Supplemental Appendix

Demand, Competition and Public Policy in the Automobile Industry

Johannes Van Biesebroeck Frank Verboven

1 Data sources

To estimate a discrete choice demand system on the automobile industry, three pieces of information are necessary: product characteristics, consumer choices or market shares, and a few pieces of aggregate information. We discuss them in turn and provide an overview in Table A.1.[1]

Product characteristics Every study that estimates the demand for passenger vehicles requires information on (car) model characteristics, including the Manufacturer Suggested Retail Price (MSRP). BLP obtained this information from the annual printed volumes of two trade press publications (*Automotive News Market Data Book, Ward's Automotive Yearbook*). For studies of the US market this remains the most popular source of information because it is inexpensive to obtain.

Early research on other countries, which tend not to be covered by the trade press, used information from a variety of consumer publications. In many countries, 'car magazines' aimed at the consumer market publish an annual catalog of vehicles for sale on their national market, see for example Verboven (1996). Some private publishers compile information on new and second-hand vehicle prices by model that also contain information on a limited number of model characteristics, such as the *Kelly Blue Book* (United States) or the *Yitzhak Levi pricebook* (Israel) used in Fershtman and Gandal (1998).

More recent studies often use more detailed information for specific variants of each model. This information often comes from companies specialized in data collection and market research, or producing reports with insights on the automotive industry. Often, a very large range of vehicle characteristics can be provided, including segment classifications, dealer invoice price, fuel consumption, standard equipment (such as A/C or traction control), brand, firm ownership information, country of assembly, platform code, etc. The two most important companies that provide information to many researcher are *JATO Analytics* and *S&P Global Mobility*, which has taken over *Polk Automotive Solutions* as well as *IHS Markit*.

[1]The original BLP data can be downloaded from `https://pyblp.readthedocs.io/en/stable/_notebooks/tutorial/blp.html` as part of the PyBLP tutorial of Conlon and Gortmaker (2020). The micro-moments tutorial on the same site provides a link to download the data used in Petrin (2002). Comparable information (to BLP) for the automobile markets in five EU countries can be found at: `https://sites.google.com/site/frankverbo/data-and-software/data-set-on-the-european-car-market`.

Table A.1: Overview of commonly used data sources

Variables	Data	Sources	Examples
(a) Product characteristics			
- Size, market segment, engine power, MSRP,...	1,2,3	Trade press publications Consumer publications Market research companies	*Automotive News, Ward's Automotive Yearbook* *Katalog Der Automobil Review, Autogids* *S&P Global Mobility, Jato Analytics*
- Transaction prices	2	Market research companies Administrative data	Undisclosed (Albuquerque and Bronnenberg, 2012) Texas DMV (Sagl, 2024)
(b) Purchases / registrations			
- Product-level sales or market share	1,3	Trade press publications Market research companies Administrative data	*Automotive News, Ward's Automotive Yearbook* *S&P Global Mobility, Jato Analytics* National vehicle registration databases
- Individual choices	2,3	National consumption survey Administrative data Company-run surveys	Consumer Expenditure Survey (CEX) Vehicle registrations by ZIP code MRI's Survey of the American Consumer (Grieco et al., 2024)
- Second-choice information	2,3	Consumer surveys	CAMIP survey for GM (Berry et al., 2004)
(c) Aggregate information			
	1,2,3	Official statistics	Number of households Income distribution

Notes: The column "Data" indicates the type of study, with the numeric values referring to: 1 = aggregate, 2 = micro, 3 = combined.

Depending on a study's focus, other information is sometimes collected from specialized sources. Vehicle reliability, which is used, for example, in Goldberg (1995), is not available in standard sources, but published for all models for sale on the US market by the consumer magazine *Consumer Reports* or the consulting company *J.D. Power* based on consumer surveys. Similar information for models in operation in Germany is published by the roadside assistance company *ADAC*.

Most studies rely on the MSRP as a price measure. Although this overestimates the actual transaction price due to discounting practices, some studies discuss why it does not necessarily generate biased estimates of price elasticities. Appropriate fixed effects would account for time-invariant country-specific or brand-specific discounting practices, whereas the unobserved demand error term may be interpreted to include remaining measurement error for price. A few studies use transaction-specific prices. Chandra et al. (2017) use data from a market research firm, which in turn obtained it from automobile dealers, to perform a reduced-form analysis. Albuquerque and Bronnenberg (2012) obtained transaction-level information from a market research firm covering 20% of all transactions in the San Diego area between 2004 and 2006. Sagl (2024) uses comparable information on almost 1.5 million individual transactions in Texas between 2011 and 2019, obtained from registrations at the Texas Department of Motor Vehicles. The last two studies use the infor-

mation to estimate an individual choice model. Although such transaction-level prices may seem to be an appealing alternative to MSRP, their use in demand estimation also requires additional assumptions because they are by definition not observed for the non-chosen models. One typical approach assumes that all consumers can purchase at the average transaction price observed in the data set.

Purchases or registrations Empirical studies of demand have used either aggregate sales data, or individual purchase data, or a combination of both.

First consider aggregate, product-level sales data. BLP construct product-level market shares as the fraction of sales in the potential market size (including the outside good). Sales per product is available for multiple years and/or countries, and comes from the two data sources that also provide information on model characteristics. Early studies usually define the product quite broadly as the car model because they rely on information from trade press publications in a printed format, where the level of detail tends to be limited. For example, for most cars, SUVs and minivans the unit of analysis is the so-called "nameplate", for instance, Toyota Camry or Ford Explorer. More recent studies that use information from market research companies often have access to vehicle sales data for variants of models that are defined at much greater detail. Variants can be distinguished by their engine size (displacement categories), fuel type (gasoline, diesel, electric), body type (sedan, station wagon, convertible), transmission (automatic, manual), drive train (FWD, AWD, 4x4), version (luxury level of trim & options, classification varies by company). Often, vehicle sales is available by month, rather than year.

In some countries, including the United States, the aggregate sales information is obtained from the manufacturers, but in many other countries, including most EU countries, the information is aggregated from administrative registrations obtained from governments. Increasingly, researchers are also able to access this administrative data directly from various governments, without having to go through middlemen. That type of information is often available for smaller geographical areas (states or postal codes) and the analysis can pool over multiple markets. For example, the province-level analysis for China in Barwick et al. (2021) uses vehicle registration information from the State Administration of Industry and Commerce. For some countries, including the case of China studied in Hu et al. (2014), this information is available from the national association of automobile manufacturers, which generally belong to the International Organization of Motor Vehicle Manufacturers (OICA).

Next, consider data on individual vehicle purchases rather than aggregate market shares. Such data may sometimes come from the above-mentioned administrative records collected by governments. A main advantage of such individual registration data is that they cover the universe of transactions. The limitation is that they typically include only a narrow set of demographic variables (postal code, gender, age), and these are not available in combination. Alternatively,

individual-level information may come from surveys which cover a (representative) sample of consumers. The Consumer Expenditure Survey (CEX) of the United States is a popular example, see for example Goldberg (1995). Similarly, researchers rely on propriety surveys by market research firms or organized by the OEMs themselves.

Studies that add information on second choices to help identify cross-product substitution patterns, for example, Berry et al. (2004) and Grieco et al. (2024), invariably rely on consumer surveys with such stated second choices. Naturally, this type of information is of interest to car manufacturers and market research data firms organize such surveys.

Studies that combine aggregate data with individual data (referred to as data type 3 in Table A.1) do not necessarily require access to the entire data set of individual choice purchases. It is sufficient to have access to relevant micro-moments of this data set, which may sometimes even be reported in public documents. Micro-moments include the average value of consumer characteristics conditional on a certain purchase decision, or sales or market shares conditional on certain consumer characteristics. For example, in his study of the impact of the minivan Petrin (2002) relies on the average household size by market segment (such as the minivan), without requiring full access to the survey data. Conversely, when studying the value of dealer proximity, Nurski and Verboven (2016) make use of market shares conditional on consumer locations (postal codes).

Aggregate information Finally, in order to estimate a demand model, one needs aggregate information on the total potential market size taken from official government statistics, which is often approximated by the total number of households. Many studies impose heterogeneity on the price sensitivity that varies with income, as indicated in Table A.2. The model is estimated by drawing simulated consumers from the empirical income distribution.[2] The observed distribution of annual mileage driven can be used similarly to allow for heterogeneity in the preference for fuel efficiency.

[2]Naturally, many other sources than government statistics can be used to infer the relevant income distribution. For example, Barwick et al. (2021) uses annual surveys conducted by the Ford Motor Company to obtain information on the conditional income distribution specifically of new car buyers.

2 Tables summarizing the covered literature

Table A.2: Car demand studies: instruments and covered heterogeneity

Authors	Year	Data	Spec.	Price	Power	Size	Fuel	Other	Fixed Effects	IVs
Hashmi & Van Biese.	2016	1	L	no						BLP
Li, Tong, Xing, Zhou	2017	1	L	no					model*mkt	incl. FE
Verboven	1996	1	NL	no		1		1	origin	BLP
Fershtman & Gandal	1998	1	NL	no		1				BLP+cost
Fershtman et al.	1999	1	NL	no		1				BLP+cost
Goldberg & Verboven	2001	1	NL	α/y_i		1		1	origin*ctry	BLP+cost
Brenkers & Verboven	2006	1	NL	α/y_i		5+		1	model	BLP
Moral & Jaumandreu	2007	1	NL	no		1				other
Tovar	2012	1	BST	no		1		1	firm	BLP+cost
Adamou et al.	2014	1	NL	no		1		1		BLP
Park & Rhee	2014	1	NL	no					brand	BLP+tax
Blonigen et al.	2017	1	NL	no		1			segment	BLP
Durrmeyer & Samano	2018	1	NL	no		1			firm	BLP
Zhou & Li	2018	1	NL	no				1	model*mkt	other
Hu, Yin & Zhao	2023	1	NL	no				1	brand*mkt	BLP+cost
Grigolon	2021	1	ONL	α/y_i		5+			model	BLP+cost
BLP	1995	1	RCL	α/y_i	1	1	1	1		BLP
Berry et al.	1999	1	RCL	α/y_i		1	1			BLP
Sudhir	2001	1	RCL	no	1	1			origin	BLP
Kitano & Ohashi	2009	1	RCL	α/y_i	1			1		BLP+cost
Grigolon & Verboven	2014	1	RCL	α/y_i	1	2	1		model	BLP
Hu, Xiao & Zhou	2014	1	RCL	σ^α					origin	BLP+cost
Huse & Lucinda	2014	1	RCL	σ^α					brand	BLP
Reynaert & Verboven	2014	1	RCL	σ^α	1	1			model	BLP+cost
Guajardo et al.	2016	1	RCL	σ^α	1	1	1		brand	BLP
Ohashi & Toyoma	2017	1	RCL	α/y_i	1					BLP
Cosar et al.	2018	1	RCL	α_h	1		1		brand*mkt	BLP+cost
Grigolon et al.	2018	1	RCL	α/y_i	1	1			model	BLP+cost
Miravete et al.	2018	1	RCL	α/y_i	1	1	1		brand	BLP
Reynaert	2021	1	RCL	α/y_i			1	1	model	BLP+cost
Shen et al.	2021	1	RCL	α/y_i	1	1	1		origin	BLP
Springel	2021	1	RCL	σ^α					model	BLP+cost
Guo & Xiao	2023	1	RCL	α_h, σ^α	1		1	1	brand	BLP
Ale-Chilet et al.	2025	1	RCL	σ^α				1	model*body	BLP+cost
Remmy	2025	1	RCL	α/y_i			1	1	firm	differentiation
Goldberg	1995	2	NL	α_h				5+	brand	FE
Goldberg	1998	2	NL	α_h				5+	brand	FE
Bachman et al.	2023	2	NL	α_h		3		1		BLP
Train & Winston	2007	2	RCL	α_h, σ^α	1		1	5+	brand	BLP
Langer	2011	2	RCL	α_h, σ^α	1		2	5+	model	BLP
Albuquerque & Bronn.	2012	2	RCL	α_h		1		1		BLP
Thomassen	2017	2	RCL	α_h	3	1	3	5+	model	FE
Murry & Zhou	2020	2	RCL	α_h, σ^α		1	1	5+		BLP
Huse & Koptyug	2022	2	RCL	no			1		model	FE
Sagl	2024	2	RCL	α_h, σ^α		1	1	5+	model*trim	FE

(continued on next page)

(continued from previous page)

Authors	Year	Data	Spec.	Price	Power	Size	Fuel	Other	Fixed Effects	IVs
Petrin	2002	3	RCL	α_h	1		1	5+		BLP
Berry et al.	2004	3	RCL	α_h, σ^α	3		1	5+	brand	BLP
Beresteanu & Li	2011	3	RCL	α_h	1	5	1	4	product	BLP
Schiraldi	2011	3	RCL	α/y_i				1		other
Li, Xiao & Liu	2015	3	RCL	α_h, σ^α	1	1	1		firm	BLP+cost
Nurski & Verboven	2016	3	RCL	α_h, σ^α	1	3	3	3	origin	BLP
Murry	2017	3	RCL	α_h, σ^α				4	brand	BLP
Li	2018	3	RCL	σ^α	1			1	model	FE
Wollmann	2018	3	RCL	no				5+		BLP+cost
D'Haultfœuille et al.	2019	3	RCL	α_h, σ^α	2	2		5+	brand	BLP+cost
Barwick, Cao & Li	2021	3	RCL	σ^α			1	1		BLP+cost
Durrmeyer	2022	3	RCL	α_h, σ^α	1	1	2	1	brand	BLP
Moraga et al.	2023	3	RCL	α_h	2			5+		cost
Wang & Xing	2023	3	RCL	α_h, σ^α			1	2	firm*year	subs.+cost
Barwick, Kwon & Li	2024	3	RCL	α_h, σ^α				2	br*eng*mkt	cost
Grieco et al.	2024	3	RCL	α_h	1	4	1	5+	brand	cost
Heid, Remmy & Reynaert	2024	3	RCL	α^α			3		firm	

Notes: This table provides background information for Table 1 and Table 2 in the main text. Studies are sorted by the entries in columns Data-Spec.-Year-Authors (in that order). The values in the column "*Data*" refers to: 1 = aggregate, 2 = micro, 3 = combined. The column "*price*" indicates how heterogeneity for price is incorporated: not (no), through empirical income distribution (α/y_i), through demographic interactions (α_h), or through the standard deviation of the random coefficient (σ^α). The columns "*power*", "*size*", "*fuel*", and "*other*" indicate how many heterogeneity coefficients for these characteristics are statistically significant; nested logit parameters enter under "*size*" if nests are segments, and under "*other*" if nests refer to country of origin or fuel engine type. The column "*IVs*" indicates the instruments used to account for price endogeneity: BLP (traditional or variants); cost (exchange rate, input*characteristic interactions); or other (for example lagged prices). The column "*fixed effects*" refers to brand, brand*market interaction, model, or model*market interactions (empty cells imply either no or higher level interactions such as firm or segment).

Table A.3: Other studies on the automotive industry

Authors	Year	Data	Section	Method	Dep.Var.	Note
Bresnahan	1981	1	Supply side	vertical diff.	q	one-dimensional heterogeneity, no ξ_j
Mertens & Ginsburgh	1985	1	Supply side	reduced-form	p	hedonic price regression
Verboven	2002	1	Supply side	engine demand	q, p	differenced-out product demand
Knittel	2011	1	Supply side	reduced-form	X	MC curve trading off characteristics
Goldberg	1996	2	Supply side	reduced-form	p	dealer discount
Scott Morton et al.	2003	2	Supply side	reduced-form	p	compares offline & online markets
Anderson et al.	2015	2	Supply side	reduced-form	X	choice of brand
Chandra et al.	2017	2	Supply side	reduced-form	p	compares margins within model-trim
Bresnahan	1987	1	Competition	vertical diff.	q	one-dimensional heterogeneity, no ξ_j
Bertomeu et al.	2021	1	Competition	reduced-form	q^{future}	features of production forecasts
Feenstra & Levinsohn	1995	1	Trade	vertical diff.	p	multi-dimensional heterog., no ξ_j
Head & Mayer	2019	1	Trade	CES demand	q	prod. location & market entry choices
Chen et al.	2021	1	Trade	reduced-form	q	subsidy as explanatory variable
Sun et al.	2021	1	Trade	reduced-form	q	logit, but price absorbed by FE
Klier & Linn	2010	1	Environment	reduced-form	q	gasoline price as explanatory variable
Li, Linn & Spiller	2013	1	Environment	diff-in-diff	q	Canada as control
Allcott & Wozny	2014	1	Environment	reduced-form	p	logit, $\ln(s_j/s_0)$ included in ξ_j
Jacobsen & van Bent.	2015	1	Environment	reduced-form	$p, scrap$	gasoline price as instrument for p
Klier & Linn	2015	1	Environment	reduced-form	q	vehicle tax as explanatory variable
Busse et al.	2013	2	Environment	reduced-form	q, p	new and second-hand cars
Grigolon et al.	2024	2	Environment	reduced-form	$usage$	mileage and fueling choices
Bento et al.	2009	3	Environment	dynamic model	q, $scrap$	replacement by new or second-hand
Jacobsen	2013	3	Environment	dynamic model	$q, p, scrap$	adds supply side to Bento et al. (2009)
Gillingham et al.	2021		Environment	reduced-form	p	natural experiment on fuel eff. report
Esteban & Shum	2007	1	Conclusions	dynamic model	p	inverse demand; linear-quadratic
Gillingham et al.	2022	1	Conclusions	dynamic model	$q, p, scrap$	new and second-hand cars
Attanasio	2000	2	Conclusions	dynamic model	q	car inventory demand
Busse et al.	2006	2	Conclusions	diff-in-diff	p	pass-through of discounts by dealer
Gavazza et al.	2014	2	Conclusions	dynamic model	q, p	second-hand car market
Bertolotti et al.	2023	3	Conclusions	dynamic model	pq, p, X	expenditures and quality upgrading

Notes: This table contains papers that are relevant for the three policy questions we study, but do not use a BLP demand model. Studies are sorted by the entries in columns Section-Data-Year-Authors (in that order). The column "*Data*" refers to: 1 = aggregate, 2 = micro, 3 = combined. The column "*Dep. Var.*" refers to: q = quantity or market share, p = price, $scrap$ = scrappage rate, X = other product characteristics.

Table A.4: Own-price elasticities and markups in BLP demand studies

Author	Year	Data	Spec.	Elasticity			Markup		
				Avg.	Range	Range of:	Avg.	Range	Range of:
Goldberg	1995	1	NL		[-1.1, -5.2]	segment avgs		[15.1, 69.0]	segment*origin avgs
Langer	2011	1	NL		[-6.4, -17.8]	4 demogr. groups		[25.6, 35.3]	4 demogr. groups
Hu, Xiao & Zhou	2014	1	RCL	-3.0	[-2.8, -3.2]	25-75 pctiles	34.8		(Bertrand conduct)
Murry & Zhou	2020	1	RCL	-4.1	[-1.4, -10.7]	all products	29		
Sagl	2024	1	RCL	-7.3			21		
Verboven	1996	2	NL		[-5.0, -15.0]	sel. products		[8.1, 19.9]	segment avgs
Goldberg & Verboven	2001	2	NL		[-3.6, -6.5]	country*origin avgs		[16.0, 40.0]	country*origin avgs
Brenkers & Verboven	2006	2	NL	-7.6	[-10.9, -4.5]	segment avgs		[16.0, 24.0]	country avgs
Durrmeyer & Samano	2018	2	NL	-3.2, -2.5			29, 42		
Hu, Yin & Zhao	2023	2	NL	-11.8			19	[11.3, 28.0]	overall range
BLP	1995	2	RCL		[-3.1, -6.8]	13 sel. products		[15.9, 32.8]	13 sel. products
Ohashi & Toyoma	2017	2	RCL		[-1.6, -7.2]	segment avgs		[35.0, 65.0]	segment avgs
Miravete et al.	2018	2	RCL		[-2.5, -3.9]	segment avgs		[44.3, 28.7]	segment avgs
Cosar et al.	2018	2	RCL		[-14.1, -15.7]	5 sel. products		[6.0, 12.4]	firm*country avgs
Ale-Chilet et al.	2021	2	RCL	-3.3			33		
Shen et al.	2021	2	RCL		[-3.5, -8.1]	8 sel. products		[15.2, 22.5]	avg imported, domestic
Beresteanu & Li	2011	3	RCL	-8.4	[-3.7, -11]	11 sel. products	17.7	[15.7, 38.6]	11 sel. products
Li, Xiao & Liu	2015	3	RCL		[-3.1, -4.5]	sel. products		[22.3, 32.4]	sel. products
Nurski & Verboven	2016	3	RCL	-3.1	[-1.5, -5.0]	10–90pctiles	43	[21.0, 70.0]	10-90pctiles
Murry	2017	3	RCL	-4.9	[-4.4, -6.2]	sel. products	20	[15, 22]	dealer avgs
DDF	2019	3	RCL	-5.6	[-3.9, -6.2]	avg by demo group	20.6	[17.5, 28.5]	avg by demo group
Moraga et al.	2023	3	RCL	-5.7	[-4.4, -6.7]	9 sel. products	32.2	[15.9, 23.9]	9 sel. products
Grieco et al.	2024	3	RCL	-5.1	[-4.2, -5.8]	annual avgs		[18.5, 32.5]	median 2018 vs 1980

Notes: This table reports average and ranges of own-price elasticities and percentage markups, as reported in various BLP demand studies. Only studies that report information on both elasticities and markups are retained. Studies are sorted by the entries in columns Data-Year-Authors (in that order). The two numbers in Durrmeyer and Samano (2018) refer to France and the United States. The numbers in Alé-Chilet et al. (2025) refer to median instead of average.

References

Albuquerque, P. and B. J. Bronnenberg (2012). Measuring the impact of negative demand shocks on car dealer networks. *Marketing Science 31*(1), 4–23.

Barwick, P. J., S. Cao, and S. Li (2021). Local protectionism, market structure, and social welfare: China's automobile market. *American Economic Journal: Economic Policy 13*(4), 112–151.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy 112*(1), 68–105.

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

Conlon, C. and J. Gortmaker (2020). Best practices for differentiated products demand estimation with PyBLP. *RAND Journal of Economics 51*(4), 1108–1161.

Fershtman, C. and N. Gandal (1998). The effect of the Arab boycott on Israel: The automobile market. *RAND Journal of Economics 29*(1), 193–214.

Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the US automobile industry. *Econometrica 63*(4), 891–951.

Grieco, P. L., C. Murry, and A. Yurukoglu (2024). The evolution of market power in the US automobile industry. *Quarterly Journal of Economics 139*(2), 1201–1253.

Hu, W.-M., J. Xiao, and X. Zhou (2014). Collusion or competition? Interfirm relationships in the Chinese auto industry. *Journal of Industrial Economics 62*(1), 1–40.

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

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy 110*(4), 705–729.

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

Verboven, F. (1996). International price discrimination in the European car market. *RAND Journal of Economics 27*(2), 240–268.