Hybrid choice modelling allowing for reference-dependent preferences

The case of alternative-fuel vehicles in Denmark

Stefan L. Mabit, Elisabetta Cherchi, Anders F. Jensen, DTU Transport,
and Jørgen Jordal-Jørgensen, Cowi, Denmark

Prepared for Danish Choice Modelling Day, December 5th 2012

DTU Transport
Department of Transport

Outline

- Research question
- Survey and data
- Modelling approach
- Estimation results and validation
- Discussion

Research question

- Attitudes, perceptions and loss aversion have been shown to affect choice in many contexts.

- The hybrid choice model has become a popular tool to include attitudes and perceptions in discrete choice models. Models allowing for reference-dependent preferences have been used to test whether individuals are loss averse.

- Here we investigate:
  - How do the combined inclusion of latent variables and allowing for reference-dependent preferences improve modelling of choice behaviour?

  - How do the latent variable – appreciation of car features – and reference-dependent cost preferences affect the modelling of preferences related to alternative-fuel vehicles (AFVs)?

Data – Survey

- Data on choices among conventional and various alternative-fuel vehicles were collected from August 2007 to December 2007 in an Internet survey.

- A random sample was taken from the Danish population of new-car buyers. The recent purchase is used as the reference vehicle.

- The sample has 2107 respondents and each respondent completed 4-8 stated choices. There is a total of 14694 observations.

- We use data collected in 2008 from the same survey as a validation sample.

Data – choice experiment

- Alternatives
  - Conventional - base pollution
  - Hybrid - pollution at 50% of base
  - Bio-diesel - base pollution except CO_2 at 50% of base
  - Electric - no pollution

- Attributes
  - price (DKK),
  - annual cost (DKK),
  - acceleration time (sec.),
  - range (kms), and
  - a service dummy

- Choices were conditional on non-listed car attributes, e.g. car type.

Data – choice experiment

- The design was pivoted around the attributes of the reference purchase and the attributes in the experiments were generated using a uniform distribution around the reference values.

- The experiments were binary choices among two of the four fuel types.

- The latent variable ”appreciation of car features” was found as the most influential factor in a factor analysis of the indicators. We use indicators measuring importance on a scale from 1-6 of

  – Road position
  – Driving enjoyment
  – Car likeability
  – Car noticeable
  – Comfort
  – Design

Data – choice experiment example

Spil 1 --- Traditionelle biler overfor elbiler --- side 1 af 4

De skal regne med at begge biler er fuldt ud lige driftsikre og anvendelige.

De kan få vist yderligere information om bilernes teknologi og miljoforhold ved at klikke på den blå understregede tekst.

Marker den onskede bil ved at klikke på feltet nederst under den valgte kolonne.

Traditionel dieselbil	Elbil med batterier (Samme størrelse)
Anskaffelsespris: 196.300 Kr.	Anskaffelsespris: 173.800 Kr.
Årlige omkostninger til drift, vedligehold og brændstof: 21.600	Årlige omkostninger til drift, vedligehold og brændstof: 24.500
Rækkevidde: 900 km på en tank	Rækkevidde: 850 km på en tank
Acceleration 0-100 km/t: 12.1 sec	Acceleration 0-100 km/t: 13.4 sec
Serviceaftale til dækning af service og vedligehold, samt reparationer der ikke er dækket af garantien, er inkluderet i de årlige driftomkostninger. Inklusive gratis lånebil	Serviceaftale til dækning af service og vedligehold, samt reparationer der ikke er dækket af garantien, er inkluderet i de årlige driftomkostninger. Inklusive gratis lånebil
Forurening som en almindelig dieselbil (med filter)	Ingen forurenende udstødning
Jeg foretrækker denne bil ○	**Jeg foretrækker denne bil** ○

Omkostninger til drift, vedligehold og brændstof er beregnet ud fra den årskorsel der er angivet tidligere i sporgeskemaet

Pause | Forrige side | Næste side

Modelling

- We use the hybrid choice model framework, see e.g. Walker (2001), and extend it to allow for reference-dependent preferences.

- The model includes two structural equations

$$U_{in} = f(x_{in}, x^*{}_n; \beta_n) + \varepsilon_{in}, \varepsilon \; is \; IID \; EV \; type \; 1$$

$$x^*{}_n = Bw_n + u_n, u \sim N(0, \sigma^2)$$

and a measurement equation for each indicator

$$I_n = \alpha + \Lambda x^*{}_n + \nu_n, \nu \sim N(0, \theta^2)$$

- We treat each indicator as continuous and constrain the coefficients of the first indicator for identification ($\alpha_1 = 0, \Lambda_1 = 1$).

Modelling - specification

- We specify the choice model as

$$U_{in} = \beta_{n,1}{}' x_{in,1} + \beta_{n,p}(x_{in,p} - x_{n,ref}) \exp(\eta_n * sign((x_{in,p} - x_{n,ref})) + \varepsilon_{in}$$

where

$x_{in,1}$, $x_{in,p}$ are the attributes

$x_{n,ref}$ is the reference price for individual n

$$\beta_{n,l} = \sum \beta_l{}^k s_{n,k} + \gamma_l x^*{}_n$$

$$s_n = w_n = \left\{ \begin{matrix} male, age, children, income, single, \\ worker, commute\ distance, car\ usage, \\ diesel\ ref.veh., ref.veh.finance \end{matrix} \right\}$$

Modelling - models

- We estimate
  - Model 1: An MNL model with systematic heterogeneity
  - Model 2: Model 1 extended to allow for reference-dependence in cost preferences
  - Model 3: Model 1 extended to include the latent variable ACF
  - Model 4: Model 2 and 3 combined
- All models are estimated in PythonBiogeme using numerical integration

Overall model statistics

- The overall estimation result for the four models were

	Model 1	Model 2 – ref. dep	Model 3 – lat. var.	Model 4 – lat. var. + ref. dep.
DoF	21	24	52 (36)	55 (39)
Final global fct.			-127135	-127093
Final LL	-7910	-7855	-7869	-7829
Choice model ρ^2	0.221	0.227	0.224	0.227

Estimation results - base model

- The most important coefficients from Model 1 are presented in the table.

Variable	Estimate	z test
Hybrid ASC	0.51	8.0
Bio-diesel ASC	0.20	3.1
Electric ASC	0.75	9.5
Electric ASC * male	-0.34	-5.6
AFV ASC * worker	0.20	2.8
Acceleration * male	-0.94	-3.3
Acceleration * a30	-0.98	-2.1
Annual cost	-0.52	-9.1
Range	0.96	12.7
Price	-1.72	-22.7
Price * a60	0.23	2.2
Price * children	-0.22	-2.5
Price * short commute	0.27	2.8
Price * high income	0.36	4.4
Price * single	-0.66	-4.6

Estimation results

- Here we present the coefficient that are added in Models 2-4

	Model 2 – ref. dep.		Model 3 – lat. var.		Model 4 – lat. var. + ref. dep.	
η	0.27	5.1			0.23	4.5
η * diesel	0.14	2.6			0.15	3.1
η * loan	-0.13	-2.5			-0.16	-3.1
Hybrid ASC * ACF			0.25	3.6	0.26	3.7
Bio-diesel ASC * ACF			0.21	3.1	0.21	3.2
Electric ASC * ACF			0.59	7.5	0.60	7.4
Acceleration * ACF			0.51	2.0	0.53	2.1
Annual cost * ACF			-0.24	-2.2	-0.24	-2.3
Range * ACF			0.20	1.6	0.19	1.5
Price * ACF			-0.73	-7.6	-0.64	-6.8

WTP and elasticity

- Statistics calculated using sample enumeration

	Sample	No. obs.	WTP annual cost	WTP operation range	Elasticity, price
Model 1	All	14694	3.8	107.9	-2.2
	Loss	3912	3.7	104.3	-2.6
	Gain	5329	3.8	109.4	-1.8
Model 2	Loss	3912	2.9	82.5	-3.4
	Gain	5329	5.0	145.2	-1.3
Model 3	All	14694	3.7	104.6	-2.1
	Loss	3912	3.7	101.1	-2.6
Model 4	Gain	5329	3.7	105.9	-1.8
	Loss	3912	3.0	83.9	-3.2
	Gain	5329	4.7	133.9	-1.4

Validation on hold-out sample

- We validated the models on a hold-out sample of 2510 observations from January 2008 and 105845 observations from January to April 2008, and 18739 observation from January to July 2008.

2008		Model 1	Model 2 – ref. dep	Model 3 – lat. var.	Model 4 – lat. var. + ref. dep.
	DoF	21	24	36	39
Jan	Final LL	-1322	-1320	-1311	-1310
Jan	Choice model ρ^2	0.228	0.227	0.225	0.224
Jan-Apr	Final LL	-5683	-5679	-5644	-5646
Jan-Apr	Choice model ρ^2	0.223	0.223	0.226	0.225
Jan-Jul	Final LL	-10146	-10144	-10087	-10092
Jan-Jul	Choice model ρ^2	0.217	0.217	0.221	0.220

Discussion and comments

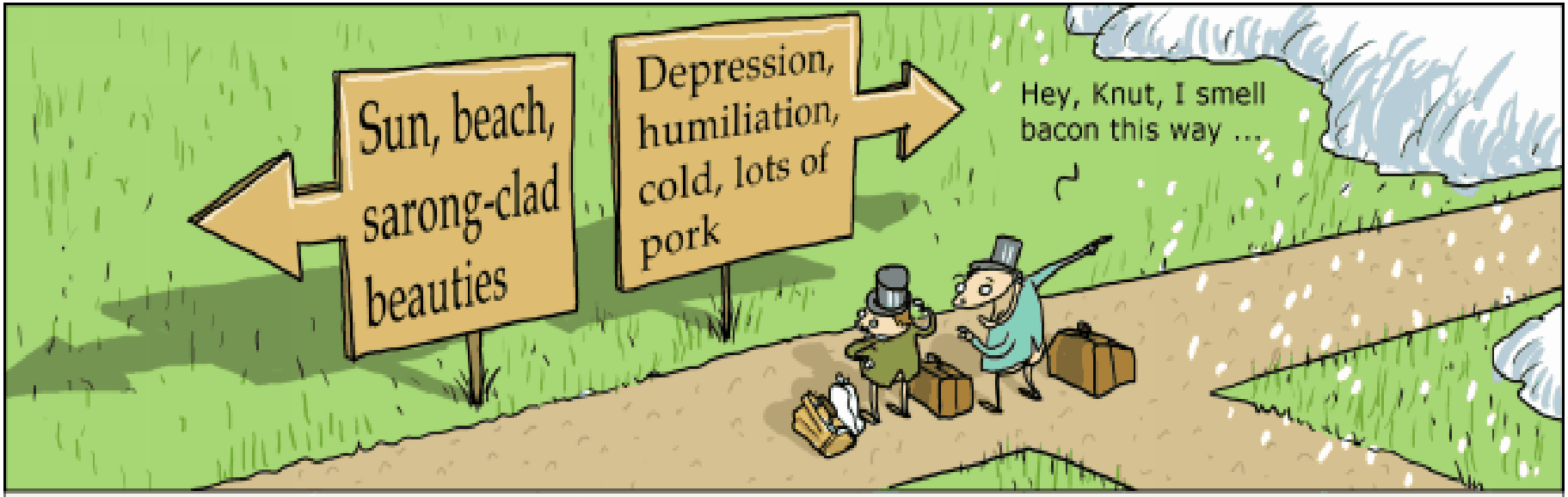
- We included explanatory variables by testing whether they were significant in both structural equations. The latent variable model was estimated alone to decide the specification.

- We do not acknowledge the panel dimension of our data. This will probabaly lower t tests by a factor around 2.

- We have applied a simple validation method and will test other validation methods in the future.

- We should test whether the indicator should be treated as discrete.

Thank you for listening



The founders of Scandinavia

- Source: http://wumocomicstrip.com/