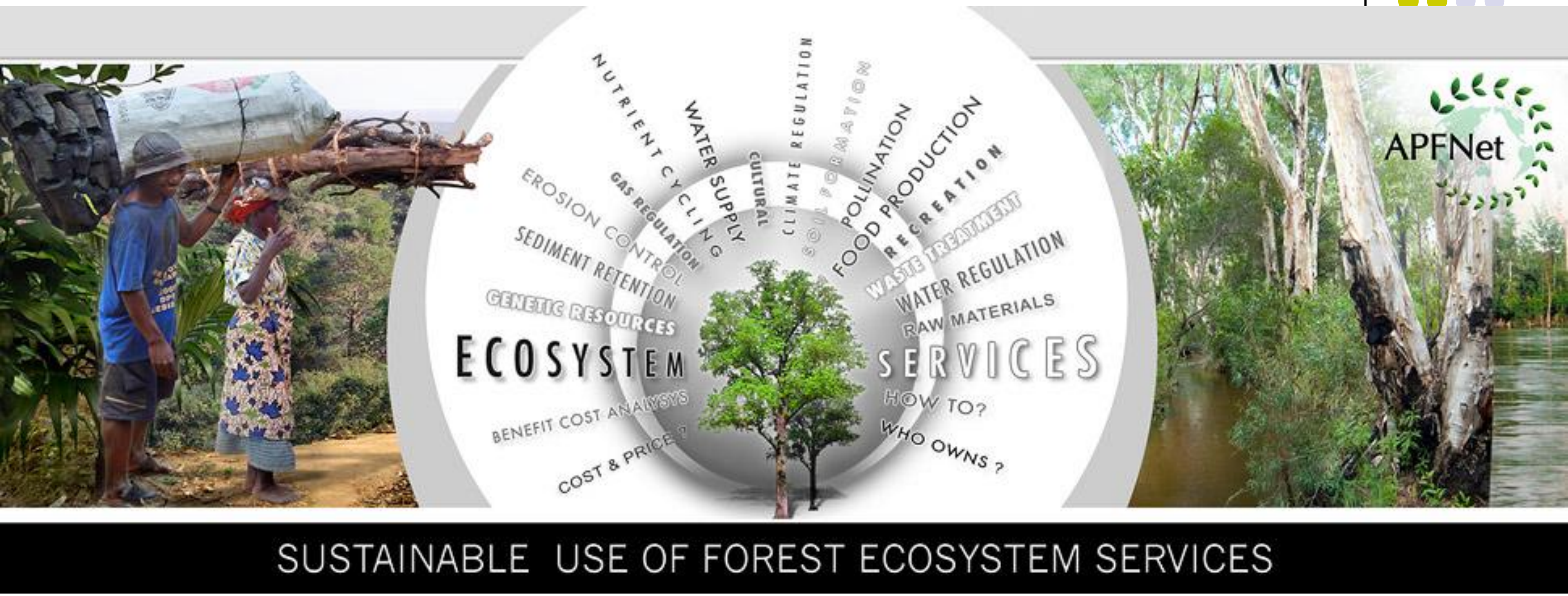


DISCRETE CHOICE EXPERIMENT (DCE)

Synopsis

Discrete choice modeling or experiment (DCM/DCE) is the standard method of analyzing individual choice behavior and market demand. Applications include transportation mode, brand choice, recreation demand, natural resource conservation, telecommunications services, health services, appliance choice, location decisions and a wide variety of other settings in many diverse fields.

Recent advances in tools and methods have been used to model individual behavior and to analyze people choice in natural resource conservation in response to various attributes or characteristics of a given area in available choice sets and choice characteristics.

Outline

- Background of Choice Model
- Theoretical Framework - RUM
- Approach of Choice Model
- Common Design Stage
- Data Analysis

What is Discrete Choice Model

- Origin – Conjoint Analysis
- A broader term, which involves the use of other statistical techniques (ANOVA tabels). Is not necessary grounded in random utility theory.
- From the mid 1990s: increasing interest in this preference based technique within environmental and health economics
- Renamed: **discrete choice experiments (DEC)or discrete choice modelling (DCM)**

What is Discrete Choice Model

- DCM – encompasses a range of SP techniques
- Take a similar approach to value environmental benefits
- Include the followings:
  - Choice Experiment
  - Contingent Ranking
  - Contingent Rating
  - Paired Comparison

What CM tells us

- Which attributes are significant determinants of values of people preference
- The implied ranking of these attributes among relevant population(s)
- The value of changing more than one of the attributes at once
- Total economic value of resource use

Choice Modelling Approaches

- Choice Experiments
- Contingent Ranking
- Contingent Rating
- Paired comparison

Choice Experiments

- Choice between two alternatives vs. status quo (base line)
- Presented with a series of alternatives
- Welfare consistent – YES
  - **Force respondents to trade-off**
  - **Respondents can opt for status quo**
  - **Can use econometric technique parallel to rational, probabilistic choice (random utility model)**
  - **Can derive estimate of compensating and equivalent surplus**

Example: Choice Experiment (Mangrove Forest)

Attributes	Option 1 (Satus Quo)	Option 2	Option 3
Forest Area	Decrease by 14%	No change	Increase by 24%
Direct Employment	Increase by 3%	No change	Increase by 3%
No of Migratory Birds	Decline by 3%	No change	No Change
Visitation Rates	No change	Increase by 5%	Increase by 50%
Annual Contribution	**RM0**	**RM50**	**RM10**

Prefer 1 ❐ **Prefer 2 ❐** **Prefer 3 ❐**

Contingent Ranking

- Rank a series of tasks / alternative options
- Welfare consistent estimate – depends
- Each alternative is characterized by a number of attributes, offered at different levels across options
- Respondents are asked to rank the options according to their preference
- One option – must be in the agent’s currently feasible choice set
- This can be done - include “do noting’ option – to interpret the results in standard welfare economic terms

Example: Choice Experiment (Mangrove Forest)

Attributes	Option 1 (Satus Quo)	Option 2	Option 3
Forest Area	Decrease by 14%	No change	Increase by 24%
Direct Employment	Increase by 3%	No change	Increase by 3%
No of Migratory Birds	Decline by 3%	No change	No Change
Visitation Rates	No change	Increase by 5%	Increase by 50%
Annual Contribution	**US$0**	**US$50**	**US$10**

Your Ranking: 1 ___ 2 ___ 3 ___

Contingent Rating

- Score alternative scenarios on a scale of 1-10
- Welfare consistent estimate – doubful
- Presented a number of scenarios one at a time and are asked to rate one individually on a semantic or numeric scale
- Use a series of questions
- **Data collected are rating scores**

Contingent Rating - Example

Characteristics	Option1
Native woodland	*500 ha protected*
Heather moorland	*1200 ha protected*
Lowland by meadow	*200 ha protected*
Cost per household per yr in additional tax	*US$50*

Please tick one box only

1	2	3	4	5	6	7	8	9	10

Very Low Preference

Very High Preference

Contingent Rating - Example

Characteristics	Option 1	Option 2
Native woodland	*500 ha protected*	*700 ha protected*
Heather moorland	*1200 ha protected*	*No protection*
Lowland by meadow	*200 ha protected*	*300 ha protected*
Cost per household per yr in additional tax	*RM50*	*RM20*

Please tick one box only

1	2	3	4	5	6	7	8	9	10

Strongly Prefer Choice A

Strongly Prefer Choice B

Paired Comparisons

- Score pairs of scenarios on similar scale
- Welfare consistent estimate – doubtful
- Respondents are asked to choose their preference alternative out of two choices and indicate the strength of their preference in numeric or semantic scale
- Combines elements of CE and rating
- Can use computer to generate choice sets within each CM approach

Is DCE basically different from CVM?

- No
- If the contingent valuation is based on closed – ended payment bids- then it is essentially a variant of DCE
- Open-ended CVM is different-because is more deterministic in nature, whereas open-ended CVM and DCE are based on probabilistics

A probabilistic utility function

- As the analyst we realise that we cannot define the utility function precisely-but we can disclose (some) factors which influence individuals’ decisions.
- $Y=h(x,\varepsilon)$. If we know all explanatory components in a utility function it is in essence deterministic.
- But ε is in random utility theory defined as the unobserved component; in realisation that all behavior cannot be explained and/or we cannot identify all explanatory variables.

Theoretical Framework

The “link” between ε, V and p

- $U=V+\varepsilon$
- We estimate V
- If the respondent prefers alternative 1 to alternative 2, we assume: $U_1>U_2$
- But because of “randomness” (measured by ε) all respondents will not choose alternative 1 each time- the same respondent may also choose differently if asked a consecutive number of times
- We observe-the probability of choice, and on the basis of this observation we estimate V

Why “randomness”?

- Omitted explanatory variables
- Preferences vary across respondents due to heterogeneity in taste. We establish an “average preference function” on the basis of groups of respondents choices. Although we may perform sub-group analyses-some heterogeneity will remain.
- Individual are not fully rational-due to limited cognitive abilities they reduce the burden of choice by use of “rules of thumb”. The choices are a result of “bounded rationality

The “link” between ε, V and p

- The probability that respondent n chooses alternative i to alternative

  $P=Prob(V_{ni}+ \varepsilon_{ni}>V_{nj}+ \varepsilon_{nj})$

- When $\varepsilon_{nj}- \varepsilon_{ni}<_{nj}$ alternative i will be chosen

Logit/probit

- One can assume different distributions of the unobserved conponent ε in the utility function
- Logit: as assumes an extreme value distribution (a gumbel distribution). It is symmetric, but more flat than the normal distribution. Probit: assumes a normal distribution
- Logit (binary or multinomial) often applied because it can be handled mathematically (and not by similation as probit). It does however entail som restrictive assumptions-e.g. it assumes the independence of irrelevant anternatives (IIA).

Random utility theory

$$\partial V = V_{ni} - V_{nj}$$

- If we observe that 50% choose alternative i, what can say about dV?
- If more than 50% choose i?
- If less than 50% choose i?
- If ε is assumed to follow a gumbel distribution (logit):

$$\partial V = \ln\left(\frac{p}{1-p}\right) \text{ or } \quad p = \frac{e^{\partial V}}{(1+e^{\partial V})}$$

Utility measured on a cardinal scale

- P is a function of V_i-V_j and ε. Assuming a certain distribution of ε, $dV=V_i-V_j$ is derived on the basis of p. The greater the probability of choosing V_i rather than V_j-the greater the difference in utility between the two options.
- Since p is defined on a cardinal scale-so is the difference $\partial V = V_i - V_j$

The discrete choice model

- The basic model which is typically applied is the linear additive model:
- $V=\alpha+ \beta$ *attribute1+ β*attribute2 +γ*price
- Additivity: assumes no interaction between the utility derived from one attribute and the score on another attribute
- Linearity: we assume constant marginal utility on all variables (including the cost variable)
- One can in principle construct non-linear and multiplicative models-but the design must be constructed accordingly

Importance of design(I)

- * Design refers to the planning of which observations to make and the best to provide possible inference from data

Importance of design(II)

- Properties of design-statistical properties

  -Avoid collinearities and non-estimable

  confounding affects

  -Efficiently capture of the relevant effects

  -The design determines the degree of

  precision obtainable in the estimations

Importance of design(III)

- Properties of design-the respondents' ability to understand the task

  -Realism and relevance to the respondents

  -Realism of the tradeoffs

  -Is the task understandable?

  -Does the exercise contain the attributes relevant to the respondent?

  -The cognitive challenge for the respondents

Step in a choice experiment

1. Description of the decision problem
2. Attribute and level selection
3. Experimental design development
4. Questionnaire development
5. Sampling and data collection
6. Estimation procedures
7. Policy analysis and welfare calculation

Description of the decision problem

- Identify the problem
- List and rank the problem
- Consider the problem in analysis

Eg: conservation vs development

Attribute and level selection

Terminology

- Attributes

  -Factors that define the choice possibilities

  -Service attributes:the attributes describing the gains

  -Cost/price/time attribute(s) describing alternative costs

  -Denoted $x_{i,}$ i={i,….I}

- Attribute levels

  -The values that each attribute can take

  -Denoted $x_{ij,}$ i={1,2….j}

  -E.g price {0,50,100,200}

Identification of attributes

- Survey development and administration process

Focus groups → -Determine decision frame,attributes and level

Pretest → -Assess responses/cognition

Pilot test → -Assess survey administration

Data collection → -What are the answer

Identification of attributes

- Identification of attributes

  -Relevance to the decision problem and policy maker(s)

  -Relevance to the respondent- must have meaning to people who answer the exercise

  -Response rate

  -Taking the exercise serious

Identification of attributes

- Sources to identify attributes
  - **Literature review**
  - **Attributes defined in theories**
  - **Expert opinion**
  - **Single and focus group interviews with potential respondents**
  - **Attributes define by policy maker**
  - **Previous questionnaires**

Identification of attributes

- Selection of attributes
  - Policy needs
  - ‘Importance’ of attributes
  - Theoretical relevance
  - Causally prior attributes should be avoided

    E.g is pollution important or is it merely a stepping stone to achieving what is valued

- Mutual dependence between attributes should be avoided

  E.g . Distance to work and possibility to have a home office

- The importance of the cost attribute

Identification of attributes

- Selection of attributes

  -Omitted attributes- cause biased information or heterogeneity in data because of difference in the respondents' assumptions

Identification of attributes

- Number of attributes

  - Restricted by design problems and respondents' cognitive capability

  -Pragmatic suggestions of maximum number: 6-8

Identification of attributes

- Attribute levels

  -Qualitative

  e.g. Colour, Brand, site, presence of a facility or not etc

-Quantitative in absolute terms or changes from the current situation

  e.g. number of birds, risk reduction, reduction in number of particles, life expectancy etc.

Identification of attributes

- Attribute level

  -Realism of the range of levels

  -Range of the levels enabling compensatory decision making-the loss in one bundle ($x_1, x_2 \ldots . x_N$) can be compensated by giving more of at least one attribute($x'_1, x'_2 \ldots . x'_N$)

  -The range of the levels of the cost attribute is of particular importance

Identification of attributes

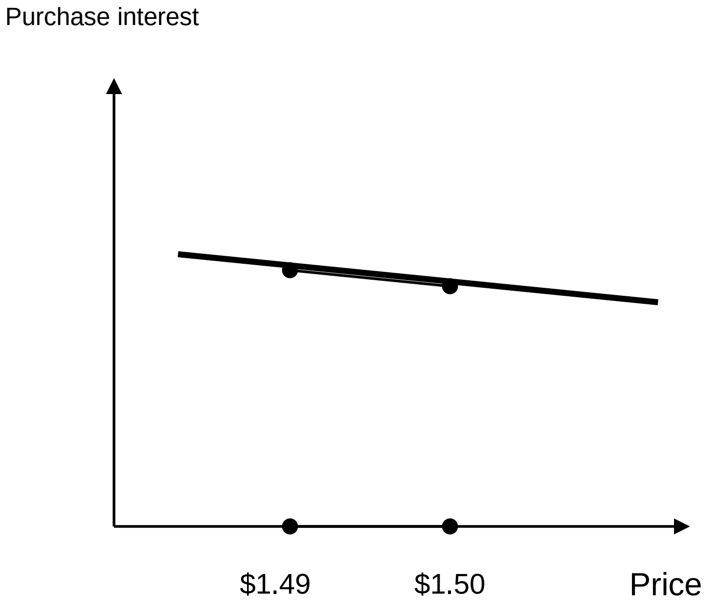
- The range of the cost attribute

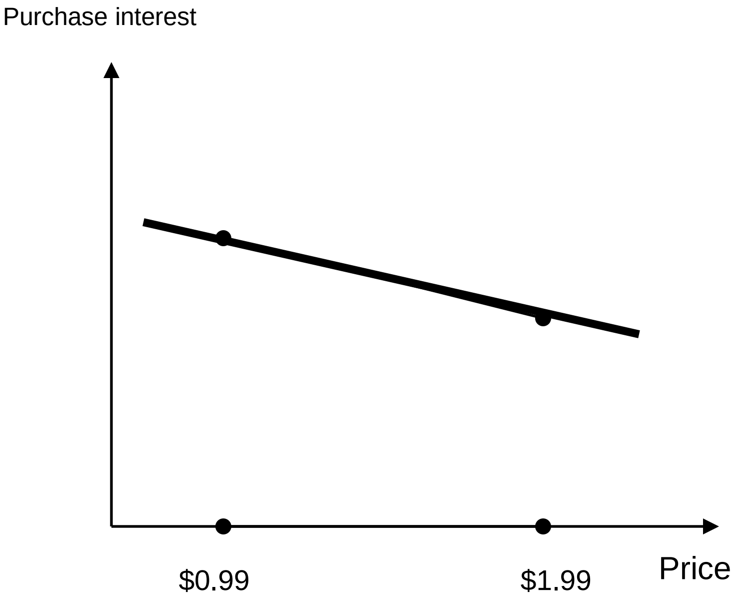
  -Too low range makes respondents insensitive to costs(price effect is minimal)-may overestimates WTP

  -Too large costs makes respondents dominant or they opt-out-may underestimate WTP

Identification of attributes

- The range of cost attribute- illustration





Identification of attributes

- Attributes levels

  -'Completely unaceptable levels' meaning that respondent reject an alternative no matter ho attractive an alternative might be

  -If respondents are consistent in their adhererence to completely unaceptable levels, it may be reasonable to remove this level

  -If not, the specification of the respondents' utility model may be compromised

  -Study results- respondents are paying attention to completely unaceptable levels - but in a more compensatory way

Identification of attributes

- Attributes levels
  - Variations in a attribute levels(but not range) influence sensitivity of WTP estimates
  - Increased number of attribute levels seems to increase the relative importance of these particular attributes. The proportion exhibiting dominant preference is similar
  - “Range differences in our study did not appear to affect, or at least had little effect on (5) logistic regression model parameters…; or (7) error variance

Alternative or Option

- <u>Alternatives</u>

  Combination of attribute levels

  Alternative I = {x_{1j} , x_{2j} , x_{3j}}

  E.g . alternative I ={red,160km/h,100.000},

  2 ={blue,175km/h,120.000},

- <u>sets</u>

  Combinations of alternatives

Attribute	**Alternative A**	**Alternative B**	**Alternative C**
X_1	**Red**	**Blue**	**Red**
X_2	**160km/h**	**175km/h**	**190km/h**
x_3	**100.000**	**120.000**	**140.00**

- Alternatives
  - Combination of attribute levels
  - Alternative I = $\{x_{1j},\ x_{2j}, x_{3j}\}$
  - E.g. alternative I ={red,160km/h,100.000},
  - 2 ={blue,175km/h,120.000},
- sets
  - Combinations of alternatives

Design Alternatives

Atribute	Status Quo	Proposed Alternative 1	Proposed Alternative 2
No of endangered species			
Visitor day per annum			
Cost to you			

Alternatives must be

- Mutually exclusive
  - If confusion arises, choice categories – A and B only
- Choice set must be exhaustive
  - If presenting choice scenario – heating in house (gas , electricity, wood) should we include no heating? Or Exclude respondents who do not have heating
- No of alternative must be finite
  - No of levels on attributes must be finite
  - Ex: 0, 1, or 2 more cars in he houshold

Alternatives: queries

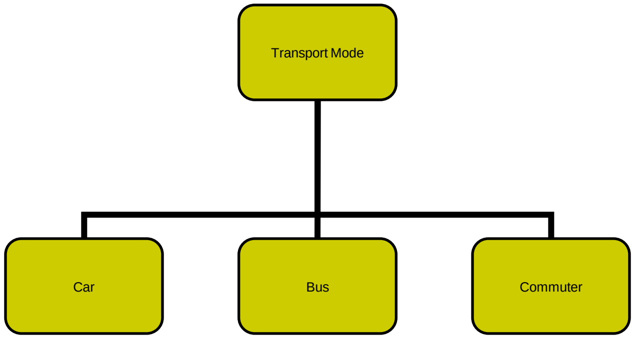
- Which type of alternatives to include?
  - Status Quo
  - Opt-out (non of the presented alternatives) – should include when relevant
- No of alternatives presented in choice scenario: 2, 3 or 4? (Health economics – 2, environment – 3)
- Scenario should be HOW and WHERE, Not WHETHER or NOT.

Number of alternatives

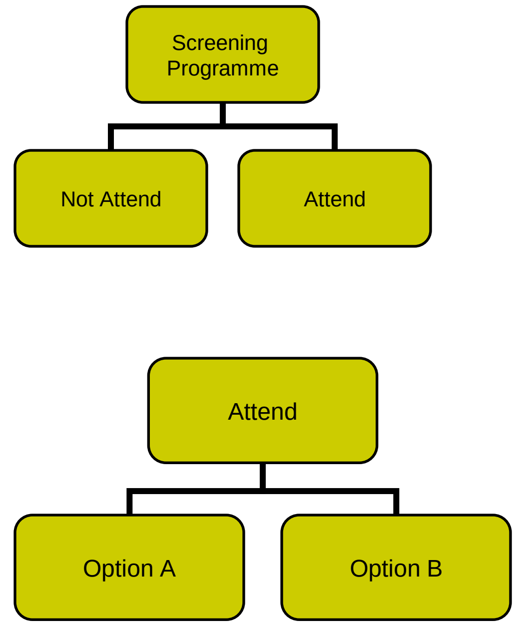
- If more than 2 alternatives: multinomial logit
- If 2 alternatives: binary logit
- Two models have underlying characteristics
- Multinomial logit (but not logit) requires that Independence of Irrelevant Alternatives Criterion (IIA) is fulfilled
- IIA – unobserved factors are uncorrelated over alternatives
- Defn IIA: Ratio of probabilities choice between two alternatives is unaffected by the introduction of a third alternative

Preference Structure

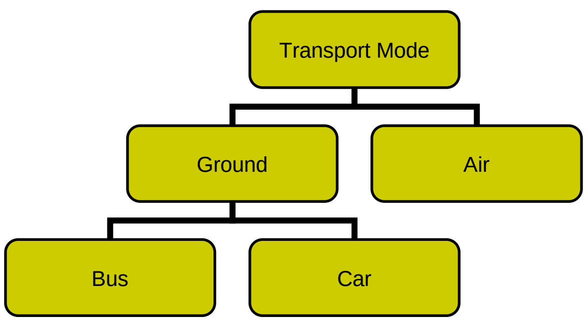
Multinomial Logit



Separate Binary Logit



Nested Logit



Number of alternatives

- If more than 2 alternatives: multinomial logit
- If 2 alternatives: binary logit
- The two models have different underlying characteristics
- Multinomial logit (but not logit) requires that the Independence of Irrelevant Alternatives criterion is fulfilled

Questionnaire formats

- Conjoint analysis (umbrella term)
  - Discrete choice experiments
  - Contingent ranking
  - Contingent rating
  - Paired comparison

Questionnaire formats

- Discrete choice experiments

  -Two or more alternatives in a choice set

  - More alternatives in a choice set provides more information but may be a cognitive burden
  - Respondent should pick one-the alternative which provides the highest utility

A DCE choice scenario

	Treatment A	Treatment B
Number of treatment per day	Two per day	Two per day
Number of products in use per day	Two products	One product
Improvement- non visual effect	Little improvement	Big improvement
Improvement-visual effect	Itching and irritation reduced	Itching and irritation reduced
Side effects	none	none
Price	RM100	RM500

Questionnaire formats

- Discrete choice experiments-example

Attributes	Job A	Job B
Salary	2,200.00	3,000.00
Working hours	8	10
Employer	Public	Private
Critical allowance	Yes	No
	Prefer A ❐	Prefer B ❐

Questionnaire formats

- Discrete choice experiments

  -What if the respondents do not want any of the jobs?

  -Should the respondent have an opt-out alternative?

Questionnaire formats

- Discrete choice experiments-example

Attributes	Job A	Job B	
Salary	2,200.00	3,000.00	
Working hours	8	10	
Employer	Public	Private	
Critical allowance	Yes	N0	
	Prefer A ❒	Prefer B ❒	I don’t want any of the jobs ❒

Questionnaire formats

- Discrete choice experiments-example with more than two

Attributes	**Job A**	**Job B**	**Job C**
Salary	2,200.00	3,000.00	3,500.00
Working hours	8	10	12
Employer	Public	Private	Private
Critical allowance	Yes	No	No
	Prefer A ❒	**Prefer B** ❒	**Prefer C** ❒

Questionnaire formats

Attributes	Option 1 (Satus Quo)	Option 2	Option 3
Forest Area	Decrease by 14%	No change	Increase by 24%
Direct Employment	Increase by 3%	No change	Increase by 3%
No of Migratory Birds	Decline by 3%	No change	No Change
Visitation Rates	No change	Increase by 5%	Increase by 50%
Annual Contribution	**US$0**	**US$50**	**US$10**

Prefer 1 ❐ **Prefer 2 ❐** **Prefer 3 ❐**

Questionnaire formats

- Discrete choice experiments
  - What is the problem with the example in the previous slide?

Questionnaire formats

- Typical arguments used in favor of discrete choice experiments
  - Simulates market situation-only the best product is chosen
  - Low cognitive burden compared to the other elicitation formats
  - Consistent with welfare economic theory-random utility theory

Questionnaire Format: Contingent Ranking

Attributes	Option 1 (Satus Quo)	Option 2	Option 3
Forest Area	Decrease by 14%	No change	Increase by 24%
Direct Employment	Increase by 3%	No change	Increase by 3%
No of Migratory Birds	Decline by 3%	No change	No Change
Visitation Rates	No change	Increase by 5%	Increase by 50%
Annual Contribution	**US$0**	**US$50**	**US$10**

Your Ranking: 1 ___ 2 ___ 3 ___

Questionnaire Format: Contingent Rating - Example

Characteristics	Option 1	Option 2
Native woodland	*500 ha protected*	*700 ha protected*
Heather moorland	*1200 ha protected*	*No protection*
Lowland by meadow	*200 ha protected*	*300 ha protected*
Cost per household per yr in additional tax	*RM50*	*RM20*

Please tick one box only

1	2	3	4	5	6	7	8	9	10

Strongly Prefer Choice A

Strongly Prefer Choice B

Opt-out alternative

- When is the opt-out alternative relevant?
  - The respondent may not prefer to take up a certain programme whatever the level of attributes of the programme
  - E.g: Expensive, risky new alternatives or technologies frequently exhin=bit high levels of non-choice
  - Want to estimate the probability that respondent choose nothing rather than solely trade-off the attributes
  - “ ideally one wants to choose experiments to mimic the actual choice situation faced by the individuals closely as possible”

Opt-out alternative

- Types of Opt-out
  - ‘No-purchase ‘option
  - ‘Current situation’- ‘current purchase’ option

Opt-out alternative

- Failure to include an opt-out alternative may bias the estimates of participation-overstate the likelihood that they would actually participate in a programme or purchase a good
- Implying biased estimates of welfare measures(WTP estimates)
  - Potential overestimation of weight for service attributes
  - Potential underestimation of disutility of the payment attribute

Opt-out alternative

- In which situations should the opt-out alternative be included?
- Examples
  - Screening programmes
  - Air pollution programmes
  - Third party payers' choice of reimbursement scheme for providers

Opt-out alternative is not(has not been) standard in health economics

Opt-out alternative

- Split sample study-Recreational saltwater Fishing sites
  - One group receives DCE including an opt-out option (no-trip option)
  - In the other group the op-out option is specified as an alternative fishing site of the respondents' choice

Efficiency of design (I)

- Full factorial design
  - All combination of the attribute levels
  - E.g. two attributes with three level and two attributes with two levels: $3^2 2^2 = 36$ combinations
  - All main effects, two-way interactions, and all higher-order interactions are estimable and uncorrelated
  - However, the number of effects increase exponentially with increasing number of attributes and attributes level

Efficiency of design (II)

- Fractional factorial design
  - Fewer runs than the full-factorial design
  - Cost: Some effects are confounded or aliased by other effect and cannot be estimated
  - Some effects may even be perfect linear combinations of non-estimable effects
  - Non-estimable confounding effects are assumed to be zero or negligible-if this is not the case, estimates may be biased
  - E.g. main effects or two-way interaction may be aliased by higher-order effects

Efficiency of design (III)

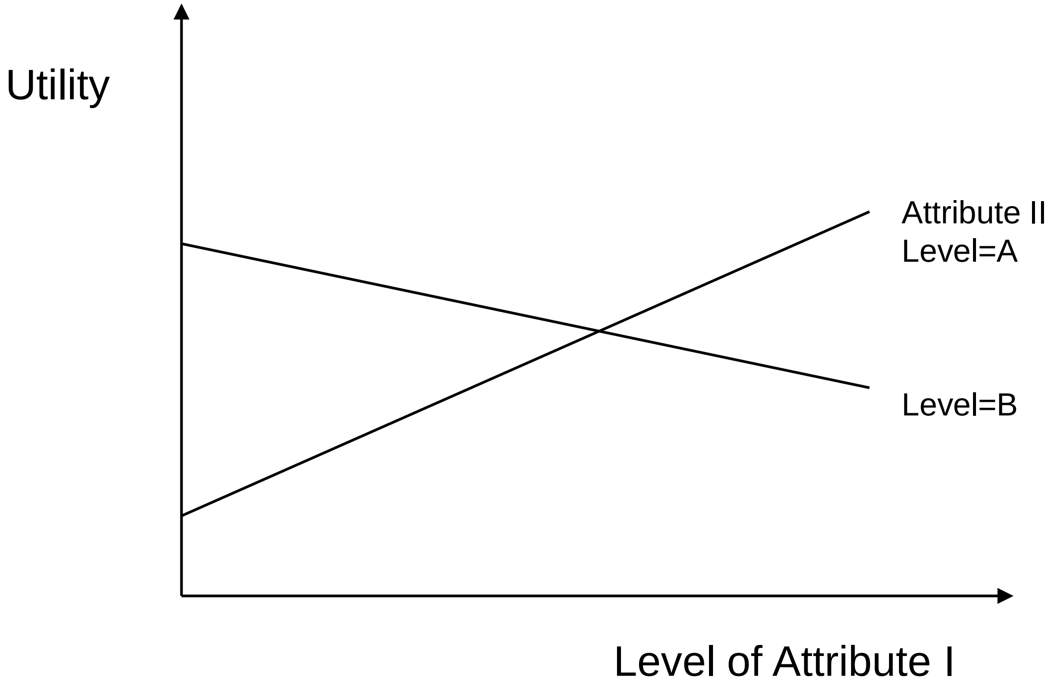
- Effects – a difference in treatment means relative to a comparison, such as the grand mean
- Main effects – the difference in the means of each level of a particular attribute or the grand mean
- Interaction effects – interaction between attributes (two-way or higher-ordered interaction)
  - Interaction occur if the respondents' preference for levels of one attribute depends of a second attribute
  - E.g. distance to a recreational site and the presence of possibility to stay over night

Efficiency of design (IV)

- No interaction effects - illustration

Efficiency of design (V)

- Interaction effects - illustration



Efficiency of design (VI)

- Efficiency criteria
  - Orthogonality
  - Level balance
- Efficiency criteria in for pairing of alternative
  - Minimal overlap
  - Utility balance

Efficiency of design (VII)

- Orthogonality
  - The parameter estimates are uncorrelated meaning that each estimate is independent of the other terms in the model
  - Orthogonality implies that the coefficients will have minimum variance
  - Orthogonal designs are available for only a relatively small number of problems
  - Degree of orthogonality is important for parameter estimation but other factors should not be ignored

Design of alternatives (I)

- Selection of alternatives
  - Number of alternative > number of parameters to be estimated
  - Optimal design sizes: smallest integer >= number of parameters and devidable by number of attributes level
  - E.g. $2^3 3^4$ design: 12 parameters, saturated design
    - Should be devidable by 2, 3, 2x2, 2x3, 3x

Design of alternatives (II)

- Alternative can be used individually asking whether the respondent will accept an alternative or not
- Accept an alternative – yes or no
- Respondents' assess the utility of one alternative relative to the utility of an implicit 'do nothing' or status quo alternative
- This is often not satisfactory when the 'do nothing' alternatives is not realistic or the status quo alternative is inadequately
- More alternative provide more information

Design of alternatives (III)

- The respondents are cognitively not able to choose one from the total number of alternatives
- Choice sets with a fraction of the alternatives should be generated, but how?

Design of choice sets (I)

- Combination of alternatives into choice sets, but how?
- Additional efficiency criteria-Minimal overlap:
  - Differences between alternatives
  - The respondents choose by looking at contrasts etween attribute level within a choice set
  - ‘Minimal overlap means that probability that an attribute level repeats itself in each choice set should be as small as possible”

Design of choice sets (II)

- Minimal overlap- illustration :
  - Some overlap

Attributes	Alternative A	Alternative B	Alternative C
X_1	1	1	0
X_2	2	4	2
X_3	10	20	10

Design of choice sets (III)

- Minimal overlap- illustration :
  - No overlap

Attributes	Alternative A	Alternative B	Alternative C
X_1	1	0	-1
X_2	2	4	2
X_3	10	20	10

Sample size (I)

- Indication of the minimum sample size(number of observations)with information from pilot study
  - **Standard errors and estimates of coefficients from pilot study**

$$N = \frac{2\sigma^2[\ Z_{1-\alpha} + Z_{1-\beta}]^2}{(\mu_1 - \mu_2)^2}$$

N= (α = 0.05 ,β = 0.80)

$$N = \frac{2\sigma^2[\ 1.96 + 1.28\ ]^2}{(\mu_1 - \mu_2)^2}$$

Sample size (II)

- Examples- the minimum sample size providing significant coefficients
- Attribute I : $\sigma = 0.3505$, $\mu=0.0907$, $H_0(\mu) = 0$

$$N = \frac{2 \times 0.3505^2 [1.96 + 1.28]^2}{(0.0907)^2} = 314$$

- Attribute II : $\sigma=0.7050$, $\mu=0.0306$, $H_0(\mu)=0$

$$N = \frac{2 \times 0.7050^2 [1.96 + 1.28]^2}{(0.0306)^2} = 11{,}145$$

Sample size (III)

- Difference between two subgroups or changes in coefficients
- Examples (0.25% change in coefficients)

  - Attribute I : σ=0.1605, μ^1 = 0.1266

$$N = \frac{2 \times 0.1605^2 [1.96 + 1.28]^2}{(0.1266-0.15825)^2} = 540$$

-Attribute II : σ= 0.0075, μ^1 = 0.0032

$$N = \frac{2 \times 0.0032^2 [1.96 + 1.28]^2}{(0.0075 - 0.009375)^2} = 60$$

What do CE data look like

Respondent	Alternative	Choice	Dosage	Cost	Age
1	1	0	1	50	44
1	2	1	1	25	44
1	3	0	2	30	44
1	1	1	1	30	44
1	2	0	2	25	44
1	3	0	2	50	44
1	1	0	2	25	44
1	2	0	1	30	44
1	3	1	1	50	44
2	1	0	1	50	50
2	2	1	1	25	50

Econometric Aspects of Discrete Choice Model

- DCM
  - Based on the notion of latent (unobserved) preference for goods or actions (recreation, job, transport, medical treatnment, etc)
  - Two type of choices
    - Binary or Dichotomous (YES/NO)
    - Multiple choice (three or more alternatives)

Binary Choice

- Two alternatives: A and B to be chosen among.
  - Choice between A and not A.
  - Latent preference for A: Yi, for individual i=1,….,n
- If Yi > 0: Positive preference for A: Choose A, otherwise not choose A (i.e. B)
- Preference unobserved – we observed only Y=1 if A chosen (i.e. Yi > 0) and Y=o if B chosen (i.e. Yi < 0)

Binary Choice

- Preference linearly determined by
  - Characteristics of the individual (age, education, income, gender, race, etc)
  - Characteristics of choices faced by invidual (salary, ticket price, etc0. This must vary for individuals (otherwise captured by constant term – uncontrolled variation)
- Model: $Yi = X\beta + \mu$
- Prob (Y=1) = Prob(Yi>0) = Prob ($X\beta + \mu>0$)=F($X\beta + \mu$)
- F is distribution function

Econometric Aspects of Discrete Choice Model

- DCM
  - Based on the notion of latent (unobserved) preference for goods or actions (recreation, job, transport, medical treatnment, etc)
  - Two type of choices
    - Binary or Dichotomous (YES/NO)
    - Multiple choice (three or more alternatives)