Open KB Seminar 29/4/2024

Latent variables, measurement models and bibliometric data

Peter van den Besselaar

Vrije Universiteit Amsterdam & DZHW Berlin

Latent and manifest variables

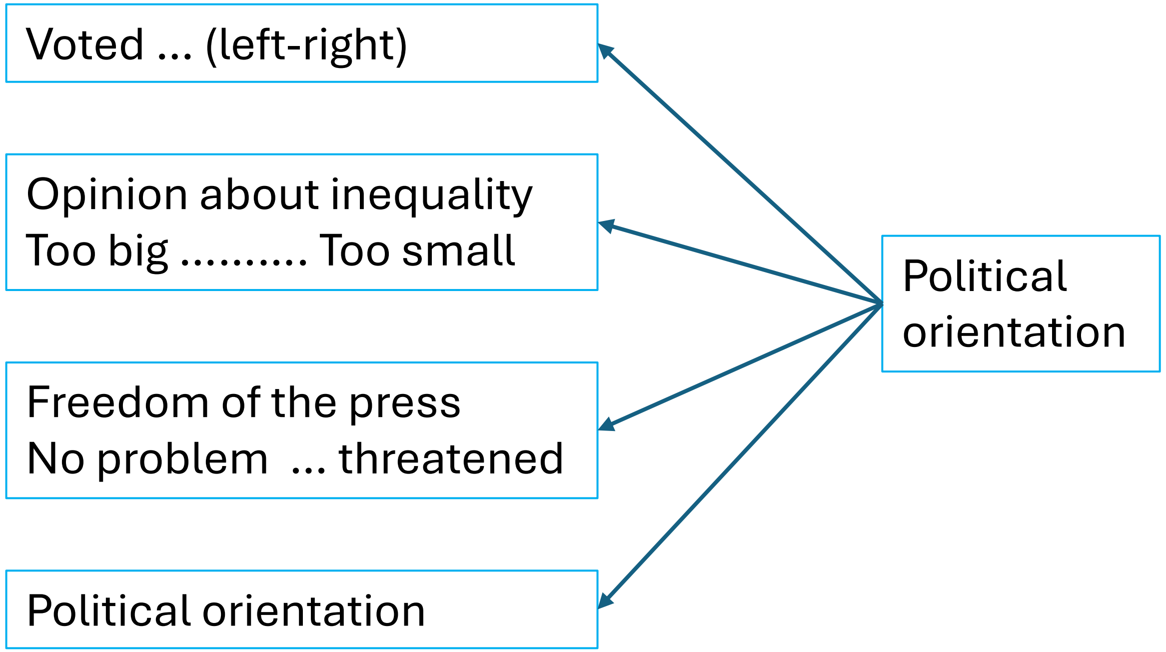
- Does political orientation influence opinion about international science cooperation?

- Measuring political orientation:
  - Cannot be observed directly => latent variable
  - Need for observables related to political orientation => manifest variables
  - E.g., through a survey

- Asking “what did you vote in the previous election” is not enough.
  - Floating voters, faulty memory, voting is only one aspect of political orientation

Measuring latent variables

- Researchers use a set of items (survey question) related to political orientation to cover the concept correctly (validity)
  - On inequality, taxation, abortion, freedom of the press, gender issues and so on
- These are observable through a survey: manifest variables
- Latent variable ‘political orientation’ causes the scores on the observable items



The score on the latent variable

- Is the latent variable one dimension? => factor analysis
  - In case of political orientation, one finds generally two dimensions
  - Socio-economic: left-right
  - Cultural: liberal-conservative

- Is it still one scale or are it two scales: scale analysis
  - Cronbach's Alpha

- Score:
  - Sum of scores on the items
  - Regression on the factor loadings
  - SEM: estimating simultaneously the structural model and the measurement model

Bibliometric indicators

- Bibliometric indicators are implicitly used as direct measures of latent variables

- Nr of publications as quality but later abandoned
- Nr of citations as quality but later changed into impact
- Derivates from these as ????

- But in many studies, these indicators remain used as proxies for performance and - implicitly - for quality

The empirical problem

- Explaining e.g., gender bias in grant success needs merit or performance indicators
  - (Bias is a deviation from merit)

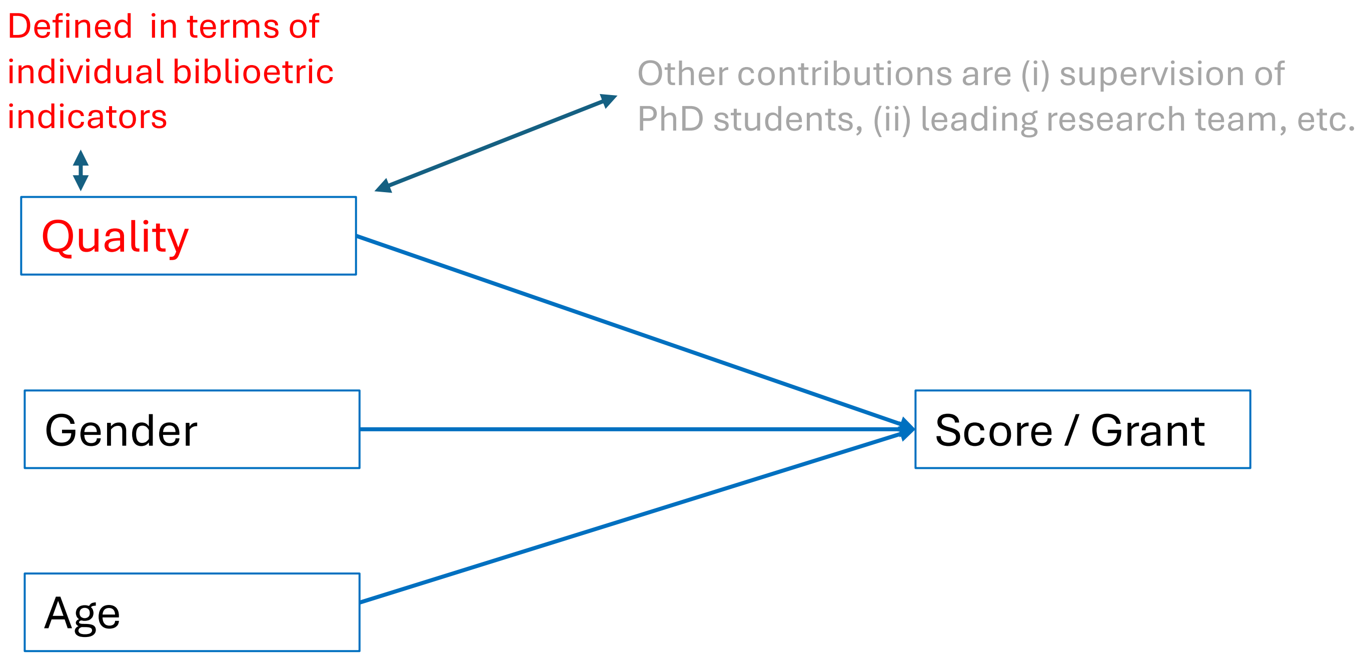
- For some panels, P10 is a significant predictor, for others P5, or Frac P
  - This is not a behavioral issue, as these are not known by the panelists.
  - Using all of course does not work (=> Multicollinearity)
  - And defining merit in three ways without a good reason is neither a solution

Model

Operationalization of contributions to science:

- Every bibliometric operationalization works only in some but not in other panels.
- Ad hoc choice to support the hypothesis in the different panels.

Unclear definition (validity problem) and reliability problem.



Learning from the social and behavioral sciences

- Distinguishing between the latent variables and manifest variables
- List of (bibliometric) items that can be assumed as valid operationalization of the relevant concepts (eg scholarly quality)
  - Excluding for the moment other relevant items (and data)
- Using those for a reliable measurement

Some properties of the indicators

- Many indicators
- Indicators correlate (high) but not perfect: all cover a part of the concept
- Holds also for the differences between datasets P_{scopus} vs P_{WoS}

Define and operationalize quality

- Scientific quality of a researcher:
    - The ability to contribute to science, given age and academic age:
    - Higher quality => more contributions

- Items that measure contributions:
    - Number of contributions to science
    - Impact of those contributions
    - Number of high impact contributions
    - Share of top contributions
    - Access with the contributions to high prestige journals
    - Etc.

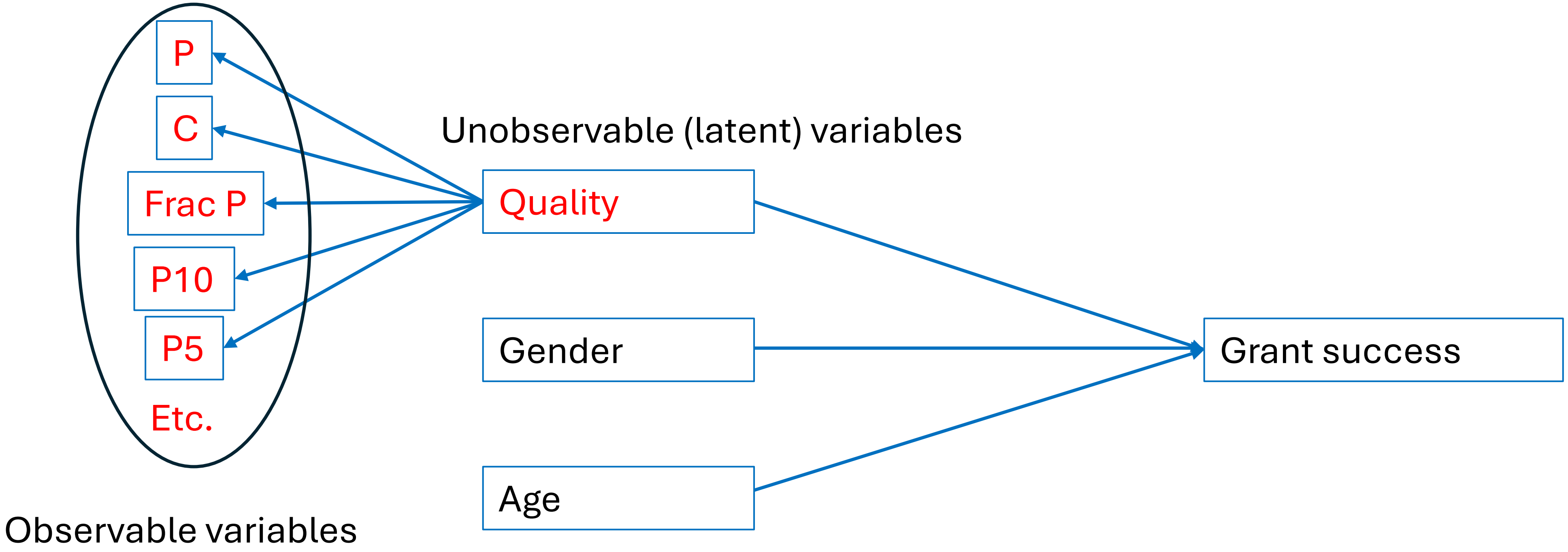
    - Independence (own contributions)
    - Broad versus narrow coverage of research topics

    - Age
    - Academic age

Measurement models: several options

- Average of the (standardized) relevant items
- Factor analysis of the items and reliability analysis
- Structural equations: simultaneous estimates of laten variables and the main model (SEM)

SEM


P
C
Frac P
P10
P5
Etc.
Unobservable (latent) variables
Quality
Gender
Age
Grant success
Observable variables

Example 1: Contribution to science (quality)

- Data
  - All Scival bibliometric indicators
  - Factor analysis -> latent dimensions of quality
  - Reliability analysis of the scales
  - Stability analysis

  - Items to be validated in interviews with committee members by interviews

The Scival indicators

	Abbreviation	Indicator
1	P	Total publications
2	P frac	Total publications, fractional counting
3	C	Citations
4	C frac	Citations, fractional counting
5	C/P	Citations per publication
6	FWCI Sum	Sum field weighted citation impact
7	FWCI Average	Average field weighted citation impact
8	P10%	Number top 10% cited papers
9	P10% FN frac	Number top 10% cited papers, field normalized, fractional counted
10	P10% share	Share top 10% cited papers
11	PP10%	Share top 10% cited papers
12	PP10% FN	Share top 10% cited papers, field normalized,
13	PP10% FN frac	Share top 10% cited papers, field normalized, fractional counted
14	SJR	SJR
15	SNIP	Average SNIP
16	Citescore	Citescore

Three factors

- Total impact
  - size dependent
  - Shows that number and impact of contributions are one dimension
  - C-alpha: 0.914
- Reputation
  - C-alpha: 0.873
- Relative impact
  - size independent
  - C-alpha: 0.853
- N: about 2600

Pattern matrix	Total impact	Journal impact	Relative impact
P*	0.994		
P frac	0.952		0.376
P10%	0.844		
P10% FN	0.837		
FWCI sum	0.815		
P10% FN frac	0.773		
C frac	0.747		
C	0.747		
SJR		0.951	
SNIP average		0.928	
Citescore		0.916	
PP10% FN			-0.839
FWCI average			-0.828
P10% sgare			-0.706
C/P			-0.697

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization
Rotation converged in 9 iterations.
* all items are panel-based z-scores
|loadings| < .30 not shown

Second case

Smaller set of indicators

N = about 2600

C-alphas > 0.80

Table 2: The bibliometric performance indicators.

Pattern matrix	Relative impact (1)	Total impact (2)	Journal impact (3)
PP10% FN	0.980		
PP10% FN frac	0.950		
FWCI average	0.719		
PP10%	0.670		
C/P	0.593		
P frac		0.919	
P10% FN frac		0.801	
C frac		0.709	
SJR			0.942
SNIP average			0.703

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.
All items are panel-based z-scores
|loadings| < .30 not shown

Use case

Van den Besselaar & Mom (2014), Is there gender bias In awarding cum laude for The PhD thesis. *Scientometrics*

Table 9 CL by gender, faculty and quality (PhD period)

	B	SE	Wald	df	Significance	Exp(*B*)	95% CI for Exp(*B*)	
							Lower	Upper
Women vs. men	- 0.858	0.246	12.161	1	< .001	0.424	0.262	0.687
Faculty			29.196	5	< .001			
Faculty(1)	- 0.763	0.272	7.859	1	0.005	0.466	0.274	0.795
Faculty(2)	- 1.733	0.411	17.801	1	< .001	0.177	0.079	0.395
Faculty(3)	- 0.072	0.289	0.061	1	0.805	0.931	0.528	1.641
Faculty(4)	- 2.324	0.771	9.081	1	0.003	0.098	0.022	0.444
Faculty(5)	- 1.08	0.765	1.993	1	0.158	0.34	0.076	1.521
PhD year	- 0.045	0.026	3.036	1	0.081	0.956	0.91	1.006
Relative impact	- 0.096	0.122	0.627	1	0.428	0.908	0.715	1.153
Total impact	0.609	0.084	52.338	1	< .001	1.839	1.559	2.169
Journal impact	0.527	0.106	24.898	1	< .001	1.694	1.377	2.084
Independence	- 0.011	0.004	6.855	1	0.009	0.989	0.981	0.997
Team size	0.03	0.133	0.051	1	0.822	1.03	0.794	1.338
Constant	- 2.226	0.202	121.235	1	< .001	0.108		

Logistic regression; all PhD students in the selected faculties; impact indicators calculated over the t - 3 to t + 3 period

$N = 2253$; Nagelkerke $R^2 = 0.203$

Example 2: Independence

- Independency is important in the mid-career: becoming independent from the early career (PhD) environment

- An independent researcher has developed independence from the early career environment (supervisors)

- Example:
  - ERC defines independence as there should be at least one paper without the supervisor.
  - Why would this be a valid and reliable indicator?
  - Again: independence as a manifest, not a latent variable

Aspects of independence

- An own network, different from the early career environment (1)
- An own oeuvre => papers without the former supervisors (2)
- An own research agenda => topics where the supervisors have not been active (3)

Van den Besselaar & Sandstrom (2019) *PLoS ONE*
Moeller, Van den Besselaar, Mom (2022) *Proceedings STI*

Operationalization

- In ego-network of researcher should the former supervisors not be central:
  - Low eigenvector centrality (EC)
  - High clustering coefficient (CC)
- Researcher should have publications not co-authored with former supervisors
  - Share of own publications among all publications of researcher (SOP)
- Researcher should be active in other fields than the former supervisor
  - Share of own (Scopus) themes (SOT)
- Score could be: Independence = average((1-EC) + CC + SOP + SOT)

Example 3: Cognitive mobility

- Movement between research topics
  - Nr of new topics over the career (1)
  - Distribution of papers over topics (2)
    - New topics can have a low of high number of papers
    - Incidental work on a topic, or substantial
  - Distance between topics: within one or within more discipline (3)

- Cognitive mobility as average of those three scores

Mom, Moeller, Van den Besselaar, *Proceedings ISSI* 2024

Lessons learned

- Using indicators as items leads to more valid and reliable indicators
  - Tested it in a variety of studies, with similar results

- Substantially:
  - Number and impact of contributions are <u>one</u> quality dimension
    - –> versus the quantity-quality discussion
  - Relative impact did not work: Very different groups have high relative impact:
    - Top and bottom performers!

Conclusions

- Many variables we are interested in are latent

- Bibliometric (and other) data can be used as items to measure the latent variables (if well defined)

- We showed a few examples, and applying those suggests that this is a fruitful approach

- But this is a developing research agenda

Thanks for your attention

Questions, Comments?