Some unsolved problems in structural equation models – Incorporating formative indicators into covariance-based SEM

Dirk Temme
Lutz Hildebrandt

Humboldt University Berlin

Meeting SFB 649 "Economic Risk"

Motzen, May 8-10, 2008

Discussion on the "correct" specification of measurement models

- Until the end of the 1990s, scale development in marketing research mainly based on the "Churchill paradigm" (*Churchill* 1979) → reflective measurement model as default → "LISREL" approach (e. g., *Anderson/Gerbing* 1982).
- "Churchill paradigm" has also been applied to theoretical concepts in management and organizational research (e. g., "strategies", "market orientation") as well (*Shook et al.* 2004).
- Recently, (mis-)specification of measurement models "hot topic" in marketing:
  - *Diamantopoulos/Winklhofer* (2001)
  - *Rossiter* (2002)
  - *Jarvis/MacKenzie/Podsakoff* (2003)
  - *Albers/Hildebrandt* (2006)

→ many constructs should better be modelled in a formative manner

→ treating them as reflective can produce grossly misleading results

Formative measurement model

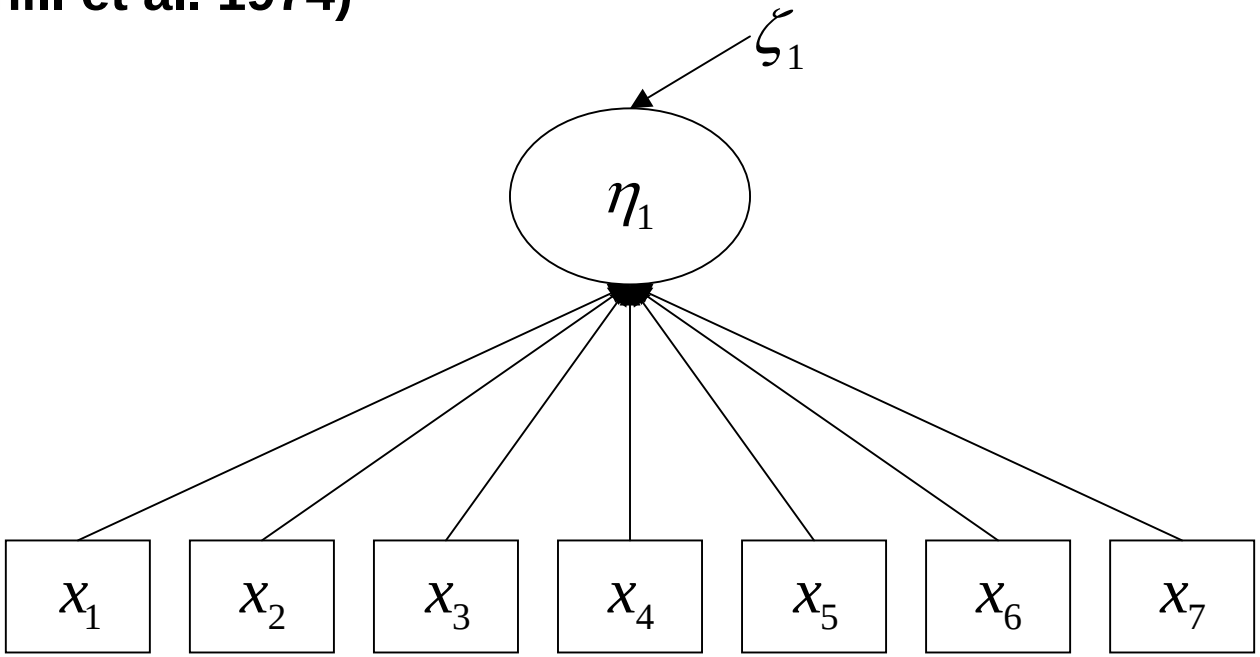
a) Formative construct operationalized as a *composite variable* (CV; *McDonald* 1996):

$$\eta_j = \gamma_{j1} x_1 + \gamma_{j2} x_2 + \ldots + \gamma_{jQ_j} x_{Q_j}$$

b) Formative construct operationalized as a *composite latent variable* (CLV; *Jarvis et al.* 2003):

$$\eta_j = \gamma_{j1} x_1 + \gamma_{j2} x_2 + \ldots + \gamma_{jQ_j} x_{Q_j} + \zeta_j$$

Example of a formative construct – Job satisfaction (Churchill et al. 1974)



x_1 : The job

x_2: Fellow workers

x_3: Supervision

x_4: Company policy and support

x_5: Pay

x_6: Promotion and advancement

x_7: Customers

Estimation of formative measurement models – Alternative approaches

- Discussion on measurement models (as well as new software programs, see *Temme/Kreis/Hildebrandt* 2008 for an overview) has stimulated renewed interest in partial least squares (PLS, *Wold* 1966) analysis (e.g., *Bliemel et al.* 2005; *Vinci et al.* 2008, Handbook on PLS & Marketing).

- Advantages of partial least squares compared to covariance structure analysis:
  - identification of formative measurement models not an issue
  - scores for (formative) constructs are a core element in parameter estimation
  - no distributional assumptions
  - small samples

Wide-spread opinions about specifying formative measurement models

- "The PLS method handles both types of indicators [i.e. formative and reflective indicators], whereas other path analytical methods (e.g., LISREL, EQS) can handle only reflective indicators." (*Dellande/Gilly/Graham* 2004, JoM, p. 84)
- "[In covariance structure analysis] all indicators must be treated in a reflective manner where they are causally influenced by an underlying construct." (*Chin* 1998, p. 301)
- "We chose PLS because of the formative nature of the higher-order value construct." (*Ulaga/Eggert* 2006, JoM, p. 129)
- "A PLS analysis is most appropriate when the model incorporates both formative and reflective indicators […]." (*White/Varadarajan/Dacin* 2003, JoM, p. 71)
- "Because the latent variable scores are determinate, PLS can be used to model formative indicators […]." (*Reinartz/Krafft/Hoyer* 2004, JoMR, p. 298)

Estimation of formative measurement models – Alternative approaches (cont.)

- But: Covariance structure analysis ("LISREL") can handle formative constructs too (e.g., MIMIC models, *Jöreskog/Goldberger* 1975)!
- Disadvantages of PLS compared to covariance structure analysis:
  - parameter estimates are known to be biased for all practical purposes (e.g. *Dijkstra* 1983),
  - does not test for model constraints → no theory testing,
  - no global fit criteria (except GoF, *Tenenhaus et al.* 2005),
  - assumes that formative constructs can be measured without error → construct operationally defined, no surplus meaning,
  - does not separate between "direct" and "indirect" effects of explanatory variables on a formative construct.
- → Good reasons to use "LISREL" provided that the model can be identified.

Outline

1. Motivation
2. Identification of recursive SEM including formative constructs – A new approach
3. Specification of SEM including endogenous formative constructs
   3.1 Exogenous vs. endogenous indicators for endogenous CLV
   3.2 Empirical study
4. Conclusions

Identification conditions (e.g., *Temme* 2006)

Basic identification conditions:

t - rule : $t \leq \frac{1}{2}(p+q)(p+q+1)$

Scaling rule:

1. Fix the path from a formative indicator to the CLV (e. g., to 1).
2. Fix the path from the CLV to one of its reflective indicators or to a latent variable measured by reflective indicators.
3. Fix the variance of the CLV to 1.

Further identification conditions (as stated in the literature):

- specifying at least two additional reflective indicators (MIMIC models, *Jöreskog/Goldberger* 1975) or, alternatively
- specifying direct effects on at least two unrelated reflective latent variables (*MacCallum/Browne* 1993; *Bollen/Davis* 1994; *Jarvis/MacKenzie/Podsakoff* 2003)

Smallest identified formative measurement model

$$\eta_1^F = 1 \cdot x_1 + \gamma_{12} x_2 + \cdots + \gamma_{1Q_1} x_{Q_1} + \zeta_1,$$

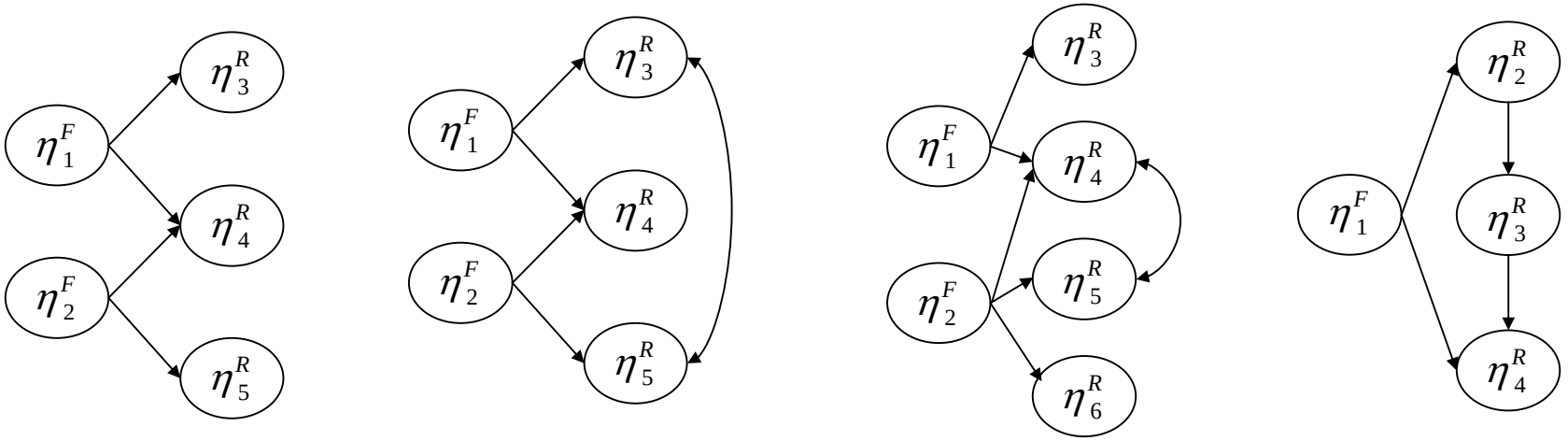
$$\eta_2^R = \beta_{21} \eta_1^F + \zeta_2,$$

$$\eta_3^R = \beta_{31} \eta_1^F + \zeta_3,$$

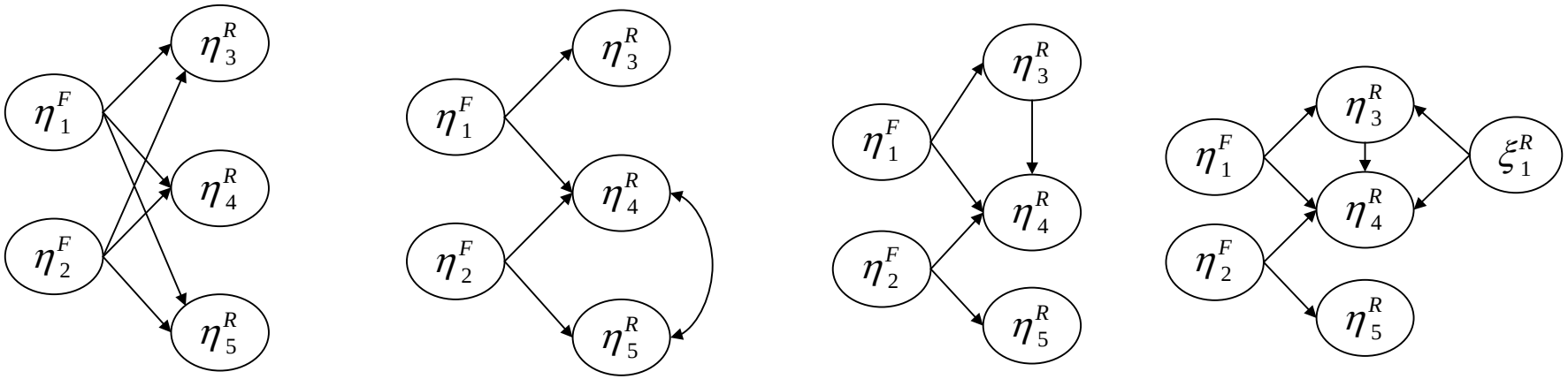
where η_2^R and η_3^R are reflective constructs which have identified measurement models (e.g., three reflective indicators each).

More complex models

Identified models



Non-identified models

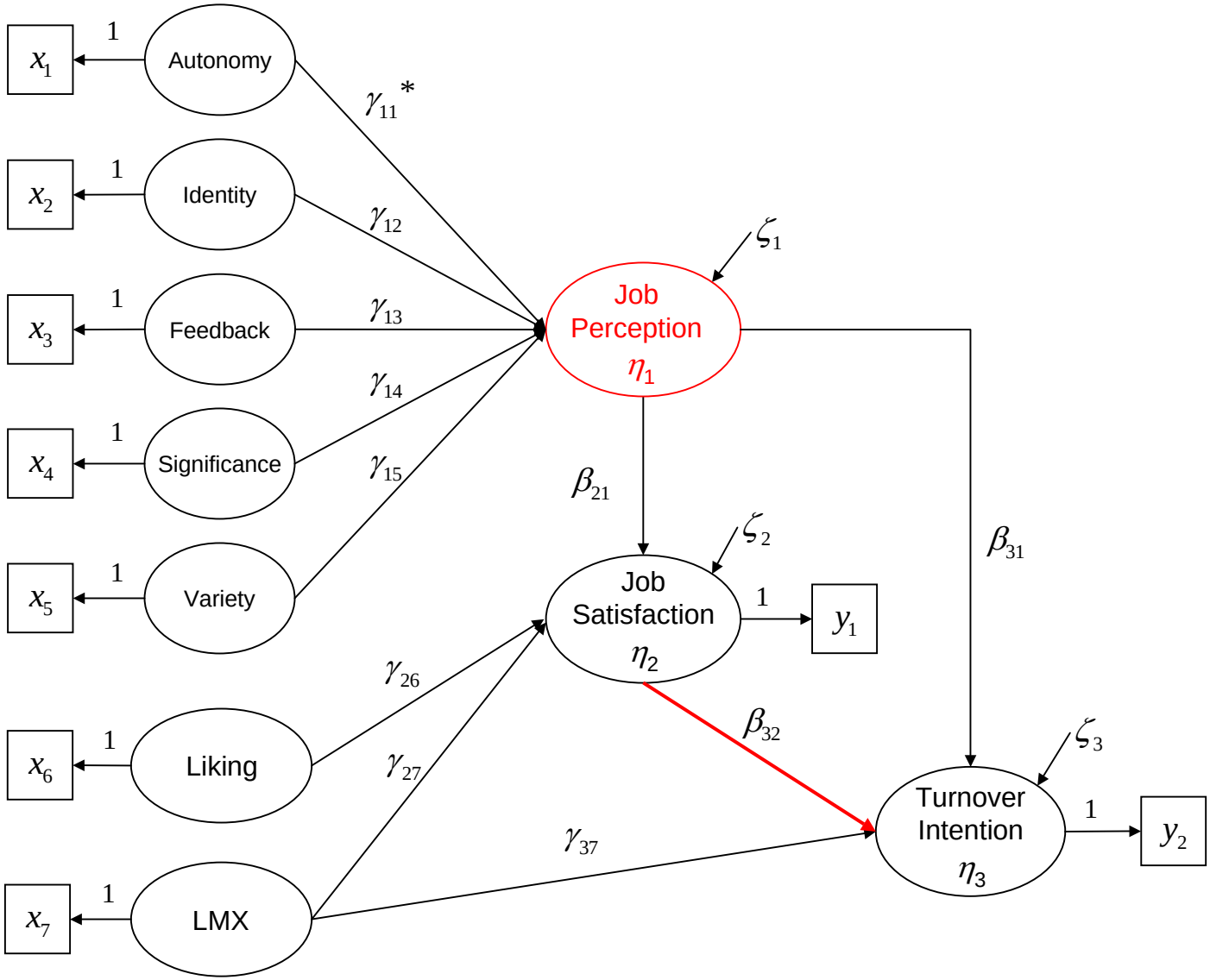


F = formative measurement model
R = reflective measurement model

Further identification options?

Does identification of formative constructs always require "at least two unrelated latent constructs with reflective indicators" (*Jarvis/MacKenzie/Podsakoff* 2003)?

Example: Model by *Law/Wong* (1999)



* fixed to 1

Further identification options? (cont.)

"[…] the error term for the job perception construct was not identified […]. This is because the job perception construct did not have two paths emanating from it that led to independent constructs. It had two paths leading from it, but the two constructs were causally related." (*MacKenzie/Podsakoff/Jarvis* 2005, p. 716)

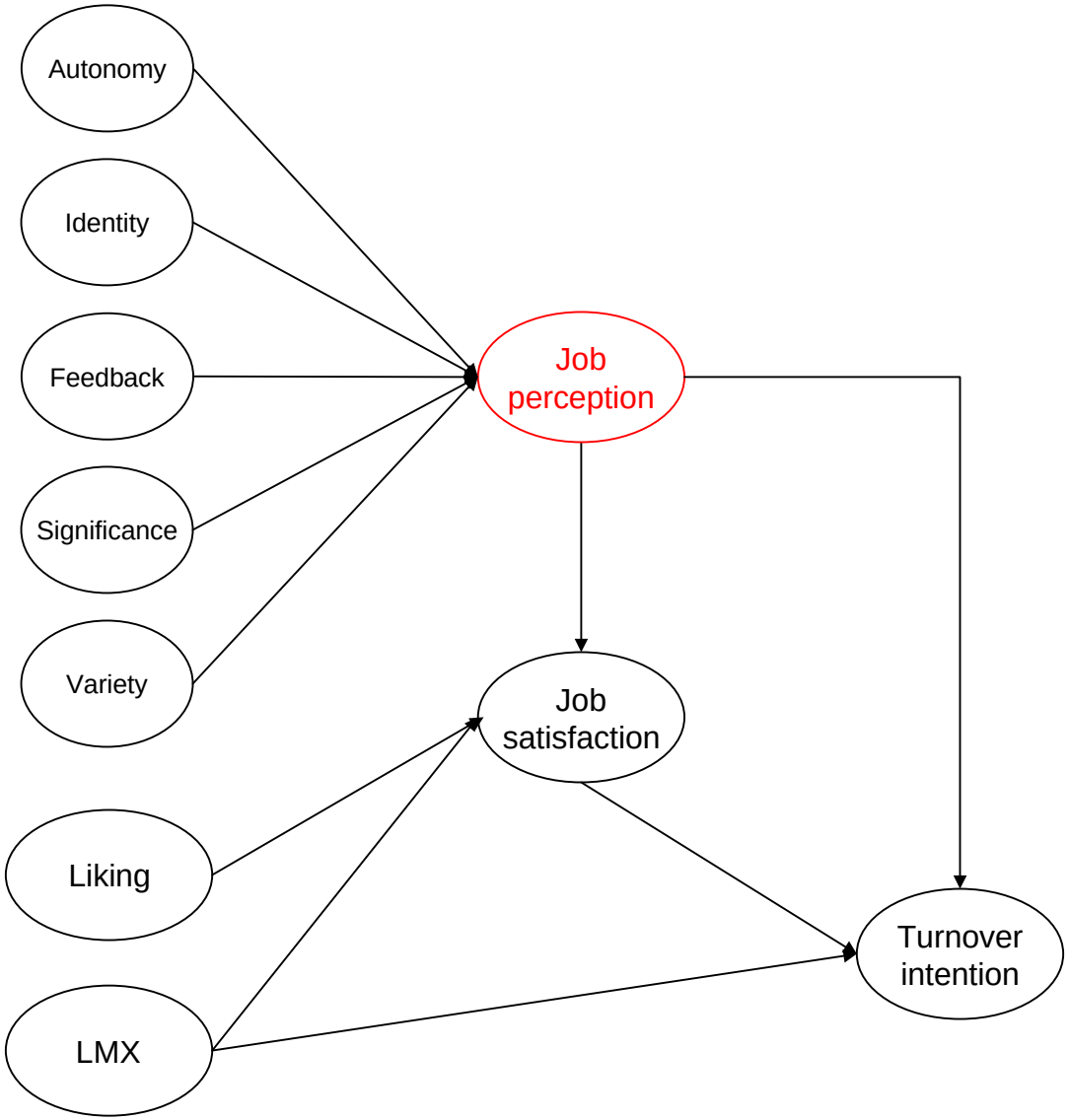
<u>But:</u> Model by *Law/Wong* (1999) is identified!

Assessing identification of models with structural relations between directly influenced latent constructs

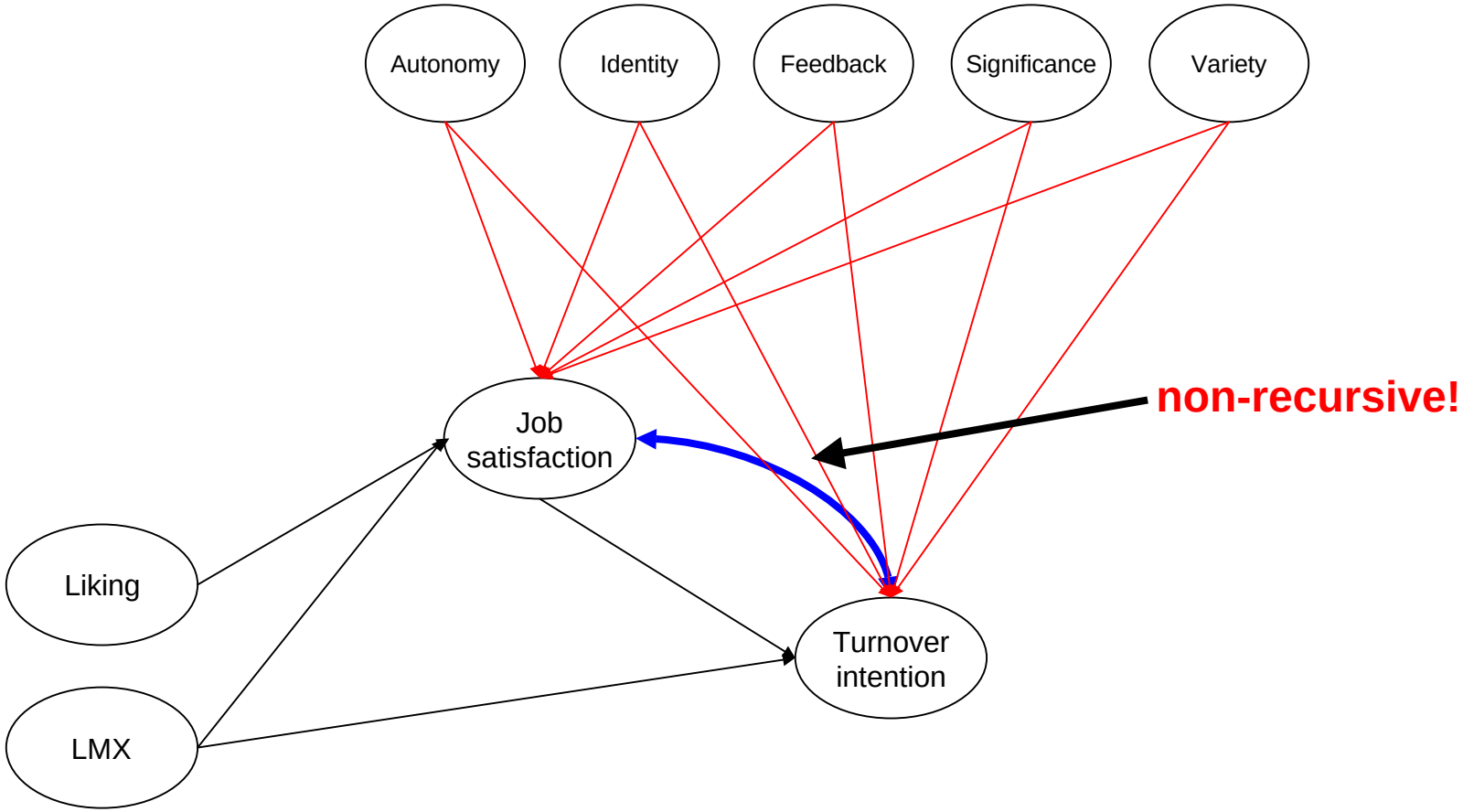
Three-step graphically-oriented approach (see also *Cantaluppi* 2002):

1. Transform the original graph into a graph without CLV.
2. Check identification of the transformed graph.
3. Check that the parameters of the original model can be unambigously derived from the parameters of the transformed model.

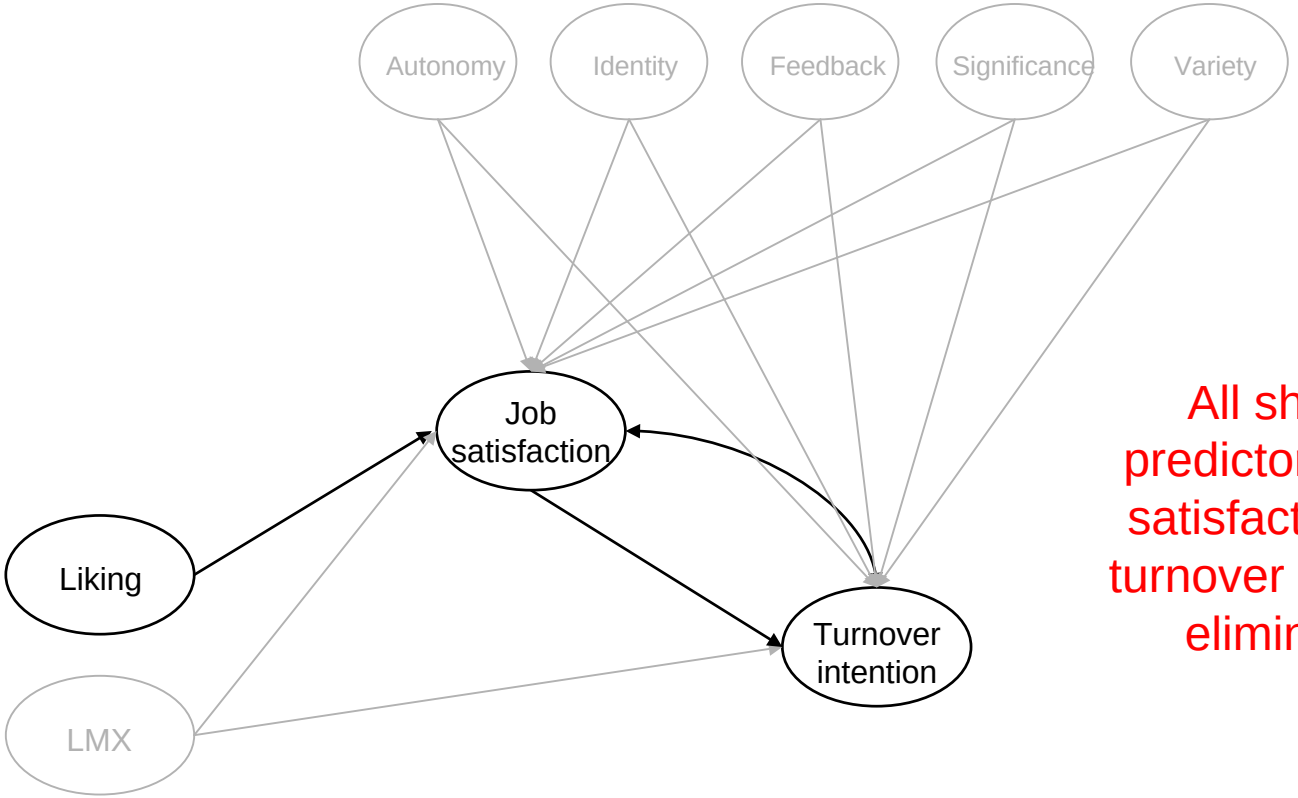
Step 1: START (original model)



Step 1: END (transformed model)



Step 2: Graphic proof of identification (*Rigdon* 1995) → Transformed model is identified



All shared predictors of job satisfaction and turnover intention eliminated

Step 3: Algebraic proof of identification → Original model is identified

$$\eta_2 = 1 \cdot \beta_{21} x_1 + \gamma_{12}\beta_{21} x_2 + \gamma_{13}\beta_{21} x_3 + \gamma_{14}\beta_{21} x_4 + \gamma_{15}\beta_{21} x_5 + \gamma_{26} x_6 + \gamma_{27} x_7$$

$$+ (\beta_{21}\zeta_1 + \zeta_2) \rightarrow$$ All coefficients are identified

$$\eta_3 = 1 \cdot \beta_{31} x_1 + \gamma_{12}\beta_{31} x_2 + \gamma_{13}\beta_{31} x_3 + \gamma_{14}\beta_{31} x_4 + \gamma_{15}\beta_{31} x_5 + \gamma_{16}\beta_{31} x_6 + \gamma_{37} x_7$$

$$+ \beta_{32}\eta_2 + (\beta_{31}\zeta_1 + \zeta_3) \rightarrow$$ All coefficients are identified

$$\left.\begin{aligned}
\psi^*_{23} &= \beta_{21}\beta_{31}\psi_{11} \rightarrow \psi_{11} = \psi^*_{23} / \beta_{21}\beta_{31} \\
\psi^*_{22} &= \beta^2_{21}\psi_{11} + \psi_{22} \rightarrow \psi_{22} = \psi^*_{22} - \beta^2_{21}\psi_{11} \\
\psi^*_{33} &= \beta^2_{31}\psi_{11} + \psi_{33} \rightarrow \psi_{33} = \psi^*_{33} - \beta^2_{31}\psi_{11}
\end{aligned}\right\}$$ All error variances are identified

Outline

1. Motivation
2. Identification of recursive SEM including formative constructs – A new approach
3. Specification of SEM including endogenous formative constructs
   - 3.1 Exogenous vs. endogenous indicators for endogenous CLV
   - 3.2 Empirical study
4. Conclusions

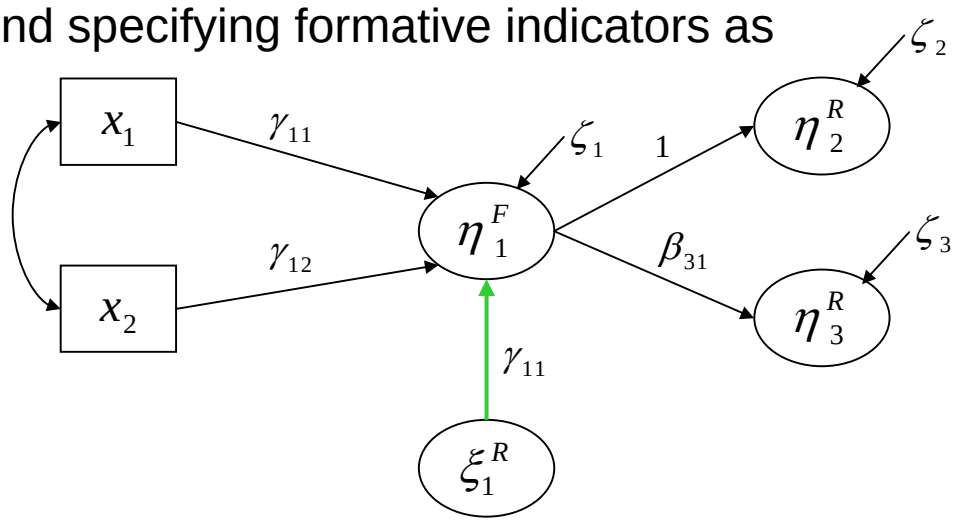
Types of formative CLV and their specification

1. Quasi exogenous CLV: Only influenced by their formative indicators.

   → Covariances between formative indicators and other exogenous variables are set free.

2. Endogenous CLV: Influenced also by other variables (e. g., exogenous latent variables representing reflective constructs).

   → Those few methodological articles discussing endogenous CLV (e. g., *MacCallum/Browne* 1993) recommend specifying formative indicators as exogenous variables.



Issues resulting from specifying *exogenous* indicators for *endogenous* CLV

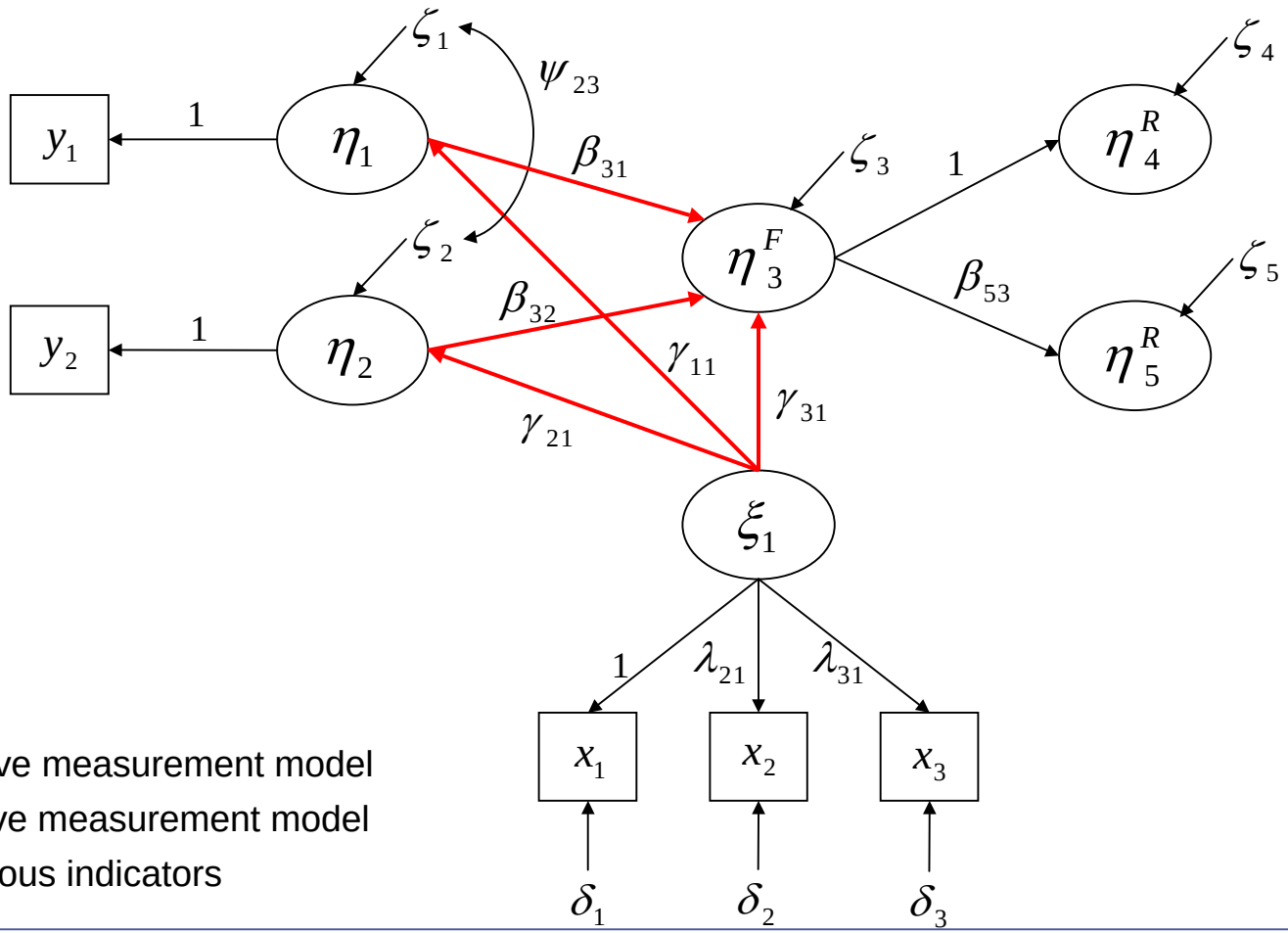
- Conceptual issue: How can explanatory variables influence an endogenous CLV without influencing its formative indicators (e. g., how can a salary increase positively impact on an employee's job satisfaction without improving his/her satisfaction with pay)?

- Substantial issue: Specifying the formative indicators of an endogenous CLV as exogenous variables under- or overestimates the influence of explanatory variables on the formative construct.

→ Formative indicators need to be specified a being endogenous

Specification including endogenous formative indicators (model A1)

Total effect of ξ_1 on $\eta_3 : \gamma_{31} + \gamma_{11}\beta_{31} + \gamma_{21}\beta_{32}$



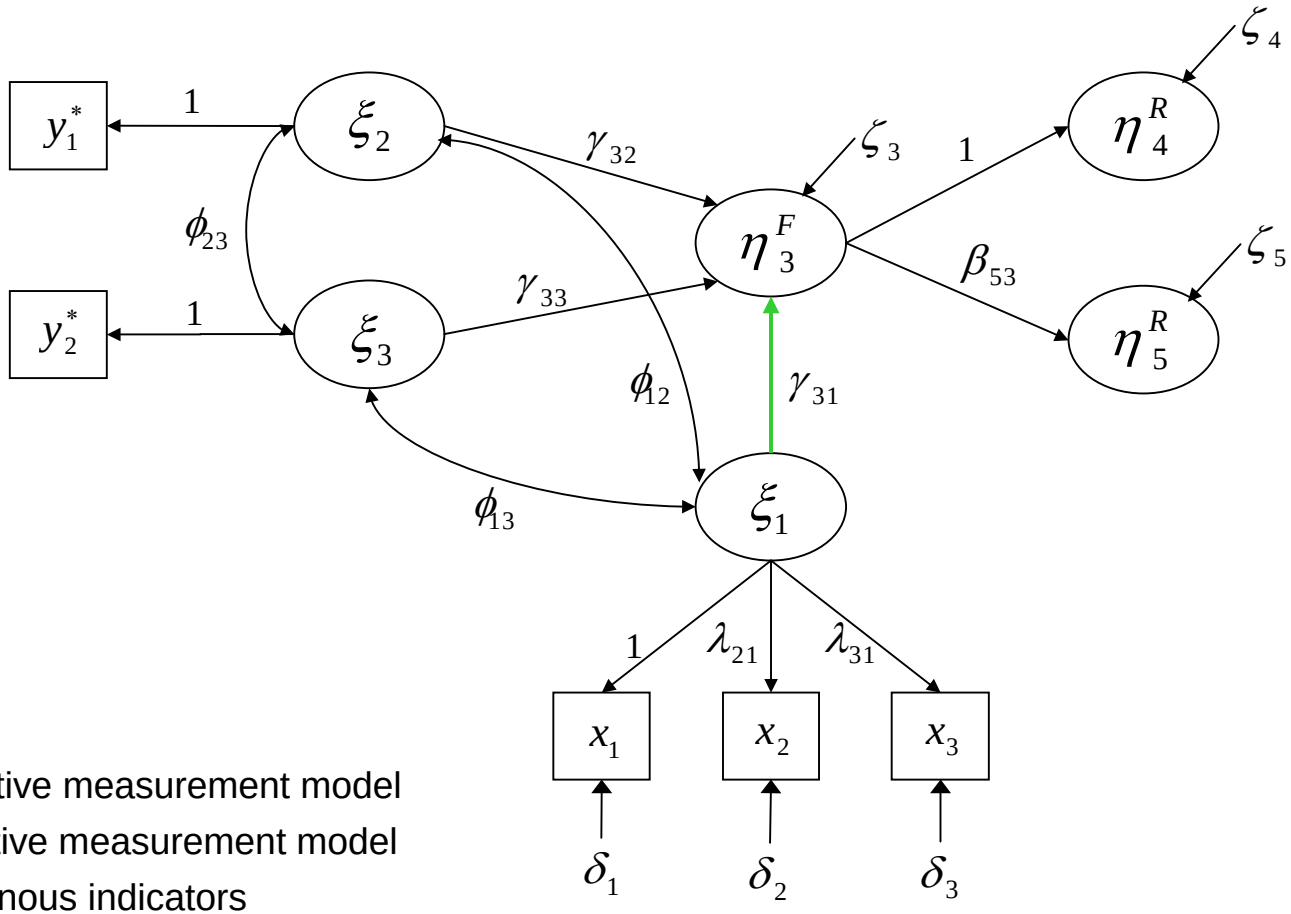
F = formative measurement model

R = reflective measurement model

* = exogenous indicators

Specification including exogenous formative indicators (model A2)

Total effect of ξ_1 on η_3: γ_{31}



F = formative measurement model

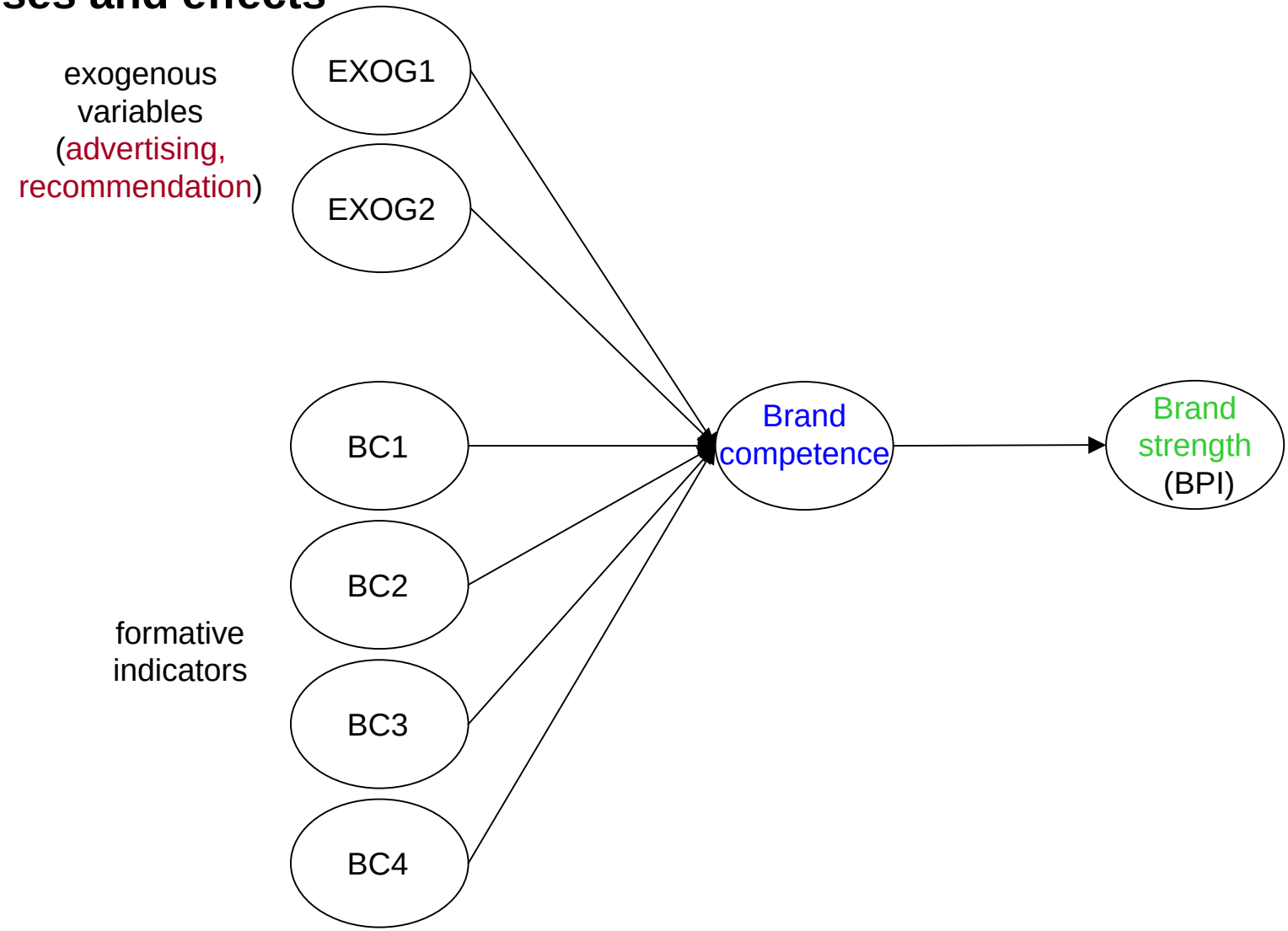
R = reflective measurement model

$*$ = exogenous indicators

Empirical study on the causes and effects of brand competence

- *Brand competence*: "A competent brand is one that has the ability to solve a consumer's problem and to meet his or her need." (*Lau/Lee* 2000)
- Data: Survey (GfK Market Research Nuremberg) on brands for *fixatives* and *denture cleansers*, *N* = 261 consumers in Germany
- Operationalization of *brand competence* by four formative indicators (e. g., "stands for hygiene and cleanliness", "stands for attractiveness and beauty")
- *Causes* of *brand competence*:
  - *Advertsing* ("really advertises a lot")
  - *Recommendation* ("is recommended by dentists")
- *Effect* of *brand competence*: *brand strength (BPI)*

Structural equation model of brand competence – causes and effects



Results – Fit measures

- Model 1: Formative indicators of *brand competence* specified as endogenous variables.
- Model 2: Formative indicators of *brand competence* specified as exogenous variables.
- Model 3: Modification of model 1 (constraining two coefficients to zero).

Model	χ^2	df	p	RMSEA (90 % K. I.)	p	CFI	TLI	RMR^s	R^2 Brand competence	R^2 Brand strength
1	158.05	87	0.000	0.053 (0.038 – 0.067)	0.350	0.987	0.982	0.045	0.61	0.39
2	identical to model 1									
3	158.66	89	0.000	0.052 (0.037 – 0.066)	0.399	0.987	0.983	0.044	0.61	0.39

Direct effects of the exogenous variables

Explanatory variable	Para-meter	Model 1	Model 2	Model 3
Advertising (EXOG1)	γ_{11}	0.077 *0.102*	–	0.076 *0.101*
	γ_{21}	0.133*** *0.160*	–	0.131*** *0.158*
	γ_{31}	0.081 *0.099*	–	0.080 *0.098*
	γ_{41}	0.005 *0.005*	–	–
	γ_{51}	0.155*** *0.261*	0.155*** *0.261*	0.148*** *0.250*
Recommendation (EXOG2)	γ_{12}	0.097* *0.120*	–	0.097* *0.121*
	γ_{22}	0.194*** *0.220*	–	0.194*** *0.220*
	γ_{32}	0.088 *0.102*	–	0.089 *0.102*
	γ_{42}	0.160** *0.161*	–	0.162*** *0.163*
	γ_{52}	–0.030 *–0.048*	–0.030 *–0.048*	–

direct effects on formative indicators

direct effect on brand competence

direct effects on formative indicators

direct effect on brand competence

*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.10$; standardised parameter estimates in italic.

Total and indirect effects

	Model 1	Model 2	Model 3
Indirect effects			
advertising → brand competence	0.056** *0.094*	–	0.054** *0.092*
recommendation → brand competence	0.073** *0.116*	–	0.072** *0.114*
Total effects			
advertising →			
brand competence	0.211*** *0.356*	0.155*** *0.261*	0.203*** *0.341*
brand strength	0.193*** *0.222*	0.142*** *0.163*	0.186*** *0.214*
recommendation →			
brand competence	0.043 *0.068*	–0.030 *–0.048*	0.072** *0.114*
brand strength	0.039 *0.043*	–0.028 *–0.030*	0.066** *0.071*

*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.10$; standardised parameter estimates in italic.

Outline

1. Motivation
2. Identification of recursive SEM including formative constructs – A new approach
3. Specification of SEM including endogenous formative constructs
   3.1 Exogenous vs. endogenous indicators for endogenous CLV
   3.2 Empirical study
4. Conclusions

Conclusions

- Many theoretical constructs in marketing or management research may be specified as being formative (instead of reflective).
- “LISREL” approach often mistaken for not being able to handle formative constructs → “LISREL” can handle formative constructs given model identification.
- PLS “solves” the identification problem but suffers from various shortcomings compared to the “LISREL” approach.
- Identification rules for formative measurement models in “LISREL” have been incomplete: New types of models formerly considered as nonidentified can now be identified.
- For endogenous formative constructs, indicators (or components) should be specified as being endogenous → using exogenous indicators under- or overestimates the effect of explanatory variables on the formative construct; in PLS, only (biased) total effects of the explanatory variables can be estimated.