# SOS3003
# **Applied data analysis for social science**
## Lecture note 09-2009

Erling Berge
Department of sociology and political science
NTNU

# Today

Literature

- Hamilton, Lawrence C. 2008. A Low-Tech Guide to Causal Modelling. http://pubpages.unh.edu/~lch/causal2.pdf
- Principal components and factor analysis
  – Hamilton Ch 8 p249-282
- Also see:
  Winship, Chrisopher, and Stephen L. Morgan 1999 "The Estimation of Causal Effects from Observational Data", Annual Review of Sociology Vol 25: 659-707

# Causal analysis

- Experiment
  - Randomized causal impacts ("treatment") provide precise causal conclusions about effects ("response") if there is significant differences in the mean response (effect)
  - Experiments can be impossible to achieve due to
    - Practical conditions
    - Economic constraints
    - Ethical judgements
- Instead one tries to obtain quasi-experiments
  - Using for example regression analysis

# Model of causal effects <sup>Ref.:</sup>

- Research using observations utilize concepts from experimental design

  - "Treatment", "Stimulus"
  - "Effect", "Outcome"

Ref.:

Winship, Chrisopher, and Stephen L. Morgan 1999 "The Estimation of Causal Effects from Observational Data", Annual Review of Sociology Vol 25: 659-707

# Experiments allocate "cases" randomly to one of two groups:

- **TREATMENT (T)**
  with observation
    - before treatment
    - after treatment

- **CONTROLL (C)**
  with observation
    - before non-treatment
    - after non-treatment

# The counterfactual hypothesis for the study of causality

- Individual "i" can a priori be assumed selected for one of two groups
    - Treatment group, T, or control group, C.
- Treatment, t, as well as non-treatment, c, can a priori be given to individuals both in the T- and C-group
- In reality we are able to observe t only in the T-group and c in the C-group

## Modelling of causal effects:
### The counterfactual hypothesis (1)

- There are for each individual "i" four possible outcomes
  - $Y_i(c,C)$ or $Y_i(t,C)$; if allocated to a control group
  - $Y_i(c,T)$ or $Y_i(t,T)$ ; if allocated to a treatment group

  - Only $Y_i(c,$ **given that "i" is a member of C)** or
  - $Y_i(t,$ **given that "i" is a member of T)** can be observed for any particular individual

## Modelling of causal effects:
### The counterfactual hypothesis (2)

More formally one may write the possible outcomes for person no i:

|  | Treatment: t | Non-treat.: c |
|---|---|---|
| T-group | $\mathbf{Y^t_i \in T}$ | $Y^c_i \in T$ |
| C-group | $Y^t_i \in C$ | $\mathbf{Y^c_i \in C}$ |

# Modelling of causal effects:
## The counterfactual hypothesis (3)

- Then the causal effect for individual i is

- $\quad \delta_i = Y_i(t) - Y_i(c)$

- Only one of these two quantities can be observed for any given individual
- This leads to the "counterfactual hypothesis"

# The counterfactual hypothesis: concluding

- "The main value of this counterfactual framework is that causal inference can be summarized by a single question: Given that the $\delta_i$ cannot be calculated for any individual and therefore that $Y^t_i$ and $Y^c_i$ can be observed only on mutually exclusive subsets of the population, what can be inferred about the distribution of the $\delta_i$ from an analysis of $Y_i$ and $T_i$ ?" (Winship and Morgan 1999:664)

## Modelling of causal effects: from individual effects to population averages

- We can observe
  $Y_i (c \,|i \in C)$, but not $Y_i (t \,|i \in C)$

- The problem may be called a problem of missing data
- Instead of individual effects we can estimate average effects for the total population

# Modelling of causal effects (1)

- Average effects can be estimated, but usually it involves difficulties
- One assumption is that the effect of the treatment will be the same for any given individual independent of which group the individual is allocated to
- This, however, is not self-evident

# Modelling of causal effects (2)

The counterfactual hypothesis assumes:

- That changing the treatment group for one individual do not affect the outcome of other individuals (no interaction)
- That treatment in reality can be manipulated (e.g. sex can not be manipulated)

Spring 2010     © Erling Berge 2010     13

# Modelling of causal effects (3)

- One problem is that in a sample the process of allocating person no i to a control or treatment group may affect the estimated average effect (*the problem of selection*)
- In some cases, however, the interesting quantity is <u>the average effect for those who actually receive the treatment</u>

Spring 2010     © Erling Berge 2010     14

# Modelling of causal effects (4)

- It can be shown that there are two sources of bias for the estimates of the average effect
1. An established difference between the C- and T- groups
2. The treatment works in principle differently for those allocated to the T-group compared to those in the C-group
   – To counteract this one has to develop models of how people get into C- and T-groups (**selection models**)

# Modelling of causal effects (5)

- A general class of methods that may be used to estimate causal effects are the **regression models**
- These are able to "control for" observable differences between the C- and T- groups, but not for unequal response to treatment

# Causal modelling

- "path analysis" or "structural equations modelling" go back to the 60ies
- Jöerskog and Sörbom: LISREL
  - Use maximum likelihood to estimate model parameters maximising fit to the variance-covariance matrix
  - Commonly available in statistical packages
    - Covariance structural modelling
    - Structural equation modelling
    - Full information maximum likelihood estimation

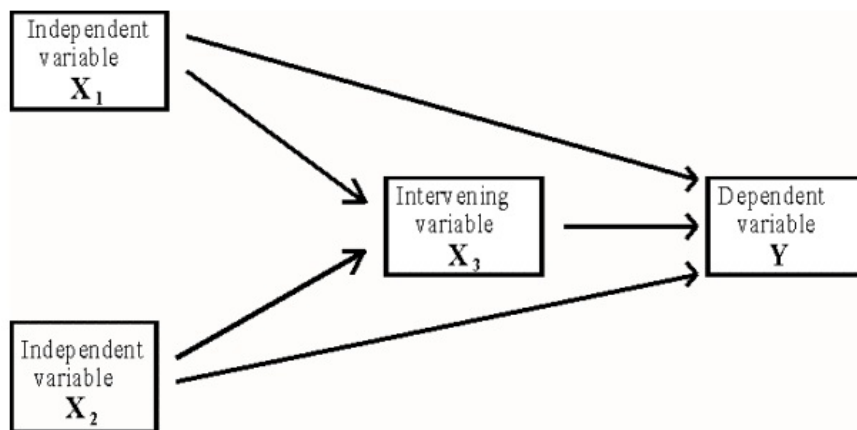Spring 2010                    © Erling Berge 2010                    17

# Low-Tech approach

- Uses OLS to do simple versions of the structural equations models
- The key assumption is the causal ordering of variables. In survey data this ordering is supplied by theory
- The causal diagram visualize the order of causation:
  - Causality flows from left to right
  - Intervening variables give rise to indirect effects
  - "reverse causation" creates problems

Spring 2010                    © Erling Berge 2010                    18
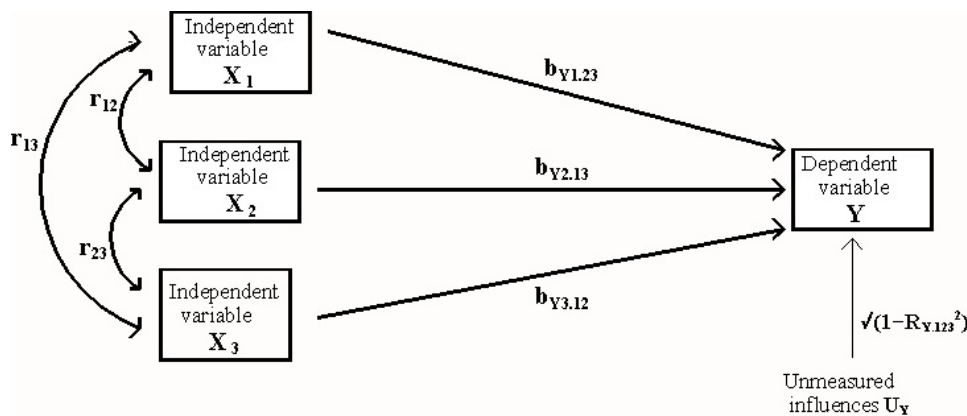
# Low-Tech causal modelling
## Figure 1

# Multiple regression as a causal model
## Figure 2

Ref.:
http://www.svt.ntnu.no/iss/Erling.Berge/

Spring 2010

## Quantities in the diagram

| | |
|---|---|
| $r_{12}$, $r_{13}$, $r_{23}$ | Pearson correlations among x-variables |
| $b_{Y1.23}$, etc. | Usually a standardized regression coefficient ("beta weight") taken from the regression of Y on $X_1$, and "." means controlled for $X_2$, $X_3$ |
| $R_{Y.123}^2$ | Coefficient of determination $R^2$ from the regression of Y on $X_1$, $X_2$, $X_3$ |
| $\sqrt{1-R_{Y.123}^2}$ | Is an estimate of unmeasured influences called error term or disturbance |

Spring 2010 © Erling Berge 2010 21
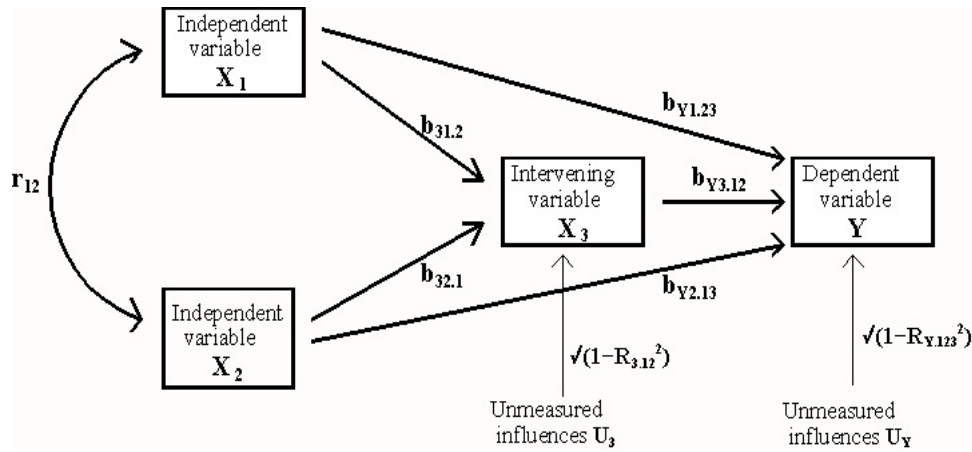
# Multiple regression

- All assumptions and all problems apply as before
  - Note in particular that error terms must be uncorrelated with included x-variables (all relevant variable are included)
- If some of the X-es are intervening in figure 2 the model is too simple, but it matters only if we are interested in causality

Spring 2010 © Erling Berge 2010 22

© Erling Berge 2010

11

# Path coefficients
# Figure 3

# New elements in figure 3

| | |
|---|---|
| $b_{31.2}$, $b_{32.1}$ | Standardized regression coefficients ("beta weight") from the regression of $X_3$ on $X_1$ controlled for $X_2$ and from the regression of $X_3$ on $X_2$ controlled for $X_1$ |
| $R_{3.12}^2$ | Coefficient of determination ($R^2$) from the regression of $X_3$ on $X_1$ and $X_2$ |
| $\sqrt{\{1-R_{3.12}^2\}}$ | The error term from the regression of $X_3$ on $X_1$ and $X_2$ |

# The structural model of figure 3

- $\hat{Y} = b_{Y1.23}X_1 + b_{Y2.13}X_2 + b_{Y3.12}X_3$
- $\hat{X}_3 = b_{31.2}X_1 + b_{32.1}X_2$

- In structural equations variables and coefficients are standardized
- That means that variables have an average of 0 and a standard deviation of 1 and that coefficients vary between -1 and +1

## Figure 5: the regression of $X_3$ on $X_1$ and $X_2$
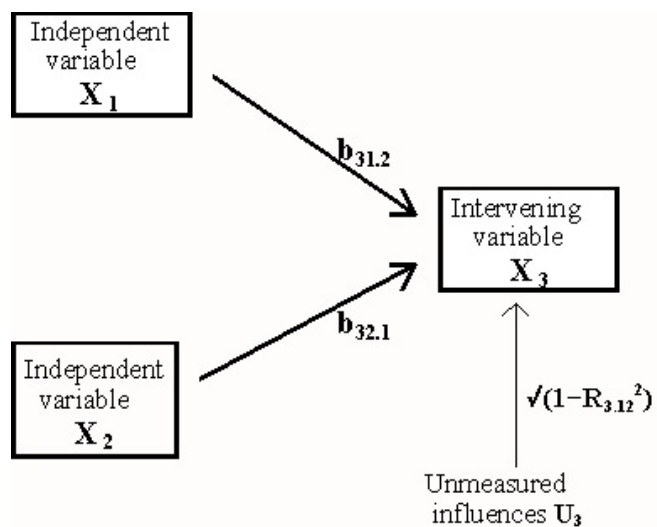
# Direct, Indirect and Total Effects

- *Indirect effects* equal the product of coefficients along any series of causal paths that link one variable to another
- *Total effects* equal the sum of all direct and indirect effects linking two variables

Indirect effects as products of path coefficients

- $\hat{Y} = b_{Y1.23}X_1 + b_{Y2.13}X_2 + b_{Y3.12}X_3$
- $\hat{X}_3 = b_{31.2}X_1 + b_{32.1}X_2$
- Means that we have
- $\hat{Y} = b_{Y1.23}X_1 + b_{Y2.13}X_2 + b_{Y3.12}X_3$
- $= b_{Y1.23}X_1 + b_{Y2.13}X_2 + b_{Y3.12}(b_{31.2}X_1 + b_{32.1}X_2)$
- $= b_{Y1.23}X_1 + b_{Y2.13}X_2 + b_{Y3.12}b_{31.2}X_1 + b_{Y3.12}b_{32.1}X_2$
- $= (b_{Y1.23} + b_{Y3.12}b_{31.2})X_1 + (b_{Y2.13} + b_{Y3.12}b_{32.1})X_2$

- Compare compound coefficients to the diagram

# Structural model
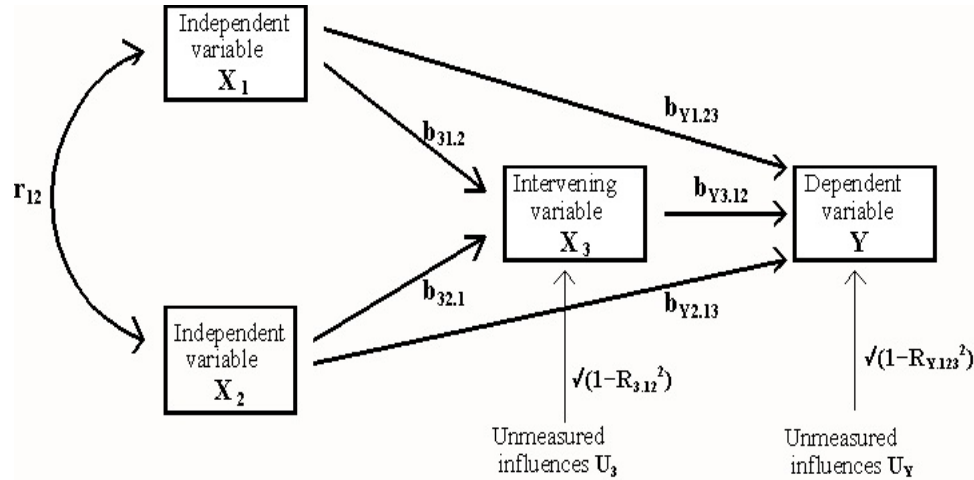
# Path Coefficients

- $X_1$ to Y: $\mathbf{b_{Y1.23}}$ (regression coefficient of Y on $X_1$, controlling for X2 and X3)
- $X_2$ to Y: $\mathbf{b_{Y2.13}}$ (regression coefficient of Y on $X_2$, controlling for $X_1$ and X3)
- $X_3$ to Y: $\mathbf{b_{Y3.12}}$ (regression coefficient of Y on $X_3$, controlling for $X_1$ and $X_2$)
- $X_1$ to $X_3$: $\mathbf{b_{31.2}}$ (regression coefficient of $X_3$ on $X_1$, controlling for $X_2$)
- $X_2$ to $X_3$: $\mathbf{b_{32.1}}$ (regression coefficient of $X_3$ on $X_2$, controlling for $X_1$)

# Direct effects

| $X_1$ to Y: $\mathbf{b_{Y1.23}}$ | regression coefficient of Y on X1, controlling for X2 and X3 |
|---|---|
| $X_2$ to Y: $\mathbf{b_{Y2.13}}$ | regression coefficient of Y on X2, controlling for X1 and X3 |
| $X_3$ to Y: $\mathbf{b_{Y3.12}}$ | regression coefficient of Y on X3, controlling for X1 and X2 |
| $X_1$ to $X_3$: $\mathbf{b_{31.2}}$ | regression coefficient of X3 on X1, controlling for X2 |
| $X_2$ to $X_3$: $\mathbf{b_{32.1}}$ | regression coefficient of X3 on X2, controlling for X1 |

Spring 2010                    © Erling Berge 2010                    31

# Indirect and total effects

| Indirect effects | |
|---|---|
| $X_1$ to Y, through $X_3$: | $\mathbf{b_{31.2} \times b_{Y3.12}}$ |
| $X_2$ to Y, through $X_3$: | $\mathbf{b_{32.1} \times b_{Y3.12}}$ |
| Total effects | |
| $X_1$ to Y: | $\mathbf{b_{Y1.23} + (b_{31.2} \times b_{Y3.12})}$ |
| $X_2$ to Y: | $\mathbf{b_{Y2.13} + (b_{32.1} \times b_{Y3.12})}$ |

Spring 2010                    © Erling Berge 2010                    32

# Additions to multiple regressions

- We learn something new if the indirect effects are large enough to have substantial interest
- More than two steps of causation tends to become very weak
  - 0.3*0.3*0.3 = 0.027
  - 0.3 standard deviation change in causal variables leads to a 0.027 standard deviation change in the dependent variable

Spring 2010 © Erling Berge 2010 33

# Variables and measurement

- All interval scale variables used in multiple regression (including non-linear transformed variables and interaction terms) can be included in structural equations models
- But interpretation becomes tricky when variables are complex. Conditional effect plots are very useful
- Robust, quantile, logit, and probit regression should not be used
- Categorical variables should not be used as intervening variables
- Scales or index variables can be used as usual in OLS regression

Spring 2010 © Erling Berge 2010 34

**Concluding on structural equations modelling**

- Including factors from factor analysis as explanatory variables make it possible to approximate a LISREL type analysis
- If assumptions are true LISREL will perform a much better and provides more comprehensive estimation, but too often assumptions are not true. Then the low-tech approach has access to the large toolkit of OLS regression for diagnostics and exploratory methods testing basic assumptions and discovering unusual data points
- Simple diagnostic work sometimes yields the most unexpected, interesting and replicable findings from our research

Spring 2010       © Erling Berge 2010       35

Principal components and factor analysis

- Principal components and factor analysis are both methods for data reduction
- They seek underlying dimensions that are able to account for the pattern of variation among a set of observed variables
- Principal components analysis is a transformation of the observed data where the idea is to explain as much as possible of the observed variation with a minimum number of components

Spring 2010       © Erling Berge 2010       36

# Factor analysis

- Estimates coefficients on - and variable values of - unobserved variables (Factors) to explain the co-variation among an observed set of variables
- The assumption is that a small set of the unobserved factors are able to explain most of the co-variation
- Hence factor analysis can be used for data reduction. Many variables can be replaced by a few factors

Spring 2010 © Erling Berge 2010 37

# Factor analysis

- $Z_k = \ell_{k1}F_1 + \ell_{k2}F_2 + \ldots + \ell_{kj}F_j + \ldots + \ell_{kJ}F_J + u_k$
  - k = 1, 2, 3, … , K
- Symbols
  - K observed variables, $Z_k$ ; k=1, 2, 3, … , K
  - J unobserved factors, $F_j$ ; j=1, 2, 3, … , J where J<K
  - For each variable there is a unique error term, $u_k$, also called unique factors while the F factors are called common factors
  - For each factor there is a **<u>standardized</u>** regression coefficient, $\ell_{kj}$, also called factor loading; k refers to variable no, j refers to factor no. An index denoting case no has been omitted here.

Spring 2010 © Erling Berge 2010 38

# Correlation of factors

- Factors my be correlated or uncorrelated
    - Uncorrelated: they are then called **orthogonal**
    - Correlated: they are then called **oblique**
- Factors may be rotated
    - Oblique rotations create correlated factors
    - Orthogonal rotations create uncorrelated factors

# Principal components

- Represents a simple transformation of variables. There are as many principal components as there are variables
- Principal components are uncorrelated

- $Z_k = \ell_{k1}F_1 + \ell_{k2}F_2 + \ldots + \ell_{kj}F_j + \ldots + \ell_{kK}F_K$
- If the last few principal components explain little variation we can retain J<K components. Thus Principal Components also can be used to reduce data.
- $Z_k = \ell_{k1}F_1 + \ell_{k2}F_2 + \ldots + \ell_{kj}F_j + \ldots + \ell_{kJ}F_J + v_k$

    where J<K and
    the residual $v_k$ has small variance and consist of the discarded principal components

## Principal components vs factor analysis

- Principal components analysis attempts to explain the observed variation of the variables
- Factor analysis attempts to explain their intercorrelations
- Use principal components to generate a composite variable that reproduce the maximum variance of observed variables
- Use factor analysis to model relationships between observed variables and unobserved latent variables and to obtain estimates of latent variable values
- The choice between the two is often blurred, to some degree it is a matter of taste

Spring 2010                              © Erling Berge 2010                              41

## The number of principal components

- K variables yield K principal components
- If the first few components account for most of the variation, we can concentrate on them and discard the remaining
- The eigenvalues of the standardized correlation matrix provides a guide here
- Components are ranked according to eigenvalues
- A principal component with an eigenvalue $\lambda<1$ accounts for less variance than a single variable
- Thus we discard components with eigenvalues below 1
- Another criterion for keeping components is that each component should have substantive meaning

Spring 2010                              © Erling Berge 2010                              42

## Eigenvalues and explained variance

- In a covariance matrix the sum of eigenvalues equals the sum of variances.
- In a correlation matrix this = K (the number of variables) since each standardized variable has a variance of 1
- Thus the sum of eigenvalues of the principal components
- $\lambda_1 + \lambda_2 + \lambda_3 + \ldots + \lambda_K = K$ and
- $\lambda_j / K$ = proportion of variance explained by component no j

Spring 2010                              © Erling Berge 2010                              43

# Uniqueness and communality

- If K-J components are discarded and we have only J factors
- $Z_k = \ell_{k1}F_1 + \ell_{k2}F_2 + \ldots + \ell_{kj}F_j + \ldots + \ell_{kJ}F_J + v_k$
- And an error term $v_k$
- The variance of the error term is called the uniqueness of the variable
- Communality is the proportion of a variable's variance shared with the components
- Communality = $h_k^2$ = 1 - Uniqueness = $\Sigma_j \lambda_{kj}^2$ , j=1,…, J ; k = variable number

Spring 2010                              © Erling Berge 2010                              44

# Rotation to simple structure

- The idea is to transform (rotate) the factors so that the loadings on each components make it easier to interpret the meaning of the component
- If the loading are close either to 1 or -1 on one factor and close to 0 on all others the structure is simpler to interpret: we rotate to "simple structure". The rotated factors fit data equally well but are simpler to interpret
- Rotations may be
  - Orthogonal  (rotation method typically: varimax)
  - Oblique      (rotation method typically: oblimin, promax)

Spring 2010                              © Erling Berge 2010                              45

# Why rotate?

- Underlying unobserved dimensions may in theory be seen as correlated
- Allowing correlated factors may provide even simpler structure than uncorrelated factors, thus easier to interpret
- All rotations fit data equally well
- Hence the one chosen depends on a series of choices done by the analyst
- Try different methods to see if results differ

Spring 2010                              © Erling Berge 2010                              46

# SPSS output

- For rotated factor solutions with correlated factors SPSS provides two matrixes for interpretation
- The pattern matrix provides the direct regression of the variables on the factors. The coefficients tells about the direct contribution of a factor in explaining the variance of a variable. Due to the correlations of the factors there are also indirect contributions
- The structure matrix provides the correlations between the variables and the factors

# Factor scores

- Both principal components and factor analysis may be used to compute composite scores called factor scores
- Recall that variables and factors are assumed to be related like
  - $Z_k = \ell_{k1}F_1 + \ell_{k2}F_2 + \ldots + \ell_{kj}F_j + \ldots + \ell_{kK}F_K$
- Then it is possible to find values $c_{ij}$ making
  - $\hat{F}_j = c_{1j}Z_1 + c_{2j}Z_2 + \ldots + c_{kj}Z_j + \ldots + c_{Kj}Z_K$
- The coefficients $c_{ij}$ are the factor score coefficients. They come from the regression of the factor $F_j$ on the variables

# Methods for extracting factors

- Principal factor analysis
  - The original correlation matrix **R** is replaced by **R\***
    where the original 1-values of the diagonal has been
    replaced by estimates of the communality (the shared
    variance)
  - The factors extracted tries to explain the co-variance
    or correlations among the variables.
  - The unexplained variance is attributed to a unique
    factor (error term). The uniqueness may reflect
    measurement error or something that this variable
    measure that no other variable measure
  - The most common estimate of communality is $R_k^2$ the
    coefficient of determination from the regression of $Z_k$
    on all other variables

Spring 2010　　　　　　　　　© Erling Berge 2010　　　　　　　　　49

# How many factor should we retain?

- In principal component analysis factors with
  eigenvalues above 1 is recommended
- In principal factor analysis factors with
  eigenvalues above 0 is recommended
- Procedure:
  - Extract initial factors or components
  - Rotate to simple structure
  - Decide on how many factors to retain
  - Obtain and use scores for the retained factors,
    ignoring discarded factors

Spring 2010　　　　　　　　　© Erling Berge 2010　　　　　　　　　50

# Concluding (1)

- Principal components
  - transformation of the data, not model based. Appropriate if goal is to compactly express most of the variance of k variables. Minor components (perhaps all except the first) may be discarded and viewed as a residual.

- Factor analysis
  - Estimates parameters of a measurement model with latent (unobserved) variables.

Spring 2010 © Erling Berge 2010 51

# Concluding (2)

- Types of factor analysis
  - Principal factoring – principal components of a modified correlation matrix R* in which communality estimates ($R_k^2$) replace "1" on the main diagonal
    - Principal factoring without iteration
    - Principal factoring with iteration
  - Maximum likelihood estimation – significance tests regarding number of factors and other hypotheses, **assuming multivariate normality**

Spring 2010 © Erling Berge 2010 52

# Concluding (3)

- Rotation
  - If we retain more than one factor rotation simplifies structure and improves interpretability
    - Orthogonal rotation (varimax) maximum polarization given uncorrelated factors
    - Oblique rotation (oblimin, promax) further polarization by permitting interfactor correlations. The results may be more interpretable and more realistic than uncorrelated factors
- Scores
  - Factor scores can be calculated for use in graphs and further analysis, based on rotated or unrotated factors and principal components

Spring 2010 © Erling Berge 2010 53

# Concluding (4)

- Factor analysis is based on correlations and hence as affected by non-linearities and influential cases as in regression
  - Use scatter plots to check for outliers and non-linearities
  - In maximum likelihood estimation this becomes even more important since it assumes multivariate normality making it even less robust than principal factors

Spring 2010 © Erling Berge 2010 54

# Principal components of trust in Malawi

- Survey of 283 households in 18 villages in Malawi, 2007
- There are 8 related questions asked in one group
- Are there 1, 2 or more underlying dimensions shaping the attitudes expressed?
- Analysis of correlations (not co-variances)
- The questions:

Spring 2010 © Erling Berge 2010 55

**M3 Would you say you trust all, most, some or just a few people in the following groups?** (All=1 – None=5)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| a | Your family members | All | Most | Some | Only a few | None | Do not know |
| b | Your relatives | All | Most | Some | Only a few | None | Do not know |
| c | Your village | All | Most | Some | Only a few | None | Do not know |
| d | People from outside the village | All | Most | Some | Only a few | None | Do not know |
| e | People of same ethnic group | All | Most | Some | Only a few | None | Do not know |
| f | People from outside ethnic group | All | Most | Some | Only a few | None | Do not know |
| g | People from same church/mosque | All | Most | Some | Only a few | None | Do not know |
| h | People *not* from same church/mosque | All | Most | Some | Only a few | None | Do not know |

Spring 2010 © Erling Berge 2010 56

# Trust in Malawi: descriptive

**Descriptive Statistics**

|  | Mean | Std. Deviation | Analysis N |
|---|---|---|---|
| M3.a. Trust in family members | 1.60 | .935 | 266 |
| M3.b. Trust in relatives | 2.12 | 1.136 | 266 |
| M3.c. Trust in people in own village | 2.69 | 1.090 | 266 |
| M3.d. Trust in people outside the village | 3.28 | 1.118 | 266 |
| M3.e. Trust in people of same ethnic group | 2.90 | 1.082 | 266 |
| M3.f. Trust in people outside ethnic group | 3.26 | 1.098 | 266 |
| M3.g. Trust in people from same church/mosque | 2.39 | 1.062 | 266 |
| M3.h. Trust in people not from same church/mosque | 3.02 | 1.197 | 266 |

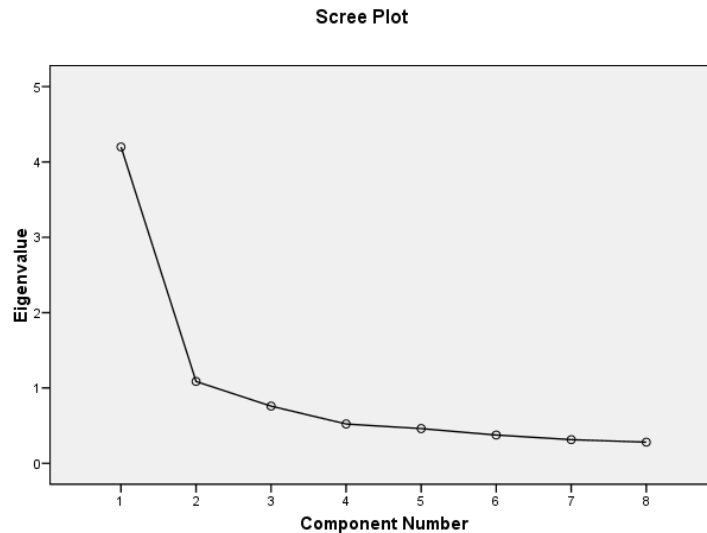# Trust in Malawi: correlation of variables

**Correlation Matrix**

|  | M3.a. Trust in family members | M3.b. Trust in relatives | M3.c. Trust in people in own village | M3.d. Trust in people outside the village | M3.e. Trust in people of same ethnic group | M3.f. Trust in people outside ethnic group | M3.g. Trust in people from same church/mosque | M3.h. Trust in people not from same church/mosque |
|---|---|---|---|---|---|---|---|---|
| M3.a. Trust in family members | 1.000 | .500 | .416 | .236 | .370 | .316 | .422 | .305 |
| M3.b. Trust in relatives | .500 | 1.000 | .496 | .315 | .363 | .353 | .424 | .292 |
| M3.c. Trust in people in own village | .416 | .496 | 1.000 | .482 | .588 | .573 | .465 | .430 |
| M3.d. Trust in people outside the village | .236 | .315 | .482 | 1.000 | .526 | .610 | .233 | .469 |
| M3.e. Trust in people of same ethnic group | .370 | .363 | .588 | .526 | 1.000 | .702 | .504 | .643 |
| M3.f. Trust in people outside ethnic group | .316 | .353 | .573 | .610 | .702 | 1.000 | .430 | .618 |
| M3.g. Trust in people from same church/mosque | .422 | .424 | .465 | .233 | .504 | .430 | 1.000 | .536 |
| M3.h. Trust in people not from same church/mosque | .305 | .292 | .430 | .469 | .643 | .618 | .536 | 1.000 |

# Trust in Malawi: number of factors

**Scree Plot**

# Trust in Malawi: factor/ component matrix

**Component Matrix** [a]

| | Component | |
|---|---|---|
| | 1 | 2 |
| M3.a. Trust in family members | .588 | .586 |
| M3.b. Trust in relatives | .624 | .532 |
| M3.c. Trust in people in own village | .776 | .080 |
| M3.d. Trust in people outside the village | .675 | -.398 |
| M3.e. Trust in people of same ethnic group | .832 | -.221 |
| M3.f. Trust in people outside ethnic group | .816 | -.330 |
| M3.g. Trust in people from same church/mosque | .690 | .265 |
| M3.h. Trust in people not from same church/mosque | .757 | -.262 |

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

# Trust in Malawi: loadings on orthogonal factors

| Rotated component matrix | Unrotated components | | Orthogonal varimax | |
|---|---|---|---|---|
| **Variables** | **F1** | **F2** | **F1** | **F2** |
| M3.a. Trust in family members | .588 | .586 | .117 | .821 |
| M3.b. Trust in relatives | .624 | .532 | .178 | .800 |
| M3.c. Trust in people in own village | .776 | .080 | .572 | .531 |
| M3.d. Trust in people outside the village | .675 | -.398 | .779 | .089 |
| M3.e. Trust in people of same ethnic group | .832 | -.221 | .798 | .324 |
| M3.f. Trust in people outside ethnic group | .816 | -.330 | .850 | .228 |
| M3.g. Trust in people from same church/mosque | .690 | .265 | .391 | .627 |
| M3.h. Trust in people not from same church/mosque | .757 | -.262 | .762 | .246 |

# Trust in Malawi: communalities

**Communalities**

| | Extraction |
|---|---|
| M3.a. Trust in family members | .689 |
| M3.b. Trust in relatives | .671 |
| M3.c. Trust in people in own village | .609 |
| M3.d. Trust in people outside the village | .614 |
| M3.e. Trust in people of same ethnic group | .741 |
| M3.f. Trust in people outside ethnic group | .774 |
| M3.g. Trust in people from same church/mosque | .546 |
| M3.h. Trust in people not from same church/mosque | .641 |

Extraction Method: Principal Component Analysis.

# Trust in Malawi: explained variance

**Total Variance Explained**

| Component | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 4.199 | 52.487 | 52.487 | 3.071 | 38.387 | 38.387 |
| 2 | 1.087 | 13.582 | 66.069 | 2.215 | 27.681 | 66.069 |

Extraction Method: Principal Component Analysis.

Spring 2010                    © Erling Berge 2010                    63

# Trust in Malawi: oblique factors pattern matrix

| Rotated component matrix | varimax (orthogonal) | | oblimin | | promax | |
|---|---|---|---|---|---|---|
| **Variables** | F1 | F2 | F1 | F2 | F1 | F2 |
| M3.a. Trust in family members | .117 | **.821** | -.087 | **.868** | -.145 | **.901** |
| M3.b. Trust in relatives | .178 | **.800** | -.014 | **.826** | -.067 | **.855** |
| M3.c. Trust in people in own village | **.572** | **.531** | **.493** | **.414** | **.476** | **.409** |
| M3.d. Trust in people outside the village | **.779** | .089 | **.838** | -.133 | **.864** | -.170 |
| M3.e. Trust in people of same ethnic group | **.798** | .324 | **.797** | .120 | **.806** | .093 |
| M3.f. Trust in people outside ethnic group | **.850** | .228 | **.881** | -.001 | **.899** | -.036 |
| M3.g. Trust in people from same church/mosque | .391 | **.627** | .268 | **.573** | .237 | **.582** |
| M3.h. Trust in people not from same church/mosque | **.762** | .246 | **.779** | .045 | **.792** | .016 |

Spring 2010                    © Erling Berge 2010                    64

## Trust in Malawi: oblique factors structure matrix

| Rotated component matrix | varimax | | oblimin | | promax | |
|---|---|---|---|---|---|---|
| **Variables** | **F1** | **F2** | **F1** | **F2** | **F1** | **F2** |
| M3.a. Trust in family members | .117 | .821 | .327 | .826 | .351 | .821 |
| M3.b. Trust in relatives | .178 | .800 | .380 | .819 | .403 | .817 |
| M3.c. Trust in people in own village | .572 | .531 | .690 | .649 | .702 | .671 |
| M3.d. Trust in people outside the village | .779 | .089 | .775 | .267 | .771 | .306 |
| M3.e. Trust in people of same ethnic group | .798 | .324 | .854 | .500 | .857 | .537 |
| M3.f. Trust in people outside ethnic group | .850 | .228 | .880 | .419 | .880 | .460 |
| M3.g. Trust in people from same church/mosque | .391 | .627 | .541 | .700 | .557 | .712 |
| M3.h. Trust in people not from same church/mosque | .762 | .246 | .800 | .416 | .801 | .452 |

## Trust in Malawi: correlation of components

### Component Correlation Matrix

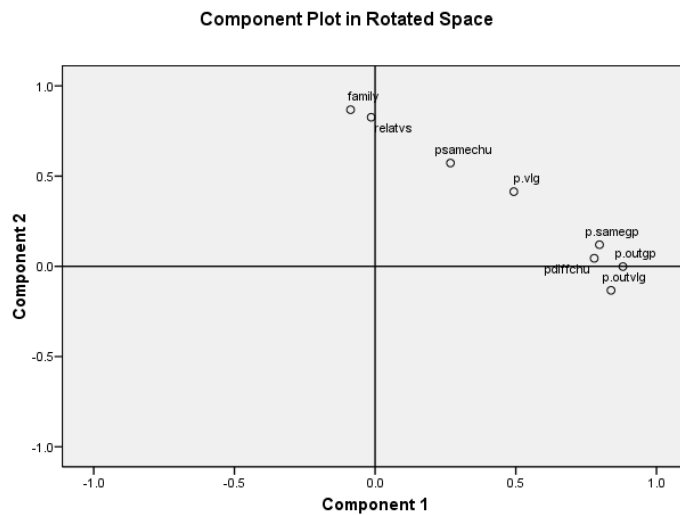| Component | 1 | 2 |
|---|---|---|
| 1 | 1.000 | .477 |
| 2 | .477 | 1.000 |

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
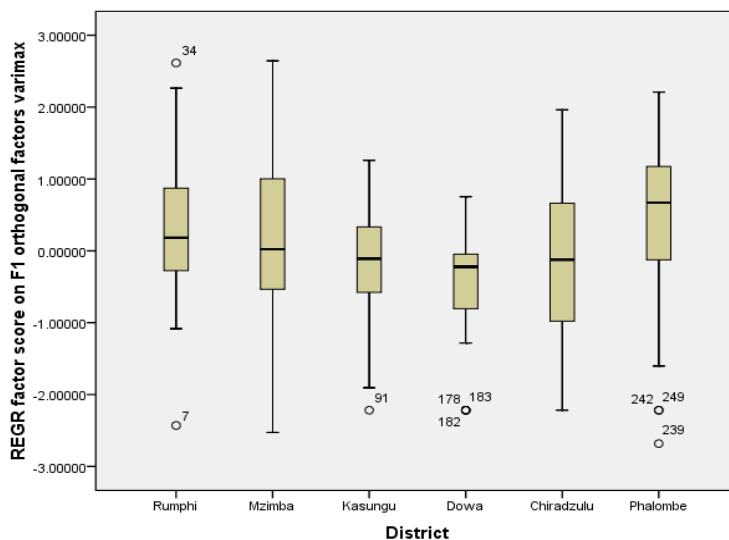
# Trust in Malawi: variables in component plot

**Component Plot in Rotated Space**

# Trust in Malawi: Orthogonal Factor 1 by district

# Trust in Malawi: Orthogonal Factor 2 by district
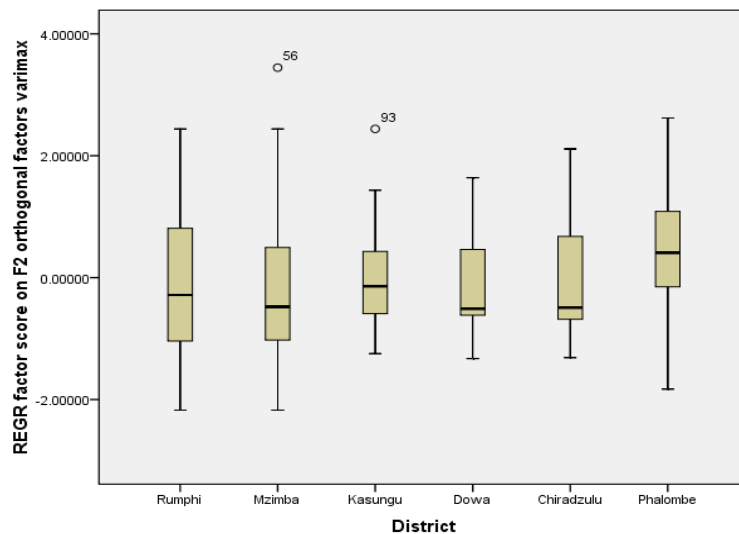
# Trust in Malawi: Orthogonal factors by district

**Case Processing Summary**

| | | Cases | | | | | |
| | | Valid | | Missing | | Total | |
| | District | N | Percent | N | Percent | N | Percent |
|---|---|---|---|---|---|---|---|
| REGR factor score on F1 orthogonal factors varimax | Rumphi | 43 | 95.6% | 2 | 4.4% | 45 | 100.0% |
| | Mzimba | 37 | 82.2% | 8 | 17.8% | 45 | 100.0% |
| | Kasungu | 47 | 95.9% | 2 | 4.1% | 49 | 100.0% |
| | Dowa | 49 | 98.0% | 1 | 2.0% | 50 | 100.0% |
| | Chiradzulu | 46 | 93.9% | 3 | 6.1% | 49 | 100.0% |
| | Phalombe | 44 | 97.8% | 1 | 2.2% | 45 | 100.0% |

**Case Processing Summary**

| | | Cases | | | | | |
| | | Valid | | Missing | | Total | |
| | District | N | Percent | N | Percent | N | Percent |
|---|---|---|---|---|---|---|---|
| REGR factor score on F2 orthogonal factors varimax | Rumphi | 43 | 95.6% | 2 | 4.4% | 45 | 100.0% |
| | Mzimba | 37 | 82.2% | 8 | 17.8% | 45 | 100.0% |
| | Kasungu | 47 | 95.9% | 2 | 4.1% | 49 | 100.0% |
| | Dowa | 49 | 98.0% | 1 | 2.0% | 50 | 100.0% |
| | Chiradzulu | 46 | 93.9% | 3 | 6.1% | 49 | 100.0% |
| | Phalombe | 44 | 97.8% | 1 | 2.2% | 45 | 100.0% |