Work Procedures for Empirical Research (Regression and SEM)
--- Step by step to go from data to statistical models and to good explanation

*Alex Liu ****

This is a summary of a step-by-step approach to empirical research that was developed and used by Dr. Alex Liu – it covers the research process from data preparation and model specification to the estimation of statistical models including structural equation models and to a good explanation of research findings.

1) Check data structure to ensure a good understanding of the data
   - Is the data a cross sectional data? is implicit timing incorporated?
   - are categorical variables used?

2) Check sampling
   - What is the population and sampling method?

3) Check missing values
   - don't know or forget as an answer may be recoded as neutral
   - OR treated as a special category
   - some variables may have a lot of missing values
   - to recode some variables as needed

4) Study qualitative background
   - Use this study to form some hypotheses
   - to select some key dependent variables as the effects

5) Conduct some descriptive studies to begin telling stories
   - use comparing means and crosstabulations
   - check variability of some key variables (st dev and variance)

6) Select groups of ind variables (exogenous variables)
   - as candidates of causes

7) Basic descriptive statistics
   - mean, st dev and frequencies for ALL vars

8) Measurement work
   - study dimensions of some measurements (EFA exploratory factor analysis may be useful here)
   - may form measurement models

9) Local models
   - Identify sections out from the whole picture
   - To explore about relationship

- use crosstabulations
- graphical plots
- use logistic regression
- use linear regression

10) Conduct some partial correlation analysis to help model specification
11) Propose structural equation models by using results of (9)
- identify main structures and sub structures
- Connect measurements with structure models

12) Initial fits
- use SPSS to create data sets for LISREL or Mplus
- programming in LISREL or Mplus

13) Model modification
- Use SEM results (mainly model fit indices) to guide
- Re-analyze partial correlations

14) Diagnostics
- distribution
- residuals
- curves

15) Final model estimation may be reached here
- if not repead step 13 and 14

16) Explaining the model (causal effects identified and quantified)

Obtaining good statistical models is usually not the end of a research. Deriving causal inference or causal explanation is often needed, in order to make the research useful.

In this step, as a kind of common practice, theoretical knowledge and logic and common sense are frequently used, but often in an implicit and informal way.