AN ENHANCED MEASUREMENT MODEL OF PERCEPTION OF COMFORT IN PUBLIC TRANSPORTATION

Aurélie Glerum

Michel Bierlaire

LATSIS 2012

1st European Symposium on Quantitative Methods in Transportation Systems
6th September 2012

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

OUTLINE

Introduction & motivation

The data

- RP survey
- Adjective quantification survey

The integrated model framework

- Discrete choice model
- Latent variable model
- Quantification model

Application example

- Quantification model
- Integrated model
- Validation of the integrated model

Conclusion

INTRODUCTION & MOTIVATION

Recent developments in discrete choice modeling (DCM)

- Choice cannot only be explained by economic indicators (travel duration, price of a trip, etc.)
- **Psychological constructs (attitudes, perceptions, etc.) play important role in choice behavior**: need to be integrated in an appropriate way into DCMs.
- Framework handling this issue:
  hybrid choice model (HCM) framework
  (Walker, 2001; Ben-Akiva et al., 2002)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

INTRODUCTION & MOTIVATION

Hybrid choice model (HCM): DCM with latent constructs.

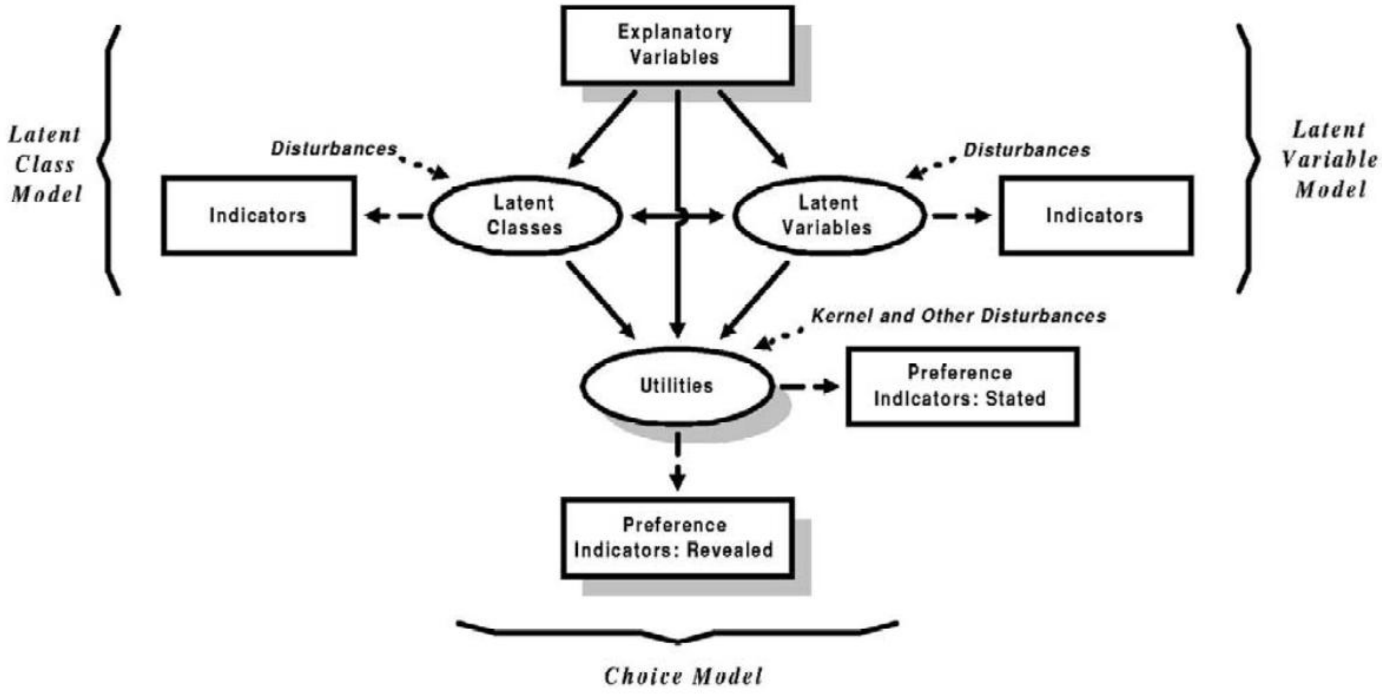


Figure extracted from Walker and Ben-Akiva, 2002.

INTRODUCTION & MOTIVATION

Hybrid choice model (HCM): DCM with latent constructs.

In this research: focus on the **integration of choice model and latent variable model**

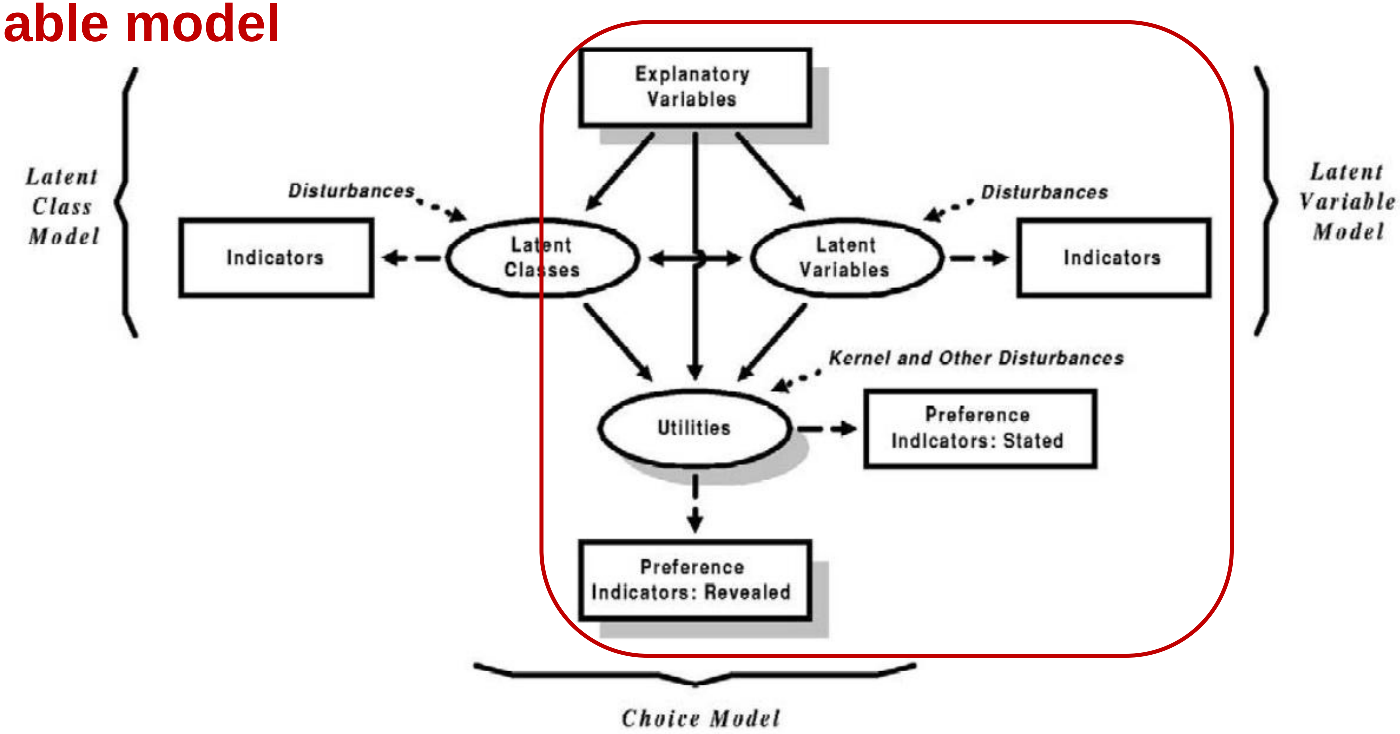


Figure extracted from Walker and Ben-Akiva, 2002.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

INTRODUCTION & MOTIVATION

Issues related to the integration of latent variables into choice models:

1. **Measurement of latent variable**

   ⟹ How to obtain the most realistic and accurate measure of a perception?

2. **Integration of the measurement into the choice model**

   ⟹ How to incorporate this information in the choice modeling framework?

INTRODUCTION & MOTIVATION

1. **Measurement of latent variable:**

- Use of **opinion statements**
  Five-point Likert scale

  } Usual way in literature (Likert, 1932; Bearden and Netemeyer, 1999)

- **Recent technique** developed in **social sciences:**

  Respondents report **adjectives** characterizing a variable of interest (Kaufmann et al., 2001; Kaufmann et al., 2010)

  Reflects **spontaneous** perceptions of individuals
  (≠ survey designer's conception of the perception)

INTRODUCTION & MOTIVATION

2. Integration of the measurement into the choice model:

- Structural equation model (SEM) framework used to characterize latent variable and relate it to its measurement indicators (e.g. Bollen, 1989).
- Latent variable model embedded into DCM ⟹ HCM framework
- Integration of measurements into HCM framework:
  - Easy for models with opinion statements
  - Needs an **additional modeling step** for model with adjectives

INTRODUCTION & MOTIVATION

Purpose of the research:

Develop an HCM that uses adjectives as measurements of latent construct

Steps:

1. Collection of choice data & psychometric data in the form of adjectives

2. Quantification of adjectives:

    1. Survey to obtain ratings of adjectives
    2. Quantification model
    3. Integration of the quantification model into the HCM framework

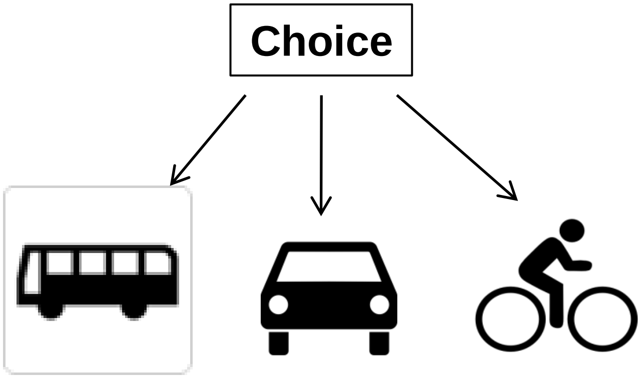
THE DATA

Two surveys:

- Revealed preferences (RP) survey
- Survey with evaluators (adjective quantification survey)

RP survey

- Mode choice study
- Conducted between 2009-2010 in low-density areas of Switzerland
- Conducted with PostBus (major bus company in Switzerland, operates in low-density areas)
- Info on all trips performed by inhabitants in one day:
  - Transport mode
  - Trip duration
  - Cost of trip
  - Activity at destination
  - Etc.
- **1763 valid questionnaires** collected



THE DATA

RP SURVEY

Adjective data for perception of transport modes:

For each of the following transport modes, give three adjectives that describe them best according to you.

		Adjective 1	Adjective 2	Adjective 3
1	The car is:			
2	The train is:			
3	The bus, the metro and the tram are:			
4	The post bus is:			
5	The bicycle is:			
6	The walk is:			

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DATA

RP SURVEY

Adjective data for perception of transport modes:

For each of the following transport modes, give three adjectives that describe them best according to you.

		Adjective 1	Adjective 2	Adjective 3
1	The car is:	**convenient**	**comfortable**	**expensive**
2	The train is:	**relaxing**	**punctual**	**restful**
3	The bus, the metro and the tram are:	**fast**	**frequent**	**cheap**
4	The post bus is:	**punctual**	**comfortable**	**cheap**
5	The bicycle is:	**stimulating**	**convenient**	**cheap**
6	The walk is:	**healthy**	**relaxing**	**independent**

THE DATA

RP SURVEY

Extraction of information on perceptions

1. Classification into themes:

- Perception of cost
- Perception of time
- Difficulty of access
- Flexibility
- Comfort, etc.

2. Focused on adjectives related to one theme only and one mode only:

Comfort in public transportation (PT)

Comfort
hardly full
packed
bumpy
comfortable
hard
irritating
tiring
unsuitable with bags
uncomfortable
bad air
…

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DATA

RP SURVEY

Extraction of information on perceptions

1. Classification into themes:

- Perception of cost
- Perception of time
- Difficulty of access
- Flexibility
- Comfort, etc.

2. Focused on adjectives related to one theme only and one mode only:

Comfort in public transportation (PT)

⇨ **LATENT VARIABLE WE STUDY**

Comfort
hardly full
packed
bumpy
comfortable
hard
irritating
tiring
unsuitable with bags
uncomfortable
bad air
…

TRANSP-OR

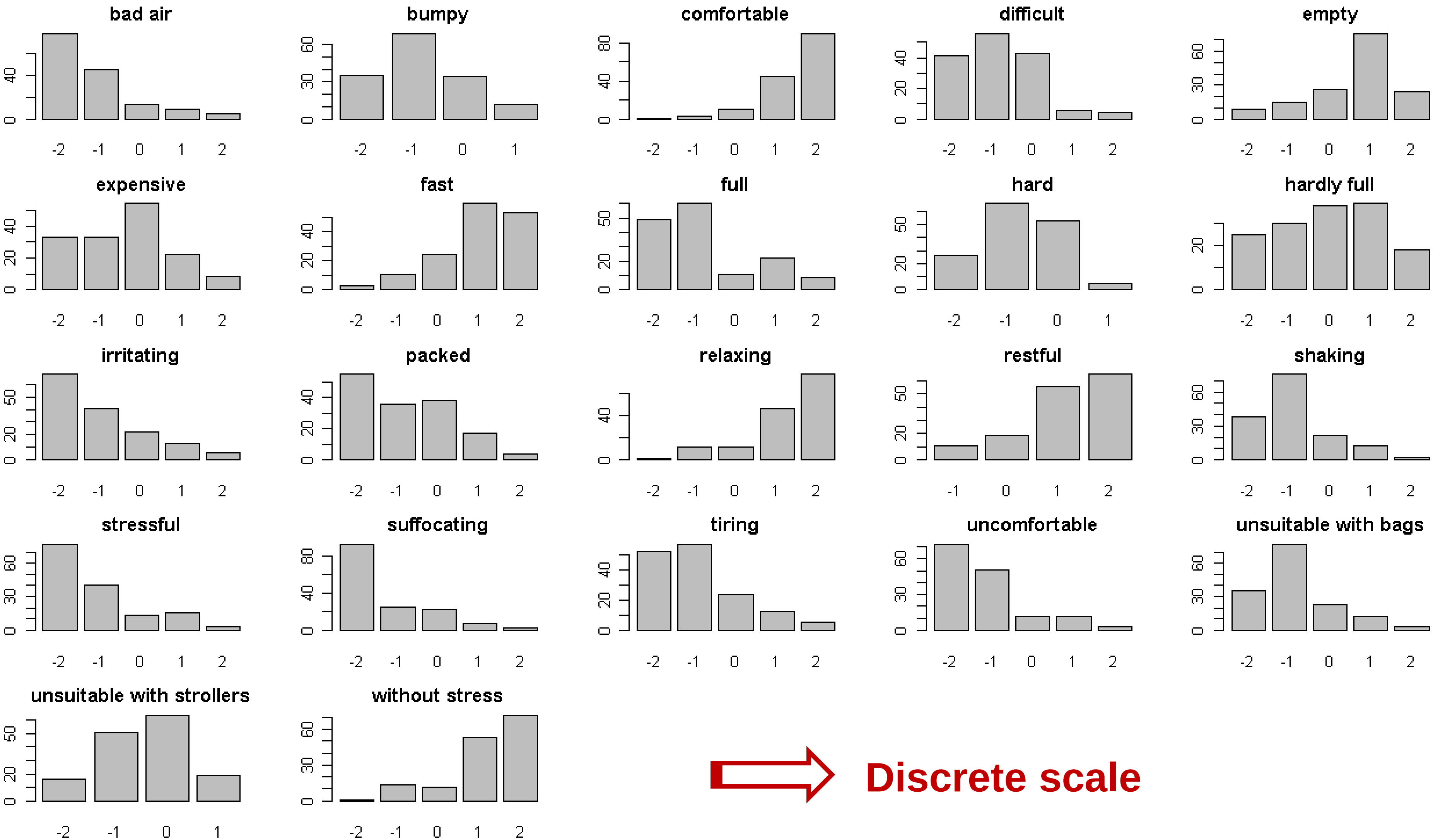
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Adjective quantification survey

- Asked external evaluators to rate the adjectives on scale of comfort.
- Two scales:
  - Discrete scale: ratings from -2 to 2.
  - Continuous scale: ratings from -1000 to 1000.
- Number of evaluators: 277

THE DATA

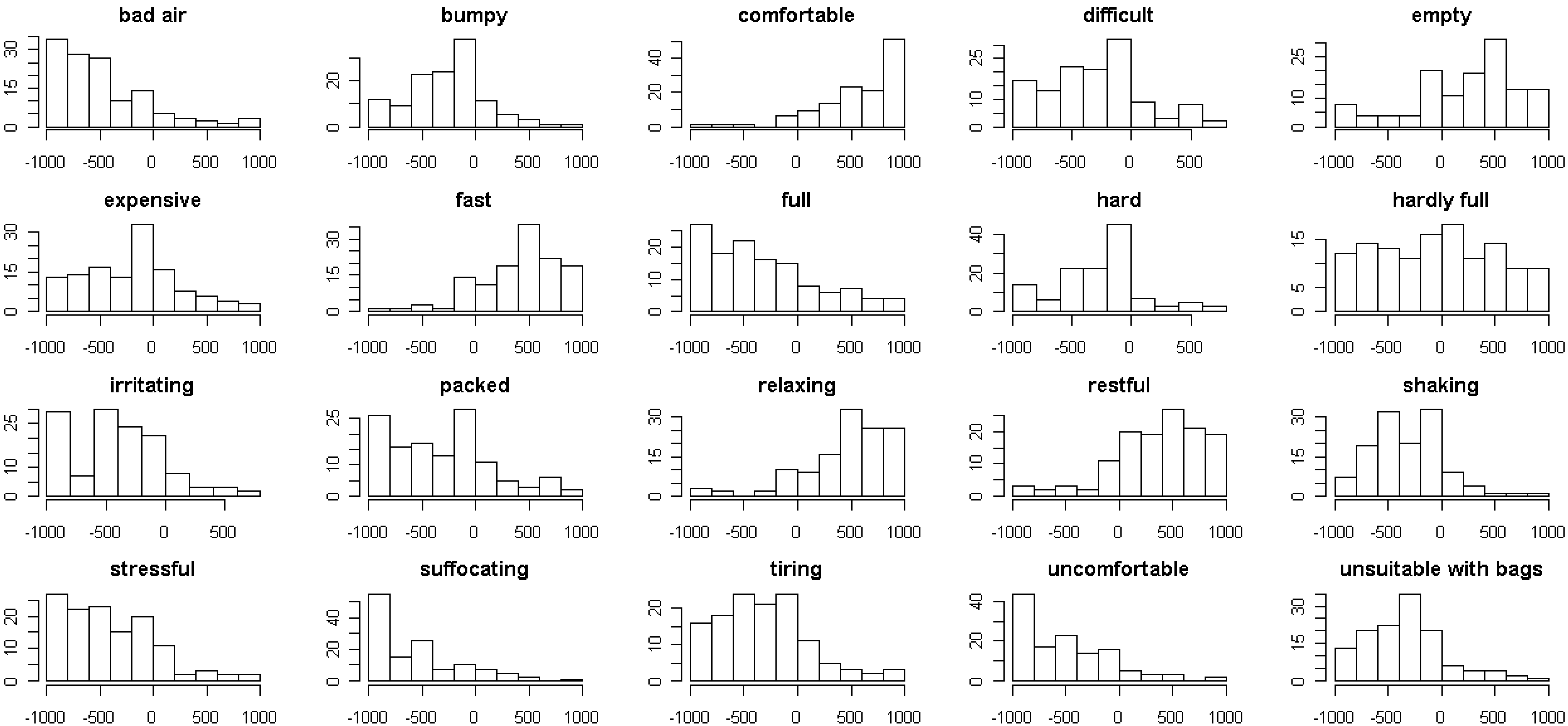
ADJECTIVE QUANTIFICATION SURVEY

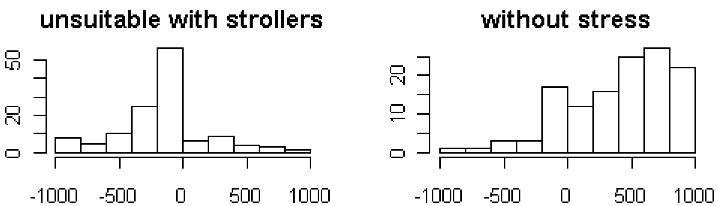


Discrete scale

THE DATA

ADJECTIVE QUANTIFICATION SURVEY





⇒ **Continuous scale**

THE INTEGRATED MODEL FRAMEWORK

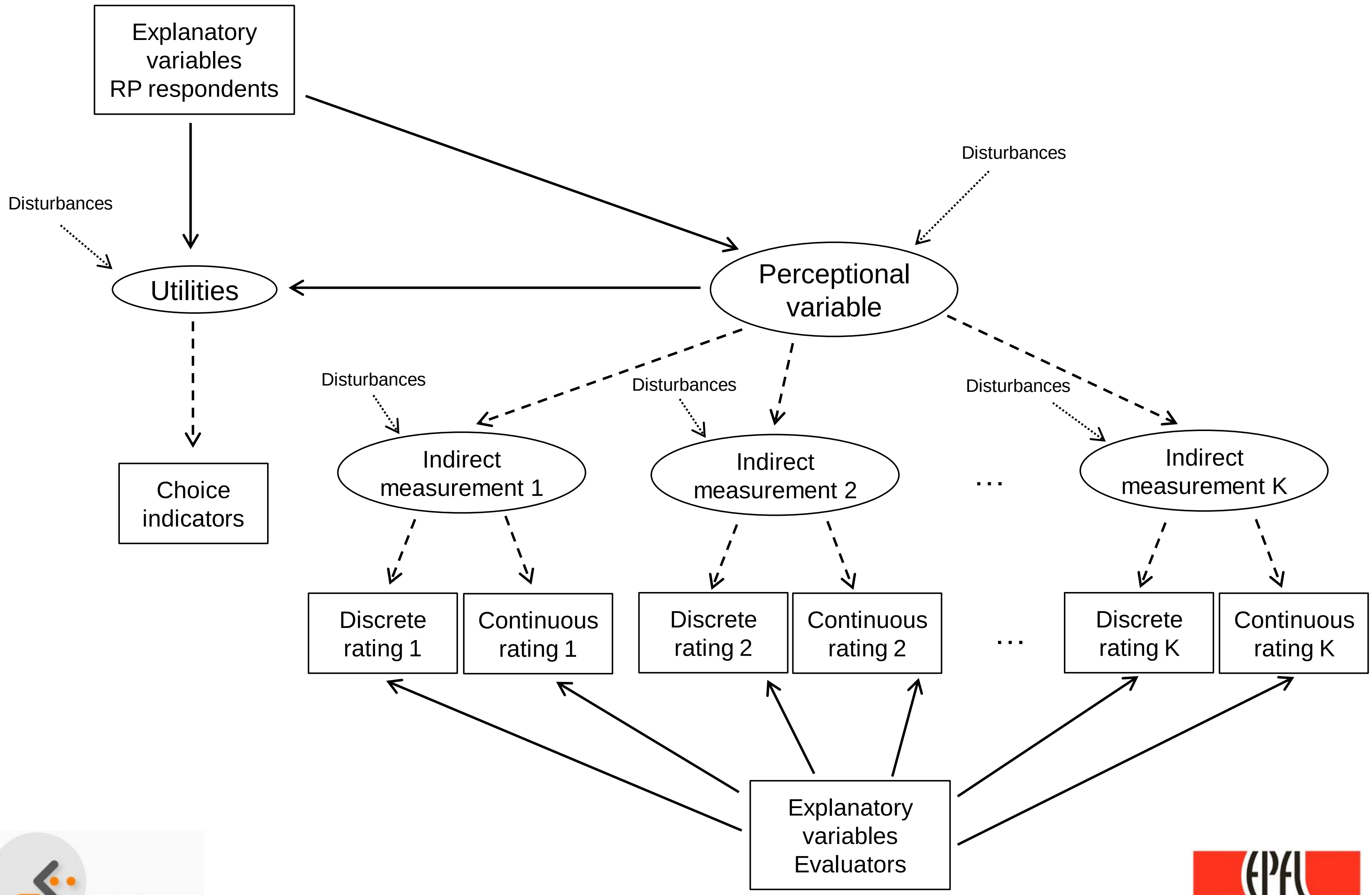
Purpose of the developed HCM:

Assess impact of perception on choice.
Using adjective data $\Longrightarrow$ need following integrated framework.

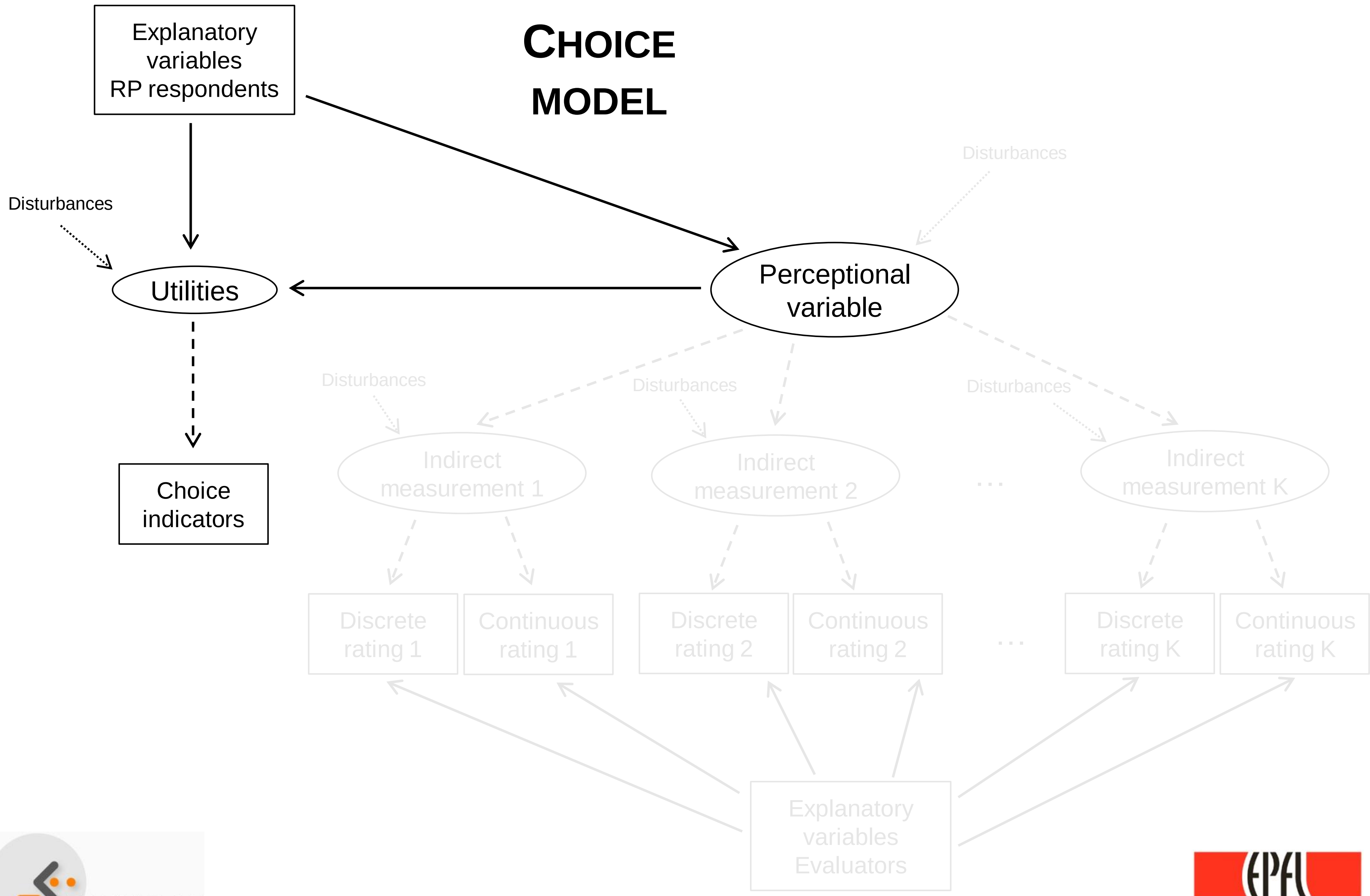
Framework involves three components:

- Discrete choice model
- Latent variable model for the perception
- Quantification model for the indicators of the latent variable

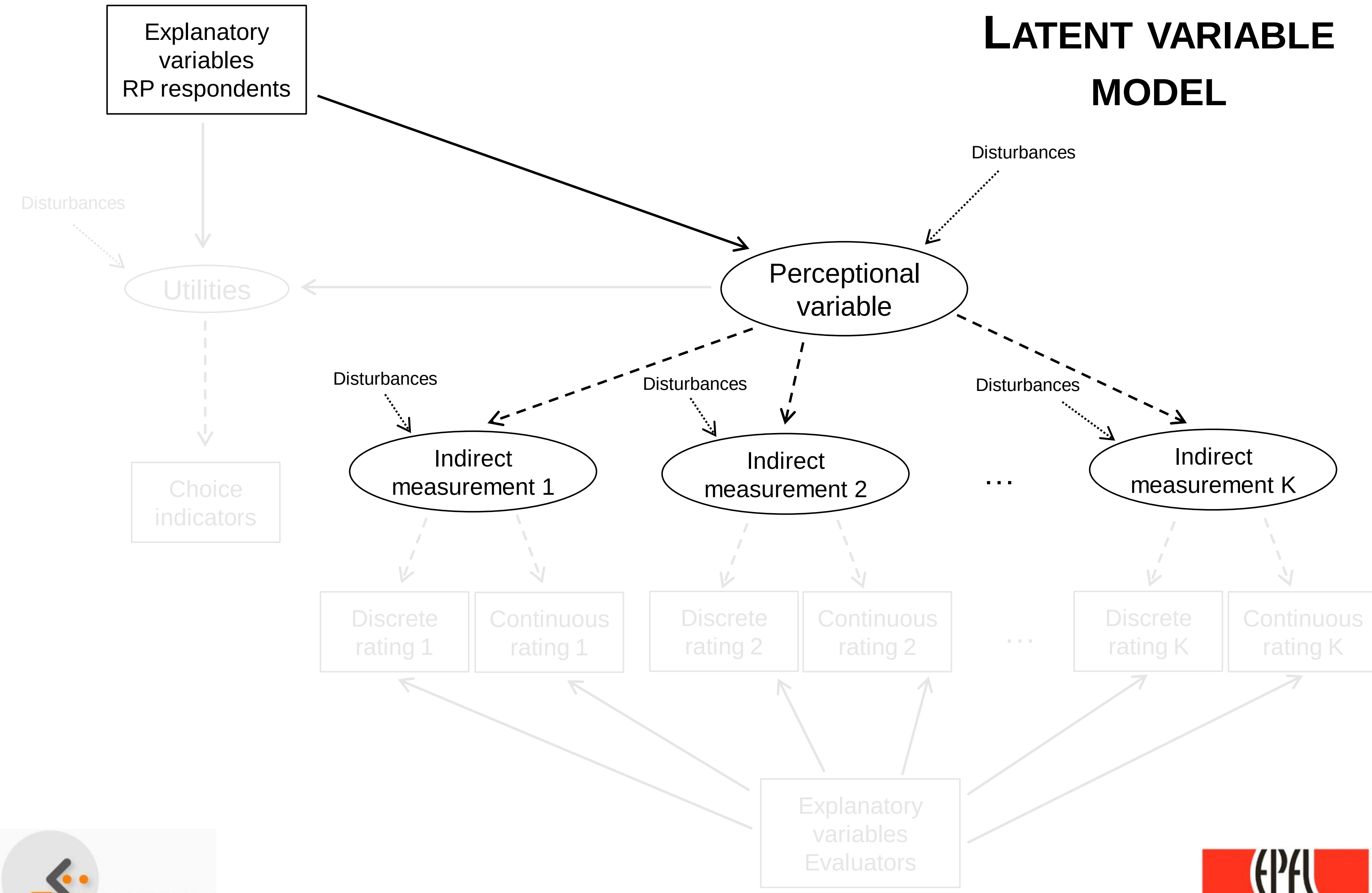
THE INTEGRATED MODEL FRAMEWORK



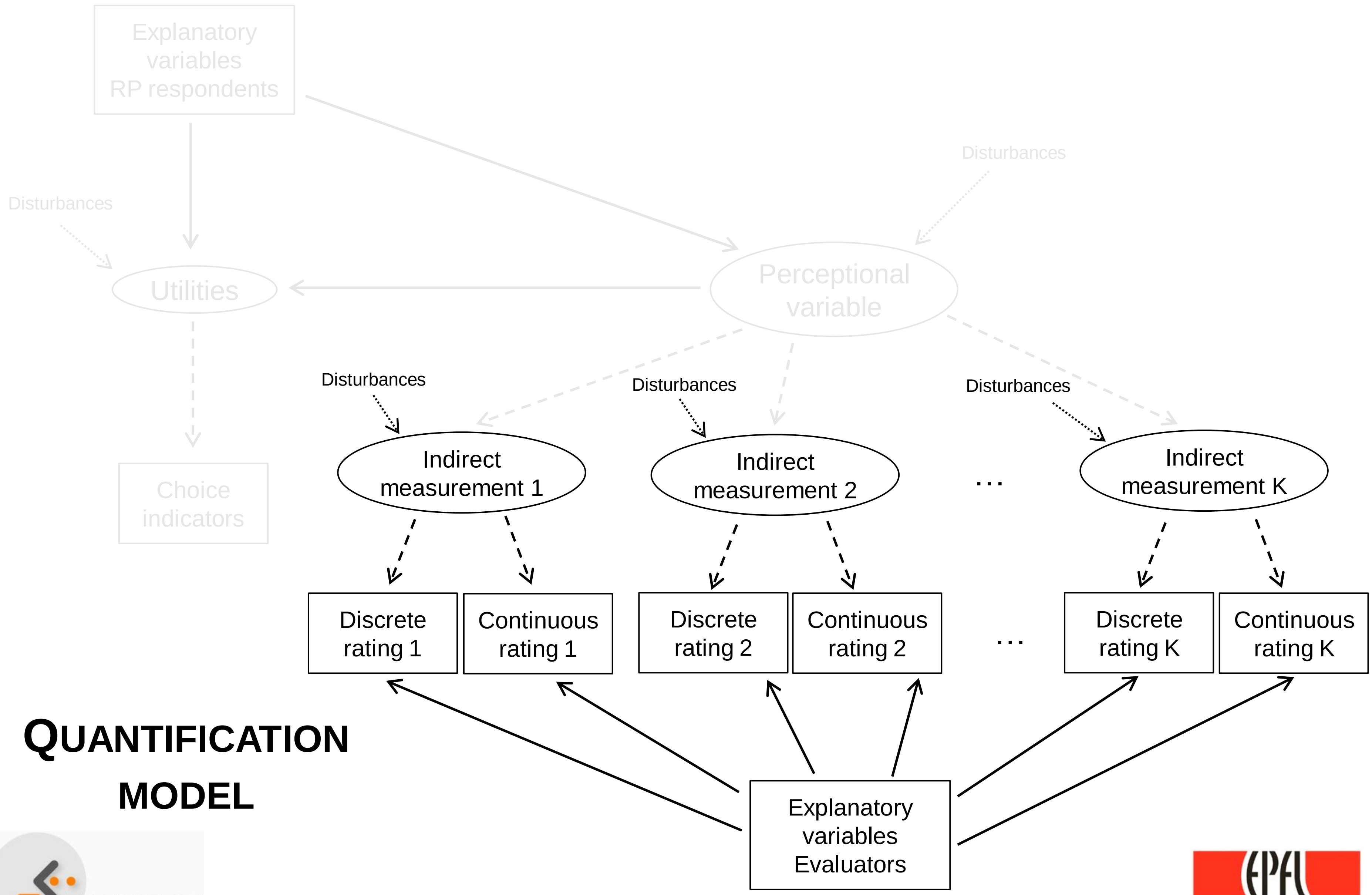
THE INTEGRATED MODEL FRAMEWORK



THE INTEGRATED MODEL FRAMEWORK

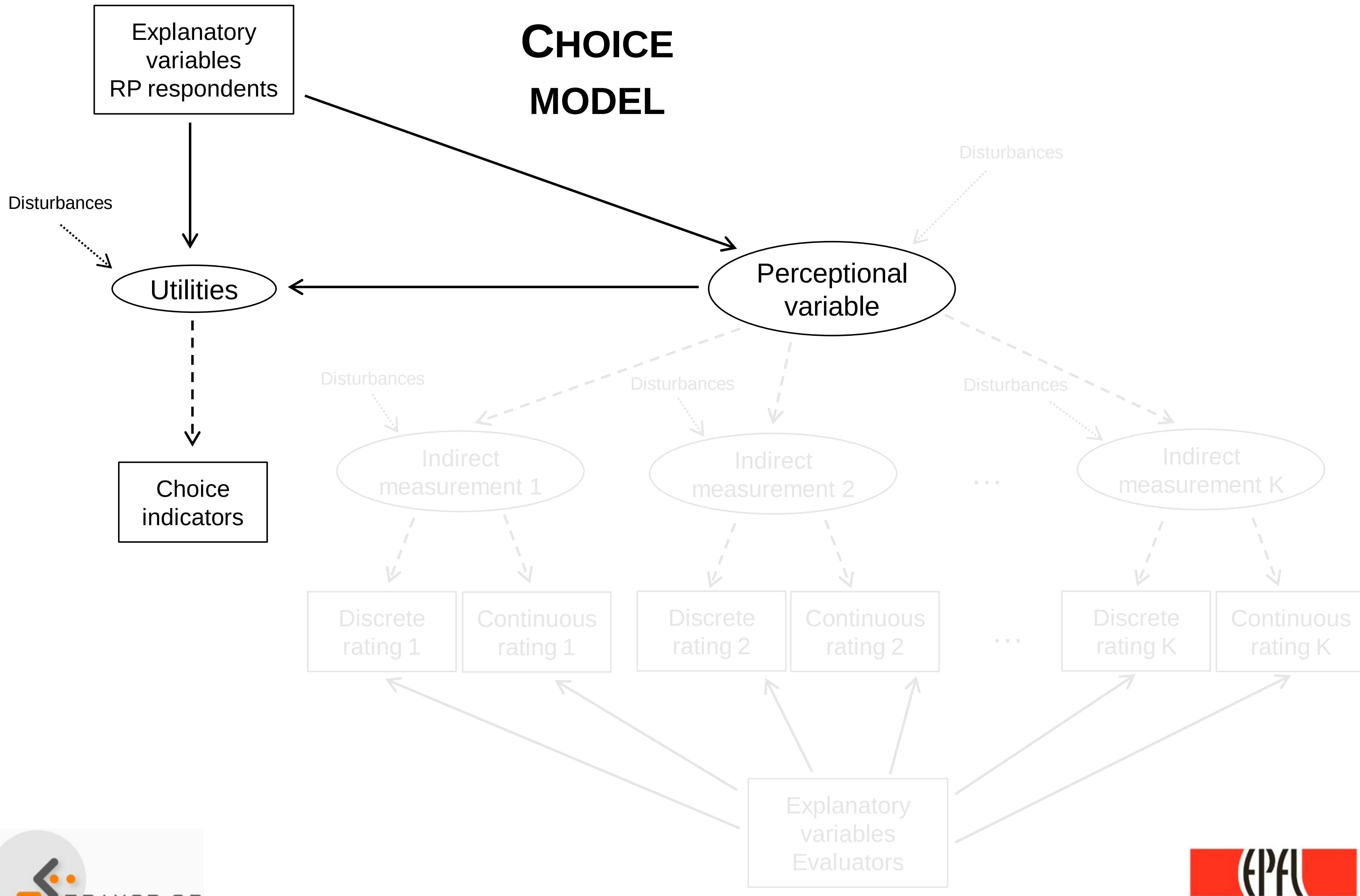


THE INTEGRATED MODEL FRAMEWORK



QUANTIFICATION MODEL

THE INTEGRATED MODEL FRAMEWORK



TRANSP-OR

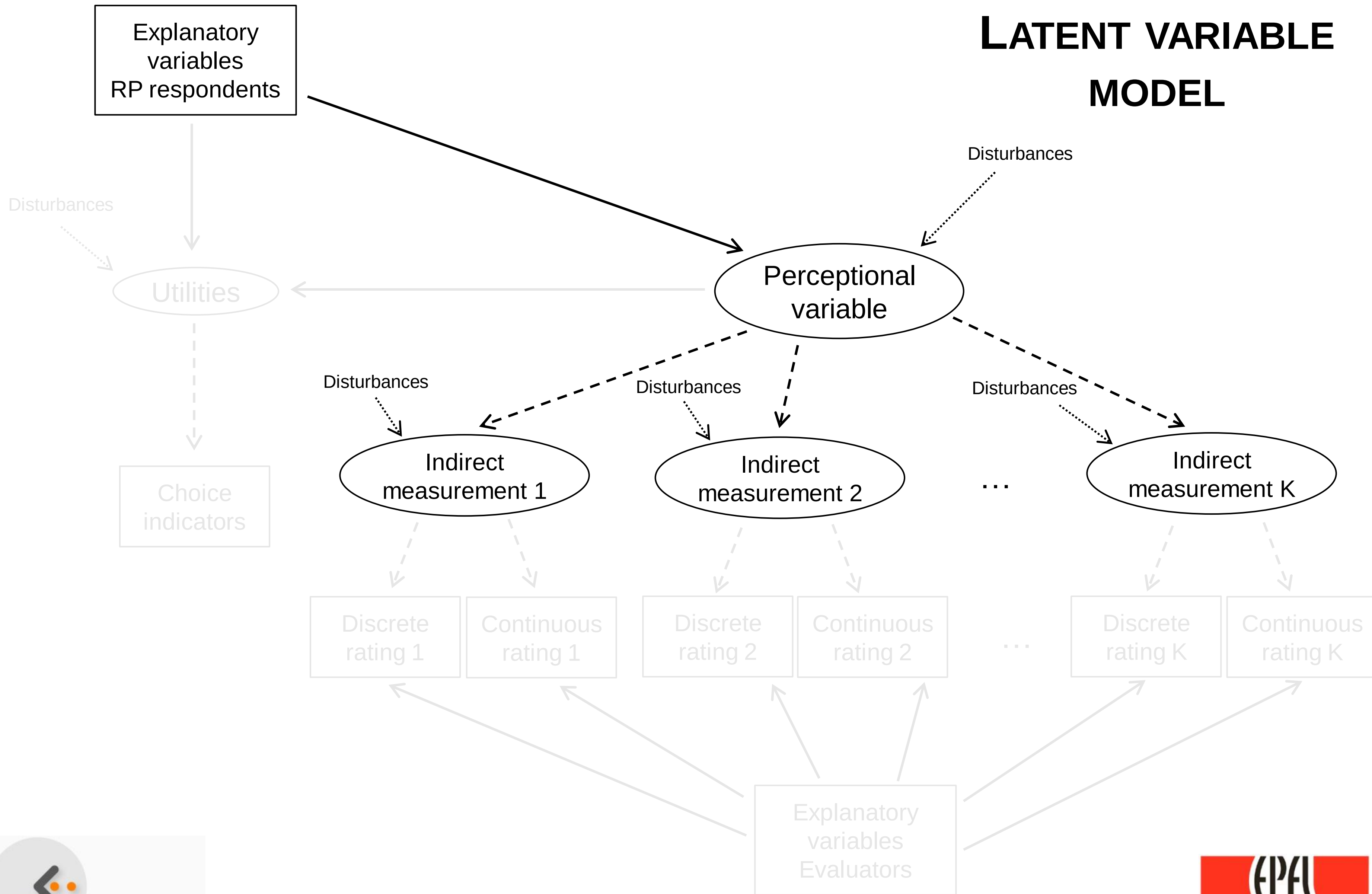
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Discrete choice model is standard:

$$U_{in} = V(X_{in}, X_n^*; \beta) + \varepsilon_{in} \qquad \text{with } \varepsilon_{in} \sim EV(0,1)$$

THE INTEGRATED MODEL FRAMEWORK

LATENT VARIABLE MODEL



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE INTEGRATED MODEL FRAMEWORK

Latent variable model of perception (SEM):

Structural equation:

$$X_n^* = h(X_n; \mu) + \omega_n, \qquad \text{with } \omega_n \sim \mathcal{N}(0, \sigma_\omega)$$

Measurement equation:

$$I_{kn}^* = r_k(X_n^*; \eta_k) + \upsilon_{kn}, \qquad \text{with } \upsilon_{kn} \sim \mathcal{N}(0, \sigma_k)$$

THE INTEGRATED MODEL FRAMEWORK

Latent variable model of perception (SEM):

Structural equation:

$$X_n^* = h(X_n; \mu) + \omega_n, \qquad \text{with } \omega_n \sim \mathcal{N}(0, \sigma_\omega)$$

Measurement equation:

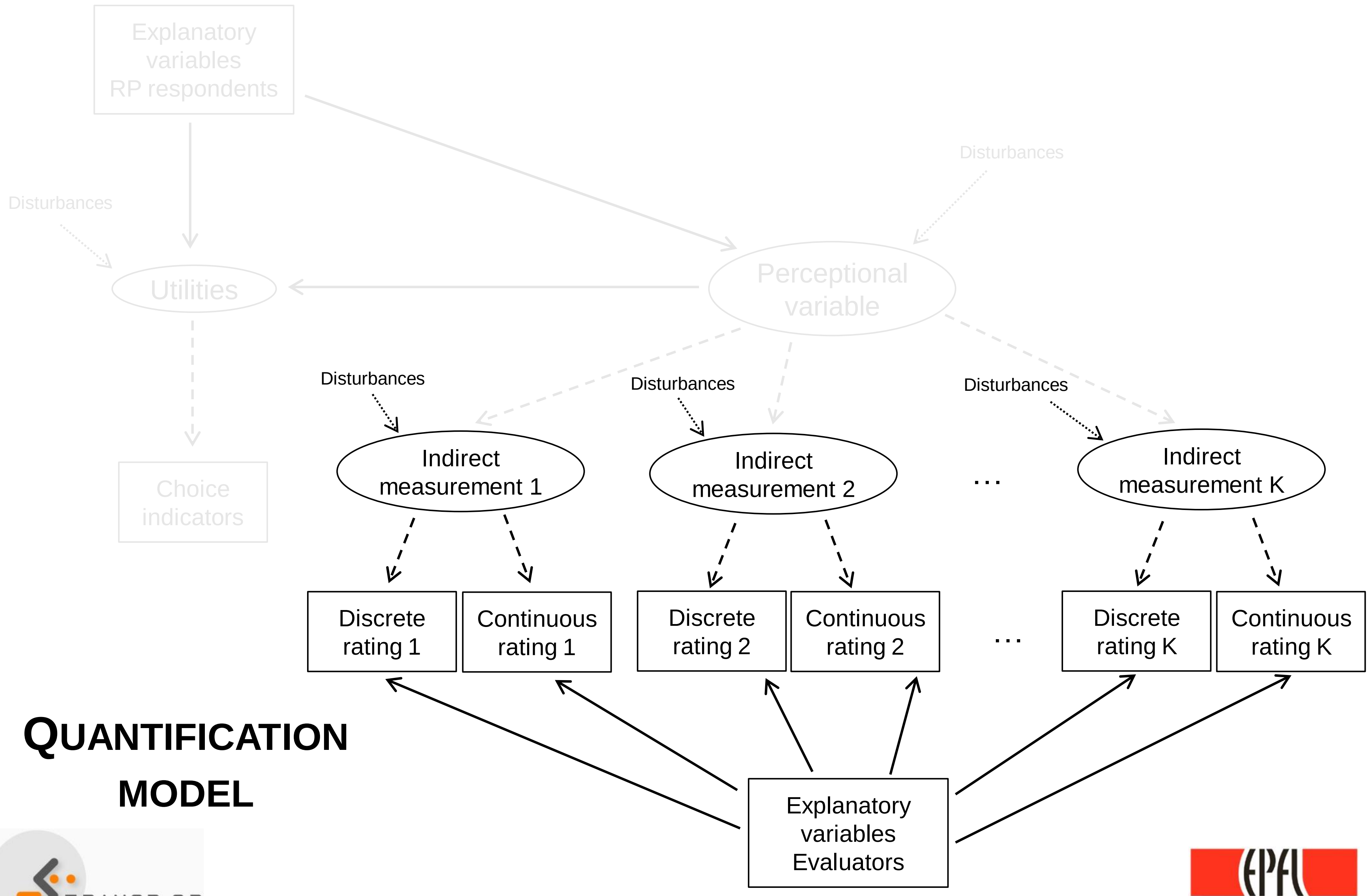
$$\boxed{I_{kn}^*} = r_k(X_n^*; \eta_k) + \upsilon_{kn}, \qquad \text{with } \upsilon_{kn} \sim \mathcal{N}(0, \sigma_k)$$

Unobservable score of indicator *k* for individual *n* ⇒ **Indirect measurement** of perception X_n^*, which is treated as a latent variable

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE INTEGRATED MODEL FRAMEWORK



QUANTIFICATION MODEL

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE INTEGRATED MODEL FRAMEWORK

Quantification model (SEM):

Structural equation:

Score of adjective *l* by individual *m*

$$\boxed{J^*_{lm}} = c_l + \delta_\gamma, \qquad \text{with } \delta_\gamma \sim \mathcal{N}(0, \sigma_\gamma)$$

Measurement equation:

Discrete:

$$\tilde{J}^D_{lm} = \lambda_D \cdot J^*_{lm} + \beta^D_{Xl} \cdot X_m + \delta_D, \text{ with } \delta_D \sim \text{Logistic}(0,1)$$

$$J^D_{lm} = \begin{cases} -2 & \text{if } -\infty < \tilde{J}^*_{lm} \leq \tau_{1l} \\ -1 & \text{if } \tau_{1l} < \tilde{J}^*_{lm} \leq \tau_{2l} \\ 0 & \text{if } \tau_{2l} < \tilde{J}^*_{lm} \leq \tau_{3l} \\ 1 & \text{if } \tau_{3l} < \tilde{J}^*_{lm} \leq \tau_{4l} \\ 2 & \text{if } \tau_{4l} < \tilde{J}^*_{lm} \leq +\infty \end{cases}$$

Continuous:

$$J^C_{lm} = \alpha_C + \lambda_C \cdot J^*_{lm} + \beta^C_{Xl} \cdot X_m + \delta_C, \text{ with } \delta_C \sim \mathcal{N}(0, \sigma_C)$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Quantification model (SEM):

Structural equation:

Adjective-specific constant to be estimated

$$J^*_{lm} = \boxed{c_l} + \delta_\gamma, \qquad \text{with } \delta_\gamma \sim \mathcal{N}(0, \sigma_\gamma)$$

Measurement equation:

Discrete:

$$\tilde{J}^D_{lm} = \lambda_D \cdot J^*_{lm} + \beta^D_{Xl} \cdot X_m + \delta_D, \text{ with } \delta_D \sim \text{Logistic}(0,1)$$

$$J^D_{lm} = \begin{cases} -2 & \text{if } -\infty < \tilde{J}^*_{lm} \leq \tau_{1l} \\ -1 & \text{if } \tau_{1l} < \tilde{J}^*_{lm} \leq \tau_{2l} \\ 0 & \text{if } \tau_{2l} < \tilde{J}^*_{lm} \leq \tau_{3l} \\ 1 & \text{if } \tau_{3l} < \tilde{J}^*_{lm} \leq \tau_{4l} \\ 2 & \text{if } \tau_{4l} < \tilde{J}^*_{lm} \leq +\infty \end{cases}$$

Continuous:

$$J^C_{lm} = \alpha_C + \lambda_C \cdot J^*_{lm} + \beta^C_{Xl} \cdot X_m + \delta_C, \text{ with } \delta_C \sim \mathcal{N}(0, \sigma_C)$$

Quantification model (SEM):

Structural equation:

$$J^*_{lm} = c_l + \delta_\gamma, \qquad \text{with } \delta_\gamma \sim \mathcal{N}(0, \sigma_\gamma)$$

Measurement equation:

Discrete:

$$\tilde{J}^D_{lm} = \lambda_D \cdot J^*_{lm} + \boxed{\beta^D_{Xl} \cdot X_m} + \delta_D, \text{ with } \delta_D \sim \text{Logistic}(0,1)$$

$$J^D_{lm} = \begin{cases} -2 & \text{if } -\infty < \tilde{J}^*_{lm} \leq \tau_{1l} \\ -1 & \text{if } \tau_{1l} < \tilde{J}^*_{lm} \leq \tau_{2l} \\ 0 & \text{if } \tau_{2l} < \tilde{J}^*_{lm} \leq \tau_{3l} \\ 1 & \text{if } \tau_{3l} < \tilde{J}^*_{lm} \leq \tau_{4l} \\ 2 & \text{if } \tau_{4l} < \tilde{J}^*_{lm} \leq +\infty \end{cases}$$

Continuous:

$$J^C_{lm} = \alpha_C + \lambda_C \cdot J^*_{lm} + \boxed{\beta^C_{Xl} \cdot X_m} + \delta_C, \text{ with } \delta_C \sim \mathcal{N}(0, \sigma_C)$$

- Socio-economic information of the evaluator is introduced into measurement equation.
- Heterogeneity in response behavior is handled.

THE INTEGRATED MODEL FRAMEWORK

QUANTIFICATION MODEL

Estimation of the quantification model alone:

- Likelihood for an adjective *l*:

$$\mathcal{L}_l = \prod_{m=1}^{M} \int_{J^*_{lm}} f(J^C_{lm} | J^*_{lm}, X_m; \alpha_C, \lambda_C, \beta^C_X, \sigma_C) f(J^D_{lm} | J^*_{lm}, X_m; \lambda_D, \beta^D_X, \tau_1, \tau_2, \tau_3, \tau_4) f(J^*_{lm} | c_l, \sigma_\gamma) dJ^*_{lm}$$

- Score of adjective *l* by individual *m* is inferred.

$$\boxed{\hat{J^*_{lm}}} = c_l, \forall m$$

- The obtained scores are then introduced as measurements of the perceptional variable.

THE INTEGRATED MODEL FRAMEWORK

Integration of the 3 model components:

- Simultaneous estimation of the DCM and LVM of perception

- Likelihood

$$\mathcal{L} = \prod_{n=1}^{N} \int_{X_n^*} \prod_{i=1}^{I} P(y_{in}|X_{in}, X_n^*; \beta)^{y_{in}} \cdot f(X_n^*|X_n; \mu, \sigma_\omega) \cdot \prod_{k=1}^{K} f(\hat{I_{kn}^*}|X_n^*; \eta_k; \sigma_k) dX_n^*$$

APPLICATION EXAMPLE

Specification

Structural equation:

$$J^*_{lm} = c_l + \delta_\gamma, \qquad \text{with } \delta_\gamma \sim \mathcal{N}(0, \sigma_\gamma)$$

Measurement equations:

Discrete $$\tilde{J}^D_{lm} = \lambda_D \cdot J^*_{lm} + \boxed{\beta^D_{\text{Educ},l} \cdot \text{Educ}_m} + \delta_D, \text{ with } \delta_D \sim \text{Logistic}(0, 1)$$

$$J^D_{lm} = \begin{cases} -2 & \text{if } -\infty < \tilde{J}^*_{lm} \leq \tau_{1l} \\ -1 & \text{if } \tau_{1l} < \tilde{J}^*_{lm} \leq \tau_{2l} \\ 0 & \text{if } \tau_{2l} < \tilde{J}^*_{lm} \leq \tau_{3l} \\ 1 & \text{if } \tau_{3l} < \tilde{J}^*_{lm} \leq \tau_{4l} \\ 2 & \text{if } \tau_{4l} < \tilde{J}^*_{lm} \leq +\infty \end{cases}$$

Observation from exploratory analysis:

Evaluators with higher education level give higher scores.

Continuous $$J^C_{lm} = \alpha_C + \lambda_C \cdot J^*_{lm} + \boxed{\beta^C_{\text{Educ},l} \cdot \text{Educ}_m} + \delta_C, \text{ with } \delta_C \sim \mathcal{N}(0, \sigma_C)$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

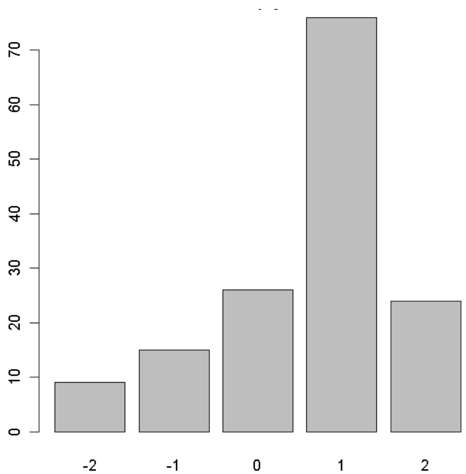
APPLICATION EXAMPLE

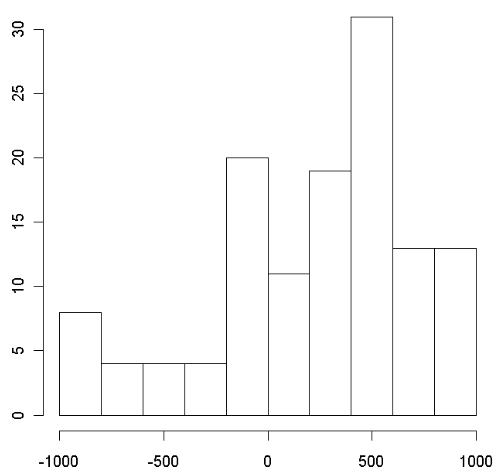
QUANTIFICATION MODEL

Model estimated for all 22 adjectives:

- Separate estimation for each adjective
- Results consistent with expectations

Example: empty





Name	Value	*t*-test
c_{empty}	0.348	29.52
$\beta^{C}_{\text{Educ, empty}}$	0.245	24.29
$\beta^{D}_{\text{Educ, empty}}$	0.372	2.08
$\sigma^{C}_{\text{empty}}$	-2.74	-29.32
$\tau_{1,\text{ empty}}$	-2.72	-7.3
$\delta_{1,\text{ empty}}$	1.23	3.99
$\delta_{2,\text{ empty}}$	1.16	5.49
$\delta_{3,\text{ empty}}$	2.85	10.21

Loglikelihood: - 373

- Constants have expected signs: adjectives related to comfort have + signs.
- Results from exploratory analysis confirmed:
  the higher the level of education, the higher the scores in absolute value.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

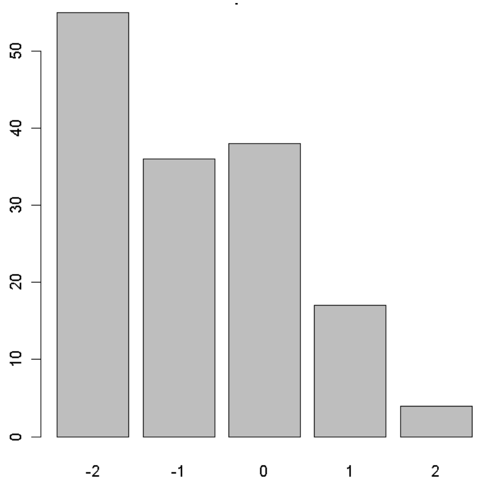
APPLICATION EXAMPLE

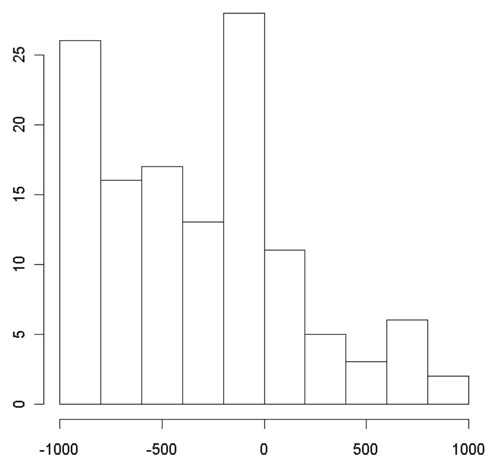
QUANTIFICATION MODEL

Model estimated for all 22 adjectives:

- Separate estimation for each adjective
- Results consistent with expectations

Example: packed





Name	Value	*t*-test
c_{packed}	-0.547	-25.46
$\beta^{C}_{\text{Educ, packed}}$	-0.237	-18.34
$\beta^{D}_{\text{Educ, packed}}$	-0.447	-2.54
$\sigma^{C}_{\text{packed}}$	-2.62	-24.2
$\tau_{1,\text{ packed}}$	-1.43	-6.36
$\delta_{1,\text{ packed}}$	1.23	6.64
$\delta_{2,\text{ packed}}$	1.68	6.77
$\delta_{3,\text{ packed}}$	1.93	3.99

Loglikelihood: - 380

- Constants have expected signs: adjectives related to discomfort have - signs.
- Results from exploratory analysis confirmed:
  the higher the level of education, the higher the scores in absolute value.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

APPLICATION EXAMPLE

INTEGRATED MODEL

Estimation results for the DCM and LVM of perception

Discrete choice model

Name	Value	*t*-test
ASC_{PT}	-0.161	-0.8
ASC_{PMM}	0.42	2.28
β_{Cost}	-0.0653	-8.1
β_{TimePT}	-0.0208	-7.15
$\beta_{TimeCar}$	-0.0323	-9.45
$\beta_{Distance}$	-0.235	-11.44
$\beta_{Work, PT}$	-0.0441	-0.19
$\beta_{Work, PMM}$	-0.575	-2.6
$\beta_{Language, PT}$	-0.0507	-0.17
$\beta_{Language, PMM}$	0.964	3.55
$\beta_{PerceptionComfortPT}$	1.32	4.4

Latent variable model of perception (structural equation)

Name	Value	*t*-test
$b_{meanImageConfortTP}$	7.59	10.41
$b_{regionLanguage}$	-0.726	-2.51
$b_{age<50}$	-1.15	-5.06
b_{actif}	-1.15	-4.72
$b_{voiture}$	-0.727	-3.2

Loglikelihood of the HCM: - 4355

APPLICATION EXAMPLE

Estimation results for the DCM and LVM of perception

Discrete choice model

Name	Value	*t*-test
ASC_{PT}	-0.161	-0.8
ASC_{PMM}	0.42	2.28
β_{Cost}	-0.0653	-8.1
β_{TimePT}	-0.0208	-7.15
$\beta_{TimeCar}$	-0.0323	-9.45
$\beta_{Distance}$	-0.235	-11.44
$\beta_{Work, PT}$	-0.0441	-0.19
$\beta_{Work, PMM}$	-0.575	-2.6
$\beta_{Language, PT}$	-0.0507	-0.17
$\beta_{Language, PMM}$	0.964	3.55
$\beta_{PerceptionComfortPT}$	1.32	4.4

Latent variable model of perception (structural equation)

Name	Value	*t*-test
$b_{meanImageConfortTP}$	7.59	10.41
$b_{regionLanguage}$	-0.726	-2.51
$b_{age<50}$	-1.15	-5.06
b_{actif}	-1.15	-4.72
$b_{voiture}$	-0.727	-3.2

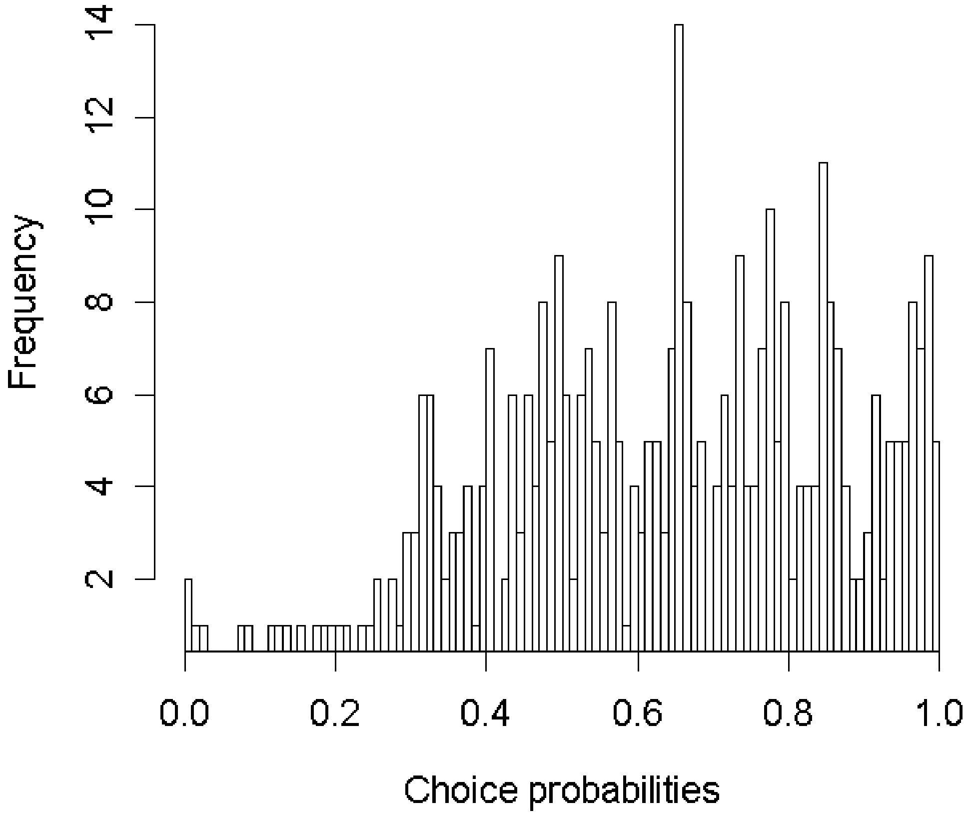
Loglikelihood of the HCM: - 4355

For individuals with a better perception of comfort in PT, the impact of an increase in travel time is less strong.

APPLICATION EXAMPLE

VALIDATION OF THE INTEGRATED MODEL

Model estimation on 80% data and application on 20% data.



Choice probabilities generally well predicted.

CONCLUSION

Main findings:

- Alternative approach to measure perceptions
- Main advantage over classical opinion statements: spontaneity of respondents captured.
- Difficulty: code and integrate these measurements in choice model.

  The proposed model:

  1. Quantifies adjectives
  2. Accounts for subjectivity inherent to quantification method:
     - Uses a fairly large sample of evaluators
     - Account for bias linked to different education levels
- Importance of including individual-level information in measurement component of an LVM in HCM.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

CONCLUSION

Next steps:

- Further validation: comparison of the prediction power of the presented HCM with HCMs including ratings of individual evaluators.

- Estimate the quantification model parts relative to each adjective simultaneously.

Thanks!