WORKBOOK

DISCRETE CHOICE EXPERIMENT

How-to measure what people prefer:
Health preference research to optimize health-related
interventions

Associate Professor Jason J. Ong
Warittha Tieosapjaroen

HOPE

HOW TO MEASURE WHAT PEOPLE PREFER:

Health preference research to optimize health-related interventions

There is growing interest in how to better measure and integrate end-user preferences into the design and delivery of health-related interventions and programs. Health preference research is an exciting field using quantitative statistical methods to study choices made by individuals or groups to understand how choices are made and to forecast future choices.

This practical one-day workshop aims to introduce discrete choice experiments (DCEs) to students, researchers, clinicians, and policymakers. No prior knowledge of DCEs or health preference is assumed. We will cover theory but spend a significant time to provide "hands-on" experience in the steps of conducting a DCE so they can apply the methodology to their respective settings.

ASSOCIATE PROFESSOR JASON J. ONG is a sexual physician and health economist with expertise in conducting DCEs. He has conducted DCEs in 24 countries to date. This includes evaluating the preferences for HIV testing among men who have sex with men (MSM) in China, HIV testing in youths living in Nigeria, and HIV self-testing among MSM in Australia.

WARITTHA TIEOSAPJAROEN is a PhD candidate at Monash University. She has conducted DCEs related to weight gain in people living with HIV and preferences for pre-exposure prophylaxis (PrEP) in 16 countries in the Asia-Pacific region.

HOPE NETWORK: **HOPE-EndHIV.com**

AIMS OF THE WORKSHOP

- Introduce health preference research with a focus on DCE
- Describe the steps of conducting a DCE
- Share examples of DCEs in health research

AGENDA

8:45-9:00	Arrival
9:00-9:30	Welcome and Introductions
9:30-10:30	Overview of how to conduct DCEs
10:30-11:00	Morning break
11:00-12:00	How to identify the right attributes / levels
12:00-13:00	Lunch
13:00-14:00	Experimental design and presentation of the choice sets
14:00-15:00	How to analyse choice data
15:00-15:30	Afternoon break
15:30-16:30	Application of choice data
16:30-17:00	Questions and Answers + Next Steps (Resources)

WIFI

Click on "FREE_WiFi@SuntecSingapore" in the list of connection choices

CONTENT

*Session 1 is an introduction

SESSION 2:
OVERVIEW OF HOW TO CONDUCT DISCRETE CHOICE EXPERIMENTS
SMALL GROUP DISCUSSION

You are interested in designing a PrEP program and want to use a DCE to measure preferences.

Your choice scenario:
You have chosen to use PrEP.
If these are your available options, which one would you choose?

Tasks:
- Name your group
- Which population(s) do you want to focus on?
- How do you identify relevant attributes/levels? (what methods can you use)?
- List all likely attributes that would influence someone using PrEP

SESSION 3:
HOW TO IDENTIFY THE RIGHT ATTRIBUTES/ LEVELS
SMALL GROUP DISCUSSION

You plan to embed your DCE questions within a survey for your target population.

TASKS

- What other questions would you include in your survey?
- How will you collect survey responses?
    - Discuss the pros/cons of online vs. paper vs. interviewer-assisted
- How will you recruit?

SESSION 4:
EXPERIMENTAL DESIGN AND PRESENTATION OF CHOICE SETS

No tasks

This will be a live demonstration of the NGENE program

SESSION 5:
HOW TO ANALYSE CHOICE DATA

OBJECTIVES

5.1 Understand your DCE data

5.2 Understand DCE syntax

5.3 Demonstrate to Nlogit

5.4 Understand DCE outputs

5.1 Understand DCE data

An example of a DCE question

	A	**B**	**Opt out**
Type of PrEP	Oral long-acting PrEP	Injectable PrEP	x
Service location	Pharmacy	Hospital	x
Cost	$AU 25	Free	x
Side effects	Rare chance of kidney problems	Mild	x
Visit frequency	Every 6 months	Every year	x
Extra services	STI testing	None	x
Which choice do you prefer?	○	○	○

An example of DCE data

	pid	totalrow	choicev	question	cset	altij	type	location	cost	sideeffects
1	168	18	0	1	3	1	2	6	4	4
2	168	18	1	1	3	2	1	1	1	3
3	168	18	0	1	3	3	.	.	.	.
4	168	18	0	2	3	1	2	1	4	2
5	168	18	1	2	3	2	1	3	3	3
6	168	18	0	2	3	3	.	.	.	.
7	168	18	1	3	3	1	5	2	1	4
8	168	18	0	3	3	2	3	4	3	1
9	168	18	0	3	3	3	.	.	.	.
10	168	18	1	4	3	1	5	5	1	4

Important variables
- choiceV = choice indicator (0=not chosen, 1=chosen)
- cset = number of alternatives in each choice situation
- altij = Alternative indicator (1=alternative A, 2=alternative B, 3=opt out)

Other variables
- pid = participant ID
- totalrow= total number of rows for each participant
- question = question number
- type, location, cost, side effects, freq, extra are attributes that are measured in the DCE.

5.2 DCE syntax

This is a basic syntax for MNL model.
Nlogit
  ;lhs=choicev,cset,altij
  ;choices=A,B,C
  ;check data
  ;model:
  U(A,B)=Type1*Type1+Type2*Type2+Type3*Type3+Type4*Type4+
  Cost1*Cost1+Cost2*Cost2+Cost3*Cost3+
  SE1*SE1+SE2*SE2+SE3*SE3+SE4*SE4/
  U(C)=neither
  $

What you need to know to create a DCE syntax for your DCE data
Choices = alternative names (A, B and C for the PrEP DCE.lpj)

U(A,B) = utility function of alternative A and B

Type1*Type1 = Parameter*Variable
              *Parameter is named by you
              *Variable is from the dataset you analyse
Variable = attributes presented to respondent.
Parameter = estimated values that represent underlying preferences that individuals have
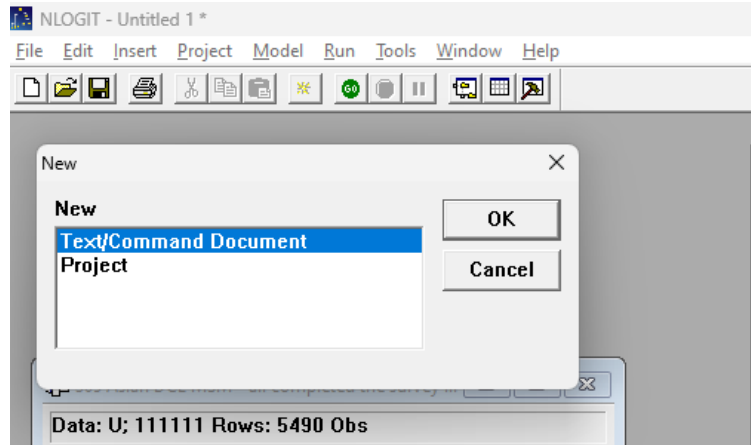for each level of variables

; = begin the command
/ = separate utility function
$ = end of command

5.3 Introduction to Nlogit

1. Download and install Nlogit
2. Open the Nlogit program
3. Click ⬜. A new window will pop up. Choose 'Project' then 'ok'

4. Load your project file (.lim) and you will see the window below. Please check whether all your variables are imported.
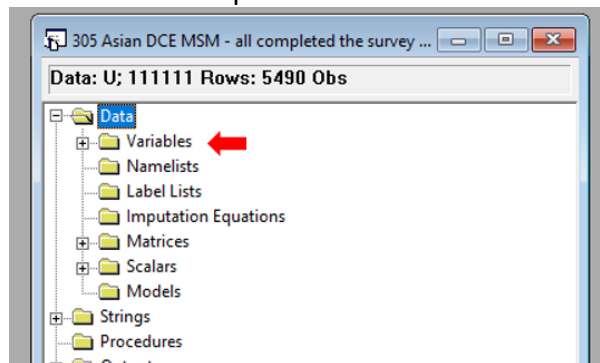
5. Click ⬜. A new window will pop up. Choose 'Text/Command Document' then 'ok'. A new window will pop up.
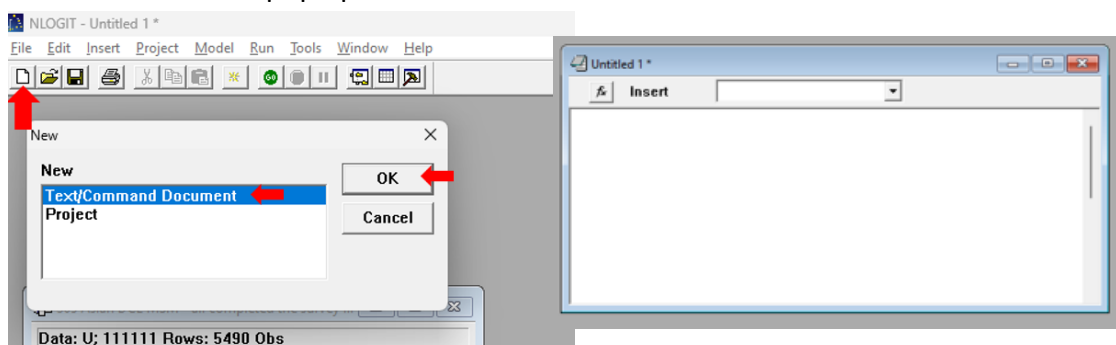
6. Copy and paste the syntax you have prepared into the command window.

7. Select all the code that you want to run and click

8. The output window will pop up

5.4 Understand DCE outputs

5.4.1 Components of MNL outputs

- **Information about the sample**

```
+----------------------------------------------------------+
| Inspecting the data set before estimation.               |
| These errors mark observations which will be skipped.    |
| Row Individual = 1st row then group number of data block |
+----------------------------------------------------------+
No bad observations were found in the sample
```
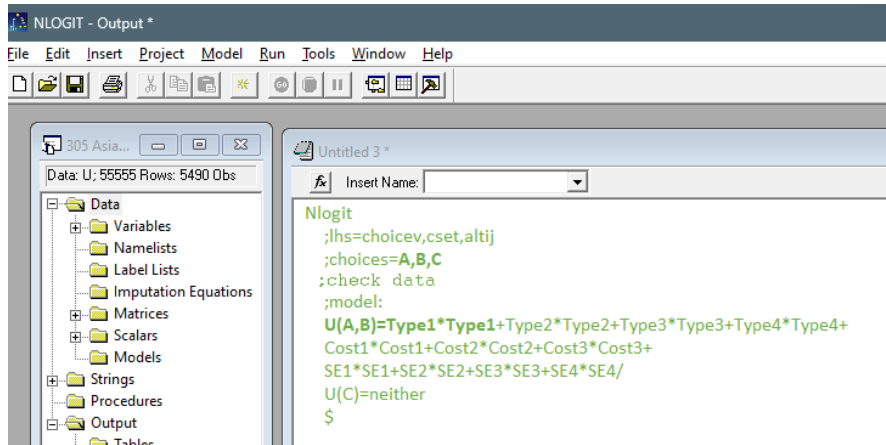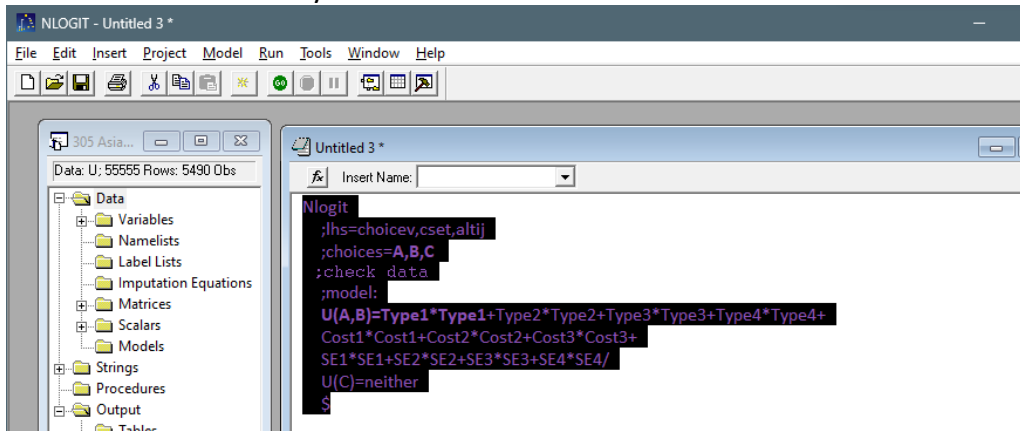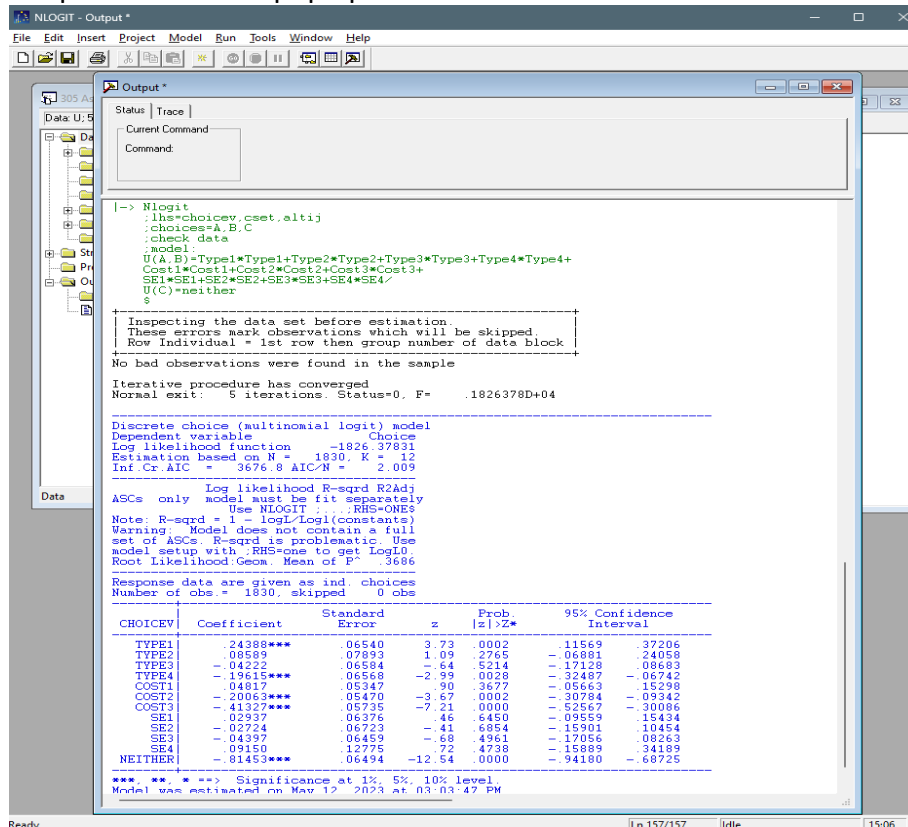
*Nlogit cannot include bad observations (e.g., missing data on an attribute) in the analysis. If 'bad observations were found in the sample' pops up, you will need to check your data again.

- **Normal or abnormal convergence**

```
Iterative procedure has converged
Normal exit:   5 iterations. Status=0, F=     .1826378D+04
```

- **Model summary statistics and report**

Lower Log likelihood function and **AIC/N** indicate a better model fit.

Number of obs. indicates the number of observations included in this analysis.

```
-----------------------------------------------------------------------------
Discrete choice (multinomial logit) model
Dependent variable               Choice
Log likelihood function     -1826.37831
Estimation based on N =   1830, K =  12
Inf.Cr.AIC  =   3676.8 AIC/N =    2.009
---------------------------------------
          Log likelihood R-sqrd R2Adj
ASCs  only  model must be fit separately
          Use NLOGIT ;...;RHS=ONE$
Note: R-sqrd = 1 - logL/Logl(constants)
Warning:  Model does not contain a full
set of ASCs. R-sqrd is problematic. Use
model setup with ;RHS=one to get LogL0.
Root Likelihood:Geom. Mean of P^  .3686
---------------------------------------
Response data are given as ind. choices
Number of obs.= 1830, skipped    0 obs
--------+--------------------------------------------------------------------
        |                Standard          Prob.      95% Confidence
 CHOICEV| Coefficient     Error        z   |z|>Z*        Interval
--------+--------------------------------------------------------------------
   TYPE1|    .24388***     .06540     3.73  .0002     .11569     .37206
   TYPE2|    .08589        .07893     1.09  .2765    -.06881     .24058
   TYPE3|   -.04222        .06584     -.64  .5214    -.17128     .08683
   TYPE4|   -.19615***     .06568    -2.99  .0028    -.32487    -.06742
   COST1|    .04817        .05347      .90  .3677    -.05663     .15298
   COST2|   -.20063***     .05470    -3.67  .0002    -.30784    -.09342
   COST3|   -.41327***     .05735    -7.21  .0000    -.52567    -.30086
    SE1|    .02937        .06376      .46  .6450    -.09559     .15434
    SE2|   -.02724        .06723     -.41  .6854    -.15901     .10454
    SE3|   -.04397        .06459     -.68  .4961    -.17056     .08263
    SE4|    .09150        .12775      .72  .4738    -.15889     .34189
 NEITHER|   -.81453***     .06494   -12.54  .0000    -.94180    -.68725
--------+--------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on May 12, 2023 at 03:03:47 PM
-----------------------------------------------------------------------------
```

5.4.2 Output interpretation

- **Coefficient**

  A positive coefficient indicates desired attribute level, and a negative coefficient indicates undesired attribute level.

Cost levels	Variables	Coefficient	Desirable level
Free	-(COST1+COST2+COST3)	0.56573	Most desirable
Low	COST1	0.04817	
High	COST2	-0.20063	
Very High	COST3	-0.41327	Least desirable

- **z**

  Z indicates the statistically significant difference of the coefficient value

z	*	P-value
>1.65 to <1.96	*	<0.10
1.96 to <2.58	**	<0.05
≥2.58	***	<0.01

- **Relative importance**

Attribute	Levels	Coefficient	Range of attributes	Relative Importance*100
Type of PrEP	Daily oral	-0.0914	0.24388-(-0.19615)=0.44003 **(the highest coefficient – the lowest coefficient)**	0.44003/(0.44003+0.97900+0.14116)=28.2 Range of attributes/sum of the ranges
	On-demand	0.24388		
	Injectable	0.08589		
	Long-acting oral	-0.04222		
	Implant	-0.19615		
Cost	Free	0.56573	0.56573-(-0.41327)=0.97900	0.97900/(0.44003+0.97900+0.14116)=62.7
	Low	0.04817		
	High	-0.20063		
	Very High	-0.41327		
Side effects	No	-0.04966	0.09150-(-0.04966)=0.14116	0.14116/(0.44003+0.97900+0.14116)=9.05
	Interactions with other medications	0.02937		
	Mild	-0.02724		
	Rare chance of kidney problems	-0.04397		
	Mild pain at the injection site	0.09150		

Range of attribute = the highest coefficient – the lowest coefficient for each attribute
Relative importance = range of attribute for each attribute/sum of range of attribute for all attributes

Practice 5.1

Please create a DCE syntax using the data provided below

Antiretroviral A and B have side effects as mentioned below. Choose the antiretroviral that you prefer more.

	Antiretroviral A	**Antiretroviral B**
Nausea	Not all all	Once a week
diarrhea	Three times a week	Not at all
headache	Not at all	Once a week
	○	○

* There are two DCE questions in this survey

----

	pid	Totalrow	choicev	question	cset	altij	nausea	diarrhea	headache
1	111	4	1	1	2	1	1	2	3
2	111	4	0	1	2	2	2	3	1
3	111	4	1	2	2	1	3	2	1
4	111	4	0	2	2	2	2	1	3
5	112	4	0	1	2	1	1	3	2
6	112	4	1	1	2	2	3	1	2
7	112	4	0	2	2	1	2	1	3
8	112	4	1	2	2	2	3	2	1
9	113	4	1	1	2	1	2	1	3
10	113	4	0	1	2	2	1	2	3
11	113	4	1	2	2	1	1	3	2
12	113	4	0	2	2	2	2	3	1

Nlogit
;

14

Practice 5.2
You ran an MNL model for your DCE on PrEP use among MSM in three Asian countries.

```
+----------------------------------------------------------+
| Inspecting the data set before estimation.               |
| These errors mark observations which will be skipped.    |
| Row Individual = 1st row then group number of data block |
+----------------------------------------------------------+
No bad observations were found in the sample

Iterative procedure has converged
Normal exit:   5 iterations. Status=0, F=     .1671696D+05


----------------------------------------------------------------------
--
Discrete choice (multinomial logit) model
Dependent variable               Choice
Log likelihood function    -16716.96429
Estimation based on N =  17034, K =  12
Inf.Cr.AIC  =  33457.9 AIC/N =    1.964
-----------------------------------------
          Log likelihood R-sqrd R2Adj
ASCs  only  model must be fit separately
             Use NLOGIT ;...;RHS=ONE$
Note: R-sqrd = 1 - logL/Logl(constants)
Warning:  Model does not contain a full
set of ASCs. R-sqrd is problematic. Use
model setup with ;RHS=one to get LogL0.
Root Likelihood:Geom. Mean of P^  .3748
-----------------------------------------
Response data are given as ind. choices
Number of obs.= 17034, skipped    0 obs
--------+-------------------------------------------------------------
        |                     Standard          Prob.      95% Confidence
 CHOICEV| Coefficient        Error         z    |z|>Z*       Interval
--------+-------------------------------------------------------------
   TYPE1|     .18423***       .02209     8.34   .0000      .14092    .22753
   TYPE2|    -.10623***       .02681    -3.96   .0001     -.15878   -.05367
   TYPE3|     .20243***       .02177     9.30   .0000      .15975    .24510
   TYPE4|    -.33720***       .02297   -14.68   .0000     -.38222   -.29219
   COST1|     .15201***       .01789     8.50   .0000      .11695    .18707
   COST2|    -.09151***       .01816    -5.04   .0000     -.12710   -.05593
   COST3|    -.70162***       .02060   -34.05   .0000     -.74200   -.66124
     SE1|    -.08038***       .02159    -3.72   .0002     -.12269   -.03807
     SE2|    -.05410**        .02237    -2.42   .0156     -.09795   -.01025
     SE3|    -.28835***       .02231   -12.93   .0000     -.33207   -.24463
     SE4|     .16032***       .04255     3.77   .0002      .07691    .24372
 NEITHER|    -.68532***       .02080   -32.95   .0000     -.72608   -.64456
--------+-------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on May 12, 2023 at 04:33:40 PM
----------------------------------------------------------------------
```

Please add coefficients to this table using the output provided above

Attribute	Levels	Coefficient	Range of attribute	Relative Importance*100
Type of PrEP	Daily oral			
	On-demand			
	Injectable			
	Long-acting oral			
	Implant			
Cost	Free			
	Low			
	High			
	Very High			
Side effects	No			
	Interactions with other medications			
	Mild			
	Rare change of kidney problems			
	Mild pain at injection			

Please report the DCE result

The participants preferred..

16

TASKS

- Share with your group members a possible DCE you could construct within the next 12 months in your area of interest.
- What is your next step?

ANSWERS
Question 5.1
Nlogit
  ;lhs=choicev,cset,altij
  ;choices=A,B
  ;check data
  ;model:
  U(A,B)=Naus1* Naus1+ Naus2* Naus 2+
  Diar1*Diar 1+ Diar 2*Diar 2+
  Head1*Head1+Head2*Head2
  $

*Please note that we use effect coded data, meaning that the base level will be -1. Please see the effect code table below for your information.

	Naus1	Naus2
Not at all	-1	-1
Once a week	1	0
Three times a week	0	1

For more information regarding effect code, please see page 119 'Dummy and effect coding' via this link

https://www.researchgate.net/publication/227390290_Applied_Choice_Analysis

Question 5.2

Attribute	Levels	Coefficient	Range of attribute	Relative Importance*100
Type of PrEP	Daily oral	0.05677	0.54	0.22
	On-demand	0.18423		
	Injectable	-0.10623		
	Long-acting oral	0.20243		
	Implant	-0.3372		
Cost	Free	0.64112	1.34	0.55
	Low	0.15201		
	High	-0.09151		
	Very High	-0.70162		
Side effects	No	0.26251	0.55	0.23
	Interactions with other medications	-0.08038		
	Mild	-0.0541		
	Rare chance of kidney problems	-0.28835		
	Mild pain at the injection site	0.16032		

The participants preferred long-acting oral PrEP, followed by on-demand and daily oral PrEP. They disliked most implant PrEP. (You should use the same structure to report preferences for cost and PrEP side effects) Cost was the most important driver of PrEP use.