# Incorporating scale heterogeneity in Latent class analysis: site choice applications

Marco Boeri Edel Doherty Danny Campbell Alberto Longo

Gibson Institute for Land, Food and Environment and
Institute for a Sustainable World
Queen's University Belfast

Workshop on Discrete Choice Models 2011, EPFL Lausanne

# Presentation Outline

# Presentation Outline

# Introduction

- The mixed logit model (ML) is one of the central advances (through accommodating unobserved heterogeneity) in choice modelling.
- Types of heterogeneity (observed and unobserved) are:
  - heterogeneity in preferences;
  - heterogeneity in scale (heteroscedasticity).
- Most of the analysis with ML to date incorporates either preference heterogeneity or heteroscedasticity
- However *“Addressing only one source of heterogeneity negates the fact that true choice behaviour is likely to be in some middle ground with some variation in scale and some in taste”* (Thiene and Scarpa, 2010).

# Introduction

- The mixed logit model (ML) is one of the central advances (through accommodating unobserved heterogeneity) in choice modelling.
- Types of heterogeneity (observed and unobserved) are:
  - heterogeneity in preferences;
  - heterogeneity in scale (heteroscedasticity).
- Most of the analysis with ML to date incorporates either preference heterogeneity or heteroscedasticity
- However *"Addressing only one source of heterogeneity negates the fact that true choice behaviour is likely to be in some middle ground with some variation in scale and some in taste"* (Thiene and Scarpa, 2010).

# Introduction

- The mixed logit model (ML) is one of the central advances (through accommodating unobserved heterogeneity) in choice modelling.
- Types of heterogeneity (observed and unobserved) are:
  - heterogeneity in preferences;
  - heterogeneity in scale (heteroscedasticity).
- Most of the analysis with ML to date incorporates either preference heterogeneity or heteroscedasticity
- However *"Addressing only one source of heterogeneity negates the fact that true choice behaviour is likely to be in some middle ground with some variation in scale and some in taste"* (Thiene and Scarpa, 2010).

# Introduction

- Aim of the study:
  - Compare different approaches to include scale heterogeneity in LC and RPL models in WTP-space.
- We build on and combine:
  - The heteroscedastic multinomial logit model (Swait and Adamowicz, 2001; Swait, 2006);
  - The scale-adjusted Latent class (Magidson and Vermunt 2007; Hensher et al 2011);
  - The discrete mixtures of continuous distributions (Bujosa et al 2010; Greene and Hensher 2010).
- Formalising the heteroscedastic mixed logit (HML) model.

# Introduction

- Aim of the study:
  - Compare different approaches to include scale heterogeneity in LC and RPL models in WTP-space.
- We build on and combine:
  - The heteroscedastic multinomial logit model (Swait and Adamowicz, 2001; Swait, 2006);
  - The scale-adjusted Latent class (Magidson and Vermunt 2007; Hensher et al 2011);
  - The discrete mixtures of continuous distributions (Bujosa et al 2010; Greene and Hensher 2010).
- Formalising the heteroscedastic mixed logit (HML) model.

# Introduction

- Aim of the study:
  - Compare different approaches to include scale heterogeneity in LC and RPL models in WTP-space.
- We build on and combine:
  - The heteroscedastic multinomial logit model (Swait and Adamowicz, 2001; Swait, 2006);
  - The scale-adjusted Latent class (Magidson and Vermunt 2007; Hensher et al 2011);
  - The discrete mixtures of continuous distributions (Bujosa et al 2010; Greene and Hensher 2010).
- Formalising the heteroscedastic mixed logit (HML) model.

# Presentation Outline

# Methodology

Starting from the generic Utility function (RUM - McFadden, 1974)

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

We reparameterise $V(\beta, x_{nit})$ in WTP-space (Train and Weeks, 2005):

$$V_{nit}(\alpha, w, x_{nit}) = -(\lambda\alpha)\, p_{nit} + (\lambda\alpha w)' X_{nit}$$

Why WTP-space?

# Methodology

Starting from the generic Utility function (RUM - McFadden, 1974)

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

We reparameterise $V(\beta, x_{nit})$ in WTP-space (Train and Weeks, 2005):

$$V_{nit}(\alpha, w, x_{nit}) = -(\lambda\alpha)\, p_{nit} + (\lambda\alpha w)' X_{nit}$$

Why WTP-space?

# Rational for working in WTP-space

- Adopting continuous random parameter, the spread of the WTP distributions can be directly tested (Thiene and Scarpa, 2009);
- The welfare results are reported directly in the models;
- Importantly, in this paper, WTP-space allows us:
  - to directly compare estimates from continuous and discrete mixture representations;
  - to demonstrate the importance of including the scale parameter even if utility is parameterised in WTP-space.

Indeed, under this specification the WTP estimates are scale free, however the model itself is not.

# Rational for working in WTP-space

- Adopting continuous random parameter, the spread of the WTP distributions can be directly tested (Thiene and Scarpa, 2009);
- The welfare results are reported directly in the models;
- Importantly, in this paper, WTP-space allows us:
  - to directly compare estimates from continuous and discrete mixture representations;
  - to demonstrate the importance of including the scale parameter even if utility is parameterised in WTP-space.

Indeed, under this specification the WTP estimates are scale free, however the model itself is not.

# Rational for working in WTP-space

- Adopting continuous random parameter, the spread of the WTP distributions can be directly tested (Thiene and Scarpa, 2009);
- The welfare results are reported directly in the models;
- Importantly, in this paper, WTP-space allows us:
  - to directly compare estimates from continuous and discrete mixture representations;
  - to demonstrate the importance of including the scale parameter even if utility is parameterised in WTP-space.

Indeed, under this specification the WTP estimates are scale free, however the model itself is not.

# Rational for working in WTP-space

- Adopting continuous random parameter, the spread of the WTP distributions can be directly tested (Thiene and Scarpa, 2009);
- The welfare results are reported directly in the models;
- Importantly, in this paper, WTP-space allows us:
  - to directly compare estimates from continuous and discrete mixture representations;
  - to demonstrate the importance of including the scale parameter even if utility is parameterised in WTP-space.

Indeed, under this specification the WTP estimates are scale free, however the model itself is not.

# Heteroscedastic Multinomial Logit Model

$$L\left(y_n|Z_n, p_n, x_n\right) = \prod_{t=1}^{T_n} \frac{\exp\left[\mu_{ni}\left(Z_{ni}|\lambda_{ni}\right) \cdot V_{ni}\left(p_{ni}, X_{ni}|\alpha, w\right)\right]}{\sum_{j=1}^{J} \exp\left[\mu_{nj}\left(Z_{nj}|\lambda_{ni}\right) \cdot V_{nj}\left(p_{nj}, X_{nj}|\alpha, w\right)\right]}$$

- Scale heterogeneity can be:
  - observed - discrete (Swait and Adamowicz, 2001);
  - unobserved - continuous (Breffle and Morey, 2000).
- Preference homogeneity.

Heterogeneity in taste can be introduced with latent class analysis, as well as with RPL models, in WTP-space.

# Heteroscedastic Multinomial Logit Model

$$L\left(y_n|Z_n, p_n, x_n\right) = \prod_{t=1}^{T_n} \frac{\exp\left[\mu_{ni}\left(Z_{ni}|\lambda_{ni}\right) \cdot V_{ni}\left(p_{ni}, X_{ni}|\alpha, w\right)\right]}{\sum_{j=1}^{J} \exp\left[\mu_{nj}\left(Z_{nj}|\lambda_{ni}\right) \cdot V_{nj}\left(p_{nj}, X_{nj}|\alpha, w\right)\right]}$$

- Scale heterogeneity can be:
  - observed - discrete (Swait and Adamowicz, 2001);
  - unobserved - continuous (Breffle and Morey, 2000).
- Preference homogeneity.

Heterogeneity in taste can be introduced with latent class analysis, as well as with RPL models, in WTP-space.

# Heteroscedastic Mixed Logit models

Starting from the HMNL we describe and compare different Heteroscedastic Mixed Logit (HML) models:

$$Prob_{nit} = \prod_{t=1}^{T_n} \int L_{nit}\left(\kappa, \phi, \vartheta\right) f\left(\vartheta|\theta\right) \mathrm{d}\vartheta,$$

$L_{nit}$ is the HMNL where preferences that **can** be heterogeneous.

- $\kappa$ represents fixed parameters across the ML (homogeneity)
- $\phi$ represents the observed heterogeneity (taste, scale or both)
- $\vartheta$ represents the unobserved heterogeneity (taste, scale or both)
- $\theta$ describes the density function $f(\vartheta)$.

# Heteroscedastic Mixed Logit models

Starting from the HMNL we describe and compare different Heteroscedastic Mixed Logit (HML) models:

$$Prob_{nit} = \prod_{t=1}^{T_n} \int L_{nit}\left(\kappa, \phi, \vartheta\right) f\left(\vartheta|\theta\right) \mathrm{d}\vartheta,$$

$L_{nit}$ is the HMNL where preferences that **can** be heterogeneous.

- $\kappa$ represents fixed parameters across the ML (homogeneity)
- $\phi$ represents the observed heterogeneity (taste, scale or both)
- $\vartheta$ represents the unobserved heterogeneity (taste, scale or both)
- $\theta$ describes the density function $f(\vartheta)$.

# HML: observed scale heterogeneity

- Heterogeneity linked to socio-economic characteristics:

$$\mu_g\left(Z_{n|g}|\lambda_g\right) = 1 + \sum_{g=1}^{G} \gamma_g \eta_g$$

- Heterogeneity associated to characteristic of choice situation:

$$\mu_{cs}\left(Z_{n|cs}|\lambda_{cs}\right) = 1 + \sum_{cs=1}^{CS} \gamma_{cs} \eta_{cs}$$

where $\gamma$ is a dummy variable representing each group.

# HML: unobserved scale heterogeneity

- Heterogeneity probabilistically described by a discrete mixing distribution:

$$\mu_s(\lambda_s) = 1 + \sum_{s=1}^{S} \pi_s \eta_s$$

where $\sum_{s=1}^{S} \pi_s = 1$ and $\pi_s > 0 \forall s$

- Heterogeneity probabilistically described by a continuous distribution:

$$\mu_n(\lambda_{ni}) = \int \lambda f(\lambda) \mathrm{d}\lambda.$$

# HML: our analysis

We compare models for representing observed or unobserved scale heterogeneity and unobserved taste heterogeneity in WTP-space:

- RPL model;
- 3 LC model specifications:
  - unobs. scale adjusted LC;
  - obs. scale adjusted LC;
  - discrete mixing (for WTP) of continuous distributions (cost-scale).

# HML: our analysis

- Unobs. scale adjusted LC

$$\Pr\left(y_n|p_n,x_n\right)=\sum_{c=1}^{C}\pi_c\left[\prod_{t=1}^{T_n}\left(\sum_{s=1}^{S}\pi_{s|c}\frac{\exp\left[\mu_{s|c}\left(\lambda_{s|c}\right)\cdot V_{cni}\left(p_{ni},X_{ni}|\alpha_c,w_c\right)\right]}{\sum_{j=1}^{J}\exp\left[\mu_{s|c}\left(\lambda_{s|c}\right)\cdot V_{cnj}\left(p_{nj},X_{nj}|\alpha_c,w_c\right)\right]}\right)\right]$$

- Obs. scale adjusted LC:

$$\Pr\left(y_n|p_n,x_n\right)=\sum_{c=1}^{C}\pi_c\left[\prod_{t=1}^{T_n}\left(\frac{\exp\left[\mu_{g|c}\left(Z_{g|c}|\lambda_{g|c}\right)\cdot V_{cni}\left(p_{ni},X_{ni}|\alpha_c,w_c\right)\right]}{\sum_{j=1}^{J}\exp\left[\mu_{g|c}\left(Z_{g|c}|\lambda_{g|c}\right)\cdot V_{cnj}\left(p_{nj},X_{nj}|\alpha_c,w_c\right)\right]}\right)\right]$$

# HML: our analysis

- RPL model in WTP-space:

$$\Pr\left(y_n|\Omega, p_n, x_n\right) = \int_{(\phi)} \prod_{t=1}^{T_n} \frac{\exp\left(-\left(\beta_{\lambda\alpha}\right) p_{nit} + \left(\beta_{\lambda\alpha} w\right)' x_{nit}\right)}{\sum_{j=1}^{J} \exp\left(-\left(\beta_{\lambda\alpha}\right) p_{njt} + \left(\beta_{\lambda\alpha} w\right)' x_{njt}\right)} f\left(\phi|\, \Omega\right) d\phi$$

- Discrete mixing (for preferences) of continuous distributions (cost-scale)

$$\Pr\left(y_n|p_n, x_n\right) = \sum_{c=1}^{C} \pi_c \left[\prod_{t=1}^{T_n} \int_{\beta_c} \left(\frac{\exp\left[-\left(\beta_c\right) p_{nit} + \left(\beta_c w_c\right)' x_{nit}\right]}{\sum_{j=1}^{J} \exp\left[-\left(\beta_c\right) p_{njt} + \left(\beta_c w_c\right)' x_{njt}\right]}\right)\right] f\left(\beta_c|\, \Omega_c\right) d\beta_c$$

where $\beta_c = \beta_{\lambda\alpha|c}$ (cost and scale are perfectly correlated).

# Presentation Outline

Introduction

Methodology
- Rational for working in WTP-space
- Heteroscedastic Multinomial Logit Model
- Heteroscedastic Mixed Logit models

Results
- The case study
- Comparing the models
- Validations

Conclusions

# The case study

To compare the HML models we used stated preference (SP) data on walkers site choice in Ireland

- The data is from a discrete choice experiment aimed at eliciting values for recreational walks on farmland in Ireland.
- Data collected from a representative sample of the Irish adult population
- A total of 5,460 observations from 470 respondents used for the analysis.

# Example of choice card

| | Bog walk | Field walk | River walk | Choose none |
|---|---|---|---|---|
| Length | 3-4 hours | 1-2 hours | 1-2 hours | |
| Car Park | Yes | No | No | I would not choose any of the walks. I would stay at home. |
| Fence | ⊢ | No | Yes | |
| Path and Signage | Yes | No | Yes | |
| Distance | 40 km | 10 km | 80 km | |

# Comparing the models

| Model | LogLik. | K | $\bar{\rho}^2$ | $\chi^2$ | BIC | AIC | 3AIC | crAIC |
|---|---|---|---|---|---|---|---|---|
| MNL | -6882.982 | 9 | 0.101 | 1561.084 | 13843.703 | 13783.964 | 13792.964 | 13784.316 |
| HMNL | -6863.299 | 10 | 0.103 | 1600.450 | 13812.974 | 13746.598 | 13756.598 | 13747.067 |
| HML | -6486.8 | 10 | 0.152 | 2353.448 | 13059.976 | 12993.600 | 13003.600 | 12994.069 |
| RPL - Cost fixed | -6247.31 | 13 | 0.183 | 2832.428 | 12606.909 | 12520.620 | 12533.620 | 12521.591 |
| 2 cl. LC | -6181.596 | 15 | 0.191 | 2963.856 | 12492.757 | 12393.192 | 12408.192 | 12394.643 |
| 3 cl. LC | -5836.88 | 21 | 0.236 | 3653.288 | 11855.150 | 11715.760 | 11736.760 | 11719.544 |
| 4 cl. LC | -5796.234 | 27 | 0.240 | 3734.580 | 11825.684 | 11646.468 | 11673.468 | 11654.283 |
| 2 cl. obsHLC | -6180.8 | 17 | 0.191 | 2965.448 | 12508.440 | 12395.600 | 12412.600 | 12397.669 |
| 3 cl. obsHLC | -5825.869 | 24 | 0.237 | 3675.310 | 11859.041 | 11699.738 | 11723.738 | 11705.296 |
| 4 cl. obsHLC | -5702.303 | 31 | 0.252 | 3922.442 | 11672.373 | 11466.606 | 11497.606 | 11478.283 |
| 2 cl. probHLC | -6030.245 | 19 | 0.211 | 3266.558 | 12224.605 | 12098.490 | 12117.490 | 12101.330 |
| 3 cl. probHLC | -5799.313 | 27 | 0.240 | 3728.422 | 11831.842 | 11652.626 | 11679.626 | 11660.441 |
| 4 cl. probHLC* | -5690.141 | 35 | 0.253 | 3946.766 | 11682.599 | 11450.282 | 11485.282 | 11466.923 |
| RPL | -6034.622 | 14 | 0.211 | 3257.804 | 12190.171 | 12097.244 | 12111.244 | 12098.439 |
| 2 cl. MLC | -6068.022 | 17 | 0.206 | 3191.004 | 12282.884 | 12170.044 | 12187.044 | 12172.113 |
| 3 cl. MLC | -5802.464 | 24 | 0.240 | 3722.120 | 11812.231 | 11652.928 | 11676.928 | 11658.486 |
| 4 cl. MLC* | -5762.685 | 31 | 0.244 | 3801.678 | 11793.137 | 11587.370 | 11618.370 | 11599.047 |

Note: Observed heteroscedasticity is retrieved based on respondents from Rural and Urban (Baseline) areas
* denotes an unidentified model.

# Comparing the models

| Model | LogLik. | K | $\bar{\rho}^2$ | $\chi^2$ | BIC | AIC | 3AIC | crAIC |
|---|---|---|---|---|---|---|---|---|
| MNL | -6882.982 | 9 | 0.101 | 1561.084 | 13843.703 | 13783.964 | 13792.964 | 13784.316 |
| HMNL | -6863.299 | 10 | 0.103 | 1600.450 | 13812.974 | 13746.598 | 13756.598 | 13747.067 |
| HML | -6486.8 | 10 | 0.152 | 2353.448 | 13059.976 | 12993.600 | 13003.600 | 12994.069 |
| RPL - Cost fixed | -6247.31 | 13 | 0.183 | 2832.428 | 12606.909 | 12520.620 | 12533.620 | 12521.591 |
| 2 cl. LC | -6181.596 | 15 | 0.191 | 2963.856 | 12492.757 | 12393.192 | 12408.192 | 12394.643 |
| 3 cl. LC | -5836.88 | 21 | 0.236 | 3653.288 | 11855.150 | 11715.760 | 11736.760 | 11719.544 |
| 4 cl. LC | -5796.234 | 27 | 0.240 | 3734.580 | 11825.684 | 11646.468 | 11673.468 | 11654.283 |
| 2 cl. obsHLC | -6180.8 | 17 | 0.191 | 2965.448 | 12508.440 | 12395.600 | 12412.600 | 12397.669 |
| 3 cl. obsHLC | -5825.869 | 24 | 0.237 | 3675.310 | 11859.041 | 11699.738 | 11723.738 | 11705.296 |
| 4 cl. obsHLC | -5702.303 | 31 | 0.252 | 3922.442 | 11672.373 | 11466.606 | 11497.606 | 11478.283 |
| 2 cl. probHLC | -6030.245 | 19 | 0.211 | 3266.558 | 12224.605 | 12098.490 | 12117.490 | 12101.330 |
| 3 cl. probHLC | -5799.313 | 27 | 0.240 | 3728.422 | 11831.842 | 11652.626 | 11679.626 | 11660.441 |
| 4 cl. probHLC* | -5690.141 | 35 | 0.253 | 3946.766 | 11682.599 | 11450.282 | 11485.282 | 11466.923 |
| RPL | -6034.622 | 14 | 0.211 | 3257.804 | 12190.171 | 12097.244 | 12111.244 | 12098.439 |
| 2 cl. MLC | -6068.022 | 17 | 0.206 | 3191.004 | 12282.884 | 12170.044 | 12187.044 | 12172.113 |
| 3 cl. MLC | -5802.464 | 24 | 0.240 | 3722.120 | 11812.231 | 11652.928 | 11676.928 | 11658.486 |
| 4 cl. MLC* | -5762.685 | 31 | 0.244 | 3801.678 | 11793.137 | 11587.370 | 11618.370 | 11599.047 |

Note: Observed heteroscedasticity is retrieved based on respondents from Rural and Urban (Baseline) areas
* denotes an unidentified model.

# Comparing the models

| Model | LogLik. | K | $\bar{\rho}^2$ | $\chi^2$ | BIC | AIC | 3AIC | crAIC |
|---|---|---|---|---|---|---|---|---|
| MNL | -6882.982 | 9 | 0.101 | 1561.084 | 13843.703 | 13783.964 | 13792.964 | 13784.316 |
| HMNL | -6863.299 | 10 | 0.103 | 1600.450 | 13812.974 | 13746.598 | 13756.598 | 13747.067 |
| HML | -6486.8 | 10 | 0.152 | 2353.448 | 13059.976 | 12993.600 | 13003.600 | 12994.069 |
| RPL - Cost fixed | -6247.31 | 13 | 0.183 | 2832.428 | 12606.909 | 12520.620 | 12533.620 | 12521.591 |
| 2 cl. LC | -6181.596 | 15 | 0.191 | 2963.856 | 12492.757 | 12393.192 | 12408.192 | 12394.643 |
| 3 cl. LC | -5836.88 | 21 | 0.236 | 3653.288 | 11855.150 | 11715.760 | 11736.760 | 11719.544 |
| 4 cl. LC | -5796.234 | 27 | 0.240 | 3734.580 | 11825.684 | 11646.468 | 11673.468 | 11654.283 |
| 2 cl. obsHLC | -6180.8 | 17 | 0.191 | 2965.448 | 12508.440 | 12395.600 | 12412.600 | 12397.669 |
| 3 cl. obsHLC | -5825.869 | 24 | 0.237 | 3675.310 | 11859.041 | 11699.738 | 11723.738 | 11705.296 |
| 4 cl. obsHLC | -5702.303 | 31 | 0.252 | 3922.442 | 11672.373 | 11466.606 | 11497.606 | 11478.283 |
| 2 cl. probHLC | -6030.245 | 19 | 0.211 | 3266.558 | 12224.605 | 12098.490 | 12117.490 | 12101.330 |
| 3 cl. probHLC | -5799.313 | 27 | 0.240 | 3728.422 | 11831.842 | 11652.626 | 11679.626 | 11660.441 |
| 4 cl. probHLC* | -5690.141 | 35 | 0.253 | 3946.766 | 11682.599 | 11450.282 | 11485.282 | 11466.923 |
| RPL | -6034.622 | 14 | 0.211 | 3257.804 | 12190.171 | 12097.244 | 12111.244 | 12098.439 |
| 2 cl. MLC | -6068.022 | 17 | 0.206 | 3191.004 | 12282.884 | 12170.044 | 12187.044 | 12172.113 |
| 3 cl. MLC | -5802.464 | 24 | 0.240 | 3722.120 | 11812.231 | 11652.928 | 11676.928 | 11658.486 |
| 4 cl. MLC* | -5762.685 | 31 | 0.244 | 3801.678 | 11793.137 | 11587.370 | 11618.370 | 11599.047 |

Note: Observed heteroscedasticity is retrieved based on respondents from Rural and Urban (Baseline) areas
* denotes an unidentified model.

# Comparing the models

| Model | LogLik. | K | $\bar{\rho}^2$ | $\chi^2$ | BIC | AIC | 3AIC | crAIC |
|---|---|---|---|---|---|---|---|---|
| MNL | -6882.982 | 9 | 0.101 | 1561.084 | 13843.703 | 13783.964 | 13792.964 | 13784.316 |
| HMNL | -6863.299 | 10 | 0.103 | 1600.450 | 13812.974 | 13746.598 | 13756.598 | 13747.067 |
| HML | -6486.8 | 10 | 0.152 | 2353.448 | 13059.976 | 12993.600 | 13003.600 | 12994.069 |
| RPL - Cost fixed | -6247.31 | 13 | 0.183 | 2832.428 | 12606.909 | 12520.620 | 12533.620 | 12521.591 |
| 2 cl. LC | -6181.596 | 15 | 0.191 | 2963.856 | 12492.757 | 12393.192 | 12408.192 | 12394.643 |
| 3 cl. LC | -5836.88 | 21 | 0.236 | 3653.288 | 11855.150 | 11715.760 | 11736.760 | 11719.544 |
| 4 cl. LC | -5796.234 | 27 | 0.240 | 3734.580 | 11825.684 | 11646.468 | 11673.468 | 11654.283 |
| 2 cl. obsHLC | -6180.8 | 17 | 0.191 | 2965.448 | 12508.440 | 12395.600 | 12412.600 | 12397.669 |
| 3 cl. obsHLC | -5825.869 | 24 | 0.237 | 3675.310 | 11859.041 | 11699.738 | 11723.738 | 11705.296 |
| 4 cl. obsHLC | -5702.303 | 31 | 0.252 | 3922.442 | 11672.373 | 11466.606 | 11497.606 | 11478.283 |
| 2 cl. probHLC | -6030.245 | 19 | 0.211 | 3266.558 | 12224.605 | 12098.490 | 12117.490 | 12101.330 |
| 3 cl. probHLC | -5799.313 | 27 | 0.240 | 3728.422 | 11831.842 | 11652.626 | 11679.626 | 11660.441 |
| 4 cl. probHLC* | -5690.141 | 35 | 0.253 | 3946.766 | 11682.599 | 11450.282 | 11485.282 | 11466.923 |
| RPL | -6034.622 | 14 | 0.211 | 3257.804 | 12190.171 | 12097.244 | 12111.244 | 12098.439 |
| 2 cl. MLC | -6068.022 | 17 | 0.206 | 3191.004 | 12282.884 | 12170.044 | 12187.044 | 12172.113 |
| 3 cl. MLC | -5802.464 | 24 | 0.240 | 3722.120 | 11812.231 | 11652.928 | 11676.928 | 11658.486 |
| 4 cl. MLC* | -5762.685 | 31 | 0.244 | 3801.678 | 11793.137 | 11587.370 | 11618.370 | 11599.047 |

Note: Observed heteroscedasticity is retrieved based on respondents from Rural and Urban (Baseline) areas
* denotes an unidentified model.

# Comparing the models

| Model | LogLik. | K | $\bar{\rho}^2$ | $\chi^2$ | BIC | AIC | 3AIC | crAIC |
|---|---|---|---|---|---|---|---|---|
| MNL | -6882.982 | 9 | 0.101 | 1561.084 | 13843.703 | 13783.964 | 13792.964 | 13784.316 |
| HMNL | -6863.299 | 10 | 0.103 | 1600.450 | 13812.974 | 13746.598 | 13756.598 | 13747.067 |
| HML | -6486.8 | 10 | 0.152 | 2353.448 | 13059.976 | 12993.600 | 13003.600 | 12994.069 |
| RPL - Cost fixed | -6247.31 | 13 | 0.183 | 2832.428 | 12606.909 | 12520.620 | 12533.620 | 12521.591 |
| 2 cl. LC | -6181.596 | 15 | 0.191 | 2963.856 | 12492.757 | 12393.192 | 12408.192 | 12394.643 |
| 3 cl. LC | -5836.88 | 21 | 0.236 | 3653.288 | 11855.150 | 11715.760 | 11736.760 | 11719.544 |
| 4 cl. LC | -5796.234 | 27 | 0.240 | 3734.580 | 11825.684 | 11646.468 | 11673.468 | 11654.283 |
| 2 cl. obsHLC | -6180.8 | 17 | 0.191 | 2965.448 | 12508.440 | 12395.600 | 12412.600 | 12397.669 |
| 3 cl. obsHLC | -5825.869 | 24 | 0.237 | 3675.310 | 11859.041 | 11699.738 | 11723.738 | 11705.296 |
| 4 cl. obsHLC | -5702.303 | 31 | 0.252 | 3922.442 | 11672.373 | 11466.606 | 11497.606 | 11478.283 |
| 2 cl. probHLC | -6030.245 | 19 | 0.211 | 3266.558 | 12224.605 | 12098.490 | 12117.490 | 12101.330 |
| 3 cl. probHLC | -5799.313 | 27 | 0.240 | 3728.422 | 11831.842 | 11652.626 | 11679.626 | 11660.441 |
| 4 cl. probHLC* | -5690.141 | 35 | 0.253 | 3946.766 | 11682.599 | 11450.282 | 11485.282 | 11466.923 |
| RPL | -6034.622 | 14 | 0.211 | 3257.804 | 12190.171 | 12097.244 | 12111.244 | 12098.439 |
| 2 cl. MLC | -6068.022 | 17 | 0.206 | 3191.004 | 12282.884 | 12170.044 | 12187.044 | 12172.113 |
| 3 cl. MLC | -5802.464 | 24 | 0.240 | 3722.120 | 11812.231 | 11652.928 | 11676.928 | 11658.486 |
| 4 cl. MLC* | -5762.685 | 31 | 0.244 | 3801.678 | 11793.137 | 11587.370 | 11618.370 | 11599.047 |

Note: Observed heteroscedasticity is retrieved based on respondents from Rural and Urban (Baseline) areas
* denotes an unidentified model.

# Conditional WTP

| | Min. | 1st qu | Median | Mean | 3rd qu | Max. |
|---|---|---|---|---|---|---|
| SCALE PARAMETER | -0.632 | -0.348 | -0.346 | -0.163 | -0.250 | 0.767 |
| | | | | | | |
| long walk | -1.799 | -1.798 | -0.424 | -0.757 | -0.054 | -0.003 |
| car park | -0.014 | 0.022 | 0.071 | 1.429 | 3.837 | 3.840 |
| fence | 0.000 | 0.009 | 0.061 | 2.552 | 6.900 | 6.904 |
| path and signage | -0.069 | -0.010 | 0.139 | 2.417 | 6.523 | 6.528 |
| | | | | | | |
| hill | - | - | - | 0.097 | - | - |
| bog | - | - | - | 0.074 | - | - |
| field | - | - | - | 0.098 | - | - |
| river | - | - | - | 0.123 | - | - |

# Elasticities based on conditional parameters

**LONG WALK**

| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|---|---|---|---|---|---|---|
| hill | -1.518 | -0.41320 | -0.01475 | -0.353 | 0.00000 | 0.00000 |
| bog | -1.540 | -0.47560 | -0.04362 | -0.384 | 0.00000 | 0.00000 |
| field | -1.573 | -0.44160 | -0.02349 | -0.388 | 0.00000 | 0.00000 |
| river | -1.562 | -0.41330 | -0.02172 | -0.365 | 0.00000 | 0.00000 |

**CAR PARK**

| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|---|---|---|---|---|---|---|
| hill | -0.013 | 0 | 0 | 0.54690 | 0.05690 | 3.26600 |
| bog | -0.013 | 0 | 0 | 0.52390 | 0.05160 | 3.28700 |
| field | -0.013 | 0 | 0 | 0.63530 | 0.10760 | 3.40700 |
| river | -0.013 | 0 | 0 | 0.58800 | 0.06272 | 3.37700 |

**FENCE**

| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|---|---|---|---|---|---|---|
| field | 0.000 | 0 | 0 | 1.017 | 0.046 | 6.127 |
| river | 0.000 | 0 | 0 | 0.9193 | 0.0265 | 6.0720 |

**PATH AND SIGNAGE**

| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|---|---|---|---|---|---|---|
| hill | -0.068 | 0 | 0 | 1.22300 | 1.63300 | 5.55300 |
| bog | -0.068 | 0 | 0 | 1.17100 | 0.79540 | 5.58800 |
| field | -0.068 | 0 | 0 | 1.36900 | 3.63300 | 5.79200 |
| river | -0.068 | 0 | 0 | 1.28000 | 1.90700 | 5.74100 |

# Presentation Outline

# Conclusions

- This paper compared RPL and LC models in WTP space to simultaneously incorporate both scale and preference heterogeneity.
- The results show the importance of including a scale parameter even in models estimated in WTP-space.
- We find that:
  - Using LC models accommodating observed scale heterogeneity has some advantages.
  - More flexible models (ProbHLC, MLC) are more prone to identification and confounding problems.
- This paper is part of a more comprehensive study on more flexible Latent class models estimated with the new version of pythonbiogeme combined with the use of PERL.

# Questions???

*Contact information:*

**Marco Boeri**
Gibson Institute for Land, Food and Environment,
Queen's University Belfast.
Email: mboeri01@qub.ac.uk.

www.qub.ac.uk/sites/GibsonInstitute/