KATHOLIEKE UNIVERSITEIT
LEUVEN

Some *methodological* challenges of *cross-national* social research with *hierarchical* survey data: Reflections on conceptual and measurement validity

Jaak Billiet

Emeritus professor KU Leuven

(CeSO/ISPO)

Higher School of Economics,
Moscow, Russia,
April 11-14, 2017

jaak.billiet@kuleuven.be

Introduction: explaining the title and aim

Methodological problems:

- conceptual (*focus in this presentation)*
- measurement

Quantitative comparative research =

- 4 main types of designs X subtypes within
- cross-country *(focus)*

Hierarchical data

- micro level and higher level data

(*with survey data at micro level)* *(focus)*

Aim

- Considerations for improving
  - **conceptual validity** = conceptual *(focus)*

    operationalisation step 1: *theoretical concepts, costrucs design, and assumptions*
  - **measurement validity** = technical

    operationalisation step 2 ,and measurement

- **Main idea behind:**

  Conceptual validity is…
  - often *underappreciated* in empirical studies…
  - is related to low explanatory power in hierarchical models and…
  - invalid inferences *(related to measurement issues)*

Outline

1. Concepts, classifications, designs: ***short overview of designs***
2. A typical example *(ESS 2008)* as ***steppingstone*** for…
3. **Validity considerations** related to the '*steppingstone' example*
4. **Main conclusion**

Ideas & examples developed in:

Billiet, J. (2013). "**Quantitative methods with survey data in comparative research**", pp. 264-300 in: P. Kenneth (Ed.) *A Handbook of Comparative Social Policy*. (2nd Edition). Cheltenham: Edward Elgare.

Billiet, J., Meuleman, B. (2014). "**Some methodological challenges of quantitative cross-national social policy research**", pp 289-303 in: van Oorschot W., Peeters H., Boos K. (Eds.)*, Invisible social security revisited: essays in honour of Jos Berghman*, Chapt. 16. Leuven: Lannoo Campus

Billiet, J. Meuleman, B. & Davidov, E. (2015). *"**Some methodological challenges of cross-national social research: conceptual and measurement validity**", pp. 99-129 i*n: Pawel B. Sztabinski, Henryk, Domanski & Franciszek Sztabinski (Eds.) (pp. 99-129). *Hopes and Anxieties in Europe: Six waved of the European Social survey*. Peter Lang: Frankfurt am Main.

Billiet, J. (2015). "**What does measurement mean in a survey context?**" Chapter 14 in Sage *Handbook of Survey Research.*

1.
Concepts, classifications, designs

Classification of research designs *(Dogan & Rokkan, 1969 modified)*

1. Six criteria for cross-classification of designs:
   - **level** of **data:** *individual (micro) – territorial (macro; context)*
   - **Origin** of the measures: *direct (primary) indirect (derived)*

     origin => level where variables are **measured**
   - **Purpose** of analysis: *propositional – descriptive*
   - **Level** of **dependent variable**: *lower – higher*
   - **Nature** of higher level unit: *1 package variable – set of variables*
   - **Focus** of analysis: *one level – more than one*

Classification of research designs…

most relevant combinations…

1. **level** of units by **origin** of the data (Dogan & Rokkan, 1969:4-5; Kendall & Lazarsfeld, 1950: 195-196)

   goes back to *American Soldier*…

   level: individual / territorial (context)…

   origin: derived / direct (primary)

*M*ost common practice these days…

context data measured by aggregation of micro information (surveys) ***indirect*** *measures, and official statistics (also derived measures???) i.c.* ***upward*** *measurement of means, percentages… aggregated = transferred to context level); (possibly SE estimation problems (bias) sometimes* ***downward*** *measurement (danger: ecological Phalacy*

Classification of esearch designs…

2. **purpose** of analysis by **nature** of society (or *context*)

(Scheuch, 1986: 176-200)

purpose: propositional / descriptive

= *aim of the research project*

nature: country

- as historical unit = as *package variable*

or

- as *sets of variables*

*as **Package variable**: no specification of specific effects;*
common design in past before multi-level
(countries as dummy's with reference country 0)

Classification of research design

3. **level** of dependent (*outcome*) variable by **focus of analysis** (Dogan & Rokkan, 1969: 6-9)

Level of dependent variable
= individual / territorial (*higher level*)
focus = number of levels: one or more levels

proposed cross-classification on these dimensions:
(*next figure "level of measurement within each cell)*
= useful classification

Table 1 Classification of research designs by level of dependent variable, focus of analysis, and, within each cell, origin of the data

Level of dependent variable	Focus of analysis	
	One level	*Two or more levels*
Individual level	Type I *Either*: Individual-level data (e.g. from surveys) treated without reference to the territorial contexts. *Or*: territorial aggregate data used to analyse variation between individual.	Type III *Either*: individual-level data used jointly with contextual data of territorial units. *Or*: aggregate/global data used to test interaction between levels.
Higher level	Type II Aggregate/global data for territorial units used to describe and account for variations at territorial level.	Type IV *Either*: joint use of individual-level/aggregate/global data to test *sources of change* in territorial structure. *Or*: aggregate/global data used to test *interaction* between levels.

Source: Dogan and Rokkan, 1969: 6 (modified).

Usefulness of the design characteristics

Designs **rely** on …

- **Research question**
- **Theoretical expectations**
- **Available data**
- **Assumptions about measures and relations**

Crucial to make background assumptions **explicit** in case of designs with two or more levels (designs **III** & **IV)**

- in particular design **IV**

when focus is on **higher level outcomes**

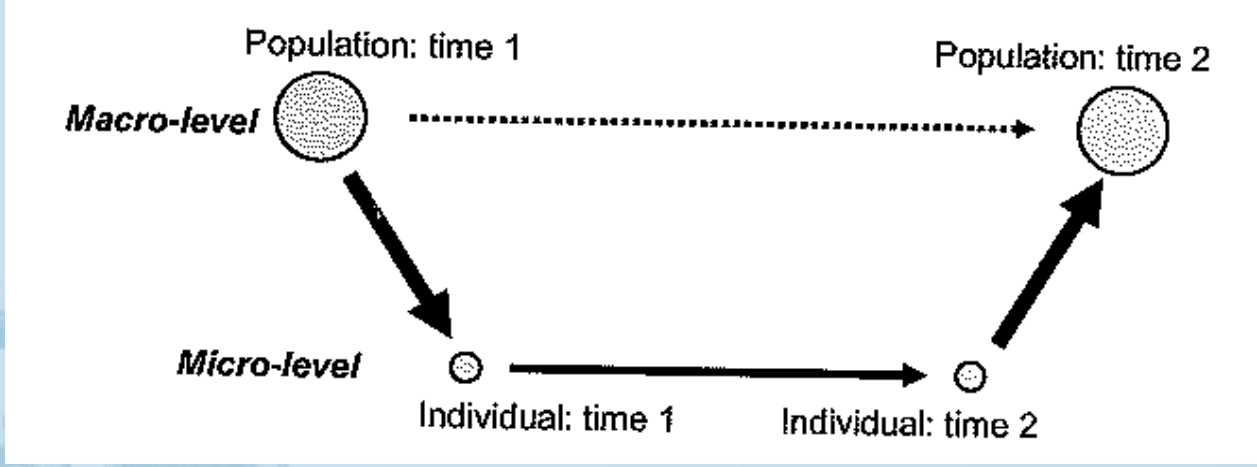
- based on theoretical background of

"methodological individualism" *(Coleman, 1990)*

Relevance of theoretical assumptions behind designs

Basic idea of methodological individualism :
"society is constituted by the interaction between individuals but cannot be reduced to it"

This view is formalized in the "*bathube diagram*" *("Coleman diagram")*



Bathtube model *(Oaks, 2009:4)*

This diagram clarifies the **bridge assumptions** and **transformations** rules behind the micro-macro link in ML research

Relevance of theoretical assumptions behind designs

Is micro foundation of macro relations always necessary?

It depends of the theoretical meaning of these relations between the variables *(constructs):*

Is required when the macro variables refer to forms of social action, social (or collective) choice, social movements

- in particular when derived measures of macro variables are **aggregate** (indirect) measures

Relevance of the design characteristics

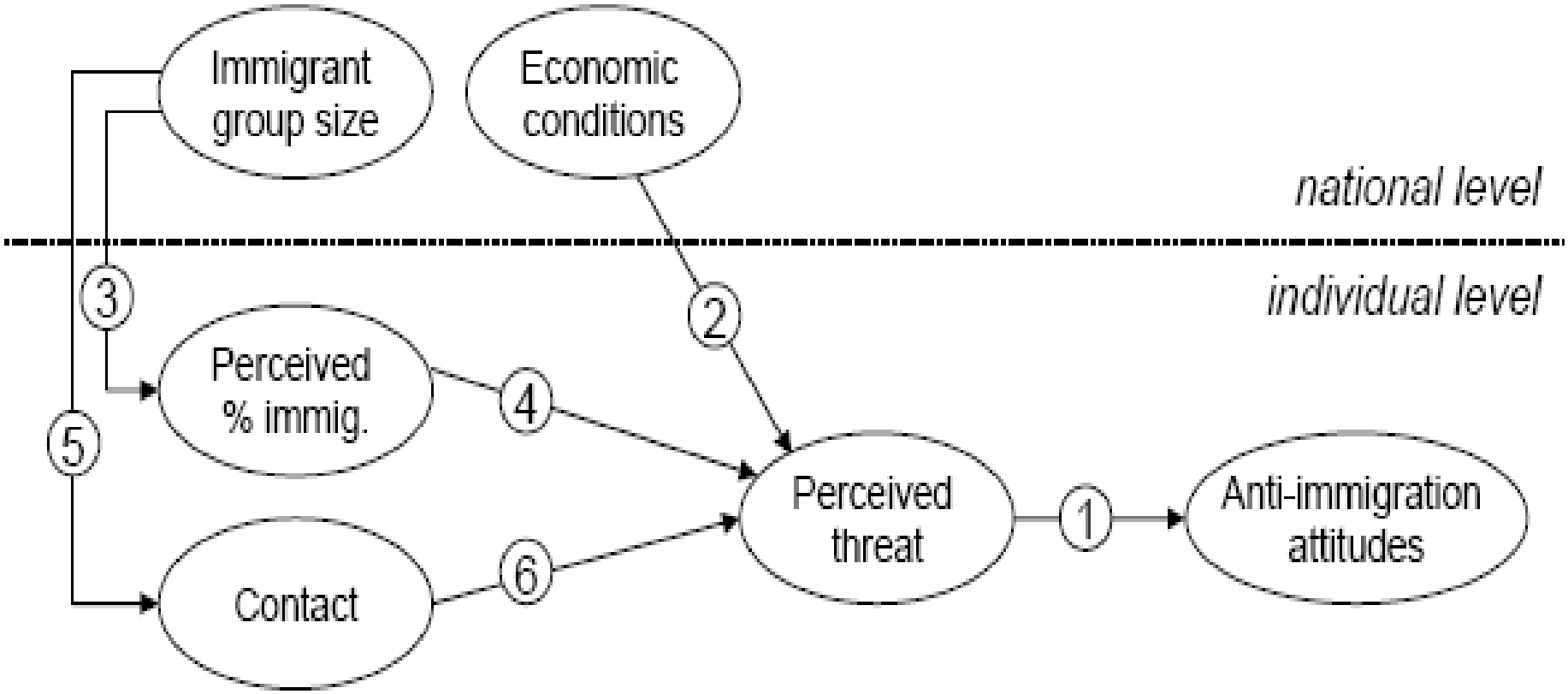
Designs have **consequences** for

- validity of **operations** *(analysis: statistical & conceptual)*
- validity of **inferences**: *statistical (BIAS in SE estimates) & theoretical misspecifications of relations)*
- validity of **conclusions** *(inferences made: statistical and theoretical)*

hint:
- always start to identify the design!!!
- show the concepts & relations in a graph
- discuss with team members in early stage

Example of type **III** design (with indirect effects)

Multilevel regression with direct and indirect effects on anti-immigration attitudes, and with effects of context (country level) variables on perceived ethnic threat at the individual level (see Meuleman, 2011



Example of type **III** design (with cross-level interaction)

Design IIIA + cross-level interactions at one point in time. Interaction effect of context factors on the relation between individual level economic condition and perceived threat

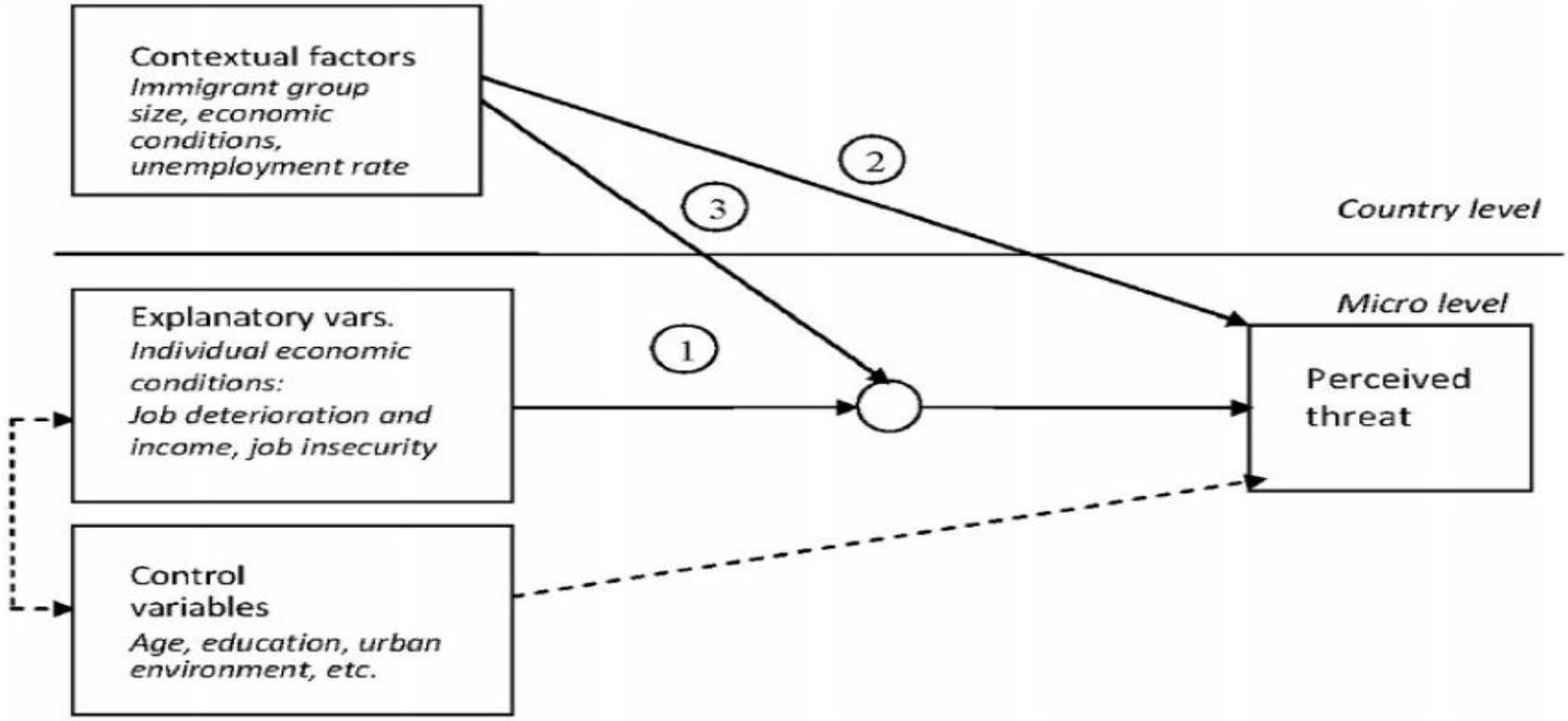


Figure 1. Expected relationships between factors at the micro level and expected between level interactions.

A note on change over **time**…

Additional implications when *change over time* is considered

Three possibilities

- **weak design:** one point in time with *retrospective* measures
- **strong design**: **panel data** at both individual level and context level (repeated measurements) = *very scarse*
- **in between:** repeated **cross-sections at individual** level + **panel data at context** (country) level

this design often used*(ESS, EVS…) (Meuleman et all, in press)*

Analysis = societal ***grow curves models:*** longitudinal effects (time variant components) + cross-sectional differences

Think about alternatives for multilevel regression:

MGSEM, Latent variables at Group level, combinations

2

A typical example

*as **steppingstone** for considerations on conceptual and measurement validity*

A typical example as *steppingstone*

- **Designs III & IV** widely applied these days
- Possible because of enormous **growth in cross-country** and **over time micro data** (ESS, EVS, Share, ISSP…) -> **survey data**
- **Cross-sectional: mostly** not real longitudinal at micro level *(panel datasets at micro level very scarce)*
- Limitation = **no strict causal** inferences possible

  *(*only say: *"causal effect is likely", "is not rejected")*

  - even with retrospective questions at micro level
  - **valid counterfactual argumentation** is vital

    is always complement of inferences based on statistical analysis *(see further)*

Example: Individual **V**alues, **C**ultural **E**mbeddedness & **A**nti-immigrant sentiments *(IVCEA)* **(Davidov, Meuleman, Schwartz, Schmidt, 2014)**

- A design **(III)** based on published multi-level cross-national analysis in 24 European *countries (ESS Round 4, 2008-9 + Survey of Schwartz, 2000-2006)*

- **Micro** variables *(direct micro-level measures: ESS)*
  Macro level (*context*) variables:
  - ***direct*** *measures based on official statistics (2008)*
  - ***derived*** *measures (aggregate) micro measures*
    in this example = from ***different survey*** *(# ESS 2008)*

- Study based on **theory** and **theoretical expectations** *(Schwartz, 2006)*

- **Steppingstone** to illustrate previous concepts and reflections

Research questions in and hypotheses *IVCEA*
(Davidov et al., 2014)

(1) What** value orientations affect attitudes towards immigration? (indicator of outcome variable = **allowing immigrants)

Value orientations: *(2 x 2 combined Schwartz values)*

H1.1. Universalism: express motivation to appreciate differences btw individuals, tolerance, care for weak and vulnerable *(Sagiv & Schwartz, 1995)*
expect: **positive effect** on allowing of immigrants

H1.2. Conformity/tradition: expect motivation to maintain the beliefs, customs of own culture, conventional expectations and norms *(Davidov, 2010)*
expect: ***negative effect*** on allowing of immigrants

Research questions and hypotheses in *IVCEA*
(Davidov et al., 2014)

(2) *What is effect of country characteristics on acceptance of immigrants?*

In line of Group Conflict *Theory:* ***relative size of immigrant population*** (%) affects feeling threatened by immigrants *(Billiet, Meuleman & de Witte, 2014)*

note: feeling threatened is = ***concept****; allowing immigrants =* ***indicator***

H2. negative effect of **size (%) of immigrant population** on allowing immigrants

(*note: somewhat reframed)*

Research questions and hypotheses in IVCEA
(Davidov et al., 2014)

(3) *What is effect of cultural (!) a context level variable "embeddedness" on feeling on the relation between value orientations and feeling threatened by immigrants? (cross-level interactions)*

Concept: *embeddedness (= cultural 'climate' of society):* characteristic of societies where people are viewed as embedded in the collectivity, maintain group solidarity, identify with norms of the group.

Research questions and hypotheses in *IVCEA* *(Davidov et al., 2014)*

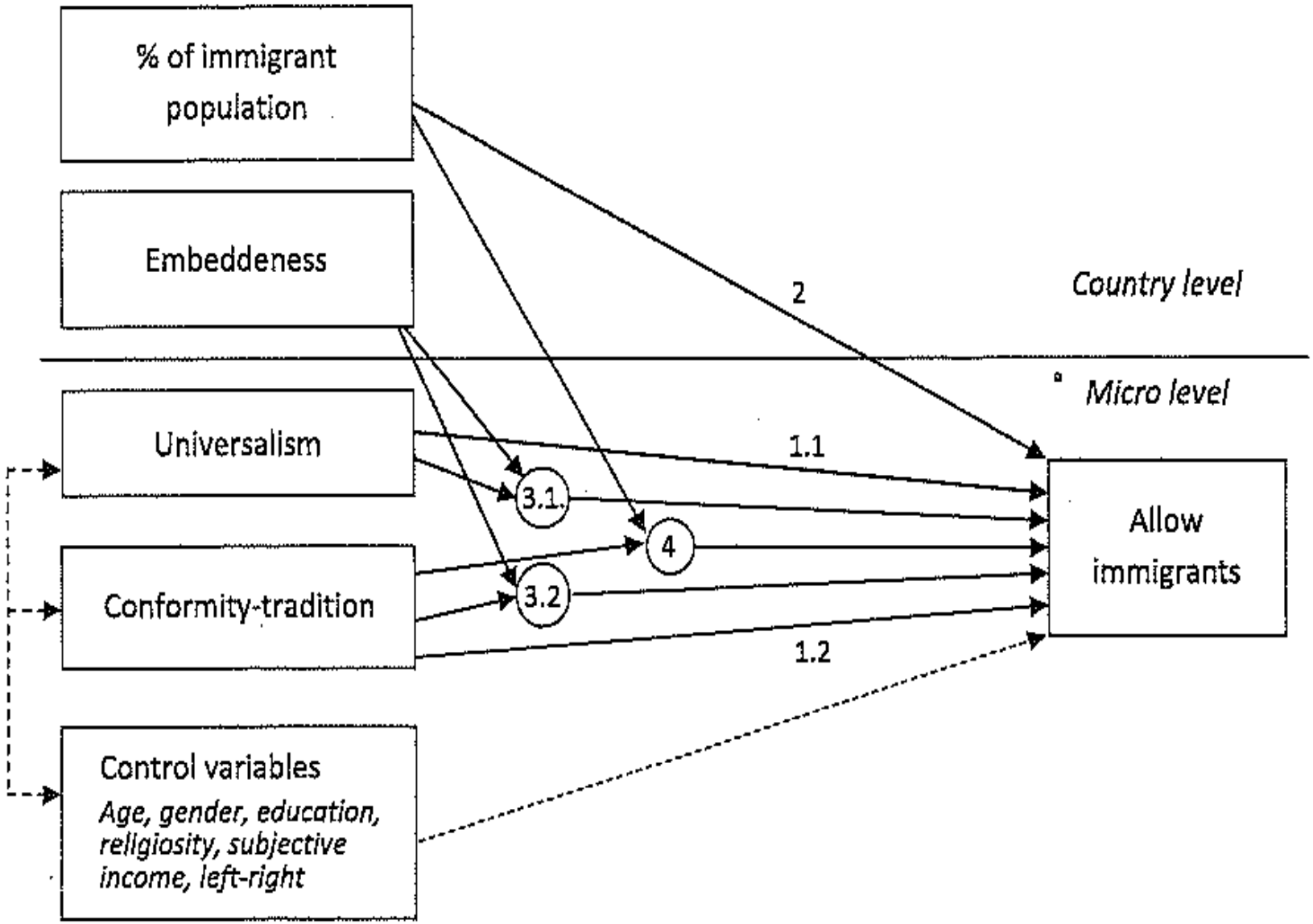
H3.1. the effect of universalism on allowing immigrants is **weaker** in cultures with high level of embeddedness.

H3.2. the effect of conformity-tradition on allowing immigrants is **weaker** in cultures with high level of embeddedness.
***Rationale* behind:** *more people follow their individual preferences in less embedded societies*

(4) How is the relation between value orientations affected by relative size of immigrant population at country level? *(cross-level interactions)*

H4. the negative effect of conformity-tradition is **stronger** in countries with higher relative size of immigrant population

Figure 5.1: Expected relations and cross-level interactions in the theoretical model of the IVCEA study (Davidov et al. 2014)



3.

Validity considerations

related to '*steppingstone' example*

Validity considerations: an overview

About…

1. Mixed measures and derived measures at context level…
2. Large theoretical distance
3. Weak theory at context level *(Hypotheses often based on available data)* *(skip)*
4. Small N at context level *(skip)*
5. Blind spots in explanatory models… *(skip)*
6. Problematic causal inferences *(skip)*

For complete overview see Billiet (2013) *(copy on request)*

Validity consideration # 1

Mixed measures at **context level** in *IVCEA* study

- global **direct** measures (*e.g. relative GDP size*)
- **derived** measures from other survey in *(nearly)* same period

- different population (*school-teachers as privileged **proxy** sample units for cultural climate?*)
- 2006 # 2008: argument pro validity: *embeddedness is stable over time* (Schwartz, 2006)

Rule:

whenever possible use ***direct*** measures for concepts at higher level

???…*to discuss…*

Validity considerations…

However ...***what to do when direct measures at higher level are not available?***

Acceptable to use derived measures from lower level?

= *upward measurement?* *(no downward measurement in type III)*

- derived by aggregation (*means, %,…*) from other surveys *(in this example of IVCEA)*
- derived by aggregation (*means, %,…*) from individual level variables of same survey

=> **double use** of same measures

conditional acceptance?

Conditions for **double use** of measures of constructs with derived upward measures

1. Constructs at *higher* (context) level have a clear and **different theoretical meaning** than constructs at *micro* level
2. The specific meaning (**conceptual validity**) at both levels is argued (made explicit)
3. Operations with indicators and constructs are **statistically** valid at both levels
4. The cross-level **transformation rules** of measures are adequate *(see "Coleman diagram")*
5. The **bridge assumptions** are specified and made plausible (see "Coleman diagram")

Note: **NO double use** of same measures in *IVCEA* example *(why?)*

Considerations when using indirect **upward** measurement

1. *Does sample distribution **reflect** the population distribution?*
2. *What are implications of **measurement error** & non-response bias in samples? (is bias transferred to context level by indirect measure; aggregate)*
3. *Are conclusions based on aggregate statistics of outcome variables biased? (i.e. means) (Croon & van van Veldhoven, 2007).*
4. *Are the measures **equivalent** btw country samples? (in-equivalence transferred to context level by indirect measure)*
5. *Are upward measured variable needed for statistical/theoretical reasons (within country deviations from mean vs between country deviation from grand mean)*

acceptability depends on these conditions questions and answers to considerations

Illustration by *IVCEA* example

Example of **valid (?) bridge assumptions** of the subjective micro level measures (opinions, beliefs, values) used to measure macro variable *"embeddeness" as a characteristic of the cultural climate (context)*

Bridge assumptions

- individual opinions seen as **collective** representations
- *individuals acquire knowledge through collective debates*
- *cultural context shapes **frames of reference** to which perceptions are related*
- *aggregate data = indication of **dominant ideas** that serve as frames*
- specific for *IVCEA*: *schoolteachers opinions are valid indicators proxy indicators for "embeddeness" (?)*

Validity consideration # 2

Theoretical distance between country level predictors and micro level dependent variable **too large** when direct measures for country characteristics are used (e.g. *GDP*)

Not clear for cultural variables: depends on theoretical meaning at country level

Hierarchical analysis (Coleman, 1966) applied in field of education = **success story** in *sociology of education*

(despite numerous problems!)

but… **often problematic in cross-nation** research

WHY*?... find the contrasts between [countries+citizens] and [schools+studens] from viewpoint of ML analysis*

4

Conclusions

"historically derived cultural diversity makes Europe a **natural laboratory** for the social sciences"*
(*Roger Jowell*[+])

at condition that researchers…

Conclusion

Technical **statistical** validity is **not sufficient** condition for conceptual (theoretical) validity

Invest more in **serious reflection** on operationalisation of *concepts*, *design, bridge assumptions,* and *transformation rules, in the operationalisation stage* **before** designing (or applying) statistical models

This will contribute to both *theoretical (conceptual)* and *measurement* validity

Ask always questions about concepts and assumptions to you desigs and made these **explicit**

THANKS