Price Discrimination and Online Sales in the Automobile Industry

Xavier D'Haultfœuille* Isis Durrmeyer†
Jean-François Fournel‡ Alessandro Iaria§

May 19, 2025

Latest version available here.

Abstract. We investigate the welfare consequences of introducing an online distribution channel in the French car industry, where most sales take place in person through car dealers relying on third-degree price discrimination. We estimate a structural model of demand with unobserved third-degree price discrimination and transportation costs related to visiting car dealers. In counterfactuals, we introduce an online distribution channel in which prices are uniform and consumers benefit from lower transportation costs. When both distribution channels are available, firms charge low online prices to attract internet-savvy consumers online, while continuing to price discriminate the less internet-savvy consumers in person. The online channel is profitable for firms, and the more it reduces transportation costs, the more profitable it is. However, the costs and benefits of the online channel are unevenly distributed among consumers, with older, wealthier, and internet-savvy consumers obtaining most of the benefits.

*CREST–ENSAE. E-mail: xavier.dhaultfoeuille@ensae.fr.
†Toulouse School of Economics and CEPR. E-mail: isis.durrmeyer@tse-fr.eu.
‡Toulouse School of Economics. E-mail: jean-francois.fournel@tse-fr.eu.
§University of Bristol and CEPR. E-mail: alessandro.iaria@bristol.ac.uk
We would like to thank Pierre Dubois and Gaston Illanes for helpful comments and suggestions. We would also like to thank the participants of various seminars and conferences. We acknowledge financial support from the European Research Council under grant ERC-2019-STG-852815 "PRIDISP" and the Agence Nationale de Recherche under grant ANR-17-EUR-0010 (Investissements d'Avenir program).

1 Introduction

Recent technological progress has facilitated online transactions for a wide variety of products and services. Doing business online has several advantages for firms and consumers. On the one hand, firms may gain access to a larger consumer base and may save the costs of establishing and maintaining a dense network of physical stores. Consumers, on the other hand, may benefit from having access to a wider variety of products and services and avoiding the potentially significant transportation costs associated with visiting physical stores for their purchases.

As consumers and firms are getting used to online marketplaces, there is some evidence suggesting that larger and more expensive products, such as cars, will also be traded online. A pioneering example is Tesla, which operates almost exclusively online. Ordering the vehicle, signing the contract, and making the payment are all happening through the company's website, and the car is delivered to the buyer's doorsteps at no extra cost provided they live within 354km (220 miles) of a Tesla distribution center. Along the same lines, Ford's CEO Jim Farley announced in 2022 a plan to move part of the downstream company's activity online, ending the traditional dealership model and selling directly to consumers at a fixed price.[1] Moving sales online and simplifying the transaction process is part of a larger plan to enforce price transparency and improve consumer convenience and overall purchase experience.[2] Other manufacturers are expected to follow suit if the examples of Tesla and Ford prove successful.

It is well documented that, very often, consumers obtain discounts over the posted prices when buying a new car in person at a car dealer. Through personal interactions, the car dealer observes the consumer's characteristics, forms an expectation about their preferences, and then offers a discount over listed prices (or valuable advantages like free upgrades or an extended warranty). We interpret discounts over listed prices as a form of third-degree price discrimination. In addition, purchasing a car in person entails transportation costs for consumers, typically associated with traveling a certain distance to reach the car dealer and the associated opportunity cost of time.

In this context, a hypothetical online distribution channel introduces a trade-off for consumers. On the one hand, by choosing the online distribution channel, consumers lose a potential discount and are bound to pay the uniform online price. On the other hand, making the transaction online involves a reduction in transportation costs as individuals

[1]Source: Phoebe Ward Howard, "Ford CEO Farley says electric vehicles will be sold 100% online, have non-negotiable price", Detroit Free Press. The full article is available here.

[2]Additional information can be found on Ford's website, see here.

can avoid some or even all the otherwise necessary visits to car dealers.

Using French data for the years 2009–2021, a period in which online car sales were essentially absent, we estimate an equilibrium model of car pricing and sales. We explicitly account for the locations of car dealers and consumers, and their equilibrium effects on firms' pricing and consumers' purchasing decisions. To achieve this, we rely on a novel dataset of car dealer locations and consumers' driving distances in France. We combine it with car registrations, by age and municipality, and various municipal-level demographics. We categorize consumers in demographic groups, based on their age and the median income in their municipality of residence. We assume that these groups are observable by car dealers, who then use this information to engage in third-degree price discrimination.

As in D'Haultfœuille et al. (2019), our model enables us to estimate unobserved price discounts for groups of consumers based on demographic characteristics. This method extends the standard demand estimation approach developed by Berry et al. (1995) to account for unobserved price discrimination. We augment D'Haultfœuille et al. (2019) to infer the extent of third-degree price discrimination in the presence of transportation costs. We model transportation costs similarly to Nurski and Verboven (2016), who assume that they are a function of the driving distance between consumers and the closest car dealer of each car model.

We take advantage of the granularity of the data to augment the standard demand- and supply-side moments with micro moments that match observed average distances with the corresponding model predictions for each demographic group. There are two key advantages to using these additional micro moments. First, we expect these micro moments to be informative about transportation costs. Second, thanks to these micro moments, our estimation method is robust to the potential endogeneity of distance without requiring additional instrumental variables. Distance could be endogenous if, for example, firms took unobserved components of preferences into account at the time of opening new car dealers.

Given our model estimates, we simulate the introduction of an online distribution channel and study its interactions with brick-and-mortar car dealers. In counterfactual experiments, we assume that firms charge a uniform price online while still offering discounts for in-person transactions.[3] We also assume that consumers face reduced transporta-

[3]Following the stated intentions of firms and the practice of Tesla (see discussion above), we favor this assumption over the alternative that firms price discriminate also in the online channel. However, for completeness, we also perform a set of counterfactuals in which firms price discriminate in both distribution channels, see Appendix C.

tion costs when buying online. Since car dealers are important for after-sale services, maintenance, and repairs, we believe that transportation costs may still matter when shopping online, but to a lesser extent than for in-person transactions. We consider various levels of transportation cost reductions in the counterfactual experiments. In the most extreme case, transportation costs are eliminated, implicitly assuming that consumers do not expect future interactions with car dealers. Throughout the analysis, we account for consumer heterogeneity in their propensity to shop online relying on a survey of attitudes toward online purchases by demographic group.

Our results can be summarized by two key findings. Our first finding is that, when all consumers can access the online channel, committing to a uniform online price reduces the extent of in-person discounts. In this scenario, firms have an incentive to set a uniform price in both channels. This makes the online channel unambiguously better for most consumers, since transportation costs are lower online. As a result, we observe a large transfer from the in-person to the online channel and a market expansion.

When, instead, some consumers are restricted in their ability to shop online, the results are different. We observe two forces at play. On the one hand, firms want to continue to offer in-person discounts to extract more surplus from those consumers who cannot access the online channel. On the other hand, the competitive pressure from the online channel instead leads firms to set a uniform price for those consumers who have access to both channels. When the online channel provides a small reduction in transportation costs, the second effect dominates and firms drastically reduce in-person discounts for most consumers. In contrast, when the reduction in transportation costs is large, firms charge low online prices to divert unrestricted consumers online, while continuing to price discriminate the consumers who cannot purchase online. In some sense, the level of transportation cost reductions accruing from shopping online dictates firms' ability to separate the market between consumers who are captive to the in-person channel and those who can take advantage of both channels.

Our second finding relates to the welfare effects of introducing an online channel. We find that price discrimination, taken in isolation, benefits only some consumers, typically the younger and the less wealthy, as they receive discounts over the list prices. The aggregate effect on consumer surplus is a small loss, and the aggregate effect on industry profits is a small gain, in the range of 1%. In contrast, transportation costs are detrimental to all consumers and firms. Reducing transportation costs by 25% already generates a larger increase in consumer surplus than that implied by eliminating price discrimination and a greater increase in industry profits than that implied by price discrimination.

Eliminating transportation costs altogether yields much larger welfare gains (around five times larger), meaning that price discrimination plays a somewhat less prominent role than transportation costs in this industry. An online distribution channel is generally profitable for car manufacturers, as it leads to market expansion and the purchase of more expensive vehicles. Industry profits increase by up to 4% when the online channel brings reductions in transportation costs of at least 25%. In terms of consumer surplus, the online distribution channel instead gives rise to heterogeneous effects, with winners and losers. Consumers who tend to benefit from the introduction of an online channel are older, wealthier, and internet savvy. These consumers would typically not receive in-person discounts in any case, and thus the possibility of saving on transportation costs makes them unambiguously better off. For other consumers, the introduction of an online channel brings either overall losses or only very marginal gains.

Similarly to other empirical papers in the literature (see Nurski and Verboven, 2016), we assume that the network of car dealers remains fixed when the online distribution channel is introduced. Although potential adjustments to the network of car dealers may be important, we maintain this assumption because of the practical infeasibility of incorporating an additional layer of endogenous network formation into an already rich structural model of unobserved third-degree price discrimination and spatial differentiation. In this sense, our results should be interpreted as a collection of short-run responses to the introduction of the online distribution channel. However, to provide some insight into this mechanism, we also consider counterfactuals in which the 5%, 10%, and 20% lowest performing car dealers of each brand exit the market. We find that our main results overstate the gains in consumer surplus and profits by, for example, 1.5% and 1.6%, respectively, compared to a case with 10% exit of car dealers.

In a final set of counterfactuals, we also allow for the possibility that the introduction of an online channel enables firms (car manufacturers) to bypass the double marginalization of car dealers (see Brenkers and Verboven, 2006). In line with intuition, in such a case firms' "effective" marginal costs would decrease, leading equilibrium prices to also decrease, overall car sales to increase, and, as a consequence, both consumer surplus and industry profit to increase. Importantly, the fact that in this case industry profit would increase means that, in theory, there could be ways of redistributing industry profit so to keep car dealers at least as well off as in the scenario with double marginalization.

Related literature. Our research contributes to several branches of the literature. First, it relates to a growing literature on the welfare effects of e-commerce, such as Brown and Goolsbee (2002) on the impact of comparison websites on insurance prices in the US and Morton et al. (2001) on car referral websites (a precursor of online sales in the car industry). Similarly to Pozzi (2013), Fan et al. (2018), and Forman et al. (2009), our study shows that the coexistence of an in-person and an online channel can generate welfare gains through both increased price competition and reductions in transportation costs. However, it also highlights important distributional effects, showing that firms and a small group of inelastic consumers can obtain most of the benefits. Along the lines of Huang and Bronnenberg (2023) and Brynjolfsson et al. (2003), our analysis also illustrates that these forces are closely related to the variety of products available to consumers, highlighting novel mechanisms through which an online distribution channel can limit the ability to price discriminate of brick-and-mortar stores.

Second, we contribute to the literature on price discrimination and price dispersion in retail markets. Seminal work by Corts (1998) and Thisse and Vives (1988) and recent work by Iaria and Wang (2021) and Rhodes and Zhou (2024) provide evidence that price discrimination can intensify competition, can benefit consumers (in the aggregate) and, in some cases, decrease profits in oligopolistic industries. Previous empirical studies on the car industry investigated price discrimination based on consumer demographics, see Ayres and Siegelman (1995), Goldberg (1996), Harless and Hoffer (2002), and Chandra et al. (2017). These studies find contrasting evidence linking price dispersion to demographics (typically gender and race). We provide novel evidence on the relationship among transaction prices, income, and age (we find no relationship with gender), and more broadly on the relationship between price dispersion and spatial differentiation (for related evidence on a homogeneous product, see Miller and Osborne, 2014).

Third, our paper is closely related to recent work studying price discrimination through the lens of structural models, such as D'Haultfœuille et al. (2019) on the French car industry and Sagl (2024) on the trucking industry in Texas. Sagl (2024) finds that most of the observed price dispersion can be explained by consumer unobservables (or soft information), as opposed to demographics. His analysis leverages consumer-level transaction prices and repeated purchases over time. D'Haultfœuille et al. (2019) instead rely on list prices and recover unobserved transaction prices resulting from third-degree price discrimination based on consumer demographics. We contribute to this literature by proposing a unified framework that incorporates (potentially unobserved) transaction prices and spatial differentiation, and that can be used to investigate the relationship between the two in oligopolistic industries with differentiated products.

Fourth, our work relates to recent papers studying price personalization in online markets. Shiller (2020) studies price personalization for Netflix subscription plans based on browsing histories, while Dubé and Misra (2023) study price personalization for a digital firm based on observable consumer characteristics. These studies leverage the vast amount of information available online to investigate the consequences of price personalization for a firm. Instead, we focus on the consequences of online sales for an oligopolistic industry with differentiated products and a long tradition of brick-and-mortar stores. Car manufacturers use the online channel to enforce price transparency rather than personalization, allowing us to deepen our understanding of oligopolistic pricing behavior when both online and in-person distribution channels coexist.

Fifth, we contribute to the literature investigating the role of dealer networks in the car industry. A growing strand of this literature takes the perspective of costly search (e.g., Moraga-González et al., 2023; Murry and Zhou, 2020; Yavorsky et al., 2021), where consumers need to personally visit car dealers to learn about some of the features of car models (or about their very existence), essentially adding them to their consideration sets, and where search costs depend on the distance to car dealers. For the separate identification of search from utility (necessary for counterfactual analyses), the implementation of these structural models typically rules out price discrimination, especially in cases such as ours in which consumer-level transaction prices are not observed (Moraga-González et al., 2023). As the main objective of this paper is to investigate the relationship between price discrimination and spatial differentiation, we follow the route of augmenting the structural model of unobserved price discrimination by D'Haultfœuille et al. (2019) with transportation costs, leaving the important question of also incorporating a search dimension into the framework for future research (for evidence on the relevance of each of these dimensions, see Scott Morton et al., 2011).

Finally, our work is closely related to Duch-Brown et al. (2023), who study the interaction between online and in-person sales in the portable PC industry in Europe. In their application, price dispersion occurs in the online market as a result of geoblocking restrictions on cross-border transactions. They show that banning these restrictions results in the convergence of prices to a unique European-level price for each product sold online. We uncover a similar mechanism when the online channel brings small transportation cost reductions. However, we also show that when the reductions in transportation costs are large, firms tend to direct internet-savvy consumers to that channel with advantageous uniform online prices, but continue also to price discriminate the other consumers in the in-person channel. More broadly, we contribute to Duch-Brown et al. (2023) by specifically investigating the roles of spatial differentiation and transportation costs in the transition to market integration promoted by an online distribution channel.

2 Model

2.1 Demand

We incorporate transportation costs in the model of (unobserved) third-degree price discrimination by D'Haultfœuille et al. (2019). Throughout, we assume that each consumer belongs to one of D mutually exclusive groups based on their observed demographics, and that firms price discriminate by offering different prices to consumers from different groups. We enrich the model explicitly considering that consumers are spatially distributed and face heterogeneous distances to car dealers. This implies heterogeneous transportation costs when purchasing a car of a given brand.

Consider consumer i belonging to demographic group d (an age and income group) and living in municipality m (a town in France). We omit the time subscript for simplicity. Their indirect utility from purchasing car model $j = 1, ..., J$ is

$$U_{ijdm} = \underbrace{X_j' \beta_d + \alpha_d p_{jd} + \xi_{jd}}_{\delta_{jd}} + \underbrace{X_j'(\Pi_d \mathrm{dem}_{dm} + \Sigma_d \nu_i)}_{\mu_{jdm}(\nu_i)} + \gamma_d \mathrm{dist}_{jm} + \epsilon_{ijdm}, \tag{1}$$

where X_j is a vector of observed car characteristics that is invariant across groups (e.g., horsepower), p_{jd} is the (unobserved) transaction price faced by group d, and ξ_{jd} captures the average indirect utility of the car characteristics unobserved by the econometrician. Note that both p_{jd} and ξ_{jd} do not vary geographically. As we discuss in Section 2.3, these restrictions relate to data availability and identification in the context of unobserved transaction prices.

The term $\mu_{jdm}(\nu_i)$ captures individual-level deviations in the preferences for X_j from the group average $X_j'\beta_d$. The vector dem_{dm} collects observable average demographics (e.g., income, household size, an urban indicator) specific to the individuals of group d living in municipality m, while ν_i is a vector of random coefficients.[4] Variable dist_{jm} is the average driving distance from municipality m to the closest dealer selling car model j.[5] Variable ϵ_{ijdm} is an idiosyncratic error term assumed to be distributed extreme value type I. Indirect utility (1) is analogous to that in D'Haultfœuille et al. (2019), with the exception of the additional term $\gamma_d \mathrm{dist}_{jm}$, which represents the transportation cost for

[4] We also investigated specifications with heterogeneous coefficients on price within each demographic group d, but we did not obtain qualitatively different results. Given the higher computational complexity of these specifications (more on this below), we then opted for the simpler indirect utility (1).

[5] As we explain in detail in Section 3.3, we compute the driving distance from each "housing" building (i.e., we exclude the buildings in which people do not live, such as airports or other businesses) in municipality m to the closest car dealer of model j and then compute dist_{jm} as the average across all such distances.

consumers of group d living in municipality m of traveling to the closest car dealer selling j. The demand parameters $(\beta_d, \alpha_d, \Pi_d, \Sigma_d, \gamma_d)$ are allowed to be heterogeneous across demographic groups.

As detailed in Section 3.3, we observe car purchases at the level of the municipality m (a town in France) by demographic group d. Given indirect utility (1), the probability that a consumer in demographic group d and municipality m purchases car model j is

$$s_{jdm}(\delta_d, \Pi_d, \Sigma_d, \gamma_d) = \int \frac{\exp(\delta_{jd} + \mu_{jdm}(\nu_i) + \gamma_d \text{dist}_{jm})}{1 + \sum_{k=1}^{J} \exp(\delta_{kd} + \mu_{kdm}(\nu_i) + \gamma_d \text{dist}_{km})} dF(\nu_i), \qquad (2)$$

where $\delta_d = (\delta_{1d}, ..., \delta_{Jd})$ and F is the distribution of the random coefficients ν_i.

As in D'Haultfœuille et al. (2019), we assume that each transaction price p_{jd} is chosen at the national level (more details on this below). To obtain the national-level market share of car model j for group d, we average (2) over municipalities:

$$s_{jd}(\delta_d, \Pi_d, \Sigma_d, \gamma_d) = \sum_{m \in \mathcal{M}} w_{dm} \cdot s_{jdm}(\delta_d, \Pi_d, \Sigma_d, \gamma_d), \qquad (3)$$

where $\mathcal{M}$ collects all municipalities. The relative weight $w_{dm} \equiv M_{dm}/M_d$ measures the incidence of demographic group d in municipality m relative to the country, where M_{dm} and M_d are the observed numbers of consumers of group d in municipality m and throughout the country, respectively, with $\sum_{m \in \mathcal{M}} M_{dm} = M_d$.

2.2 Supply

We consider a Bertrand-Nash price-setting game in which firms are able to implement third-degree price discrimination by choosing different prices for each demographic group $d = 1, ..., D$. Each firm $f = 1, ..., F$ selects a menu of national prices $p_j = (p_{j1}, ..., p_{jD})$ for each car model j they sell by maximizing the national-level profit function

$$\pi_f = \sum_{d=1}^{D} \phi_d \sum_{j \in \mathcal{J}_f} s_{jd}(p_d) \cdot (p_{jd} - c_{jd}), \qquad (4)$$

where $\mathcal{J}_f$ is the collection of car models sold by firm f, $p_d = (p_{1d}, ..., p_{Jd})$ is the vector of all prices for group d, $s_{jd}(p_d)$ is the market share defined in (3) (where we highlight its dependence on p_d), and c_{jd} is the marginal cost of car model j for group d. Variable $\phi_d \equiv M_d/M$ is the observed group-specific population weight, where M_d is the national population of group d and M is the national population (over all groups).

The system of J first-order conditions associated with demographic group d are

$$p_d = c_d - \widetilde{\mathcal{D}}_d(p_d, \mathcal{H})^{-1} s_d, \tag{5}$$

where c_d is the vector of group-specific marginal costs of all car models, s_d is the vector of group-specific market shares for all car models, $\widetilde{\mathcal{D}}_d(p_d) = \mathcal{H} \odot \mathcal{D}_d(p_d)$, $\mathcal{H}$ is the ownership matrix, and $\mathcal{D}_d(p_d)$ is the matrix of derivatives of s_d with respect to p_d, with typical element (j,k) equal to $\partial s_{kd}/\partial p_{jd}$.

2.3 Identification and estimation

Identification of the unobserved transaction prices. Compared to standard demand models, the identification of $\theta = (\theta_1, ..., \theta_D)$, where $\theta_d \equiv (\beta_d, \alpha_d, \Pi_d, \Sigma_d, \gamma_d)$, presents the additional complication that the group-specific transaction prices $p(\theta) = (p_1(\theta), ..., p_D(\theta))$ are not observed by the econometrician. Following D'Haultfœuille et al. (2019), we address this complication by relying on both demand and supply restrictions to jointly identify preference parameters θ and transaction prices $p(\theta)$.[6]

We make the following assumptions, which are sufficient for the identification of the unobserved transaction prices (see details in D'Haultfœuille et al., 2019).

A1. Observability of national-level group-specific market share s_{jd} for each j and d.

A2. Constant marginal costs across demographic groups, $c_{jd} = c_j$ for any j and d.

A3. Relevance of list prices. For each car model j, we assume that the list price $\bar{p}_j$ satisfies $\bar{p}_j = \max\{p_{j1}, ..., p_{jD}\}$: it is the highest transaction price the demographic groups can pay at any car dealer.

Intuitively, assumptions A1-A3 allow us to back out, for any given value of the demand parameters θ, the transaction prices $p(\theta)$ that rationalize both demand and supply. With these, we are back to a standard model in which all prices are observed.

First, given assumption A1 and following Berry (1994), for a given group d and some value of $(\Pi_d, \Sigma_d, \gamma_d)$, we obtain $\delta_d(\Pi_d, \Sigma_d, \gamma_d) = (\delta_{jd}(\Pi_d, \Sigma_d, \gamma_d))_{j=1,...,J}$ by inverting the system of J national-level market share equations given by (3). Second, we obtain the transaction prices corresponding to θ. To this end, note that the first-order conditions

[6] Note that, even if transaction prices were observed, without further "extrapolating" assumptions, the prices that consumers face for the car models they do *not* purchase remain unobserved and a similar procedure would still be needed.

(5) and assumptions A2-A3 imply

$$\bar{p}_j = c_j - \min \left\{ \left[\widetilde{\mathcal{D}}_1{}^{-1} s_1 \right]_j, ..., \left[\widetilde{\mathcal{D}}_D{}^{-1} s_D \right]_j \right\}, \tag{6}$$

with $[x]_j$ denoting the j-th element of vector x. In the absence of random coefficient on prices, the matrices $\widetilde{\mathcal{D}}_1, ..., \widetilde{\mathcal{D}}_D$ do not depend on transaction prices: they only depend on $\varsigma \equiv (\alpha_d, \Pi_d, \Sigma_d, \gamma_d)_{d=1,...,D}$ through $(\alpha_d, \delta_d(\Pi_d, \Sigma_d, \gamma_d))_{d=1,...,D}$. Then, for given value of ς, each transaction price can be obtained as

$$p_{jd}(\varsigma) = \bar{p}_j + \min \left\{ \left[\widetilde{\mathcal{D}}_1^{-1} s_1 \right]_j, ..., \left[\widetilde{\mathcal{D}}_D^{-1} s_D \right]_j \right\} - \left[\widetilde{\mathcal{D}}_d^{-1} s_d \right]_j, \tag{7}$$

where we let the dependence of $\widetilde{\mathcal{D}}_d$ on ς implicit.

Assumption A1 relates to data availability and it is the main reason for our modeling choice that p_{jd} and ξ_{jd} vary at the level of (j, d) rather than at the more disaggregate level of, say, (j, d, r), where r denotes a specific car dealer. In order to handle unobserved transaction prices at this level of detail, one would need to observe the specific car dealer r in which each consumer of group d purchased car model j. In other words, one would need precise measures of the market shares at the level of (j, d, r), which are currently unavailable. Even having access to disaggregate r-specific purchase data, because of the limited number of sales of each car dealer, market shares would be imprecisely calculated, with severe consequences in terms of measurement error which cannot be easily addressed in nonlinear structural models (see Freyberger (2015); Gandhi et al. (2019) and the discussion below).

Demand-side moments. We compute the empirical counterpart of the following moment conditions

$$\mathbb{E} \left[Z'_{jd} \xi_{jd} \right] = 0, \quad d = 1, ..., D, \tag{8}$$

with Z_{jd} a group-specific vector of instruments. To do this, we first compute $\xi_{jd}(\beta_d, \varsigma) = \delta_{jd}(\Pi_d, \Sigma_d, \gamma_d) - X'_j \beta_d - \alpha_d p_{jd}(\varsigma)$ and then consider the empirical moment condition $g_1(\theta) = (g_{11}(\beta_1, \varsigma), ..., g_{1D}(\beta_D, \varsigma))$, where

$$g_{1d}(\beta_d, \varsigma) = \frac{1}{J} \sum_{j=1}^{J} Z'_{jd} \xi_{jd}(\beta_d, \varsigma) = 0. \tag{9}$$

As usual, while transaction prices are endogenous by construction, we assume the observed characteristics X_j to be exogenous. Valid instruments can be obtained as functions of the exogenous characteristics of car model j and those of other car models.

Supply-side moments. We assume that the log of marginal cost c_j is a linear combination of observed car characteristics X_j, cost shifters W_j, and an unobserved cost shock ω_j, such that

$$\ln(c_j) = X_j'\lambda_1 + W_j'\lambda_2 + \omega_j, \tag{10}$$

where we assume that (X_j, W_j) are uncorrelated with respect to ω_j and to $(\xi_{jd})_{d=1,\dots,D}$. We compute the empirical counterpart of the following moment conditions

$$\mathbb{E}\left[Z_{jS}'\omega_j\right] = 0, \tag{11}$$

with Z_{jS} a vector of supply-side instruments. The associate supply-side moment conditions are

$$g_2(\varsigma, \lambda) = \frac{1}{J}\sum_{j=1}^{J} Z_{jS}'\omega_j(\varsigma, \lambda) = 0, \tag{12}$$

where $\lambda = (\lambda_1, \lambda_2)$ and $\omega_j(\varsigma, \lambda)$ can be computed using (6) and (10). Again, valid instruments Z_{jS} can be obtained as functions of (X_j, W_j) and of (X_k, W_k), $k \neq j$.

Micro moments. We complement moment conditions (8) and (11) with micro moments that help identify the nonlinear parameters $(\Pi_d, \Sigma_d, \gamma_d)$. To do this, we take advantage of the fact that we observe purchase probabilities at the demographic group-by-municipality level, s_{jdm}. We illustrate here the idea of these micro moments for the variable dist_{jm} and follow a similar procedure for the demographics dem_{dm} and their interactions with X_j. The full list of micro moments is then presented in Section 3.5.

We construct micro moments for dist_{jm} by matching the observed and predicted average distance between consumers and car dealers of the purchased car models. In particular, we specify $g_3(\Pi, \Sigma, \gamma) = (g_{31}(\Pi_1, \Sigma_1, \gamma_1), \dots, g_{3D}(\Pi_D, \Sigma_D, \gamma_D))$, with

$$\begin{aligned} &g_{3d}(\Pi_d, \Sigma_d, \gamma_d) \\ &= \frac{\sum_{j=1}^{J}\sum_{m\in\mathcal{M}} w_{dm}\cdot s_{jdm}(\Pi_d, \Sigma_d, \gamma_d)\cdot \text{dist}_{jm}}{\sum_{j=1}^{J}\sum_{m\in\mathcal{M}} w_{dm}\cdot s_{jdm}(\Pi_d, \Sigma_d, \gamma_d)} - \frac{\sum_{j=1}^{J}\sum_{m\in\mathcal{M}} w_{dm}\cdot s_{jdm}\cdot \text{dist}_{jm}}{\sum_{j=1}^{J}\sum_{m\in\mathcal{M}} w_{dm}\cdot s_{jdm}}, \end{aligned} \tag{13}$$

where s_{jdm} is the observed market share of group d for product j in municipality m. The corresponding market share predicted by the model is instead obtained evaluating equation (2) at any given $(\Pi_d, \Sigma_d, \gamma_d)$: $s_{jdm}(\Pi_d, \Sigma_d, \gamma_d) = s_{jdm}(\delta_d(\Pi_d, \Sigma_d, \gamma_d), \Pi_d, \Sigma_d, \gamma_d)$, where we obtain $\delta_d(\Pi_d, \Sigma_d, \gamma_d)$ by inverting the system of J national-level market shares in (3). Note that for any given $(\Pi_d, \Sigma_d, \gamma_d)$, we have $s_{jd}(\Pi_d, \Sigma_d, \gamma_d) = s_{jd}$ and the denominators on the right-hand side of (13) are equal. However, in general, only at the true value of $(\Pi_d, \Sigma_d, \gamma_d)$, say $(\Pi_d^0, \Sigma_d^0, \gamma_d^0)$, we also have $s_{jdm}(\Pi_d^0, \Sigma_d^0, \gamma_d^0) = s_{jdm}$.

Three points are worth noting. First, if we only have a few moment conditions in (8) and (11), we may be unable to identify $(\Pi_d, \Sigma_d, \gamma_d)$. In such cases, additional moments, such as the micro moments we propose, become necessary to identify $(\Pi_d, \Sigma_d, \gamma_d)$. Moreover, the instruments in (8) and (11) may lack the power to estimate with sufficient precision the distance parameter γ_d (for example). Instead, we expect the micro moments in (13) to be informative about γ_d as, intuitively, the observed average distance between consumers and the car dealers of the purchased models is monotonic in γ_d.

Second, these micro moments remain valid even if distance dist_{jm} is endogenous. By this, we mean that distances could be correlated with $(\xi_{jd})_{j,d}$: for example, if firms partially or fully observed $(\xi_{jd})_{j,d}$ at the moment of deciding where to locate their car dealers. As mentioned above, at the true value of $(\Pi_d, \Sigma_d, \gamma_d)$, say $(\Pi_d^0, \Sigma_d^0, \gamma_d^0)$, $\delta_{jd}(\Pi_d, \Sigma_d, \gamma_d)$ will equal its true value $\delta_{jd}(\Pi_d^0, \Sigma_d^0, \gamma_d^0) = \delta_{jd}^0$, and thus $s_{jdm}(\Pi_d^0, \Sigma_d^0, \gamma_d^0) = s_{jdm}^0$. This only follows from Berry (1994)'s demand inverse and therefore holds irrespective of any dependence between the distances and the unobserved components of demand $(\xi_{jd})_{j,d}$.

Third, assuming that s_{jd} is measured without error is reasonable (and standard) given the large number of consumers in each demographic group throughout the country. However, the assumption that we also perfectly observe the municipality-level market shares s_{jdm} may be strong. We only observe a proportion on a finite sample instead of the true purchase probability (say, s_{jdm}^0), and the corresponding sample is small for small municipalities. However, we still have

$$\mathbb{E}\left[\sum_{j=1}^{J}\sum_{m\in\mathcal{M}} w_{dm}\cdot s_{jdm}\cdot \text{dist}_{jm} \;\middle|\; (\text{dist}_{jm})_{j,m}, (\xi_{jd})_{j,d}\right] = \sum_{j=1}^{J}\sum_{m\in\mathcal{M}} w_{dm}\cdot s_{jdm}^0\cdot \text{dist}_{jm}.$$

Hence, even if s_{jdm} is measured with error (but maintaining, as usual, that the $(s_{jd})_{j=1,\dots,J}$ are measured without error), at the true value of the nonlinear parameters $(\Pi_d^0, \Sigma_d^0, \gamma_d^0)$, we still obtain $\mathbb{E}[g_{3d}(\Pi_d^0, \Sigma_d^0, \gamma_d^0)] = 0$. In other words, the micro moments in (13) are robust to this form of measurement error. For the same reason, with these micro moments, "zeros" in the observed market shares at the level of the municipality do not raise any concern (Gandhi et al., 2019).

Two-step GMM Estimation. Although the parameters (θ, λ) could be simultaneously and efficiently estimated by the generalized method of moments (GMM) using all moment conditions $g(\theta, \lambda) = (g_1(\theta), g_2(\varsigma, \lambda), g_3(\Pi, \Sigma, \gamma))$, this would be computationally intensive as, at each guess of θ, one would need to solve: (i) demand inverses to determine the average group-specific indirect utilities and (ii) the system of equations (7) to determine the group-specific transaction prices (D'Haultfœuille et al., 2019).

Given the absence of a random coefficient on price in the specification of indirect utility (1), where the price coefficient only varies across demographic groups (but not within each group), we can, however, greatly simplify implementation and estimate (θ, λ) in two sequential GMM steps. Intuitively, the two sequential estimation steps "separate" the computationally intensive tasks of repeatedly solving for the demand inverses (only in the first step) and of repeatedly solving for the unobserved transaction prices (only in the second step).

First, we use the micro moments $g_3(\Pi, \Sigma, \gamma)$ to estimate the nonlinear parameters (Π, Σ, γ). Second, given $(\hat{\Pi}, \hat{\Sigma}, \hat{\gamma})$, we estimate the remaining demand parameters $(\beta, \alpha) = (\beta_d, \alpha_d)_{d=1,..,D}$ and the marginal cost parameters $\lambda = (\lambda_1, \lambda_2)$ using the remaining moment conditions $(g_1(\beta, \alpha, \hat{\Pi}, \hat{\Sigma}, \hat{\gamma}),\ g_2(\alpha, \hat{\Pi}, \hat{\Sigma}, \hat{\gamma}, \lambda))$. Compared to the usual BLP estimator, this step involves the additional computation of transaction prices using (7). Importantly, because the average group-specific utilities $\delta_d(\Pi_d, \Sigma_d, \gamma_d)$ are fully determined by $(\Pi_d, \Sigma_d, \gamma_d)$, this second step takes them as given.

As each step of this two-step GMM estimation procedure is standard, we report all computational details in Appendix B.2. To account for the estimation error in $(\hat{\Pi}, \hat{\Sigma}, \hat{\gamma})$ arising from the first-step GMM estimation, we compute the variance-covariance matrix of the second-step GMM estimator of (β, α, λ) using the formulae in Newey and McFadden (1994). Clearly, because the moment conditions $(g_1(\theta),\ g_2(\varsigma, \lambda))$ used in the second step carry information about the nonlinear parameters (Π, Σ, γ), this two-step GMM is less efficient than the alternative (and more standard) one-step GMM that uses all moment conditions $g(\theta, \lambda)$ simultaneously. As mentioned in footnote 4, after extensive investigations of more general specifications incorporating a random coefficient on price (and estimated by a one-step GMM), we obtained qualitatively similar results and, given the substantially higher computational complexity, we opted for the simpler indirect utility in (1) and the two-step GMM procedure.

3 Data and estimation results

3.1 Demographic groups definition

We divide consumers into three age categories (below 40, 40 to 59, and above 60) and two income categories (low and high income) to form six demographic groups. These groups are easily observable by car dealers and potentially associated with heterogeneous preferences, thus forming a basis for third-degree price discrimination. We do not observe consumers' income directly in the car registration data, so we assign an income category based on their age and municipality of residence. Within each age category, we evenly divide municipalities into low- and high-income classes based on the municipality-specific median income. As a consequence, consumers of the same age and living in the same municipality are assigned to the same group. However, a given municipality could be considered high income in one age category and low income in another.

To characterize the set of potential car buyers, we assemble a rich dataset from the National Institute of Statistics and Economic Studies (INSEE). Our data include yearly population censuses, income by age category, and a survey of consumers' attitudes towards online purchases.[7] This survey collects data on a representative sample of French individuals about their use of information and communication technologies, including online sales platforms. We summarize these data at the level of our demographic groups in Appendix Table A.1 and we provide additional details on the construction of the final datasets in Appendix B.

3.2 Evidence of price dispersion

In this section, we provide evidence that income and age are the most relevant observable demographics that correlate with price dispersion in the French car industry.

We combine two waves of a French survey of consumers' expenditures that contain both consumers' demographic characteristics and the transaction prices of their most recent car purchases.[8] In these surveys, car purchases are divided into new and second-hand vehicles, and we can distinguish sales that occurred at a car dealer versus sales that occurred between consumers. Whenever a consumer resold their old car in the same year, the trade-in value is also recorded. We estimate a regression of the transaction prices paid by consumers who purchased directly from a car dealer on a rich set of

[7] Source: "lil-1407 : Technologies of l'information et de la communication auprès des ménages (TIC) - 2019 (2019, INSEE)," accessed from Progedo Adisp.

[8] Source: "Enquête Budget des Familles (BDC) - 2011–2017 (2011–2017, INSEE)."

Table 1: Evidence of price dispersion

	Transaction price		Transaction price net of buyback value	
	(1)	(2)	(3)	(4)
Income	40.497***	25.916**	39.253***	28.970**
	(13.112)	(11.958)	(13.760)	(14.533)
Age	67.387***	20.156	47.276***	2.754
	(15.828)	(19.359)	(17.284)	(25.626)
Female	851.802	141.661	1,051.810	604.796
	(1,078.583)	(1,463.399)	(1,147.382)	(1,893.736)
Age × Female	−17.770	−5.874	−25.134	−13.574
	(19.998)	(26.556)	(21.727)	(32.152)
Value of down payment	0.006	0.005	0.013	0.011**
	(0.005)	(0.003)	(0.008)	(0.005)
Household: 2 pers.	−182.571	−250.977	−478.245	−707.159
	(393.500)	(633.025)	(500.774)	(815.847)
Household: 3 pers.	−688.958	−786.539	−954.493	−1,117.172
	(620.113)	(925.880)	(709.379)	(1,208.438)
Household: 4 pers.	−644.082	−1,357.238*	−1,464.877**	−2,485.373**
	(626.979)	(802.681)	(655.590)	(1,044.736)
Household: 5 pers.	−3,000.296***	−2,397.637**	−2,987.593***	−4,892.123***
	(882.021)	(1,109.743)	(962.957)	(1,589.441)
Household: 6+ pers.	−202.262	−929.100	1,490.901	−399.893
	(2,294.165)	(2,114.791)	(2,113.088)	(2,009.281)
Urban area: less than 15,000	−825.524	2,275.435	−1,235.378	2,558.358
	(1,425.966)	(1,552.297)	(2,290.406)	(3,842.866)
Urban area: 15,000–24,999	345.570	548.107	1,610.750	2,790.948
	(1,717.314)	(1,827.368)	(1,458.979)	(2,818.982)
Urban area: 25,000–34,999	−1,588.274	1,024.648	−1,243.062	1,283.947
	(1,377.100)	(2,077.463)	(1,659.565)	(3,108.214)
Urban area: 35,000–49,999	−1,733.160	−6.787	−1,418.662	1,805.762
	(1,095.743)	(1,002.154)	(1,220.085)	(1,598.266)
Urban area: 50,000–99,999	−1,316.561	−462.050	−1,999.194**	−1,382.434
	(815.740)	(1,136.611)	(785.564)	(1,128.750)
Urban area: 100,000–199,999	−822.825	195.033	−198.663	222.065
	(791.007)	(930.849)	(714.154)	(1,036.763)
Urban area: 200,000–499,999	−1,093.716	182.223	−864.898	−6.343
	(697.664)	(876.896)	(608.700)	(1,023.145)
Urban area: 500,000 or more	−1,143.714*	−366.690	−795.605	197.814
	(657.948)	(888.946)	(643.131)	(882.634)
Urban area: Paris greater metro area	−900.305	−523.017	−121.970	189.700
	(731.188)	(1,140.885)	(676.618)	(990.022)
New vehicles only	No	Yes	No	Yes
Fixed effects				
Car model × engine × new/used	Yes	Yes	Yes	Yes
Year × month of purchase	Yes	Yes	Yes	Yes
Country of origin of buyer	Yes	Yes	Yes	Yes
Observations	1,283	698	1,283	698
R-squared	0.742	0.795	0.600	0.617

Notes: This table presents the result of a regression of transaction prices on demographic characteristics of buyers, based on a survey of consumers' expenditures. We have excluded observations where the car was purchased following an insurance claim (i.e., the replacement of a damaged vehicle). Columns (1) and (3) include sales of both new and used cars, purchased at a car dealer. Columns (2) and (4) include only new cars. The buyback value represents the payment that was received by the consumer for trading in their old car. Standard errors in parenthesis are clustered at the car model × engine × new/used level. Significance: * < 0.10, ** < 0.05, *** < 0.01.

consumers' demographics. We focus on two different measures of transaction price: the transaction price paid by the consumer and the transaction price net of the buyback value. Finally, we include a rich set of fixed effects to control for product characteristics and seasonality. The results are presented in Table 1.

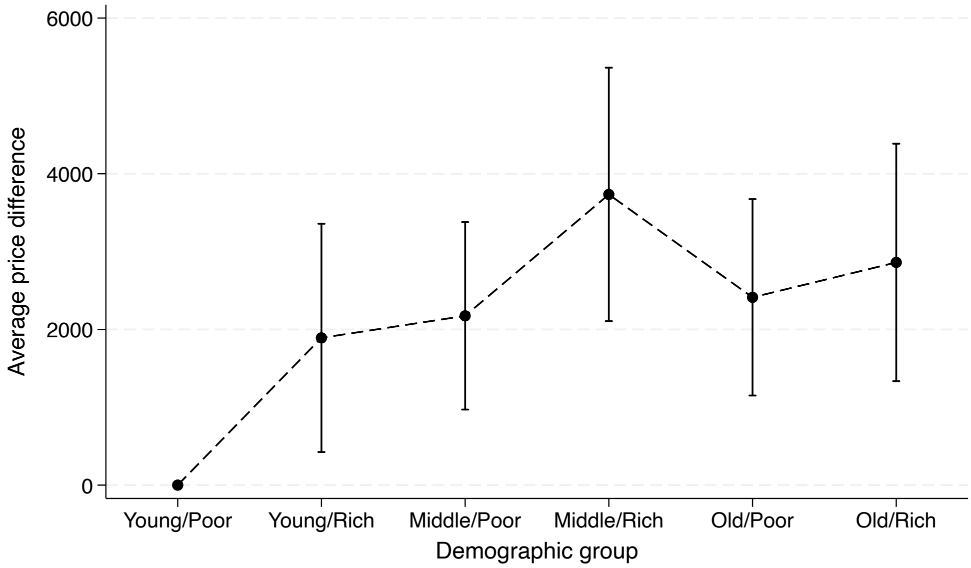
Our estimates indicate that income correlates positively with transaction prices. Since our regressions include model-by-engine-by-new/used fixed effects, this means that high-income consumers pay more on average for the same new/used car model and engine type. We cannot rule out that the effect is partly driven by the choice of additional options as these are unobserved to us, even though it is not clear in which direction this could bias our estimates. On the one hand, price dispersion could be explained by the fact that wealthier individuals purchase similar vehicles but with more expensive options. However, on the other hand, price dispersion could be underestimated if car dealers provide additional options at no cost to low-income consumers and both groups buy similar vehicles with similar options.

Our results suggest that age also correlates positively with transaction prices, although the effect is not statistically significant if we focus only on new vehicles. However, we do not find a statistically significant correlation for other observable (by car dealers) demographic characteristics, namely gender, household size, and the level of urbanity in the consumer's municipality of residence.[9] The fact that we do not find a statistically significant correlation with gender is not surprising: in most cases, purchasing a car is a decision that is taken at the level of the household, and either partner or both partners could have visited the car dealer. In this case, transaction prices are not expected to correlate with the gender of the main respondent in the survey.

To further motivate the choice of our demographic groups in the structural model, we divide the respondents in the survey by age (three groups, as defined above) and income (income above or below the median by age group) and estimate a regression of transaction prices on these demographic group indicators. Again, we control for gender, household size, level of urbanity, and a rich set of fixed effects. Figure 1 shows the estimated coefficients associated with these demographic group indicators, while the details of the regression results are reported in Appendix Table A.2. Our results reveal the presence of price dispersion in terms of income and age. Middle-aged consumers with

[9] Some specific household sizes correlate to price dispersion (four and five components); however, these specific household sizes do not seem observable by car dealers: it may perhaps be possible to infer whether a household has no children, but not whether it has five rather than six or more components. All regressions include indicator variables for the country of origin (not reported in Table 1), also mostly statistically insignificant.

Figure 1: Evidence of price dispersion among demographic groups



Notes: This figure presents the result of a regression of transaction prices on demographic group indicators, based on a survey of consumers' expenditures (see Table A.2, column 1). We exclude observations in which the vehicle was purchased following an insurance claim (i.e., the replacement of a damaged vehicle). The regression includes demographic characteristics (gender, household size, urbanity, country of origin of buyer), car model × engine × new/used and year × month fixed effects. The brackets represent the 95% confidence interval, clustered at the car model × engine × new/used level.

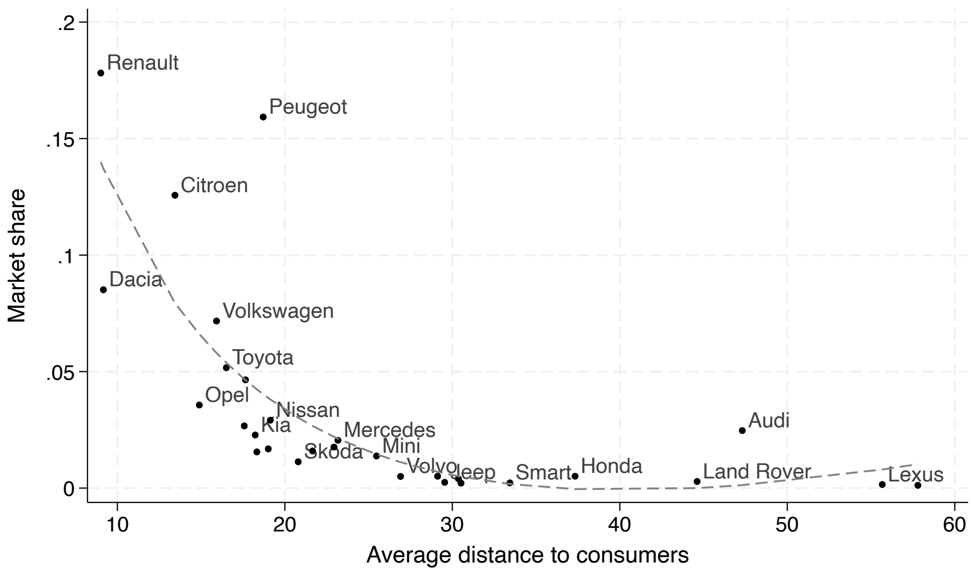
high income face the highest transaction prices, followed by old consumers with high income. In contrast, young and poor consumers pay on average less for the same car models.

3.3 Data and descriptive statistics

Distance to car dealers. We assemble a novel dataset of car dealer locations in France, collected directly from each manufacturer's website. The data were collected in 2024 and include the name of each car dealer in France, its address (which we converted to coordinates), and the associated brands sold. The data include 7,241 dealer-brand combinations.

To compute the average driving distance from consumers living in any given municipality to the car dealers of each brand, we rely on data from OpenStreetMap which include the coordinates and sizes of all buildings in France. For each municipality m, we randomly select 100 buildings without replacement that could be inhabited and calculate

Figure 2: Market share by brand and dealer proximity



Notes: This figure presents the relationship between brands' total market share and their market presence. The market share of each brand is computed as the ratio of its total sales to the total sales of all brands between 2009 and 2021. Market presence is measured as the average distance to consumers over the same period. The dashed line represents the fitted values of a fractional polynomial regression. Additional descriptive statistics are reported in Table A.3.

the driving distance from these buildings to each car dealer.[10] We then compute the average driving distance from m to every car dealer selling j by weighing each building-specific driving distance by the size of the building. Doing so, we obtain a list of average driving distances from m to all car dealers selling j: $\text{dist}^1_{jm}, \text{dist}^2_{jm}, \text{dist}^3_{jm}, \ldots$ Finally, we set $\text{dist}_{jm} = \min\{\text{dist}^1_{jm}, \text{dist}^2_{jm}, \text{dist}^3_{jm}, \ldots\}$, the average driving distance from m to the closest car dealer selling j.

Driving distances are obtained in two steps. First, we use the coordinates of buildings and car dealers to calculate the linear distances for all pairs. Second, for a subset of building-dealer combinations, we perform 969,455 queries on TomTom's API[11] to recover effective driving distances and fit fractional polynomial regressions to convert all linear distances into driving distances.

We provide an overview of the importance of each brand and its proximity to consumers in Figure 2. We plot the aggregate market share of each brand (excluding the outside op-

[10] We exclude all buildings of size below 25 sq. meters or above 500 sq. meters from the set of buildings from which we sample.

[11] See https://developer.tomtom.com/.

tion) against the average driving distance to consumers (over all demographic groups).[12] Brands with a large share of the market are typically located closer to consumers: they operate at more locations which increase proximity to potential buyers. We provide additional information on the market presence of brands in Appendix Table A.3.

Car registrations. We obtain information on all new car registrations in France, between 2009 and 2021, from AAA Data.[13] The data are aggregated at the municipality-level and by age group (in increments of 5 years). There are on average 1,350 inhabitants per municipality and there are 35,296 municipalities in Metropolitan France (Mainland European France).

For each municipality-by-age group, sales are recorded at the level of the brand (29 brands), model (372 models), engine type (gas, diesel, electric, plug-in, hybrid), and body trim (sedan, convertible, station wagon). The data include common car attributes such as horsepower, weight, CO_2 emissions, and fuel consumption, as well as the list price. These car attributes are collected by AAA Data from car manufacturers' catalogs. We complement the dataset with annual average fuel and electricity prices to construct a measure of driving cost (in euros per 100km). Finally, we obtain the market segment (e.g., subcompact, compact, SUV, etc.) of each car model from Jato Dynamics.[14]

We define a product as a combination of a brand, a model, an engine type, and a body trim. After aggregating by product, demographic group, and year, the final dataset includes 4,975 observations over 13 years. Whenever the data are available at a more disaggregate level than our product definition, we keep the characteristics of the most frequently purchased option. List prices are adjusted to be net of fees and rebates tied to the French Feebate Program.[15] Both list prices and driving costs are deflated to 2018 euros. We encountered some missing observations on key car characteristics (namely, horsepower for electric vehicles). In these cases, we filled the missing values with additional data from the French National System of Vehicle Registration (SIV).

Descriptive statistics on car sales are presented in Table 2. In the first panel, we break-down sales by product and demographic group. Groups 4 and 6, representing high-income consumers, aged 40 years old or older, purchase on average more than twice the number of vehicles than other groups.

[12] A similar figure can be obtained by plotting market shares against the number of car dealers of each brand.

[13] Source: https://www.aaa-data.fr/.

[14] Source: https://www.jato.com.

[15] The French Feebate Program offers incentives to promote low-emission vehicles, based on engine type and tailpipe emissions.

Table 2: Car characteristics

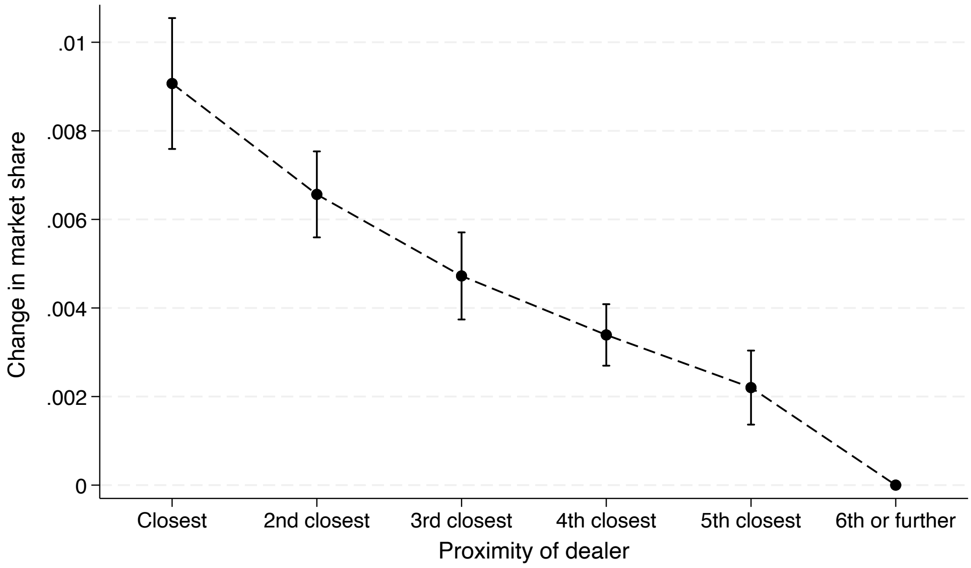
Description	Mean	Std. dev.	10th pct.	Median	90th pct.	Observations
Sales						
Group 1: Young/Poor	283	651	9	66	736	4,975
Group 2: Young/Rich	242	561	10	60	594	4,975
Group 3: Middle/Poor	393	850	22	111	968	4,975
Group 4: Middle/Rich	707	1,463	46	213	1,787	4,975
Group 5: Old/Poor	282	739	10	55	649	4,975
Group 6: Old/Rich	745	1,776	33	174	1,700	4,975
Distance to dealers, km						
Group 1: Young/Poor	15.9	18.3	3.0	7.9	39.8	627,473
Group 2: Young/Rich	14.8	15.1	2.9	10.3	31.3	590,208
Group 3: Middle/Poor	19.2	19.8	3.1	10.7	46.0	582,610
Group 4: Middle/Rich	15.3	16.2	3.2	10.1	32.4	610,218
Group 5: Old/Poor	22.7	19.0	3.4	18.9	47.7	579,101
Group 6: Old/Rich	14.2	15.9	3.0	8.4	32.1	617,178
Car characteristics						
Net list price, in €	22,640	9,581	12,718	20,400	33,591	4,975
Horsepower, in kW	75.5	22.9	51.0	70.0	103.0	4,975
Weight, in kg	1,736	274	1,418	1,700	2,080	4,975
Fuel cost, in €/100km	6.5	1.6	4.8	6.5	8.4	4,975
Fuel consumption, in L/100km	4.7	0.9	3.7	4.7	5.9	4,975
Gasoline	0.49	0.50	0	0	1	4,975
Diesel	0.45	0.50	0	0	1	4,975
Electric	0.02	0.12	0	0	0	4,975
Plug-in hybrid	0.00	0.07	0	0	0	4,975
Hybrid	0.04	0.19	0	0	0	4,975
Sedan	0.73	0.44	0	1	1	4,975
Convertible	0.01	0.08	0	0	0	4,975
Station wagon	0.27	0.44	0	0	1	4,975

Notes: Sales are aggregated at the national level, by product, year, and demographic group. The associated statistics are unweighted. Distance to dealers is the driving distance to the closest dealer of each brand by demographic group, and the associated statistics are weighted by brand importance and municipal-level group-specific populations. All other statistics are weighted by total sales (over all groups). All monetary values are in 2018 euros.

The second panel reports statistics related to how far car dealers are from consumers in terms of driving distance. These statistics are not weighted by group-specific sales. Instead, we weigh them by brand importance (total sales of each brand) and group-specific municipality-level populations. We do this to preserve comparability across groups. Consumers belonging to groups 3 and 5 (low-income, aged 40 or older) live significantly further away from car dealers than other groups; a large share of these consumers live in rural areas compared to other groups. In addition, car dealers are on average slightly closer to high-income than to low-income municipalities.

Finally, the third panel presents a summary of the car characteristics included in the utility specification of our demand model. Since all consumers face the same set of products, we provide a common set of statistics, weighted by total sales.

Figure 3: Market share advantage from car dealer proximity



Notes: This figure presents the estimates of a regression of market shares at the level of the brand, year, and municipality, on brand proximity indicators. The regression controls for municipality and brand × year fixed effects. The brackets represent the 95% confidence interval, clustered at the municipality level. The calculation of market shares excludes the outside option and pools sales from all products within the same brand and all demographic groups.

3.4 Evidence of transportation costs

We provide evidence that the distance from car dealers, which we use as a proxy for transportation costs, matters to consumers. We estimate a regression of the market share of each brand at the municipality level on proximity indicators for car dealers, controlling for municipality and brand-by-year fixed effects. Figure 3 plots the estimated coefficients of the proximity indicators and shows that, according to intuition, being geographically closer to consumers is positively correlated with market shares. The figure illustrates that the market share of, say, Renault is larger in municipalities where it is the closest dealer versus the second closest dealer, the second closest dealer versus the third, the third closest dealer versus the fourth, and so on.

3.5 Estimation results

We now present the estimation results of our model. As mentioned above, we define products as brand, model, engine type, and body trim combinations. We consider each year to be a different market and set the potential market for each demographic group to be one-quarter of the number of households in that group, by year. We include the list price (net of rebates and fees, in €10,000), the horsepower (in 100kW), the weight (in 1,000kg), the fuel cost (in euros per 100km), and fixed effects for the various engine types and body trims. Finally, we include the driving distance to the closest car dealer of each brand (in 10km) and fixed effects for the brand and year. The vector of characteristics X_j includes a constant, common to all $j \neq 0$, interacted with $\beta_d^{\text{const}} + \Pi_d^{\text{const}}\text{dem}_{dm} + \sigma_d \nu_i$, where dem_{dm} includes average income, average household size and an urban indicator, and ν_i is a scalar random coefficient that is normally distributed with standard deviation σ_d.[16] The inclusion of this flexible intercept plays an important role in allowing different consumers to have different substitution patterns toward the outside option, which is crucial in our counterfactuals to avoid overstating the potential market expansion induced by the introduction of an online distribution channel.

Our marginal cost specification includes horsepower, weight, fuel consumption (in liters per 100km), fixed effects for the engine type and body trim, and a time trend. We include two cost shifters. First, we interact the average yearly price of several key inputs (steel, iron, plastics, and aluminum) with the car's weight to compute a single input price index, similarly to D'Haultfœuille et al. (2019). We assume cars are made of 56% steel, 8% iron, 8% plastics, 10% aluminum, and 18% other materials not captured by the index. Second, we follow Grieco et al. (2023) and use the real exchange rate interacted with the car's country of origin as an additional cost shifter.[17] The real exchange rate is meant to capture differences in the cost of labor for each brand. Finally, we lag both cost shifters by one year to reflect planning horizons.

We deal with price endogeneity using both demand-side and supply-side moments. For the demand-side moments, the same set of instruments is used for all demographic groups (such that $Z_{jd} = Z_j, d = 1, ..., D$). In addition to X_j, these instruments include the sums of exogenous characteristics of competitors' products. The chosen characteristics are horsepower, weight, and fuel cost. We also include the number of products sold by competitors, the number of products sold by competitors that have the same engine type, and the number of products sold by competitors that have the same body trim. We specify the supply-side instruments Z_{jS} in a similar fashion, using horsepower, weight, fuel consumption, both cost shifters, number of competing products, number of competing products that have the same engine type, and number of competing products that have the same body trim.

[16] We experimented with the inclusion of additional random coefficients interacted with other elements of X_j, but did not obtain any significant estimates for the corresponding Σ_d. We therefore present our results for this more parsimonious specification that only includes a random coefficient on the constant and denote the associated group-specific standard deviation as σ_d instead of Σ_d.

[17] The real exchange rate is taken from Penn World Tables 10.0, `pl_con`. See Feenstra et al. (2015).

As discussed in Section 2.3, our estimator is robust to potential endogeneity of the observed distances without requiring additional instruments. More details on the empirical specification and each estimation step can be found in Appendix B.2. We present the estimates of the first-step GMM (the nonlinear parameters) in Table 3 and of the second-step GMM (the linear parameters) in Table 4.

The third panel of Table 3 summarizes the 702 micro moments used for the estimation of the nonlinear parameters $(\Pi_d, \sigma_d, \gamma_d)_{d=1,\dots,6}$ (9 "types" of micro moments $\times$ 6 demographic groups $\times$ 13 years). The first two panels of Table 3 show that the distance coefficients and those of the observed demographics are accurately estimated from these micro moments and have the expected signs. All demographic groups dislike traveling farther away to purchase cars, with the older and poorer consumers obtaining the largest disutity from traveling. This suggests that older and poorer consumers face additional constraints that make traveling more costly. Wealthier and larger households are more likely to purchase a car, while consumers living in urban municipalities are less likely to own one, perhaps because of the (relative) lack of parking and/or the availability of better public transport.

Table 4 shows significant heterogeneity in price sensitivities across demographic groups. Price sensitivities vary between -3.149 (young/poor) and -1.997 (old/rich), and the associated (median) own-price elasticities range from -5.93 to -4.14, see Table 5. Two intuitive patterns emerge with respect to these price sensitivities. First, within each age group, high-income consumers are less price sensitive than low-income consumers. Second, price sensitivities are ranked with respect to age (within income categories): older consumers have the lowest price sensitivities (groups 5 and 6), followed by middle-aged consumers (groups 3 and 4), and younger consumers are the most price sensitive (groups 1 and 2). By comparing both dimensions, we find that age is a more important determinant of price sensitivity than income.

We combine the estimated distance and price parameters to compute consumers' willingness to pay to reduce travel distance by one kilometer. Our estimated willingness to pay ranges from €18.1 to €27.7. These estimates highlight that young consumers have the lowest willingness to pay to reduce the distance from car dealers, while middle-aged/rich and old/poor have the highest. It is important to keep in mind that what we call "transportation costs" should be interpreted in a broad sense to encompass the burden of all visits to car dealers involved in the purchase of a car. In fact, these can factor in both visits prior to the purchase (e.g., the customer went for a test drive) and

Table 3: Estimates, nonlinear parameters (first-stage GMM)

	Demographic group					
	Young/Poor	Young/Rich	Middle/Poor	Middle/Rich	Old/Poor	Old/Rich
Nonlinear parameters γ_d and σ_d						
Distance	-0.059	-0.053	-0.072	-0.054	-0.061	-0.055
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Constant $\times$ ν_i	-0.275	0.086	-0.229	0.069	0.185	-0.169
	(0.163)	(0.144)	(0.136)	(0.160)	(0.155)	(0.167)
Nonlinear parameters Π_d						
Constant $\times$ Income	0.414	0.304	0.572	0.272	0.584	0.345
	(0.105)	(0.045)	(0.135)	(0.029)	(0.282)	(0.039)
Constant $\times$ Household size	0.423	0.360	0.344	0.650	0.281	0.431
	(0.016)	(0.021)	(0.016)	(0.020)	(0.018)	(0.018)
Constant $\times$ Urban	-0.434	-0.376	-0.314	-0.286	-0.407	-0.395
	(0.020)	(0.014)	(0.023)	(0.014)	(0.020)	(0.014)
Horsepower $\times$ Income	0.075	0.368	0.009	0.268	-0.619	0.267
	(0.462)	(0.229)	(0.657)	(0.154)	(1.298)	(0.193)
Horsepower $\times$ Urban	0.586	0.557	0.458	0.323	0.390	0.414
	(0.094)	(0.064)	(0.113)	(0.069)	(0.102)	(0.072)
Weight $\times$ Urban	-0.888	-0.774	-0.714	-0.615	-0.801	-0.815
	(0.060)	(0.057)	(0.058)	(0.056)	(0.077)	(0.061)
Micro moments						
$\mathbb{E}(\text{Distance} \mid j \neq 0)$	1.575	1.466	1.828	1.465	2.153	1.357
	[1.585]	[1.464]	[1.832]	[1.466]	[2.153]	[1.353]
$\mathbb{E}(\text{Income} \mid j \neq 0)$	1.709	2.441	1.816	2.516	1.890	2.457
	[1.705]	[2.440]	[1.813]	[2.525]	[1.894]	[2.465]
$\mathbb{E}(\text{Age} \mid j \neq 0)$	2.699	2.756	4.954	4.945	7.048	7.011
	[2.702]	[2.758]	[4.950]	[4.937]	[7.041]	[7.000]
$\mathbb{E}(\text{Female} \mid j \neq 0)$	0.493	0.496	0.505	0.507	0.544	0.549
	[0.498]	[0.499]	[0.509]	[0.508]	[0.545]	[0.545]
$\mathbb{E}(\text{Household size} \mid j \neq 0)$	2.251	2.462	2.253	2.406	2.259	2.310
	[2.249]	[2.455]	[2.253]	[2.407]	[2.256]	[2.307]
$\mathbb{E}(\text{Urban} \mid j \neq 0)$	0.365	0.267	0.293	0.254	0.148	0.305
	[0.366]	[0.268]	[0.297]	[0.253]	[0.146]	[0.303]
$\mathbb{C}(\text{Horsepower}, \text{Income} \mid j \neq 0)$	1.259	1.850	1.384	1.974	1.420	1.890
	[1.263]	[1.854]	[1.386]	[1.971]	[1.419]	[1.886]
$\mathbb{C}(\text{Horsepower}, \text{Urban} \mid j \neq 0)$	0.272	0.205	0.230	0.189	0.106	0.233
	[0.270]	[0.203]	[0.227]	[0.196]	[0.108]	[0.229]
$\mathbb{C}(\text{Weight}, \text{Urban} \mid j \neq 0)$	0.619	0.468	0.518	0.443	0.252	0.514
	[0.622]	[0.461]	[0.514]	[0.440]	[0.248]	[0.517]
Number of micro moments	702					
Value of first-step GMM objective function	0.272					

Notes: Distance is in 10km, income in €10,000, horsepower in 100kW, weight in 1,000kg. The micro moments used in estimation are at the demographic group $\times$ year level. We report the micro moments implied by the model averaged over markets, with their observed counterparts in square brackets.

Table 4: Estimates, linear parameters (second-stage GMM)

	Demographic group						Cost function
	Young/Poor	Young/Rich	Middle/Poor	Middle/Rich	Old/Poor	Old/Rich	$\ln(c_j)$
Linear parameters α_d							
Price	-3.149	-2.916	-2.614	-2.332	-2.432	-1.997	
	(0.417)	(0.408)	(0.400)	(0.390)	(0.400)	(0.382)	
Linear parameters β_d, λ_1, *and* λ_2							
Constant	-14.625	-13.322	-13.441	-12.353	-13.720	-10.802	-1.270
	(0.467)	(0.447)	(0.433)	(0.408)	(0.451)	(0.417)	(0.408)
Horsepower	5.681	4.502	4.672	3.492	5.314	2.576	0.577
	(0.292)	(0.268)	(0.252)	(0.226)	(0.250)	(0.221)	(0.400)
Weight	3.387	3.197	2.774	2.440	2.472	2.053	0.632
	(0.299)	(0.279)	(0.264)	(0.242)	(0.277)	(0.251)	(0.390)
Fuel cost	-0.209	-0.190	-0.143	-0.137	-0.147	-0.149	
	(0.039)	(0.037)	(0.033)	(0.031)	(0.033)	(0.030)	
Fuel consumption							0.019
							(0.400)
Diesel	0.548	0.295	0.641	0.367	0.226	-0.252	0.158
	(0.181)	(0.169)	(0.157)	(0.151)	(0.158)	(0.149)	(0.036)
Electric	1.130	1.222	0.992	1.097	0.070	-0.158	0.405
	(0.451)	(0.414)	(0.384)	(0.344)	(0.368)	(0.324)	(0.039)
Plug-in	-0.075	-0.115	0.196	0.169	-0.081	-0.261	0.189
	(0.423)	(0.403)	(0.381)	(0.354)	(0.373)	(0.327)	(0.011)
Hybrid	0.632	0.519	0.719	0.531	0.561	0.378	0.168
	(0.234)	(0.221)	(0.216)	(0.206)	(0.216)	(0.204)	(0.098)
Convertible	0.084	0.157	-0.016	0.085	-0.035	0.039	0.038
	(0.126)	(0.120)	(0.119)	(0.113)	(0.128)	(0.119)	(0.031)
Wagon	-0.187	-0.207	-0.145	-0.059	-0.779	-0.710	0.260
	(0.224)	(0.223)	(0.195)	(0.183)	(0.200)	(0.174)	(0.023)
Input price index							-0.160
							(0.382)
Real exchange rate							0.123
							(0.075)
Trend							0.018
							(0.069)
Willingness-to-pay (γ_d/α_d)	18.767	18.149	27.737	23.250	24.975	27.533	
	(2.678)	(2.866)	(4.391)	(4.230)	(4.295)	(5.604)	
Observations	4,975						
Value of second-step GMM objective function	3206.0						

Notes: The demand-side specification includes (non-group specific) brand and year fixed effects. Price is in €10,000, Horsepower is in 100kW, Weight is in 1,000kg, Fuel cost is in €/100km, and Fuel consumption is in L/100km. Distance is the driving distance to the nearest retailer of each brand, in 10km. Willingness to pay, in €/km, is computed as γ_d/α_d $(\times 1,000)$ for each demographic group. Standard errors are computed using the second-step correction formulae in Newey and McFadden (1994).

Table 5: Estimated own-price elasticities

Description	Mean	Std. dev.	10th pct.	Median	90th pct.	Observations
Group 1: Young/Poor	-6.61	3.05	-10.13	-5.93	-3.40	4,975
Group 2: Young/Rich	-6.19	2.82	-9.45	-5.57	-3.23	4,975
Group 3: Middle/Poor	-5.66	2.53	-8.58	-5.10	-3.01	4,975
Group 4: Middle/Rich	-5.16	2.25	-7.77	-4.66	-2.80	4,975
Group 5: Old/Poor	-5.34	2.35	-8.06	-4.82	-2.87	4,975
Group 6: Old/Rich	-4.57	1.93	-6.80	-4.14	-2.54	4,975

Notes: To maintain comparability, statistics are computed using a set of uniform weights $w_j = \sum_d \phi_d s_{jd} / \sum_j \sum_d \phi_d s_{jd}$ that are common across demographic groups.

Table 6: Estimated transaction prices

Description	Mean	Std. dev.	10th pct.	Median	90th pct.	Observations
Transaction price (€)						
Group 1: Young/Poor	20,980	9,669	10,816	18,865	32,157	4,975
Group 2: Young/Rich	21,239	9,669	11,082	19,132	32,422	4,975
Group 3: Middle/Poor	21,656	9,664	11,531	19,553	32,856	4,975
Group 4: Middle/Rich	22,138	9,661	12,028	20,036	33,326	4,975
Group 5: Old/Poor	21,955	9,660	11,815	19,863	33,142	4,975
Group 6: Old/Rich	22,883	9,656	12,768	20,797	34,087	4,975
Discount (€)						
Group 1: Young/Poor	1,903	52	1,840	1,918	1,964	4,975
Group 2: Young/Rich	1,645	49	1,585	1,660	1,705	4,975
Group 3: Middle/Poor	1,228	35	1,185	1,231	1,275	4,975
Group 4: Middle/Rich	746	26	716	742	782	4,975
Group 5: Old/Poor	928	20	901	930	954	4,975
Group 6: Old/Rich	0	0	0	0	0	4,975
Discount (%)						
Group 1: Young/Poor	9.66	3.67	5.52	9.22	14.92	4,975
Group 2: Young/Rich	8.35	3.17	4.76	7.98	12.88	4,975
Group 3: Middle/Poor	6.23	2.35	3.57	5.97	9.60	4,975
Group 4: Middle/Rich	3.78	1.43	2.16	3.62	5.81	4,975
Group 5: Old/Poor	4.71	1.78	2.70	4.49	7.26	4,975
Group 6: Old/Rich	0	0	0	0	0	4,975

Notes: All monetary values are converted to 2018 euros. The demographic group that is estimated to pay the list price is group 6 for all products. Statistics are computed using a set of uniform weights $w_j = \sum_d \phi_d s_{jd} / \sum_k \sum_d \phi_d s_{kd}$ that are common across demographic groups.

expected future visits (e.g., maintenance, after-sale services, etc.).[18]

One of the key features of our approach is the estimation of the unobserved transaction prices paid by different demographic groups as the (potential) result of third-degree price discrimination. We summarize these estimated transaction prices in Table 6. To remove the effect of group-specific sales from the reported statistics, we calculate all statistics using a single set of weights based on the total sales of each car model (i.e., aggregated over groups). In line with the results of D'Haultfœuille et al. (2019), we find evidence in support of third-degree price discrimination. Group 6 (old/rich) is the demographic group estimated to always pay the observed list price for all car models. This follows from the fact that old/rich consumers are the most price inelastic among all groups.

[18] We note that our estimates are significantly lower than those by Nurski and Verboven (2016), which find a willingness to pay of €112 per kilometer for the Belgian car market in 2011-2012. Although it is difficult to pin down the exact reason for these different estimates, there are several contributing factors, from the different structural models and estimation methods (e.g., different ways of dealing with endogenous distance) to the different data used for estimation (e.g., the average distance between consumers and car dealers in Belgium is substantially smaller than in our data, 11.7km vs. 16.3km).

However, for the other demographic groups, discounts can be significant: the average discount ranges from 3.8% to 9.7%, corresponding to €746 and €1,903, respectively. Consistent with intuition, given the highest price elasticity, the consumers of group 1 (young/poor) are those estimated to receive the largest discounts.

3.6 Price discrimination versus transportation costs

Before moving on to the introduction of an online distribution channel, we perform a few counterfactual experiments to shed some light on the relationship between price discrimination and transportation costs. The results of these counterfactuals are presented in Table 7 and Table 8.

We focus on three broad sets of counterfactuals. First, we consider a case in which firms cannot price discriminate among consumers. Second, we consider a case in which price discrimination is possible, but consumers face reduced transportation costs. We consider various levels of transportation cost reductions which allow consumers to still value car dealer proximity. Finally, we consider a case in which price discrimination is not possible and consumers face reduced transportation costs. This counterfactual coincides with our description of a world in which all sales occur online.

First, we discuss price discrimination. We focus on the first two columns of Table 7 and the first two rows of Table 8 and compare a counterfactual without price discrimination to the baseline. Most consumers (mostly young and low-income groups) benefit from a discount over the non-discriminatory prices. This increases their total purchases and their gain in consumer surplus ranges between €8.5 and €16.2 per consumer per year. Differently, consumers who are middle-aged or old and rich (groups 4 and 6) pay higher prices under price discrimination and reduce their purchases of all car models. Their respective losses in consumer surplus are €3.4 and €55 per consumer per year.

Overall, because of the large share of total purchases from these older and wealthier groups, price discrimination decreases consumer surplus by €3.5 per consumer per year. Whether or not price discrimination is profitable remains an empirical question in oligopolistic settings. We find that industry profits increase by around 1.3%, suggesting that—holding everything else constant—price discrimination is not tremendously profitable for car manufacturers in the French market. Since the gains from price discrimination are relatively small for firms and the losses are relatively small on average for consumers, we conclude that price discrimination is mostly redistributive (shifting surplus from old/rich to others) in the French market.

Second, we discuss transportation costs and prices. We compare the set of counterfactuals with price discrimination and reduced transportation costs to the baseline. Reducing transportation costs (with or without price discrimination) does not seem to particularly affect pricing decisions (see panel 1 of Table 7). When consumers face reduced transportation costs, firms capture the associated gains through market expansion rather than increased prices. This suggests that the pass-through of transportation costs to prices is small.

Third, we discuss consumers' responses to a reduction in transportation costs. We again contrast the set of counterfactuals with price discrimination and reduced transportation costs with the baseline. As transportation costs are reduced, consumers gradually switch to more expensive car models (see panel 2 of Table 7) and to car dealers located on average farther away (see panel 4 of Table 7). In some sense, consumers save on transportation costs and "reinvest" part of these savings by spending more on better car models and purchasing from car dealers located farther away. Our model predicts that eliminating transportation costs entirely can lead to an increase in consumer surplus and profits in the range of 10%.

Table 7: Effect of price discrimination and transportation costs on consumers' purchases

	Baseline	No discr.	Reduced transportation costs				No discr. + Reduced transportation costs			
			-25%	-50%	-75%	-100%	-25%	-50%	-75%	-100%
Transaction prices, uniform weights										
Group 1: Young/Poor	21,685		21,685	21,685	21,685	21,686				
Group 2: Young/Rich	21,941		21,941	21,942	21,942	21,942				
Group 3: Middle/Poor	22,355		22,356	22,356	22,357	22,358				
Group 4: Middle/Rich	22,834		22,834	22,835	22,836	22,837				
Group 5: Old/Poor	22,658		22,660	22,661	22,662	22,664				
Group 6: Old/Rich	23,582		23,583	23,584	23,585	23,587				
Uniform		22,793					22,792	22,791	22,791	22,790
Transaction prices, sales-weighted										
Group 1: Young/Poor	20,536		20,582	20,632	20,684	20,741				
Group 2: Young/Rich	22,216		22,250	22,287	22,325	22,365				
Group 3: Middle/Poor	22,202		22,279	22,362	22,453	22,552				
Group 4: Middle/Rich	24,075		24,124	24,177	24,233	24,294				
Group 5: Old/Poor	21,336		21,399	21,466	21,537	21,612				
Group 6: Old/Rich	23,530		23,579	23,631	23,687	23,748				
Uniform		22,546					22,596	22,651	22,709	22,771
Sales, in units										
Group 1: Young/Poor	105,752	-28,839	+2,494	+5,113	+7,865	+10,763	-26,981	-25,028	-22,974	-20,809
Group 2: Young/Rich	68,371	-13,618	+1,326	+2,701	+4,126	+5,607	-12,530	-11,402	-10,230	-9,012
Group 3: Middle/Poor	137,531	-13,545	+4,246	+8,768	+13,593	+18,751	-9,639	-5,475	-1,027	+3,736
Group 4: Middle/Rich	228,904	+2,940	+4,394	+8,956	+13,696	+18,630	+7,474	+12,183	+17,080	+22,180
Group 5: Old/Poor	114,085	-4,338	+3,559	+7,306	+11,254	+15,417	-849	+2,829	+6,709	+10,805
Group 6: Old/Rich	300,651	+44,435	+5,394	+11,000	+16,832	+22,908	+50,681	+57,172	+63,926	+70,962
All consumers	955,294	-12,965	+21,413	+43,844	+67,366	+92,076	+8,156	+30,279	+53,484	+77,862
Average distance to car models, in km										
Group 1: Young/Poor	15.97	+0.08	+0.41	+0.85	+1.31	+1.80	+0.50	+0.94	+1.40	+1.90
Group 2: Young/Rich	14.83	+0.06	+0.25	+0.52	+0.80	+1.10	+0.31	+0.58	+0.86	+1.16
Group 3: Middle/Poor	17.36	+0.04	+0.60	+1.24	+1.92	+2.65	+0.64	+1.29	+1.97	+2.71
Group 4: Middle/Rich	14.91	+0.02	+0.29	+0.60	+0.94	+1.30	+0.32	+0.63	+0.97	+1.33
Group 5: Old/Poor	21.38	+0.02	+0.50	+1.01	+1.55	+2.11	+0.52	+1.03	+1.57	+2.14
Group 6: Old/Rich	14.00	0.00	+0.31	+0.65	+1.00	+1.39	+0.31	+0.64	+1.00	+1.38

Notes: All counterfactual experiments are computed using the 2019 data only. Unless indicated otherwise, in-person sales imply price discrimination and transportation costs, while online sales imply a uniform price and reduced transportation costs. Transaction prices are in 2018 euros. "Uniform weights" are constructed using the total sales of each product in the baseline scenario, hence are fixed across demographic groups and counterfactual experiments. "Sales weights" use realized sales for each demographic group and counterfactual experiment. For sales and average distances, we report the values at baseline in the first column, and the change from baseline in the other columns.

Table 8: Effect of price discrimination and transportation costs on welfare

Counterfactual	Δ Consumer surplus, € per capita per year							Δ Profits (MM€)
	Young/Poor	Young/Rich	Middle/Poor	Middle/Rich	Old/Poor	Old/Rich	All	Total
Baseline	59.1	81.4	153.5	254.6	217.0	356.4	190.8	4,100.7
No discrimination	-16.2	-16.4	-15.4	+3.4	-8.5	+55.0	+3.5	-53.5
Reduced transportation costs								
• -25% transportation costs	+1.4	+1.6	+4.9	+5.0	+7.0	+6.6	+4.3	+91.0
• -50% transportation costs	+2.9	+3.3	+10.0	+10.3	+14.3	+13.6	+8.8	+186.3
• -75% transportation costs	+4.5	+5.0	+15.6	+15.7	+22.1	+20.8	+13.5	+286.2
• -100% transportation costs	+6.1	+6.8	+21.5	+21.4	+30.3	+28.3	+18.4	+391.2
No discr. + Reduced transp. costs								
• -25% transportation costs	-15.2	-15.1	-11.0	+8.6	-1.7	+62.8	+7.9	+36.2
• -50% transportation costs	-14.1	-13.7	-6.2	+14.0	+5.5	+70.9	+12.5	+130.2
• -75% transportation costs	-12.9	-12.3	-1.2	+19.6	+13.1	+79.4	+17.3	+228.7
• -100% transportation costs	-11.7	-10.8	+4.3	+25.5	+21.2	+88.2	+22.3	+332.1

Notes: All counterfactual experiments are computed using the 2019 data only. We report the values at baseline in the first row and the change from baseline in the other rows. Consumer surplus is in 2018 euros. Profits are in million 2018 euros.

Finally, we comment on the trade-off faced by firms between price discrimination and reductions in transportation costs. We notice that, under uniform pricing, even modest reductions in transportation costs (e.g., a 25% reduction) would result in an increase in industry profit comparable to the increase implied by price discrimination (36.2 vs. 53.5 million euros per year). Under uniform pricing, eliminating transportation costs entirely would lead to an increase in industry profit around 6.2 times larger compared to the increase implied by price discrimination (332.1 vs. 53.5 million euros). These estimates suggest that, from the perspective of car manufacturers, price discrimination has a second-order effect compared to transportation costs in the French market. This provides some suggestive evidence in support of the "stated" intentions of car manufacturers to move their businesses online, even if this meant losing some ability (or even all, as Tesla) to price discriminate.

4 Model with the online distribution channel

4.1 Introducing online sales

We extend our model to investigate a set of counterfactuals in which cars can be purchased either in person at the closest car dealer of the chosen brand (as observed in the data) or online directly from a firm's website (a distribution channel currently not observed in the data). By completing the transaction online and having the car delivered to their doorsteps, consumers would face a lower transportation cost since fewer or no visits to the car dealer would be required, and would pay the non-discriminatory uniform price posted on the website. Concluding the transaction in person instead, consumers would physically travel, potentially multiple times, to car dealers which could then offer them a personalized price possibly different from the uniform online price.

Throughout, we remain agnostic about the extent to which buying online reduces consumers' transportation costs. In extreme cases where consumers do not visit car dealers for test drives and do not value after-sale services, purchasing cars online could eliminate transportation costs entirely. In more realistic cases, some transportation costs may remain if consumers still expect to visit car dealers for, e.g., maintenance in the future. We investigate these extremes and other intermediate scenarios by repeating our analysis for different levels of reduction in transportation costs, captured by the parameter $\tau \in [0, 1]$.

We focus our attention on the case in which car dealers do not price discriminate against consumers who purchase online. We do this to mimic observed industry practice, which leans heavily toward price transparency and streamlining the transaction process (see the Introduction). For completeness, we, however, also perform a set of counterfactuals with price discrimination in both distribution channels. The results of these additional counterfactual simulations are reported in Appendix C.

Lastly, we assume that the introduction of an online distribution channel does not change the observed configuration of car dealers (i.e., no entry or exit of car dealers) or the vertical relations between car manufacturers and car dealers. These assumptions imply that our results should be interpreted as short-run responses of the industry to the introduction of an online distribution channel. Although we do not explicitly relax these assumptions in our structural model, we, however, conduct a series of additional counterfactuals in Section 5.5 to provide some insight into these potentially important long-run responses of the industry.

4.2 Demand

We extend our model from Section 2.1 in the simplest possible way that allows us to capture the key features of the online channel. For clarity, we define p^P_{jd} (previously p_{jd}) as the discriminatory in-person price paid for product j by consumers of group d, and p^O_j as the uniform online price paid for product j by all consumers. We maintain the assumption that all car models are available in both distribution channels.

We allow for the possibility that consumers belonging to different demographic groups have different propensities toward online shopping. For example, younger and wealthier consumers may be at ease with the option of purchasing a car from a firm's website, while older and less affluent consumers may not be willing or able to do so. We rely on a national survey on attitudes towards online shopping to estimate the probability ψ_d that consumers of demographic group d consider the online sales channel. We report on these probabilities both in Table 10 and in Appendix Table A.1.

We assume that a share $1-\psi_d$ of consumers of group d do not have access to the online distribution channel and have indirect utility (1). The remaining share ψ_d of consumers of group d who can access both channels, the indirect utility of purchasing j is

$$\begin{aligned} U_{ijdm} = & \underbrace{X_j'\beta_d + \alpha_d p_{jd}^P + \xi_{jd}}_{\delta_{jd}} + \mu_{jdm}(\nu_i) \\ & + \max\left\{\underbrace{\gamma_d \text{dist}_{jm}}_{\eta_{jdm}^P}, \underbrace{\alpha_d(p_j^O - p_{jd}^P) + \tau\gamma_d \text{dist}_{jm}}_{\eta_{jdm}^O}\right\} + \epsilon_{ijdm}, \end{aligned} \tag{14}$$

where $\tau \in [0,1)$ is a parameter (to be calibrated) that controls the reduction in transportation costs brought by the online channel relative to the in-person channel. Note that, as in indirect utility (1), η_{jdm}^P represents the transportation cost of purchasing car model j in person from the closest car dealer, while η_{jdm}^O represents the trade-off faced by online shopping (i.e., reduced transportation costs versus a potential discount).

Given this specification, the probability with which consumers in demographic group d and municipality m purchase car model j is

$$\begin{aligned} s_{jdm}(p_d^P, p^O) = \psi_d \cdot & \int \frac{\exp\left(\delta_{jd} + \mu_{jdm}(\nu_i) + \max\{\eta_{jdm}^P, \eta_{jdm}^O\}\right)}{1 + \sum_{k=1}^J \exp\left(\delta_{kd} + \mu_{kdm}(\nu_i) + \max\{\eta_{kdm}^P, \eta_{kdm}^O\}\right)} dF(\nu_i) \\ & + (1-\psi_d) \cdot \int \frac{\exp\left(\delta_{jd} + \mu_{jdm}(\nu_i) + \eta_{jdm}^P\right)}{1 + \sum_{k=1}^J \exp\left(\delta_{kd} + \mu_{kdm}(\nu_i) + \eta_{kdm}^P\right)} dF(\nu_i), \end{aligned} \tag{15}$$

where $p_d^P = (p_{1d}^P, ..., p_{Jd}^P)$ and $p^O = (p_1^O, ..., p_J^O)$. Note that we can equivalently express (15) as $s_{jdm} = s_{jdm}^P + s_{jdm}^O$, distinguishing between the share of in-person purchases (denoted by superscript P) and the share of online purchases (superscript O):

$$\begin{aligned} s_{jdm}^P(p_d^P, p^O) = \ & \psi_d \cdot \int \frac{\exp(\delta_{jd} + \mu_{jdm}(\nu_i) + \eta_{jdm}^P)\mathbb{1}\left\{\eta_{jdm}^P \geq \eta_{jdm}^O\right\}}{1 + \sum_{k=1}^J \exp(\delta_{kd} + \mu_{kdm}(\nu_i) + \max\{\eta_{kdm}^P, \eta_{kdm}^O\})} dF(\nu_i) \\ & + (1-\psi_d) \cdot \int \frac{\exp\left(\delta_{jd} + \mu_{jdm}(\nu_i) + \eta_{jdm}^P\right)}{1 + \sum_{k=1}^J \exp\left(\delta_{kd} + \mu_{kdm}(\nu_i) + \eta_{kdm}^P\right)} dF(\nu_i), \end{aligned} \tag{16}$$

$$s_{jdm}^O(p_d^P, p^O) = \ \psi_d \cdot \int \frac{\exp(\delta_{jd} + \mu_{jdm}(\nu_i) + \eta_{jdm}^O)\mathbb{1}\left\{\eta_{jdm}^P < \eta_{jdm}^O\right\}}{1 + \sum_{k=1}^J \exp(\delta_{kd} + \mu_{kdm}(\nu_i) + \max\{\eta_{kdm}^P, \eta_{kdm}^O\})} dF(\nu_i). \tag{17}$$

Averaging (16) and (17) over municipalities, we obtain the national-level market shares of group d for car model j from sales channel $\ell \in \{P, O\}$

$$s^{\ell}_{jd}(p^P_d, p^O) = \sum_{m \in \mathcal{M}} w_{dm} \cdot s^{\ell}_{jdm}(p^P_d, p^O). \tag{18}$$

4.3 Supply

Similar to the model described in Section 2.2, we consider a Bertrand-Nash price-setting game in which every firm f chooses a menu of transaction prices for in-person sales $p^P_j = (p^P_{j1}, ..., p^P_{jD})$ and the non-discriminatory online price p^O_j for each j they sell by maximizing the national-level profit function

$$\pi_f(p^P, p^O) = \sum_{d=1}^{D} \phi_d \sum_{j \in \mathcal{J}_f} s^P_{jd}(p^P_d, p^O) \cdot (p^P_{jd} - c_j) + \sum_{d=1}^{D} \phi_d \sum_{j \in \mathcal{J}_f} s^O_{jd}(p^P_d, p^O) \cdot (p^O_j - c_j), \tag{19}$$

where $p^D = (p^P_1, ..., p^P_D)$ and the national-level group-specific market shares $s^P_{jd}(p^P_d, p^O)$ and $s^O_{jd}(p^P_d, p^O)$ correspond to (18) for $\ell \in \{P, O\}$.

4.4 Solving the model

The model with online sales is difficult to solve in practice. This is due to the maximum operator in the indirect utility (14), which leads to discontinuities in the resulting purchase probabilities (16)-(17). This causes traditional numerical routines for the maximization of (19) to fail, as small price changes can cause discontinuous changes to the system of first-order conditions (see also Duch-Brown et al., 2023).

To avoid this problem, we implement a methodology similar to that proposed by Duch-Brown et al. (2023). The idea is to approximate the mixed logit model implied by indirect utility (14) by a mixed nested logit in which each j belongs to a nest and where each of these J nests includes the two distribution channels: in person (j, P) and online (j, O). In other words, consumers first choose which of the J car models (or the outside option) to purchase (i.e., they choose the "nest"), and then choose whether to purchase that car model in person or online. In this case, the indirect utility of purchasing car model j from $\ell \in \{P, O\}$ does not involve any maximum operator and we have

$$U^{\ell}_{ijdm} = \delta_{jd} + \mu_{jdm}(\nu_i) + \eta^{\ell}_{jdm} + \zeta_{ijdm} + (1 - \sigma)\epsilon^{\ell}_{ijdm}, \tag{20}$$

where both ϵ^{ℓ}_{ijdm} and $\zeta_{ijdm} + (1 - \sigma)\epsilon^{\ell}_{ijdm}$ are distributed extreme value type I, ζ_{ijdm} is common to both distribution channels $\ell \in \{P, O\}$ of car model j, and parameter

$\sigma \in [0, 1)$ (Cardell, 1997). Importantly, when σ tends to 1, the two sales channels become perfect substitutes, and the mixed nested logit market share implied by indirect utility (20) converges to that of the mixed logit implied by indirect utility (14).

We solve for the optimal price vector $p^* = (p_1^{P*}, ..., p_D^{P*}, p^{O*})$ by adapting the ζ-markup algorithm suggested by Morrow and Skerlos (2011). In practice, we cannot maximize profit function (19) at $\sigma = 1$, as the mixed nested logit market shares are not defined. We instead compute market shares for values of σ numerically close to 1 and rely on extrapolation to approach the limit as $\sigma \to 1$. For more details, see Appendix B.3.

5 Counterfactual simulations

In this section, we present our main counterfactual results. We simulate scenarios in which an online distribution channel is introduced in the French car industry. All counterfactuals are performed on our 2019 data. Unless otherwise mentioned, consumers purchasing in person at car dealers can receive discounts as a result of price discrimination and incur full transportation costs. Consumers who purchase online instead pay the uniform price posted on the car manufacturer's website and reduce their transportation costs by a factor of $1 - \tau$.

We consider four different levels of transportation cost reductions, that is $\tau \in \{0.75, 0.50, 0.25, 0\}$. In addition, we consider two cases of consumers' attitudes toward online shopping. First, a case where everyone can use both distribution channels without restrictions. We use this to benchmark the forces at play. Then, we consider a more realistic scenario in which some consumers never purchase online. The propensity to shop online, indicated by the parameter $\psi_d \in [0, 1]$, is estimated using a survey of consumers' attitudes toward online shopping, as described in Section 3.1.

5.1 Unrestricted access to the online distribution channel

The results of our counterfactuals with unrestricted access to the online channel are presented in Table 9. These results highlight three main patterns.

The first pattern concerns sales. The online channel introduces a market expansion that varies from around 2% to 9% (depending on the reduction in transportation costs), mostly driven by the largest buyers, especially the consumers in group 6 (old/rich). The online channel provides these consumers with an opportunity to reduce both the price they pay (the uniform online price is lower than the discriminatory price they paid in person at baseline) and to reduce their transportation costs. Therefore, the implied

Table 9: Effect of online channel with unrestricted access

	Baseline	Transp. costs red. from online channel			
		-25%	-50%	-75%	-100%
Transaction prices, uniform weights					
Group 1: Young/Poor	21,685	21,628	21,558	21,600	22,145
Group 2: Young/Rich	21,941	21,765	21,524	22,055	21,986
Group 3: Middle/Poor	22,355	22,780	23,085	22,975	22,911
Group 4: Middle/Rich	22,834	23,086	23,084	22,973	22,909
Group 5: Old/Poor	22,658	23,087	23,084	22,973	22,909
Group 6: Old/Rich	23,582	23,084	23,082	22,971	22,906
Online		23,090	23,091	22,978	22,912
Transaction prices, sales-weighted					
Group 1: Young/Poor	20,536	20,560	19,117	16,629	14,504
Group 2: Young/Rich	22,216	22,029	19,560	18,851	14,567
Group 3: Middle/Poor	22,202	19,117	23,816	21,783	16,222
Group 4: Middle/Rich	24,075	23,641	24,243	22,203	17,395
Group 5: Old/Poor	21,336	20,844	22,054	21,420	16,147
Group 6: Old/Rich	23,530	21,400	22,272	21,638	16,419
Online		28,404	25,750	23,969	23,595
Sales, in units					
Group 1: Young/Poor	105,752	+4,382	+13,457	+13,069	-3,633
Group 2: Young/Rich	68,371	+6,140	+15,096	+1,798	+6,186
Group 3: Middle/Poor	137,531	-10,151	-15,266	-6,855	-136
Group 4: Middle/Rich	228,904	-7,472	-3,101	+8,389	+16,498
Group 5: Old/Poor	114,085	-7,502	-3,910	+1,346	+7,008
Group 6: Old/Rich	300,651	+33,798	+40,786	+51,487	+63,075
All consumers	955,294	+19,195	+47,062	+69,234	+88,998
Prop. of online sales					
Group 1: Young/Poor	0	0.004	0.107	0.349	0.616
Group 2: Young/Rich	0	0.017	0.147	0.509	0.618
Group 3: Middle/Poor	0	0.857	0.985	0.996	0.998
Group 4: Middle/Rich	0	0.975	0.990	0.997	0.998
Group 5: Old/Poor	0	0.983	0.994	0.998	0.998
Group 6: Old/Rich	0	0.970	0.992	0.996	0.997
Average distance to car models, in km					
Group 1: Young/Poor	15.97	+0.45	+0.26	+0.46	+0.55
Group 2: Young/Rich	14.83	+0.21	+0.04	+0.07	-0.31
Group 3: Middle/Poor	17.36	+0.12	+1.36	+2.14	+2.96
Group 4: Middle/Rich	14.91	+0.30	+0.70	+1.08	+1.50
Group 5: Old/Poor	21.38	+0.53	+1.02	+1.68	+2.39
Group 6: Old/Rich	14.00	+0.33	+0.67	+1.06	+1.52

Notes: All counterfactual experiments are computed using the 2019 data only. Unless indicated otherwise, in-person sales imply price discrimination and transportation costs and online sales imply a uniform price and reduced transportation costs. Transaction prices are in 2018 euros. “Uniform weights” are constructed using the total sales of each product in the baseline scenario, hence are fixed across demographic groups and counterfactual experiments. “Sales weights” use realized sales for each demographic group and counterfactual experiment. For sales and average distances, we report the values at baseline in the first column, and the change from baseline in the other columns.

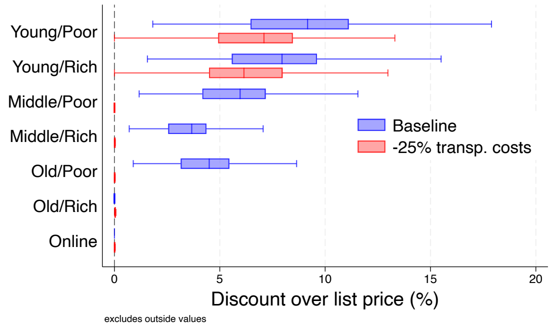
increase in indirect utility generates substitution away from the outside option. For other demographic groups, the price effect and the reduction in transportation costs go in opposite directions, so the impact on sales is smaller in magnitude (and can be negative for some groups). In particular, relative to the baseline, consumers in group 3 (middle-aged/poor) always face higher prices both in person and online and consequently purchase fewer vehicles for any level of reduction in transportation costs.

The second pattern concerns consumers' attitude towards the online channel. We note that once the online channel is available, most sales are diverted away from the traditional in-person channel. Importantly, this is true even for small decreases in transportation costs. As an example, a decrease in transportation costs by 25% leads most consumers in groups 3-6 to purchase cars online. If transportation costs are eliminated altogether, then also more than 60% of the purchases of groups 1 and 2 happen online (the only consumers who still purchase in person are those with a distance from car dealers very close to zero). In line with the preliminary evidence in Section 3.6, we find that consumers purchase car models from the online channel that would otherwise be sold by car dealers that are farther away from where they live.

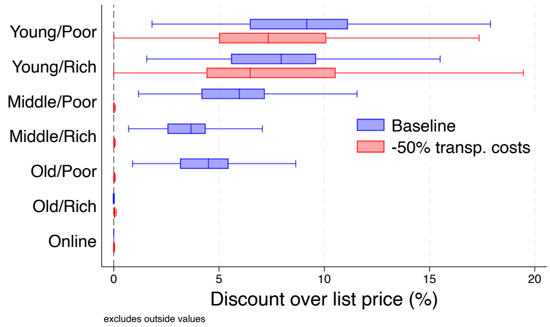
The last pattern concerns price dispersion in the traditional in-person channel. We find that once the online channel with a uniform price becomes available, firms have an incentive to reduce the amount of price discrimination in the in-person channel (see panel 1 of Table 9). This is a result of the competitive pressure the online channel exerts on the traditional in-person channel. Due to lower transportation costs, firms benefit by redirecting most consumers to the online channel, which increases overall sales. Since online prices are restricted to be uniform, firms then set a similar price in both channels, which makes the online channel unambiguously better for consumers.

To better visualize these changes in price dispersion, Figure 4 plots the distribution of discounts in the in-person channel both at baseline and when the online channel is available. For any level of reduction in transportation costs, firms continue to offer discounts to young consumers (the most price elastic consumers), as for them price plays a more important role than distance to car dealers. In contrast, even for small reductions in transportation costs, discounts are completely eliminated for the other demographic groups (which are less price elastic and more sensitive to distance).

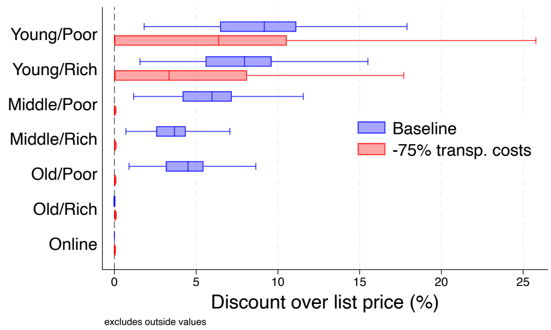
Figure 4: Price dispersion with unrestricted online access



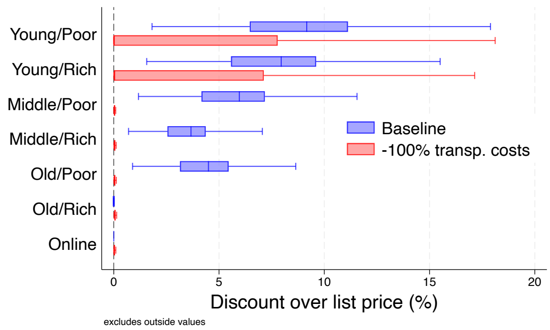
(a) -25% transportation costs



(b) -50% transportation costs



(c) -75% transportation costs



(d) -100% transportation cost

Notes: These figures illustrate price dispersion in the in-person channel when consumers have unrestricted access to the online channel, for varying transportation cost reductions as per Table 9. Price dispersion is represented as a discount over the list price, in percentage points.

5.2 Restricted access to the online distribution channel

We now turn to our preferred specification. Since we cannot identify a preference parameter for the online channel in our data (as there were essentially no online sales during our sample period), we enrich our counterfactual model by restricting access to the online channel for a subset of consumers based on additional survey data. Table 10 reports the propensity to shop online by demographic group, which we calibrate using a survey of consumers' attitudes toward online shopping. We construct these probabilities as the proportion of consumers in each demographic group who bought (anything) online in the year prior to the survey. A share ψ_d of consumers have access to both the in-person and online channels (the unrestricted consumers). Instead, it is assumed that the remaining share $(1-\psi_d)$ of consumers is captive to the traditional in-person channel (the captive or restricted consumers).

Table 10: Propensity to shop online, in 2019

Demographic group	Propensity (ψ_d)	Observations
Group 1: Young/Poor	0.770	9,545
Group 2: Young/Rich	0.894	12,131
Group 3: Middle/Poor	0.505	14,098
Group 4: Middle/Rich	0.770	19,614
Group 5: Old/Poor	0.146	20,034
Group 6: Old/Rich	0.464	17,724

We perform a similar set of counterfactuals as in the previous section, with varying levels of transportation cost reductions accruing from the online channel, and where consumers are restricted to shop in person according to the propensities in Table 10. The results of these counterfactuals are presented in Table 11. Restricting some consumers to the in-person channel limits the expansion in sales observed in the previous set of counterfactuals in Table 9, as the proportion of consumers who shop online decreases significantly. For example, among the consumers of group 5 (old/poor) who purchase a car, fewer than 20% do it online. Similarly, for group 6 fewer than 60% of the purchased cars are bought online. As a result, the average distance to the closest car dealers that sell the purchased cars does not increase as much as in the case of unrestricted consumers, since many are still buying in person.

Compared to the results in Table 9, when some consumers are captive to the in-person channel, we observe less "convergence" between in-person and online prices. We notice an interesting pattern: some consumers actually pay a premium over the online uniform price when purchasing in person (e.g., groups 4-6 in column (5) of Table 11). For some demographic groups, firms set in-person prices that are higher than the online price, so that the unrestricted consumers in these groups purchase online, while more surplus is extracted from the captive consumers.

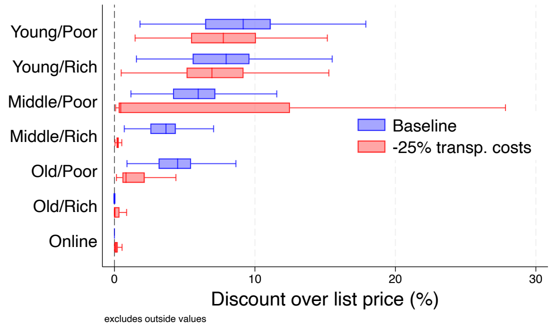
From the perspective of firms, there are two opposing forces at play. The first force is the competitive pressure from the online channel, which leads firms to set in-person prices very close to the online uniform price. With reduced transportation costs, a uniform price makes the online channel unambiguously better and directs unrestricted consumers to that channel, allowing firms to expand the market (as seen in Table 9). The second opposing force concerns captive consumers. For these consumers, firms would still like to extract more surplus through price discrimination, potentially setting different in-person prices than the online price. However, since unrestricted and captive consumers coexist in the market and face the same prices (i.e., price discrimination in that dimension is

Table 11: Effect of online channel with restricted access

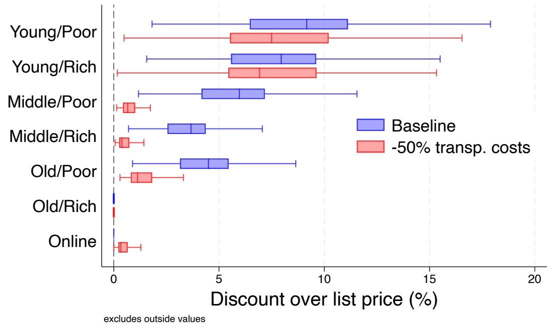
	Baseline	Transp. costs red. from online channel			
		-25%	-50%	-75%	-100%
Transaction prices, uniform weights					
Group 1: Young/Poor	21,685	21,671	21,702	21,931	22,261
Group 2: Young/Rich	21,941	21,774	21,613	21,984	21,976
Group 3: Middle/Poor	22,355	22,189	22,977	22,796	22,694
Group 4: Middle/Rich	22,834	23,239	23,020	22,835	22,734
Group 5: Old/Poor	22,658	22,946	22,907	22,791	22,728
Group 6: Old/Rich	23,582	23,259	23,322	23,619	23,808
Online		23,261	23,028	22,831	22,721
Transaction prices, sales-weighted					
Group 1: Young/Poor	20,536	20,400	19,275	18,548	18,250
Group 2: Young/Rich	22,216	22,046	19,762	19,106	16,077
Group 3: Middle/Poor	22,202	19,736	22,862	22,588	22,371
Group 4: Middle/Rich	24,075	22,697	24,308	24,186	23,893
Group 5: Old/Poor	21,336	21,446	21,556	21,459	21,384
Group 6: Old/Rich	23,530	22,469	22,771	23,236	23,572
Online		28,144	25,530	23,596	23,312
Sales, in units					
Group 1: Young/Poor	105,752	+1,343	+2,221	-4,345	-11,224
Group 2: Young/Rich	68,371	+5,884	+11,923	+2,488	+4,118
Group 3: Middle/Poor	137,531	+9,382	-15,305	-6,335	-526
Group 4: Middle/Rich	228,904	-14,336	-615	+13,846	+23,060
Group 5: Old/Poor	114,085	-6,755	-5,773	-2,165	+716
Group 6: Old/Rich	300,651	+19,742	+32,173	+32,240	+33,450
All consumers	955,294	+15,260	+24,624	+35,729	+49,594
Prop. of online sales					
Group 1: Young/Poor	0	0.011	0.111	0.387	0.570
Group 2: Young/Rich	0	0.019	0.165	0.504	0.629
Group 3: Middle/Poor	0	0.251	0.506	0.564	0.584
Group 4: Middle/Rich	0	0.687	0.804	0.818	0.823
Group 5: Old/Poor	0	0.067	0.137	0.168	0.191
Group 6: Old/Rich	0	0.383	0.518	0.549	0.577
Average distance to car models, in km					
Group 1: Young/Poor	15.97	+0.03	+0.07	+0.32	+0.85
Group 2: Young/Rich	14.83	+0.24	+0.12	+0.24	+0.05
Group 3: Middle/Poor	17.36	-0.71	+0.94	+1.43	+1.87
Group 4: Middle/Rich	14.91	+0.45	+0.67	+0.97	+1.26
Group 5: Old/Poor	21.38	+0.02	+0.27	+0.48	+0.65
Group 6: Old/Rich	14.00	+0.30	+0.28	+0.51	+0.80

Notes: All counterfactual experiments are computed using the 2019 data only. Unless indicated otherwise, in-person sales imply price discrimination and transportation costs and online sales imply a uniform price and reduced transportation costs. Access to the online channel in columns (2)-(5) is restricted according to a survey of online purchases, see Table A.1. Transaction prices are in 2018 euros. "Uniform weights" are constructed using the total sales of each product in the baseline scenario, hence are fixed across demographic groups and counterfactual experiments. "Sales weights" use realized sales for each demographic group and counterfactual experiment. For sales and average distances, we report the values at baseline in the first column, and the change from baseline in the other columns.

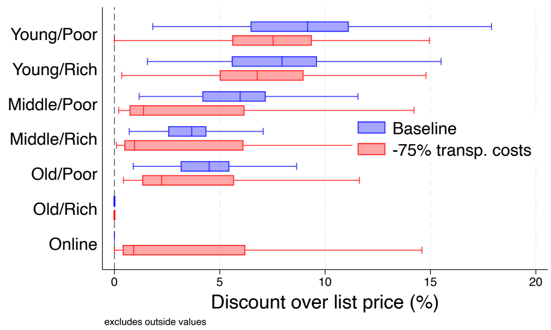
Figure 5: Price dispersion with restricted online access



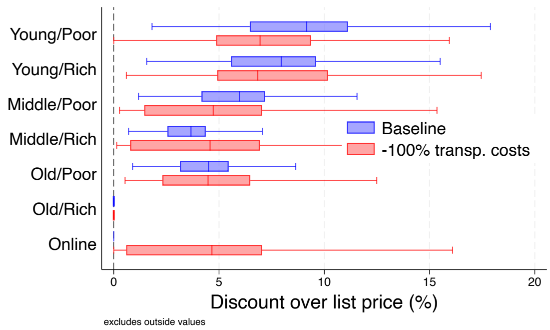
(a) -25% transportation costs



(b) -50% transportation costs



(c) -75% transportation costs



(d) No transportation cost

Notes: These figures illustrate price dispersion in the in-person channel when some consumers have restricted access to the online channel, for varying transportation cost reductions as per Table 11. Price dispersion is represented as a discount over the list price, in percentage points.

not possible), firms must strike a balance between these two opposing forces.

Figure 5 plots the distribution of discounts as a function of the reduction in transportation costs. Compared to Figure 4, we notice that the presence of captive consumers substantially changes firms' responses, especially when the online channel brings large reductions in transportation costs (panels (c) and (d)). With small reductions in transportation costs (panels (a) and (b)), we observe patterns similar to those in Figure 4: list prices are basically equivalent to online prices, young consumers continue to benefit from in-person discounts, while older consumers face the same prices in both distribution channels. Differently, with larger reductions in transportation costs, list prices are higher than online prices and all demographic groups—but the old/rich consumers—receive in-person discounts analogous to those at baseline (in relative terms).

With small reductions in transportation costs, firms offer almost no in-person discounts to older consumers in order to nudge the unrestricted ones to purchase online, which is more profitable due to market expansion. However, by doing this, firms forgo some of the surplus of older captive consumers that they could extract with price discrimination. With larger reductions in transportation costs, purchasing online is more attractive for unrestricted consumers. This provides an opportunity for firms to reintroduce some price discrimination, primarily targeted at captive consumers. In some sense, the higher the reduction in transportation costs guaranteed by the online channel, the easier it is for firms to separate captive from unrestricted consumers through price discrimination in the in-person channel. This implies that firms may benefit from increasing customer convenience (e.g., home delivery, virtual test drives, etc.), in that reduced transportation costs may facilitate the separation of captive from unrestricted consumers.

5.3 Within- versus across-firm effects

In the previous sections, we simulated counterfactuals that involve the introduction of an online distribution channel by all firms simultaneously. Here, we instead consider a counterfactual scenario in which one large car manufacturer starts selling online while its competitors are restricted to sell only in person. We choose the Nissan-Renault group (which includes Dacia and Mitsubishi) as our candidate online firm and re-evaluate counterfactual prices for all firms, for varying levels of transportation cost reductions.

The counterfactual equilibrium prices are presented in Table 12. The top panel shows that the introduction of an online channel leads the Nissan-Renault group to set prices that follow patterns similar to those in Table 11. In contrast, the bottom panel of Table 12 shows that the pricing behavior of the other firms, those without an online distribution channel, is unaffected. These results suggest that it is profitable for firms to price discriminate in the in-person channel even when a competitor starts selling online; although, in this case, firms lose sales to the competitor's online channel.

Table 12: Within versus between-firm effects on price dispersion

	Baseline	Transp. costs red. from online channel			
		-25%	-50%	-75%	-100%
Transaction prices: Nissan-Renault group					
Group 1: Young/Poor	16,468	16,465	16,459	16,508	16,493
Group 2: Young/Rich	16,725	16,679	16,638	16,607	15,786
Group 3: Middle/Poor	17,148	16,334	17,784	17,760	17,802
Group 4: Middle/Rich	17,630	18,237	17,813	17,793	17,850
Group 5: Old/Poor	17,460	17,699	17,722	17,690	17,710
Group 6: Old/Rich	18,388	18,229	17,916	18,054	18,128
Online		18,274	17,815	17,796	17,862
Transaction prices: Other manufacturers					
Group 1: Young/Poor	23,664	23,664	23,664	23,664	23,664
Group 2: Young/Rich	23,921	23,921	23,921	23,921	23,921
Group 3: Middle/Poor	24,330	24,330	24,330	24,330	24,330
Group 4: Middle/Rich	24,808	24,808	24,808	24,808	24,808
Group 5: Old/Poor	24,631	24,631	24,631	24,631	24,631
Group 6: Old/Rich	25,552	25,552	25,552	25,552	25,552

Notes: This table presents transaction prices that occur when only the Nissan-Renault group introduces an online distribution channel. Competitors are restricted to in-person transactions. All prices are computed using a uniform set of weights, $w_j = \sum_d \phi_d s_{jd} / \sum_{k \in \mathcal{J}_g} \sum_d \phi_d s_{kd}$, where $\mathcal{J}_g$ is the set of products offered by group $g = \{\text{Nissan-Renault, Other}\}$.

5.4 Welfare analysis

We turn to the welfare consequences of the introduction of an online distribution channel. We focus on scenarios in which some consumers have restricted access to the online channel as in Table 11. We investigate both aggregate (average) and distributional effects. We begin with the aggregate effects on consumer surplus in Table 13.

Consumers benefit on average from the online channel, with larger gains associated with larger reductions in transportation costs (from around 2% to around 7%). However, our results also reveal some heterogeneity with respect to the average consumer gains and losses across demographic groups. The demographic group that clearly gets the most out of the online channel is group 6, the old/rich consumers. They receive both a lower price and reduced transportation costs compared to the baseline (where they always pay the list price and the full transportation costs). Their surplus can increase by as much as 12% under the most favorable reductions in transportation costs. For other demographic groups, results vary depending on the level of transportation cost reductions.

Figure 6 plots the distribution of consumer surplus by demographic group and separately

Table 13: Effect of online channel on welfare

Counterfactual	Δ Consumer surplus, € per capita per year							Δ Profits (MM€)
	Young/Poor	Young/Rich	Middle/Poor	Middle/Rich	Old/Poor	Old/Rich	All	Total
Baseline	59.1	81.4	153.5	254.6	217.0	356.4	190.8	4,100.7
• -25% transportaiton costs	+0.8	+7.1	+10.7	-16.4	-13.2	+24.4	+3.6	-8.0
• -50% transportaiton costs	+1.3	+14.4	-17.4	-0.7	-11.3	+39.8	+6.9	+43.4
• -75% transportaiton costs	-2.5	+3.0	-7.2	+15.9	-4.2	+40.0	+9.7	+105.1
• -100% transportaiton costs	-6.3	+5.0	-0.6	+26.6	+1.4	+41.7	+12.9	+165.2

Notes: All counterfactual experiments are computed using the 2019 data only. All counterfactual experiments correspond to those in Table 11, where it is assumed that some consumers have restricted access to the online channel. Consumer surplus is in 2018 euros. Profits are in million 2018 euros.

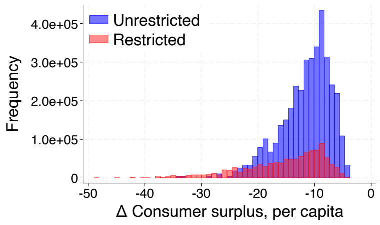
for unrestricted and captive consumers. Figure 6 focuses on the case in which transportation costs are completely eliminated, while the other cases of transportation cost reductions are presented in Appendix Figures A.1, A.3, and A.5. In line with intuition, Figure 6 shows that, overall, unrestricted consumers tend to gain more (or lose less) from the introduction of an online channel. Most young consumers, especially those who are poor and captive to the in-person channel, experience large decreases in consumer surplus. A similar pattern, even though on a more negative scale, can be observed for middle-aged (both poor and rich) and old/poor consumers.

In contrast, the distribution of consumer surplus of old/rich consumers follows a different pattern. Although Table 13 suggests that, on average, old/rich consumers always benefit from the online channel, Figure 6 instead clarifies that those who are captive to the in-person channel, in fact, mostly lose out. Only the unrestricted old/rich consumers benefit from the introduction of an online channel. These results suggest that, although aggregate consumer surplus increases on average (across all demographic groups), most of the benefit indeed accrues to the unrestricted old/rich consumers, while the others either gain little or lose out.

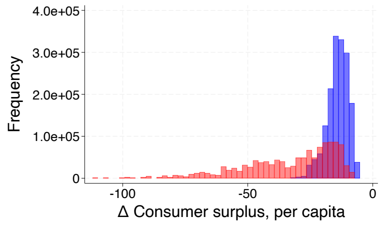
Finally, we consider both industry and car dealers' profits. The last column of Table 13 shows that, in aggregate, industry profits increase for any reduction in transportation costs above 25%. When transportation costs are completely eliminated in the online channel, profits increase by around €165 millions (4%).

Figure 7 presents a breakdown of the changes in profits by brand and focuses on the case in which transportation costs are completely eliminated. The other cases of transportation cost reductions are presented in Appendix Figures A.2, A.4, and A.6. In panel (a), we consider all profits earned by a brand, both from online and in-person sales; while in panel (b) we only consider the profits from in-person sales. We order brands on the x-axis by their market presence, starting with Renault, which operates the largest car dealer network and is the "closest" to consumers. In line with our previous results, brands

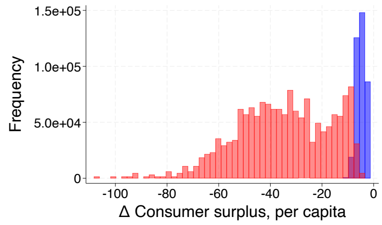
Figure 6: Change in per capita consumer surplus (no transportation costs)



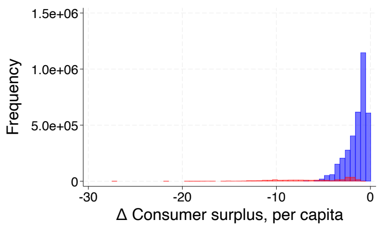
(a) Young/Poor



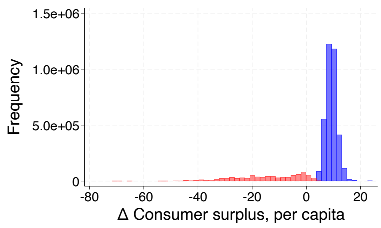
(b) Middle/Poor



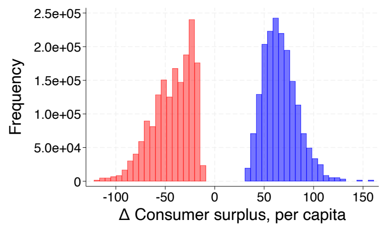
(c) Old/Poor



(d) Young/Rich



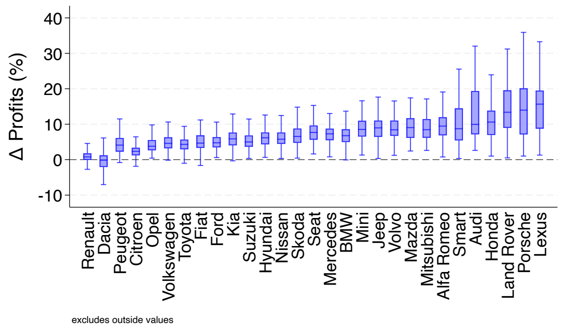
(e) Middle/Rich



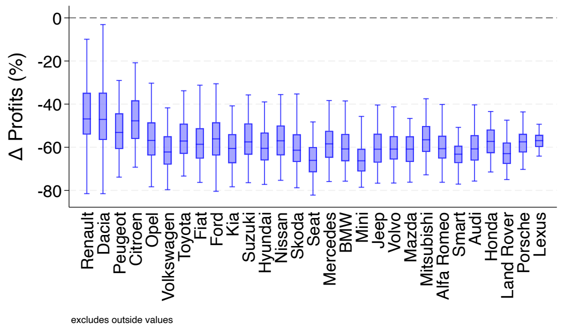
(f) Old/Rich

Notes: These figures plot the change in consumer surplus by demographic group from introducing an online channel as per Table 11, column (5). Consumer surplus is the average per capita consumer surplus at the level of the municipality, and its distribution is weighted by group-specific populations.

Figure 7: Change in brand-level profits (no transportation costs)



(a) Online + in-person sales



(b) In-person sales only

Notes: These figures plot the change in profits at the brand level from introducing the online channel as per Table 11, column (5). Brands are ordered on the x-axis by the total number of car dealers, in decreasing order.

farther away from consumers experience larger increases in profits (in relative terms). Once transportation costs are reduced or removed altogether, consumers respond by shifting some of their purchases toward car models they like more, which sometimes are only sold by car dealers located further away from where they live.

Panel (b) of Figure 7 focuses solely on the profits from the in-person channel. As expected, the introduction of an online channel diverts a large part of sales from car dealers, resulting in large losses for the in-person channel—in the range of 50% or more. These findings raise several questions about the future of "physical" car dealers and their relationship with car manufacturers. We investigate some of these questions next.

5.5 Exit of car dealers and double marginalization

To conclude our analysis, we assess the welfare consequences of two of our maintained assumptions. The first is that the online channel does not lead to exit of car dealers and the second that it does not change the vertical relations between car dealers and car manufacturers, in particular the potential for double marginalization.

A plausible consequence of the introduction of an online distribution channel is to drive out of business some car dealers. If this were the case, our estimates could overstate the associated welfare gains. Since we do not model entry decisions explicitly, we proceed by closing a certain number of car dealers and re-evaluating welfare in this new environment. We base this investigation on a scenario with restricted access to the online channel for some consumers and $\tau = 1$ (as in Table 11, column 5). We consider three counterfactuals: closing the 5%, 10%, and 20% least profitable car dealers, respectively. Detailed counterfactual results are reported in Appendix Table A.4, while we summarize the implied welfare changes in the central rows of Table 14.

These experiments reveal that our estimates of consumer surplus are relatively robust to the exit of car dealers. We focus on the most extreme scenario where 25% of the car dealers go out of business after the introduction of the online channel. In this case, the change in yearly consumer surplus ranges from a loss of €1.5 to a gain of €2.6 per capita depending on the demographic group. The average decrease in surplus is around €0.5 per consumer per year. Industry profits decrease by around €15 millions per year in this worst-case scenario. Meanwhile, the average gain in consumer surplus when no car dealer exits is around €13 and the increase in industry profits is around €165 millions (Table 13). Closing the 20% least profitable dealers thus limits by around 4% the realized gains in consumer surplus and by 9% the realized increase in industry profits. For a more realistic market reallocation of 10% of car dealer exit, these welfare effects are mitigated by 1.5% and 1.6%, respectively.

Another plausible consequence of the introduction of an online channel is that car manufacturers could try to bypass the "middleman" and sell directly to consumers. Following Brenkers and Verboven (2006), we assume that bypassing car dealers leads to cost sav-

Table 14: Effect on welfare with exit of car dealers and cost efficiencies

Counterfactual	Δ Consumer surplus, € per capita per year							Δ Profits (MM€)
	Young/Poor	Young/Rich	Middle/Poor	Middle/Rich	Old/Poor	Old/Rich	All	Total
Baseline (Table 11, column 5)	52.8	86.3	153.0	281.2	218.4	398.1	203.7	4,265.9
Exit of car dealers								
• -5% fewer car dealers	-0.2	-0.3	0.0	+0.0	-0.1	0.0	-0.1	-0.4
• -10% fewer car dealers	-0.2	-0.2	-0.2	-0.2	-0.4	-0.3	-0.2	-2.7
• -20% fewer car dealers	-0.1	+2.6	-1.0	-1.1	-1.5	-1.4	-0.5	-14.8
Cost efficiencies + online delivery cost								
• -5% marginal cost	+4.8	+0.6	+18.7	+35.6	+9.1	+32.1	+18.1	+392.5
• -10% marginal cost	+22.1	+32.4	+53.1	+111.5	+22.6	+87.7	+57.9	+1,206.8

Notes: All counterfactual experiments are computed using the 2019 data only. All counterfactuals consider a scenario in which some consumers are captive to the in-person channel and where the online channel brings a 100% reduction in transportation costs, as per Table 11, column (5). The first row reports our baseline welfare estimates as per Table 11, column (5). Other rows report welfare differences with respect to this baseline scenario. In the first set of counterfactuals, we reduce the market presence of brands by closing 5%, 10%, or 20% of the least profitable car dealers, respectively. In the second set of counterfactuals, we reduce marginal costs by 5% or 10%, and we impose a €400 delivery cost on online sales. Consumer surplus is in 2018 euros. Profits are in million 2018 euros.

ings, in the sense that car manufacturers will base their pricing decisions on marginal costs that are lower than the wholesale prices under double marginalization. Since we do not model vertical relations or wholesale prices explicitly, we assume that selling directly to consumers entails a small reduction in marginal costs, in the range of 5-10%.[19]

In this case, we also assume that car manufacturers must incur a cost for delivering their vehicles to the consumers' doorsteps. To evaluate these delivery costs, we used an online platform specialized in car deliveries, Shiply.com,[20] and asked quotes for various vehicle deliveries for a selection of city pairs in France (shortest distance inquired: 75km, longest distance inquired: 250km). Since all quotes were between €300 and €500 for a single car delivery, we choose an average of €400, which amounts to around 2% of the estimated average marginal cost (see footnote 19). Detailed counterfactual results are reported in Appendix Table A.4, while we summarize the implied welfare changes in the bottom rows of Table 14.

In line with intuition, if the online channel also allowed car manufacturers to save on intermediate costs (such as double marginalization), equilibrium prices would decrease, overall car sales would increase and, as a consequence, both consumer surplus and industry profits would substantially increase. In relation to this scenario, our baseline results that keep marginal costs unchanged (implying that double marginalization also remains unaltered) would underestimate the overall benefits of the introduction of an online channel. Importantly, the fact that industry profits would substantially increase (by around 9% in the case of a 5% marginal cost reduction) means that, in theory, there could be ways of redistributing profits so to guarantee that car dealers are as well off

[19]We take this range of marginal cost reductions from the analysis of vertical relations in the European car market by Brenkers and Verboven (2006), who estimate it to be around 7–8%. Since our estimates suggest an average marginal cost of €18,427, a 5% (10%) reduction corresponds to around €900 (€1800).

[20]Source: https://www.shiply.com.

as in the scenario with double marginalization (or even better off). That is, car dealers could be asked not to charge any margin on car sales, be more than fully compensated with lump sum transfers, and car manufacturers would still make more profits than in the absence of an online channel.

6 Conclusion

In this paper, we investigate the consequences of the introduction of an online distribution channel in the French car industry. We focus on the case in which car manufacturers sell online at a fixed price advertised on their websites, but can discriminate via car dealers by offering personalized discounts based on buyers' observable characteristics. We propose a structural model of oligopolistic competition with differentiated products, unobserved third-degree price discrimination, and transportation costs to study equilibrium prices and the associated welfare effects.

We show that, when all consumers can access the online channel, committing to a uniform online price reduces the extent of in-person discounts, as firms try to avoid cannibalization of their online sales. Differently, when some consumers are captive to purchasing in person, firms charge low online prices to divert internet-savvy consumers online, while continuing to price discriminate in person and to extract most of the surplus of captive consumers.

In terms of welfare, we find that the introduction of an online distribution channel benefits a small portion of consumers while harming the others. These gains and losses depend on three factors: the personalized discounts consumers received before the online channel became available, the reduction in transportation costs that the online channel provides, and whether consumers can take advantage of the online channel. In general, adding an online channel has a redistributive effect on consumer surplus, benefiting older and wealthier consumers who are internet-savvy at the expense of everyone else. Finally, we find that selling online induces market expansion and increases industry profits.

Our analysis is subject to two important caveats. First, we assume that the introduction of an online distribution channel does not alter car dealer networks. Second, we assume that selling directly to consumers online does not affect the vertical relations between car manufacturers and car dealers. Although we probe the robustness of our main results with respect to these important dimensions, our structural model is not fully equipped to deal with these additional complications, and we leave a thorough investigation of these mechanisms to future research.

References

Ayres, I. and Siegelman, P. (1995). Race and gender discrimination in bargaining for a new car. *The American Economic Review*, pages 304–321.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.

Brenkers, R. and Verboven, F. (2006). Liberalizing a distribution system: the european car market. *Journal of the European Economic Association*, 4(1):216–251.

Brown, J. R. and Goolsbee, A. (2002). Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of political economy*, 110(3):481–507.

Brynjolfsson, E., Hu, Y., and Smith, M. D. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management science*, 49(11):1580–1596.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, 13(2):185–213.

Chandra, A., Gulati, S., and Sallee, J. M. (2017). Who loses when prices are negotiated? an analysis of the new car market. *The Journal of Industrial Economics*, 65(2):235–274.

Conlon, C. and Gortmaker, J. (2025). Incorporating micro data into differentiated products demand estimation with pyblp. *Journal of Econometrics*, page 105926.

Corts, K. S. (1998). Third-degree price discrimination in oligopoly: All-out competition and strategic. *The RAND Journal of Economics*, 29:306–323.

D'Haultfœuille, X., Durrmeyer, I., and Février, P. (2019). Automobile prices in market equilibrium with unobserved price discrimination. *The Review of Economic Studies*, 86(5):1973–1998.

Dubé, J.-P. and Misra, S. (2023). Personalized pricing and consumer welfare. *Journal of Political Economy*, 131(1):131–189.

Duch-Brown, N., Grzybowski, L., Romahn, A., and Verboven, F. (2023). Evaluating the impact of online market integration—evidence from the eu portable pc market. *American Economic Journal: Microeconomics*, 15(4):268–305.

Fan, J., Tang, L., Zhu, W., and Zou, B. (2018). The alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce. *Journal of International Economics*, 114:203–220.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the penn world table. *American economic review*, 105(10):3150–3182.

Forman, C., Ghose, A., and Goldfarb, A. (2009). Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management science*, 55(1):47–57.

Freyberger, J. (2015). Asymptotic theory for differentiated products demand models with many markets. *Journal of Econometrics*, 185(1):162–181.

Gandhi, A., Lu, Z., and Shi, X. (2019). Estimating demand for differentiated products with zeroes in market share data.

Goldberg, P. K. (1996). Dealer price discrimination in new car purchases: Evidence from the consumer expenditure survey. *Journal of Political Economy*, 104(3):622–654.

Grieco, P., Murry, C., and Yurukoglu, A. (2023). The evolution of market power in the us automobile industry. *The Quarterly Journal of Economics*.

Harless, D. W. and Hoffer, G. E. (2002). Do women pay more for new vehicles? evidence from transaction price data. *American Economic Review*, 92(1):270–279.

Huang, Y. and Bronnenberg, B. J. (2023). Consumer transportation costs and the value of e-commerce: Evidence from the dutch apparel industry. *Marketing Science*, 42(5):984–1003.

Iaria, A. and Wang, A. (2021). An empirical model of quantity discounts with large choice sets. *Available at SSRN 3946475*.

Liu, Q. and Pierce, D. A. (1994). A note on gauss—hermite quadrature. *Biometrika*, 81(3):624–629.

Miller, N. H. and Osborne, M. (2014). Spatial differentiation and price discrimination in the cement industry: evidence from a structural model. *The RAND Journal of Economics*, 45(2):221–247.

Moraga-González, J. L., Sándor, Z., and Wildenbeest, M. R. (2023). Consumer search and prices in the automobile market. *The Review of Economic Studies*, 90(3):1394–1440.

Morrow, W. R. and Skerlos, S. J. (2011). Fixed-point approaches to computing bertrand-nash equilibrium prices under mixed-logit demand. *Operations research*, 59(2):328–345.

Morton, F. S., Zettelmeyer, F., and Silva-Risso, J. (2001). Internet car retailing. *The Journal of Industrial Economics*, 49(4):501–519.

Murry, C. and Zhou, Y. (2020). Consumer search and automobile dealer colocation. *Management Science*, 66(5):1909–1934.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245.

Nurski, L. and Verboven, F. (2016). Exclusive dealing as a barrier to entry? evidence from automobiles. *The Review of Economic Studies*, 83(3):1156–1188.

Pozzi, A. (2013). The effect of internet distribution on brick-and-mortar sales. *The RAND Journal of Economics*, 44(3):569–583.

Rhodes, A. and Zhou, J. (2024). Personalized pricing and competition. *American Economic Review*, 114(7):2141–2170.

Sagl, S. (2024). Dispersion, discrimination, and the price of your pickup. *Working paper.*

Scott Morton, F., Silva-Risso, J., and Zettelmeyer, F. (2011). What matters in a price negotiation: Evidence from the us auto retailing industry. *Quantitative Marketing and Economics*, 9:365–402.

Shiller, B. R. (2020). Approximating purchase propensities and reservation prices from broad consumer tracking. *International Economic Review*, 61(2):847–870.

Thisse, J.-F. and Vives, X. (1988). On the strategic choice of spatial price policy. *The American Economic Review*, pages 122–137.

Varadhan, R. and Roland, C. (2008). Simple and globally convergent methods for accelerating the convergence of any em algorithm. *Scandinavian Journal of Statistics*, 35(2):335–353.

Yavorsky, D., Honka, E., and Chen, K. (2021). Consumer search in the us auto industry: The role of dealership visits. *Quantitative Marketing and Economics*, 19:1–52.

A Additional Tables and Figures

Table A.1: Demographic characteristics by group

Description	Mean	Std. dev.	10th pct.	Median	90th pct.	Observations
Group 1: Age < 40, *Income* $=$ *Low*						
Median income	16,745	2,680	13,052	17,201	19,795	211,397
Average age	26.9	1.0	25.8	26.9	28.0	211,397
Share of women	0.499	0.035	0.468	0.503	0.524	211,397
Average household size	2.22	0.27	1.92	2.16	2.58	211,397
Urban	0.457	0.498	0	0	1	211,397
Shop online (ψ_1)	0.770	0.421	0	1	1	9,541
Share of population (ϕ_1)	0.247					13
Group 2: Age < 40, *Income* $=$ *High*						
Median income	24,840	4,246	20,832	23,362	31,646	189,754
Average age	27.5	1.1	26.4	27.5	28.7	189,754
Share of women	0.503	0.040	0.462	0.509	0.536	189,754
Average household size	2.35	0.26	1.97	2.38	2.67	189,754
Urban	0.459	0.498	0.000	0.000	1.000	189,754
Shop online (ψ_2)	0.894	0.308	0.000	1.000	1.000	12,131
Share of population (ϕ_2)	0.145					13
Group 3: Age $\in [40, 60)$, *Income* $=$ *Low*						
Median income	17,891	2,248	15,019	18,388	20,237	195,694
Average age	49.5	0.7	48.9	49.5	50.2	195,694
Share of women	0.510	0.037	0.469	0.517	0.540	195,694
Average household size	2.23	0.27	1.92	2.18	2.59	195,694
Urban	0.367	0.482	0	0	1	195,694
Shop online (ψ_3)	0.505	0.500	0	1	1	14,098
Share of population (ϕ_3)	0.151					13
Group 4: Age $\in [40, 60)$, *Income* $=$ *High*						
Median income	25,252	5,179	20,858	23,661	31,944	205,510
Average age	49.4	0.7	48.7	49.4	50.1	205,510
Share of women	0.511	0.031	0.474	0.516	0.540	205,510
Average household size	2.33	0.25	1.96	2.36	2.65	205,510
Urban	0.354	0.478	0	0	1	205,510
Shop online (ψ_4)	0.770	0.421	0	1	1	19,614
Share of population (ϕ_4)	0.185					13
Group 5: Age ≥ 60, *Income* $=$ *Low*						
Median income	18,842	1,568	16,796	19,081	20,600	192,170
Average age	70.5	1.2	69.1	70.5	71.8	192,170
Share of women	0.545	0.050	0.484	0.553	0.596	192,170
Average household size	2.24	0.25	1.94	2.21	2.55	192,170
Urban	0.196	0.397	0	0	1	192,170
Shop online (ψ_5)	0.146	0.353	0	0	1	20,034
Share of population (ϕ_5)	0.093					13
Group 6: Age ≥ 60, *Income* $=$ *High*						
Median income	24,673	4,711	20,932	23,313	29,837	208,943
Average age	70.1	1.0	68.9	70.2	71.2	208,943
Share of women	0.551	0.040	0.500	0.558	0.592	208,943
Average household size	2.25	0.26	1.92	2.23	2.59	208,943
Urban	0.401	0.490	0	0	1	208,943
Shop online (ψ_6)	0.464	0.499	0	0	1	17,724
Share of population (ϕ_6)	0.179					13

Notes: Statistics concerning the median income, age, household size, the share of women, and the level of urbanity are weighted by municipal-level group-specific populations. Statistics concerning the propensity to shop online are weighted by survey weights. For future reference, we denote the probability to shop online by ψ_d (see Section 5). We report a simple year-over-year average of the group-specific population shares, denoted by ϕ_d.

Table A.2: Evidence of price dispersion, by demographic group

	Transaction price		Transaction price — buyback value	
	(1)	(2)	(3)	(4)
Group 1: Age < 40, Income = Low	— Base category (omitted) —			
Group 2: Age < 40, Income = High	1,891.780**	1,672.480*	2,268.324***	2,303.590*
	(744.064)	(910.848)	(710.450)	(1,215.509)
Group 3: Age $[40, 60)$, Income = Low	2,174.785***	1,891.699**	2,766.616***	1,472.772
	(611.006)	(899.649)	(708.275)	(1,307.562)
Group 4: Age $[40, 60)$, Income = High	3,734.329***	3,482.132**	2,961.747***	1,845.544
	(826.167)	(1,418.426)	(853.483)	(1,763.667)
Group 5: Age $\geq g0$, Income = Low	2,412.376***	1,843.251*	1,498.397**	80.061
	(640.156)	(959.263)	(690.785)	(1,148.006)
Group 6: Age ≥ 60, Income = High	2,861.136***	811.124	1,725.900**	-477.457
	(773.938)	(832.586)	(803.734)	(1,130.036)
Female	-306.879	-571.473	-510.496	-516.426
	(346.764)	(487.459)	(372.094)	(621.099)
Value of down payment	0.007	0.006*	0.012	0.011**
	(0.005)	(0.003)	(0.008)	(0.005)
Household: 2 pers.	63.960	-297.050	-133.333	-561.202
	(387.543)	(598.280)	(479.113)	(820.017)
Household: 3 pers.	-151.740	-942.284	-551.323	-1,268.167
	(602.009)	(815.275)	(660.687)	(1,140.575)
Household: 4 pers.	117.470	-1,524.256*	-1,019.702*	-2,687.011**
	(548.932)	(785.119)	(550.327)	(1,132.520)
Household: 5 pers.	-2,166.131***	-1,871.581	-2,437.147***	-4,361.964***
	(794.271)	(1,161.155)	(874.917)	(1,666.250)
Household: 6+ pers.	1,709.236	-1,236.283	3,163.307	-315.518
	(3,021.293)	(1,943.145)	(2,911.596)	(1,961.040)
Urban area: less than 15,000	-1,016.912	2,739.753*	-1,576.286	3,032.271
	(1,326.101)	(1,429.939)	(2,297.180)	(3,751.639)
Urban area: 15,000–24,999	288.247	1,084.762	1,611.884	2,909.341
	(1,581.361)	(1,614.560)	(1,441.358)	(2,866.254)
Urban area: 25,000–34,999	-1,493.065	1,596.119	-1,166.079	1,902.540
	(1,370.082)	(1,903.312)	(1,654.036)	(2,956.974)
Urban area: 35,000–49,999	-1,823.739	-132.945	-1,590.025	1,627.714
	(1,187.328)	(975.182)	(1,281.310)	(1,626.837)
Urban area: 50,000–99,999	-1,339.540	-528.954	-2,156.797***	-1,535.573
	(837.413)	(1,113.736)	(795.857)	(1,116.164)
Urban area: 100,000–199,999	-760.895	316.345	-196.287	325.665
	(803.541)	(876.063)	(699.684)	(998.607)
Urban area: 200,000–499,999	-943.257	215.036	-700.908	84.523
	(713.562)	(835.990)	(617.439)	(1,019.786)
Urban area: 500,000 or more	-971.024	-55.535	-626.253	373.406
	(642.346)	(828.815)	(620.582)	(873.457)
Urban area: Paris greater metro area	-583.891	-67.373	280.185	748.892
	(719.717)	(991.942)	(661.801)	(919.323)
New vehicles only	No	Yes	No	Yes
Fixed effects Car model × engine × new	Yes	Yes	Yes	Yes
Year of purchase × month	Yes	Yes	Yes	Yes
Country of origin of buyer	Yes	Yes	Yes	Yes
Fstat	5.41	1.47	3.96	1.61
Pr(Fstat) > F	< 0.001	0.205	0.002	0.164
Observations	1,283	698	1,283	698
R-squared	0.740	0.801	0.600	0.620

Notes: This table presents the result of a regression of transaction prices on demographic group indicators and other demographic characteristics of buyers, based on a survey of consumers' expenditure. We have excluded observations where the car was purchased following an insurance claim (i.e., the replacement of a damaged vehicle). Columns (1) and (3) include sales of both new and used cars, purchased at a car dealer. Columns (2) and (4) include only new car purchases. The buyback value represents the payment that was received by the consumer for trading in his old car. The F-statistic tests for the hypothesis that the coefficients on the group indicators are jointly zero. Standard errors in parenthesis are clustered at the car model × engine × new/used level. Significance: * < 0.10, ** < 0.05, *** < 0.01.

Table A.3: Car dealer's market presence, by brand

Brand	Stores	Market share	Distance to consumers, in km				
		(%)	Mean	Std. dev.	10th pct.	Median	90th pct.
Renault	1132	17.8	9.01	8.79	2.19	5.35	21.52
Dacia	1101	8.5	9.16	8.98	2.23	5.45	21.80
Peugeot	401	15.9	18.71	16.14	3.62	13.34	42.60
Citroen	396	12.6	13.44	12.09	3.38	8.51	30.57
Opel	351	3.6	14.89	13.56	3.80	9.28	33.84
Volkswagen	303	7.2	15.92	14.48	3.32	9.86	37.07
Toyota	264	5.2	16.50	14.76	3.69	10.21	37.82
Fiat	260	2.7	17.58	15.56	4.01	11.08	41.08
Ford	252	4.6	17.65	15.65	4.13	11.01	40.54
Jeep	225	0.2	29.54	26.80	4.92	19.51	70.21
Kia	215	2.3	18.23	16.42	3.96	11.15	42.68
Suzuki	212	1.6	18.33	16.11	4.10	11.68	42.22
Alfa Romeo	208	0.4	30.36	27.74	5.02	20.11	72.18
Hyundai	198	1.7	19.01	17.04	4.23	11.90	42.90
Nissan	191	2.9	19.13	16.77	4.05	12.18	44.32
Skoda	186	1.1	20.80	18.27	4.40	13.36	48.54
Seat	157	1.6	21.66	18.79	4.67	14.58	48.77
Mercedes	152	2.1	23.17	19.63	4.55	16.70	52.84
BMW	152	1.8	22.93	20.67	4.58	14.98	53.57
Mini	133	1.4	25.47	23.15	4.84	16.70	60.18
Volvo	116	0.5	26.91	24.03	5.10	17.97	63.52
Mazda	112	0.5	29.12	27.44	5.33	19.20	68.06
Mitsubishi	109	0.2	30.52	28.29	5.27	19.01	72.68
Smart	102	0.2	33.43	30.80	4.91	21.82	80.24
Audi	88	2.5	47.31	40.29	7.26	36.33	102.03
Honda	76	0.5	37.33	33.34	5.95	24.94	86.5
Land Rover	62	0.3	44.61	40.50	6.46	30.78	106.2
Porsche	45	0.1	57.80	49.54	7.55	45.10	132.6
Lexus	42	0.2	55.7	48.28	6.38	44.40	127.5
TOTAL	7,241	100					

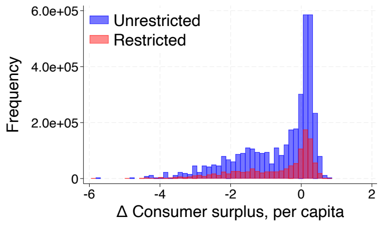
Notes: Brands are ordered by their market presence, defined by their total number of dealers. The market share is computed as each brand's sales over total sales. Distance to consumers is computed over all demographic groups, weighted by their respective municipal-level populations.

Table A.4: Exit of car dealers and cost efficiencies

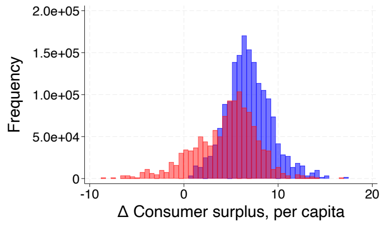
	Baseline	Exit of car dealers			Cost efficiencies	
		-5%	-10%	-20%	-5%	-10%
Transaction prices, uniform weights						
Group 1: Young/Poor	22,261	22,267	22,252	22,229	22,172	21,810
Group 2: Young/Rich	21,976	21,984	21,979	21,944	22,267	21,860
Group 3: Middle/Poor	22,694	22,692	22,694	22,703	22,472	22,388
Group 4: Middle/Rich	22,734	22,733	22,736	22,747	22,844	22,868
Group 5: Old/Poor	22,728	22,727	22,729	22,734	22,676	22,674
Group 6: Old/Rich	23,808	23,807	23,802	23,799	23,635	23,580
Online	22,721	22,720	22,722	22,735	22,161	21,263
Transaction prices, sales-weighted						
Group 1: Young/Poor	18,499	18,526	18,553	18,519	18,045	18,630
Group 2: Young/Rich	15,666	15,604	15,716	15,220	16,151	11,977
Group 3: Middle/Poor	22,352	22,347	22,340	22,327	21,848	22,345
Group 4: Middle/Rich	23,273	23,273	23,280	23,323	23,250	24,312
Group 5: Old/Poor	21,416	21,412	21,409	21,400	21,298	21,376
Group 6: Old/Rich	23,464	23,471	23,476	23,476	23,542	23,596
Online	23,673	23,663	23,660	23,794	24,000	24,720
Sales, in units						
Group 1: Young/Poor	93,647	-139	-41	+54	+9,159	+39,738
Group 2: Young/Rich	68,234	-124	-97	+748	+4,526	+29,458
Group 3: Middle/Poor	136,418	-7	-177	-825	+16,622	+46,391
Group 4: Middle/Rich	250,605	+31	-167	-891	+31,401	+95,925
Group 5: Old/Poor	114,603	-25	-196	-785	+4,763	+11,530
Group 6: Old/Rich	332,212	+9	-124	-962	+27,010	+70,600
All consumers	995,719	-255	-802	-2,661	+93,481	+293,642
Prop. of online sales						
Group 1: Young/Poor	0.570	0.574	0.573	0.572	0.739	0.797
Group 2: Young/Rich	0.629	0.633	0.632	0.622	0.886	0.860
Group 3: Middle/Poor	0.584	0.585	0.585	0.582	0.618	0.685
Group 4: Middle/Rich	0.823	0.823	0.824	0.823	0.854	0.880
Group 5: Old/Poor	0.191	0.192	0.192	0.192	0.233	0.275
Group 6: Old/Rich	0.577	0.577	0.577	0.578	0.596	0.637
Average distance to car models, in km						
Group 1: Young/Poor	1.12	+0.01	+0.02	+0.08	+0.07	+0.13
Group 2: Young/Rich	1.02	+0.02	+0.04	+0.08	+0.09	+0.13
Group 3: Middle/Poor	1.95	+0.01	+0.04	+0.14	+0.04	+0.14
Group 4: Middle/Rich	1.42	+0.02	+0.05	+0.14	+0.03	+0.10
Group 5: Old/Poor	1.67	+0.01	+0.03	+0.09	+0.00	+0.03
Group 6: Old/Rich	1.00	+0.01	+0.03	+0.09	+0.02	+0.05

Notes: All counterfactual experiments are computed using the 2019 data only. The baseline is a scenario with some consumers captive to the in-person channel and no transportation costs for the online channel, as per Table 11, column (5). The other counterfactuals also maintain these assumptions. Transaction prices are in 2018 euros. "Uniform weights" are constructed using the total sales of each product observed in the data, hence are fixed across demographic groups and counterfactual experiments. "Sales-weights" use realized sales for each demographic group and counterfactual experiment. For sales and average distances, we report the values at baseline in the first column, and changes from baseline in the other columns. In the first set of counterfactuals, we reduce the market presence of brands by closing 5%, 10%, or 20% of the least profitable car dealers, respectively. In the second set of counterfactuals, we reduce marginal costs by 5% or 10%, and we impose a €400 delivery cost on online sales.

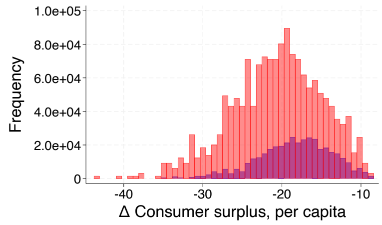
Figure A.1: Change in consumer surplus from online channel (-25% transportation costs)



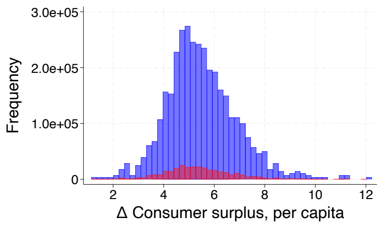
(a) Young/Poor



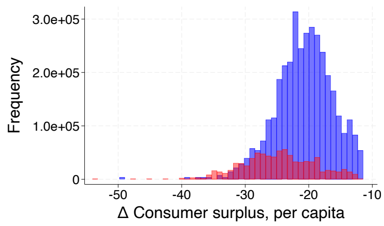
(b) Middle/Poor



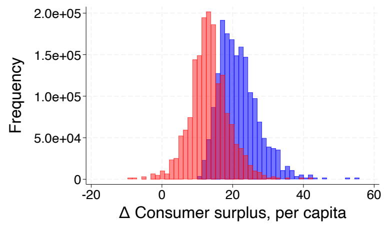
(c) Old/Poor



(d) Young/Rich



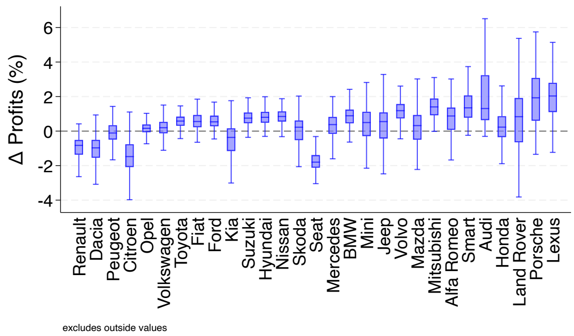
(e) Middle/Rich



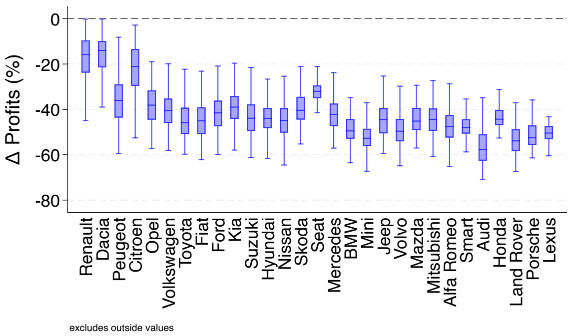
(f) Old/Rich

Notes: These figures plot the change in consumer surplus by demographic group from introducing an online channel as per Table 11, column (2). Consumer surplus is the average per capita consumer surplus at the level of the municipality, and its distribution is weighted by group-specific populations.

Figure A.2: Change in brand-level profits from online channel (-25% transportation costs)



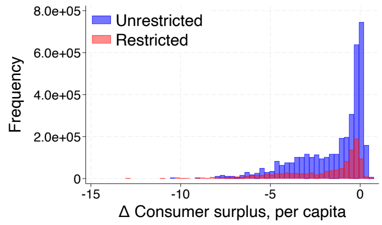
(a) Online + in-person sales



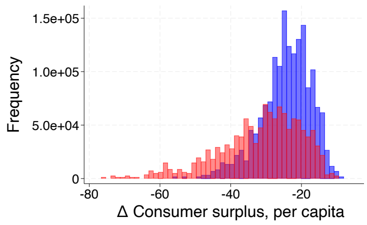
(b) In-person sales only

Notes: These figures plot the change in profits at the brand level from introducing the online channel as per Table 11, column (2). Brands are ordered on the x-axis by the total number of car dealers, in decreasing order.

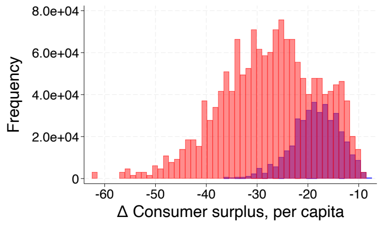
Figure A.3: Change in consumer surplus from online channel (-50% transportation costs)



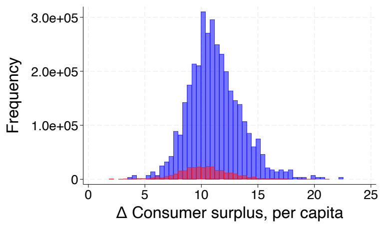
(a) Young/Poor



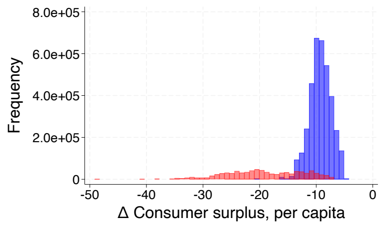
(b) Middle/Poor



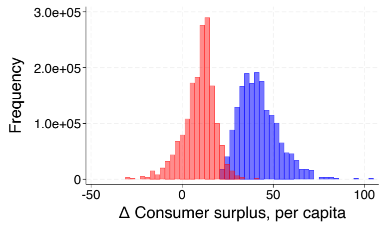
(c) Old/Poor



(d) Young/Rich



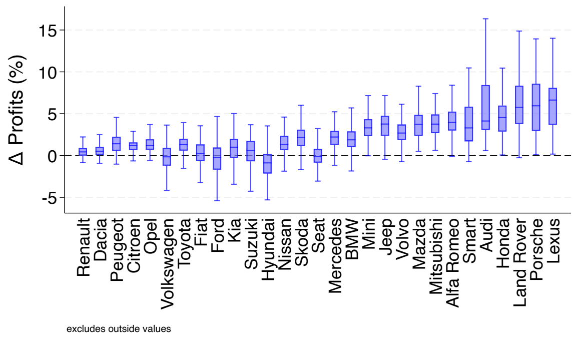
(e) Middle/Rich



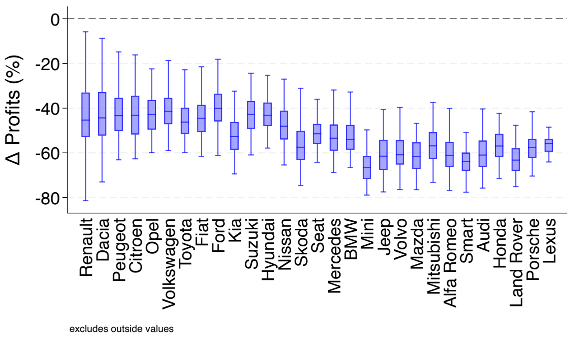
(f) Old/Rich

Notes: These figures plot the change in consumer surplus by demographic group from introducing an online channel as per Table 11, column (3). Consumer surplus is the average per capita consumer surplus at the level of the municipality, and its distribution is weighted by group-specific populations.

Figure A.4: Change in brand-level profits from online channel (-50% transportation costs)



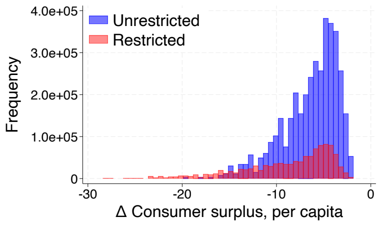
(a) Online + in-person sales



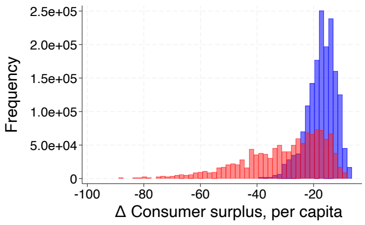
(b) In-person sales only

Notes: These figures plot the change in profits at the brand level from introducing the online channel as per Table 11, column (3). Brands are ordered on the x-axis by the total number of car dealers, in decreasing order.

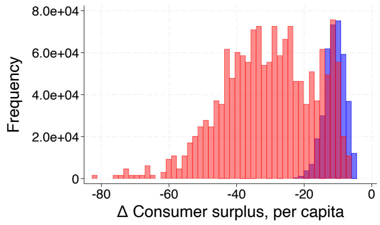
Figure A.5: Change in consumer surplus from online channel (-75% transportation costs)



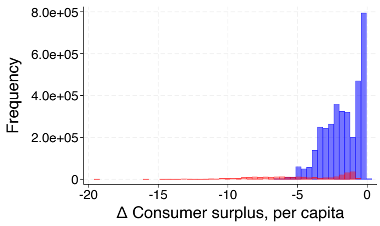
(a) Young/Poor



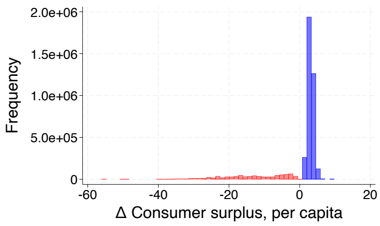
(b) Middle/Poor



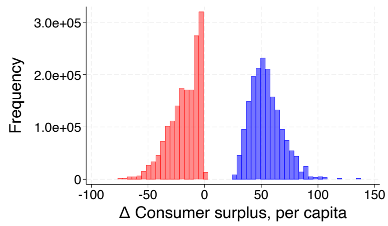
(c) Old/Poor



(d) Young/Rich



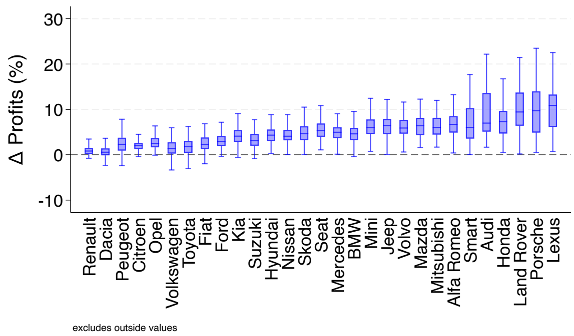
(e) Middle/Rich



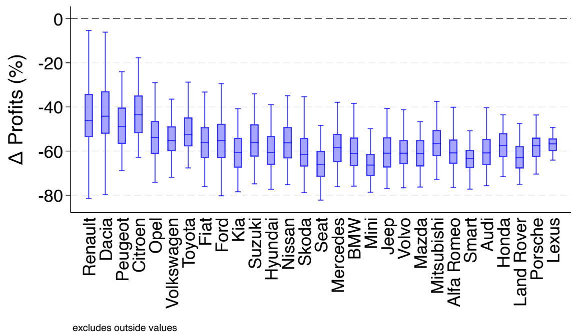
(f) Old/Rich

Notes: These figures plot the change in consumer surplus by demographic group from introducing an online channel as per Table 11, column (4). Consumer surplus is the average per capita consumer surplus at the level of the municipality, and its distribution is weighted by group-specific populations.

Figure A.6: Change in brand-level profits from online channel (-75% transportation costs)



(a) Online + in-person sales



(b) In-person sales only

Notes: These figures plot the change in profits at the brand level from introducing the online channel as per Table 11, column (4). Brands are ordered on the x-axis by the total number of car dealers, in decreasing order.

B Computational details

In this section, we provide additional computational details related to the data, the two-step estimation routine, and the counterfactual simulations.

B.1 Additional details on the data

Construction of demographic groups. We provide more details on the construction of demographic groups. We collect data from two sources, a population survey by municipality and age group, available every five years, and an income survey by municipality and age group, available yearly. Both datasets are available from the Institut National de la Statistique et des Études Économiques (INSEE).[21] We use the following age categorization: young (39 or younger), middle aged (between 40 and 59 included) and old (60 or older). Within age category, we divide municipalities into two evenly sized groups, high- and low-income, according to the median income reported in the income files. Since income is reported for finer increments in age than our age categories, we use a population-weighted average of the median income within age groups and municipalities to assign an income group. In some cases, for very small municipalities, income is not reported separately by age. We then assign the median income of the municipality to all age groups. We drop a small number of municipalities that are too small to report income at all (along with the associated car sales). We obtain six demographic groups, described in Table B.1 below.

Table B.1: Demographic groups definition

Group	Definition	
Group 1	Age 39 or younger	Income in bottom half of age-specific distribution
Group 2	Age 39 or younger	Income in top half of age-specific income distribution
Group 3	Age between 40 and 59	Income in bottom half of age-specific distribution
Group 4	Age between 40 and 59	Income in top half of age-specific distribution
Group 5	Age 60 or older	Income in bottom half of age-specific distribution
Group 6	Age 60 or older	Income in top half of age-specific distribution

Demographic characteristics. We use the population census and income files described above to construct the demographic group-by-municipality average characteristics of consumers. A summary of these consumer characteristics by demographic group is available in Table A.1.

The population files include both the population by municipality and age and the number

[21] Source: https://www.insee.fr/.

of households. We use the number of households (divided by 4) to define market size, and we compute the average household size using the ratio of population to the number of households. The median income is computed as described above and the average age can be approximated with the population data, taking the midpoint of age intervals (5-year increments) and using population weights. Although income and age are used to define the groups, we find that the within-group variation (at the demographic group-by-municipality level) is informative of consumers preferences. The population files offer a breakdown of populations by municipality, gender, and age, allowing us also to compute the share of women by demographic group and municipality.

We merge these data to a survey of population densities, also available at INSEE. Population density is available at the municipality level as a categorical variable indicating whether a given municipality is urban, suburban, or rural. Since not all demographic groups can be found in all municipalities, we define the indicator variable "urbanity" at the demographic group-by-municipality level. Finally, we use a survey of attitudes towards online shopping to determine the propensity to shop online, by demographic group (see Section 5.2 for details).

Construction of the car data. Our car data come from AAA data, which collects data on all car registrations in France. We obtain all new car registrations between 2009 and 2021. The data are aggregated at the level of the car model (a product), age group (in increments of 5 years), and municipality. We merge these data (using the age and the municipality of residence) to our data on consumer demographics to recover demographic groups based on the age and income of buyers, as described above. We create two main datasets. The first is aggregated at the level of the brand-model-engine-year-demographic group; this is our aggregated dataset used for the estimation of the linear parameters. The second dataset is aggregated at the level of the brand-model-engine-year-demographic group-municipality; this is our disaggregated dataset used to compute micro moments and estimate the nonlinear parameters.

We keep the 29 most prominent brands, and keep products with a net price (adjusted for the French Feebate Program) below €100,000. The car data include list prices and some common car characteristics such as horsepower. Horsepower and fuel consumption are not available for electric vehicles in the data. We impute the missing horsepower using an alternative data source and set the fuel consumption of electric vehicles to their fuel-equivalent electricity consumption. We compute fuel costs using various fuel prices interacted with fuel consumption, depending on the engine type (e.g., diesel prices for diesel engines). Finally, we obtain each vehicle's marketing segment (e.g., compact, SUV, etc.) and the country of origin of each model (e.g., the location of the plant that produces each model) from Jato Dynamics. Summary statistics about sales and the main car characteristics are presented in Table 2.

We exclude Tesla from the set of manufacturers under consideration. There are two reasons for this. First, Tesla represents a very small share of total sales for several years of our data. Second, Tesla did not operate a physical network of dealers in France before 2020. There is also ample anecdotal evidence that Tesla does not price discriminate against consumers: consumers buy the car on the website at the posted price. Consequently, the inclusion of Tesla in the analysis would significantly complicate the model and its estimation, while only explaining a very small fraction of sales.

Construction of car dealership data. We obtain the location of all car dealers in France from manufacturers' websites in early 2024, for all brands under consideration except Smart, for which we were unable to scrape the website reliably. In that case, we obtained the addresses and names of car dealers from AutoConcession,[22] an online registry of French car dealers. This registry is not comprehensive for all brands: as an example, it contains about half of the Renault dealers we observe on Renault's website. For other brands, such as Mercedes-Benz, both datasets almost coincide. Since Smart is a brand associated with Mercedes-Benz, we assume that the data on AutoConcession are accurate for Smart also (several Mercedes-Benz dealers also sell Smart).

The data include the full name of dealers, by brand, their addresses, and the type of services offered (e.g., sale of new vehicles, sale of used vehicles, service, etc.). We remove all dealers that do not sell new vehicles. To compute the coordinates of each car dealer, we use OpenCage Geocoding API[23] to recover longitudes and latitudes from dealers' addresses. Most requests on OpenCage API returned a match at the street level or better, however, a small number returned a match at the postal code or city levels. We manually fix these low-quality matches using Google Maps.[24].

Our final dataset includes 7,241 car dealers, selling 29 brands. Summary statistics for these car dealers are available in Table A.3. We report on the number of dealers by brand, their combined market share by brand, and their proximity to consumers. Proximity is defined as the driving distance from consumers' location to the closest dealer of each brand. Additional details on the computation of these driving distances can be found in Section 3.3.

[22]Source: https://www.autoconcession.fr.
[23]https://opencagedata.com.
[24]Source: https://www.google.com/maps/.

Table B.2: Model notation

Notation	Description
i	Individuals
j	Products
d	Demographic groups
m	Municipalities
t	Markets (years)
f	Firms
P	In-person channel
O	Online channel
M_{dmt}	Number/set of individuals in municipality m, demographic group d, and market t
M_{dt}	Number/set of individuals in demographic group d and market t
M_t	Number/set of individuals in market t
$\mathcal{M}$	Set of all municipalities
J_t	Number/set of products available in market t
$\mathcal{J}_{ft}$	Set of products offered by firm f in market t
D	Number/set of demographic groups
T	Number/set of markets (years)
α_d	Price sensitivity of group d
β_d	Preference parameters for car characteristics of group d
γ_d	Distance sensitivity of group d
Π_d	Parameters for demographic interactions of group d
Σ_d	Parameters for random coefficients of group d
θ_d	Set of all demand-side parameters $(\alpha_d, \beta_d, \gamma_d, \Pi_d, \Sigma_d)$
λ	Cost function parameters (λ_1, λ_2)
ψ_{dt}	Share of consumers from demographic group d in market t that have access to the online channel
τ	Transportation cost reduction for online sales
σ	Nesting parameters, online vs in-person channel
p_{jdt}	Price of product j for group d in market t (discriminatory price)
$\overline{p}_{jt}$	List price of product j in market t
c_{jt}	Marginal cost of product j in market t
s_{jdmt}	Market share of product j for group d in municipality m and market t
s_{jdt}	Market share of product j for group d in market t
p^P_{jdt}	In-person price of product j for group d in market t
p^O_{jt}	Online (uniform) price of product j in market t
s^P_{jdmt}	In-person market share of product j for group d in municipality m and market t
s^O_{jdmt}	Online market share of product j for group d in municipality m and market t
s^P_{jdt}	In-person market share of product j for group d in market t
s^O_{jdt}	Online market share of product j for group d in market t
ϕ_{dt}	Share of consumers in market t that belong to demographic group d (M_{dt}/M_t)
w_{dmt}	Share of consumers in demographic group d and market t that live in municipality m (M_{dmt}/M_{dt})

B.2 Additional details on estimation

Notation and specification details. Table B.2 summarizes the notation used throughout the paper. Table B.3 provides additional details on the variables used in the model specification that we estimate. Unless indicated otherwise, car characteristics are used in both specifications of demand and of the marginal cost function. We include horsepower,

Table B.3: Specification details

Variable	Description
Car characteristics	
Price	Price, net of French feebate program, in 10,000 2018 euros (Demand only)
Distance	Driving distance, in 10km (Demand only)
Horsepower	Horsepower, in 100kW
Weight	Curb weight, in 1,000kg
Fuel cost	Cost for driving 100km, in 2018 euros (Demand only)
Fuel consumption	Fuel consumption, in L / 100km (Supply only)
Diesel	=1 if Diesel
Electric	=1 if Electric
Plug-in hybrid	=1 if Plug-in hybiid
Hybrid	=1 if Hybrid
Station wagon	=1 if Station Wagon
Convertible	=1 if Convertible
Trend	Time trend (Supply only)
Demographics (Demand only)	
Median income	Median income, in 10,000 2018 euros
Average age	Average age, in 10 years
Share of women	Share of women, in percentage
Household size	Population / Number of households
Urban	=1 if Urban
	Notes: All demographics are demeaned and are at demographic group-by-municipality level.
Cost shifters (Supply only)	
Input price index	Composite price index based on steel price (56%), polypropylene price (8%), iron price (8%), and aluminum price (10%), interacted with vehicle weight (Supply only)
Real exchange rate	Penn World Table 10.0, `pl_con`, see Grieco et al. (2023) (Supply only)
	Notes: Both cost shifters are lagged one period to reflect planning horizons.
Instruments	
Demand-side	(1) Sum of characteristics of competitors using horsepower, weight, fuel cost
	(2) Number of competitors' products
	(3) Number of competitors' products with same engine type
	(4) Number of competitors' products with same body trim
	Note: Demand-side instruments are the same for all demographic groups.
Supply-side	(1) Sum of characteristics of competitors using horsepower, weight, fuel consumption, input price index, real exchange rate
	(2) Number of competitors' products
	(3) Number of competitors' products with same engine type
	(4) Number of competitors' products with same body trim

curb weight, fuel cost or fuel consumption, and indicator variables for the engine type and body type. We take list prices as given and estimate transaction prices along with the model parameters. Driving distance to the closest dealer of product j is included in the demand specification and captures the transportation costs incurred by consumers to purchase product j. As emphasized in the main text, these transportation costs should be interpreted as inclusive of all interactions between the consumer and the car dealer, e.g., visiting the car dealer for a test drive and future visits related to service and maintenance. Demographic characteristics are used in the estimation of demand and for constructing micro moments. We include the (average) median income, the average age, the share of women, the household size, and an indicator variable for urban municipalities. Intuitively, these demographic variables help us capture heterogeneous preferences across demographic groups and municipalities, while the random coefficients allow us to capture heterogeneity within demographic group and municipality. Cost shifters enter the estimation of the marginal cost function and include an input price index and the real exchange rate between France and the country of origin of the vehicle, defined by the location of the manufacturing plant that produces each model. Both cost shifters are lagged by one year to reflect planning horizons. We also report on the demand- and supply-side instruments used in the estimation, which are constructed following Berry et al. (1995). Additional details are available in Table B.3.

We estimate our structural model based on our data, which only include in-person sales. We introduce the online distribution channel only in counterfactual experiments. The model is estimated using a two-step GMM that estimates the nonlinear and the linear parameters sequentially. We find that this allows for greater computational tractability and a much faster estimation, since we can effectively separate the tasks of iterating over the nonlinear parameters (first step) and of solving for the transaction prices (second step). We provide details on the estimation routine in what follows. To simplify notation, we omit the time subscript from the following expressions, even though estimation is performed over 13 years of data.

Estimation of the nonlinear parameters. The estimation of the nonlinear parameters $(\Pi_d, \Sigma_d, \gamma_d)$ relies on the micro moments discussed in Sections 2.3 and 3.5. We provide an example using the distance to dealers, dist_{jm}, and the construction of the other micro moments follows the same logic. We specify $g_3(\Pi, \Sigma, \gamma) = (g_{31}(\Pi_1, \Sigma_1, \gamma_1),$ ..., $g_{3D}(\Pi_D, \Sigma_D, \gamma_D))$, with

$$\begin{aligned} & g_{3d}(\Pi_d, \Sigma_d, \gamma_d) \\ & \quad = \frac{\sum_{j=1}^{J} \sum_{m \in \mathcal{M}} w_{dm} \cdot s_{jdm}(\Pi_d, \Sigma_d, \gamma_d) \cdot \text{dist}_{jm}}{\sum_{j=1}^{J} \sum_{m \in \mathcal{M}} w_{dm} \cdot s_{jdm}(\Pi_d, \Sigma_d, \gamma_d)} - \frac{\sum_{j=1}^{J} \sum_{m \in \mathcal{M}} w_{dm} \cdot s_{jdm} \cdot \text{dist}_{jm}}{\sum_{j=1}^{J} \sum_{m \in \mathcal{M}} w_{dm} \cdot s_{jdm}}, \end{aligned} \tag{21}$$

where s_{jdm} is the observed market share of group d for product j in municipality m and $w_{dm} \equiv \frac{M_{dm}}{M_d}$ are group-by-municipality specific population weights.

The corresponding market share predicted by the model is obtained by solving the following equation at any given $(\Pi_d, \Sigma_d, \gamma_d)$,

$$s_{jdm}(\Pi_d, \Sigma_d, \gamma_d) = \int \frac{\exp\left(\delta_{jd}(\Pi_d, \Sigma_d, \gamma_d) + \mu_{jdm}(\Pi_d, \Sigma_d, \nu_i) + \gamma_d \text{dist}_{jm}\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_{kd}(\Pi_d, \Sigma_d, \gamma_d) + \mu_{kdm}(\Pi_d, \Sigma_d, \nu_i) + \gamma_d \text{dist}_{km}\right)} dF(\nu_i), \tag{22}$$

where we obtain $\delta_{jd}(\Pi_d, \Sigma_d, \gamma_d)$ by inverting the system of J national-level market shares. We perform this demand inverse using the SQUAREM algorithm, see Varadhan and Roland (2008), and a tight convergence threshold of `1e-14`.

Note that the integral in equation (22) is of dimension one, as we include only a scalar random coefficient on the intercept while the demographic characteristics dem_{dm} and the distance dist_{jm} are observed. We approximate the integral in (22) using a Gauss-Hermite quadrature with 10 nodes, see Liu and Pierce (1994).

The moment conditions $g_{3d}(\Pi_d, \Sigma_d, \gamma_d)$ are computed using a random sample of 3,000 municipalities (per demographic group), representing about 10% of all municipalities in France. To avoid sampling the same municipality twice and preserve regional representativity, we draw the sample of municipalities using systematic sampling and we replace the population weights w_{dm} in (21) by appropriate weights $\widetilde{w}_{dm}$ that take into account our sampling procedure.

The GMM estimator for this first step is

$$(\widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}) = \underset{\Pi,\Sigma,\gamma}{\arg\min} \quad g_3(\Pi, \Sigma, \gamma)' \, \mathbf{W} \, g_3(\Pi, \Sigma, \gamma), \tag{23}$$

where $\mathbf{W}$ is an appropriate weighting matrix. We first obtain consistent estimates of (Π, Σ, γ), say $(\widetilde{\Pi}, \widetilde{\Sigma}, \widetilde{\gamma})$, using as weighting matrix the identity matrix, then we compute the optimal weighting matrix $\mathbf{W} = \mathbf{W}(\widetilde{\Pi}, \widetilde{\Sigma}, \widetilde{\gamma})$, following the best practices described in Conlon and Gortmaker (2025), and finally recalculate (23) to obtain $(\widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma})$.

Our final specification includes distance, income, household size, an urban indicator, and interactions between income and horsepower, urban and horsepower, and urban and weight. We include the associated micro moments, as well as micro moments based on the share of women and the average age. These two additional sets of moments are useful to identify the parameters on the random coefficient, Σ_d, while the other moments are useful to identify the parameters on distance, γ_d, and the parameters on the demographics, Π_d. We define the micro moments at the demographic group-by-year level, so our first-step GMM estimator includes 702 moments (9 "types" of moments $\times$ 6 groups $\times$ 13 years).

Our first-step GMM yields the parameter estimates $(\widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma})$ and the mean utilities $(\delta_j(\widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}))_{j=1,...,J}$. We take these parameter estimates as given in the second-step GMM estimator we describe next.

Estimation of the linear parameters. In the second-step GMM estimator of the linear parameters (α, β, λ), we concentrate out (β, λ) to simplify implementation. We perform the following steps. For given $(\widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma})$ and value of the price coefficients α, we calculate the marginal costs and the transaction prices from the FOCs of the firms:

$$c_j(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}) = \overline{p}_j - \max_d \left\{ \left[\mathcal{D}_d(\alpha_d, \widehat{\Pi}_d, \widehat{\Sigma}_d, \widehat{\gamma}_d)^{-1} \cdot s_d(\widehat{\Pi}_d, \widehat{\Sigma}_d, \widehat{\gamma}_d) \right]_j \right\}$$

and

$$p_{jd}(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}) = c_j(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}) + \left[\mathcal{D}_d(\alpha_d, \widehat{\Pi}_d, \widehat{\Sigma}_d, \widehat{\gamma}_d)^{-1} \cdot s_d(\widehat{\Pi}_d, \widehat{\Sigma}_d, \widehat{\gamma}_d) \right]_j,$$

which do not depend on β and λ. Given these and $(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma})$, we calculate $\widehat{\beta}(\alpha) = (\widehat{\beta}_d(\alpha))_{d=1,...,D}$ and $\widehat{\lambda}(\alpha)$ from the following OLS regressions:

$$\delta_{jd}(\widehat{\Pi}_d, \widehat{\Sigma}_d, \widehat{\gamma}_d) - \alpha_d \, p_{jd}(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}) = X_j' \beta_d + \xi_{jd}$$

and

$$\ln \left(c_j(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}) \right) = X_j' \lambda_1 + W_j' \lambda_2 + \omega_j.$$

Given the residuals of these OLS regressions, we then construct the following demand-side moment conditions:

$$g_{1d}(\widehat{\beta}_d(\alpha), \alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}) = \frac{1}{J} \sum_{j=1}^{J} Z_j' \, \xi_{jd}(\widehat{\beta}_d(\alpha), \alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma})$$

for $d = 1, ..., D$, and supply-side moment conditions:

$$g_2(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}, \widehat{\lambda}(\alpha)) = \frac{1}{J} \sum_{j=1}^{J} Z_{jS}' \, \omega_j(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}, \widehat{\lambda}(\alpha)).$$

Denote the stacked vector of demand- and supply-side moments as $g(\alpha, \widehat{\beta}(\alpha), \widehat{\lambda}(\alpha)) \equiv (g_{11}(\widehat{\beta}_1(\alpha), \alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma})', \; ..., \; g_{1D}(\widehat{\beta}_D(\alpha), \alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma})', \; g_2(\alpha, \widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma}, \widehat{\lambda}(\alpha))')'$. Finally, our second-step GMM estimator is

$$\widehat{\alpha} = \arg\min_{\alpha} \quad g(\alpha, \widehat{\beta}(\alpha), \widehat{\lambda}(\alpha))' \, \mathbf{W} \, g(\alpha, \widehat{\beta}(\alpha), \widehat{\lambda}(\alpha)),$$

where $\mathbf{W} = \text{diag}(\mathbf{W}_1, ..., \mathbf{W}_D, \mathbf{W}_S)$ is a block-diagonal weighting matrix. In practice, since our instruments do not vary across demographic groups, we define $\mathbf{W}_1 = ... = \mathbf{W}_D = (Z' Z)^{-1}$ for the demand-side moments and $\mathbf{W}_S = (Z'_S Z_S)^{-1}$ for the supply-side moments, where $Z = (Z_j)_{j=1,...,J}$ and $Z = (Z_{jS})_{j=1,...,J}$.

As explained in Section 2.3, we account for the estimation error in $(\widehat{\Pi}, \widehat{\Sigma}, \widehat{\gamma})$ arising from the first-step GMM estimator by implementing the variance-covariance correction in Newey and McFadden (1994).

Zero market shares. In the data, we observe a few products with a national market share equal to zero for some demographic group-by-year combination. The traditional approach would be to assume that these products were not available in these markets and remove these products from the choice set. However, this is not realistic in our context: if a product was purchased by some demographic group in a municipality, it must be that it was available to other groups in that municipality as well.

To circumvent this problem, we follow D'Haultfœuille et al. (2019) and compute observed market shares as

$$s_{jd} = \frac{q_{jd} + 0.5}{M_d},$$

where q_{jd} is the total quantity of car model j purchased by demographic group d and M_d is the market size of group d (defined as the number of households of group d divided by 4). As a robustness check, we also re-estimate the model by removing the products with zero market shares from the choice set of all demographic groups. We find that the estimated coefficients are statistically unaffected by this change.

B.3 Solving for prices with online sales

Indirect Utilities. In what follows, we discuss how we compute counterfactual prices when an online distribution channel is introduced. As discussed in Section 4.4, the idea is to approximate the probabilities of purchase implied by our model with a mixed nested logit. Remember that the indirect utility of consumer i from purchasing car model j from distribution channel $\ell \in \{P, O\}$ is given by

$$U^{\ell}_{ijdm} = \delta_{jd} + \mu_{jdm}(\nu_i) + \eta^{\ell}_{jdm} + \zeta_{ijdm} + (1 - \sigma)\epsilon^{\ell}_{ijdm},$$

where $\zeta_{ijdm} + (1 - \sigma)\epsilon^{\ell}_{ijdm}$ and ϵ^{ℓ}_{ijdm} are both distributed as extreme value type I.

Market shares. Consider consumers of demographic group d who live in municipality m. For these consumers, the channel ℓ-specific purchase probability of product j is

$$\begin{aligned} s^{\ell}_{jdm} &= \int s^{\ell}_{jdm}(\nu_i)\, dF(\nu_i), \\ &= \int s_{\ell dm|j}(\nu_i) \cdot s_{jdm}(\nu_i)\, dF(\nu_i), \end{aligned}$$

where

$$s_{\ell dm|j}(\nu_i) = \frac{\exp\left(\frac{\delta_{jd}+\mu_{jdm}(\nu_i)+\eta^{\ell}_{jdm}}{1-\sigma}\right)}{\sum_{\ell\in\{P,O\}} \exp\left(\frac{\delta_{jd}+\mu_{jdm}(\nu_i)+\eta^{\ell}_{jdm}}{1-\sigma}\right)},$$

$$s_{jdm}(\nu_i) = \frac{\exp\left(\mathrm{IV}_{jdm}(\nu_i)\right)}{1+\sum_k \exp\left(\mathrm{IV}_{kdm}(\nu_i)\right)},$$

and $\mathrm{IV}_{jdm}(\nu_i)$ denotes the inclusive value for any given ν_i

$$\mathrm{IV}_{jdm}(\nu_i) = (1-\sigma)\ln\left[\sum_{\ell\in\{P,O\}} \exp\left(\frac{\delta_{jd}+\mu_{jdm}(\nu_i)+\eta^{\ell}_{jdm}}{1-\sigma}\right)\right].$$

The integral over ν_i is calculated using a Gauss-Hermite quadrature with 10 nodes, see Liu and Pierce (1994). Finally, aggregating over municipalities, we obtain the (d,ℓ)-specific market share

$$s^{\ell}_{jd} = \sum_{m\in\mathcal{M}} w_{dm} \cdot s^{\ell}_{jdm},$$

where $\mathcal{M}$ is the set of all municipalities and each $w_{dm} \equiv \frac{M_{dm}}{M_d}$ is a demographic group-by-municipality population weight.

Derivatives. We list the derivatives of the market shares with respect to prices that enter the first-order conditions of the firms. Note that all prices in the in-person and online distribution channels affect all market shares. Given that each "nest" j contains the two distribution channels through which product j can be purchased, the derivatives of the in-person market shares are

$$\frac{\partial s^P_{jd}}{\partial p^P_{jd}} = \frac{\alpha_d}{1-\sigma} \sum_{m\in\mathcal{M}} w_{dm} \int s^P_{jdm}(\nu_i) \left(1 - \sigma\, s_{Pdm|j}(\nu_i) - (1-\sigma)\, s^P_{jdm}(\nu_i)\right) dF(\nu_i),$$

$$\frac{\partial s^P_{jd}}{\partial p^P_{kd}} = -\alpha_d \sum_{m\in\mathcal{M}} w_{dm} \int s^P_{jdm}(\nu_i)\, s^P_{kdm}(\nu_i)\, dF(\nu_i),$$

$$\frac{\partial s^P_{jd}}{\partial p^O_{j}} = -\frac{\alpha_d}{1-\sigma} \sum_{m\in\mathcal{M}} w_{dm} \int s^P_{jdm}(\nu_i) \left(\sigma\, s_{Odm|j}(\nu_i) + (1-\sigma)\, s^O_{jdm}(\nu_i)\right) dF(\nu_i),$$

$$\frac{\partial s^P_{jd}}{\partial p^O_{k}} = -\alpha_d \sum_{m\in\mathcal{M}} w_{dm} \int s^P_{jdm}(\nu_i)\, s^O_{jdm}(\nu_i)\, dF(\nu_i).$$

The derivatives of the online market shares are calculated in a similar way.

Counterfactuals. We adapt the methodology in Morrow and Skerlos (2011) to our framework. We begin by stacking the price vectors, the marginal cost vectors, and the market share vectors to solve for the equilibrium prices in one step:

$$\mathbf{p} = \begin{bmatrix} p^P_1 \\ \cdots \\ p^P_D \\ p^O \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} c \\ \cdots \\ c \\ c \end{bmatrix}, \qquad \mathbf{s} = \begin{bmatrix} \phi_1 s^P_1 \\ \cdots \\ \phi_D s^P_D \\ \sum_d \phi_d s^O_d \end{bmatrix}, \tag{24}$$

where $p^P_d = (p^P_{1d}, ..., p^P_{Jd})'$, $p^O = (p^O_1, ..., p^O_J)'$, $c = (c_1, ..., c_J)'$, $s^P_d = (s^P_{1d}, ..., s^P_{Jd})'$, and $s^O_d = (s^O_{1d}, ..., s^O_{Jd})'$. Note that in $\mathbf{s}$ the market share of each demographic group is multiplied by the share of consumers in that demographic group, so that multiplying the full vector of market shares by the total population yields total sales by demographic group. Equilibrium prices cannot be solved separately for the various demographic groups (as when the online distribution channel is not present), as the uniform online price affects the in-person prices (and vice-versa) through the derivatives above.

Let $\mathcal{H}$ be the ownership matrix, $\mathcal{D}^{\ell,\kappa}_d$ be the $J \times J$ matrix with element (j,k) equal to $\partial s^\ell_{kd}/\partial p^\kappa_{jd}$ for $(\ell,\kappa) \in \{P,O\}^2$, and $\widetilde{\mathcal{D}}^{\ell,\kappa}_d = \mathcal{H} \odot \mathcal{D}^{\ell,\kappa}_d$. Then, we can compute the matrix of demand derivatives for the stacked vectors defined in (24) as

$$\widetilde{\mathcal{D}}(\mathbf{p}) = \begin{bmatrix} \phi_1 \widetilde{\mathcal{D}}_1^{P,P} & 0 & 0 & 0 & 0 & 0 & \phi_1 \widetilde{\mathcal{D}}_1^{O,P} \\ 0 & \phi_2 \widetilde{\mathcal{D}}_2^{P,P} & 0 & 0 & 0 & 0 & \phi_2 \widetilde{\mathcal{D}}_2^{O,P} \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \phi_D \widetilde{\mathcal{D}}_D^{P,P} & \phi_D \widetilde{\mathcal{D}}_D^{O,P} \\ \phi_1 \widetilde{\mathcal{D}}_1^{P,O} & \phi_2 \widetilde{\mathcal{D}}_2^{P,O} & \dots & \dots & \dots & \phi_D \widetilde{\mathcal{D}}_D^{P,O} & \sum_d \phi_d \widetilde{\mathcal{D}}_d^{O,O} \end{bmatrix}. \tag{25}$$

Solving for counterfactual prices amounts to a straightforward fixed-point iteration, based on Morrow and Skerlos (2011), on the stacked system of first-order conditions:

$$\mathbf{p} = \mathbf{c} + \Lambda(\mathbf{p})^{-1} \cdot \widetilde{\Gamma}(\mathbf{p}) \cdot (\mathbf{p} - \mathbf{c}) - \Lambda(\mathbf{p})^{-1} \cdot \mathbf{s},$$

where $\widetilde{\mathcal{D}}(\mathbf{p}) = \Lambda(\mathbf{p}) - \widetilde{\Gamma}(\mathbf{p})$ as in Morrow and Skerlos (2011).

Extrapolation. A computational issue that arises when approximating market shares and their derivatives during the price optimization routine is that $s_{jd}^{\ell} \to \frac{\infty}{\infty}$ for $\sigma \to 1$ using conventional software packages (i.e., both the numerator and the denominator "blow up" past the threshold for infinity which is around 1e700). This prevents the evaluation of counterfactual prices in the limit as $\sigma \to 1$.

To avoid this issue, we rely on linear extrapolation. We evaluate counterfactual prices for two values of σ close to 1 but still numerically manageable, say $\sigma_1 = 0.95$ and $\sigma_2 = 0.96$, and then approximate counterfactual prices at $\sigma \approx 1$ by

$$p_{jd}^{\ell}(\sigma \approx 1) = \lim_{\sigma \to 1} p_{jd}^{\ell}(\sigma) \approx p_{jd}^{\ell}(\sigma_2) + \frac{1 - \sigma_2}{\sigma_2 - \sigma_1} \cdot \left(p_{jd}^{\ell}(\sigma_2) - p_{jd}^{\ell}(\sigma_1) \right). \tag{26}$$

Given these approximated prices $(p_{jd}^{\ell}(\sigma \approx 1))_{j,d}$ for $\ell \in \{P, O\}$, we calculate the corresponding market shares in (16) and (17) (that is, using the $\max\{\cdot, \cdot\}$ formulation).

C Price discrimination in the online channel

We consider counterfactuals in which firms can price discriminate in both the online and in-person distribution channels. While this does not align with firms' stated intentions about online sales (see Introduction), we acknowledge that they could in principle price discriminate also online, as data on consumers are readily available online. For example, consumer demographics could be inferred from browsing histories by third-party data brokers and resold to car manufacturers. The results are presented in Table B.4.

Table B.4: Online and offline price discrimination

	Baseline	Transp. costs red. from online channel			
		-25%	-50%	-75%	-100%
Transaction prices: In-person channel					
Group 1: Young/Poor	21,685	21,685	21,685	21,685	21,685
Group 2: Young/Rich	21,941	21,941	21,942	21,942	21,942
Group 3: Middle/Poor	22,355	22,355	22,356	22,356	22,357
Group 4: Middle/Rich	22,834	22,834	22,835	22,835	22,836
Group 5: Old/Poor	22,658	22,658	22,659	22,659	22,659
Group 6: Old/Rich	23,582	23,582	23,583	23,583	23,584
Transaction prices: Online channel					
Group 1: Young/Poor		21,685	21,685	21,685	21,685
Group 2: Young/Rich		21,941	21,942	21,942	21,942
Group 3: Middle/Poor		22,355	22,356	22,356	22,357
Group 4: Middle/Rich		22,834	22,835	22,835	22,836
Group 5: Old/Poor		22,658	22,659	22,660	22,660
Group 6: Old/Rich		23,582	23,583	23,583	23,584

Notes: This table presents the results from a counterfactual experiment in which all manufacturers can price discriminate both in person and online. Column (1) presents the baseline scenario without online sales. Columns (2)-(5) present counterfactuals in which some consumers are restricted to purchase in person, for different levels of reductions in transportation costs. Transaction prices are weighed by a uniform set of weights constructed from the total sales of each product in the baseline.

Unsurprisingly, in this scenario, firms set almost the same prices in both distribution channels. Prices are slightly increasing as transportation costs reduce in the online channel (just a few euros for the most price-inelastic groups). We interpret this as an attempt on the part of firms to extract the consumer surplus associated with the reduction in transportation costs. This can be seen, for example, by comparing the baseline prices with those in column (5). Regardless of the amount of reduction in transportation costs, the extent of price discrimination in these counterfactuals remains very similar to that observed in the baseline scenario.