Heterogeneity in preferences

Stephane Hess
stephane.hess@gmail.com

Outline

1. Including heterogeneity

2. Interpretation with heterogeneity

Including heterogeneity

Including heterogeneity

What is heterogeneity?

- Heterogeneity is another word for variability
- In a choice modelling context, this relates to:
  - differences across people
  - differences across choices for same person
    - e.g. different settings, different points in time
- Main focus is in differences in sensitivities (e.g. cost sensitivity)
- Heterogeneity in preferences can lead to differences in choice outcomes

Mathematical fit

- Incorporating heterogeneity in our models increases flexibility
- Ability to better explain choices
- Models with heterogeneity invariably obtain better model fit

Including heterogeneity

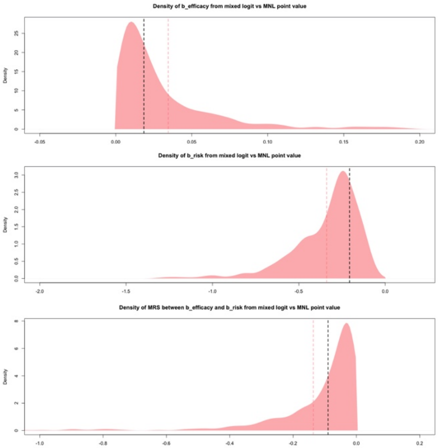
Behavioural insights

- Provides insights into why some people make specific choices
  - e.g. underlying health conditions reducing the likelihood of using active transport
- Can ensure a more robust appraisal of cost vs benefits
  - e.g. if we know how relative importance of time and cost varies across travellers
- Can help in the context of shaping a more efficient provision of transport services
  - e.g. make a variety of services available, where the mix is such that it can help cover the widest range of preferences
- Can help shape strategies to nudge behaviour
  - e.g. identify rasons for lack of use of active transport

Including heterogeneity

What if I'm only interested in the mean?

- Offering multiple services may not be possible
- May conclude that only aggregate preferences are of interest
- Aggregate preferences from a model allowing for heterogeneity may be different (and less biased) than those from a model assuming homogeneity



Including heterogeneity

Why do we get bias?

- Not accounting for heterogeneity increases the amount of noise in the model

$$
\begin{aligned}
U_{jnt} &= \sum_{k=1}^{K} \beta_{n,k}\, x_{jnt,k} + \varepsilon_{jnt} \\
&= \sum_{k=1}^{K} \beta_{k}\, x_{jnt,k} + \varepsilon_{jnt} + \sum_{k=1}^{K} \left(\beta_{n,k} - \beta_{k}\right) x_{jnt,k}
\end{aligned}
$$

- Unobserved part of utility now correlated with deterministic part, breaking core assumption of additive in errors random utility models, potentially leading to bias
- Even if not *interested* in heterogeneity, will have less bias by accounting for it

Including heterogeneity

Where is the heterogeneity?

taste heterogeneity

- differences in relative sensitivities to individual attributes
- differences in baseline preferences for different options

scale heterogeneity

- differences across individuals in amount of noise (from analyst's perspective)

process heterogeneity

- differences in how information is processed, and what decision rule is used

- In practice, difficult/impossible to disentangle, main focus is taste heterogeneity

Including heterogeneity

Deterministic heterogeneity: link to observed information

Option 1: discrete segmentations, with separate models

- e.g. male *vs* female, business *vs* leisure
- same as a joint model with segment-specific parameters
- assumes that differences exist in sensitivities to all attributes

Option 2: differences only for some attribute-covariate pairs

- interaction with categorical variables
  - e.g. interaction with gender, implying different sensitivities for men and women
- interactions with continuous covariates
  - e.g. continuous interaction with income, implying different sensitivity for each possible income

Including heterogeneity

Discrete segmentations

- Divide population into mutually exclusive subsets
- Estimate separate parameters for different segments
  - separate models
  - separate parameters within the same model
- Potential segmentations
  - occupation
  - income class
  - age group
  - gender

Including heterogeneity

Linear interactions

- Examples with income
- Assumes a unit elasticity for income: $V_{jn} = \dots + \beta_C \frac{C_{jn}}{inc_n} + \dots$
  - real effect might be smaller
- Or linear shift in baseline preferences: $V_{jn} = \delta_i + \Delta_j inc_n + \dots$
  - real effect might be non-linear

Including heterogeneity

Non-linear interactions

- For use with continuous attributes
- inc_n = income of respondent n
- $\overline{inc}$ = reference income for sample, e.g. average
- $V_{jn} = \dots + \beta_C \left(\frac{inc_n}{\overline{inc}}\right)^{\lambda_{inc,C}} C_{jn} + \dots$
- $\lambda_{inc,C}$ gives elasticity of cost sensitivity in relation to income
  - $\lambda_{inc,C} > 0$: increasing cost sensitivity
  - $\lambda_{inc,C} < 0$: decreasing cost sensitivity
  - $\lambda_{inc,C} = 0$: no interaction
- β_C gives cost sensitivity at mean income

Including heterogeneity

Normalisation with a continuous covariate

	Bus	Train
Travel time (T)	45 min	30 min
Travel cost (C)	£7	£12
Wifi	Available for £1	Available for free
Traveller age	43 years	

	Bus	Train
Travel time (T)	45 min	30 min
Travel cost (C)	£7	£12
Wifi	Available for £1	Available for free
Traveller age	65 years	

- Can estimate all continuous interactions
- But still need normalisations for interactions with categorial variables

$$\begin{aligned} V_{a,n} = {} & \delta_A + \Delta_{\delta_A} age_n \\ & + (\beta_T + \Delta_{\beta_T} age_n) T_A \\ & + (\beta_C + \Delta_{\beta_C} age_n) C_A \\ & + (\beta_{wifi_1} + \Delta_{\beta_{wifi_1}} age_n)(W_A == 1) + (\beta_{wifi_2} + \Delta_{\beta_{wifi_2}} age_n)(W_A == 2) \end{aligned}$$

Including heterogeneity

Normalisation with a categorical covariate

	Bus	Train
Travel time (T)	45 min	30 min
Travel cost (C)	£7	£12
Wifi	Available for £1	Available for free
Traveller gender	female	

	Bus	Train
Travel time (T)	45 min	30 min
Travel cost (C)	£7	£12
Wifi	Available for £1	Available for free
Traveller gender	male	

- Why is this the same as two separate models?
- Need normalisation for interactions too now

$$
\begin{aligned}
V_{a,n} = {} & \delta_A + \Delta_{\delta_A} female_n \\
& + (\beta_T + \Delta_{\beta_T} female_n) T_A \\
& + (\beta_C + \Delta_{\beta_C} female_n) C_A \\
& + (\beta_{wifi_1} + \Delta_{\beta_{wifi_1}} female_n)(W_A == 1) + (\beta_{wifi_2} + \Delta_{\beta_{wifi_2}} female_n)(W_A == 2)
\end{aligned}
$$

- With multiple covariates, we need to think about interactions between covariates too

Including heterogeneity

Misattributing deterministic heterogeneity

- Model estimation will use whatever flexibility we give it to try and improve fit
- Allowing heterogeneity in only some attributes risks lower fit and misattribution

Number of individuals	500		500		500		500	
Number of modelled outcomes	5000		5000		5000		5000	
Estimated parameters	2		3		3		4	
LL(final)	-2167.405		-2083.096		-2061.137		-2060.733	
Adj.Rho-square (0)	0.374		0.3981		0.4044		0.4042	
AIC	4338.81		4172.19		4128.27		4129.47	
BIC	4351.84		4191.74		4147.82		4155.54	
	estimate	Rob.t-ratio(0)	estimate	Rob.t-ratio(0)	estimate	Rob.t-ratio(0)	estimate	Rob.t-ratio(0)
efficacy	0.0199	11.78	0.0279	16.64	0.0202	14.1	0.0212	11.78
risk	-0.3651	-25.18	-0.3647	-26.37	-0.2958	-23.36	-0.3025	-20.12
efficacy interaction for men			-0.021	-12.79			-0.0029	-0.99
risk interaction for men					-0.2033	-13.33	-0.1816	-6.57

- Ignoring heterogeneity in β_k can lead to bias in β_l
  - Bad idea to keep the cost coefficient fixed!

Apollo

Scale heterogeneity

- Idea that the choices of some individuals are more deterministic than others (from the perspective of the analyst)
- Try and link this to observed characteristics

Including heterogeneity

Allowing for scale heterogeneity

- Example with male *vs* female:

$$P_n(i,n) = (male_n) \cdot \frac{e^{\mu_{male} V_{n,i}}}{\sum_{j=1}^{J} e^{\mu_{male} V_{n,j}}} + (female_n) \cdot \frac{e^{\mu_{female} V_{n,i}}}{\sum_{j=1}^{J} e^{\mu_{female} V_{n,j}}}$$

- Need to normalise one scale parameter
- Assume we set $\mu_{female} = 1$
  - estimate μ_{male}
  - if $\mu_{male} > 1$, modelled choice processes for men are more deterministic (opposite applies if $\mu_{male} < 1$)
  - scale of β parameters relates to part of the sample where the scale is normalised to 1, i.e. in our case women
  - while relative values of β parameters are influenced by both segments

Including heterogeneity

Scale heterogeneity: warning

- Impossible to fully disentangle scale heterogeneity from other heterogeneity
- For example, men might be more risk sensitive and risk might be most important attribute
  - Then scale heterogeneity might pick that up
- Safest approach is to allow for heterogeneity in all individual parameters

Including heterogeneity

Misattributing heterogeneity

- Allowing for scale heterogeneity but not heterogeneity in individual attributes
- Fake evidence of scale heterogeneity!

Number of individuals	500		500	
Number of modelled outcomes	5000		5000	
Estimated parameters	4		3	
LL(final)	-2060.733		-2077.792	
Adj.Rho-square (0)	0.4042		0.3996	
AIC	4129.47		4161.58	
BIC	4155.54		4181.14	
	estimate	Rob.t-ratio(0)	estimate	Rob.t-ratio(0)
efficacy	0.0212	11.78	0.0126	10.41
risk	-0.3025	-20.12	-0.255	-19.99
efficacy interaction for men	-0.0029	-0.99		
risk interaction for men	-0.1816	-6.57		
scale parameter for men			2.087	17.45

Including heterogeneity

Data requirements

- Capturing heterogeneity places additional demands on data
- Need rich data, with trade-offs that allow us to observe different outcomes for different people
- Heterogeneity in sensitivities alone may not lead to different outcomes if stimuli are not strong enough
- Focus on categorical variables in many fields, e.g., health studies, makes capturing heterogeneity especially hard
  - Proliferation of parameters

Within people *vs* between people effects

- May find that cost sensitivity is lower for higher income people
  - but data is generally cross-sectional
  - not clear that cost sensitivity would change in same way if a person's income increases
- Need longitudinal data in order to separate within-people from between-people effects

Interpretation with heterogeneity

Interpretation with heterogeneity

MRS with deterministic heterogeneity (categorical)

- Interaction with and gender and purpose

$$V_{in} = \ldots + \left(\beta_{tt} + \beta_{tt,female} \cdot z_{female,n}\right) TT_{in} \\ + \left(\beta_{cost} + \beta_{cost,business} \cdot z_{business,n}\right) C_{in}$$

- Trade-off becomes person-specific (4 groups)

$$VTT_n = \frac{\beta_{tt} + \beta_{tt,female} \cdot z_{female_n}}{\beta_{cost} + \beta_{cost,business} \cdot z_{business_n}}$$

```
Estimates:
                    Estimate
asc_car              0.00000
asc_bus             -2.36476
asc_air             -0.93027
asc_rail            -1.01880
b_tt                -0.01338
b_tt_female      -6.9615e-04
b_access            -0.01855
b_cost              -0.08524
b_cost_business      0.05522
b_no_frills          0.00000
b_wifi               0.98081
b_food               0.40660
```

```
VTT(in £/hr)
        Leisure Business
Male       9.42    26.76
Female     9.91    28.15
```

Interpretation with heterogeneity

MRS with deterministic heterogeneity (continuous)

- Additional interaction between income and cost sensitivity:

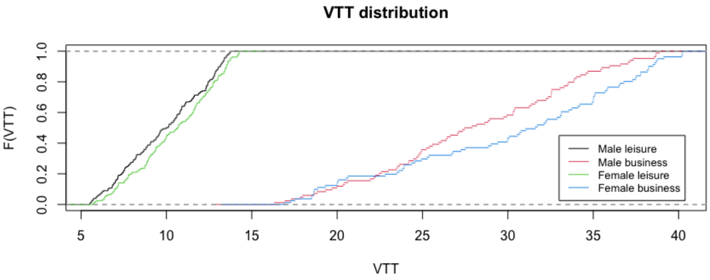
$$V_{in} = \ldots + \left(\beta_{tt} + \beta_{tt,female} \cdot z_{female,n}\right) TT_{in}$$
$$+ \left(\beta_{cost} + \beta_{cost,business} \cdot z_{business,n}\right) \cdot \left(\frac{income_n}{\overline{income}}\right)^{\lambda_{income}} \cdot C_{in}$$

```
Estimates:
                   Estimate
asc_car             0.00000
asc_bus            -2.57432
asc_air            -1.10547
asc_rail           -1.14259
b_tt               -0.01447
b_tt_female      -5.1631e-04
b_access           -0.01846
b_cost             -0.08501
b_cost_business     0.05491
b_no_frills         0.00000
b_wifi              1.01211
b_food              0.41897
lambda_income      -0.58221
```

Interpretation with heterogeneity

Trade-off becomes person-specific

$$VTT_n = \frac{\beta_{tt} + \beta_{tt,female} \cdot z_{female_n}}{\left(\beta_{cost} + \beta_{cost,business} \cdot z_{business_n}\right) \cdot \left(\frac{income_n}{\overline{income}}\right)^{\lambda_{income}}}$$



Summary

Summary

Key points from this class

- Incorporating heterogeneity can lead to gain in fit and insights, as well as bias reduction
- Model comparison is crucial part of model building

Suggested reading

- Train, K.E. (2009), Discrete Choice Methods with Simulation, Cambridge University Press, free online access `https://eml.berkeley.edu/books/choice2.html`
  - Chapters 3 and 8

Questions?

Apollo

www.ApolloChoiceModelling.com

The most flexible choice modelling software (up to a probability)