that there are some counterfactual predictions that cannot be identified without other information. For example, an upgrade to public transportation infrastructure would change the utility of the outside good, $u_\tau(\o)$, but we are able to identify only how consumers evaluate the utility of cars relative to the outside good. We would need independent information on the incremental willingness to pay for an upgrade to public transport infrastructure in order to make counterfactual predictions of these sorts of policy changes.

D. Alternative Estimation Approaches

One drawback of the maximum likelihood approach to estimation is the requirement of car ownership transition data rather than the typical market share data that are commonly used to estimate vehicle choice models. In principle, we could use McFadden's (1989) method of simulated moments to match the market shares in the outer loop of the DNFXP algorithm. To this extent, we could use the stationary ownership distribution $q(\theta)$ or $q_\tau(\theta)$ implied by the model equilibrium to match the observed aggregate or consumer-type-specific market shares. However, as discussed in the previous section, market share data alone are generally insufficient to identify the structural parameters without stronger functional form assumptions or utilizing additional data.

It follows from the identification argument in the previous section that our method can estimate models with agnostic and flexible specifications of consumer preferences. Moreover, the DNFXP approach allows for direct modeling of price endogeneity by capturing the functional dependence of equilibrium prices on heterogeneous consumer preferences for the observed and unobserved car characteristics. This contrasts our method with the well-known BLP method (Berry, Levinsohn, and Pakes 1995), which numerically inverts the mapping from expected discounted utility of different trading decisions (i, a) to its aggregate market shares q_{ia}

this is an additional source of identification. Identification can be assisted by placing stronger restrictions on households' utilities, $\{u_\tau(i, a)\}$, and in sec. VI we assume that these functions are quadratic in car age. With these additional restrictions (which are probably much stronger than absolutely necessary), we can identify all of the buyer-side transaction costs (again assuming $T_s(i, a) = 0$, as in theorem 6) and also the accident probabilities $\{\alpha(i, a)\}$ even when both accidents and secondary prices P are unobserved.

[34] See Kalouptsidi et al. (2021), who establish conditions where counterfactual predictions from dynamic single-agent discrete-choice models are identified even though agent preferences are only partially identified.

and then uses the method of instrumental variables to regress the inverted market shares to the prices P, which are endogenous right-hand-side variables. In contrast, we offer an instrument-free estimation approach.[35]

As we see in the next section, our model delivers predictions of secondary-market prices and accident rates that are quite reasonable. Thus, direct full-information maximum likelihood estimation with nested numerical solution of the equilibrium for each trial value of the parameter vector θ enables us to overcome serious econometric challenges that incomplete estimation approaches such as BLP are unable to deal with.[36]

VI. Analysis of Danish Car Tax Policy

In this section, we use the DNFXP maximum likelihood estimator to structurally estimate a version of our model with eight types of households and four different car types of 25 ages, using a data set provided by Statistics Denmark that follows the car holdings of all Danish households from 1997 to 2008. This data set contains nearly 39 million observations, yet we show that the 131 structural parameters of the model can be estimated in a matter of minutes using an ordinary laptop computer. We then show how the estimated model can be used to make counterfactual predictions that may be crucial for analysis of car tax policy in Denmark. We also extend the model by incorporating driving. This also enables us to study the effects of additional taxes, such as fuel taxes, and account for environmental and congestion impacts of hypothetical policy changes. We show that it is possible to raise tax revenue and consumer surplus while reducing CO_2 emissions by lowering registration taxes and raising fuel taxes.

We simulate the effects of reforms similar to ones that have been under consideration in Denmark, which shift taxation from the purchase of new cars to the use of cars by increasing the fuel tax. We compare the predictions from our estimated equilibrium model to those obtained from a model that does not account for equilibrium in the used-car market and instead assumes a proportional change in prices of new and used cars. Predictions that fail to account for equilibrium responses are unable to accurately capture behavioral responses to this policy change. They overestimate the change in fleet composition, compared to an equilibrium analysis where used-car prices, scrappage rates, and holdings of cars respond

[35] We are not the first to use an instrument-free full-solution likelihood-based approach to identify and estimate the structural parameters. Yang, Chen, and Allenby (2003) used a Bayesian approach for inference of the structural parameters of a static equilibrium model of simultaneous supply and demand.

[36] We refer to BLP as "incomplete" in the sense that it focuses on estimation of demand-side parameters (e.g., preferences for different cars) while avoiding nested numerical solution for equilibrium prices in order to directly model this endogeneity.

endogenously. Paradoxically, predictions from nonequilibrium models are too extreme, overpredicting tax revenue gains, for example.

A. *Incorporating Driving*

The primary reason to own a car is to drive it, and thus far we have ignored this important aspect of car ownership. Let x denote the number of kilometers a consumer chooses to travel in a period, and let p_j denote the price per kilometer traveled for a car of type j. This equals the price of fuel (e.g., Danish kroner [DKK] per liter) divided by the car's fuel efficiency (kilometers per liter of fuel) for car type j.

Let $u_\tau(j, a, x)$ be the utility a consumer of type τ obtains from owning a car of type j and age a and driving it (on average) for x kilometers during the period. We make a simplifying assumption that the probability of an accident and other physical deterioration in an automobile is independent of driving, x, and is instead only a function of car type j and car age a. The benefit of this assumption is that driving becomes a static subproblem of the consumer's overall dynamic trading problem. The optimal amount of driving $x_\tau(j, a, p_j)$ then simply maximizes the driving utility net of monetary cost, $u_\tau(j, a, x) - \mu_\tau x p_j$. Substituting $x_\tau(j, a, p_j)$ back into the utility function $u_\tau(j, a, x)$, we obtain the indirect utility $u_\tau(j, a, p_j)$ for owning a car of age a that incorporates the individual's optimal choice of driving. Assuming that p_j is time invariant, the resulting model falls within the specification of section III.

To allow for discrepancies between the theoretical optimal amount of driving and actual data on kilometers traveled by different cars between (biannual) inspections in Denmark, we treat $x_\tau(j, a, p_j)$ as planned driving by the consumer at the start of each period. Actual driving is subject to ex post unexpected events during the period that cannot be predicted exactly. We represent by ζ the net ex post effect of these unexpected driving needs on the marginal utility of driving, resulting in an ex post utility specification of the following form:

$$\begin{aligned} u_\tau(j, a, x, \zeta) &= \psi_\tau(j, a) + (\gamma_\tau(j, a) + \zeta)x - \frac{\phi_{\tau,j}}{2}x^2 - \mu_\tau p_j x, \\ \gamma_\tau(j, a) &= \gamma_{\tau,j,0} + \gamma_{\tau,j,1}a, \end{aligned} \tag{43}$$

where the first component does not depend on x and can be considered as the utility of owning a car apart from driving it and therefore affects only trading behavior. We assume that $\psi_\tau(j, a)$ is a quadratic in the age of the car to fit the overall market shares across car types and ages, so we have $\psi_\tau(j, a) = \psi_{\tau,j,0} + \psi_{\tau,j,1}a + \psi_{\tau,j,2}a^2$.

The optimal ex post level of driving $x_\tau(j, a, p_j, \zeta)$ implied by this structure is

$$x_\tau(j, a, p_j, \zeta) = \frac{1}{\phi_{\tau,j}}\left(-\mu_\tau p_j + \gamma_{\tau,j,0} + \gamma_{\tau,j,1} a + \zeta\right). \tag{44}$$

Note that conditional on type-specific coefficient $\phi_{\tau,j}$, the parameters in equation (44) can be estimated by regressing the observed kilometers traveled to the cost of driving and household and vehicle characteristics. However, the regression only partially identifies a subset of structural parameters as ratios involving the parameter $\phi_{\tau,j}$, the coefficient governing the level of diminishing marginal utility from driving.[37]

When we substitute the expression for optimal ex post driving (eq. [44]) back into the utility function (eq. [43]) and take expectations over the ex post shocks ζ, we obtain a specification for the ex ante indirect utility of car ownership that is a quadratic in age. By a slight redefinition, the parameters $(\gamma_{\tau,j,0}, \gamma_{\tau,j,1}, \gamma_{\tau,j,2})$ also subsume the first and second moments of the unobserved ex post shock to utility ζ. Thus, besides the marginal utility of money parameter μ_τ, there are a total of six unknown parameters for each consumer type τ and car type j, which are $\theta_{\tau,j} = (\psi_{\tau,j,0}, \psi_{\tau,j,1}, \psi_{\tau,j,2}, \gamma_{\tau,j,0}, \gamma_{\tau,j,1}, \phi_{\tau,j})$. In appendix D, we show that the six parameters in $\theta_{\tau,\,j}$ are just-identified in terms of the six corresponding "semi–reduced-form" parameters, three for the linear driving equation and three for the ex ante expected indirect utility of owning car (j, a), which is a quadratric in a after taking expectations of the ex post preference shock ζ: $E\{u_\tau(j, a, x(j, a, p_j, \zeta))\} = u_{\tau,j,0} + u_{\tau,j,1} a + u_{\tau,j,2} a^2$. This implies that we can estimate the model in two steps: first, we estimate separate linear driving regressions for each (τ, j) combination to identify the three ratios $(-\mu_\tau/\phi_{\tau,j}, \gamma_{\tau,j,0}/\phi_{\tau,j}, \gamma_{\tau,j,1}/\phi_{\tau,j})$. Then, we use the DNFXP algorithm to estimate and identify the four parameters $(\mu_\tau, u_{\tau,j,0}, u_{\tau,j,1}, u_{\tau,j,2})$ of the implied quadratic expected indirect utility function. Using these seven estimated parameters, appendix D shows that we can solve for all seven of the structural parameters $(\mu_\tau, \psi_{\tau,j,0}, \psi_{\tau,j,1}, \psi_{\tau,j,2}, \gamma_{\tau,j,0}, \gamma_{\tau,j,1}, \phi_{\tau,j})$. It is critical to fully identify all the underlying structural parameters in order to make counterfactual predictions involving changes in the fuel price paid by consumers, p, which in turn change the per-kilometer cost of driving, p_j, of the different types of cars $j \in \{1, \ldots, J\}$.

B. Estimation Results and Model Fit

Though there are hundreds of different makes and models of cars sold in Denmark, for this analysis we aggregated them into four car types differentiated by their fuel economy and pollution levels ("green" for more fuel efficient, environmentally friendly cars and "brown" for others) and

[37] We ignore the restriction $x(j, a, p_j, \zeta) \geq 0$, which implies that ζ must be a truncated normal distribution.

based on car weight ("heavy" vs. "light"). We divided Danish households into eight groups τ, depending on (a) whether they were singles or couples, (b) whether the distance to work was short or long, and (c) whether the household was rich or poor. The precise criteria for defining these groups are detailed in appendix E.

We estimated a linear specification for preferences, including driving, as discussed in the previous section. Household preferences for cars decrease with age but at a diminishing rate, and there is heterogeneity in preferences for the different types of cars. We also estimated household-specific quasi-linear price sensitivity parameters μ_τ for each of the eight types τ. By dividing the estimated coefficient $u_{\tau,j,0}$ for household τ's utility of a new car of type j by μ_τ, we obtain a measure of willingness to pay for one period's use of a new car in DKK. For example, we estimate that a rich couple with low work distance is willing to pay (e.g., rent for 1 year) a new, light, brown car for 36,704 DKK (or about US$5,580), compared to 32,682 DKK for the corresponding poor household. In general, we find that, on the basis of revealed choices, (1) rich households are willing to pay more for any type of car than poor households; (2) all households preferred the heavy cars to the light ones and brown cars to green ones, resulting in the following preference ordering: heavy brown $\succ$ heavy green $\succ$ light brown $\succ$ light green; (3) willingness to pay for cars by high-work-distance households exceeds that of low-work-distance ones; and (4) couples generally have higher willingness to pay for cars than singles. Given that there are a total of 131 parameters in the model, we refer the reader to appendix E for details on the maximum likelihood parameter estimates and standard errors.

The estimated model also has reasonable implications for driving (see table A3): households with high work distances drive much more than those with low ones, and more so for the rich. The estimated model implies fuel price elasticities between -0.10 and -0.60 across households. This is relatively close to the results of Gillingham and Munk-Nielsen (2019), who find an average elasticity of -0.30 using a wide array of regression specifications.

The estimated equilibrium model provides a good fit to the observed distribution of car holdings and successfully captures key features of Danish households: (1) poor households are significantly more likely not to own a car than rich ones, (2) couples are more likely to own cars than singles, (3) high-work-distance households are relatively more likely to own cars than those with low work distance, and (4) a large fraction of households (40%) do not own cars. The last is in part due to the excellent public transport infrastructure and the widespread use of bicycles in Denmark but is also due to the high taxation of cars, which we analyze in more detail in the next section. Figure 4 shows how our model captures the posttrade age distribution of holdings of different cars by different

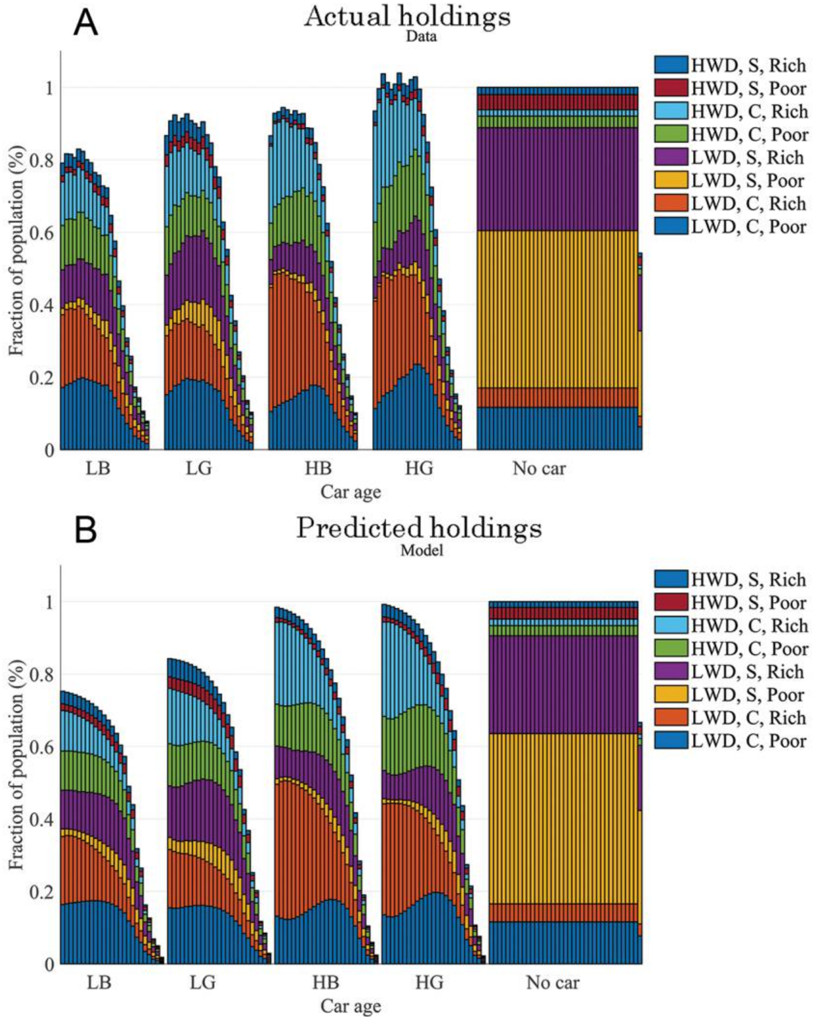

FIG. 4.—Actual and predicted holdings, by household type, car type, and car age. The graphs show the fraction of the population holding each car-age combination as well as the outside good. For the outside good, we have opted to split the bars and put them next to each other, since otherwise that option would dominate the scale of the *y*-axis. Household types: HWD = high work distance; LWD = low work distance; S = single; C = couple. Car types: LB = light, brown; LG = light, green; HB = heavy, brown; HG = heavy, green. Within each of the four car types, car ages go from 0 to 24.

households, including the "hand-me-down chain" from rich to poor consumers. For example, note that for low-work-distance singles (the dark blue and red regions at the bottom of the bar graphs), the rich households (colored red) are relatively more concentrated in holding newer cars of each type, whereas the poor households are more concentrated in holding older cars.

As we noted in section V.A, our maximum likelihood estimation does not attempt to directly fit the holdings distributions, which we previously denoted q_τ for each household type τ. Since figure 4 plots the actual and predicted posttrade ownership distribution, $q\Omega(P)$, it also involves a comparison of the implied stationary distribution from our model, $q_\tau(\hat{\theta})$, to the nonparametric estimates q_τ from the data. Though our model slightly underpredicts holdings of new light cars and overpredicts holdings of new heavy cars, overall we think that the model provides a remarkably good overall fit to over 800 nontargeted probabilities shown in figure 4 using a fairly parsimonious model with 131 parameters.[38]

As we noted above, the Danish register data contain information on nearly 80,000 car ownership transition probabilities, so the next figures provide some information on the model's ability to fit these transitions. Figure 5 illustrates the model's ability to capture the probability of car purchases as well as the probability of keeping the existing car. Figure 5*A* plots the conditional probability that households purchase cars of a given age. The model closely tracks the observed purchase patterns at the aggregate level: households are more likely to buy a new car rather than any of the used ones, and purchase probabilities decline as cars approaches the scrap age. When comparing these purchase probabilities for each of the eight household types (results not shown), the model also closely tracks observed purchase probabilities and mimics the overall pattern from figure 4 that rich, couple, high-work-distance households are the types most likely to buy newer and larger (more expensive) cars.

Figure 5*B* focuses on the 60% of Danish households that do own cars and plots the conditional probability of keeping their existing car as a function of its age. The model is generally able to match the overall level of probability of keeping a car but also reveals an aspect of the data that our model is unable to capture well: we see that in the data, the probability of keeping a car is very high in the first couple of years and then gradually falls with the age of the car, whereas our model predicts only a more modest decrease until a drop around age 15.

We conclude our presentation of the estimation results with figure 6, which illustrates the model's predictions of quantities that we do not directly observe in our data set from Statistics Denmark. As we noted, our data allow us to observe scrappage of cars but not directly accidents leading to scrappage, since another source of scrappage is voluntary scrappages by

[38] There are relatively few cars that are more than 20 years old in our data set, and as a result of measurement issues relating to the oldest cars discussed in app. E, we eliminated cars over 22 years old from our estimation sample. Thus, the histograms for the data in fig. 4 are truncated at age 22.

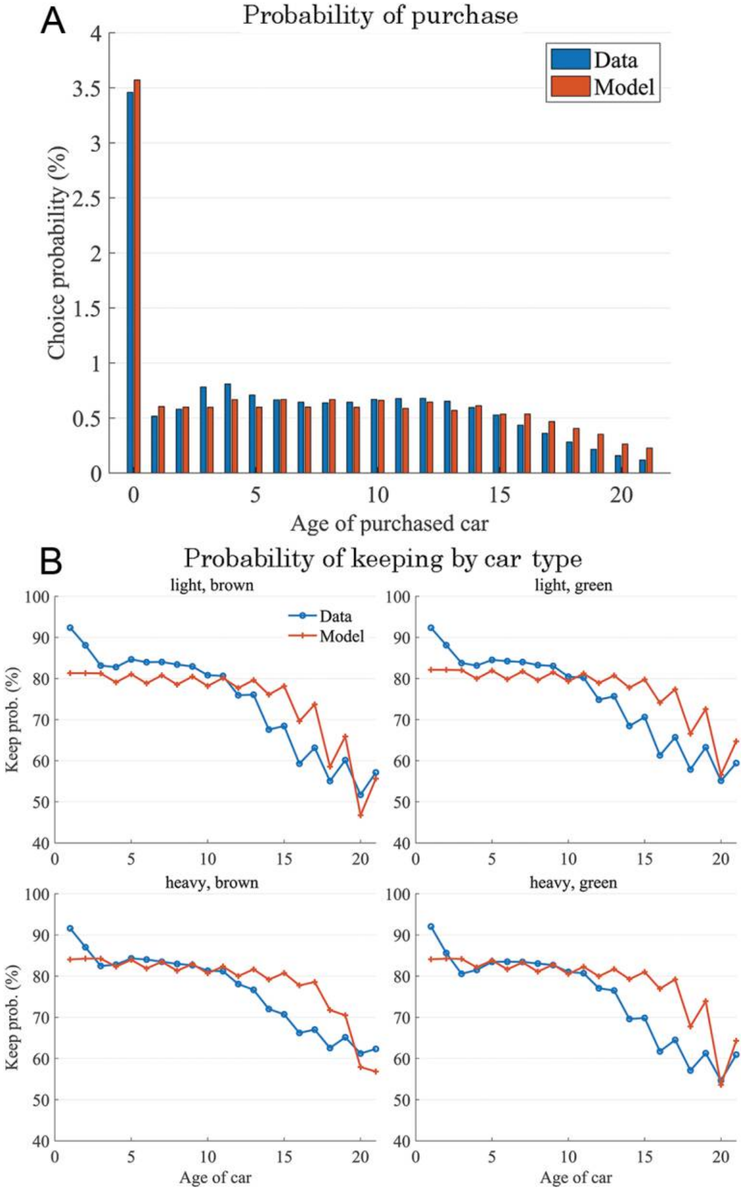

FIG. 5.—Actual and predicted probability of keeping and purchase.

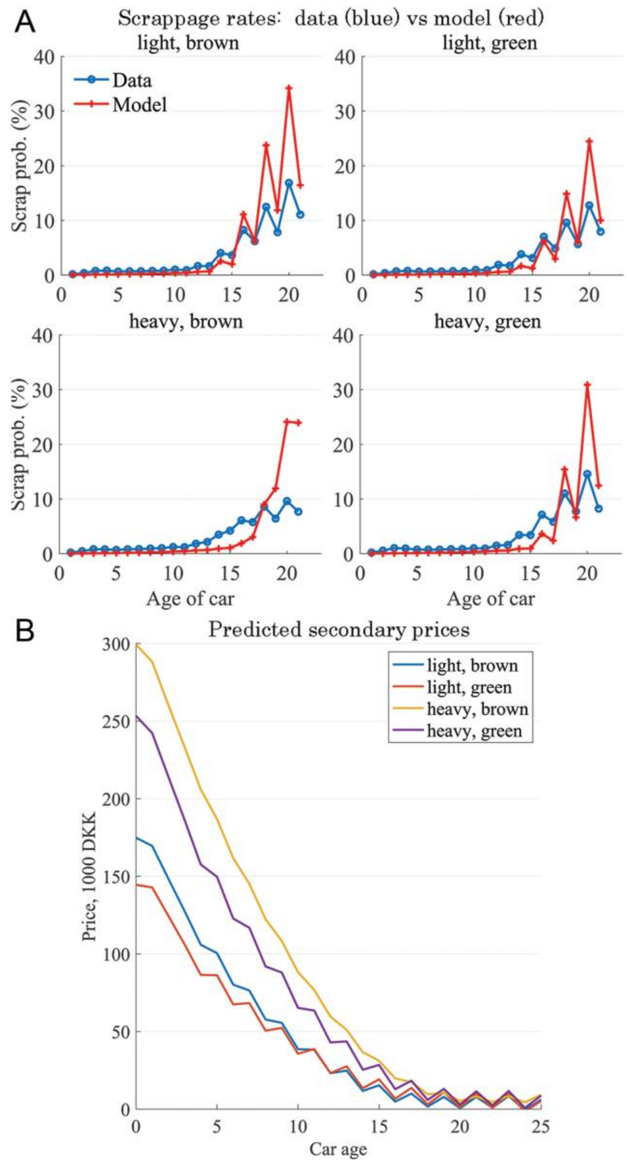

FIG. 6.—Zig-zag patterns in scrappage and predicted equilibrium prices. *A* shows the fit of the scrappage rates for each car and age (averaged over households weighted by the equilibrium ownership distribution), and *B* shows the equilibrium prices.

households, such as cars that are still drivable but may require expensive repairs to enable them to pass safety checks that are required before they can be sold to another household. As we described in section V.A, we are able to identify and estimate accident probabilities using our structural estimation approach, even though we do not observe accidents. The estimated parameters shown in table A4 imply that accident rates of the light cars are generally higher than those of the heavy ones and that accident rates rise quickly with age after cars are 15 years old but are negligible when cars are new.

Figure 6*A* displays the overall probability that a car is scrapped by the age and type of the car. We do directly observe when scrappages occur in the Danish Register data, so in this graph we can compare the model predictions (red lines) to the data (blue lines). Here we see the curious "zig-zag" pattern in scrappage rates that we noted in section I: Danish cars are much more likely to be scrapped at even ages than at odd ages. We have verified that this effect is real and is not an artifact of how and when scrappages are recorded. It is due to the very strict biannual car inspections in Denmark that occur at even ages once cars are 4 years old or older. If the inspection reveals mechanical, safety, or emissions problems, the owner is required to repair them in order to continue driving the car. We believe that Danes find these costs to be onerous and thus that they are more likely to scrap rather then keep or sell their cars if they have problems when they are sufficiently old. We capture this effect in our model by including an even-age dummy in the utility of car ownership (skipping $a = 2$). The estimation results reveal big negative estimated coefficients for these dummy variables, with an estimated disutility that is typically 15%–50% as large as the estimated single-period utility the household obtains from owning a brand new car of the same type.

Finally, figure 6*B* shows the estimated secondary prices of the four types of cars in our model. As we noted in section V.A, even though we do not directly observe these prices, we can compute them for any trial value of the structural parameters and use the substantial number of other moments in the data to identify both the structural parameters θ and the implied secondary prices, $P_{i,a}(\theta)$, for all four car types i and car ages a.[39] We first note that the rate of decline of our used-car prices is broadly consistent with external evidence. We have limited data on suggested annual discount rates for used cars from the Danish Used Car Dealer Association, which suggest that prices should fall by 13% per year, on average. Our model solution implies that prices fall, on average, 14% per year

[39] We do observe prices of new cars, so we use the average new-car prices $\bar{P}_i$ of the different makes and models in the four aggregate type groupings of cars as data rather than estimating these as additional parameters (see table A2). The observed new-car prices help to "tie down" secondary prices.

for three of the four car types and 11% per year for light green cars. Thus, the overall magnitude of depreciation in our results is quite similar to the best data available.

Second, we note that the zig-zag pattern in scrappages is also present in the secondary prices predicted by our model. The effect on secondary-market prices is a natural consequence of the estimated disutility our model predicts that Danish households experience during the even-aged years when their car is subject to inspection. As mentioned above, independent evidence from the limited data we have on actual used-car transactions prices suggest that the zig-zag pattern in secondary prices that our model predicts is a real phenomenon in Denmark.

C. *Counterfactual Policy Analysis*

In this final section, we carry out counterfactual predictions, using our estimated equilibrium model of the Danish auto market. We focus on the effects of changes in the Danish new-car (registration) tax and fuel tax. As we noted in section I, Denmark has one of the highest new-car taxes in the world. The tax is progressive, with a rate of 105% of the retail price of the car in the first bracket (for the cars priced up to 81,000 DKK, approximately equal to US$16,000 at the time, excluding VAT) and a rate of 180% of the retail price in the second bracket. There is also a VAT of 25% applied to the wholesale price of a new car, before the calculation of the additional new-car registration tax. Appendix E provides further details about car taxation policy in Denmark, which has been subject to vigorous political debate and a few reforms in recent years.[40]

In order to make the counterfactual predictions, which involve tax policy changes that affect new-car prices $\bar{P}_i$ and thus used-car prices $P_{i,a}$, as well fuel prices p that affect the price per kilometer driven for different car types, p_i, we need estimates of the "deep structural parameters" described in section VI.A, where we introduced driving into the model. These deep parameters are $\theta_\tau = (\mu_\tau, \{\psi_{\tau,j,0}, \psi_{\tau,j,1}, \psi_{\tau,j,2}, \gamma_{\tau,j,0}, \gamma_{\tau,j,1}, \phi_{\tau,j}\})$, which differ for each household type τ where the other parameters, except the marginal utility of money μ_τ, also differ by car type j. Recall that the ψ parameters reflect the pure utility of ownership (independent of any driving) for different cars, whereas the γ and ϕ parameters capture the utility from driving.

Once we have identified the deep structural parameters, we can systematically vary the registration tax rate (which affects the gross of tax new-car prices $\bar{P}_i$ and thus also equilibrium secondary prices $P_{i,a}$), as well

[40] After the end of our sample, the registration tax rate was reduced across two separate reforms, in combination with changes to the treatment of electric vehicles (which were not present during our sample period).

as the after-tax fuel price p that affects the cost per kilometer driven p_j of the different car types j. For each alternative policy we consider, we calculate the counterfactual equilibrium, which also enables us to evaluate consumer welfare as well as the impact on overall tax revenues received by the Danish government. Specifically, we analyze a proposed policy involving cutting the new-car registration tax in half while increasing the fuel tax to offset the revenue loss from reduced new-car taxes. Intuitively, this policy shifts taxation from the purchases of cars to their usage, with the intention to offset some of their harmful externalities. We assumed that the social cost of carbon is US$50/ton to estimate the marginal external social costs of driving, using results from Transport DTU (2010) that also include negative externalities from congestion, accidents, noise, and local air pollution.[41]

The overall purpose of this analysis is to illustrate the added value of using an equilibrium model to inform car tax policy. We want to illustrate how a policy maker would design a revenue-neutral reform that shifts taxation away from car registrations and toward fuel and usage. Specifically, we want to compare a sophisticated policy maker to a "naïve" policy maker using a nonequilibrium model. As mentioned above, a naïve way of handling the used-car market in a nonequilibrium framework is to assume proportional changes in used- and new-car prices. The baseline tax system for new cars is a two-part linear system with a kink, after which the marginal tax rate increases. Thus, the registration tax is progressive. We analyze a reform where both the low and high rates are cut in half.

We assume 100% pass-through of taxes to new-car prices, which is consistent with our assumption in our estimation of perfectly elastic supply for new cars as a result of Denmark being a small open economy with no automobile manufacturing. We are not aware of any studies of pass-through in the Danish new-car market, but full pass-through aligns with some studies in the US new-car market, such as Sallee (2011), but differs from others (Busse, Silva-Risso, and Zettelmeyer 2006). Assuming full pass-through to new-car prices, cutting the low and high rates in half results in new-car prices falling by between 25.6% and 27.6% for the four car types. In the nonequilibrium setting, which we refer to as "naïve, expected," we assume that the prices of used cars of all vintages fall by the same percentages relative to the baseline equilibrium. However, while it is easier to make predictions, the naïve approach fails to account for the endogenous adjustment of car trading and secondary prices to the

[41] See also Winston and Yan (2021), who analyze US data and find that traffic congestion can increase the demand for larger, less fuel-efficient cars. Thus, congestion pricing, i.e., taxes or tolls that can mitigate congestion, "could reduce the vehicle fatality rate, generating $25 billion in annual benefits, and could improve vehicle fleet fuel efficiency, generating roughly $10 billion in annual operating cost savings" (196).

change in fuel prices and new-car prices. Thus, we consider the following four scenarios.

1. Baseline: the model is solved for equilibrium prices and calibrated under status quo Danish tax policy as of 2008.
2. Naïve, expected: a naïve policy maker assumes that used-car prices will fall by the same proportion as the corresponding new-car price for each car type; that is, the market is not in equilibrium. The policy maker raises fuel taxes until revenue is equivalent to the baseline. We calculate individual household welfare using these prices even though the used-car market is not in equilibrium.
3. Naïve, realized: this is the equilibrium outcome that would result from the fuel tax policy enacted by the naïve policy maker above. That is, the market is in equilibrium here, and used-car prices are set to equate supply and demand, but tax revenue is not equal to the baseline.
4. Sophisticated: these are the predictions of a sophisticated policy maker who correctly predicts the endogenous equilibrium responses to tax policy changes. That is, the used-car prices are such that the market is in equilibrium, and fuel taxes are set so that the total tax revenue is equal to the baseline tax policy scenario.

The outcomes under the four different policy scenarios are presented in table 1, and the resulting car prices are in figure 7. In the "Naïve, Expected" scenario, the policy maker is guided by a naïve expectation of proportional pass-through. Lowering registration taxes results in an increase in tax revenue, so in order to achieve revenue equivalence, the policy maker increases fuel taxes from 57% of the price at the pump in the baseline to 76%.[42] According to this nonequilibrium model, that should achieve revenue equivalence at 9,391 DKK per household annually, but with a much younger car fleet, where the average car age falls from 6.5 to 3.1 years. However, the scenario "Naïve, Realized" shows what will actually happen once used-car prices adjust to equilibrate the market. First, we note from figure 7 that used-car prices fall more than proportionally, which correspondingly results in car ages falling by less than predicted, only to 4.3 years. In other words, the naïve model predicts a much-too-strong movement toward newer cars, which results in excess demand for newer cars. As a result, registration tax revenue crumbles,

[42] Alternatively, the policy maker could also lower fuel taxes to achieve revenue equivalence. However, the required reduction implies a virtual abolishment of fuel taxes, which we judge to be less realistic in practice. Nevertheless, this illustrates the complexities in policy design with Laffer curve effects.

TABLE 1
Policy Simulation Results

	Baseline	Naïve, Expected	Naïve, Realized	Sophisticated
	A. Policy Choice Variables			
Registration tax (bottom rate)	1.050	.525	.525	.525
Registration tax (top rate)	1.800	.900	.900	.900
Fuel tax (share of pump price)	.573	.761	.761	.732
	B. Exogenous Prices			
Light, brown car (000 DKK)	174.902	129.532	129.532	129.532
Light, green car (000 DKK)	144.551	107.532	107.532	107.532
Heavy, brown car (000 DKK)	299.452	214.048	214.048	214.048
Heavy, green car (000 DKK)	253.397	182.796	182.796	182.796
Fuel (DKK/L)	8.322	14.885	14.885	13.243
	C. Outcomes			
Social surplus (000 DKK)	9.382	11.281	8.439	10.203
Total tax revenue (000 DKK)	9.391	9.391	7.452	9.391
Fuel tax revenue (000 DKK)	4.282	5.184	4.983	6.224
Car tax revenue (000 DKK)	5.110	4.207	2.468	3.167
Non-CO_2 externalities (000 DKK)	6.751	3.385	3.281	4.711
Externalities (000 DKK)	7.374	3.702	3.586	5.157
Consumer surplus (000 DKK)	7.364	5.592	4.573	5.969
CO_2 (tons)	2.148	1.094	1.052	1.537
Driving (000 km)	10.861	5.446	5.279	7.580
E(car age)	6.507	3.080	4.336	5.417
Pr(no car)	.367	.535	.534	.418

Note.—In the baseline scenario, the institutional parameters conform to the data for 2008. In the column "Naïve, Expected," the two rates for the new-car tax are both cut in half, and then fuel taxes are increased until tax revenue is equal to the baseline. Used-car prices are assumed to fall by the same percentage as new-car prices (i.e., 100% pass-through from the new- to the used-car market). In "Naïve, Realized," the new-car taxes and fuel taxes are as in "Naïve, Expected," but we solve for the equilibrium used-car prices. In "Sophisticated," we change the fuel tax so that the total tax revenue is equal to the baseline, each time solving for the used-car price equilibrium.

resulting in total tax revenue falling short of the intended equivalence target, from 9,391 to 7,452 DKK per household.

Instead, the column "Sophisticated" shows that the policy maker would have to increase fuel taxes only until they make up 73% of the price at the pump if she takes into account the endogenous responses in the used-car market.

The general takeaway message is that a nonequilibrium model produces much greater movement in new-car sales, and thus in registration taxes, than an equilibrium model can sustain in flow equilibrium. This means that the policy maker will expect greater effects from changes in registration taxes than will actually come to pass. The analysis so far has focused on comparing the decisions made under the guidance of a nonequilibrium model, as opposed to our equilibrium model. If we consider first the implications of the tax policy in the "Sophisticated" column, we note that the

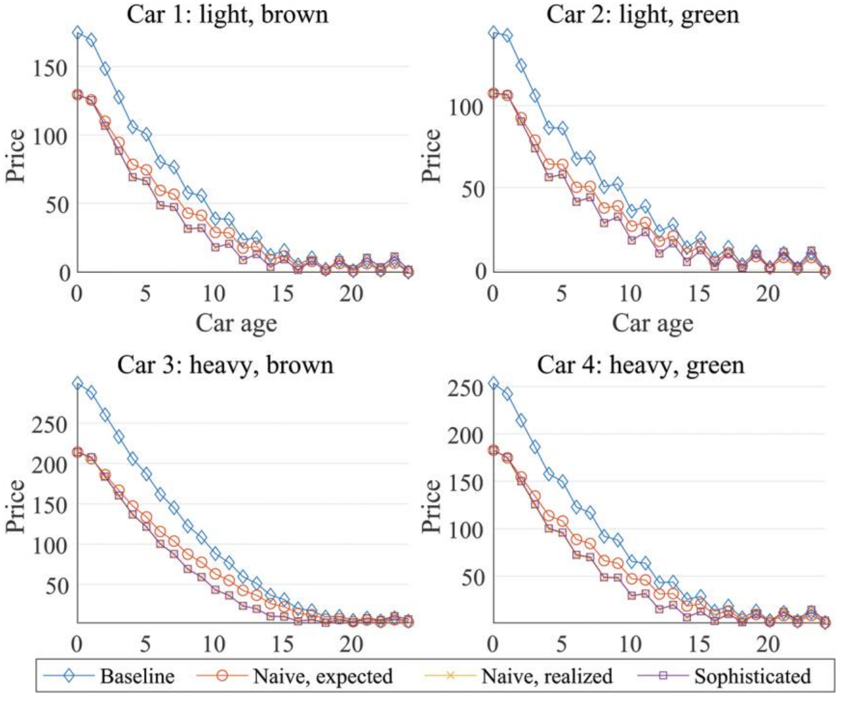

FIG. 7.—Equilibrium prices. Each panel shows the prices for the corresponding car type; the four lines represent the four different scenarios (see table 1 for descriptions).

reform succeeds in raising total societal welfare. This is because, although consumer surplus falls slightly from 7,364 to 5,969 DKK, driving-related externalities fall by more, as total driving falls from 10,861 to 7,580 km annually in response to the much higher cost of driving. Thus, there are clear welfare differences between revenue-equivalent combinations of the two tax rates on purchase and use.

We next consider whether there are policy options that can raise welfare without harming tax revenue or increasing emissions. To do this, figure 8 shows contour lines on a plot of the registration tax against the fuel tax.[43] The axes are scaled so that the point (1, 1) denotes the baseline tax levels and the green area to the southeast of the baseline levels represents combinations of the two tax rates that result in lower emissions, higher welfare, and higher tax revenue—a "win-win-win" situation.[44]

[43] See app. E and, specifically, online fig. 1, for the three-dimensional graphs these contours are taken from.

[44] The x- and y-axes show the tax rates for fuel and car registrations, respectively, normalized so that the baseline is 1. Each line represents the contour lines for one of three outcomes; i.e., combinations of the two tax rates where the outcome is kept constant and

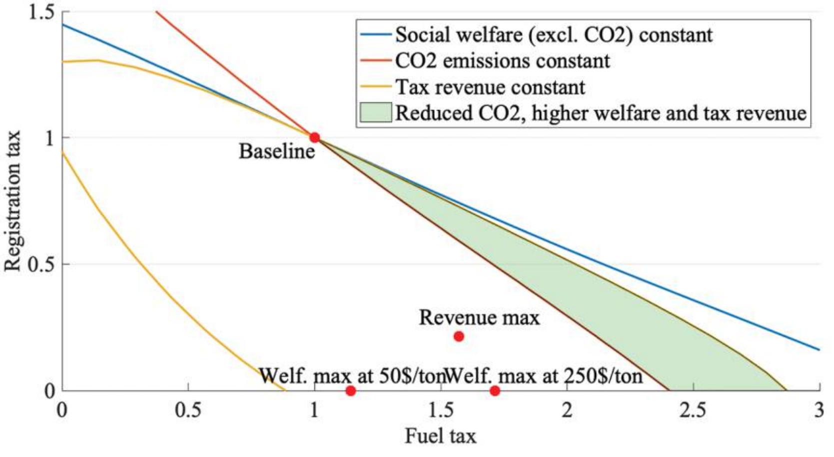

FIG. 8.—Welfare, revenue, and environmental effects of different tax policies.

Notably, the baseline (current situation) is quite far from the tax revenue–maximizing level, which has a much lower registration tax but a higher fuel tax. The welfare-maximizing tax combinations set the registration tax to zero, instead relying solely on the fuel tax to target externalities. The complex interaction between the two car taxes illustrates the importance of jointly modeling the purchase and driving decisions in an equilibrium framework and the possibility of using such a model to optimize tax policy to meet policy-maker objectives.

VII. Conclusion

We have introduced a computationally tractable model of equilibrium in the primary and secondary markets for automobiles that allows for flexible specifications of preferences and consumer heterogeneity and transaction costs. Our work was inspired by the early static discrete-choice models

equal to the value in the baseline configuration, occurring at the point (1,1). The three outcomes are tax revenue, CO_2 emissions, and social welfare (excluding the external costs of CO_2 emissions). Moving from the baseline in the direction of the origin implies an increase in all three outcomes, although tax revenue will eventually start to decline again (because in the baseline, both tax rates are above the top point of the Laffer curve). Four points are depicted on the graph in red: first, the baseline, (1,1); second, the top point of the Laffer "curve," where overall tax revenue is maximized; and third, two points that show the overall social welfare–maximizing policies: one under a CO_2 price of \$50/ton and one at a higher price of \$250/ton (a price recently suggested by the Danish environmental council). Not surprisingly, a higher price results in a higher fuel tax, but still a zero tax on car purchases.

of equilibrium in the automobile market pioneered by Manski, Sherman, and Berkovec and the subsequent efforts to extend their models to include dynamics and transaction costs and model equilibrium price setting in the primary market by Rust, Stolyarov, Gavazza, Lizzeri, Roketskiy, and Esteban and Shum. We believe that our framework is promising for empirical applications and policy analysis, and in future work we plan to further extend and apply it in a number of directions.

One of these directions is ongoing work (Gillingham et al. 2019) to structurally estimate an "overlapping-generations" version of our model using Danish register data to allow for a realistic counterfactual analysis of vehicle tax reform in Denmark. Another direction is to extend the model to include Bertrand-Nash equilibrium in the primary market for autos in addition to competitive equilibrium in the secondary markets. Estimation of the model with the primary market requires a triply nested version of NFXP, but the payoff is that we can use the model to relax our assumption of 100% pass-through of new-car taxes to retail prices as well as to predict merger counterfactuals.

Another direction would relax the assumption of stationarity and extend our definition of equilibrium to allow for macroeconomic shocks that can capture the pronounced "waves" often found in the age distribution of vehicles (Adda and Cooper 2000). We are comparing different solution concepts in terms of computational tractability and empirical realism, including the "temporary equilibrium" concept of Grandmont (1977), the "sufficient statistic" approach of Krusell and Smith (1998), and a full-blown rational expectations equilibrium that takes into account the entire holdings distribution of cars as a component of the "state variables" that consumers use to predict future prices, as in Cao (2020).

A very challenging extension of our model would endogenize the characteristics of vehicles by allowing firms to invest in R&D to produce new vehicle designs. Longer-run competition on attributes will likely require a fundamentally nonstationary framework and raises questions of consumer expectations over future products. Very promising headway into this sort of analysis has been done in the pioneering work of Goettler and Gordon (2011), and it may be possible to adapt this approach into a more evolutionary model of the automobile market. A final challenging extension would be to incorporate asymmetric information in a more detailed treatment of the "microstructure" of trade in the automobile market, including endogenous intermediation of trade by car dealers as well as direct consumer transactions. Recent studies, such as Biglaiser et al. (2020), have provided new empirical insights into the microstructure of trade that are not modeled in our framework but represent important directions to pursue in the development of more detailed and realistic models of the microstructure of trade in automobiles.

References

Abbring, Jaap H., and Øystein Daljord. 2020. "Identifying the Discount Factor in Dynamic Discrete Choice Models." *Quantitative Econ.* 11(2): 471–501. https://onlinelibrary.wiley.com/doi/pdf/10.3982/QE1352.

Adda, Jérôme, and Russell Cooper. 2000. "Balladurette and Juppette: A Discrete Analysis of Scrapping Subsidies." *J.P.E.* 108 (4): 778–806.

Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Q.J.E.* 84 (3): 488–500.

Arcidiacono, Peter, and Robert A. Miller. 2011. "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity." *Econometrica* 79 (6): 1823–67. https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA7743.

Berkovec, James. 1985. "New Car Sales and Used Car Stocks: A Model of the Automobile Market." *RAND J. Econ.* 16 (2): 195–214.

Berndt, Ernst R., Bronwyn H. Hall, Robert E. Hall, and Jerry A. Hausman. 1974. "Estimation and Inference in Nonlinear Structural Models." *Ann. Econ. and Soc. Measurement* 3 (4): 653–85.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–90.

Biglaiser, Gary, Fei Li, Charles Murry, and Yiyi Zhou. 2020. "Intermediaries and Product Quality in Used Car Markets." *RAND J. Econ.* 51 (3): 905–33.

Borken-Kleefeld, Jens, and Yuche Chen. 2015. "New Emission Deterioration Rates for Gasoline Cars—Results from Long-Term Measurements." *Atmospheric Environment* 101:58–64.

Busse, Meghan, Jorge Silva-Risso, and Florian Zettelmeyer. 2006. "$1,000 Cash Back: The Pass-Through of Auto Manufacturer Promotions." *A.E.R.* 96 (4): 1253–70.

Cao, Dan. 2020. "Recursive Equilibrium in Krusell and Smith (1998)." *J. Econ. Theory* 186:104978.

Esteban, Susanna, and Matthew Shum. 2007. "Durable-Goods Oligopoly with Secondary Markets: The Case of Automobiles." *RAND J. Econ.* 38 (2): 332–54.

Gavazza, Alessandro, Alessandro Lizzeri, and Nikita Roketskiy. 2014. "A Quantitative Analysis of the Used-Car Market." *A.E.R.* 104 (11): 3668–3700.

Gillingham, Kenneth, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, and Bertel Schjerning. 2019. "A Dynamic Model of Vehicle Ownership, Type Choice, and Usage." Manuscript.

Gillingham, Kenneth, and Anders Munk-Nielsen. 2019. "A Tale of Two Tails: Commuting and the Fuel Price Response in Driving." *J. Urban Econ.* 109:27–40.

Goettler, Ronald L., and Brett R. Gordon. 2011. "Does AMD Spur Intel to Innovate More?" *J.P.E.* 119 (6): 1141–1200.

Grandmont, Jean Michel. 1977. "Temporary General Equilibrium Theory." *Econometrica* 45 (3): 535–72.

Hendel, Igal, and Alessandro Lizzeri. 1999. "Adverse Selection in Durable Goods Markets." *A.E.R.* 89 (5): 1097–115.

Hotz, V. Joseph, and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Rev. Econ. Studies* 60 (3): 497–529.

House, Christopher L., and John V. Leahy. 2004. "An *sS* Model with Adverse Selection." *J.P.E.* 112 (3): 581–614.

Kalouptsidi, Myrto, Yuichi Kitamura, Lucas Lima, and Eduardo Souza-Rodrigues. 2021. "Counterfactual Analysis for Structural Dynamic Discrete Choice Models." Discussion Paper no. 2221, Cowles Found., Yale Univ.

Konishi, Hideo, and Michael T. Sandfort. 2002. "Existence of Stationary Equilibrium in the Markets for New and Used Durable Goods." *J. Econ. Dynamics and Control* 26 (6): 1029–52.

Krusell, Per, and Anthony A. Smith Jr. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *J.P.E.* 106 (5): 867–96.

Ma, Qingyin, and John Stachurski. 2021. "Dynamic Programming Deconstructed: Transformations of the Bellman Equation and Computational Efficiency." *Operations Res.* 69 (5): 1591–607.

Magnac, Thierry, and David Thesmar. 2002. "Identifying Dynamic Discrete Decision Processes." *Econometrica* 70 (2): 801–16. https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0262.00306.

Mannering, Fred, and Clifford Winston. 1985. "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization." *RAND J. Econ.* 16 (2): 215–36.

Manski, Charles F. 1980. "Short Run Equilibrium in the Automobile Market." Discussion Paper 80.18, Falk Inst., Hebrew Univ. Jerusalem.

———. 1983. "Analysis of Equilibrium Automobile Holdings in Israel with Aggregate Discrete Choice Models." *Transportation Res. B* 17 (5): 373–89.

Manski, Charles F., and Ephraim Goldin. 1983. "An Econometric Analysis of Vehicle Scrappage." *Transportation Sci.* 17 (4): 365–75.

Manski, Charles F., and Leonard Sherman. 1980. "Forecasting Equilibrium Motor Vehicle Holdings by Means of Disaggregate Models." *Transportation Res. Record* 764:96–103.

McFadden, Daniel. 1981. "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data with Econometric Applications*, edited by Charles F. Manski and Daniel McFadden, 198–272. Cambridge, MA: MIT Press.

———. 1989. "A Method of Simulated Moments for Estimation of Discrete Models without Numerical Integration." *Econometrica* 57 (5): 995–1026.

NHTSA (National Highway Traffic Safety Administration). 2013. "How Vehicle Age and Model Year Relate to Driver Injury Severity in Fatal Crashes." Res. note, US Dept. Transportation, Washington, DC.

Rust, John. 1985a. "Stationary Equilibrium in a Market for Durable Assets." *Econometrica* 53 (4): 783–805.

———. 1985b. "When Is It Optimal to Kill Off the Market for Used Durable Goods?" *Econometrica* 54 (1): 65–86.

———. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55 (5): 999–1033.

———. 1994. "Structural Estimation of Markov Decision Processes." In *Handbook of Econometrics*, vol. 4, edited by Robert F. Engle and Daniel L. McFadden, 3081–143. Amsterdam: North-Holland.

Sallee, James M. 2011. "The Surprising Incidence of Tax Credits for the Toyota Prius." *American Econ. J. Econ. Policy* 3 (2): 189–219.

Stolyarov, Dmitriy. 2002. "Turnover of Used Durables in a Stationary Equilibrium: Are Older Goods Traded More?" *J.P.E.* 110 (6): 1390–413.

Transport DTU (Technical University of Denmark). 2010. "Transportøkonomiske Enhedspriser." Version 1.1. Discussion paper, Center for Transport Analytics, Technical Univ. Denmark.

Winston, Clifford, and Jia Yan. 2021. "Vehicle Size Choice and Automobile Externalities: A Dynamic Analysis." *J. Econometrics* 222 (1A): 196–218.

Yang, Sha, Yuxin Chen, and Greg M. Allenby. 2003. "Bayesian Analysis of Simultaneous Demand and Supply." *Quantitative Marketing and Econ.* 1 (3): 251–75.