Deliberative thinking in discrete choice experiments

A theoretical and empirical investigation

D.A. Regier[1] J. Sicsic[2] V. Watson[3]

[1]Canadian Centre for Applied Research in Cancer Control (ARCC)
Cancer Control Research, BC Cancer
School of Population and Public Health, UBC

[2]LIRAES, University Paris Descartes

[3]Health Economics Research Unit
University of Aberdeen

NOAHE Rounds, Feb 21 2018
Twitter @deanregier

Introduction

Motivations

- Resource allocation requires individuals preferences
  - Stated preferences methods when no market data available
- Discrete choice experiments (DCE)
  - Simulates market conditions and generates choice
  - Framework consistent with microeconomic utility theory (Lancaster, 1966, McFadden, 1974)

Introduction

Motivations

- Resource allocation requires individuals preferences
  - Stated preferences methods when no market data available
- Discrete choice experiments (DCE)
  - Simulates market conditions and generates choice
  - Framework consistent with microeconomic utility theory (Lancaster, 1966, McFadden, 1974)
- A criticism: responses may differ from people's real choices $\rightarrow$ **hypothetical bias** problem (Blumenschein et al, 2001)

Introduction

Literature review

Ex-ante survey designs

- Origin: contingent valuation method (CVM)
  1. Emphasize consequentiality of the choices (Loomis, 2014)
  2. "Oath statements" (Jacquement et al, 2013)
  3. "Cheap talk" (Morrison and Brown, 2009)

Ex-post calibration with **certainty scales**

- Focus of our study
- Assumption: Respondents who state they are certain about their choices (certainty 8 on a 0 to 10 scale):
  1. Are more engaged $\Rightarrow$ more consistent in their choices
  2. Are less subject to *hypothetical bias* (Ready et al, 2010)

Introduction

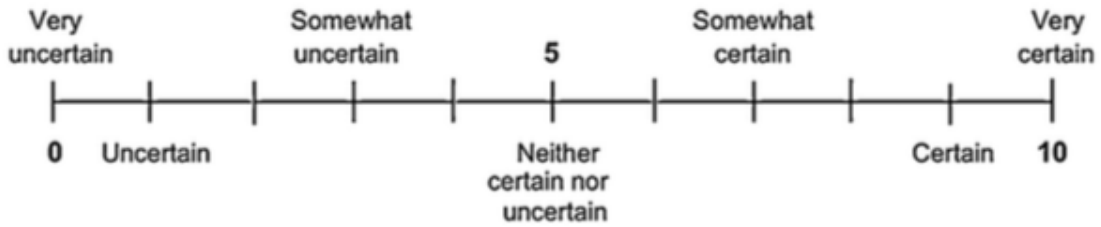
Illustration of a post-choice certainty scale

- Regier et al (2014). J Behav Exp Econ

Characteristics	Scenario A	Scenario B
Drug treatment	Doctor or nurse managed I.V. one day per month	Self or parent managed injection four days per month
Child reported pain from arthritis	Moderate	None
Side-effects	Nausea four days per month	Headache four days per month
Participation in daily activities	With some difficulty	With much difficulty
Days missed from school	Two days per month	Half-a-day per month
Cost to you	$50 per month	$2100 per month

Which would you choose? ☐ Scenario A ☐ Scenario B

How certain are you?



Introduction

Literature review

- Certainty scales are used in two ways :
  1. By re-coding uncertain responses (Li and Mattsson, 1995)
  2. By re-weighting data to favor more certain responses

CVM :

- Both strategies are effective to reduce the hypothetical bias (Murphy et al, 2005)

DCE :

- Mixed evidence (Ready et al 2010, Beck et al 2016)
- Currently no theoretical framework

Introduction

Aims and contributions

1. Propose a new theoretical framework :
   - To identify engaged respondents in DCE tasks
   - Central role of choice certainty **variability**
   - Taxonomy of choice certainty related to *dual processing theories* (using a set of testable assumptions)

Introduction

Aims and contributions

1. Propose a new theoretical framework :
   - To identify engaged respondents in DCE tasks
   - Central role of choice certainty **variability**
   - Taxonomy of choice certainty related to *dual processing theories* (using a set of testable assumptions)
2. Results are consistent with theoretical predictions
3. Alternative ex-post calibration strategies are proposed

Outline

Outline

The two case studies

1. DCE eliciting women's preferences for breast cancer screening (Sicsic et al, 2018)
2. DCE eliciting citizens's preferences for the return of incidental genomic findings (Regier et al, 2015)

Summary of the two case studies

Characteristics	**Case study 1**	**Case study 2**
Setting	France	Canada
Topic	Breast cancer screening	Genomic sequencing
Language	French	English / French
Respondents	Women (age 40-74 years)	Citizens (age 18+ years)
Sample size	812	1200
Choice tasks per respondent	8 (+1)	16
Response certainty scale	0 to 10	0 to 10
Response time	yes	yes
Monotonicity test	yes	no

The two case studies

Case study 1

Example of a breast cancer screening DCE choice task (case study 1)

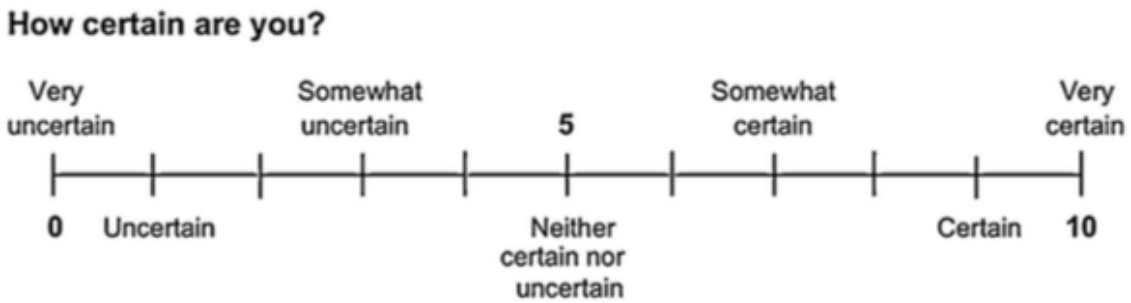
	Screening option A	Screening option B	No screening option
BC mortality	10	25	30
False-positive	200	50	0
Overdiagnosis	150	10	0
Type of screening referral	invitation letter	your doctor	none
Travel time	10 min	90 min	0 min
Number of tests	18	12	0
Out-of-pocket cost	€ 60	€ 30	€ 0
Which option would you choose ?	□	□	□

The two case studies

Case study 2

Example of return of genomic information DCE choice task (case study 2)

	Option A	Option B	No information
Disease Risk *More diseases will be identified if the lifetime risk is lower*	Diseases with a **5% lifetime risk or higher**	Diseases with a **90% lifetime risk or higher**	No information
Disease Treatability	Recommended effective medical treatment only	Recommended effective lifestyle change only	No information
Disease Severity *Health consequences of the diseases you may develop*	Very severe health consequences	Severe health consequences	No information
Carrier Status *Disease risk not affecting you but can affect your family*	Does not provide information on carrier status	Information on if your family members could be affected	No information
Cost to you	$1500	$750	$0
	Option A ☐	**Option B ☐**	**No Information ☐**



Outline

Theoretical framework

Choice certainty and task complexity

- In DCE, response certainty is related to choice task complexity (Olsen et al, 2011)
- Task complexity is proxied by the utility difference between alternatives in a choice task (Regier et al, 2014)
  - Alternatives that provide similar utility to respondents are hard to distinguish $\Rightarrow$ certainty $\searrow$
  - Alternatives that provide very different utilities to respondents are easy to distinguish $\Rightarrow$ certainty $\nearrow$
- Because **task complexity varies in a DCE**, we posit that engaged respondents should vary in their certainty

Theoretical framework

Certainty and deliberative thinking

Dual processing theories distinguish:

1. Respondents who make intuitive choices without much thought (*System 1*)
   - Experience computational limitations $\rightsquigarrow$ decision heuristics and errors of intuition (Simon, 1979):
2. Respondents using deliberative thinking (*System 2*)
   - They make rational choices based on all of the information available (Kahneman, 2003)

Theoretical implications

- Only System 2 respondents experience task complexity
- Their level of certainty should vary during the DCE as a result of varying task complexity

Theoretical framework

Assumptions to be tested

- **H1** Engaged respondents should be **on average** sufficiently certain of their choices
  - *Rational*: certain respondents offer more consistent choices (Becker et al, 2013; Dekker et al, 2016)
- **H2** In DCE, only individuals with sufficient certainty variability use rational (deliberative) decision-making
  - *Corollary*: Respondents who are *always* certain are more likely to make quick and intuitive choices

Theoretical framework

Taxonomy of choice certainty

Group label	Choice certainty		Assumptions	
	Mean	SD	Engagement in deliberative thinking	Data quality
Constantly uncertain (CU)	<=6	< $\sigma_{uncertain}$	low	low
Variably uncertain (VU)	<=6	>= $\sigma_{uncertain}$	low/moderate	low/moderate
Constantly hesitant (CH)	]6 ; 8]	< $\sigma_{hesitant}$	low	low
Variably hesitant (VH)	]6 ; 8]	>= $\sigma_{hesitant}$	high	high
Constantly certain (CC)	>8	< $\sigma_{certain}$	low	low
Variably certain (VC)	>8	>= $\sigma_{certain}$	high	high

Note: σ denotes the mean certainty standard deviation of a class. For instance, $\sigma_{uncertain}$ is the mean SD of the *uncertain* class.

- **H1** *"Engaged respondents should be **on average** sufficiently certain of their choices"*
  - It excludes the two *uncertain* groups (CU and VU)
- **H2** *"Only individuals with sufficient certainty variability use rational (deliberative) decision-making"*
  - Further excludes the CH and CC groups

Theoretical framework

Descriptive statistics

Certainty group	Case study 1 (Breast cancer screening)		Case study 2 (Genomic sequencing)	
	N	%	N	%
Constantly uncertain (CU)	113	13.9%	126	10.5%
Variably uncertain (VU)	78	9.6%	94	7.8%
Constantly hesitant (CH)	179	22.0%	290	24.2%
Variably hesitant (VH)	145	17.9%	215	17.9%
Constantly certain (CC)	166	20.4%	264	22.0%
Variably certain (VC)	131	16.2%	211	17.6%
Overall	812	100%	1200	100%

- Similar certainty distribution in case studies 1 and 2
  - The *variably uncertain* (VU) group has lowest sample size
  - The *constantly hesitant* (CH) group has highest sample size

Theoretical framework

Set of testable conjectures

- Deliberative thinking can not be inferred from the data:
  - Assumptions tested based on **5 conjectures**

Theoretical concept	**Conjecture**	**Data quality indicators**
1. Decision heuristics	The *CU, CH and CC* groups are more likely to use decision heuristics	(i) serial non-trading behaviour (ii) serial non-demanding behaviour
2. Monotonicity of preferences	The *VH and VC* groups have monotonic preferences more often	Ability to select a dominant alternative (with better attribute levels) in a specific choice task
3. Considered choices	The *VH and VC* groups offer more considered choices	(i) median response time (ii) % of "speeders" (RT<10 min)
4. Logical consistency	The choices of the *VH and VC* groups exhibit higher logical consistency	Number and % of preferene parameters in line with a priori assumptions
5. Choice consistency	The choices of the *VH and VC* groups exhibit higher choice consistency	Error variance in heteroskedastic multinomial logit models

Empirical testing

Conjecture 1 (decision heuristics)

Certainty group	Serial non-traders				Serial non-demanders			
	Case study 1		Case study 2		Case study 1		Case study 2	
	N	%	N	%	N	%	N	%
Constantly uncertain (CU)	**12**	**10,2%**	0	0,0%	6	5,3%	**27**	**21,4%**
Variably uncertain (VU)	4	5,1%	0	0,0%	1	1,3%	5	5,3%
Constantly hesitant (CH)	8	4,5%	2	0.93%	12	6,7%	25	8,6%
Variably hesitant (VH)	1	0,7%	1	0.34%	1	0,7%	5	2,3%
Constantly certain (CC)	1	0,6%	1	0.38%	**29**	**17,5%**	**135**	**51,1%**
Variably certain (VC)	1	0,8%	0	0,0%	1	0,8%	23	10,9%
Chi square test (p-value)	*p<0.0001*				*p<0.0001*		*p<0.0001*	
Total	27	3,3%	4	0.3%	50	6,2%	220	18,3%

- The results provide empirical support for conjecture 1 □

Empirical testing

Conjectures 2, 3 and 4

Certainty group	Conjecture 2. **Monotonicity of preferences**		Conjecture 3. **Response time**				Conjecture 4. **Logical consistency**			
	Case study 1		Case study 1		Case study 2		Case study 1		Case study 2	
	N	%	Median	% speeder	Median	% speeder	N	%	N	%
Constantly uncertain (CU)	**77**	**68.1%**	**12.8**	**27.4%**	**12.8**	**23,0%**	**4/8**	**50%**	**7/10**	**70%**
Variably uncertain (VU)	68	87.2%	15.6	16.7%	15.4	22.3%	5/8	63%	8/10	80%
Constantly hesitant (CH)	**147**	**68.1%**	**14.1**	**21.8%**	**14.8**	**18.9%**	7/8	88%	9/10	90%
Variably hesitant (VH)	**131**	**90.3%**	**16.3**	**8.9%**	**16.1**	**15.8%**	**8/8**	**100%**	**10/10**	**100%**
Constantly certain (CC)	**127**	**76.5%**	**13.9**	**16.7%**	**13.8**	**18.9%**	**5/8**	**63%**	**8/10**	**80%**
Variably certain (VC)	**124**	**94.7%**	**15.8**	**9.2%**	**17.2**	**11.8%**	**8/8**	**100%**	**9/10**	**90%**
P-value of independence test	*p<0.0001*[c]		*p<0.0001*	*p<0.0001*[c]	*p<0.0001*	*p=0.080*[c]	-		-	
Total	674	83,0%	14.8	16.8%	15.1	17.8%	8/8	100%	11/11	100%

- Study 1: empirical support for conjecture 2 □
- Studies 1/2: empirical support for conjectures 3 and 4 □

Empirical testing

Conjecture 5 (choice consistency)

Results of Heteroscedastic Multinomial Logit Models (Case study 1)

	HMNL 1		HMNL 2		HMNL 3	
Scale function parameters	Estimate	(SE)	Estimate	(SE)	Estimate	(SE)
Constantly uncertain (CU)	**-0.233***	**(0.137)**	-	-	-	-
Variably uncertain (VU)	-0.121	(0.139)	-	-	-	-
Constantly hesitant (CH)	*ref*		-	-	-	-
Variably hesitant (VH)	**0.277*****	**(0.098)**	-	-	-	-
Constantly certain (CC)	**-0.504*****	**(0.184)**	-	-	-	-
Variably certain (VC)	**0.508*****	**(0.095)**	-	-	-	-
Mean certainty	-	-	0.048	(0.030)	0.068**	(0.031)
SD certainty	-	-	**0.301*****	**(0.062)**	**0.325*****	**(0.075)**
Mean certainty*SD certainty	-	-	-	-	**0.211*****	**(0.049)**
# choice observations	6496		6496		6496	
# respondents	812		812		812	
Log-Likelihood	-5855.2717		-5906.5256		-5848.2509	

- Heteroskedastic MNL: scale $\nearrow \Rightarrow$ choice consistency $\nearrow$

Empirical testing

Conjecture 5 (choice consistency)

Results of Heteroscedastic Multinomial Logit Models (Case study 1)

	HMNL 1		HMNL 2		HMNL 3	
Scale function parameters	Estimate	(SE)	Estimate	(SE)	Estimate	(SE)
Constantly uncertain (CU)	**-0.233***	**(0.137)**	-	-	-	-
Variably uncertain (VU)	-0.121	(0.139)	-	-	-	-
Constantly hesitant (CH)	*ref*		-	-	-	-
Variably hesitant (VH)	**0.277*****	**(0.098)**	-	-	-	-
Constantly certain (CC)	**-0.504*****	**(0.184)**	-	-	-	-
Variably certain (VC)	**0.508*****	**(0.095)**	-	-	-	-
Mean certainty	-	-	0.048	(0.030)	0.068**	(0.031)
SD certainty	-	-	**0.301*****	**(0.062)**	**0.325*****	**(0.075)**
Mean certainty*SD certainty	-	-	-	-	**0.211*****	**(0.049)**
# choice observations	6496		6496		6496	
# respondents	812		812		812	
Log-Likelihood	-5855.2717		-5906.5256		-5848.2509	

- Heteroskedastic MNL: scale $\nearrow \Rightarrow$ choice consistency $\nearrow$
- Results of HMNL1: empirical evidence for conjecture 5 $\square$
- HMNL2/HMNL3: $\frac{\partial Consistency}{\partial \sigma_{certainty}} > 0$, $\frac{\partial^2 Consistency}{\partial \mu_{certainty} \partial \sigma_{certainty}} > 0$

Empirical testing

Conjecture 5 (choice consistency)

Results of Heteroscedastic Multinomial Logit Models (Case study 2)

	HMNL 1		HMNL 2		HMNL 3	
Scale function parameters	Estimate	(SE)	Estimate	(SE)	Estimate	(SE)
Constantly uncertain (CU)	0.060	(0.109)	-	-	-	-
Variably uncertain (VU)	0.018	(0.146)	-	-	-	-
Constantly hesitant (CH)	*ref*		-	-	-	-
Variably hesitant (VH)	**0.226***	**(0.136)**	-	-	-	-
Constantly certain (CC)	**-0.888****	**(0.397)**	-	-	-	-
Variably certain (VC)	**0.238*****	**(0.092)**	-	-	-	-
Mean certainty	-	-	**-0.038****	**(0.019)**	-0.007	(0.022)
SD certainty	-	-	0.062	(0.049)	**0.118****	**(0.051)**
Mean certainty*SD certainty	-	-	-	-	**0.090*****	**(0.027)**
# choice observations	19200		19200		19200	
# respondents	1200		1200		1200	
Log-Likelihood	-19904,959		-19970,004		-19956,269	

- Results of HMNL1: empirical evidence for conjecture 5 □
- HMNL3: $\frac{\partial Consistency}{\partial \sigma_{certainty}} > 0$, $\frac{\partial^2 Consistency}{\partial \mu_{certainty} \partial \sigma_{certainty}} > 0$

Outline

Alternative ex-post calibrations

- **Aim**: Design alternative calibration strategies of respondents' choices
- Using information on choice certainty and *variability* and assess its impact on :
  1. The precision of welfare estimates (i.e., WTA/WTP)
  2. The accuracy of welfare estimates (does hypothetical bias decrease ?)
- **Problem**: we don't observe real WTA/WTP $\Rightarrow$ we can not assess the extent of hypothetical bias

Alternative ex-post calibrations

- **Aim**: Design alternative calibration strategies of respondents' choices
- Using information on choice certainty and *variability* and assess its impact on :
  1. The precision of welfare estimates (i.e., WTA/WTP)
  2. The accuracy of welfare estimates (does hypothetical bias decrease ?)
- **Problem**: we don't observe real WTA/WTP $\Rightarrow$ we can not assess the extent of hypothetical bias
- **We posit** there is hypothetical bias
  1. Thus individuals overestimate their WTA/WTP (Ready et al, 2010, Beck et al, 2016)
  2. *Corollary*: any decrease in WTA/WTP estimates $\Rightarrow$ more plausible welfare estimates

Alternative ex-post calibrations

The re-weighting models

- The following weights enter the expression of the SLL function of the choice model

$$w_n = 1 \tag{1}$$

$$w_n = \mu_{certainty} \tag{2}$$

$$w_n = \sigma_{certainty} \tag{3}$$

$$w_n = \mu_{certainty} \times \sigma_{certainty} \tag{4}$$

- The weights are normalized such that : $\sum w_n = N$

Alternative ex-post calibrations

Results: case study 1

Willingness-to-accept (WTA)		**(1)** (w_n=1)		**(2)** (w_n= mean certainty)		**(3)** (w_n= SD certainty)		**(4)** (w_n= mean*SD certainty)	
		Estimate	*% change*	Estimate	*% change*	Estimate	*% change*	Estimate	*% change*
False-positives	Mean	50,27	*ref*	50,90	*1%*	42,86	*-15%*	43,74	*-13%*
	SE	12,70	*ref*	12,48	*-2%*	7,73	*-39%*	7,55	*-41%*
	95%CI	(25.38-75.15)	-	(26.45-75.35)		(27.71-58.01)		(28.94-58.55)	
Overdiagnosis	Mean	14,01	*ref*	14,31	*2%*	13,12	*-6%*	13,45	*-4%*
	SE	0,61	*ref*	0,60	*-1%*	0,49	*-20%*	0,47	*-23%*
	95%CI	(12.82-15.20)	-	(13.13-15.49)		(12.17-14.07)		(12.53-14.37)	
Tavel time	Mean	19,95	*ref*	20,55	*3%*	18,43	*-8%*	18,73	*-6%*
	SE	2,61	*ref*	2,67	*2%*	1,91	*-27%*	1,84	*-29%*
	95%CI	(14.83-25.08)	-	(15.31-25.79)		(14.69-22.17)		(15.11-22.34)	
Screenig tests	Mean	9,89	*ref*	10,37	*5%*	6,38	*-35%*	6,38	*-35%*
	SE	4,84	*ref*	5,11	*6%*	1,61	*-67%*	1,50	*-69%*
	95%CI	(0.40-19.38)	-	(0.35-20.39)		(3.22-9.53)		(3.44-9.33)	
# observations		6496		6496		6496		6496	
wSLL		-4687,1293		-4573,6364		-4812,3461		-4707,0675	

- In model (2) : higher WTA and higher SEs
- In models (3-4) : lower WTA and lower SEs □

Alternative ex-post calibrations

Results: case study 2

Willingness-to-pay (WTP)		**(1)** (w_n=1)		**(2)** (w_n= mean certainty)		**(3)** (w_n= SD certainty)		**(4)** (w_n= mean*SD certainty)	
		Estimate	*% change*	Estimate	*% change*	Estimate	*% change*	Estimate	*% change*
WTP carrier status	Mean	390,53	*ref*	428,44	*10%*	345,31	*-12%*	370,37	*-5%*
	SE	19,4	*ref*	21,1	*9%*	18,6	*-4%*	19,2	*-1%*
	95%CI	(352.6-428.5)	-	(387.0-469.8)		(308.9-381.7)		(332.5-408.2)	
WTP treat. of the disease (treatment + lifestyle)	Mean	202,51	*ref*	227,48	*12%*	189,71	*-6%*	207,29	*2%*
	SE	29,4	*ref*	31,4	*7%*	28,3	*-4%*	29,0	*-1%*
	95%CI	(144.9-260.1)	-	(166.0-288.9)		(134.3-245.1)		(150.5-264.1)	
WTP 40% lifetime risk or higher	Mean	274,95	*ref*	279,63	*2%*	243,29	*-12%*	251,52	*-9%*
	SE	34,0	*ref*	36,3	*7%*	33,0	*-3%*	33,8	*-1%*
	95%CI	(208.2-341.7)	-	(208.5-350.7)		(178.5-308.1)		(185.2-317.8)	
WTP 90% lifetime risk or higher	Mean	354,33	*ref*	384,91	*9%*	394,16	*11%*	419,19	*18%*
	SE	30,3	*ref*	32,6	*8%*	29,7	*-2%*	30,6	*1%*
	95%CI	(295.0-413.6)	-	(321.1-448.8)		(336.0-452.3)		(359.2-479.2)	
# observations		1200		1200		1200		1200	
wSLL		-13154,64		-12816,806		-15214,620		-15009,953	

- In model (2) : higher WTA estimates and higher SEs
- In models (3-4) : lower WTA estimates and lower SEs □ (but lower effect sizes compared to case study 1)

Outline

Conclusion

Summary of findings

- We have refined the current framework to analyse repondents' certainty in DCE non market valuation tasks
- We have shown that :
  - Respondents who are *always* certain of their choices (i) are more likely to use decision heuristics and (ii) their choices are less consistent
  - Re-weighting respondents to favor those with higher certainty **variability** improved:
    1. The precision of welfare estimates (up to +69%)
    2. The plausibility of welfare estimates (up to +35%)

Conclusion

Strengths

- Theory-based evidence
- Results are consistent in two different settings
- The alternative calibration technique is simple

Limitations

1. No revealed preference data $\Rightarrow$ difficult to interpret the results in terms of hypothetical bias reduction
2. The approach works better in one setting
   - Impact of the complexity of the experimental design, study population, age ?
3. Is the calibration strategy *too* simple ?
   - Need to assess non-linearity in the re-weighting function