Empirical Research Methods

1: The Scientific Method

1.1 – Evidence Based Management

Evidence Based Management

Evidence based management should by applied by managers to eliminate wrong beliefs and cognitive biases (too much/litte information, time pressure) in their decision making. Rational decision making relies on objective data and formal processes of analysis.

- Evidence comes from data & data interpretation > need for sufficient methodology!

1. Collect evidence: conduct studies
2. Aggregate evidance: meta-analysis
3. Translate: develop principles and guidelines for action
4. Show efficacy: evaluate guidelines

1.2 – Theory

Definitions

A theory is a causal proposition, which is testable through hypotheses.

- Causal: describing how/why
- Proposition: causal statement linking two constructs
- Construct: not directly measurable (vs. objects), e.g. intelligence, motivation, performance
  > Operationalization: defining measurables for constructs
- Hypothesis: testable statement derived from linking to measurables

Theory

Theory (set of [*corroborated*] causal conjectures) gives meaning and allows for prediction.

- Abstraction: identifying generic attributes of a specific problem
- Generalization: specific solution may solve many similar problems
- Intervention: change things

Scientific Method

- Empiricism: things need to be observable
- Objectively: data collection should be objective (almost impossible)
  > fully disclosing the methodology used
- control: data collection and presentation without bias
  > Selection of people and method *self-selected, survey*
  > Application of selected method *mean vs. median, correlation vs. causation*
  > Presentation of results *full picture vs. selected parts*

Good Theories

- Falsifiability: it is possible to conduct experiments proving the theory wrong
- Accuracy: must be good in explaining/predicting
- Parsimony: should demand for a minimum of preassumptions (if explaining similar well)

Deductions vs. Induction

Induction : Empirics → Theory Deduction : Theory → Empirics (prove)

1.3 – Value Chain of Empirical Research

① State of Current Research

- Definition of research field and research question

② Research Design

- Method on data collection
- Operationalizazion (developing measures for constructs)

③ Data Collection

- Sampling
- Pretesting

④ Data Anlaysis

- Data preparation (coding, data „cleaning")
- Descriptive & inferential statistics

⑤ Publication

- Interpretation of results
- Writing and submitting research paper

2: Research as Conversations

2.1 – Research as Conversations

Field of Research

Shared… phenomenon of interest level/object of analysis theoretical perspective

Conversations

A field of research can be compared to a group of people standing together talking. You can make a contribution that is

- Relevant: it relates to the ongoing academic conversation
- Novel: it says something that has not yet been shared
- Interesting: it is actually interesting

Conversations/Fields of Research are referenced in first paragraph(s) of papers usually.

2.2 – Making a Contribution

Interesting Questions

What is interesting for you? Are other interested in this? Conflicting views?

- Conflicting results: previous studies on same topic do not converge
- Boundary conditions: under what conditions does this theory hold
- Increase specificy: theory has not been applied here
- Study new phenomena: has not been explored before

Desingning Research Questions

- Do/Is > Yes/No: novel effects, conflict in literature, isloated variables
- What/How: studying processes (qualitative work: discover & explain relationships)
- Why: causal reasons
- To what extent: diff. in input > diff. in output (quantitative work: measure effects, corroborate relationships)

A research project needs focus & precision: only study one thing at a time.

Research Question to Theory Section

Research Question > Quantitative & Experimental Work > Qualitative Work

3: What is good Research?

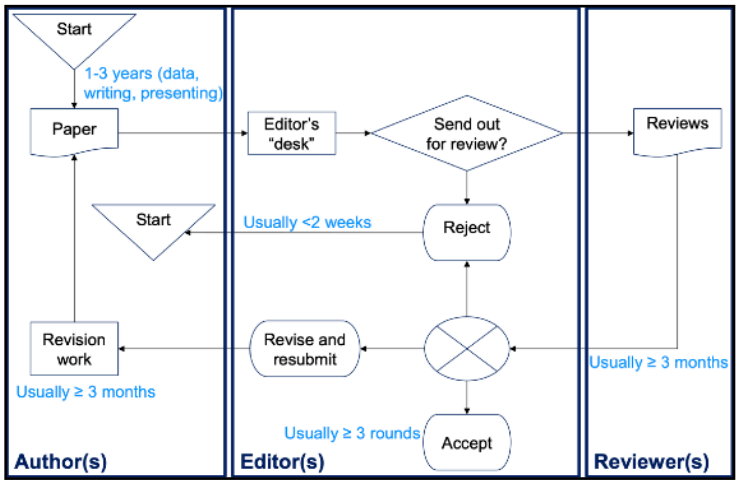
3.1 – The Publication Process

Why Publications?

- Academic debate advances through publication
- Ensure novelty & high methodological standards
  > Peer Review: submissions are forwarded to other scientists to be judged

Publication Process

- Good journals rejection rate is larger than 80% in the 1st round & about 50% in the 2nd round



Arguments on the process:

–) Reviewers never agree > but editor selects feedback
–) Bad papers still published > but minimized
–) Politicizing the process > but words get arround
–) Wasted efforts/resources > but progress if science, non-finding is a finding

- Editors (board of editors) are gatekeepers of scientific process: more effect than authoring
- Reviewing essential to scientific process, reviewers are rated too, 1 submission ≈ 3 reviews

3.2 – Measuring Impact

Tools to determine Quality

- Journal ranking: objective (impact factore) & subjective (surveys)
- Citations: derivative measures (e.g. h-index)
  > Boundary conditions: different citation style in different fields, co-authorships
- Combined measures: (discipline, age, citations, journals, graduate income)
  > Handelsblatt ranking, FT ranking, etc.

3.3 – Research Ethics

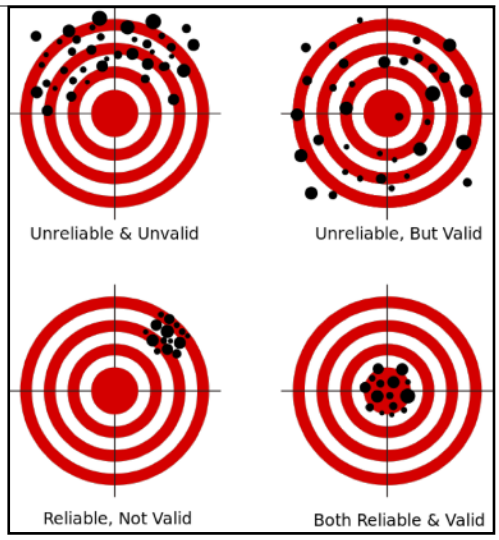
Ethics

- Scientist working for dictatorships (against mankinds best)
- Plagiarism in PhD-Thesis
- False claims to be ‚first'
- Rise of retractions (journals taking back publication of paper), perhaps due to digital technologies

Rules for Research

- Do not cause physical harm to anyone.
- Do not deceive people (anonymity vs. confidentiality).

3.4 – Validity, Correlation & Causation

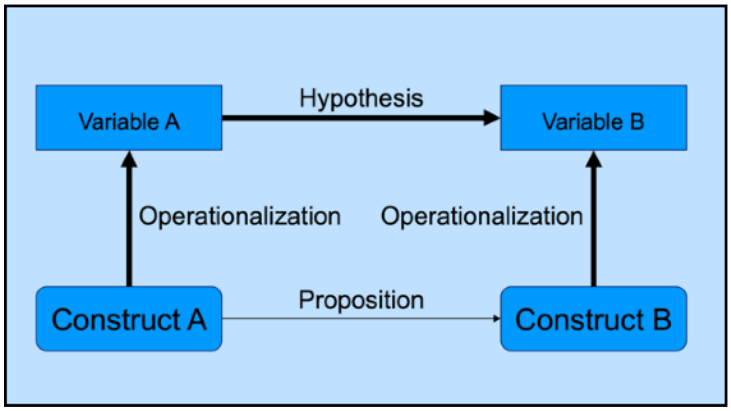


Reliability & Validity

- Reliability: Measuring again with different method will lead to same result. [dt. *Zuverlässigkeit*]
- Validity: Doing, what you're saying. (free of bias) [dt. *Gültigkeit*]
  - > Internal validity: Did the experimental treatments make a difference? Is a causal conclusion warranted?
  - > External validity: To what populations, tratment variables & measurement variables can this effect be generalized?
  - > Face validity: Does the measurable fit to the construct?
  - > Construct validity
    - convergent validity: correlated with similar scales/constructs?
    - discriminant validity: is distinct from other scales/constructs?
    - content validity: is the scale/construct really fully captured?

Correlation vs. Causation

- Correlation: there is a connection between A and B free of bias
  - > May be caused by third mechanism
  - > May be caused by selection (just look until you found)
  - > My be just a coincidence (random)
- Direction: explain, how $A \rightarrow B$ or $B \rightarrow A$
- Causation: one phenomenon is the result of the occurence of another phenomenon
  - > Argue, why this is a causation and not just a coincidence & give the direction
  - > Exclude any possible third mechanism



4: Designing good Research Projects

4.1 – Methods of Data Collection

Data Collection

Data Collection should facilitate

- Generalizability: reach as many people as you can
  > Sample Surveys, see sampling strategies
- Precision: have full control over study
  > Laboratory experiments, control 100% of the environment
- Realism: study the actual phenomenon in life (vs. lab)
  > Field studies, mostly correlation; Field experiment, study with systemativ manipulation

Methods & Conversation

State of Conversation	Nascent	Intermediate	Mature
Research Question	Open-ended inquiry on new phenomenon	Proposed relationsh. Between new & establ. Constructs	Focused question on existing constructs
Type of Data collected	Qualitative, interpred meaning	Hybrid	Quantitative, focused on measures
Method of Collection	Interviews, Field Studies	–	Lab, Surveys, Experiments

Papers can apply mixed methods to enable better results.

4.2 – Sampling

Sampling

Sampling describes to process of picking the observed group from a population while ensuring a certain validity and representativeness to allow to draw generalizable conclusions.

- Selection bias, drop-outs, survivorship bias, etc.

Propability Sampling

- Simple Random Sampling: just randomly pick units from the entire population.
- Stratified Sampling: Population is segmented into mutually exclusive subgroups/strata, from which units are randomly selected (from each stratum)
- Cluster Sampling: Population is segmented into clusters. A cluster is randomly selected and all units of that particular cluster are investigated.

Non-propability Sampling

- Convenience Sampling: units are selected at convenience of researcher (e.g. family)
  > Not generalizable, but useful information for pilot study
- Quota sampling: sampling until certain demographic variables are as demanded
- Snowball sampling: reach one unit of a hard-to-reach population and sample from surrounding
- Judgement sampling: researcher selects units from the population
- Theoretical sampling: selection of extreme/specific cases to ensure observability

5: Qualitative Research

5.1 – The Case for Qualitative Research

When to apply Qualitative Research?

- Describe a phenomenon/process, first observation in the field
- How/Why: explain the research subject
  > Theory testing: test existing theory (often quantitative)
  > Theory eloborating: pre-existing ideas (develop theory)
  > Theory generating: something, which was not been looked at before

Qualitative Methods

- Characteristics: in the field, data derived from participants perspective, flexible research design
  > Collect & interpret without prior knowledge > let data speak > inductive
- Method fulfilling criteria (potentially applied mixed)
  > Case study research: gather all information on one particular setting in field
  > Process research: observe a process
  > Ethnography: go deep into social group (will encourage real behaviour, pot. subjective)
  > In-depth interviews

Critique of Qualitative Research: Subjectivity!

- Samling: sample sizes are typically small & not really random > lack of generalizability?
- Measures: no quantitative (metric) variables
- Analysis: no statistical analysis, no defined processes

5.2 – Qualitative and Quantitative Content Analysis

Coding Data and Categorization

- Unstructured data needs to be aggregated to higher-order constructs
  > Observe variable in field > identify constructs behind variables
- Difficult: idetify patterns reliably, open to others (traceability) > multiple coders

Raw Data → Higher order themes → General dimensions

Coding

Coding is data reduction and structuring.

- Reading and marking
- Iterative aggregation of data (perhaps going back)
- Theorizing (explain beyond particular case) on top-level > theoretically relevant constructs emerge

5.3 – Case Study Research

What is Case-Based research?

Gather all available information on one (but ideally multiple cases > generalizability) particular case. Case-Research usually handels with relatively current data (to prevent bias over time).

Case Study à la Eisenhardt

1. Beginnning: definiton of research question no hyptohesis and theories
2. Case Selection: determination of population, selection of sample
   > Deliberate selection based on properties & theoretical reason (extreme situation & differing)
3. Development of Instruments & Protocols: method for data collection, cooperation with other researchers to increase confidence
   > Different methods enable triangulation, quantitative (strengthen impressions), qualitative (understand causal relationships & derive theoretical conclusions)
4. Entry into the Field: data collection and analysis (flexible data collection)
5. Data Analysis: analysis within case, search for pattern between cases
6. Hypothesis Formation: search for similarities between cases > search for explanation
7. Relating to Literature: comparison with suporting & contradicting literature
8. Finalizing: theoretical saturation

5.4 – Interviews

Interviewing Techniques

- Critical Incident: pinpoint to critical events (from interviewee's perspective)
- Storytelling: let the interviewee answer broad question
- Courtroom Questioning: concrete questions, like a prosecutor, point out inconsistencies
- Event Tracking: let the interviewee describe an event in chronological order
- Non-directive Questioning: not directly related to subject but more likely to provoke desired answer

Who to Interview?

- The Average Joe
- Persons with much knowledge but little responsibility

6: Getting Data (quant.)

6.1 – Getting Data

Written Interaction

- Postal Survey: surveys are sent out & back by mail (e-mail/fax)
  > Cost effective, no interviewer bias, large amout of data
  > Cover Letter: general topic, confidentiality/anonymity, eliminate bias
- Group Survey: survey is sent to a key contact person within a group (who distributes to group)
- Drop-Off Survey: survey is handed out to subjects - may return
- Online Survey: internet tools are used to make subjects complete the survey online
  > Inexpensive, fast, easy to modify; easy to manipulate, junk mail syndrome
  > Use ‚progress bar', check response time to filter manipulation

Personal Interaction

- In-Home/Office Interview: interview/surveying at the subjects base (pos. computer supported)
  > Suitable for non-audio information; costly
- Mall-Intercept Interview: ask subjects to participate while at mall/shopping center
  > Convenience-sample, must be fast to stay socially acceptable
- Purchase-Intercept Interview: ask subjects directly at point of purchase
- Phone Interview: interview by phone
  > Cost efficient & scalable, no non-audio information
- Computer-Assisted Telephone Interview (CATI/CATS): computers call subjects and guide them trough survey

Administered Questionaire

+) Fewer misunderstandings/inappropriate responses
+) Higher response rate
+) Greater control over environment
+) Collect additional information
–) More expensive (more people, more administration time)
–) More skilled interviewers needed

6.2 – Question Design

General

Each question asks for *only one* thing, questions are grouped by topics.

- Dos: layman terms, sensitive
- Don'ts: don't lead responses, no ambigues terms - be clear!

Closed vs. Open Questions

- Open End: free text field
- Closed End: precise questions, defined answer options (easy to analyze)

Guttman Scale

Nested levels of agreements to question (several statement = different levels)

10 Simple Rules

1. Avoid ambiguous/ambivalent wording.
2. Avoid words/expressions the respondent may not know.
3. Cover only one question in one questionaire item.
4. Avoid negative instructions (‚mark, what you disagree with').
5. Avoid overly short/long sentences (max. 20 words).
6. Avoid complex questions (split into multiple questions).
7. Avoid abbreviations.
8. Avoid suggestive questions.
9. Avoid questions that migh evoke moral obligation/social desirability effects.
10. Provide full range or response options.

6.3 – Conducting the Actual Survey

Layout

Clear, reasonable length.

Control in Survey Design

- Self-Evaluation, e.g. with back translation, let friends fill out the survey
- Pilot studies: evaluation with test sample from population.

Response Rates

Reliable response rates differ with design (online: 10%, network studies > 80%). Increase resp. rates: short questionaire, (non-)financial incentives, personalisation, etc.

6.4 – Source of all Quantitative Research

Measurement

Everything can be measured or be made measurable.

- What do I want to measure?
- How do I measure it? (reliability & validity, resource constraints, previous experience)

Source of Quantitative Data

- Primary Data: data originally collected for focal study
- Secondary Data: data that already exists
- Internal Data: proprietary & confidential data
- External Data: publically available data

Mixed: Create primary data from secondary data by yourself.

Temporality of Data & Measurement

- Cross-Sectional Data: gathered across population/sample but at *only one point at time*
- Panel: repeated observation of the same object of study over time

Variables

Format
- Nominal: no natural ordering (e.g. gender)
- Ordinal: natural ordering (e.g. education)
- Metric: distance between values is meaningful (interval: no natural zero; ratio: natural zero)

Perceivability
- Manifest: can be observed directly
- Latent: cannot be observed directly, operationalization needed

Outcome
- Dichotomous: only two variables (binary)
- Discrete vs. Continous

6.5 – Data Collection

Primary: Survey

Survey must be clear, participants must trust, common-method bias: survey must not reveal the research question of participant

1. Welcome: encourage and instruct people in fitting tone
2. Legitimation: point out importance & legitimacy „university“
3. Motivation: why should one participae > prices, money
4. Story: survey should follow a plan
5. Gain trust: ask sensitive question later

Secondary: Scales & Data

Existing scales can be captured from literature (previous studes, books, etc.)
- Scale: set of items that jointly measure a construct on different dimensions
  > Different dimensions should result in correlating results

Use existing dataset as this is easier and safes money & time.
- Evaluate secondary data precisely: validity & reliability concers, biases (you don’t know the questionaire design > blackbox)

7: Preparing Data (quant.)

7.1 – Preparing & Importing Data

Software Tools

- Stata: proprietary, efficient for large data sets, many additional packages
- R: open source, often used in teaching
- SPSS: proprietary, graphically & easy to use

Preparation of Data

- Coding: categorize responses for analysis
  > Transform survey answers, text, etc. into processable numbers (based on *code book*)
  > Missing values: Imputation (infer answers from other answers), Scales & Indices (not importan for scales consiting of multiple items), Drop Observations
- Editing: fix errors (e.g. incorrect defined variables), fix incorrect responses (e.g. always the same answer)

Important: Never overwrite the original dataset!

7.2 – Descriptive Statistics

Central Point

- Mode: the value most frequently observed (dt. *Modalwert/Modus*)
- Median: central value in ordered observations (50th percentile)
- Mean: arithmetic average

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \quad \text{population mean} \qquad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{sample mean}$$

Dispersion

Spread around central point is characterized by the variance.

$$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N} \quad \text{Variance} \qquad \sigma = \sqrt{\sigma^2} \quad \text{Standard deviation}$$

$$\max_{i=1}^{N}(x_i) - \min_{i=1}^{N}(x_i) \quad \text{Range}$$

Standard Error

Degree, to which the estimated value differs from its true value (not descriptive statistics, but inferential statistics [testing]).

$$\mathrm{SE}_{\bar{x}} = \frac{s}{\sqrt{n}}$$ where s is sample standard deviation, n is size of sample

Skewness and Kurtosis

- Skewness: how different is the distribution from the normal distribution
  > Longer tail: negative = left, positive = right
- Kurtosis: were do most of the observations happen (peak, shoulder, tail)

7.3 – Correlation Analysis

Correlation

Correlation describes the ‚relatedness' of variables.

$$\text{Cov}(x_i, y_i) = \frac{1}{n-1}(\tilde{x}_1\tilde{y}_1 + \ldots\tilde{x}_n\tilde{y}_n) \quad \text{where } \tilde{x} = \mu - x$$

$$r = \frac{\text{Cov}(x, y)}{\text{s.d.}(x) \cdot \text{s.d.}(y)} \quad \text{Pearsons's correlation coefficient, } r \in [-1; 1]$$

Use correlation matrices to get a first idea on potential correlation.

Outliers

Outliers can have severy effects on correlation, thus shey should be identified and removed.

7.4 – Simplifying Data

Data Reduction

Express the same information using less variables by reducing data.

- Factor analysis: group variables which strongly correlate
- Cluster analysis: group observations which strongly correlate in their replies

Factor Analysis

A factor is a construct grouping common underlying dimesnions in a group of variables.

Factors < Variables

Types of Factor Analysis

Exploratory factor analysis

- Select variables on the basis of prior theory, show construct validity

Exploratory factor analysis

- Uncover structure of a relatively large set of variables
- Used to show (uni)dimensionality of scale, assess reliability of a scale
- Assumption: variables are continuous, variables are normally distributed

1. Correlation matrix: examine which variables are correlated (KMO criterion)
   > Kaiser-Mayer-Olkin: must be $\geq .5$, should be $\geq .8$
2. Extract factos from variables
   > Principal component (factor) analysis: maximize explained variance, retain as much information from original variables as possible; by small number of linear combianations (principal compoenents)
   > Common factor analysis: maximize underlying correlations among underlying variables
3. Factors are rotated to maximize relationships between the variables & factors

How many factors:

- Kaiser criterion: only factors having eigenvalue > 1 are retained
- Percentage of variance criterion: achieve high specified cumultative % of variance extracted by successive factors (usually 60%)
- „Elbow" criterion: plot the eigenvalues & cut where the curves flattens

Terms: Eigenvalue (sum of squared factor loadings of one factor across all variables), Communality (sum of squared factor loadings of one variable), Factor Loading (correlation of a factor and a variable)

8: Regression Analysis

8.1 – Simple Regression & Ordinary Least Squares

Simple Linear Regression Model

The simple linear regression model aims on explaining the linear associations between an independent and an explained variable.

$Y = \beta_0 + \beta_1 X + \varepsilon$ where ε is the error (all uncaptured variables)

β_0	intercept	X	independent/explanatory variable
β_1	slope	Y	dependent/outcome variable
$e_i = Y_i - \hat{Y}_i$	residual error		

Ordinary Least Squares

OLS is an estimation technique aiming to minimize the squares of residuals.

$$\min \sum_{i=1}^{n} e_i^2$$

8.2 – Assumptions of OLS

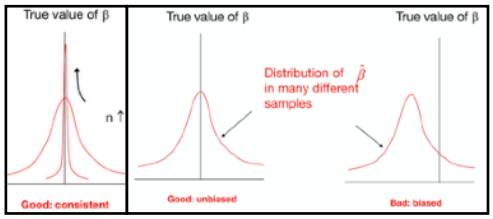
Assumptions about Dependent Variables

- Ranges from $+\infty$ to $-\infty$
- May be any real number, $\in \mathbb{R}$
- Units of measurement are constant

OLS works well for continuous, metric dependent variables (scaling is not part of assumptions).

BLUE

Best Linear Unbiased Estimator



- Regression produced the best (i.e. as closely to reality) etsimations
- Unbiased estimators of a population parameter estimate an expected value equal to the parameter.
- Unbiased estimators are consistent, if the difference betweent the estimator and the parameter grows smaller with a growing sample size

Assumptions about Independent Variables

- Linearity: the regression model is linear in parameters
- Exogenity: error terms have a mean of zero
  > If error term mean is non-zero, it is correlated with the inedependent variables > independent variables may not be independent
  > Prevents omitted variable bias
- Homoskedasticity: constant variance of error terms, otherwise model looses/gaines precision
- No Autocorrelation: zero covariance between the error terms, independent observations
- Variation & No Multicollinearity: there needs to be variance in X, independent variables must not correlate between each other > would lose a dimension (especially the higher the correlation)
- (optional) Sample Size: larg sample to ensure normally distributed error term, zero mean & var.

8.3 – Output of Regression

Output of Statistical Software

- F-Test: overall fit of the model
- Coefficient of determination R^2: how well does the mode explaint the dependent variable
- Regression coefficients: what is the realtionship between denpendent & independent variables
- Significance Level of Coefficients: how to interpret coefficients - relevant?

Variance Decompositions

$\sum (y_i - \overline{y})^2$ total variance = explained variance + residual variance, $0 \leq R^2 \leq 1$

$\sum (\hat{y}_i - \overline{y})^2$ explained variance

$\sum (y_i - \hat{y}_i)^2$ residual variance

F-Test of Overall Fit

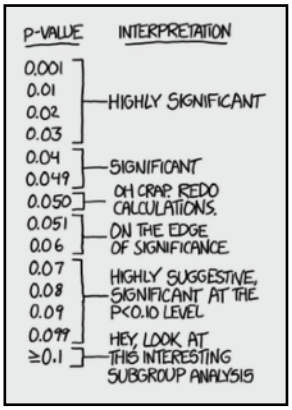
How likely is the model given the number of variables in it > reports a p-value that captures the likelihood that all coefficients are 0 in reality.

Standardizes Coefficients

Bigger coefficient does not neccessarily mean anything as variables have different scalings > standardize scales

$$\beta_k^S = \beta_k(\frac{\sigma_k}{\sigma_y})$$

8.4 – P-Values & Hypothesis Testing



Hypothesis Testing

Hypothesis testing checks for statistically significant differences from the current hypothesis. One usually tries to falsify a current hypothesis.

Test Statistics & P-Value

Use the t-distribution (with $df = n - 2$) in small samples.

$$t = \frac{\text{coefficient}}{\text{standard error}}$$

Significance level (p) can be read from a table using the calculated t.

Type I/II Errors

„Never confuse Type I and II errors again: Just remember that the boy Who Cried Wolf caused both Type I & II errors, in that order.
First everyone believed there was a wolf, when there wasn't.
Next they believed there was no wolf, when there was.
Substitute ‚effect' for ‚wolf' and you're done"

8.5 – Panel Regression

Panel Data

Panel Data contains for the same observation units data for several points in time.

- Balanced Panel: all observations units are included in all time periods
- Unbalanced Panel: observations are not available for all observations units for all time periods

Two sources of variation: Panel Data varies between units and within each units.

+) High external & internal validity (allows control of unobserved heterogeneity)
+) Contains more degrees of freedom > higher efficiency of econometric estimates
–) Data collection costly & time consuming
–) Panel mortality (missing data) > unbalanced data

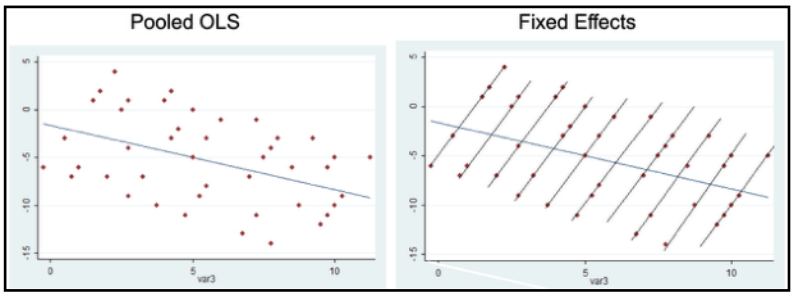
Pooled OLS

$$Y_{it} = \alpha_i + \sum_{k=1}^{m} \beta_k \cdot X_{kit} + \epsilon_{it}$$

Y dependent variable
X_k observed explanatory variables, $k \in [1; m]$
α_i unobserved effect for unit i
ϵ_{it} random error term

Simpsons Paradoxon

Trends within sub-populations can be reversed when the data are aggregated.
Pooled Regression does not consider herterogneity.



Fixed Effects

Sloped are the same for all units, but constants (intercept) differ between units

- Constants capture the combined effect of several unknown variables that are different beetween units but stable over time

8.6 – F-Test

F-Test

The F-Test can be used to assess multiple coefficients simultaneously to verify the significance of a regression model. The F-Test follows the logic of comparing signal (SSR_u) & noise (SSR_r).

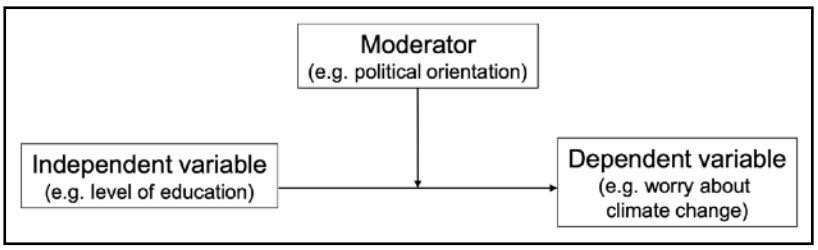
$$F \frac{SSR_r - SSR_u}{q \times (n-k-1)/SSR_u} \qquad df_1 = q, \quad df_2 = n-k-1$$

n observations
k number of independent variables in unrestricted model
q number of variables tested jointly

8.7 – Moderation

Moderation & Interaction

Conditional regression contains moderation effects captured by interaction terms (two variables multiplied)



Mathematical Base

$$Y = \beta_0 + \beta_1 X + \beta_2 Z \longrightarrow Y = \beta_0 + \beta_1 X + \beta_2 Z + \underbrace{\beta_3 XZ}_{\text{interaction term}}$$

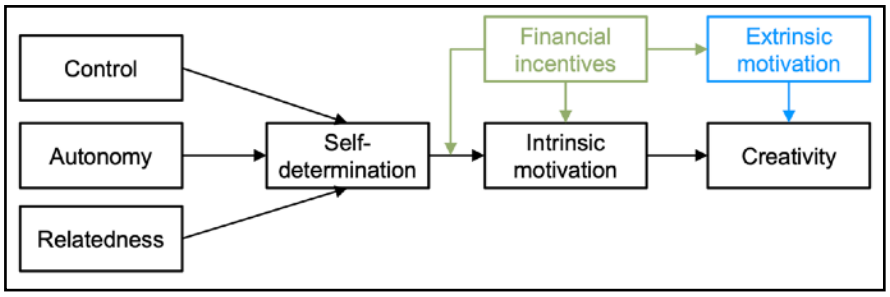
U-shaped & inverted U-shaped

Some increase in X has a postive effect on Y, a lot of increase has a negative effect.

8.8 – Limitations of OLS

Singular Causality

Any regression can only establish on causal pathway. Multi-causality may be modelled through structural equation modeling (SEM, see right) and partial least squares (PLS) methods.



Nature & Distribution of dependent Variables

OLS contains strict assumptions about the nature of the dependent variable (linear). This issue can be solved by using non-linear estimations functions (e.g. sigmoid).

Beyond Dichotomous Variables

- Categorical variables: ordered logit, odered probit
- Count variables: poison, negative binomial

8.9 – Robustness

Robustness Checks

Each paper includes a Robustness Check section.

- Test & exlude alternative hypothesis
- Sensitivity analysis
- Subsample analysis (factor analysis, sub-populations, etc.)
- Additional evidence for thesis