SIMPLE CORRECTION FOR MEASUREMENT ERRORS WITH STATA

Stata User's Conference, Nürnberg
26th June 2015
Anna DeCastellarnau
ESS-CST, Universitat Pompeu Fabra
anna.decastellarnau@upf.edu

"A simple procedure to correct for measurement errors in survey research"

Written by: Anna DeCastellarnau and Willem Saris

http://essedunet.nsd.uib.no/cms/topics/measurement/

www.europeansocialsurvey.org

	Results without corrections	Results with corrections	
Effects	**Regression coefficients**	**Regression coefficients**	
Dependent V1<-			
V2	0.263**	0.389**	+0.126
V3	0.041*	0.108**	+0.067
V4	0.290**	0.460**	+0.170
V5	0.056**	-0.066**	-0.122
V6	0.085**	0.036*	+0.049
R^2	**0.249 (25%)**	**0.561 (56%)**	

** if α<1% and * if 1%<α<5%

+0.312

- Increase in effects of more than 1 point on average
- Increase in more than factor 3 in the explained variance

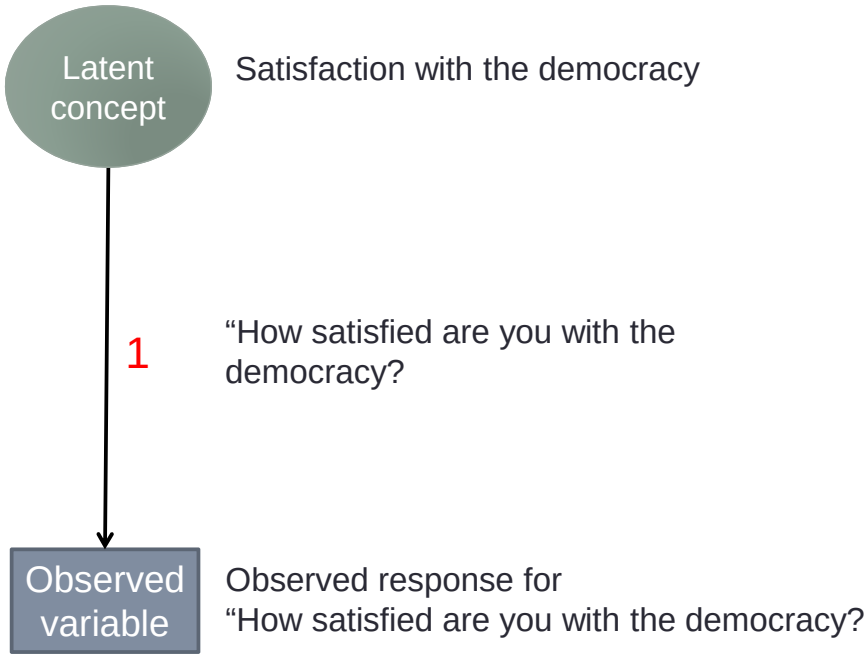
OUTLINE

Theory

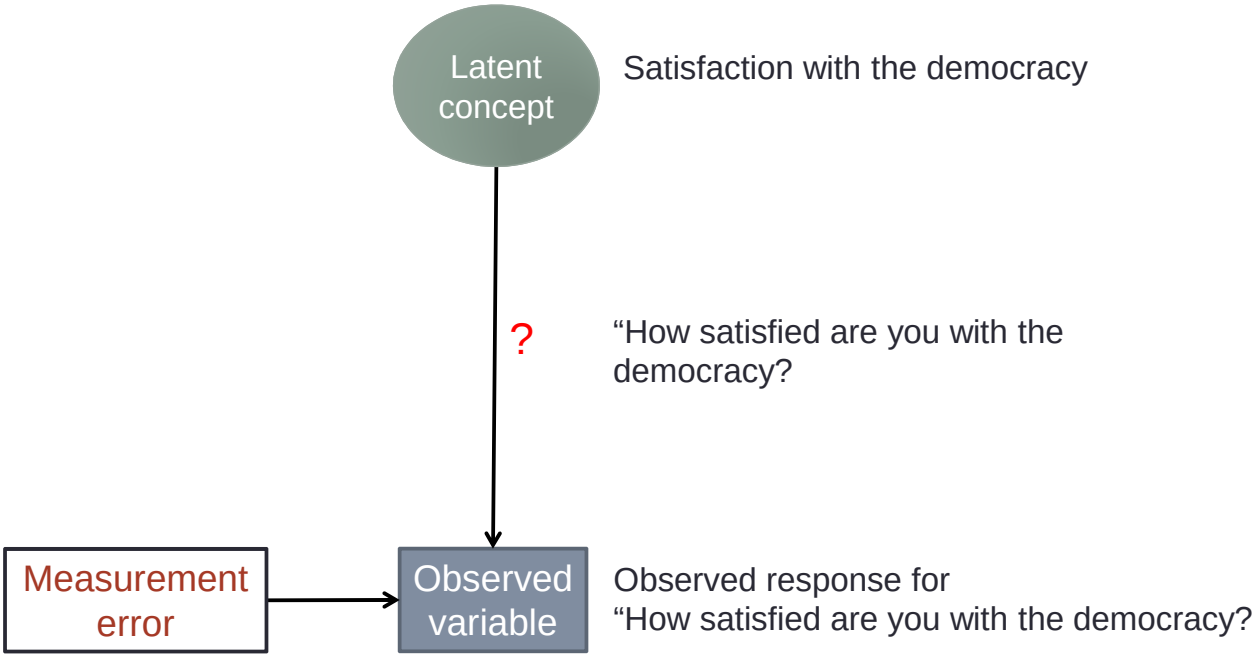
Applicability using Stata

Benefits and possibilities

WHAT DO WE MEASURE? (I)



WHAT DO WE MEASURE? (II)



WHAT IS MEASUREMENT ERROR?

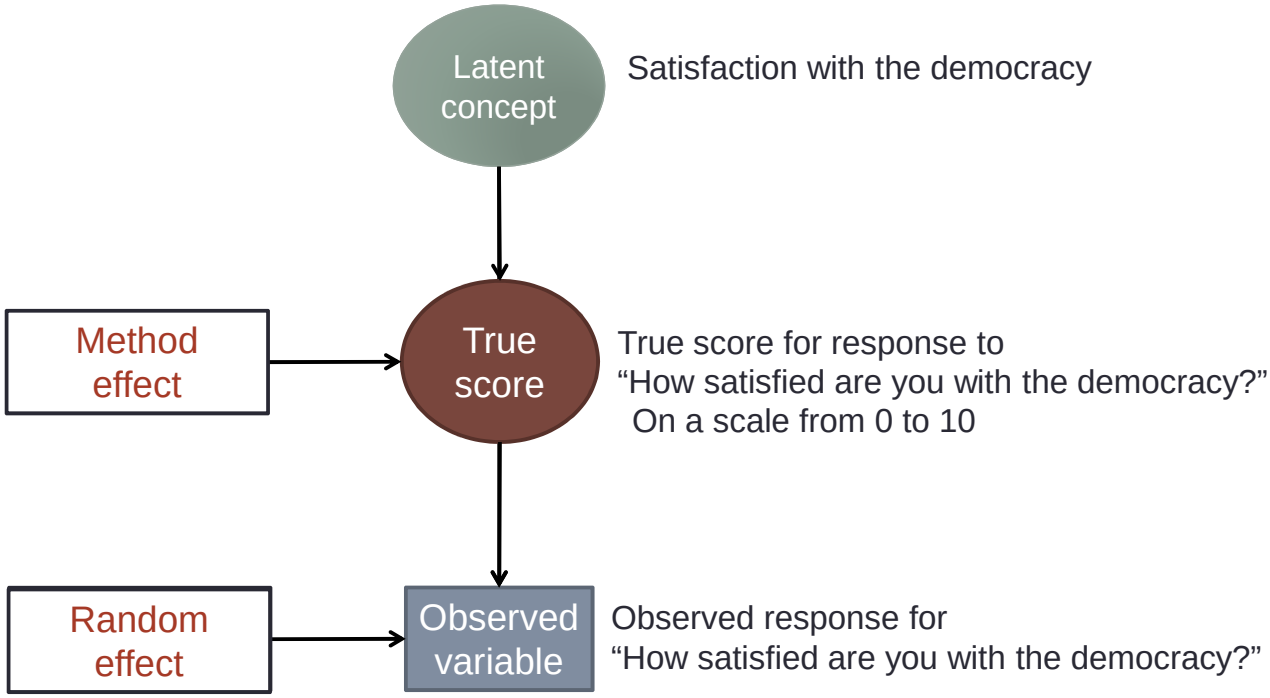
- There are two components of M.E.:

- **Random error**
  - Captures the effect of unintended and unpredictable fluctuations of the respondents, interviewers, coders, etc…

- **Systematic error or method effect**
  - Captures the effect of the reaction of the respondents to a particular formulation of a question.
  - Respondents can react differently to different formulations of questions even if the concept asked is not changed.

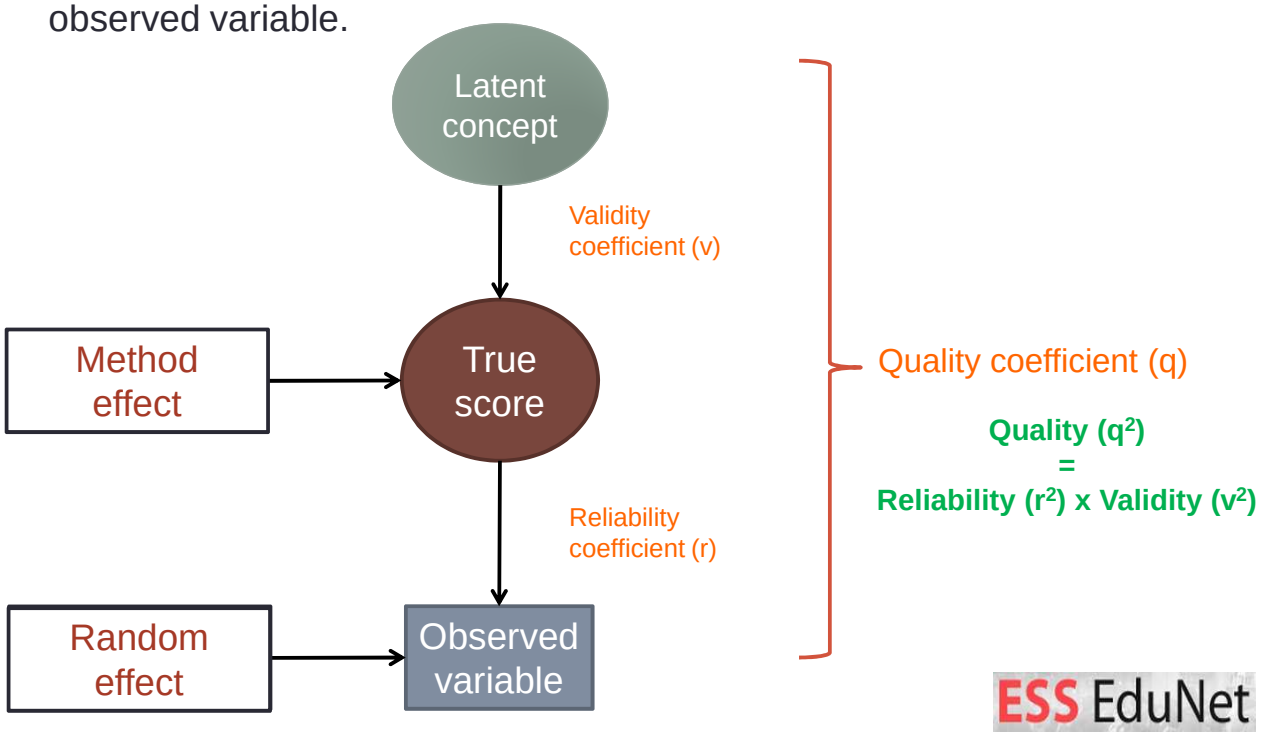
Already discussed in: Campbell and Fiske (1959), Schuman and Presser (1981) and Sudman and Bradburn (1982) and many others.

WHAT DO WE MEASURE? (II)



HOW IS THE QUALITY DEFINED?

- **Quality (q^2)** is the strength between the latent concept and the observed variable.



ESS EduNet

HOW DO WE OBTAIN QUALITY?

- **Option 1: Conduct a Multitrait-Multimethod (MTMM) experiment.**

  Already discussed in: Campbell and Fiske (1959) and Andrews (1984)

- **Option 2: Use alternative approach…**
  - **Over the last decades many MTMM data have been collected**
    - Database of:
      - 3,726 questions with quality information
      - In more than 20 countries and languages
      - From multiple surveys

European Social Survey

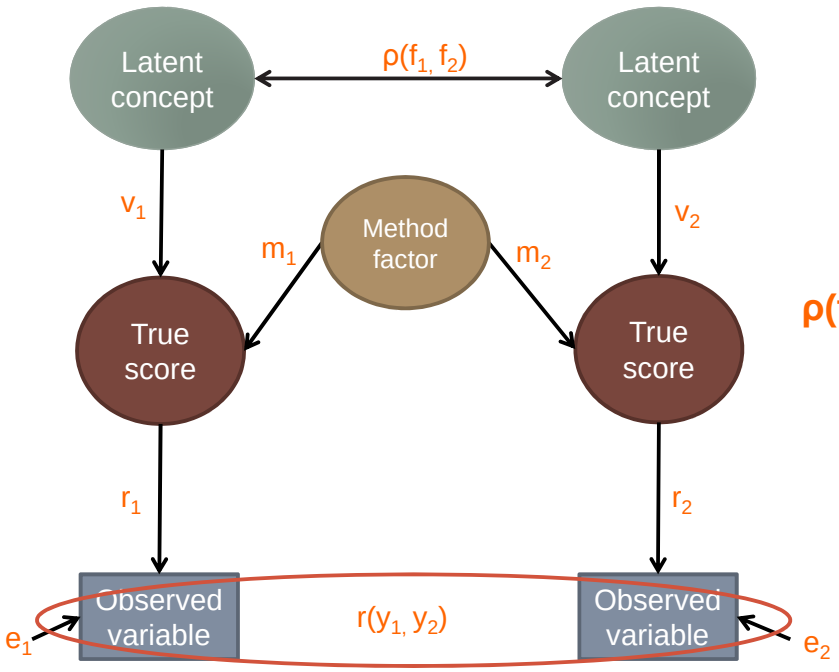
  - **The formal and linguistic characteristics of these questions were carefully coded**
    - The quality obtained from the MTMM experiments could be related to the characteristics of the survey questions.
  - **A new tool was developed:**
    - Allows to predict the quality of any survey question
    - Requires only the coding of the characteristics of the survey question
    - Provides the information about the reliability and validity
    - It is available online for free: [sqp.upf.edu](sqp.upf.edu)

SQP 2.0 Survey Quality Predictor

Already discussed in: Saris and Gallhofer (2014) and Oberski et al. (2011).

HOW CAN WE SIMPLY CORRECT FOR M.E.?

- **Correction of the observed correlation matrix**



- **Formula:**

$$r(y_1, y_2) = r_1 v_1 \rho(f_1, f_2) v_2 r_2 + r_1 m_1 m_2 r_2$$

$$\rho(f_1, f_2) = \frac{[r(y_1, y_2) - CMV_{12}]}{q_1 q_2}$$

EXAMINING THE FORMULA

$$\rho(f_1, f_2) = \frac{[r(y_1, y_2) - CMV_{12}]}{q_1 q_2}$$

- The correlation between two observed variables $r(y_1, y_2)$ is known.
- The common method variance (CMV) is the factor that decreases the over estimation of the observed correlation of those variables that share the same method.
- The CMV between two variables (CMV_{12}) is calculated as: $r_1 \cdot m_1 \cdot m_2 \cdot r_2$
- The method effect m_i can be calculated as: $\sqrt{(1 - v_i^2)}$
- The quality coefficients q_i can be calculated as: $r_i \cdot v_i$

The reliability and validity coefficients r_i and v_i can be obtained from:

OUTLINE

Theory

Applicability using Stata

Benefits and possibilities

GERMANY'S CASE ESS ROUND 6

- **Model variables:**
  - **Satdem:** Satisfaction with the democracy in Germany
  - **LRplace:** Self-placement on the left-right political scale
  - **Free:** Belief of freedom and fairness of elections in Germany
  - **Critic:** Belief of opposition parties' freedom to criticize the German government
  - **Equal**: Belief that courts treat everyone the same
  - **Income:** Household income
- **Regression model:**

$$Satdem = \alpha + \beta_L\, Lrplace + \beta_F\, Free + \beta_C\, Critic + \beta_E\, Equal + \beta_I\, Income + \zeta_S$$

ESS EduNet

ANALYSIS WITHOUT CORRECTION FOR M.E.

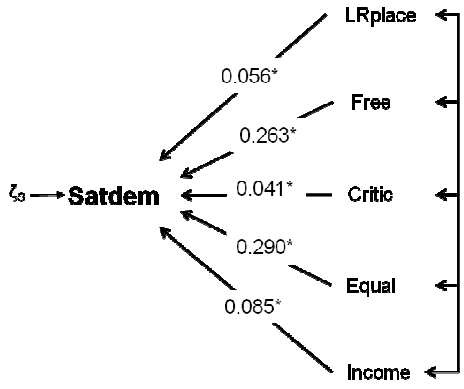
- **We can analyse our model based on the correlation matrix using…**

STATA®

```
ssd init satdem free critic equal lrplace income   /*variables*/
ssd set observations 2358                          /*observations*/
```

```
*Correlation matrix input
#delimit ;
ssd set correlations
1.000\
.3979 1.000\
.2091 .4144 1.000\
.4050 .3546 .1772 1.000\
.0733 .0051 -.0387 .0504 1.000\
.1779 .1724 .1185 .1410 .0290 1.000 ;
#delimit cr
```

```
*Regression model
sem (satdem <- free critic equal lrplace income), standardized
estat eqgof
```



- R^2: Only **25%** of the variance is explained

STEP 1: GET QUALITY INFORMATION

- We coded the characteristics of the 6 questions in our model using the **SQP 2.0 coding process**.
- The quality information is obtained:

	r	v	q	r^2	v^2	q^2	m
Satdem	0.900	0.937	0.844	0.811	0.878	0.712	0.349
Free	0.864	0.898	0.776	0.746	0.806	0.602	0.440
Critic	0.870	0.895	0.778	0.756	0.801	0.606	0.446
Equal	0.871	0.909	0.792	0.758	0.826	0.627	0.417
LRplace	0.874	0.937	0.820	0.764	0.879	0.672	0.348
Income	0.847	0.924	0.783	0.718	0.854	0.613	0.382

- Where method effect m_i is calculated as: $\sqrt{(1-v^2)}$

STEP 2: CORRECTION OF CORR MATRIX

- Observed correlation matrix without correction:

	Satdem	*Free*	*Critic*	*Equal*	*LRplace*	*Inc*
Satdem	1					
Free	0.3979	1				
Critic	0.2091	0.4144	1			
Equal	0.4050	0.3546	0.1772	1		
Lrplace	0.0733	0.0051	-0.0387	0.0504	1	
Inc	0.1779	0.1724	0.1185	0.1410	0.0290	1

$$\rho(f_1, f_2) = \frac{[r(y_1, y_2) - CMV_{12}]}{q_1 q_2}$$

- New correlation matrix corrected for measurement errors

	Satdem	*Free*	*Critic*	*Equal*	*LRplace*	*Inc*
Satdem	1					
Free	0.6075	1				
Critic	0.3184	0.4422	1			
Equal	0.6059	0.3531	0.0584	1		
Lrplace	-0.0318	0.0081	-0.0606	0.0776	1	
Inc	0.2692	0.2837	0.1945	0.2274	0.0452	1

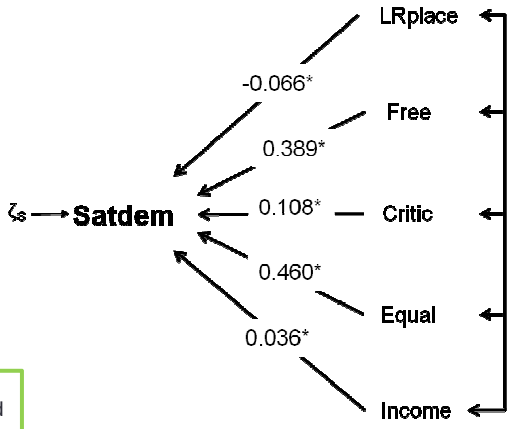
ANALYSIS WITH CORRECTION FOR M.E.

- **Analysing the new correlation matrix corrected for measurement errors using…**

```
ssd init satdem free critic equal lrplace income  /*variables*/
ssd set observations 2358                         /*observations*/
```

```
*Correlation matrix input
#delimit ;
ssd set correlations
1.000\
0.608 1.000\
0.318 0.442 1.000\
0.606 0.352 0.059 1.000\
-0.032 0.008 -0.061 0.078 1.000\
0.269 0.284 0.195 0.227 0.045 1.000;
#delimit cr
```

```
*Regression model
sem (satdem <- free critic equal lrplace income), standardized
estat eqgof
```



- R^2: Now **56%** of the variance is explained

COMPARING THE RESULTS WITH AND WITHOUT ME

	Results without corrections		Results with corrections	
Effects	**Coeff**	**E.Var**	**Coeff**	**E.Var**
Satdem <-		0.750		0.439
Free	0.263**		0.389** +0.126	
Critic	0.041*		0.108** +0.067	
Equal	0.290**		0.460** +0.170	
Lrplace	0.056**		-0.066** -0.122	
Income	0.085**		0.036* +0.49	
R^2	**0.249 (25%)**		**0.561 (56%)**	

** if α<1% and * if 1%<α<5%

OUTLINE

Theory

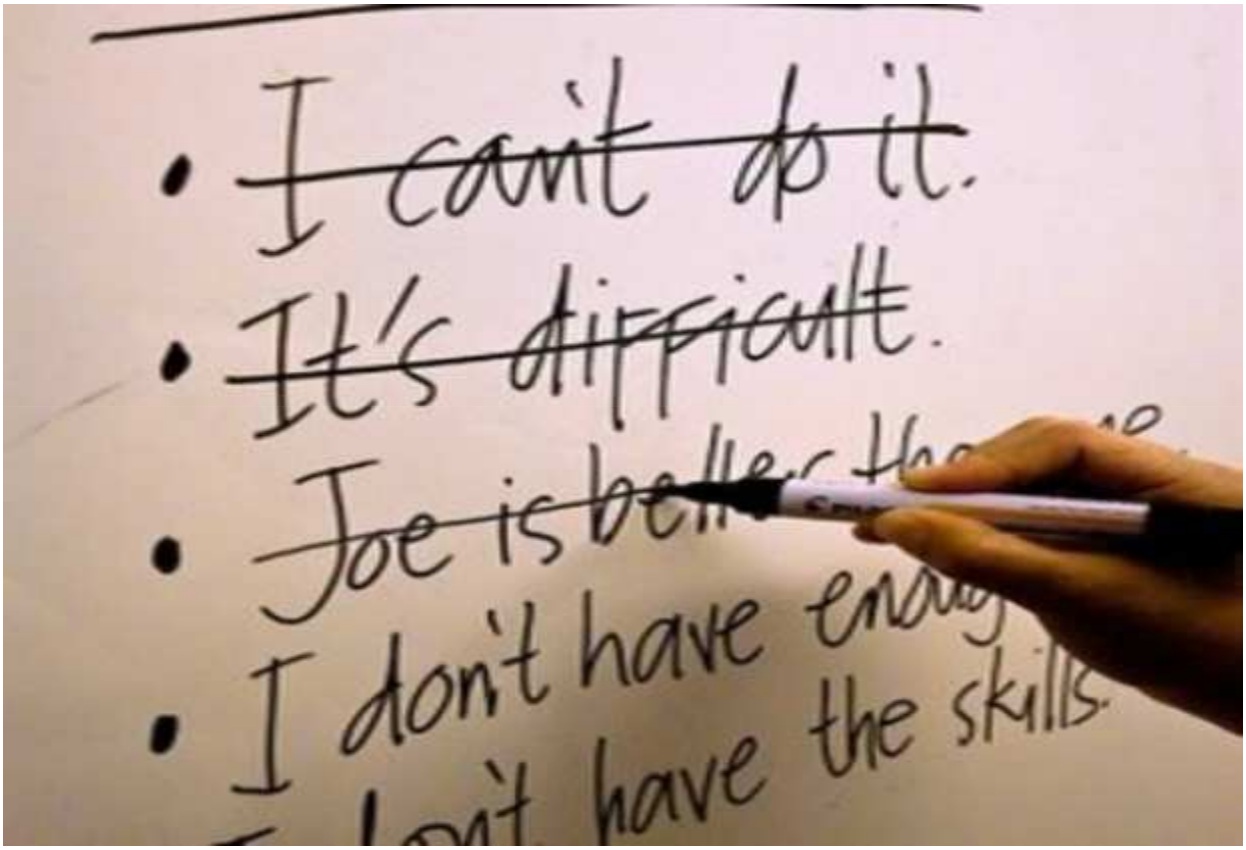
Applicability using Stata

Benefits and possibilities

Benefits and possibilities

- **Benefits:**
- Your results will be better
- The R^2 of your model will increase.
- You don't need to perform an experiment to test the quality of your measures.
- SQP is available online for free.
- Comparability across countries

- **Possibilties with Stata:**
- SEM is simple in Stata when the correlation or the covariance matrix is used.
- The covariance matrix can also be corrected for M.E. to obtain the unstandardized results.
- Different models that can be applied in Stata are illustrated in the Edunet module.



THANK YOU FOR YOUR ATTENTION!

Further information in:

"A simple procedure to correct for measurement errors in survey research"

Written by: Anna DeCastellarnau and Willem Saris

http://essedunet.nsd.uib.no/cms/topics/measurement/

anna.decastellarnau@upf.edu

References

- Andrews, F. M. (1984). Construct validity and error components of survey measures: a structural modelling approach. *Public Opinion Quarterly*, 48, 409-442.
- Campbell, D. T. and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrices. *Psychological Bulletin*, 56, 81-105.
- DeCastellarnau, A. and Saris, W. E. (2014). A simple procedure to correct for measurement errors in survey research. European Social Survey Education Net (ESS EduNet). Available in: ESS EduNet
- Oberski, D.; Gruner, T. and Saris, W. (2011). The prediction procedure of the quality of the questions based on the present data base of questions. *The development of the program SQP 2.0 for the prediction of the quality of survey questions. RECSM Working Paper 24, 71-88*.
- Saris, W. E. and Gallhofer, I. N. (2014). Design, evaluation and analysis of questionnaires for survey research. Second Edition. *Hoboken, Wiley*.
- Schuman, H. and Presser, S. (1981). Questions and answers in attitude surveys: Experiments on question form, wording and context. *New York: Academic Press*.
- Sudman, S. and Bradburn, N. M. (1982). Asking questions. *Jossey-Bass social and behavioral sciences series.*