UNIVERSITY OF JOHANNESBURG

MASTER OF COMMERCE IN COMPETITION AND ECONOMIC REGULATION

Quantitative Methods and Econometrics for application in Competition and Economic Regulation (QEC9X01)

Prof. Lukasz Grzybowski

Lecture 4: Demand Estimation

"Quantitative Techniques for Competition and Antitrust Analysis"
by Peter Davies and Eliana Garces

Chapters: 1 and 9

Why spend time on demand systems?

Understanding of demand is necessary to address the following questions:

- Knowing own and cross-price elasticities of a product is critical for definition of relevant markets.
- Comparison of pre- and post-merger prices.
- Identification of conduct.
- Price regulation, e.g., implementation of Ramsey pricing.
- …

Homogenous vs. differentiated products

- Homogenous goods is a convenient assumption but frequently violated in practice.

- Many markets feature goods that are not identical varying in quality, features, geographic location, etc. → markets of literally identical goods seem to be relatively rare, especially once differences in seller's locations and reputations are taken into account.

- In markets with differentiated products we must think of a demand system that yields the demands for each product in the market.

What data is usually used?

In most applications data of the following structure will be used for demand estimation:

- Unit of observation will be a quantity or market share of a product purchased (e.g. Toyota Lexus) together with a price for a given time period (say a day) at a location (say a store).

- Data can be enhanced with many additional information:
  - Characteristics of the product, e.g., engine size, air-conditioning.
  - Distribution of consumer characteristics, e.g., age, income.

Different approaches to demand estimation

Approaches to differentiated products demand estimation can be divided into:

- Representative agent vs. heterogeneous agent.
- Product space vs. characteristics space.

Representative agent vs. heterogeneous agent models

- Both models assume that demand can be derived from utility function of consumers.

- Representative agent model assumes that all consumers behave equally.

- Heterogeneous agent models assume that consumers differ according to certain parameters (tastes, income, age, …) and make assumptions about the distribution of these parameters (aggregation of consumers).

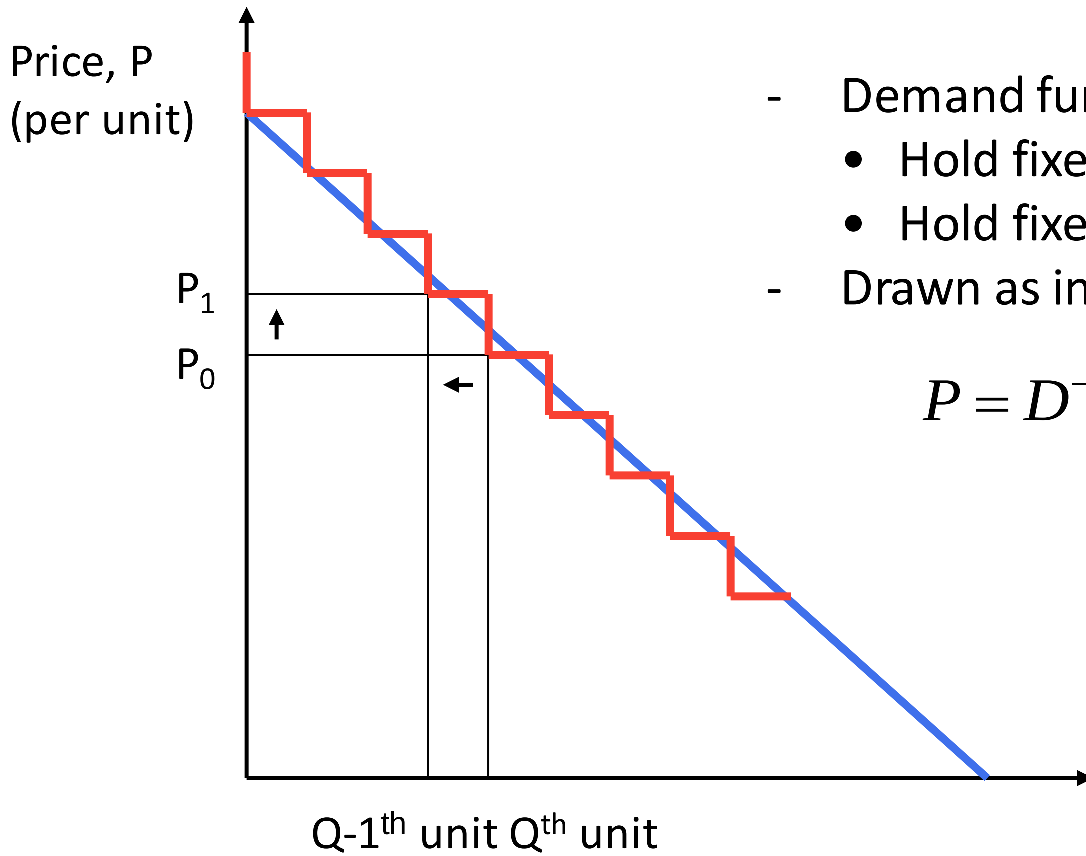
Product vs. characteristics space

- We can think of products as being:
  - a single fully integrated entity, e.g., Toyota Lexus V.
  - a collection of various characteristics, e.g., a 1500hp engine, four-wheel drive and the colour blue.

- What follows is that we can model consumers as having preferences over products or over characteristics.

- The first approach embodies the **product space** conception of goods, while the second embodies the **characteristic space** approach.

Market Demand Equations



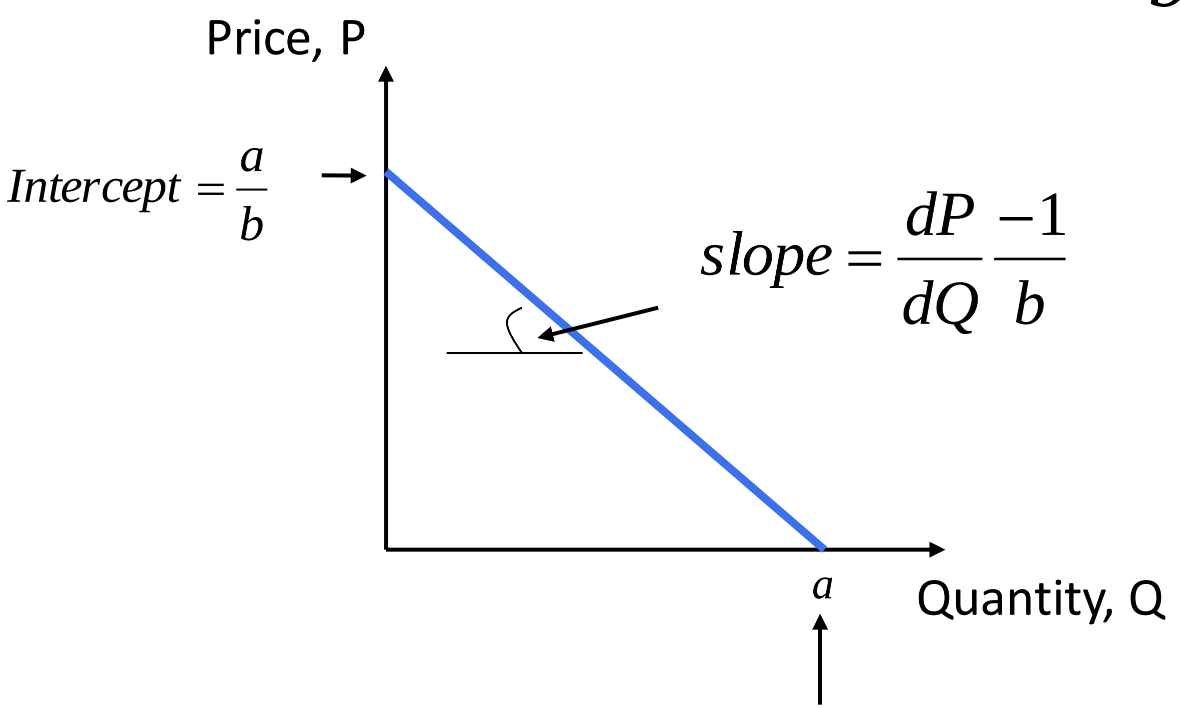
- Demand functions: $Q = D(P)$
  - Hold fixed prices of other goods
  - Hold fixed consumers incomes
- Drawn as inverse demand functions

$$P = D^{-1}(Q) = P(Q)$$

- The Demand function shows the number of units that will be sold at each price.
- As P increases from P_0, eventually one less unit is sold. P_1 is exactly the marginal value of the Q^{th} unit (to some consumer.)
- I.e., *The inverse demand curve describes consumer's marginal valuations.*

The Linear Demand Function

$$Q = a - bP, \quad \text{or} \quad P = \frac{a}{b} - \frac{1}{b}Q$$

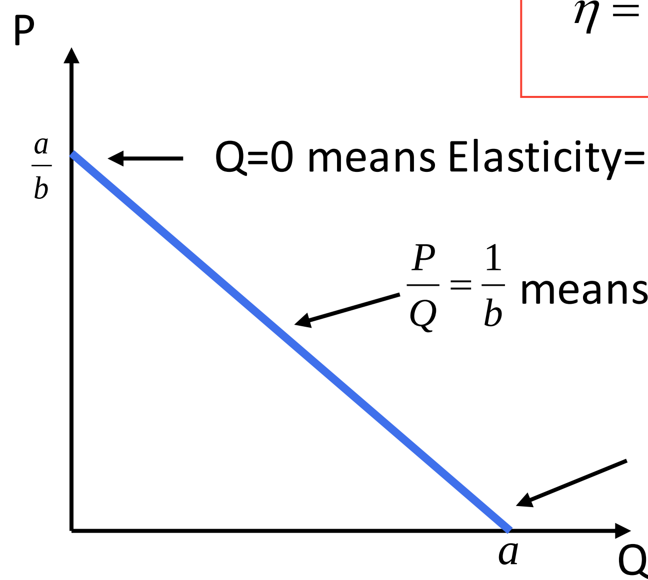


If P=0, then Q=a.
I.e., the linear demand model says you can’t even give more than ‘a’ units away!

Price Elasticities of Demand

- Recall Elasticities vary along linear demand curves since the Price Elasticity of Demand, η, is:

$$\boxed{\eta = \frac{d \ln Q}{d \ln P} = \frac{P}{Q}\frac{dQ}{dP} = \frac{P}{Q}(-b)}$$



Q=0 means Elasticity= $-\infty$

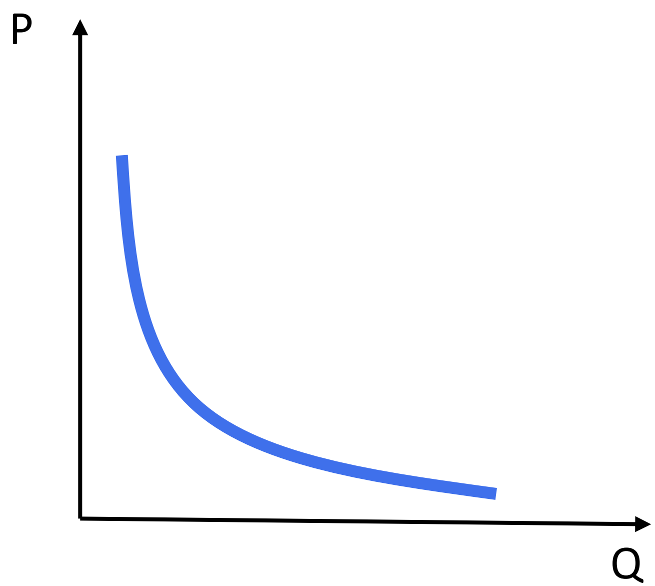
$\frac{P}{Q} = \frac{1}{b}$ means Elasticity = -1. (Recall, monopolists will always price above this point.)

P=0 means Elasticity = 0

- Elasticities increase in magnitude as we go up to top left hand corner. In practical terms:
  – It doesn't typically make sense to talk about a product having an 'elastic demand.' It can make sense to say it has an elastic demand – at current prices.
  – Many quantities we will want to measure (e.g., the profitability of a 5-10% price increase used for a SSNIP test) will depend on the magnitude of the elasticity *where we evaluate the price increase*

Log – Linear Demand Curves

- Demand: $$Q = D(P), \quad D(P) = e^{a} P^{-b}$$
- Taking natural logs: $$\ln Q = a - b \ln P$$
- Inverse Demand: $$P = \left(e^{=a} Q\right)^{-\frac{1}{b}} = e^{=a} Q^{-\frac{1}{b}}$$



Properties:

1. $$\lim_{P \to \infty} D(P) = \lim_{P \to \infty} P^{-b} = 0$$
2. $$\lim_{Q \to \infty} P(Q) = \lim_{Q \to \infty} Q^{-1/b} = 0$$
3. Constant elasticity (i.e., the exception)

$$\boxed{\eta = \frac{\partial \ln Q}{\partial \ln P} = -b}$$

Differentiated Product Demand Systems

- Consider the Two Product (hence two equation), linear, differentiated product demand system:

$$Q_1 = a_1 - b_{11}p_1 + b_{12}p_2 + c_1Y$$
$$Q_2 = a_2 + b_{21}p_1 - b_{22}p_2 + c_2Y$$

- Recall, we'll call good 2 a *substitute* for good 1 if:

$$\boxed{\frac{dQ_1}{dp_2} = b_{12} > 0}$$

- And we'll call good 2 a *complement* for good 1 if:

$$\boxed{\frac{dQ_1}{dp_2} = b_{12} < 0}$$

Differentiated Product Demand Systems

$$Q_1 = a_1 - b_{11}p_1 + b_{12}p_2 + \ldots + b_{1n}p_n + c_1Y$$
$$Q_2 = a_2 + b_{21}p_1 - b_{22}p_2 + \ldots + b_{2n}p_n + c_2Y$$
$$Q_3 = a_3 + b_{31}p_1 + b_{32}p_2 + \ldots + b_{3n}p_n + c_3Y$$
$$\ldots$$
$$Q_n = a_n + b_{n1}p_1 + b_{n2}p_2 + \ldots - b_{nn}p_n + c_nY$$

Symmetry Restrictions

- Recall from Micro 1 that *Slutsky Symmetry* says rational individual choice models will satisfy:

$$\boxed{\frac{\partial Q_1}{\partial p_2} + Q_1 \frac{\partial Q_1}{\partial Y} = \frac{\partial Q_2}{\partial p_1} + Q_2 \frac{\partial Q_2}{\partial Y}}$$

- In our particular demand system, sufficient conditions for Slutsky Symmetry will be

$$\frac{\partial Q_1}{\partial Y} = \frac{\partial Q_2}{\partial Y} = 0 \quad \text{and} \quad \frac{dQ_1}{dp_2} = \frac{dQ_2}{dp_1}$$

  – I.e., with linear demands if income effects are negligible $c_1 = c_2 = 0$, that means symmetry will be imposed by the parameter restriction:
  – This can be a very useful restriction from economic theory: $\boxed{b_{12} = b_{21}}$
    - It can mean we have fewer parameters to estimate
    - We require different data!
      – If $b_{12} = b_{21}$ we can learn about it by estimating the demand equation for good 1, $Q_1 = a_1 - b_{11}p_1 + b_{12}p_2$, i.e., data on (Q_1, p_1, p_2)
      – or alternatively by estimating the demand equation for good 2 which requires data on (Q_2, p_1, p_2)

Aggregate Demand and Symmetry Restrictions

- Suppose Coke currently sell 100 million units to 1 million customers per year whereas Virgin Cola sells 100,000 units to 10,000 customers.

- When Coke puts up its price by €0.10 then 1 million individuals will think about whether to switch some of their demand to Virgin Cola.

- But when Virgin Cola puts its price up by €0.10 then just 10,000 customers will think about whether they should switch to Coke!

- In each case, the people making the decision are different and there can be very different numbers of them.

- For each of these reasons, in general aggregate demand equations we won't expect to find symmetry and so: $\boxed{\dfrac{dQ_{Virgin}}{dp_{Coke}} \neq \dfrac{dQ_{Coke}}{dp_{Virgin}}}$

- And we will therefore often estimate: $\boxed{b_{12} \neq b_{21}}$

Product space approach: estimation issues

- In differentiated products markets there is typically a very large number of products, e.g., beers, cars, breakfast cereals, computers, etc.

- **Dimensionality:** with J products we already have J^2 parameters to estimate to get the cross-price effects alone → can be mitigated by grouping products where substitution within and across groups are treated differently, e.g., multilevel budgeting.

- **Consumer heterogeneity:** when consumers are heterogeneous, aggregate demand systems are only reasonable under strong assumption on preferences, e.g., Almost Ideal Demand System (AIDS).

Characteristics space: estimation issues

- Getting data on the relevant characteristics may be very hard as well as dealing with situations where many characteristics are relevant.

- This leads to the need for unobserved characteristics and various computational issues in dealing with them.

- Dealing choices of complements is a area of ongoing research.

Differentiated Products Models for Aggregate Data

Discrete Choice for Aggregate Data

General approach is as follows (see Berry (1994)):

1. Specify individual random utility, including the econometric error term.
2. Derive the individual choice probabilities.
3. Derive the aggregate market shares.
4. Equate this to observed aggregate market share.
5. Solve for the mean utility and hence the error term. This may be done either analytically or numerically.
6. Impose moment conditions (error term uncorrelated with instruments) to estimate the model.

Discrete choice models

- Discrete choice models describe consumer's choices among countable alternatives:
  - Alternatives must be mutually exclusive and the consumer chooses only one alternative from the choice set.
  - Choice set must be exhaustive (all possible alternatives are included).
  - The number of alternatives must be finite.

Multinomial Logit: Step 1

- Individual utility is:

$$U_{ij} = x_j \beta_i + \alpha p_j + \xi_j + \varepsilon_{ij} = \delta_j + \varepsilon_{ij}$$

- This means that individuals have the same valuation for the observed characteristics j.

- Individuals choose the product out of the J +1 products (including the outside good) that maximizes utility.

- ε_{ij} is assumed to be i.i.d. across products (and as usual individuals) and has a standard type I extreme value density function:

$$f(\varepsilon_{ij}) = \exp(-\exp(-\varepsilon_{ij}))$$

Multinomial Logit: Step 1

- Different assumptions about distribution of ε_{ij} lead to different discrete choice models.

- The scale of the utility is irrelevant because only differences in utility matters → the choices of each individual are invariant to (1) multiplication of utility by a person specific positive constant, and (2) addition to utility of any person specific number.

- A specific scale of the utility is frequently the result of some normalization of the distribution of the errors, e.g., the variance of $\varepsilon_{ij.}$

Multinomial Logit: Step 2

- Derive individual choice probability.

$$P_{ij} = \Pr(U_{ij} > U_{ik} \forall_{j \neq k})$$

- The resulting aggregate demand for product j takes the following form. This takes the standard logit form (see e.g. McFadden, 1978):

$$P_{ij}(\delta) = \frac{\exp(\delta_j)}{1 + \sum_{k=1}^{J} \exp(\delta_k)}$$

- where δ is the vector of mean utilities.

Multinomial Logit: Steps 3 & 4

- Step 3: Derive the aggregate market share function. Since all individual error has been integrated out, the choice probability function is also equal to the aggregate market share function:

$$P_{ij}(\delta) = s_j(\delta)$$

- Step 4: Equate this to the observed aggregate market share. Note that $s_j = q_j/N$, i.e., the market share is relative to the total number of consumers.

$$s_j = s_j(\delta)$$

Multinomial Logit: Step 5

- Solve for δ and hence the error term.

- The error term enters non-linearly. To make estimation feasible, model to make the error term enter linearly. Divide both terms by the market share of the outside good, and take logs to obtain:

$$s_j / s_0 = \frac{\exp(\delta_j)}{1+\sum_{k=1}^{J}\exp(\delta_k)} \frac{1+\sum_{k=1}^{J}\exp(\delta_k)}{\exp(\delta_0)} = \frac{\exp(\delta_j)}{\exp(\delta_0)} = \frac{\exp(\delta_j)}{\exp(0)}$$

$$s_j / s_0 = \exp(\delta_j)$$

$$\ln(s_j / s_0) = \delta_j = x_j\beta - \alpha p_j + \xi_j$$

Multinomial Logit: Step 6

- Impose moment conditions: main identification assumption is that characteristics other than price are uncorrelated with the error term.

- Price may be correlated so that OLS would lead to an estimate of biased towards zero or even have wrong sign)

- Additional instruments to identify the parameters:
  - cost side variables: often hard with product level data;
  - characteristics of competitors: see Berry, Levinsohn and Pakes (1995);
  - panel data also allow to use lagged variables or variables from other markets as instruments under suitable assumptions.

Multinomial Logit: Supply Side

- Though it is eventually of key interest in IO, paradoxically, in many applications no econometric analysis of the supply side is done.
- Many studies on market power with product differentiation make the following simplifying supply side assumptions:
  - constant marginal costs
  - static Bertrand-Nash equilibrium

Multinomial Logit: Supply Side

- These assumptions make it possible to limit econometric analysis to demand estimation. Based on the estimated own- and cross-price elasticities one can:
  - estimate market power (if marginal costs are known);
  - uncover marginal costs (if market conduct is known);
  - do policy counterfactuals.

- Cost estimation is thus only needed to:
  - learn about marginal cost parameters;
  - learn about conduct when marginal costs are unknown;

Multinomial Logit: Supply Side

- Suppose each firm f owns a set of products F_f and maximizes its profits by setting prices, given the prices set by other firms:

$$\max \Pi_f = \sum_{k \in F_f} (p_k - c_k) s_k(p) N$$

- The first order conditions for profit maximization are:

$$\sum_{k \in F_f} (p_k - c_k) \frac{\partial s_k(p)}{\partial p_j} + s_j(p) = 0$$

- This can be written in vector notation and inverted to yield the following system:

$$p - c = -\left(\theta^F \bullet [\nabla_p s(p)]'\right)^{-1} s(p)$$

- Markups are inversely proportional to a products 'perceived' price elasticity of demand, i.e., accounting for the fact that some of the lost sales after a price increase shift to other products in the firm's portfolio.

$$\max_f \Pi_f = \sum_{k \in F_f} (p_k - c_k) s_k(p) N$$

$$\sum_{k \in F_f} (p_k - c_k) \frac{\partial s_k(p)}{\partial p_j} + s_j(p) = 0$$

$$\begin{cases} (p_1 - c_1) \frac{\partial s_1(p)}{\partial p_1} \mathbb{1}_{11} + (p_2 - c_2) \frac{\partial s_2(p)}{\partial p_1} \mathbb{1}_{12} + \dots + (p_J - c_J) \frac{\partial s_J(p)}{\partial p_1} \mathbb{1}_{1J} + s_1(p) = 0 \\ (p_1 - c_1) \frac{\partial s_1(p)}{\partial p_2} \mathbb{1}_{21} + (p_2 - c_2) \frac{\partial s_2(p)}{\partial p_2} \mathbb{1}_{22} \dots + (p_J - c_J) \frac{\partial s_J(p)}{\partial p_2} \mathbb{1}_{2J} + s_2(p) = 0 \\ - - - - - - - - - - - \\ (p_1 - c_1) \frac{\partial s_1(p)}{\partial p_J} \mathbb{1}_{J1} + (p_2 - c_2) \frac{\partial s_2(p)}{\partial p_J} \mathbb{1}_{J2} + \dots + (p_J - c_J) \frac{\partial s_J(p)}{\partial p_J} \mathbb{1}_{JJ} + s_J(p) = 0 \end{cases}$$

$\mathbb{1}_{ij}$ indicator to ~~to~~ which firm the product belongs

$$\begin{bmatrix} (p_1 - c_1) \\ (p_2 - c_1) \\ \vdots \\ \vdots \\ (p_J - c_J) \end{bmatrix}' \cdot \begin{bmatrix} \mathbb{1}_{11} & \dots & \mathbb{1}_{1J} \\ \mathbb{1}_{21} & \dots & \mathbb{1}_{2J} \\ \vdots & & \\ \vdots & & \\ \mathbb{1}_{J1} & \dots & \mathbb{1}_{JJ} \end{bmatrix} \bullet \begin{bmatrix} \frac{\partial s_1}{\partial p_1} & \dots & \frac{\partial s_J}{\partial p_1} \\ \frac{\partial s_1}{\partial p_2} & \dots & \frac{\partial s_J}{\partial p_2} \\ \vdots & & \\ \frac{\partial s_1}{\partial p_J} & \dots & \frac{\partial s_J}{\partial p_J} \end{bmatrix}^{*} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ \vdots \\ s_J \end{bmatrix}$$

$$(p - c) \cdot \Theta^F \bullet \nabla_p s(p) = -s(p)$$

$$\boxed{(p - c) = -\left(\Theta^F \bullet \nabla_p s(p)\right)^{-1} s(p)}$$

Multinomial logit: derivatives

- Own price derivative:

$$
\begin{aligned}
\frac{\partial s_j}{\partial p_j} &= \frac{\partial\left(e^{\delta_j} / \sum_k e^{\delta_k}\right)}{\partial p_j} \\
&= \frac{e^{\delta_j}}{\sum_k e^{\delta_k}} \frac{\partial \delta_j}{\partial p_j} - \frac{e^{\delta_j}}{\left(\sum_k e^{\delta_k}\right)^2} e^{\delta_j} \frac{\partial \delta_j}{\partial p_j} \\
&= \frac{\partial \delta_j}{\partial p_j}\left(s_j - s_j^2\right) \\
&= \frac{\partial \delta_j}{\partial p_j} s_j\left(1 - s_j\right) \\
&= \alpha s_j\left(1 - s_j\right)
\end{aligned}
$$

Multinomial logit: derivatives

- Cross price derivative with respect to price of product m:

$$\frac{\partial s_j}{\partial p_m} = \frac{\partial\left(e^{\delta_j} / \sum_k e^{\delta_k}\right)}{\partial p_m}$$

$$= \frac{e^{\delta_j}}{\left(\sum_k e^{\delta_k}\right)^2} e^{\delta_m} \frac{\partial \delta_m}{\partial p_m}$$

$$= -\frac{\partial \delta_m}{\partial p_m} s_j s_m$$

$$= -\alpha s_j s_m$$

Multinomial Logit: Supply Side

- The own- and cross-price effects in the logit model are:

$$\frac{\partial s_j}{\partial p_j} = -\alpha s_j(1-s_j) \qquad \frac{\partial s_j}{\partial p_k} = \alpha s_k s_k$$

- This implies the following pricing equation for single-product firms:

$$(p_j - c_j)\frac{\partial s_j(p)}{\partial p_j} + s_j(p) = 0$$

$$-\alpha s_j(1-s_j)(p_j - c_j) + s_j = 0$$

$$p_j = c_j + \frac{1}{\alpha(1-s_j)}$$

- Can use this equation to uncover marginal cost, or estimate marginal cost function.

Multinomial logit: properties

$$P_j = \frac{\exp(\delta_j)}{1 + \sum_{k=1}^{J} \exp(\delta_k)}$$

- P_j is necessarily between 0 and 1.
- The sum of all choice probabilities adds up to 1.
- When δ_j rises, and other utilities are held constant, P_j approaches 1, and when δ_j decreases P_j approaches 0.
- The logit probability for an alternative is never exactly 0.
- A probability of exactly 1 is obtained only if the choice set consists of a single alternative.

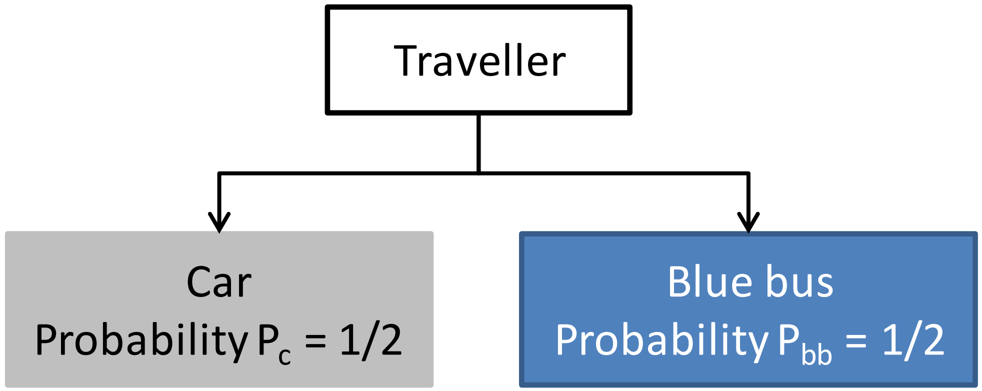
Multinomial logit: issues

1. Logit can represent systematic taste variation (related to observed characteristics of the decision maker) but not random taste variation (linked to unobserved characteristics).

2. The logit model implies proportional substitution across alternatives → to capture more flexible forms of substitution, other models are needed.

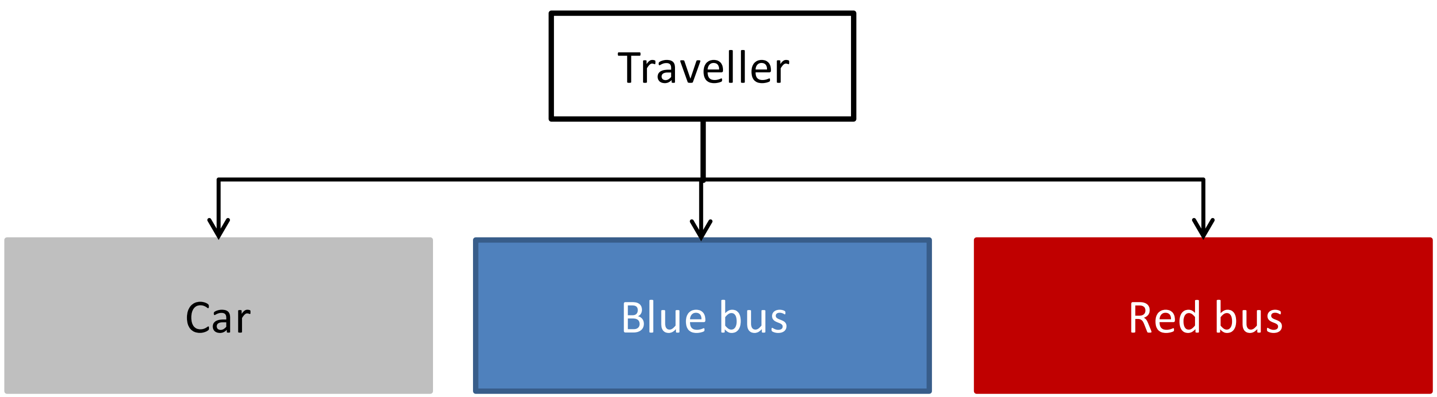
3. If unobserved factors are independent over time in repeated choice situations, then logit can capture the dynamics of repeated choice, including state-dependence. However, logit cannot handle situations where unobserved factors are correlated over time.

Multinomial logit: red bus, blue bus problem



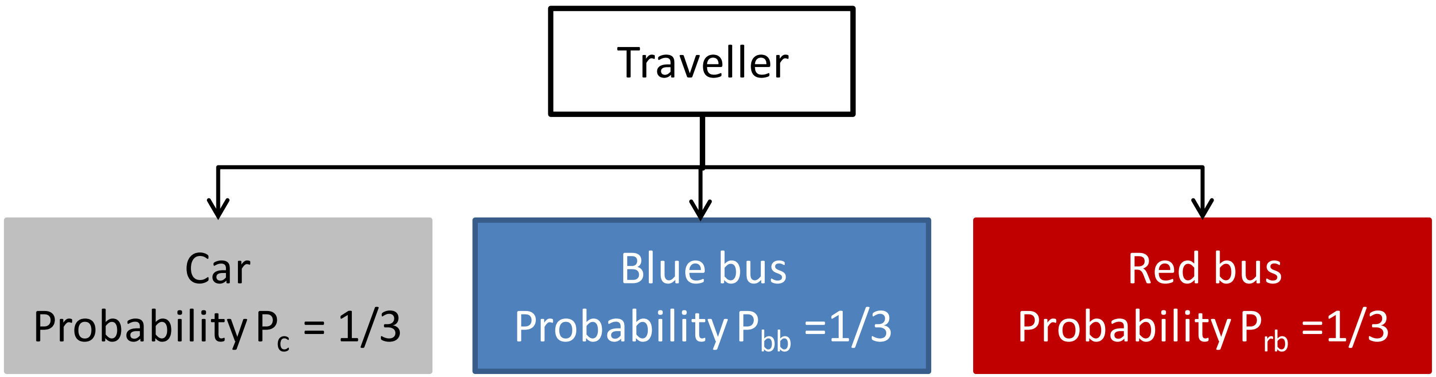
The ratio of probabilities is: $P_c / P_{bb} = 1$.

Now a red bus alternative is introduced which is equivalent to blue bus but for the color. The ratio of their probabilities is one: $P_{rb} / P_{bb} = 1$.



Multinomial logit: red bus, blue bus problem

- In the logit model the ratio P_c / P_{bb} is the same whether or not another alternative, in this case the red bus, exists. This ratio therefore remains at one.
- The only probabilities for which P_c / P_{bb} = 1 and P_{rb} / P_{bb} = 1 are P_c = P_{bb} = P_{rb} = 1/3, which are the probabilities that the logit model predicts.



- In real life, however, we would expect the probability of taking a car to remain the same when a new bus is introduced that is exactly the same as the old bus.
- We would also expect the original probability of taking bus to be split between the two buses after the second one is introduced. That is, we would expect P_c = 1/2 and P_{bb} = P_{rb} = 1/4.

Multinomial logit: issues with elasticities

- For the own and cross price elasticities we get:

$$\eta_{jm} = \frac{\partial s_j}{\partial p_m} \frac{p_m}{s_j} = \begin{cases} -\alpha p_j (1 - s_j) \text{ if } j = m \\ \alpha p_m s_m \text{ otherwise} \end{cases}$$

- Own-price elasticities are proportional to own price: the lower the price the lower the elasticity, which implies higher markups for the lower priced goods.
- Cross-price elasticities between any pair of products are entirely determined by one parameter and the market share and price of that good: consumers substitute towards other brands in proportion to market shares, regardless of characteristics (also small s_m means small elasticity).

Multinomial logit: issues with elasticities

- Example: If the price of a Lexus (price=40k, market share=.05) goes up, then the impact on demand for BMW (price=55k, market share=.01) and Yugo (price=8k, market share=.01) are the same! Our elasticities are determined by the structure of the model and not the data!

	s_1	s_2	s_3
s_1	-76	1.1	0.16
s_2	4	-108.9	0.16
s_3	4	1.1	-15.84

- Solution: relax the iid assumption, such that elasticities depend on how close products are in the characteristics space → mixed logit.

Other discrete choice models

- Different discrete choice models are obtained from different assumptions about the distribution of the unobserved portion of utility ε.

- The integral takes a closed-form only for certain specifications of $f(\varepsilon)$ → logit and nested logit have closed form expressions.

- They are derived under the assumption that the unobserved portion of utility ε is distributed iid type I extreme value and generalized extreme value, respectively.