Consumer Demand Estimation

Merino Troncoso, Carlos

UNED

7 January 2021

DEMAND ESTIMATION

1. Introduction

In this chapter I will review the main methodologies used in economics for demand estimation, focusing on recent trends such as the structural approach and machine learning techniques. As one can imagine the literature review is extensive so due to space limitations I can only provide a summarized view of each theory. Nevertheless, the interested reader has a comprehensive bibliography at the end of the chapter for extensions and examples. There is also another barrier when explaining any concept in economics. Economics is widely based on Mathematics, Statistics and Econometrics so it is not possible to explain it without its usage. As it is not possible review econometrics and mathematics in this chapter I will refer to specific texts, and an appendix will give the reader a brief summary of the main concepts. Demand is usually the first step in the study of a market. Intuitively, suppliers only start production when they identify consumer interest in a particular good. All models reviewed try to solve the problems that traditionally have embarrassed demand estimation: identification, endogeneity and simultaneity. There is no perfect solution to them, each model has its advantages and limitations and are based on assumptions that are often irreal, so the model in itself is in all cases only an approximation of demand.

2. Traditional approach to consumer behaviour

As I have mentioned above, my intention is to do a review on consumer demand covering the traditional approach developed by Marshall (1890), to more recent approaches used in Industrial Organization literature such as Nevo (2000) and a brief introduction of machine learning (2015). Demand is the most important component for empirical competitive analysis. It is not possible to quantify the change in a company's behavior if we do not have information about the potential response of consumers. Demand is based on consumer tastes and there are two empirical findings any consumer theory has to deal with: consumers tastes are hetereogeneous and products are differentiated (2000) .

Basic concepts

We recall the basic elements of consumer demand. I only provide a quick overview of the main elements of traditional consumer theory. Davies and Garcés provide an application to antitrust analysis (2009)**.**

Demand function describe the amount of a good a consumer would buy based on variables that affect this decision such as the P_i vector of consumer's price or income y .Figure 1 shows a typical example of linear demand function: p=20-2q which is generally written as $p_i = D(q_i, y)$ and is the inverse demand Marshall curve[1] (1890).

The typical negatively sloped linear demand is based on strong assumptions. It is represented as the demanded quantity for each price level of the good when all other variables that affect demand such as income levels and prices of substitute or complementary products remain fixed. All variables remain fixed except price and quantity of the good when we move along the demand curve. The slope of the demand curve indicates at each point what the consumer would be willing to reduce (increase) the amount demanded if the price increases (falls) while income and any other variable remain fixed.

[1] *"There is then one general law of demand: - The greater the amount to be sold, the smaller must be the price at which it is offered in order to find purchasers; or, in other words, the amount demanded increases with a fall in price and diminishes with a rise in price."* Alfred Marshall, Principles of Economics (eighth edition), p. 99.

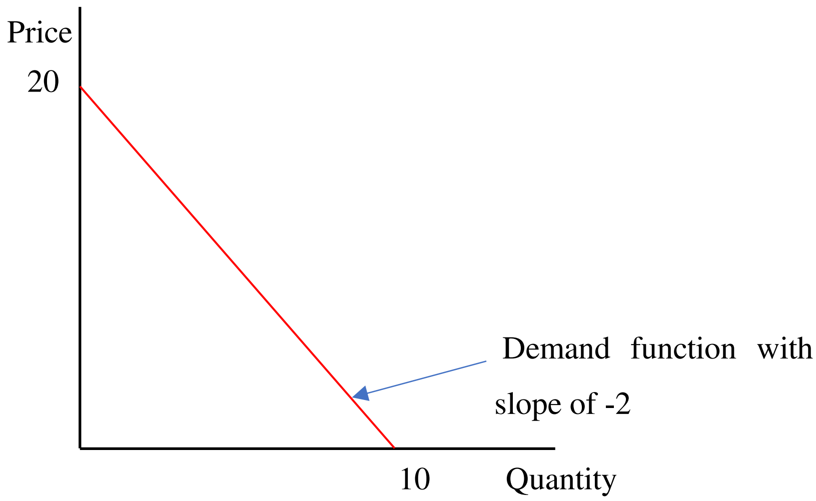


Figure 1: Inverse Demand Function

In the figure above (2009), if the price increases by 10 euros, demand will fall by 5 units. Consumers will not buy more units if the price is greater than 20 because at that level the price is greater than the value that the consumer assigns to the first unit of the good.

An interpretation of this curve is that it reflects the maximum price that the consumer wants to pay for units of the good q in question. Intuitively, consumers valuation will be lower as they own more units of the good. It is this declining marginal valuation that ensures that the demand curve typically has negative slope. Consumers will acquire a unit only if marginal valuation is greater than price so it can be said that the curve describes the consumers marginal valuation of the good.

Given this interpretation, the inverse demand curve describes the difference between the consumer valuation of each unit and the current price paid for it. This difference is called consumer surplus. At any price, we can add up the consumer surplus available in all units consumed (those with marginal valuation above price).

In a homogeneous product market, all goods are identical and perfect substitutes. This implies in theory that they all have the same price. In a market with differentiated product, they are not perfectly substitutes and prices will vary. In this market, demand for any product is determined by its price and that of potential substitutes. In practice, homogeneous markets end up being differentiated when viewed in detail. However, the assumption of homogeneity can be a reasonable approximation.

Demand functions

Demand functions are derived from consumer choice, assuming consumers maximize a utility function subject to a budget constraint. The existence of this utility function can be induced from some non-trivial assumptions. A detailed explanation of these assumptions is covered in Mas-Colell (1995). Maximizing utility is equivalent to choosing the most preferred product set given the budget constraint.

Mathematically, the problem can be represented as a consumer who chooses to maximize their welfare (utility function u) subject to disposable income y so that their total expenditure does not exceed their income:

$$Max\ u_i(q_1, q_2, \dots, q_n)$$

$$s.a. \sum_{i=1}^{n} p_i\, q_i = y$$

Where p_i and q_i are price and amounts of good, i, $u_i(q_1, q_2, \dots, q_n)$ is the individual's demand utility associated with i consuming that vector of amounts, y_i and is the individual's i disposable income.

The first order conditions of this problem are:

$$\frac{\delta u_i(q_1, q_2, \dots, q_n)}{\delta q_i} = \lambda p_i$$

$$\frac{\delta u_i(q_1, q_2, \dots, q_n)}{\delta p_i} = \lambda$$

Along with the budget constraint it gives *I+1* equations with *1+1* parameters: the quantities *I* and the value of the Lagrange multiplier, λ.

In the optimum, the first order conditions imply that the Lagrange multiplier is equal to the marginal income, which we assume constant. We assume, that consumer behavior is described by a utility function with an additive and separable q_i good, the price p_1 is standardized to *1*. The price of reference good q_1 is generally called money and its inclusion allows an interpretation of the first order conditions. In these circumstances a consumer that maximizes utility shall choose a basket of products so that the marginal profit of the last monetary unit disbursed on each product is equal to the marginal profit, i.e. 1.

In general, the solution to maximization describes the individual's problem as a function of the prices of all goods sold and consumer income. If we index the goods with i, we can write the individual demand as:

$$q_i = D(p_i, y)$$

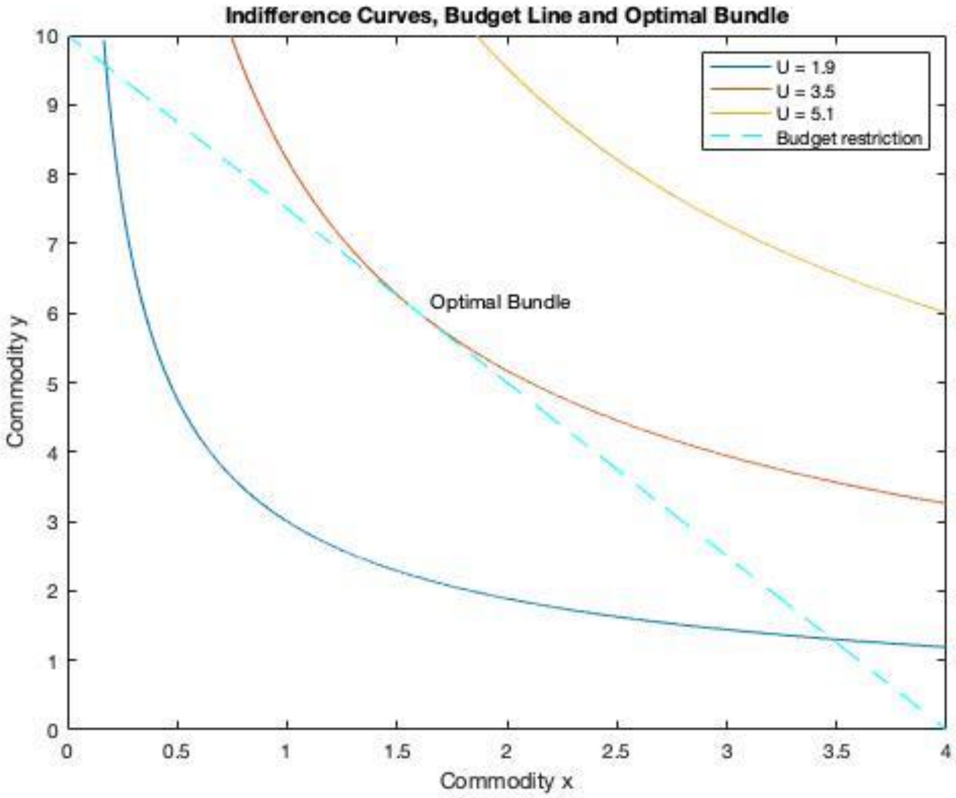
A function of demanding a product i incorporates not only the effect of its own price on the amount demanded but also the effect of the disposable income and the price of other products whose supply may affect the quantity of the goods i purchased. In Figure 1 a change in price of product i represents a movement along the curve while a change in the income or price of other substitutes represent a change or rotation of the demand curve.

The utility generated by consumption is described by the direct utility u_i function, which relates the level of utility of the goods purchased and is not observable. We know that not all consumption levels are achievable because of the budgetary constraint and because the consumer will choose that set of goods that maximizes the utility function.

The indirect utility function, $V_i\big(q_1, q_2, \dots, q_{n,}, y\big) = u_i\big(q_1(p_1, \dots, p_n, y), q_2(p_1, \dots, p_n, y)\big)$ where , V_i describes the maximum utility a consumer can obtain at any price and revenue level. The following result will be important when developing demand functions that we estimate:

For any indirect function of V_i, there is a direct U_i function that represents the same preferences on goods provided that the indirect function of profit, is continuous in prices and income, not rising in prices, not decreasing in income, quasi-convex in *(p,y)* and homogeneous grade 0 in *(p,y)*.

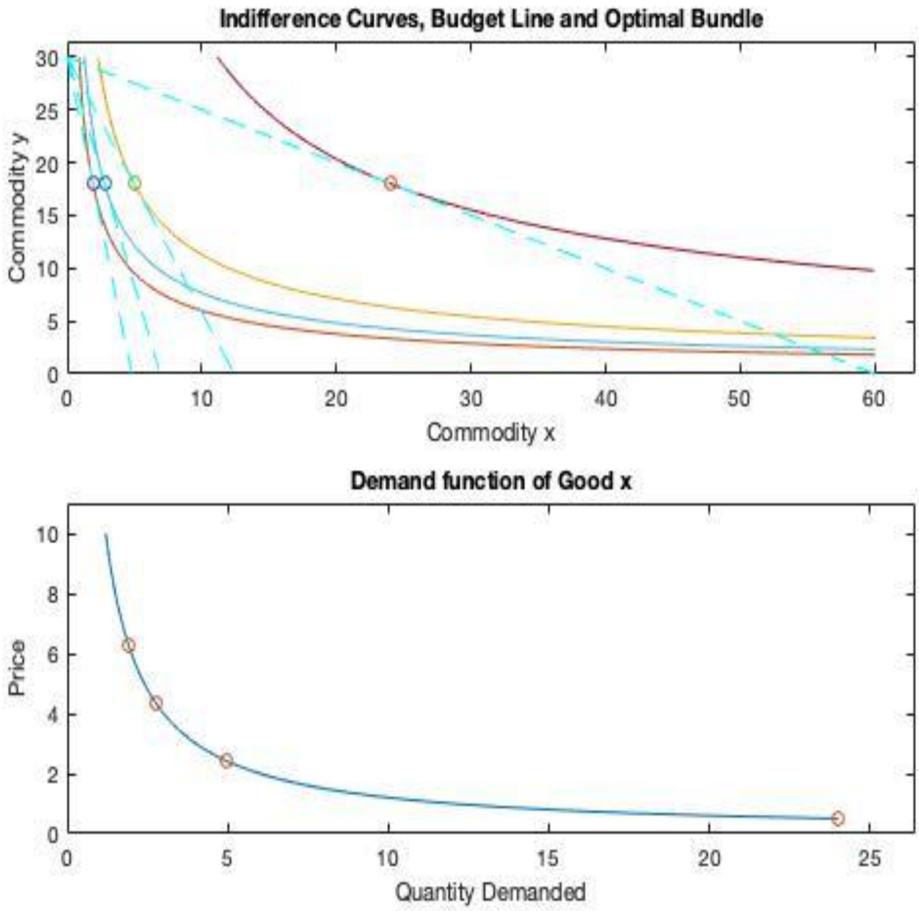
This result although it seems purely theoretical can actually be very useful in practice. In particular, it will allow you to return to our demand $q_i = D(p_i, y)$ curve without explicitly having to solve the utility maximization problem. We can see the methodology graphically, starting with the consumer problem, maximizing a utility function (in this case a Cobb Douglas subject to budgetary restriction):



The graph shows the consumer problem. Starting with a Cobb Douglas Utility Function: $x^{0.4}y^{0.6}$ with prices of 2 goods x and y, Px = 5 , Py = 2 , and an income of 20, the highest utility achieved with that income is the optimal bundle x = 1.6 , y = 6 with an utility of 3.54.

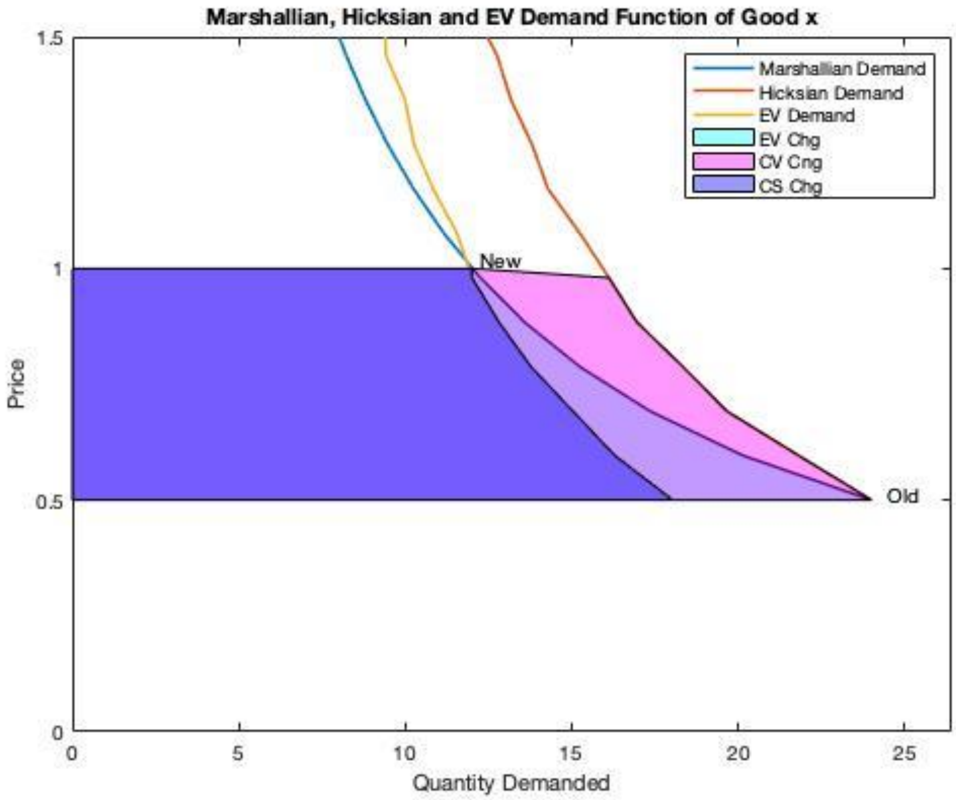
A price movement will change the quantity demanded, and so a demand function can be obtained, graphically[2]:

[2] Graphs built using Diagram Generator of Hang Qian (Iowa State University). Hang Qian (2020). Toolkit on Econometrics and Economics Teaching (https://www.mathworks.com/matlabcentral/fileexchange/32601-toolkit-on-econometrics-and-economics-teaching), MATLAB Central File Exchange. Retrieved December 28, 2020.



A demand curve is obtained as the consumer changes its optimal bundle of goods as prices changes. In this case, the starting position is Px = 0.5 , Py = 1 , Income = 30 with an initial bundle of x = 24 , y = 18 and an utility of 20.2. If the price of x increases to 1 with everything else constant Px = 1 , Py = 1 , Income = 30, the new bundle will be x = 12 , y = 18, with a lower utility of 15.3.

Finally the increase in prices reduces consumer surplus as shown in the graph below. Consumer surplus at initial bundle is 36 and at new bundle is 27.7 so the surplus reduction is 8.35. In the graph it is the area blue and magenta (trapezium with edges 0.5, 1, New, Old).



Demand Elasticity

In general, elasticity is a fundamental component of any competitive analysis as it provides a measurement of the consumer's response to a price increase. This parameter which represents demand sensibility to price changes is very relevant for firms when they set prices to maximize profits and play a central role in merger simulation.

The most useful measure of consumer sensitivity to changes in demand is demand-price elasticity[3], which measures demand sensitivity with respect to changes in prices of that good:

$$\eta_i = \frac{\%\Delta Q_j}{\%\Delta P_i} = \frac{\frac{\Delta Q_i}{Q_i}}{\frac{\Delta P_i}{P_i}}$$

Demand elasticity is the percentage change in quantity when prices increase in 1%. Marshall (1890) introduced the concept of elasticity and pointed out among his properties

[3] See USDA ERS database of demand elasticities-expenditure, income, own price, and cross price-for commodities and food products for over 100 countries in http://jaysonlusk.com/blog/2016/9/10/real-world-demand-curves.

that he had no unit of measurement, unlike prices measured in a currency or the quantities to be measured in one unit of quantity per period.

For very small variations, demand elasticity can be expressed as the slope of the demand curve multiplied by the price-to-quantity ratio. Mathematically it can be written as the derivative of the price logarithm with respect to the logarithm of the demand curve:

$$\eta_i = \frac{P_i}{Q_i}\frac{\partial Q}{\partial P} = \frac{\partial LogQ_i}{\partial Log\ P_i}$$

Demand will be elastic at one point when elasticity is greater than one at absolute value, when the change in quantity will be greater than the price increase so revenue for a seller falls if the other parameters are maintained unchanged. A demand is inelastic if its elasticity is less than one at a certain price, and means that the seller can increase their revenue by increasing prices if all other parameters remain constant. Elasticity generally depends on the price level. Therefore, it makes no sense to talk about a product with elastic or inelastic demand but rather it would have to be said that it has an elastic or inelastic demand at a certain price or volume of sales. The elasticities calculated for aggregate demand are market elasticities for a given product.

Substitute and complementary goods

The cross price-elasticity of demand (2009) shows the effect of a change in the prices of another *good k* on the demand for good *i*. A higher price of good *k* can induce some consumers to substitute purchases of good *k* for good *i*. In this case, when the consumer increases their purchases of *i* when p_k increases, we will call *i* and *k* products as substitutes. For example two earphones from two different brands are substitutes if the demand for one falls with the fall in the price of the other because consumers replace the expensive one with the cheap one. Similarly, a fall in the price of air tickets will reduce the demand for train travel, keeping the price of train tickets constant.

In conclusion, we typically want to estimate the effect of prices on the quantity demanded. To do this, one has to build a demand model in order to estimate the impact of the change in price on the quantity demanded, as we described above. An important aspect of the demand function is its curvature and how it changes when we move along the curve. The curvature of demand determines the elasticity and therefore the impact of a change in price on the quantity demanded.

The simplest possible specification is a linear demand function such as $Q = \alpha - \beta P$ or its inverse: $P = \frac{\alpha}{\beta} - \frac{1}{\beta} Q$ where α and β are model parameters. This is the demand specification often appears in economics texts[4].

Taking into account disposable income we get the following functional form:

$$y_t = \alpha - \beta p_t + \gamma y_t$$

One of the earliest empirical demand estimation of the relation between price and consumer's demand is the paper written by Working on demand of agricultural products (1925), and the first to recognize its endogeneity (1927), meaning that when predicting the quantity demanded in equilibrium, the price is endogenous as producers change their price in response to demand and consumers change their demand in response to price. What this initial papers recognize is that price and quantity are related by both supply and demand curves which shift in response to non-price variables (right figure). In order to build a demand function it is necessary to assume constant a wide variety of non-price variables such as consumer preferences, incomes, prices of other goods, etc (left figure)[5].

---

[4] The reader interested in the utility function that will result in a linear demand function will find it in http://www.its.caltech.edu/~kcborder/Notes/Demand4-Integrability.pdf

[5] See (2001): *"If the demand and supply curves shift over time, the observed data on quantities and prices reflect a set of equilibrium points on both curves. Consequently, an ordinary least squares regression of quantities on prices fails to identify—that is, trace out—either the supply or demand relationship. P.G. Wright (1928) confronted this issue in the seminal application of instrumental variables: estimating the elasticities of supply and demand for flaxseed, the source of linseed oil. Wright noted the difficulty of obtaining estimates of the elasticities of supply and demand from the relationship between price and quantity alone. He suggested (p. 312), however, that certain "curve shifters"—what we would now call instrumental variables—can be used to address the problem: "Such additional factors may be factors which (A) affect demand conditions without affecting cost conditions or which (B) affect cost conditions without affecting demand conditions." A variable he used for the demand curve shifter was the price of substitute goods, such as cottonseed, while a variable he used for the supply curve shifter was yield per acre, which can be thought of as primarily determined by the weather. …"*

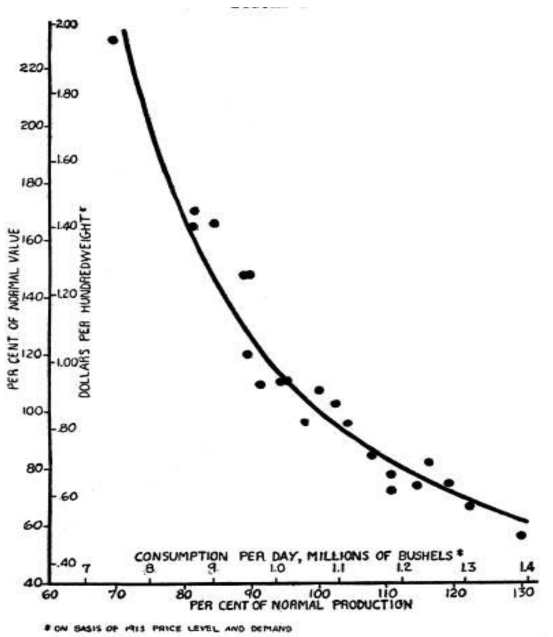


FIGURE 1. The Demand for Potatoes in the United States in Terms of Prices in St. Paul and Minneapolis

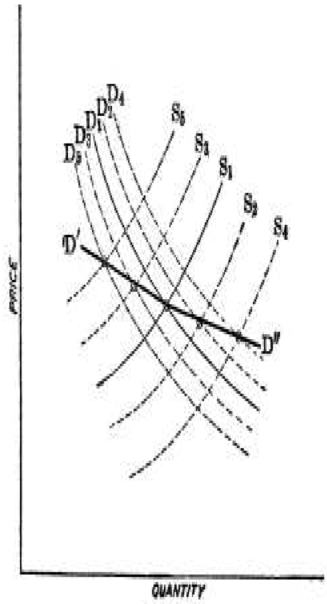


FIGURE IX

In the following section I introduce a structural model from the Industrial Organization literature which is a combination of a optimization agent problem with a statistical model. They solve some of the limitations of previous methodologies (endogeneity dimensionality, product differentiation, etc.) at the cost of greater mathematical and statistical complexity. We need to study them as they have become standard in industrial organization and antitrust economics.

3. Structural Demand Estimation: Discrete Choice Random Coefficient Logit Model Differentiated Products.

As it is explained above linear demand models are easy to work and explain small dimension consumer demand problem. The traditional approach was based in specifying a demand function for each product as a function of its own price, prices of other goods and other variables, (for example, the Rotterdam model or Almost Ideal Demand Model)[6]. Models like the above ones will produce one demand equation for each product, but if we take into account heterogeneity and differentiated products, the system of equations and so the number of parameters to estimate can become easily intractable.

Differentiated products add some difficulties that can make demand estimation intractable as parameters are a square of the number of products in the estimated market.

[6] See (2019), for an application of Almost Ideal Demand System model of Deaton and Muelbauer (1980))

If we consider for example 50 demand equations for 50 products, it will imply 2.500 parameters to estimate (50 demand equations with 50 prices each). Economists have had to deal also with consumer heterogeneity which the traditional representative consumer demand theory based (1976) on was unable to explain.

Since 1980´s, industrial organization authors developed other methodologies capable of dealing with limitations of the traditional approximation, one example being the discrete choice model approach (2000). Products are seen as bundles of characteristics and consumers have preferences over those characteristics. Demand function maps a product in a space of characteristics such as quality, accessories, brand other than price.

The discrete choice random coefficient model developed in 1995 for automobile market (1995) solved many of the previous limitations, has been extensively upgraded and applied in marketing and industrial organization[7] and has become a standard IO model (Bresnahan, 1987). This model only needs as inputs aggregate market level price and quantity data for each product and allows endogeneity[8] of prices.

Berry, Levisnton and Pakes (BLP model) estimation of automobile market is based on observed characteristics of a car such as horsepower, miles per dollar, size and air conditioning. Power over weight and miles per dollar are measures of power and fuel efficiency while air conditioning would be a measure of luxury (at that time), size is also a measure of safety.

An application of BLP is provided by Nevo (2000) on cereal market using general data available for products such as average prices, aggregate quantities, product characteristics. Product characteristics used by Nevo are calories, sodium and fiber content. As it is shown below the model is more difficult to understand but is considerably more realistic.

The first equation of the BLP model is the consumer utility function. The (indirect) utility of consumer *i* when buying product *j* in month *t* is given by equation 1:

$$u_{ijt} = \alpha_i \left(y_i - p_{jt} \right) + x_{jt}\beta_i + \xi_{jt} + \varepsilon_{ijt}$$

Where,

$$i = 1, \ldots, m \, , j = 1, \ldots, n, \, t = 1, \ldots, r$$

---

[7] See survey (2007).

[8] See Angrist and Krueger on instrumental variables to solve the endogeneity problem (2001) (prices correlated with econometric error term)

Where y_i is consumer i income, p_{jt} is the observed price of product j in month t, x_{jt} vector containing 6 characteristics of j products in month t, ξ_{jt} unobserved characteristics of product j in month t, and ε_{ijt} unobserved stochastic disturbance with mean 0. ξ_{jt} attempts to capture unobserved or unquantifiable attributes or characteristics such as brand name or promotional activity which are essential to explain the data (2000). ε_{ijt} is a stochastic term added due to the inability to explain individual preferences in a complete and deterministic way.

Parameters to estimate are now consumer i marginal utility to income, α_i, and marginal utility of each six characteristics, β_i.

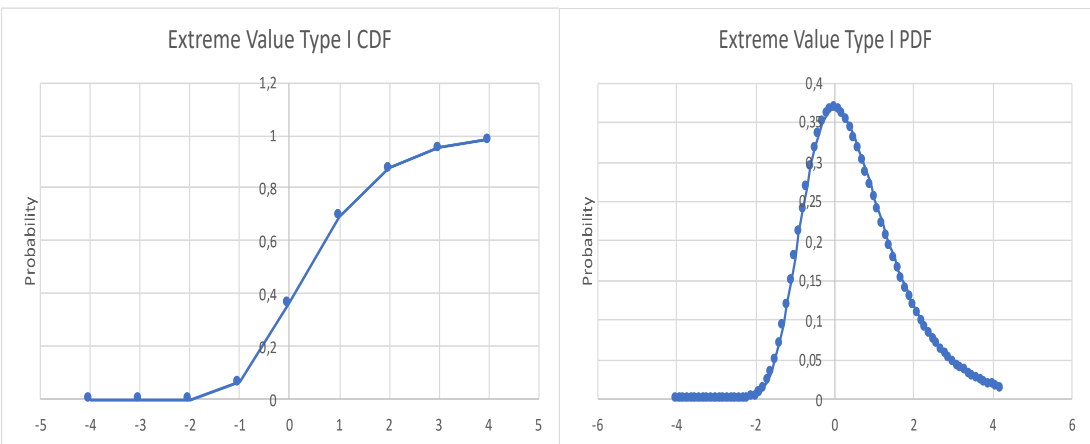
Further it is needed to make assumptions on the stochastic disturbance. Each distributional assumption will generate different models. BLP assume that ε_{ijt} follows a Gumbel or (Generalized Extreme Value Distribution Type-I). For a mean zero and scale parameter one it has the density and cumulative distribution

$$F(x) = e^{-e^{(-x)}}$$

and probability density function

$$f(x) = e^{-x}\, e^{-e^{-x}}$$

This is the limiting distribution of the maximum value of a series of draws of independent identically distributed random variables. Next Figure 2 illustrates the density (PDF) which is standardly used in logit models because its cumulative distribution (CDF) is related to the probability of x being larger than any other of a number of draws, which is like the random utility from one choice being higher than that from a number of other choices (2006).



Consumer choice is to buy j product in month t if it yield him the highest utility. We do not have data at consumer i level, we only have market shares of *j* product and only a sample of consumer characteristics but not i´s marginal utility β_i.

There are two ways to estimate equation 1, one is to use the simple multinomial logit regression if we assume all consumer have the same tastes or marginal utility for characteristics. Another way is random-coefficient logit when we let fall this assumption.

Multinomial logit estimation

We can use a simple (multinomial because there are six choices) logit if we proceed assuming that consumers have identical preferences and aggregate them:

$$u_{jt} = \alpha\left(y - p_{jt}\right) + x_{jt}\beta + \xi_{jt} + \varepsilon_{jt}$$

$$j = 1, \ldots, n, \; t = 1, \ldots, r$$

We now assume that ε_{jt} follows a Gumbel or Type I extreme value distribution (with mean zero and scale parameter one), so we conclude that we are building a multinomial logit model. We can obtain the market share of product *j* through the probability of j having the greatest utility, which occurs if ε_{jt} is high enough relative to other disturbances[9]:

$$s_{jt} = \frac{e^{x_{jt}\beta - \alpha p_{jt} + \xi_{jt}}}{1 + \sum_{k=1}^{n} e^{x_{jt}\beta - \alpha p_{jt} + \xi_{jt}}}$$

We can calculate the elasticities as the percentage change in the market share of product *j* when the price of product *k* goes up:

$$\eta_{jkt} = \frac{\%\Delta\, s_{jt}}{\%\Delta\, p_{kt}} = \frac{\delta s_{jt}}{\delta p_{kt}} \frac{p_{kt}}{s_{jt}} \begin{cases} \alpha p_{jt}\left(y_i - s_{jt}\right) & if\ j = k \\ \alpha p_{kt} s_{kt} & otherwise \end{cases}$$

[9] See (2006) pg. 12-13 for an explanation of this equation.

As Rasmussen points out this model is unrealistic in at least two points. The first one is that for small market shares, $\alpha p_{jt}(y_i - s_{jt})$ is close to α so the elasticity is close to -αp_{jt} . This implies that the model delivers low elasticity for low prices, and this implies higher markups for products with low marginal cost when it is often the other way round (higher markups for higher marginal costs e.g. luxury cars)

The second point is that cross price elasticities $\alpha p_{kt} s_{kt}$ only depend on price and market share of product *k.* This means that if there is an increase of price of k, the model predicts that consumer will equally likely substitute it for the other substitute products, (see Nevo for an explanation of the blue bus/red bus example).

Random Coefficients Logit Model

An alternative to multinomial logit is random coefficients model where it is assumed that parameters (or marginal utilities of product characteristics) are different across consumers and are determined by consumer characteristics which are a function of fixed parameters that multiply observed characteristics and unobserved random characteristics:

$$u_{ijt} = \alpha_i y_i + \delta_{jt} + \epsilon_{ijt} + u_{ijt}$$

Where, δ_{jt} is the mean utility which is a component of utility from a consumer´s choice of product j that is the same across all consumers.

$$\delta_{jt} = -\alpha p\text{-}p_{jt}\text{+}x_{jt}\beta + \xi_{jt}$$

There is the heteroskedastic disturbance, u_{ijt}, and homoscedastic disturbance i.i.d., ϵ_{ijt}.

$$\binom{\alpha_i}{\beta_i} = \binom{\alpha}{\beta} + \Pi D_i + \sum v_i\text{=}\binom{\alpha}{\beta} + \binom{\Pi_\alpha}{\Pi_\beta} D_i + \binom{\Sigma\,\alpha}{\Sigma\,\beta} \Sigma\left(v_{i\alpha}\middle|v_{i\beta}\right)$$

$\boldsymbol{D_i}$ is a vector of consumer i observable characteristics, $\boldsymbol{v_i}$, is a vector of consumer unobservable characteristics; **Π** matrix of how parameters depend on consumer observables, Σ represent how those parameters depend on observables.

We have seen that the model deals with dimensionality, consumer heterogeneity and product differentiation. Furthermore, it takes into account price endogeneity or simultaneity when estimating demand and supply functions. For this reason an instrumental variable estimation is used in the BLP model. For every endogenous explanatory variable one has to find another variable correlated with that variable but uncorrelated with the error term.

An application of this methodology is provided by Nevo in the ready-to-eat cereal market. The data consists of quantity and prices for 24 brands of a differentiated product in 47 cities over 2 quarters. Product characteristics are sugar and mushy and some demographic variables are added such as log of income, income squared, age, child. The results of this estimation are shown below (2000). In the first column appear the marginal utilities (β) and show that the average consumer shows more preference for soggy cereal but it decreases with age and income, while the mean price coefficient is negative being less sensitive for children and wealthier consumers.

Table 1: Results full Model (see Nevo (2000)

			Dem. Vbles.			
	Mean	**Std Dev.**	**Income**	**Income^2**	**Age**	**Child**
Constant	*-1,87*	*0,38*	*3,09*	*0,00*	*11.859,00*	*0,00*
	(-0,2571)	*(0,1295)*	*(1,1962)*	----	*(1,0056)*	----
Price	*-32,43*	*1,85*	*1,66*	*-0,66*	*0,00*	*11,62*
	(7,748)	*(1,0811)*	*(172,9296)*	*(8,9871)*	----	*(5,1713)*
Sugar	*0,14*	*0,00*	*-0,19*	*0,00*	*0,03*	*0,00*
	(0,2571)	*(0,0123)*	*(0,0451)*	----	*(0,0371)*	----
Mushy	*0,83*	*0,08*	*1,47*	*0,00*	*-1,51*	*0,00*
	(0,0129)	*(0,2073)*	*(0,6957)*	----	*(1,0905)*	----

GMM objective: 14.9007

MD R-squared: 0.26471

MD weighted R-squared: 0.095502

run time (minutes): 0.67607

BLP method is more flexible and requires weaker assumptions than older models and has been widely used as it allows for many different types of firm and consumer behaviour at the cost of greater complexity. The model corrects endogeneity of prices with instrumental variables, reduces the number of parameters from a considerably high number of products to five or six characteristics and is grounded on consumer theory. Other methods used in the model such as GMM (Generalized Method of Moments) for estimation or the contraction mapping used for optimization fall out of the scope of this book.

4. Machine Learning for demand estimation

We have explained the traditional models that estimate consumer demand. Firms nowadays predict consumer demand using techniques based on computer science. I describe the main models described in Bajari et al. (2015). These techniques are possible now thanks to the availability of large data sets, greater processing speed and computer efficiency, easy access to cloud computing together with the development of computer science and statistics. Bajari et al. show that some of the computer science models predict more accurately consumer demand than traditional econometric models and are preferable to instrumental variable estimation as it is difficult to find plausible instrumental variables for all models. Varian (2014) considers that the focus of Machine Learning is finding some function that provides a good prediction of y as a function of x, most notably good out-of-sample predictions, penalizing models that are excessively complex, separating data into training, testing and validation data, and tuning parameters to produce the best out-of-sample predictions. Bajari compares machine learning model performance with standard linear regression. The models analyzed are stepwise regression, where the choice of predictive variables is automated adding the ones with highest correlation with th residual. Support Vector Machines (SVM) finds a function with a deviation no greater that ε for each data value and as flat as possible. LASSO (Least Absolute Shrinkage and Selection Operator) regression estimates parameters that minimize the sum of squared residuals plus a penalty term that penalizes models of bigger size (2014). Finally, regression trees were the data is split at several points for each independent variable, until the squared prediction error (error between the predicted values and actual values squared) falls under a threshold. Bajari demonstrates, concerning the former machine learning models, that they can produce superior predictive accuracy, or lower root mean squared error (RMSE), as compared to standard linear regression or logit models.

5. Conclusion.

This chapter provides a review of demand estimation methodologies, from the traditional estimation, to modern machine learning methodologies. As we mentioned before traditional methods were ill designed to deal with endogeneity, consumer heterogeneity and product differentiation. The demand characteristic approach of structural discrete choice methodology is focused on product differentiation and has become the standard methodology. Recently, easy access to massive data and computation capacity have smoothed the path for methods based on machine learning techniques. Some of these techniques have been outlined in this chapter. When compared to traditional econometric models ML techniques provide better prediction accuracy.

Bibliography

Ackerberg, D., Benkard, C. L., Berry , S., & Pakes, A. (2007). Econometric tools for analyzing market outcomes. In *Handbook of econometrics* (pp. 4171-4276). Elsevier.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives, 15*(4), 69-85.

Bajari, P., Nekipelov, D., & Ryan, S. P. (2015). Machine Learning Methods for Demand Estimation. *American Economic Review, 105*(5), 481-85.

Belloni, A., Chernozhukov, V., & Hansen , C. (2014). High-Dimensional Methods and Inference on Structural and Treatment Effects. *Journal of Economic Perspectives, 28*(2), 1-23.

Berry, S., Levinston , J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica, 63*(4), 841-890.

Bresnahan, T. F. (1987). Competition and Collusion in the American Automobile Industry: The 1955 Price War. *The Journal of Industrial Economics, 35*(4), 457-482.

Davis, P., & Garcés, E. (2009). *Quantitative Techniques for Competition and Antitrust Analysis.* Princeton University Press.

Deaton, A., & Muellbauer, J. (1980). An Almost Ideal Demand System. *American Economic Review, 70*(3), 312-326.

Marshall, A. (1890). *Principles Of Economics; an Introductory Volume.* London: Macmillan.

Mas-Collell, A., Whinston, M. D., & Green, J. (1995). *Microeconomic Theory.* Oxford University Press.

Merino Troncoso, C. (2019). *Market power and welfare loss* . Retrieved from mpra.ub.uni-muenchen.de.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of economics & management strategy, 9*(4), 513-548.

Rasmusen, & Rasmusen, E. (2006). *The BLP Method of Demand Curve Estimation in Industrial Organization.* Retrieved from http://www.rasmusen.org/papers/blp-rasmusen.pdf

Spence, A. (1976). Product Selection, Fixed Costs, and Monopolistic Competition. *Review of Economic Studies, 43*(2), 217-235.

Varian, H. (2014). Big data:New tricks of Econometrics. *journal of Economic Perspectives, 28*(2), 3-28.

Working, E. (1927). What Do Statistical "Demand Curves" Show? *The Quarterly Journal of Economics*, 212-235.

Working, H. (1925, August 1925). The Statistical Determination of Demand Curves. *The Quarterly Journal of Economics, 39*(4), 503-543.