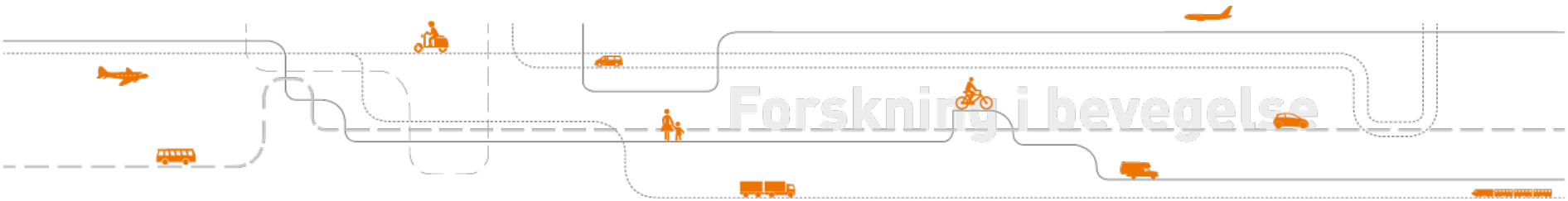
Transportøkonomisk institutt
Stiftelsen Norsk senter for samferdselsforskning

www.toi.no

The effect of travellers' attitudes and perceptions on the demand for high speed rail in Norway

Bjørn Gjerde Johansen*, Farideh Ramjerdi

*Corresponding author. E-mail: bgj@toi.no,
phone: +47 93 06 79 74



Background

- A large-scale study was recently conducted to evaluate the feasibility of high-speed rail (HSR) in Norway (Jernbaneverket, 2012).
- The study indicated that building HSR in Norway is far from economically feasible.
- However, the vast data collected provides exellent possibilities for in-depth analyses of heterogeneity of travellers.

Background

- To capture this heterogeneity, we utilize a model family called *Hybrid choice models* (see Walker, 2001; Walker and Ben-Akiva, 2002; Abou-Zeid and Ben-Akiva, 2014)
- This method focuses on explicitly modelling the decision-making process behind the modal choice
  - Personality traits/attitudes *influence the utility functions, and are modelled as* latent variables
  - *Different (unobserved) segments of individuals behave differently, which is modelled by means of* latent classes

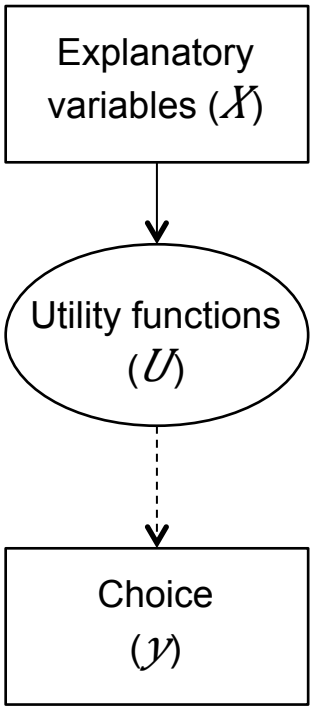
Conventional choice model

- The «black box» is filled by the latent construct «utility»
- Utility has a deterministic and a stochastic component ($V(.)$ and ε, respectively).

$$U_n = V(X_n;\beta) + \varepsilon_n$$

$$y_n = \begin{cases} 1 & if\ U_{in} \geq U_{jn}\ \forall j \in C_n \\ 0 & otherwise \end{cases}$$

$$P(y_n | X_n;\beta,\Sigma_\varepsilon) = \int_{\varepsilon} P(y_n | X_n;\beta,\varepsilon)\, f(\varepsilon|\Sigma_\varepsilon)\, d\varepsilon$$



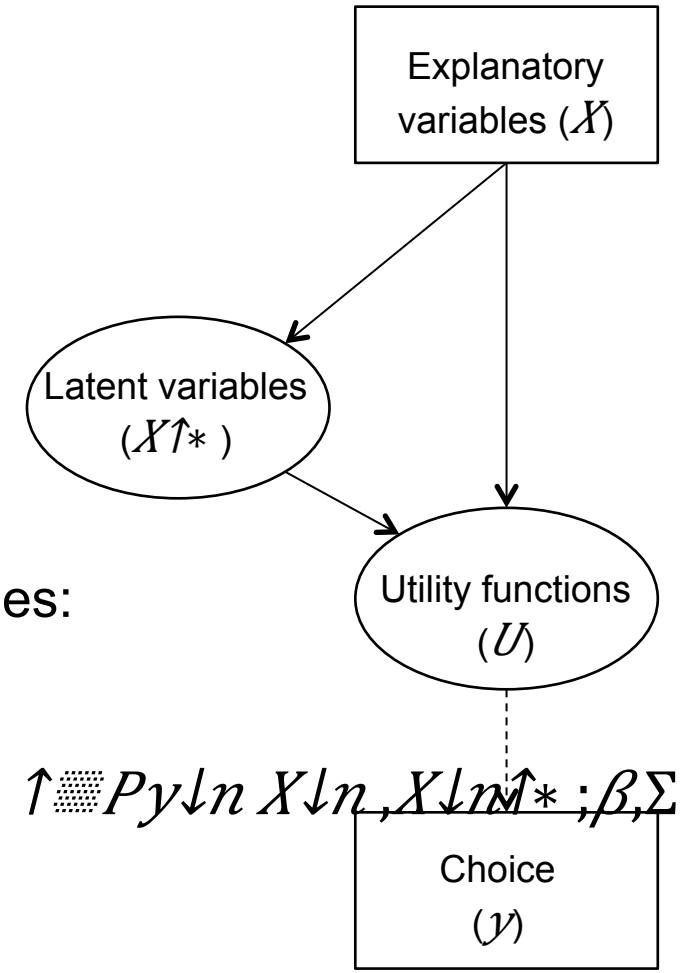
Hypothesis I

- Personality traits (*preference for comfort* and *global environmental conciousness*) affect the mode utilities, and hence the choice probabilities. Therefore, they should be included in the utility functions.
- This can be achieved by modelling the personality traits as latent variables.

Latent variables

- One more equation is added:

$$X\downarrow n\uparrow* = X\uparrow* (X\downarrow n ; \alpha) + \omega$$



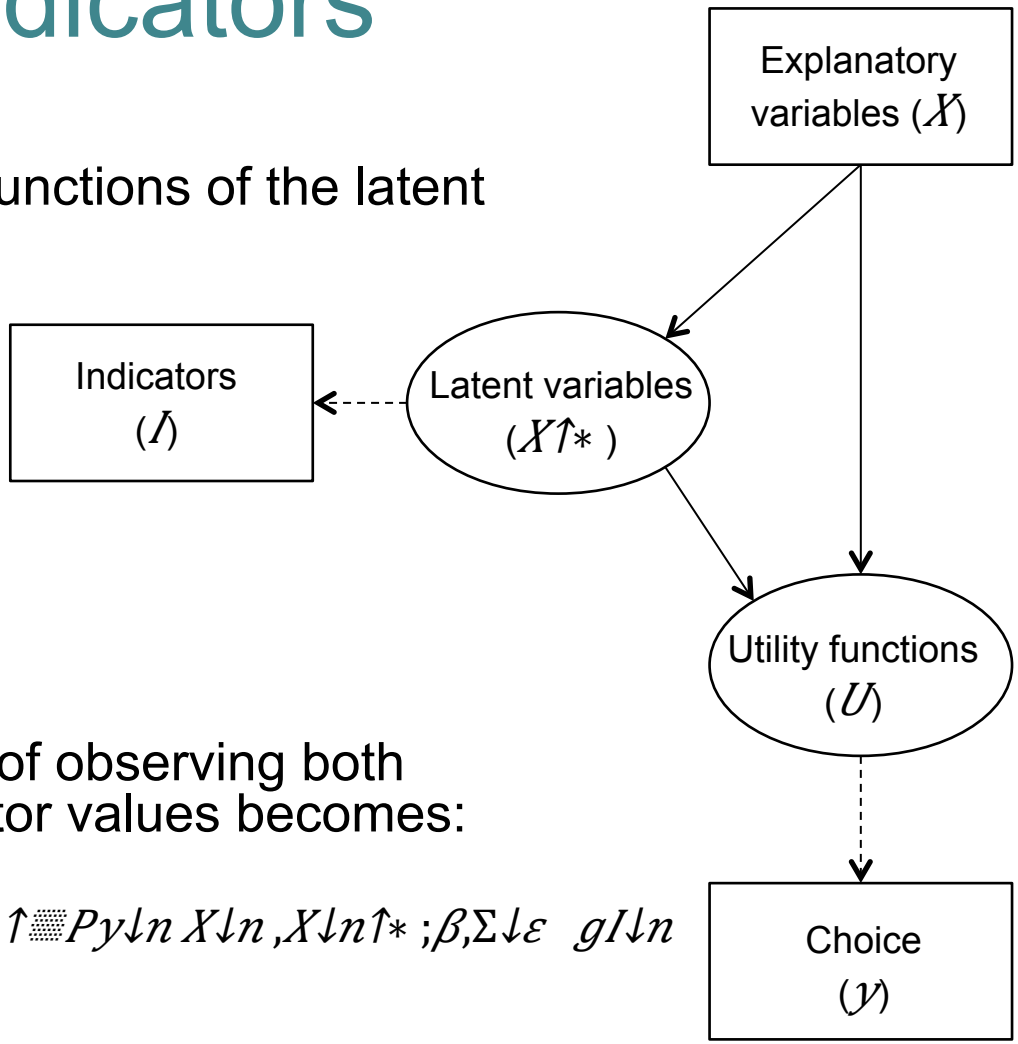
… and the choice probability becomes:

$$Py\downarrow n\, X\downarrow n ; \beta, \alpha, \Sigma\downarrow\varepsilon, \Sigma\downarrow\omega = \int X\uparrow* \uparrow Py\downarrow n\, X\downarrow n, X\downarrow n\uparrow* ; \beta, \Sigma\downarrow\varepsilon\ f$$
$$X\downarrow n\uparrow* X\downarrow n ; \alpha, \Sigma\downarrow\omega\ dX\uparrow*$$

Introducing indicators

- Indicators expressed as functions of the latent variables:

$$I_n = I(X_n^*; \lambda,) + v_n$$



- Then the joint probability of observing both the choice and the indicator values becomes:

$$,I_n\, X_n ;\beta,\alpha,\lambda,\Sigma_\varepsilon ,\Sigma_\omega ,\Sigma_v = \int X^* \uparrow\!\!\!\!\blacksquare P y_n\, X_n ,X_n^* ;\beta,\Sigma_\varepsilon \;\; g I_n * ;\lambda,\Sigma_v \;\; f X_n^*\, X_n ;\alpha,\Sigma_\omega \;\; dX^*$$

Hypothesis II

- The conventional separation between business trips and leisure trips is too rigid. By identifying latent segments (classes) of the population, one is better able to capture the underlying behavior and hence increase the predictive power of the model.

Latent classes

- Unobserved latent constructs affect class membership (analogous to utility):

$$H\downarrow n\uparrow s = H(X\downarrow n ; \gamma\uparrow s) + \tau\downarrow n\uparrow s$$

- Assuming the class membership model can be estimated, the class membership probabilities can be written as:

$$P(s\downarrow n \mid X\downarrow n ; \gamma, \Sigma\downarrow\tau)$$

- The choice probability can be written as:

$$Py\downarrow n, I\downarrow n\, X\downarrow n ; \beta, \alpha, \lambda, \gamma, \Sigma\downarrow\varepsilon, \Sigma\downarrow\omega, \Sigma\downarrow\upsilon, \Sigma\downarrow\tau = \sum s=1\uparrow S \blacksquare P(s|X\downarrow n ; \gamma, \Sigma\downarrow\tau) \int X\uparrow *$$
$$\uparrow \blacksquare Py\downarrow n\, X\downarrow n, X\downarrow n\uparrow * ; \beta\uparrow s, \Sigma\downarrow\varepsilon\uparrow s \; gI\downarrow n\, X\downarrow n\uparrow * ; \lambda, \Sigma\downarrow\upsilon \; fX\downarrow n\uparrow * \; X\downarrow n ; \alpha, \Sigma\downarrow\omega \; dX\uparrow *$$

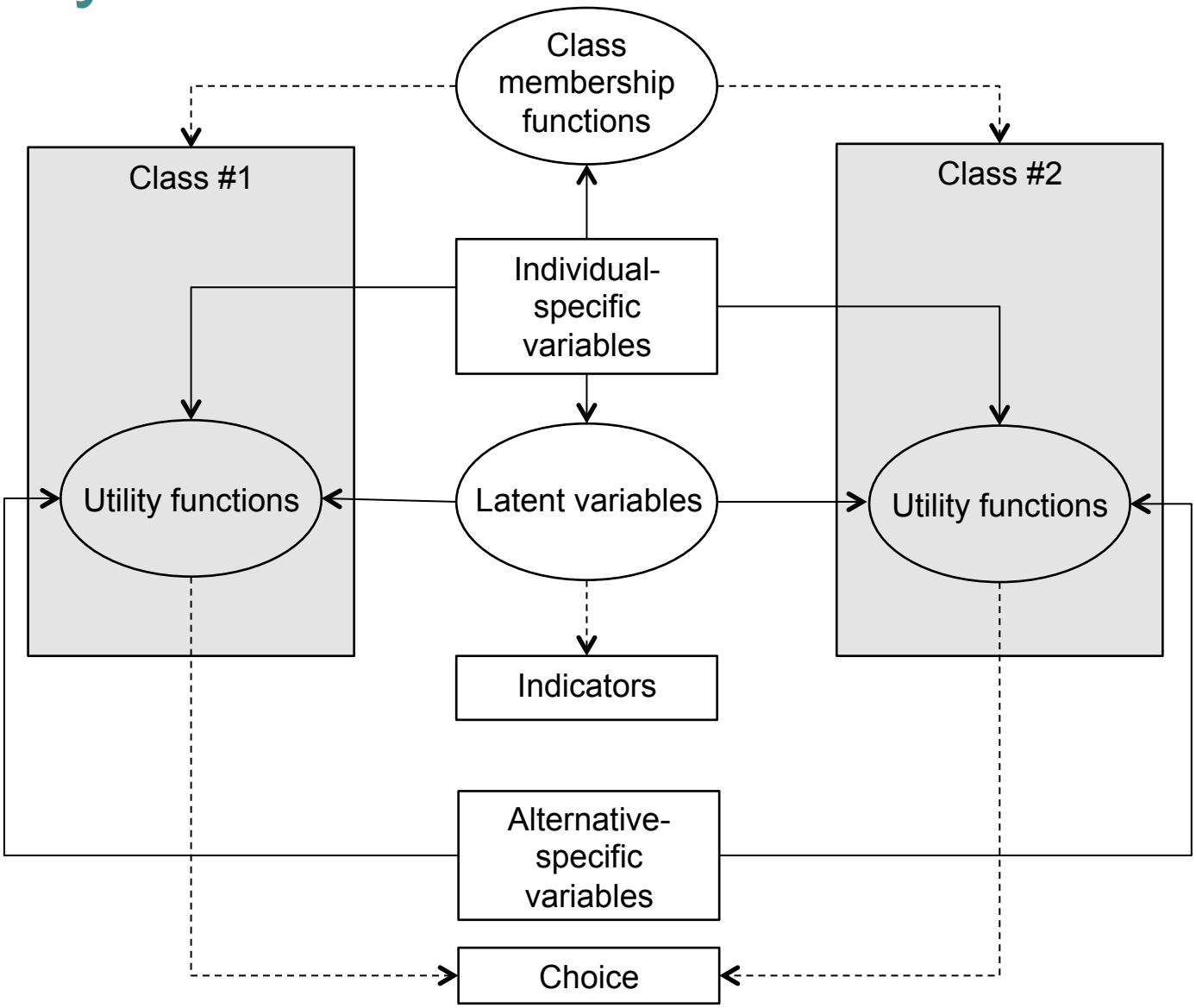
Class membership model

Choice model

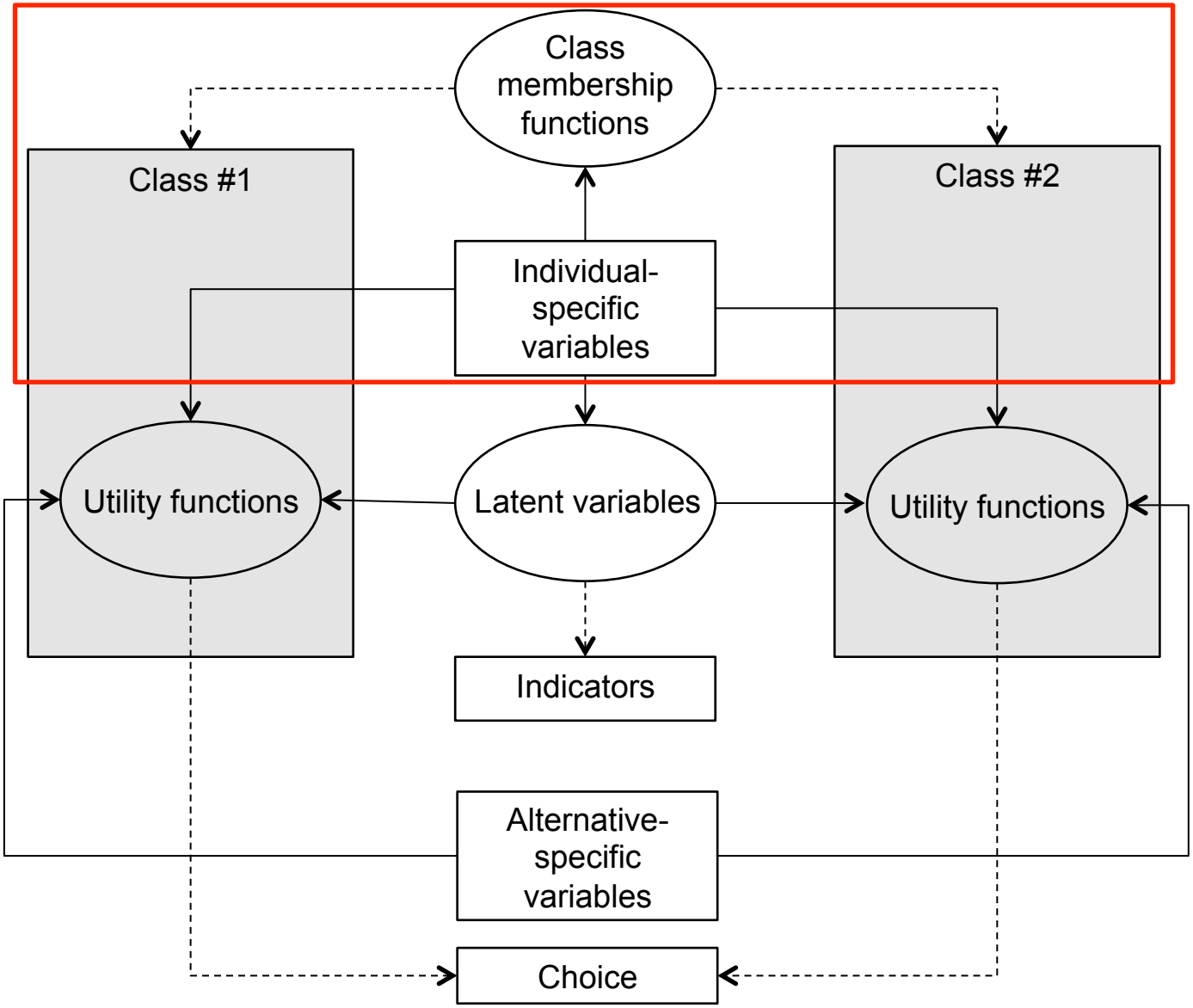
Distribution of indicators

Distribution of latent variables

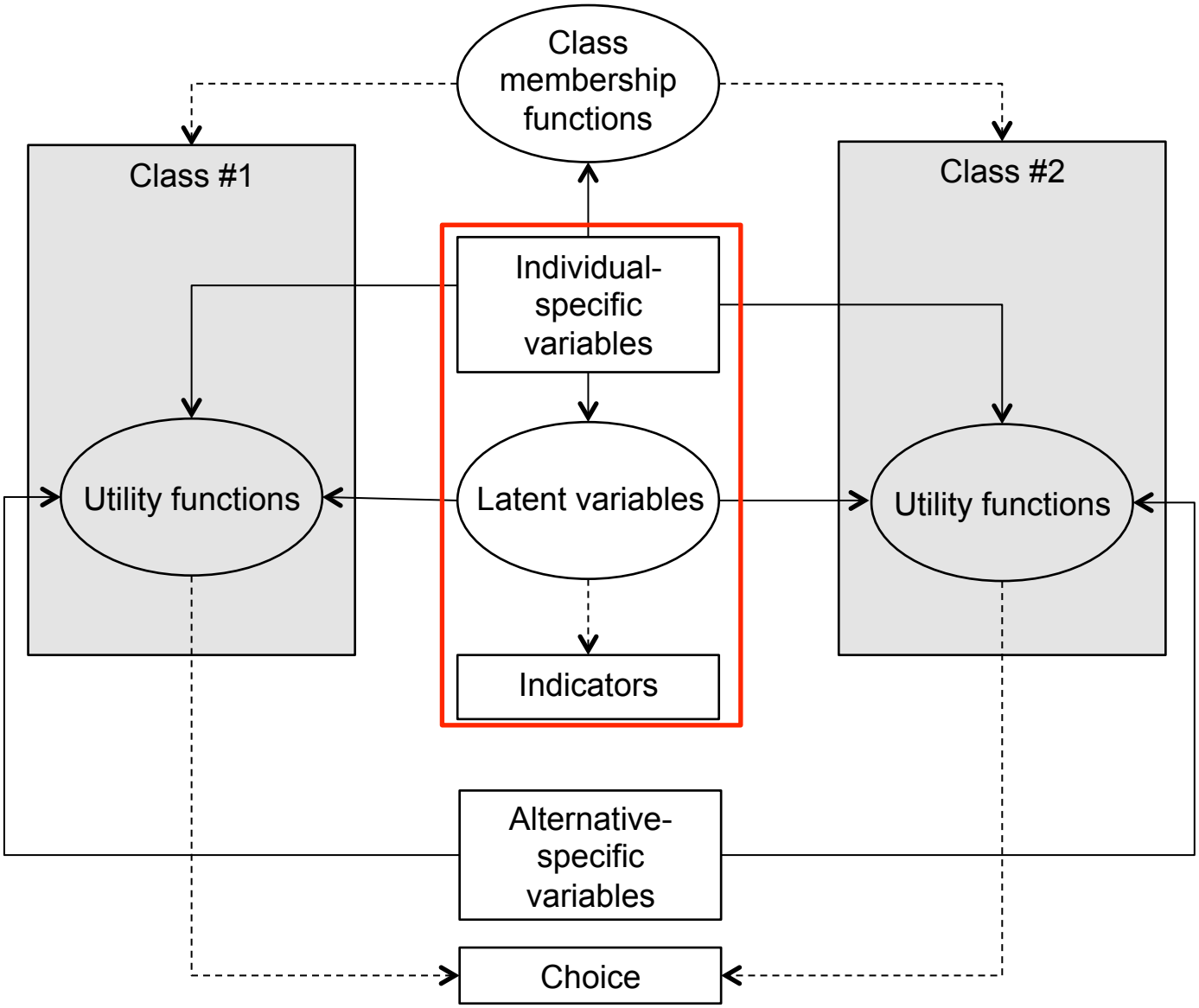
The hybrid choice model



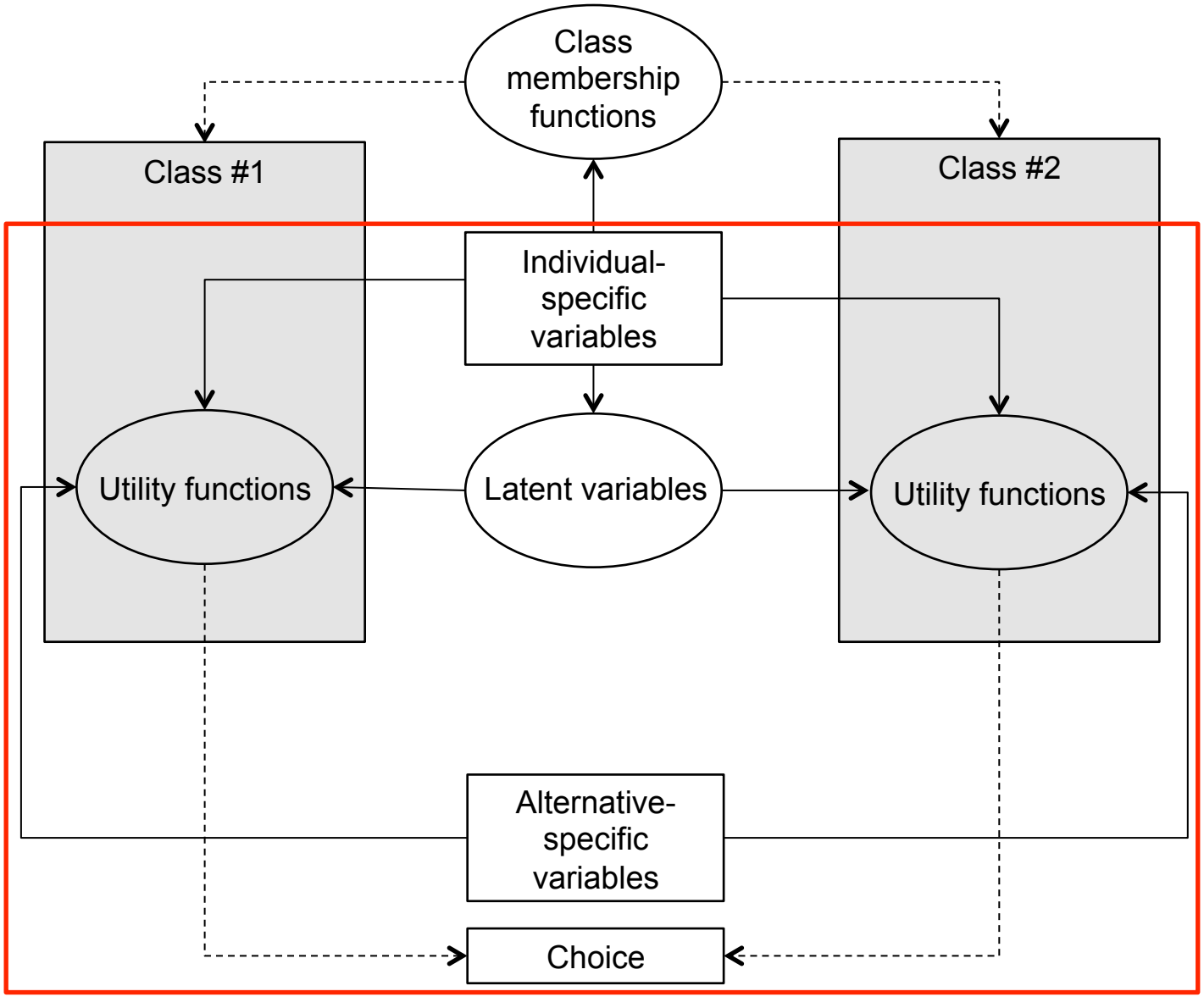
The latent class model



The latent variable model



The choice model



Weaknesses

- Large data requirements
- Latent variables explained by socio-economic characteristics
- It is difficult to find strong predictors of latent variables
- High degree of collinearity between latent variables
- Local optima
- Endogeneity of indicators

Strengths (from Abou-Zeid and Ben-Akiva, 2014)

- Accounting for unobserved taste heterogeneity
- Increased efficiency
- Increased behavioral realism
- Policy relevance

Data (Halse, 2012)

Reference trip:							
Car		Rail		Bus		Plane	
SP choices:		**SP choices:**		**SP choices:**		**SP choices:**	
Car	HSR	Rail	HSR	Bus	HSR	Plane	HSR

Attribute values:

- Time
- Cost
- Share of time in tunnel
- Departures per day
- Access time
- Egress time

Data (Johansson et al., 2006)

	Question	Target dimensions
1	*How important is it for you* to be able to control the conditions around you (air condition, noise, music)?	Comfort
2	... to be able to rest on your trip?	
3	... to be able to work on your trip?	
4	... to avoid changing the mode of transport?	
5	... to know in advance how long the trip will take?	Reliability
6	... to have little or no variation in travel time?	
7	... to avoid congestion?	
8	... to have the opportunity to shop and make other errands?	Flexibility
9	... to be able to choose departure time yourself and be able to change it in short notice?	
10	... to have a car available at the destination?	
11	... to be able to choose travel route yourself and change it on the way?	

Data (Johansson et al., 2006)

12	*How often do you* recycle batteries?	Local environmental consciousness
13s	...leave your garbage on the ground if there is no garbage can?	
14	...engage yourself to impede construction works and other activities that intervene nature?	
15	...visit unspoiled nature in order to experience it?	
16	...use a cycling helmet when you cycle?	Safety
17	...keep the speed limit when driving?	
18	...use the reflex when you walk in traffic in the dark?	
19s	...do things that are dangerous or illegal for fun?	
20s	...heat up your house so one does not have to use a sweater?	Global environmental consciousness
21	...turn off the lights before you leave the room?	
22	...bring shopping bags/used plastic bags when shopping?	
23	...do you eat dinner without meat?	

β	Class 1		Class 2		Generic	
Variables	Coef.	t-test	Coef.	t-test	Coef.	t-test
ASC_car	0.00	---	0.00	---		
ASC_air	-1.49	(-0.98)	-2.33	(-1.14)		
ASC_train	-4.08	(-2.00)	-0.429	(-0.18)		
ASC_bus	-4.05	(-1.66)	4.35	(1.67)		
ASC_HSR	-3.15	(-2.00)	0.705	(0.40)		
Time_car	-0.198	(-1.65)	-0.519	(-2.67)		
Time_air	-1.02	(-3.85)	-1.56	(-2.51)		
Time_train	-0.437	(-2.77)	-1.62	(-6.71)		
Time_bus	0.0200	(0.10)	-1.82	(-5.22)		
Time_HSR	-0.985	(-9.17)	-1.59	(-8.66)		
Cost	-0.278	(-7.17)	-0.376	(-9.30)		
Comfort_car					0.00	---
Comfort_air					0.410	(1.43)
Comfort_train					0.854	(2.30)
Comfort_bus					0.548	(1.56)
Comfort_HSR					0.915	(3.54)
Tunnel					0.00248	(0.71)
Departures					0.0373	(4.65)
Time_access_egress					-0.360	(-2.13)
Female_car					0.00	---
Female_air					-0.247	(-0.57)
Female_train					0.0183	(0.04)
Female_bus					-0.296	(-0.63)
Female_HSR					-0.194	(-0.54)
Age_car					0.00	---
Age_air					-0.400	(-0.24)
Age_train					0.363	(0.21)
Age_bus					-0.222	(-0.13)
Age_HSR					-1.69	(-1.09)

γ	Class 1		Class 2		Generic	
Variables	Coef.	t-test	Coef.	t-test	Coef.	t-test
Constant	0.00	---	-0.331	(-1.09)		
Age	0.00	---	0.405	(0.67)		
Female	0.00	---	-0.371	(-2.05)		
Business	0.00	---	-0.165	(-0.90)		

α	Class 1		Class 2		Generic	
Variables	Coef.	t-test	Coef.	t-test	Coef.	t-test
Constant					4.35	(32.80)
Age					-1.16	(-4.56)
Gender					0.245	(2.12)
Standard deviation					0.858	(8.60)

λ	Class 1		Class 2		Generic	
Indicators	Coef.	t-test	Coef.	t-test	Coef.	t-test
I_1					0.251	(3.31)
I_2					1.00	---
I_3					0.750	(4.30)
I_4					0.225	(2.69)

- Class 2 is more sensitive to time and cost
- Class 2 have a larger share of males
- An increase in comfort of one standard error increases the probability of choosing rail by 6.3% and the probability of choosing HSR by 19.6%
- The predicted value of comfort is higher for females and decreases with age
- All indicators are influenced positively (and significantly) by comfort

Conclusions

- Hypothesis I:
  - *Both personality traits (comfort and environmental consciousness) are significant. Moreover, they affect the choice probability for HSR positively, and to a larger extent than available individual-specific characteristics*

- Hypothesis II:
  - *The identified latent classes differ from the conventional separation between business and leisure trips. Moreover, the latent class model have higher explanatory power than a model in which leisure trips and business trips are separated*

Further work

- Take into account the panel structure
- Estimate the model with three classes to (try to) capture the business segment
- Include both personality traits in the same regression (by simulation?)
- How to be more confident that the solution is the global maximum?

- Other suggestions/comments?

References

- Abou-Zeid, M. and Ben-Akiva, M. 2014. “Hybrid choice models”, forthcoming in *Handbook of Choice Modelling,* Hess, S. and Daly, A. (eds), Edward Elgar Publishing Ltd.
- Halse, A. 2012*. Demand for high speed rail in Norway; Survey design and data collection.* TØI Working Paper 50067.
- Jernbaneverket. 2012 . *Høyhastighetsutredningen 2010-2012, konklusjoner og oppsummering av arbeidet i fase 3*, del 1. *Accessible at: http://www.jernbaneverket.no/PageFiles/17299/Rapport Del 1.pdf*.
- Walker, J. 2001. *Extended discrete choice models: integrated framework, flexible error structures, and latent variables*. Ph. D. thesis, Massachusetts Institute of Technology.
- Walker, J., and Ben-Akiva, M. (2002). Generalized random utility model. *Mathematical Social Sciences*, *43*(3), 303-343.