Handling resolvable uncertainty from incomplete health care scenarios - choice probabilities versus discrete choices

Morten Raun Mørkbak, University of Southern Denmark

Line Bjørnskov Pedersen, University of Southern Denmark

Riccardo Scarpa, University of Durham

Background

- DCEs has been the dominant framework within choice modelling to estimate preferences and forecast real behaviour
- Manski (1990) argues that even under best case hypothesis stated intentions will not be good predictors of future behavior.
- New (and better?) approaches are constantly emerging, e.g. best-worst, RRM and *choice probability elicitation*
- Conventional choice probability models also assume RUT but allow for uncertainty in choices
  - Some distributional assumptions can be relaxed

A short introduction to choice probability elicitation I

- Divergence between stated choices and actual choices is due to events/information relevant to choice which will be revealed in the time period between the expression of intentions and the realization of behavior (Manski 1999)

- Manski (1999) refers to this as “resolvable uncertainty”
  - Resolvable uncertainty: uncertainty about utility components that are not stated in the choice scenario but would be known in an actual choice setting
  - Unresolvable uncertainty: i.e. utility components that remain unknown in the actual choice situation

- Divergence is due to the necessarily incomplete information setting provided to subjects at the stage of elicitation of choice intentions

- Manski (1999) refers to this as DCEs being ”incomplete scenarios”

A short introduction to choice probability elicitation II

- In the standard DCE framework the issue of incomplete scenarios is typically handled by assuming that all that remains undescribed in the characterization of alternatives is equal across alternatives and respondents.
- This is a very pragmatic, and potential naïve and poorly credible assumption.
- Because of cognitive limitations and incomplete information it is impossible to include all potential characteristics of an alternative in a CE setting.
- Eliciting choice probabilities instead of stated choices, as proposed by Manski (1999) could potentially overcome this issue, by allowing respondents to explicitly be uncertain about their stated choice.
- It turns out that this approach might afford the additional advantage of being less econometrically demanding.

Litterature

- Blass et al. (2010), show how the elicitation of choice probabilities can empirically be fitted within the random utility framework with random utility coefficients, using data on consumers' preferences for the reliability of electricity services in Israel (IER).

- Shoyama et al. (2013) used the same approach together with a standard DCE for eliciting public preference for land-use scenarios in Kushiro watershed in northern Japan (LUP).

- Herriges et al. (2011) use the 2009 Iowa Lake Survey to administer a split treatment in terms of information provision (low and high) and preference elicitation method (preferred choice versus probability of choice) (WP).

- The approach has lately been used within labor economics (but not in a DCE framework) – studying e.g. college students' major choices (Arcidiacono et al. 2012 JEconmetrics; Wiswall and Zafar 2015 RES).

Objective

- The objective is to contribute to this growing literature by providing a case study in health care grounded on the work of Manski (1999) and implemented by Blass et al. (2010)

- We aim at comparing the elicited subjective choice probabilities approach with the more standard DCE approach

Econometric analyses – stated discrete choices

Lancaster consumer theory and random utility theory:

$$U_{ntk} = V(x_{ntk}, \beta) + \varepsilon_{ntk}$$

Choice probability:

$$P_{ntk} = \frac{e^{\beta' x_{ntk}}}{\sum_i^J e^{\beta' x_{nti}}}$$

Maximum Likelihood estimation

Econometric analyses – Elicited choice probabilities I

Still Lancaster consumer theory and random utility theory, but extending it with the uncertainty created by the incomplete alternatives:

$$U_{ntk} = V(x_{ntk}, \beta) + \varepsilon^{r}_{ntk} + \varepsilon^{u}_{ntk}$$

Random utility framework - individual *n* forms a subjective distribution Q_{nt} which provides us with a subjective choice probability:

$$q_{ntk} = Q_{nt}[x_{ntk}\beta_n + \varepsilon^{u}_{ntk} > x_{nti}\beta_n + \varepsilon^{u}_{nti}, \qquad for\ all\ i \neq k]$$

Allows us to estimate the subjective random utility model (still assumming ε to be iid):

$$q_{ntk} = \frac{e^{x_{ntk}\beta_n}}{\sum_{i=1}^{J} e^{x_{nti}\beta_n}}$$

Econometric analyses – Elicited choice probabilities II

The linear mixed logit probability is obtained by making a log-odds transformation:

$$ln\left(\frac{q_{ntk}}{q_{nt1}}\right) = (x_{ntk} - x_{nt1})\beta_n = (x_{ntk} - x_{nt1})m + u_{ntk}, \qquad k = 2, \dots \dots, J$$

Where the alternative *k=1* is chosen and $\beta_n = m + \sigma_n$, $u_{ntk} = (x_{ntk} - x_{nt1})\ \sigma_n$.

Without loss of generalization set $E(\sigma)=0$, which then will provide us with $m=E(\beta)$, $E(u|x)=0$, thus the linear mixed logit probability transforms into a linear mean regression model:

$$E\left[ln\left(\frac{q_{ntk}}{q_{nt1}}\right)|x\right] = (x_{ntk} - x_{nt1})m$$

Econometric analyses – Elicited choice probabilities III

Problems if respondents tend to rounding off probabilities.

- if it takes place near 0 and 1 - the log-odds being very sensitive near the boundaries (0 and 1), and in the extreme case will end up provide log-odds of either plus or minus infinity.

The inference problem created by rounding off small numbers can be resolved by assuming symmetry (and in absence of rounding), we have the following linear median regression model, which can be estimated by using Least Absolute Deviations (LAD):

$$M\left[ln\left(\frac{q_{ntk}}{q_{nt1}}\right)|x\right] = (x_{ntk} - x_{nt1})m$$

Estimation of WTP is also straight forward using elicited choice probabilities and equals the negative ratio between any given characteristic and the price/cost

Methods
- case

- We use stated discrete choices and elicited choice probabilities in a randomized split survey design

- We study postgraduate medical students' preferences for establishing in rural general practice in Denmark.

- This has been a topic of concern that has been addressed in the health economics literature a number of times (see e.g. Pedersen and Nexøe (2016) for a short overview).

- Stated choice experiments have been used to shed lights on the issue (see e.g. Holte et al. (2015) and Li et al. (2014)), whereas one other study has used best worst scaling (Günther et al. 2010).

Methods
- focus group interview (n=8)

Attributes identified in focus group interview
Included in the choice experiment
Population
Number of GPs in the practice
Control over working hours
Distance to leisure activities, school and childcare
Job security for partner
Early bonus
Excluded in the choice experiment
Professional development
Workload
Professional collaboration
Distance to carreer options
Number of on call duties
Collaboration with other general practices
Procedurally work
Continuity in care
Time for each patient

- A minimum of nine attributes candidate to transform into resolvable uncertainty or unresolvable uncertainty in the choice probability models.

Methods
- survey design

Attributes	Levels
Population	Below 2000 inhabitants 2000-5000 inhabitants 5000-10000 inhabitants 10000-20000 inhabitants
Number of GPs in the practice	1 GP (you) 2 GPs 3-4 GPs
Control over working hours	Low degree High degree
Distance to leisure activities, school and day care	Cycling distance Requires car / public transport
Job security for partner in local area	Low High
Distance to closest family	Cycling distance Short car ride Long car ride
Yearly bonus	0 DKK 50000 DKK 150000 DKK 300000 DKK

- Discrete choices

Valgsæt 1 af 12	Praksis A	Praksis B
Indbyggertal	10.000 - 20.000	5.000 - 10.000
Antal læger i praksis	2	2
Graden af kontrol over arbejdstimer	Høj grad	Lav grad
Afstand til familiære forhold (fritid, skole, børnepasning)	Kræver bil / offentlig transport	Kræver bil / offentlig transport
Jobsikkerhed for partner i nærområde	Høj	Lav
Afstand til nærmeste familie	Lang biltur	Cykelafstand
Årlig bonus	50.000 kr	300.000 kr

Hvilken praksis foretrækker du?

- Praksis A
- Praksis B
- Jeg vægter begge praksis lige højt

- Choice probabilities

Valgsæt 1 af 12	Praksis A	Praksis B
Indbyggertal	10.000 - 20.000	5.000 - 10.000
Antal læger i praksis	2	2
Graden af kontrol over arbejdstimer	Høj grad	Lav grad
Afstand til familiære forhold (fritid, skole, børnepasning)	Kræver bil / offentlig transport	Kræver bil / offentlig transport
Jobsikkerhed for partner i nærområde	Høj	Lav
Afstand til nærmeste familie	Lang biltur	Cykelafstand
Årlig bonus	50.000 kr	300.000 kr

Hvor stor er sandsynligheden for, at du vælger praksis A eller B?

Methods
- data collection

- The questionnaire was distributed to postgraduate medical students in Denmark in October 2015 using internet forums specifically established for and used by medical students at the four universities in Denmark educating doctors.

- The link to the questionnaire was shared in the groups three times during the data collection process.

- In total, 316 respondents answered the questionnaire, of whom 167 answered the discrete choice questions, and 149 answered the choice probability questions.

Analyses

- Test for successful randomization
- Descriptive comparison of choices (A vs B)
- Comparison of WTPs and test for differences
- Hit rates and hold-out samples

Results
- test for successful randomisation I

	Split		Pearson Chi2 test
	Discrete choice	Probability	
Gender			
Men	28%	28%	
Women	71%	72%	0.956
University			
Aarhus University	58%	47%	
Aalborg University	15%	17%	
Copenhagen University	12%	25%	
University of Southern Denmark	15%	12%	0.022
Length of study			
<= 4 years	41%	38%	
> 4 years	59%	62%	0.230
Marital status			
Single	25%	28%	
Married	12%	9%	
Cohabiting	43%	46%	
Have a partner - not cohabiting	20%	16%	
Do not know	1%	1%	0.751

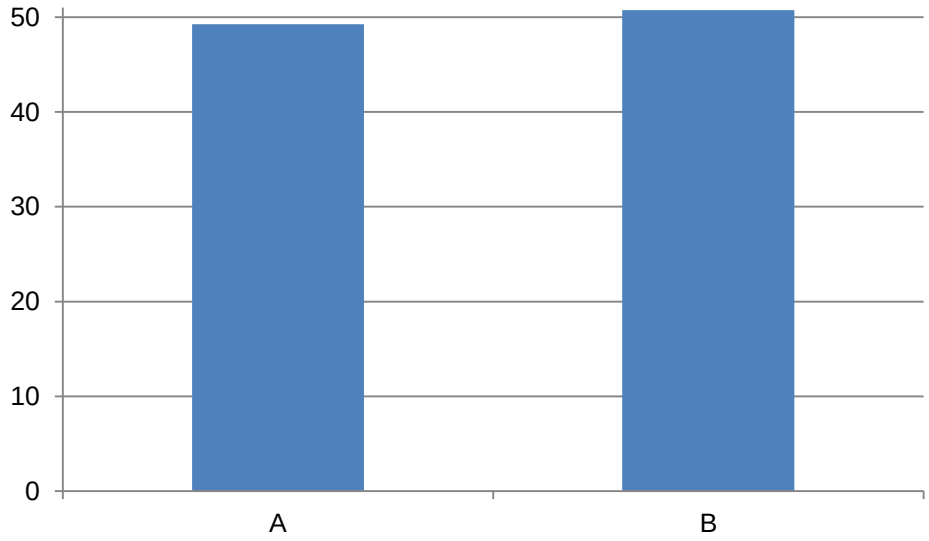
Results
- test for successful randomisation II

Do you have a state education loan?			
Yes	37%	42%	
No	62%	58%	
Do not wish to disclose	1%	0%	0.284
Do you have or have you had a study-related job during your education?			
Yes	83%	85%	
No	16%	13%	
Do not wish to disclose	1%	1%	0.639
What speciality do you expect to choose after becoming MD?			
Respondents could choose between 38 different specialties, where general practice was one of them (not displayed here)			0.783
How probably do you consider it to be that you become a general practitioner?			
Very unlikely	10%	11%	
Unlikely	17%	17%	
Neither or	24%	27%	
Likely	35%	28%	
Very likely	15%	17%	0.720
Could you consider taking a job in a rural area without getting economically compensated?			
Yes	52%	49%	
No	28%	27%	
Do not know	20%	24%	0.713

Results
- Descriptive comparison of choices I

Preferred alternatives (stated discrete choices)

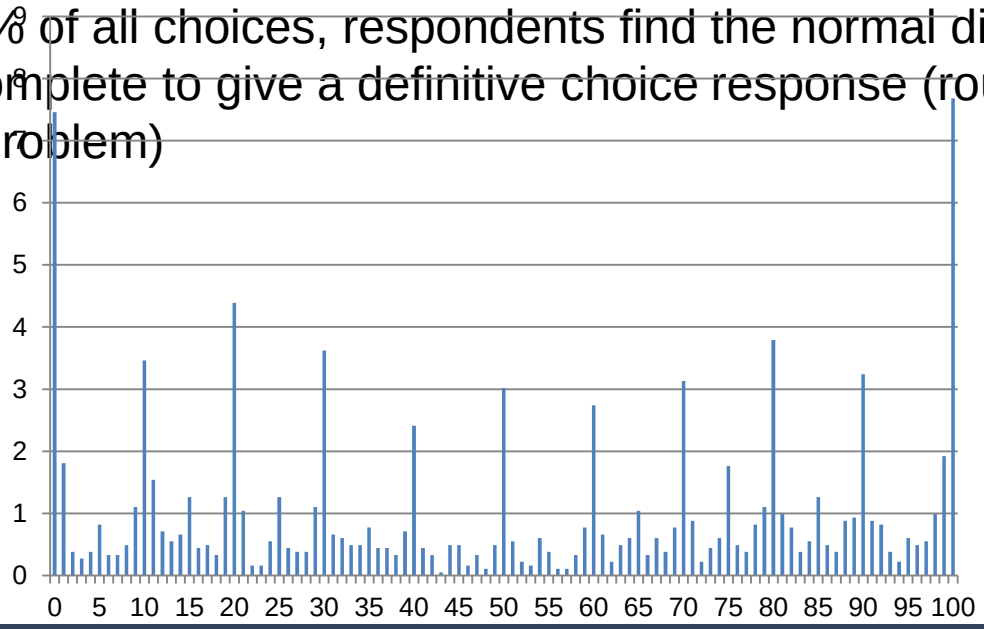
- Alternative A: 49.27 %
- Alternative B: 50.73 %



Results
- Descriptive comparison of choices II

Elicited Choice Probabilities

- A probability below 50 % (choosing Alternative A) was chosen 47.78 % of the time
- A probability above 50 % (choosing Alternative B) was chosen 49.07 % of the time
- The 50/50 option was chosen 3.02 % of the time
- In almost 85% of all choices, respondents find the normal discrete choice scenario incomplete to give a definitive choice response (rounding at 1 and 0 not a major problem)



Results – Mean WTP

	Choice	Probability
	WTP	WTP
pop2	284	1089*
pop25	163	902*
pop510	123	217*
gp2	-357	-241
gp34	-778	-552
control	-367	-156*
schoolbike	-420	-473
jobhigh	-238	-143
familybike	-300	127*
familycar	-282	-742*

WTP (in 1000 DKK)

Results
- Predictions - Hit rates

Discrete choices vs. elicited choice probabilities

Predictions from:	Actual choices		Hold out sample (15%)
	Preferred alternative split	Probability split	
Choice	77.90%	74.87%	78.90%
Probability	76.24%	74.42%	74.30%

Points for discussion and further direction

Discrete choice models versus probability models

- Probability models allow for uncertainty in choices – makes it more realistic?
- Probability models require weaker parametric assumptions and is easier to implement
- WTPs different; Probability models are almost as good at predicting stated choices

Further analyses – less restrictive assumptions

- *(Z)OIB and Maximum Score Estimation* on choice probability split
- Analyses on other cases within health economics

Directions for future research

- Do people understand probabilities? Maybe a problem for some groups within society?
- Which choice probability model (LAD, ZOIB, MSE) should be recommended?
- External validity of the choice probability appraoch?

Thank you

E-mail: mrm@sam.sdu.dk

Twitter: @mr_moerkbak

Twitter: @COHERE_SDU