

CHOICE MODELING IN R

CLASS 10

Wiktor Budziński

CHOICE MODELING

HYBRID CHOICE MODELS



HYBRID CHOICE MODELS

Preferences can be often driven by the variables which we cannot directly observe

- Risk-aversion
- Perceived quality of the good
- Comfort
- Attitude toward environment

Psychologists and sociologists created different scales that can be used to measure such covariates

- Items in these scales are usually measured on Likert scale
 - It is often not straightforward to introduce such variables directly into the model
-

HYBRID CHOICE MODELS

In psychology researchers usually use factor analysis to identify certain unobserved variables based on some observed measurements of them

Hybrid choice models (HCM) combine the discrete choice framework with factor analysis

- They are sometimes called Integrated Choice Latent Variables (ICLV) models

Formally they are a part of a broader class of Structural Equation Modelling

- Most SEMs utilize a linear structure of models
- As choice models are intrinsically nonlinear, the HCM literature is kind of separate
- Usual the “structure” of HCM will be much simpler, even though the estimation process will be much more complex

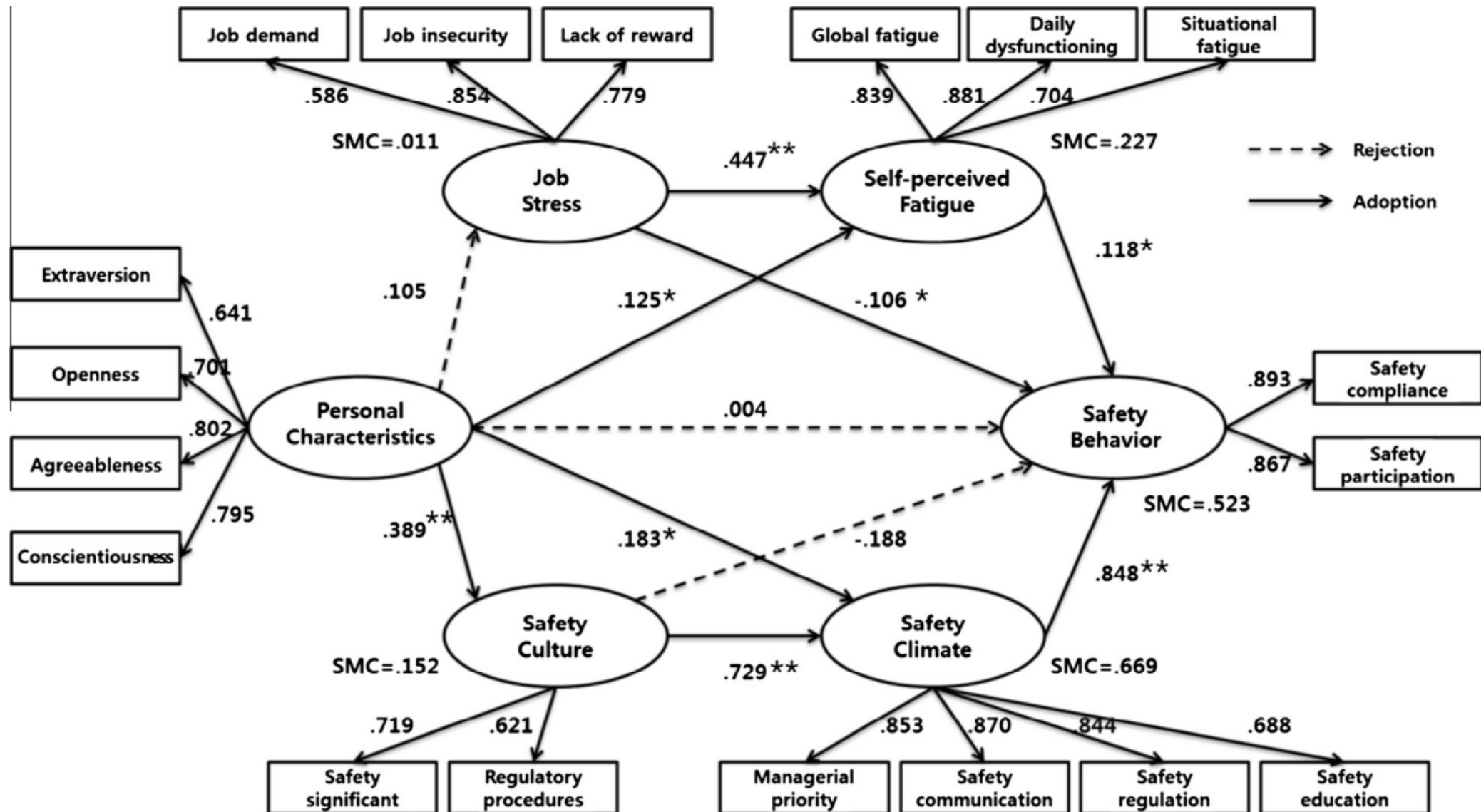


Fig. 2. Structural model of safety behavior of temporary construction workers. Note: Path estimates are standardized coefficients. * $p < .05$ and ** $p < .01$ (two-tailed set).

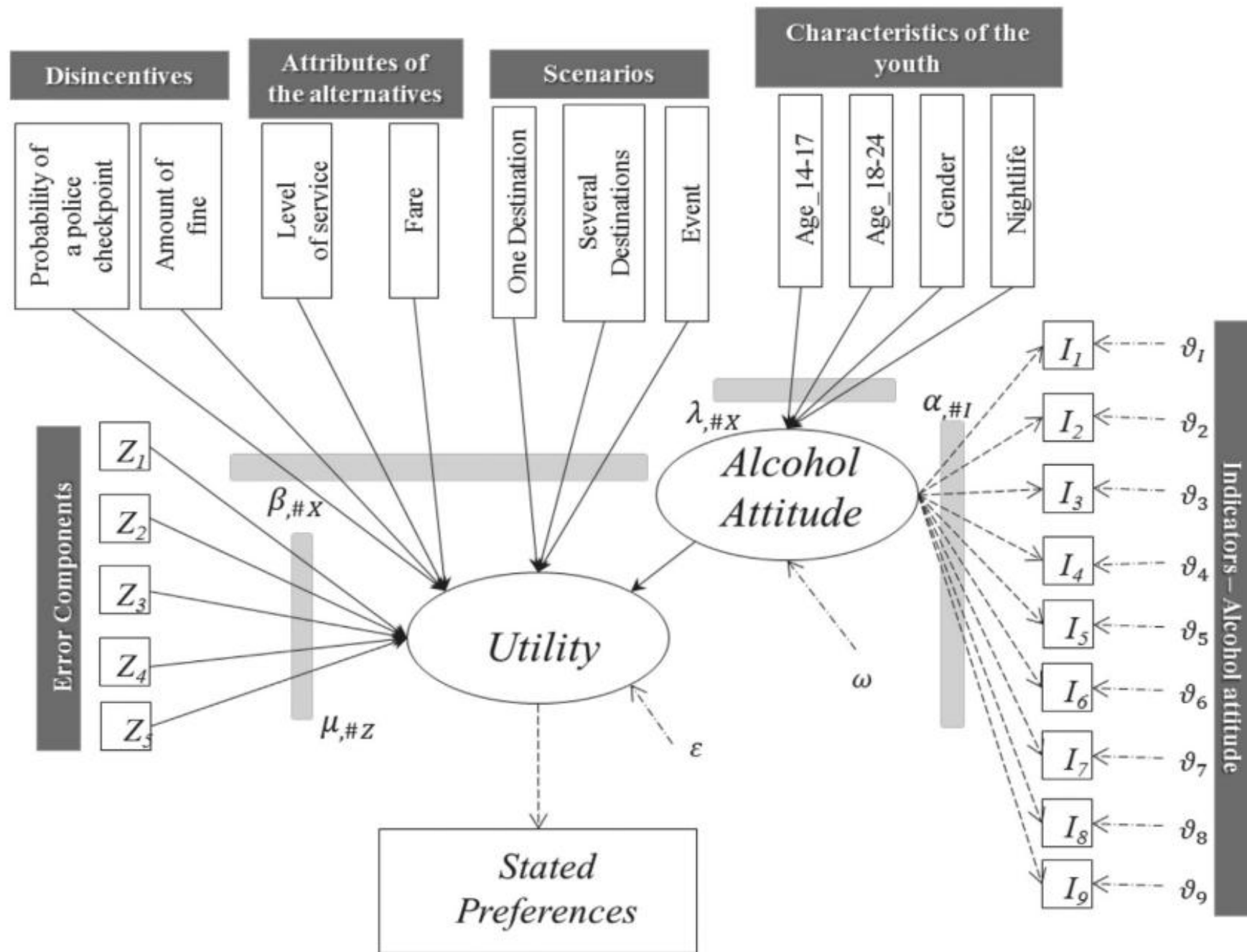
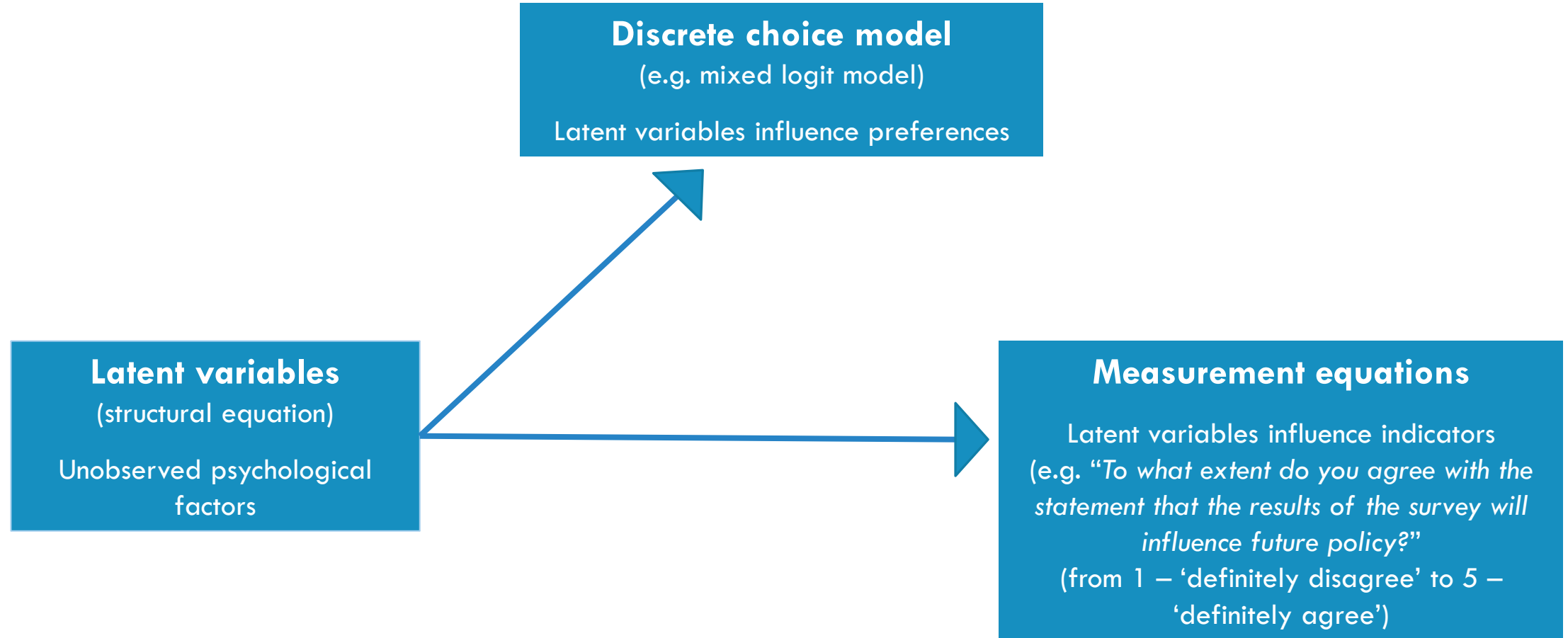


Fig. 2. Full Path Diagram for the Hybrid Choice Model.

HYBRID CHOICE MODELS



HYBRID CHOICE MODELS

The exact functional form of the model with vary depending on the type of the latent factor and the type of indicator variables

Main two types of latent factors are attitudes and perceptions

- Attitudes are usually individual-specific, similar to socio-demographic variables. For example, attitude towards environment
- Perceptions are often alternative-specific, more like attributes. For example, perceived quality.

Indicator variables are most often measured on Likert scale, so order probit/logit specific would be the best choice

- Linear model is sometimes used as a simplifying assumption

CASE STUDY 1: EXPECTED QUALITY OF THE WINE

We will use simulated data inspired by this study

- Palma, D., de Dios Ortúzar, J., Rizzi, L. I., Guevara, C. A., Casaubon, G., & Ma, H. (2016). **Modelling choice when price is a cue for quality: a case study with Chinese consumers.** Journal of Choice Modelling, 19, 24-39.

Consumers care about the quality of the wine, but usually there is little information available about it when making a choice

- Based on past experiences
- Some rating websites/apps
- Using some heuristics

To the researcher the quality measure that consumer uses is unobserved

Wine A		Wine B		Wine C	
	Winery: Don Juan Family <i>A family vineyard</i>		Winery: Union <i>The meeting of two worlds</i>		Winery: Hacienda Cachapoal <i>New World's finest</i>
	Grape variety: Red Blend		Grape variety: Syrah		Grape variety: Cabernet Sauvignon
	Price: RMB 200		Price: RMB 100		Price: RMB 300

3. Using a scale from 1 to 5, where 1 means "I strongly disagree" and 5 means "I strongly agree", please indicate your level of agreement with the following statements *

	1	2	3	4	5
I believe wine A is excellent *	☐	☐	☐	☐	☐
I believe wine B is excellent *	☐	☐	☐	☐	☐
I believe wine C is excellent *	☐	☐	☐	☐	☐

CASE STUDY 1: EXPECTED QUALITY OF THE WINE

The model is following:

$$U_{ijt} = \beta_1 X_{ijt} + \beta_2 Q_{ijt}^* + \beta_3 P_{ijt} + \varepsilon_{ijt}$$

$$Q_{ijt}^* = \alpha_1 X_{ijt} + \alpha_2 P_{ijt} + \eta_{ijt}$$

$$Rating_{ijt}^k = \mu_k + \lambda_k Q_{ijt}^* + v_{ijt}^k$$

The rating variables should not be introduced directly into the model because of the measurement error

For identification usually some additional constraints are needed

- Error term in equation for Q has constant variance

CASE STUDY 1: EXPECTED QUALITY OF THE WINE

The model is following:

$$U_{ijt} = \beta_1 X_{ijt} + \beta_2 Q_{ijt}^* + \beta_3 P_{ijt} + \varepsilon_{ijt}$$

$$Q_{ijt}^* = \alpha_1 X_{ijt} + \alpha_2 P_{ijt} + \eta_{ijt}$$

$$Rating_{ijt}^k = \mu_k + \lambda_k Q_{ijt}^* + v_{ijt}^k$$

The rating variables U_{ijt} and Q_{ijt}^* are the dependent variables of the measurement equation.

For identification usual

- Error term in equation for

Let's go to  

CASE STUDY 2: ATTITUDE TOWARDS ENVIRONMENT

In environmental economics it is likely that preferences are driven by the general attitudes towards environment, ecology etc.

To control for that we can use some socio-psychological scales, like, for example, New Ecological Paradigm (NEP)

1. We are approaching the limit of the number of people the earth can support
2. Humans have the right to modify the natural environment to suit their needs
3. When humans interfere with nature it often produces disastrous consequences
4. Human ingenuity will insure that we do NOT make the earth unlivable
5. Humans are severely abusing the environment
6. The earth has plenty of natural resources if we just learn how to develop them
7. Plants and animals have as much right as humans to exist
8. The balance of nature is strong enough to cope with the impacts of modern industrial nations
9. Despite our special abilities humans are still subject to the laws of nature
10. The so-called “ecological crisis” facing humankind has been greatly exaggerated
11. The earth is like a spaceship with very limited room and resources
12. Humans were meant to rule over the rest of nature
13. The balance of nature is very delicate and easily upset
14. Humans will eventually learn enough about how nature works to be able to control it
15. If things continue on their present course, we will soon experience a major ecological catastrophe

CASE STUDY 2: ATTITUDE TOWARDS ENVIRONMENT

We will use simulated data that replicate the most basic SP setting

	Alternative 1 (current policy)	Alternative 2	Alternative 3
Quality improvement	0%	50%	100%
Annual cost	0 USD	15 USD	50 USD
Your choice			

CASE STUDY 2: ATTITUDE TOWARDS ENVIRONMENT

The model is following:

$$U_{ijt} = \beta_{ASC} SQ_{ijt} + \beta_1 Q_{ijt} + \beta_2 P_{ijt} + \beta_{ASC_LV} SQ_{ijt} LV_i + \beta_{Q_LV} Q_{ijt} LV_i + \varepsilon_{ijt}$$

$$LV_i = \alpha X_i + \eta_i$$

$$Ind_{ik}^* = \lambda_k Q_{ijt}^* + \vartheta_{ik}$$

Ind_{ik}^* is connected to answers to the indicator variables by ordered logit model

The usual constrain in such a model is that the variance of eta is one

CASE STUDY 2: ATTITUDE TOWARDS ENVIRONMENT

The model is following:

$$U_{ijt} = \beta_{ASC} SQ_{ijt} + \beta_1 Q_{ijt} + \beta_2 P_{ijt} + \beta_{ASC_LV} SQ_{ijt} LV_i + \beta_{Q_LV} Q_{ijt} LV_i + \varepsilon_{ijt}$$

$$LV_i = \alpha X_i + \eta_i$$

$$Ind_{ik}^* = \lambda_k Q_{ijt}^* + \vartheta_{ik}$$

Ind_{ik}^* is connected to answers to the indicator variables by ordered logit model

The usual constrain in such

Let's go to  