Monograph:
Hole, A.R. (2010) A discrete choice model with endogenous attribute attendance. Working Paper. Department of Economics, University of Sheffield ISSN 1749-8368

Sheffield Economic Research Paper Series 2010006

Sheffield Economic Research Paper Series

SERP Number: 2010006

ISSN 1749-8368

Arne Risa Hole

A discrete choice model with endogenous attribute attendance

February 2010

Department of Economics
University of Sheffield
9 Mappin Street
Sheffield
S1 4DT
United Kingdom
www.shef.ac.uk/economics

Abstract:

This paper develops a discrete choice model in which the decision to consider an attribute in the choice process is modelled endogenously. In an application to patients' choice of general practitioner it is found that the proposed model outperforms the standard logit model in terms of goodness of fit and produces substantially different estimates of willingness to pay.

Key words: discrete choice, attribute attendance

JEL: C25

Acknowledgments:
I am grateful to Jennifer Roberts and Christine Valente for helpful comments.

A discrete choice model with endogenous attribute attendance

Arne Risa Hole*
Department of Economics
University of Sheffield

February 18, 2010

Abstract

This paper develops a discrete choice model in which the decision to consider an attribute in the choice process is modelled endogenously. In an application to patients' choice of general practitioner it is found that the proposed model outperforms the standard logit model in terms of goodness of fit and produces substantially different estimates of willingness to pay.

Keywords: discrete choice, attribute attendance

JEL codes: C25

*Department of Economics, University of Sheffield, 9 Mappin Street, Sheffield S1 4DT, UK. Tel.: +44 1142 223411, Fax: + 44 1142 223458. E-mail: a.r.hole@sheffield.ac.uk. I am grateful to Jennifer Roberts and Christine Valente for helpful comments.

1 Introduction

It is typically assumed that participants in choice experiments take all the available information into account when choosing between alternatives. Recent evidence suggests, however, that decision-makers often simplify the choice task by ignoring one or more alternative attributes (e.g. Hensher, 2006; Ryan et al., 2009). This paper contributes to the literature by developing a discrete choice model in which the decision to consider an attribute in the choice process is modelled endogenously. The proposed endogenous attribute attendance (EAA) model has a two-stage structure where in the first stage the decision-maker chooses the subset of attributes to take into account when choosing an alternative. In the second stage the preferred alternative is chosen conditional on the choice of attributes in stage one. In an application to patients' choice of general practitioner the EAA model outperforms a standard logit model in terms of goodness of fit and produces substantially different estimates of willingness to pay. This suggests that the usual assumption that decision-makers take all the available information into account when making their choices could lead to biased estimates of key policy parameters.

The proposed model has a similar structure to the independent availability model suggested by Swait and Ben-Akiva (1987) and is related to the latent class approach to modelling attribute attendance recently proposed by Scarpa et al. (2009)[1]. The important advantage of the EAA model over the latent class approach, however, is that all possible attribute subsets can be included in the model as opposed to only a limited number.

The paper is organised as follows. Section 2 describes the endogenous attribute attendance model, section 3 presents an empirical example where the EAA model is compared to the standard logit model and section 4 concludes.

[1] See also Hensher and Greene (forthcoming).

2 The model

We assume a sample of N decision-makers who have the choice between J alternatives on T choice occasions. Each choice is taken two be a two-step process in which the decision-maker first decides which attributes to take into account when comparing the available alternatives and second chooses the alternative with the best characteristics given his or her preferences. More formally, when faced with a total of K alternative attributes the respondent chooses a subset of attributes C_q to take into account when choosing an alternative. The total number of attribute subsets is given by $Q = 2^K$, which includes the set in which all attributes are included (C_Q) and the set in which none of the attributes are included (C_1). Note that C_Q corresponds to the conventional assumption that the decision-makers make use of all the available information on the alternatives when making a choice. C_1 represents the other extreme in which the respondents discard all the information about the alternatives, which implies that the choice process in the second stage is random. Conditional on the choice of attribute subset C_q in the first stage, the utility that individual n derives from choosing alternative i on choice occasion t is given by $U_{nit} = \sum_{k \epsilon C_q} x_{nit}^k \beta^k + \varepsilon_{nit}$ where x_{nit}^k represents the value of attribute k relating to alternative i on choice occasion t, β^k is the preference weight given to that attribute and ε_{nit} is a random term which is assumed to be IID extreme value.

Given these assumptions the probability that decision-maker n chooses alternative i on choice occasion t conditional on the choice of attribute subset C_q is given by the logit formula:

$$L_{nit}|C_q = \frac{\exp(\sum_{k \epsilon C_q} x_{nit}^k \beta^k)}{\sum_{j=1}^{J} \exp(\sum_{k \epsilon C_q} x_{njt}^k \beta^k)} \tag{1}$$

The probability of the observed sequence of choices conditional on C_q is

$$S_n|C_q = \prod_{t=1}^{T} L_{ni(n,t)t}|C_q \tag{2}$$

where $i(n,t)$ denotes the alternative chosen by individual n on choice occasion t.

The probability that decision-maker n takes attribute k into account in the first stage is specified as $\exp(\gamma_k' z_n)/\left[1+\exp(\gamma_k' z_n)\right]$, where z_n is a vector of observed characteristics of individual n and γ_k is a vector of parameters to be estimated. Assuming that these probabilities are independent over attributes the probability of choosing attribute subset C_q is given by:

$$H_{nC_q} = \prod\nolimits_{k\epsilon C_q} \frac{\exp(\gamma_k' z_n)}{1+\exp(\gamma_k' z_n)} \prod\nolimits_{k\notin C_q} \frac{1}{1+\exp(\gamma_k' z_n)} \tag{3}$$

H_{nC_q} could instead be modelled using a multinomial logit structure in which each alternative represents an attribute subset (Scarpa et al., 2009), but such a model becomes unpractical as the number of attributes grows.

Combining equations (2) and (3) the unconditional probability of the observed sequence of choices is

$$P_n = \sum\nolimits_{q=1}^{Q} H_{nC_q} \times S_n|C_q \tag{4}$$

The model is estimated by maximising the log-likelihood function $LL = \sum\nolimits_{n=1}^{N} \ln P_n$.

It should be noted that it is not possible to identify γ_k if $\beta^k = 0$. In other words, if the preference weight given to attribute k is zero it is not possible to estimate the probability of attending to this attribute. We will return to this issue in the following section. It should also be pointed out that it would be conceptually straightforward (but computationally intensive) to incorporate preference heterogeneity in the model by allowing the β^k parameters to be randomly distributed across decision-makers. This extension is left for future research.

3 An application to patients' choice of general practitioner

3.1 The choice experiment

To illustrate how the endogenous attribute attendance model can improve our understanding of the decision-making process we use data from a discrete choice experiment on patients' preferences for the key attributes of a primary care consultation. The attributes in the experiment (Table 1) were chosen after extensive focus group and pilot testing.

[Table 1 about here]

Each patient received a questionnaire with 8 choice sets and a limited number of questions regarding socio-demographic characteristics. In each of the choice sets the patients were presented with 2 alternative consultations with different characteristics and asked to imagine that the reason for their consultation was a heavy cold. For details regarding the experimental design and questionnaire development see Cheraghi-Sohi et al. (2008). The estimation sample consists of 3367 usable responses by 425 respondents.

3.2 Results

The results from a standard logit and a relatively simple EAA model in which the attribute attendance probabilities do not depend on individual characteristics ($z_n = 1$) are presented in Table 2. The coefficients in the two models are qualitatively similar in terms of sign and significance with the exception of the coefficient for the "doctor is warm and friendly" attribute which is insignificant in the EAA model[2]. The logit model is conclusively rejected in favour of the EAA model using a likelihood ratio test (LR-stat=719.58, P-value < 0.001).

[2] As mentioned in section 2 the γ_k parameters in the first stage are not identified when the attribute coefficient β^k is zero so the normalisation $\gamma_{warm} = 1$ was imposed for identification purposes.

[Table 2 about here]

It can be seen from the results that a substantial share of respondents ignored one or more of the attributes when making their choices. The most frequently ignored attribute was "cost of appointment" which was only taken into account by 30% of the respondents, followed by "choice of appointment times" (33%) and "doctor knows you" (57%). The waiting time and "thorough physical examination" attributes were taken into consideration more frequently, but still ignored by a substantial share of the respondents (40%). The low proportion of respondents taking cost into account must be borne in mind when interpreting the willingness to pay (WTP) estimates presented in Table 3.

[Table 3 about here]

There are noticeable differences in the WTP estimates derived from the standard logit and the EAA model. With the exception of the "doctor knows you" attribute the standard logit estimates are considerably higher. This suggests that imposing the assumption that respondents trade off all the attributes in the choice process may lead to biased estimates of willingness to pay, which confirms previous similar findings in the literature (Scarpa et al., 2009; Hensher and Greene, forthcoming). The fact that less than a third of the respondents take cost into account in the decision-making process also means that these estimates only apply to a subset of the population, as it is not possible to infer anything about the willingness to pay of the respondents who ignored the cost attribute.

4 Conclusion

This paper develops a discrete choice model in which attribute attendance is modelled endogenously. In an application to patients' choice of general practitioner it is found that the proposed model outperforms a standard logit model in terms of goodness of fit and produces different estimates of willingness to pay. This suggests that the usual assumption that decision-makers take all the available information into account when making their choices could lead to biased estimates of key policy parameters.

References

[1] Cheraghi-Sohi, S., Hole, A.R., Mead, N., McDonald, R., Whalley, D., Bower, P. and M. Roland, 2008, What patients want from primary care consultations: A discrete choice experiment to identify patients' priorities, Annals of Family Medicine 6, 107-115.

[2] Hensher, D.A., 2006, How do respondents process stated choice experiments? Attribute consideration under varying information load, Journal of Applied Econometrics 21, 861–878.

[3] Hensher, D.A. and W. Greene, Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification, Empirical Economics (forthcoming).

[4] Ryan, M., Watson, V. and V. Entwistle, 2009, Rationalising the 'irrational': A think aloud study of discrete choice experiment responses, Health Economics, 18, 321-336.

[5] Scarpa, R., Gilbride, T.J., Campbell, D. and D.A. Hensher, 2009, Modelling attribute non-attendance in choice experiments for rural landscape valuation. European Review of Agricultural Economics, 36, 151–174.

[6] Swait, J. and M. Ben-Akiva, 1987, Incorporating random constraints in discrete models of choice set generation, Transportation Research B, 21B, 91–102.

Table 1. Attributes and levels in the choice experiment

Attribute	Levels
Cost of appointment to patient	£0, £8, £18, £28
Number of days wait for an appointment	Same day, next day, 2 days, 5 days
Flexibility of appointment times	One appointment offered Choice of appointment times offered
Doctor’s interpersonal manner	Warm and friendly Formal and businesslike
Doctor’s knowledge of the patient	The doctor has access to your medical notes and knows you well The doctor has access to your medical notes but does not know you
Thoroughness of physical examination	The doctor gives you a thorough examination The doctor’s examination is not very thorough

Table 2. Estimation results

	Logit		EAA model			
			2nd stage		1st stage	
Variable	**Coef.**	**t-stat**	**Coef.**	**t-stat**	**Prob.**	**Std.err**
Cost (pounds)	-0.040	-16.36	-0.330	-13.03	0.295	0.024
Waiting time (days)	-0.367	-23.02	-1.687	-12.57	0.601	0.029
Dr knows you well	0.231	5.00	1.935	7.36	0.571	0.083
You get a choice of appointment times	0.211	4.69	0.987	2.16	0.330	0.175
Dr is warm and friendly	0.120	2.67	0.068	0.61	1	
Dr gives you a thorough physical examination	1.156	22.78	5.325	10.68	0.602	0.035
Number of responses	3367		3367			
Number of respondents	425		425			
Log likelihood	-1569.96		-1210.17			

Table 3. Willingness to pay estimates

	Logit	EAA model
Waiting time (days)	9.19 (8.09 – 10.29)	5.11 (4.62 – 5.60)
Dr knows you well	5.78 (3.31 – 8.26)	5.86 (4.59 – 7.14)
You get a choice of appointment times	5.29 (2.97 – 7.61)	2.99 (0.32 – 5.66)
Dr is warm and friendly	3.00 (0.81 – 5.20)	0.21 (-0.47 – 0.88)
Dr gives you a thorough physical examination	28.98 (25.39 – 32.57)	16.14 (14.40 – 17.87)

Note: 95% confidence intervals calculated using the delta method in parentheses.