Institute for Transport Studies
FACULTY OF ENVIRONMENT

UNIVERSITY OF LEEDS

Conflicting interpretations of respondent heterogeneity

Stephane Hess

ITS

Outline

UNIVERSITY OF LEEDS

- Introduction
- Outlying sensitivities
- Heterogeneity in information processing strategies
- Heterogeneity in decision rules

ITS

Introduction I

UNIVERSITY OF LEEDS

- Applications using discrete choice models increasingly incorporate a random treatment of unexplained respondent heteogeneity
- Mainly in the form of Mixed Multinomial Logit models, along with a handful of applications using more general GEV mixture approaches
- For application, no longer just interested in mean measures of sensitivity/WTP, but also variation in population
- Reducing amount of effort going into explaining taste heterogeneity in a deterministic manner, which would arguably be preferable for interpretation and model application (e.g. forecasting)

ITS

Introduction II

UNIVERSITY OF LEEDS

- When incorporating a random treatment of heterogeneity, analyst should aim to minimise the impact of model specification on findings
  - ⋆ Significant body of research looking at implications of distributional assumptions in these models
- Should also aim to reduce the part of heterogeneity that is picked up by random component
  - ⋆ link some heterogeneity to socio-demographic characteristics

ITS

Introduction III

UNIVERSITY OF LEEDS

- Can our findings in terms of heterogeneity be unduly influenced by specific behavioural processes?
  - ⋆ Pinjari & Bhat (2006): unexplained non-linearities in response can manifest themselves as random taste heterogeneity
  - ⋆ de Borger & Fosgerau (2008): unexplained reference dependence can lead to biased heterogeneity findings
- This presentation:
  - ⋆ what do our findings in terms of heterogeneity actually tell us?
  - ⋆ can the behaviour of a few respondents have a large influence on our overall findings?
    - ∗ in terms of mean valuations as well as heterogeneity

ITS

Outlying sensitivities I

UNIVERSITY OF LEEDS

- Work by Campbell, Hess, Rose & Scarpa (2009) on simulated data and Campbell & Hess (2009) on SP data
- Presence of a few respondents with extreme sensitivities can have a very significant influence on findings in terms of heterogeneity when making use of *standard* distributions
  - ⋆ impact also on mean valuations, which are crucial for application, policy work, etc
- Use of non-parameteric distributions or mixtures of distributions appealing in this context
  - ⋆ e.g. mixture of three continuous distributions, with a small lower and upper outlier group, along with a large central group

ITS

Outlying sensitivities II

UNIVERSITY OF LEEDS

- Findings show that such an approach can significantly reduce the extent of random heterogeneity in the remainder of the data
  - ⋆ Very difficult to use in practice, especially with large datasets and choice sets
- No attempt to explain *extreme* sensitivities
- Two possible interpretations offered in this presentation

ITS

Heterogeneous IPS I

UNIVERSITY OF LEEDS

- Growing recognition that different respondents process information presented in SP surveys in different ways
- Examples:
  - ⋆ respondents may ignore one or more of the attributes
  - ⋆ respondents may process several of the attributes jointly
- Direct questions about information processing strategies (IPS) seem to confirm this
- Lack of treatment of IPS heterogeneity is likely to affect findings in terms of overall respondent heterogeneity, as well as mean findings

ITS

Heterogeneous IPS II

UNIVERSITY OF LEEDS

- Hess & Hensher (2010)
- Data from an Australian toll road study
- Five attributes: free flow travel time, slowed down travel time, travel time variability, running costs, toll
- High rates of stated attribute ignoring
- Previous work had relied on setting attributes to zero for specific respondents
- Two problems:
  - ⋆ Endogeneity
  - ⋆ Can we really believe what respondents are telling us?

ITS

Heterogeneous IPS III

UNIVERSITY OF LEEDS

- Rather than using *stated* IPS, attempt to infer strategies from the data
- Estimate MMNL model on full data, and produce conditional distributions for each attribute for each respondent
- Assign respondents to ignoring/non-ignoring classes on the basis of probability that sensitivity for a given attribute is zero
  - ⋆ some very visible differences in group allocation
  - ⋆ also, rates much lower (e.g. 2% for tolls instead of 9%)
- Unlike with stated information approach, models show that sensitivities in the inferred ignoring group are indeed equal to zero

ITS

Heterogeneous IPS IV

UNIVERSITY OF LEEDS

- Impact on heterogeneity findings
- Use coefficient of variation as an indication

	Base model	Inferred IPS		
adj. ρ^2	0.4492	0.4696		
			Reductions	
	c.v.	c.v.	het	resp
β_{FFT}	0.98	0.75	-23.65%	-15.61%
β_{SDT}	0.72	0.68	-5.47%	-2.44%
β_{RC}	0.78	0.63	-19.39%	-5.37%
β_{TOLL}	0.57	0.57	-0.01%	-1.95%
β_{VAR}	3.97	2.76	-30.49%	-29.27%

- For three attributes, *ignoring* respondents accounted for disproportionally large share of heterogeneity

ITS

Heterogeneous decision rules I

UNIVERSITY OF LEEDS

- Various pieces of work looking at subsets of respondents behaving in a different manner
- But models tend to still assume that the actual decision rule (i.e. compensatory behaviour) still applies across respondents
- There is however also ample evidence to suggest that some respondents employ other decision rules
- One example is lexicographic behaviour
  - ⋆ choose alternative that dominates on most important attribute
  - ⋆ in case of ties, concerned alternatives are retained, and we move to the second most important attribute
  - ⋆ ordering of attributes varies across respondents

ITS

Heterogeneous decision rules II

UNIVERSITY OF LEEDS

- Lexicographic behaviour especially easy to spot in the case of surveys with only two attributes
- In surveys with more attributes, many different rules are possible (involving different orderings as well as numbers of levels)
- Presence of lexicographic respondents can not only change retrieved mean sensitivities (and WTP measures) but also inflate the retrieved degree of heterogeneity
  - ⋆ models will attempt to explain non-trading behaviour through extreme sensitivities

ITS

Heterogeneous decision rules III

UNIVERSITY OF LEEDS

- Past work:
  - ⋆ ignore presence of lexicographic behaviour
    - ∗ likely to impact on estimates
  - ⋆ remove lexicographic respondents
    - ∗ this in effect assumes that these respondents were definitely not behaving in a compensatory manner
    - ∗ but: incentives may simply not have been large enough
  - ⋆ attempt to accommodate lexicographic respondents
    - ∗ work by Rigby et *al*. uses a latent class structure where in the lexicographic class, only a single coefficient is used
      - · assumes that respondents ignored all other attributes
      - · would still give a non-zero probability to the dominated alternative
      - · remaining coefficient could become very large

ITS

Heterogeneous decision rules IV

UNIVERSITY OF LEEDS

- Proposed modelling framework: combine compensatory model with dominance based models
- Let $P_n(R_1)$ be the probability for the sequence of choices observed for respondent n, under the assumption that decision rule R_1 was used
- In a standard MNL model, we would thus have

$$P_n(R_1) = \prod_{t=1}^{T} \frac{e^{V_{j_{nt}}}}{\sum_{j=1}^{J} e^{V_j}}$$

where $V_{j_{nt}}$ is the modelled utility of the alternative chosen by respondent n in choice situation t

ITS

Heterogeneous decision rules V

UNIVERSITY OF LEEDS

- Probability for sequence of choices observed for respondent n:

$$P_n = \sum_{r=1}^{R} \pi_{n,r} P_n(R_r) \quad \text{where} \quad \sum_{r=1}^{R} \pi_{n,r} = 1 \quad \text{and} \; 0 \leq \pi_r \leq 1 \; \forall r$$

- Let us further assume that R_1 corresponds to the compensatory model, where we make use of a Mixed Logit model, such that:

$$P_n(R_1) = \int_{\beta} \prod_{t=1}^{T} \frac{e^{V_{j_{nt}}}}{\sum_{j=1}^{J} e^{V_j}} f(\beta \mid \Omega) \, \mathrm{d}\beta$$

- Remaining rules are dominance based rules

ITS

Heterogeneous decision rules VI

UNIVERSITY OF LEEDS

- Example: two attributes, time and money, where rule 2 is lexicography on time (tt), and rule 3 is lexicography on cost (tc)

$$P_n(R_2) = \prod_{t=1}^{T} \delta_{TT_{jnt}}$$

where $\delta_{TT_{jnt}}$ is equal to 1 if the travel time for the alternative chosen by respondent n in choice set t is less than that of any of the other available alternatives

$$P_n(R_3) = \prod_{t=1}^{T} \delta_{TC_{jnt}}$$

where $\delta_{TC_{jnt}}$ is defined analogously to equal to $\delta_{TT_{jnt}}$

ITS

Heterogeneous decision rules VII

UNIVERSITY OF LEEDS

- The probability under a given lexicographic rule will be equal to 1 only if every single choice for that respondent can be explained by that rule
- This will only apply for some respondents
- In models with more than two attributes, multiple rules may be able to explain the same sequence of choices (depending on the design) and some normalisation may be required
- Need to estimate parameters of choice model in class 1 as well as the probabilities for all classes
  - ⋆ could also link class allocation to socio-demographic characteristics

ITS

Heterogeneous decision rules VIII

UNIVERSITY OF LEEDS

- Application I: Danish Valuation of Travel Time Savings (VTTS) data
- Binary design, with two alternatives (time and money)
- $1,676$ respondents, each with 8 choices
- 13.66% of respondents always choose cheapest, with 5.97% of respondents always choosing fastest
- Removing these respondents has the obvious impacts on results, but is also very arbitrary
- But their simple inclusion in the models, without treatment, arguably biases the findings in terms of heterogeneity

ITS

Heterogeneous decision rules VIII

UNIVERSITY OF LEEDS

- Four different models estimated
  - ⋆ MNL model
    - ∗ no treatment of lexicography, or random heterogeneity
  - ⋆ LC model
    - ∗ lexicography accommodated solely through special classes, with no attempts to explain through taste heterogeneity
  - ⋆ MMNL model
    - ∗ no special treatment of lexicography; explained solely through taste heterogeneity
  - ⋆ LC-MMNL model
    - ∗ special classes for lexicography that *cannot* be explained through heterogeneity, but with attempts to accommodate some through heterogeneity

ITS

Heterogeneous decision rules IX

UNIVERSITY OF LEEDS

- Along with time and cost sensitivity, estimate a constant for first alternative
- Random heterogeneity accommodated through multivariate Lognormal distribution

ITS

Heterogeneous decision rules X

UNIVERSITY OF LEEDS

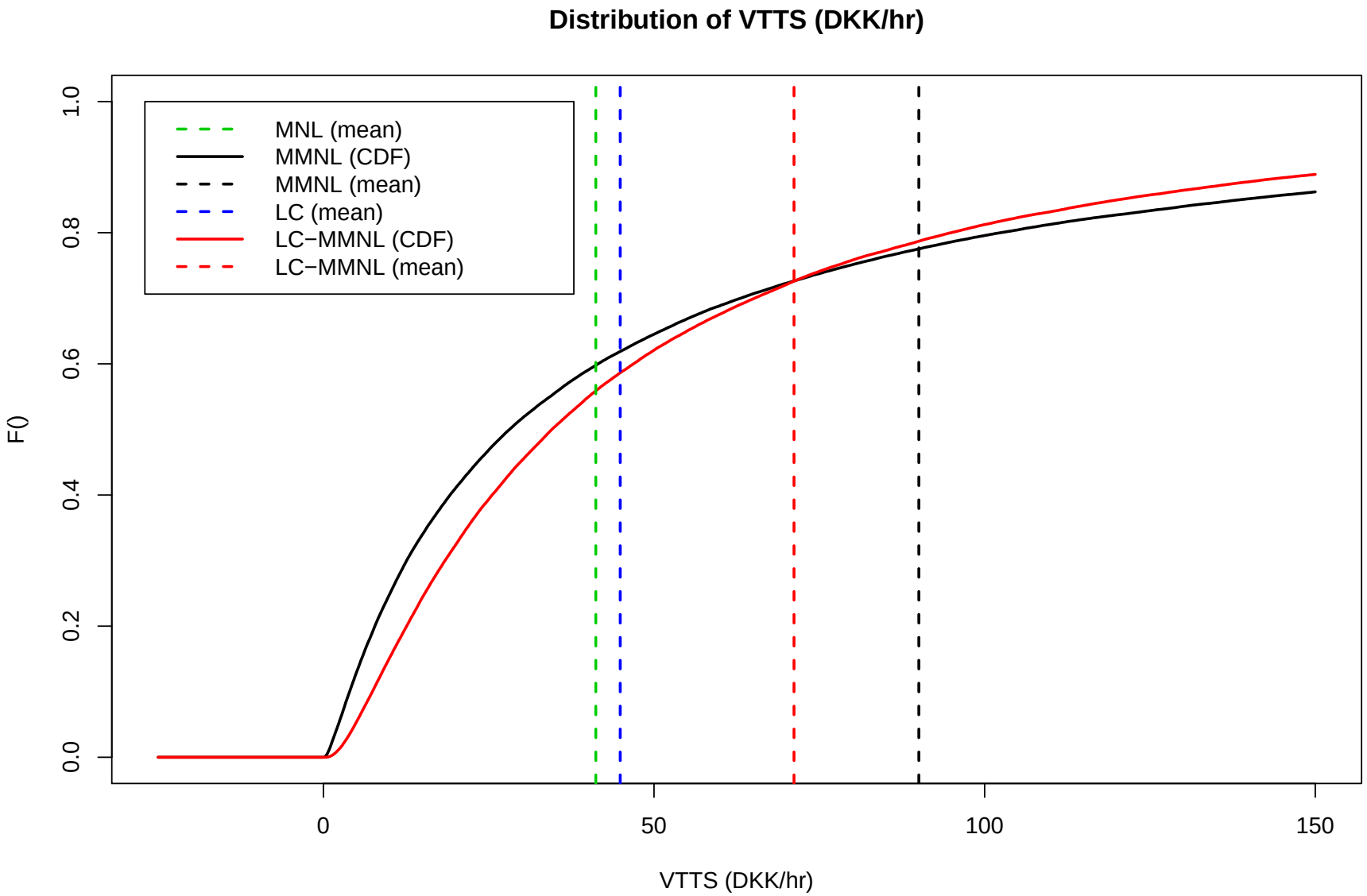
	MNL	LC	MMNL	LC-MMNL
LL	-8,925.89	-8,030.10	-7,360.77	-7,350.33
par.	3	5	5	7
adj. ρ^2	0.0393	0.1354	0.2074	0.2084
Trading	100%	81%	100%	90%
Lex-cost	0%	13%	0%	7%
Lex-time	0%	6%	0%	3%

- LC model retrieves existing size of lexicographic classes, while LC-MMNL model attempts to explain some of the behaviour through heterogeneity, without however going as far as MMNL model

ITS

Heterogeneous decision rules XI

UNIVERSITY OF LEEDS



ITS

Heterogeneous decision rules XII

UNIVERSITY OF LEEDS

- Evidence of reduced heterogeneity
- Degree of reduction very substantial compared to retrieved rates of lexicography

	VTTS					
	mean	median	st.dev.	c.v.	c.v. (β_{TT})	c.v. (β_{TC})
MNL	41.19		-	0	0	0
LC	44.89		-	0	0	0
MMNL	90.66	28.21	261.94	2.89	5.45	17.80
LC-MMNL	70.61	34.60	123.07	1.74	3.47	7.98

ITS

Heterogeneous decision rules XIII

UNIVERSITY OF LEEDS

- Application II: Survey for rail and bus commuters
- Three alternatives (first is status quo)
- Five attributes (travel time, fare, crowding, expected delay, provision of delay information system)
- Make use of constant for first two alternatives, four marginal utility coefficients, two coefficients for non-zero levels for information attribute
- Use Lognormals for four first coefficients, and Normals for final two, with full correlation structure

ITS

Heterogeneous decision rules XIV

UNIVERSITY OF LEEDS

- 5 single attribute rules (never applicable), 20 two attribute rules, 60 three attribute rules, 120 four attribute rules, and 120 five attribute rules (never applicable)

- Narrowed this down to 6 decision rules (along with compensatory)

1. fare - time - crowding
2. fare - info - reliability
3. crowding - fare - time
4. fare - crowding - info - reliability
5. crowding - reliability - fare - time
6. reliability - crowding - time - fare

ITS

Heterogeneous decision rules XV

UNIVERSITY OF LEEDS

	MNL	LC	MMNL	LC-MMNL
LL	-3,392.35	-3,245.62	-2,967.51	-2,945.76
par	8	14	29	35
adj. ρ^2	0.1589	0.1937	0.2588	0.2627
trading	100%	90.11%	100%	96.51%
fare-time-crowding		3.08%		0.00%
fare-info-reliability		0.82%		0.32%
crowding-fare-time		0.95%		0.55%
fare-crowding-info-reliability		3.04%		0.98%
crowding-reliability-fare-time		0.50%		0.44%
reliability-crowding-time-fare		1.49%		1.22%

- Much smaller rates than in Danish data
  - ⋆ was to be expected
    - ∗ survey actively encouraged trading and avoided dominance

ITS

Heterogeneous decision rules XV

UNIVERSITY OF LEEDS

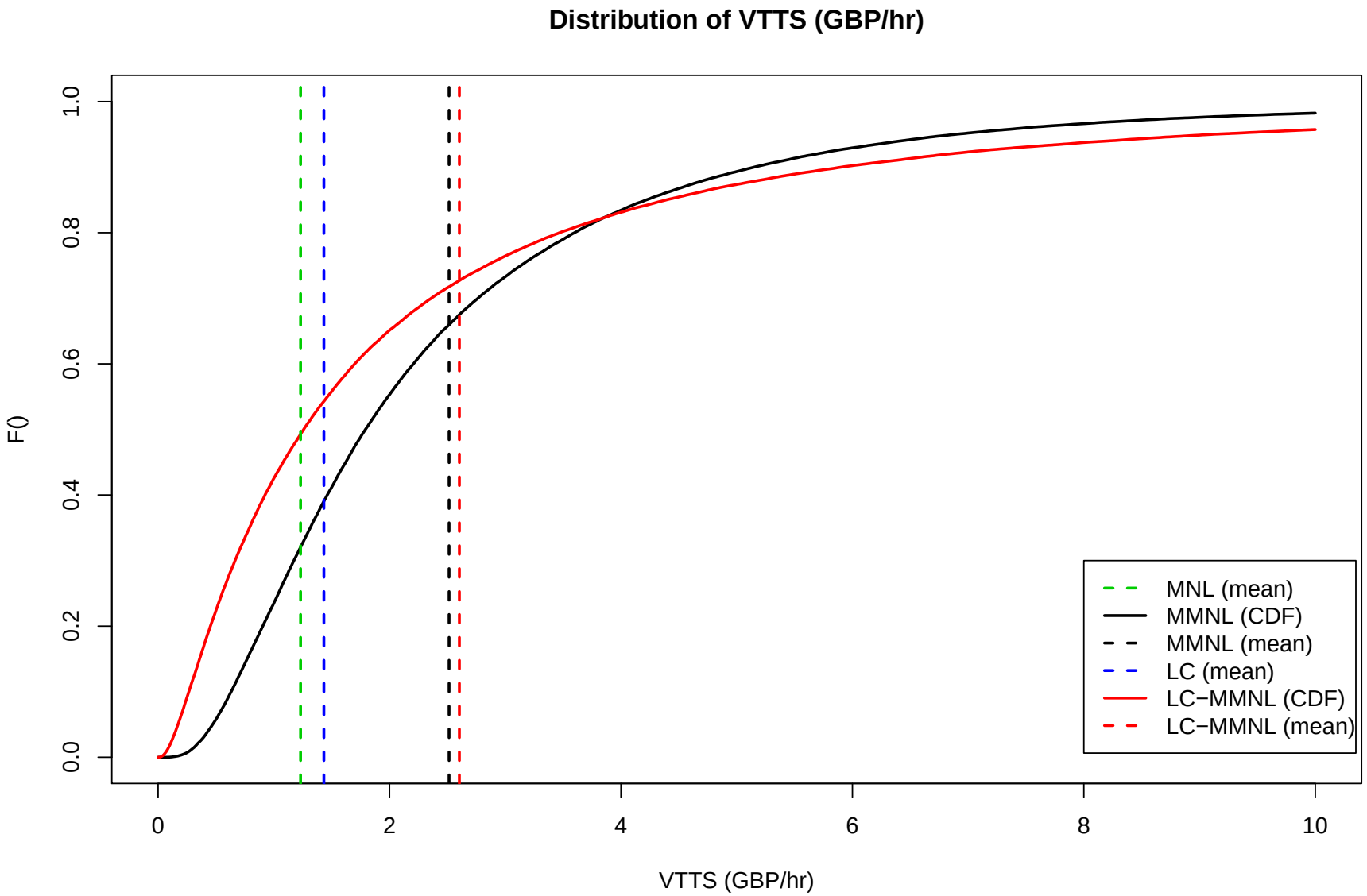
- Overall: evidence of reduced heterogeneity in WTP measures
- Exception: travel time, which was not well represented in decision rules

			MMNL			LC-MMNL		
WTP	MNL	LC	mean	median	c.v.	mean	median	c.v.
travel time (£/hr)	1.23	1.43	2.51	1.79	0.97	2.61	1.26	1.76
crowding	0.98	1.06	2.90	0.72	3.68	2.19	0.63	3.21
expected delay	0.47	0.52	2.20	0.56	3.33	1.73	0.59	2.60
info service (charged)	-0.04	-0.03	-0.09	-0.06	22.62	-0.13	-0.05	11.40
info service (free)	0.15	0.19	0.59	0.10	4.11	0.56	0.12	3.48

ITS

Heterogeneous decision rules XVI

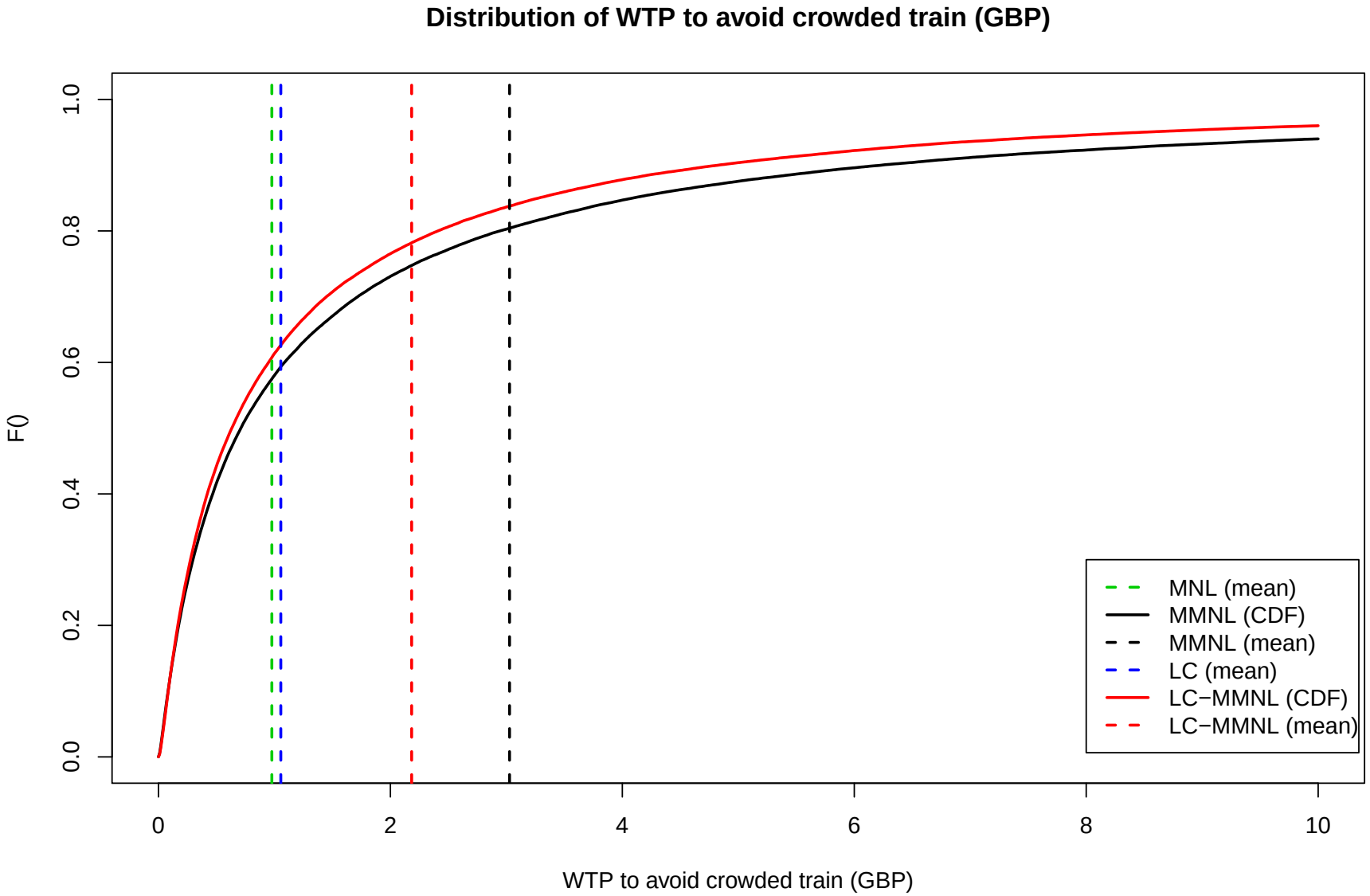
UNIVERSITY OF LEEDS



ITS

Heterogeneous decision rules XVII

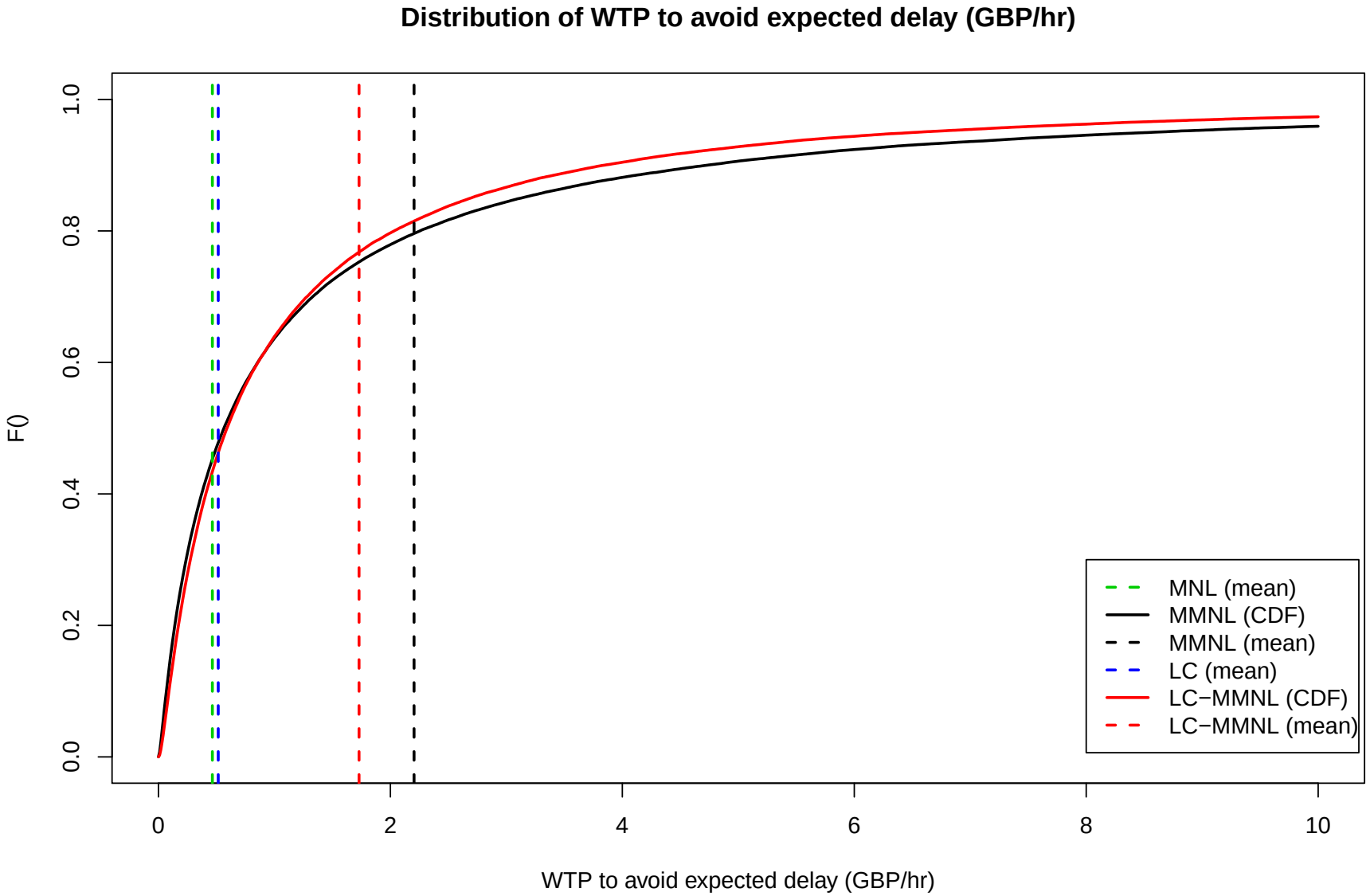
UNIVERSITY OF LEEDS



ITS

Heterogeneous decision rules XVIII

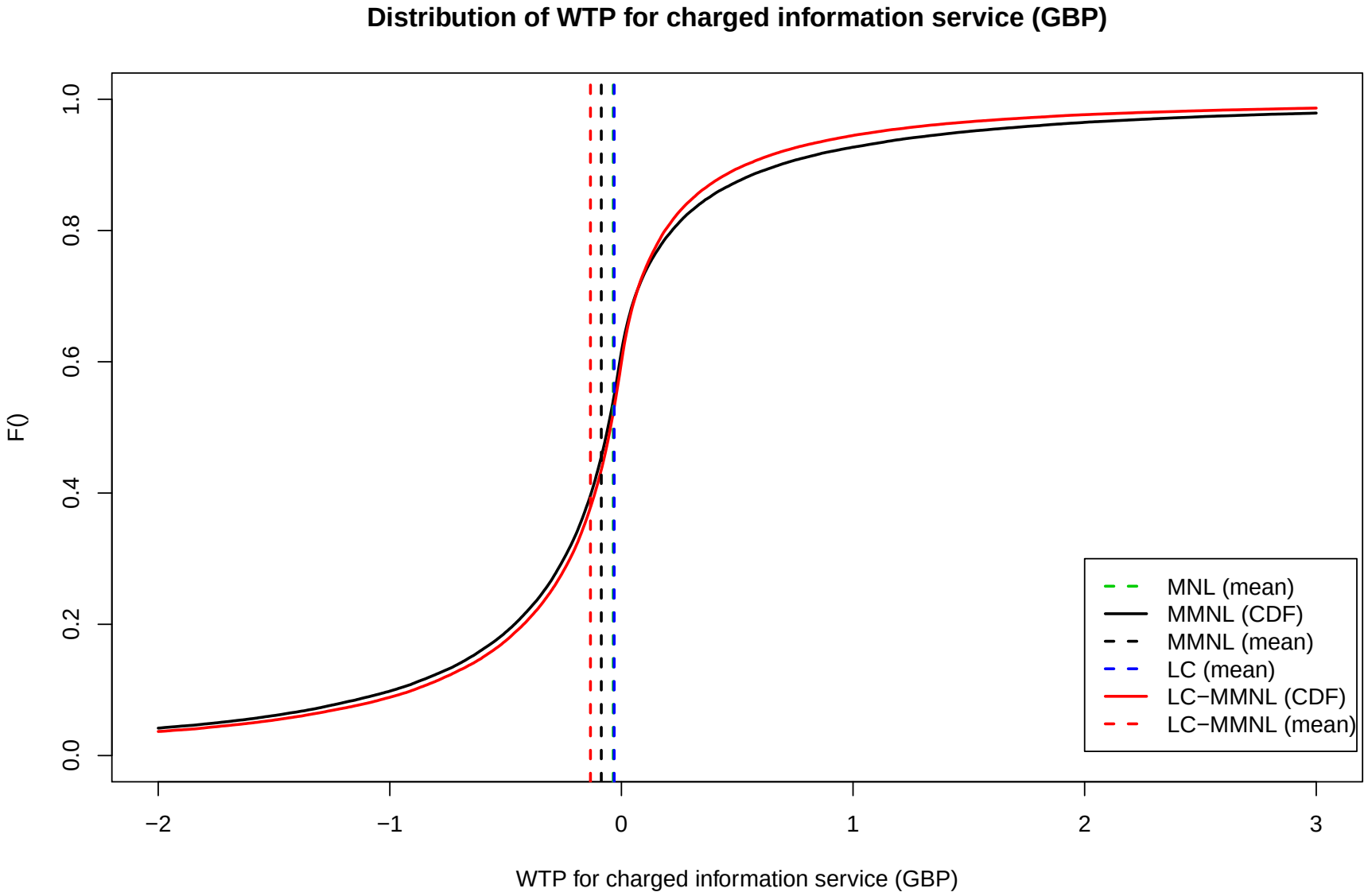
UNIVERSITY OF LEEDS



ITS

Heterogeneous decision rules XIX

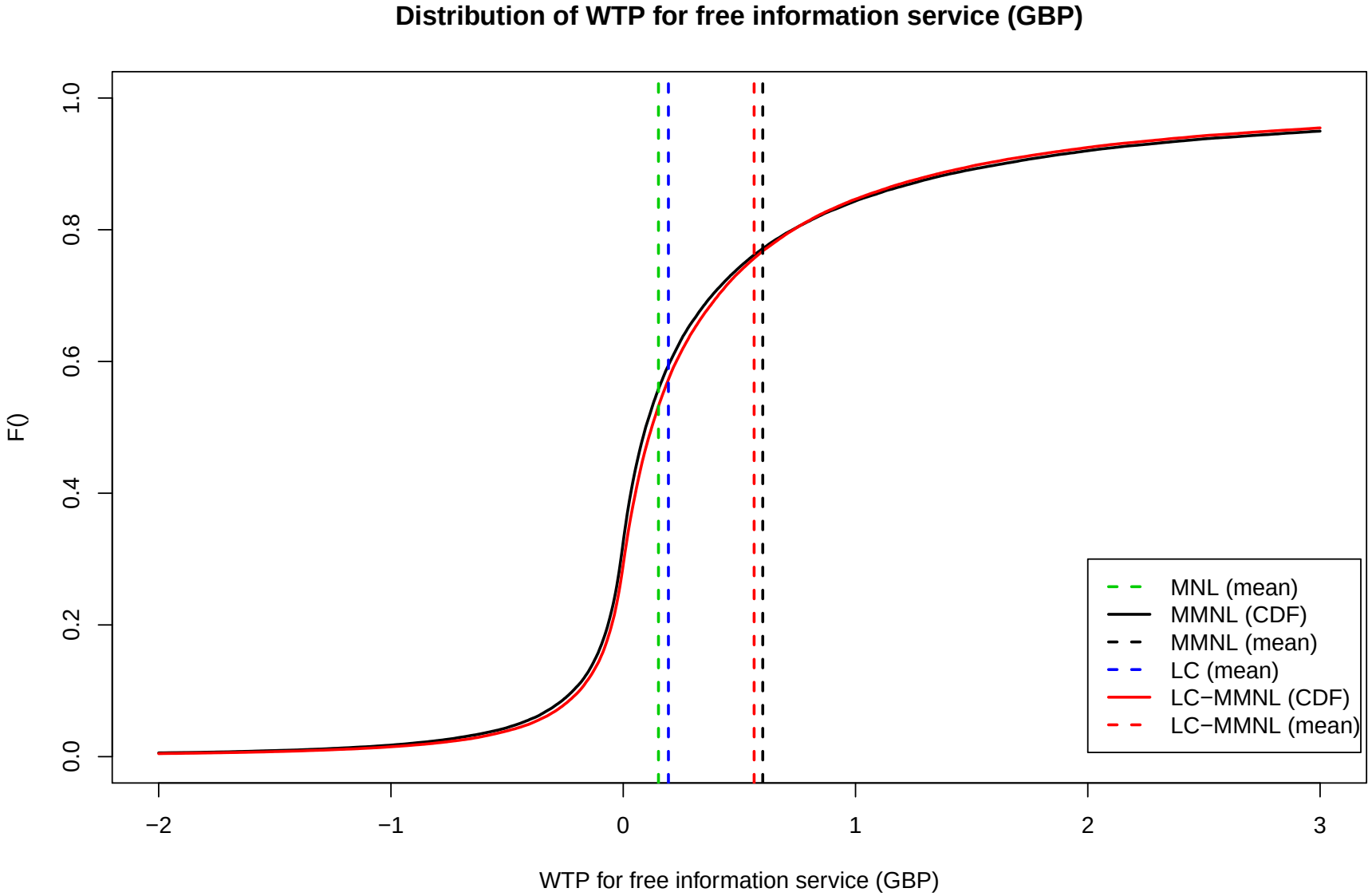
UNIVERSITY OF LEEDS



ITS

Heterogeneous decision rules XX

UNIVERSITY OF LEEDS



ITS

Heterogeneous decision rules XXI

UNIVERSITY OF LEEDS

- Evidence from two datasets
- Different scope for and rates of lexicographic behaviour
- In both datasets, mixed compensatory and dominance based model obtains significant improvements in model fit
- Evidence that small number of respondents account for large share of retrieved heterogeneity in base models
- But simply removing respondents is not advisable as their behaviour may be a reflection of sensitivities that are outside the incentives presented
- Additionally, in more complex surveys, difficult to identify lexicography in a deterministic manner

ITS

Conclusions

UNIVERSITY OF LEEDS

- Evidence that departures by some respondent from a purely compensatory approach may have a very significant effect on findings in terms of heterogeneity
- Different interpretations of what leads to extreme sensitivities
  - ⋆ Could also include political voting, effects of fatigue, etc
- Significant scope for future work, including in labelled surveys
- Could obviously accommodate some of this behaviour through very flexible distributions (e.g. spikes at zero), but not very helpful for interpretation
  - ⋆ end aim should be to explain class allocation (attitudes may help), with a view to understanding reasons behind behavioural processes

ITS

Second International Choice Modelling Conference

4 - 6 July 2011 | Oulton Hall | Yorkshire | UK

Jordan Louviere

Andrew Daly

Deadline for abstracts: 30 November 2010

Deadline for workshop proposals: 1 October 2010

www.icmconference.org.uk

UNIVERSITY OF LEEDS
Institute for Transport Studies