What If Some People Don't Care? Accommodating Low Importance Ratings for Qualitative Attributes in Our Models

Stephane Hess
Institute for Transport Studies, University of Leeds

Institute for Transport Studies
FACULTY OF ENVIRONMENT

UNIVERSITY OF LEEDS

Introduction

- Significant interest in attribute non-attendance, i.e. possibility that some respondents ignore given attributes in a survey
- Possibility that this phenomenon is restricted to SC surveys or indeed is caused by the surveys
- However, especially for qualitative attributes, there is a clear possibility that in reality, a substantial share of respondents have a zero valuation
- Not accommodating this in models may lead to misleading sample level estimates
  - in terms of heterogeneity but possibly even insignificant mean sensitivity

Methodology I

- First, let's revisit the methodology
- We have moved beyond relying on respondent reported processing strategies, as they can be unreliable and also put us at risk of endogeneity bias
- A somewhat standard approach is now to rely on latent class approaches with two classes for each coefficient, with one making use of zero sensitivity (the *"non-attendance"* class)
- Evidence in the literature shows significant shares of respondents with zero sensitivities
- This use of a latent class approach can lead to significant issues with confounding between non-attendance (i.e. zero sensitivity) and taste heterogeneity (especially very low sensitivity)
- Put forward a combined latent class - mixed logit approach that avoids this issue

Methodology II

- Existing work is based on using two classes for each coefficient within a latent class framework
  - one class with β_k estimated
  - one class with β_k fixed to zero
- With K coefficients, we have 2^K classes (all possible combinations)
- Likelihood of observed sequence of choices for respondent n

$$L\left(y_n, \beta, \pi\right) = \sum_{s=1}^{S} \pi_s \prod_{t=1}^{T} P\left(i^*_{nt} \mid \beta_s\right)$$

  where respondent n faces T choices, $\pi = \langle \pi_1, \ldots, \pi_S \rangle$, $\beta = \langle \beta_1, \ldots, \beta_S \rangle$, $S = 2^K$, and $\beta_s = \langle \beta_{s1}, \ldots, \beta_{sK} \rangle$
- Each class (and thus realisation of β_s) uses a given combination of estimated coefficients and coefficients fixed to zero

Methodology III

- Problem with this approach:
  - class at zero is likely to capture not just respondents with zero sensitivity but also respondents with low sensitivity
  - even worse, the split into two classes for each coefficient potentially just captures random heterogeneity that could equally well (or better) be captured by freely estimating the coefficients in both classes
  - estimating both coefficient values freely will invariably lead to better fit (more flexible model), but will prevent us from studying non-attendance
  - solution put forward here is to make use of a combined latent class - mixed logit model

Methodology IV

- Still use two classes for each coefficient within a latent class framework
  - one class with β_k fixed to zero, as before
  - second class estimates a random distribution for β_k
- Likelihood of observed sequence of choices for respondent n

$$L\left(y_n, \Omega, \pi\right) = \sum_{s=1}^{S} \pi_s \int_{\beta_s} \prod_{t=1}^{T} P\left(i_{nt}^{*} \mid \beta_s\right) f\left(\beta_s \mid \Omega\right) \mathrm{d}\beta_s$$

- The vector β_s still contains zero and non-zero elements, but the non-zero elements now follow a random distribution ($f\left(\beta_s \mid \Omega\right)$)
- Actual distribution for non-zero elements in β_s is kept constant across classes, but specific mix of zero and non-zero elements still varies across classes

Methodology V

- Advantage of this approach is that it will allow some (but not necessarily all) of the random heterogeneity to be captured in the randomly distributed non-zero elements within β_s
- Should reduce risk of the class with imposed zero sensitivity simply capturing heterogeneity and low sensitivity
- There is still a risk that this happens, of course

Data

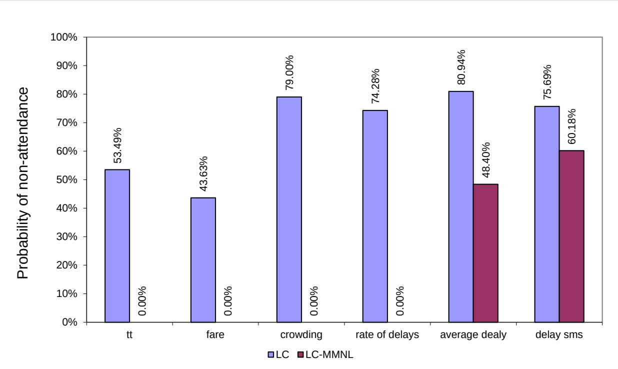
- Survey on rail and bus commuters, collected online in the UK in early 2010
- $3,680$ observations from 368 respondents
- Each scenario involves the choice between three alternatives, of which the first is a reference trip (invariant across 10 tasks)
- Six attributes
  - Travel time (minutes)
  - Fare (£)
  - Crowding (out of ten trips)
  - Rate of delays (out of ten trips)
  - Average length of delays (across delayed trips)
  - Provision of free information delay service (sms)

Model fit

	MNL	LC	MMNL	LC-MMNL
LL(0)	-4,042.89	-4,042.89	-4,042.89	-4,042.89
LL	-3,700.43	-3,161.53	-2,996.05	-2,985.82
par	8	14	14	16
adj. ρ^2	0.0827	0.2145	0.2555	0.2575

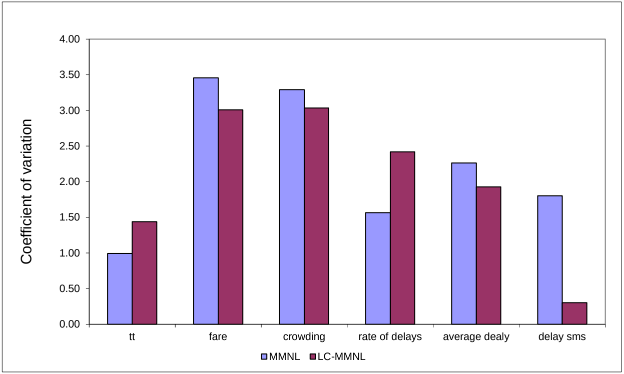
- LC model improves over MNL
- But MMNL improvement is more substantial
- LC-MMNL outperforms all other models

Class probabilities



- Significant reduction in probabilities for non-attendance classes in LC-MMNL model (four of them to zero!)
- Direct result of allowing model to capture random heterogeneity
- Non-zero probability remains for average delay and for delay

Heterogeneity in MMNL class



- Reductions for atts with positive weight for non-attendance - some of that heterogeneity is now captured by n-a classes
- Also some changes for other attributes - heterogeneity patterns are different, possibly due to increased model flexibility

Data

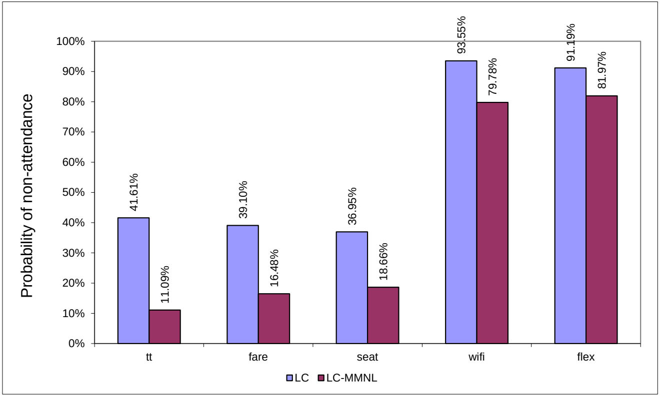
- Survey with rail commuters, from online panel collected in the UK, early 2010
- $7,968$ observations collected from 996 respondents, with 8 choices each
- Choice between 3 alternatives with attribute-levels pivoted around reference values, but no reference alternative presented
- Five attributes
  - Travel time (minutes)
  - Fare (£)
  - Guarantee of a reserved seat (yes/no dummy)
  - Provision of free wifi (yes/no dummy)
  - Ticket rebooking flexibility (yes/no dummy)

Model fit

	MNL	LC	MMNL	LC-MMNL
LL(0)	-8,753.74	-8,753.74	-8,753.74	-8,753.74
LL	-6,288.87	-5,461.53	-5,296.37	-5,250.43
par	7	14	14	17
adj. ρ^2	0.2808	0.3745	0.3934	0.3983

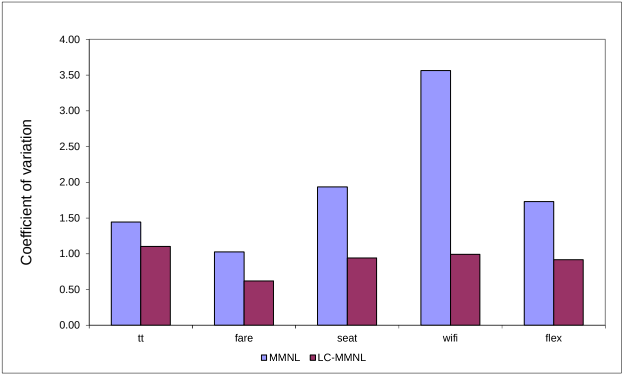
- LC model improves over MNL
- But MMNL improvement is more substantial
- LC-MMNL outperforms all other models
    - ... does this look familiar?

Class probabilities



- Significant reduction in probabilities for non-attendance classes in LC-MMNL model, especially for *important* attributes
- But remains high for two of the qualitative attributes

Heterogeneity in MMNL class



- Reductions in heterogeneity in MMNL class - some of that heterogeneity is now captured in n-a classes

Data

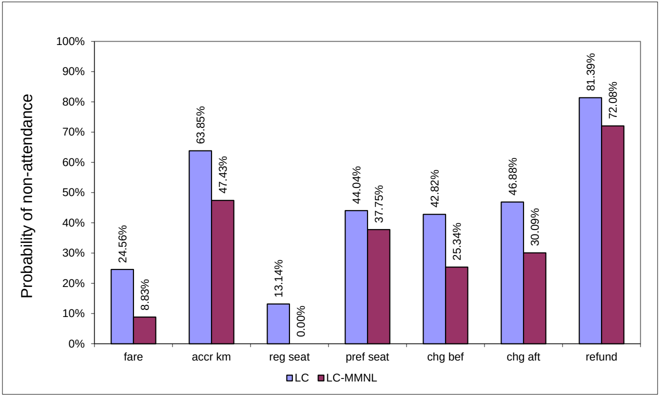
- Web-based survey with LAN (a Latin American carrier) air travellers, collected in 2009
- $9,150$ observations collected from 915 respondents
- Choice between 3 flight options
- Six attributes
  - Fare (Chilean pesos)
  - Accumulation of LANPASS kms (kms)
  - Seat reservation (no seat reservation/reservation (exc. preferential seats)/reservation (inc. preferential seats) dummy)
  - Reservation change before departure date (yes/no dummy)
  - Reservation change after departure date (yes/no dummy)
  - Ticket cancellation (yes/no dummy)

Model fit

	MNL	LC	MMNL	LC-MMNL
LL(0)	-10,711.47	-10,711.47	-10,711.47	-10,711.47
LL	-7,483.52	-6,875.05	-6,733.15	-6,681.96
par	9	16	16	22
adj. ρ^2	0.3005	0.3567	0.3699	0.3741

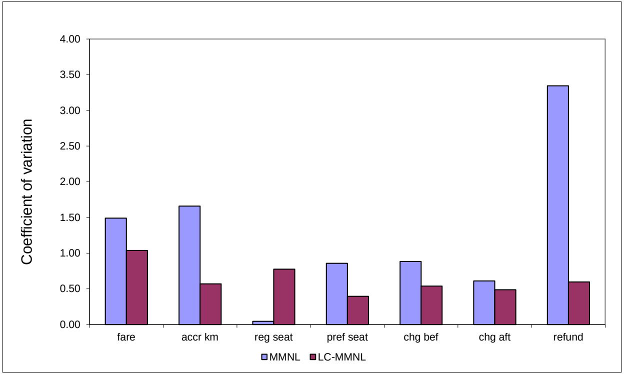
- LC model improves over MNL
- But MMNL improvement is more substantial
- LC-MMNL outperforms all other models
  - ... should be very familiar by now!

Class probabilities



- Significant reduction in probabilities for non-attendance classes in LC-MMNL model (one of them to zero!)
- But remain non-trivial for qualitative attributes

Heterogeneity in MMNL class



- Reductions for atts with non-zero weight for non-attendance - some of that heterogeneity is now captured in n-a classes
- Also some changes for reg seat, so overall patterns of heterogeneity are different

Overall conclusions I

- Substantial interest in modelling attribute non-attendance
- Should be recognised that estimates may not be caused by actual non-attendance, but by design not including sufficient scenarios in which the given attribute can influence the choice
  - especially the case for less important attributes
- This presentation has focussed on ways of allowing for non-attendance in our models
- Have shown that widely used latent class approach may be affected by confounding
- Significant reductions in shares for non-attendance classes when additionally allowing for random heterogeneity within non-zero classes
- But, rates of inferred non-attendance stay high for some of the qualitative attributes

Overall conclusions II

- Paper investigates these effects in more detail
- Results also show that:
  - in MNL models, a substantial rate of non-attendance leads to the obvious downward trend in the estimated mean sensitivity, to the point of becoming insignificant
  - in MMNL models, both mean and variance are affected
- Not recognising that qualitative attributes only matter to some individuals may yield misleading results for sample level estimates

Questions . . .