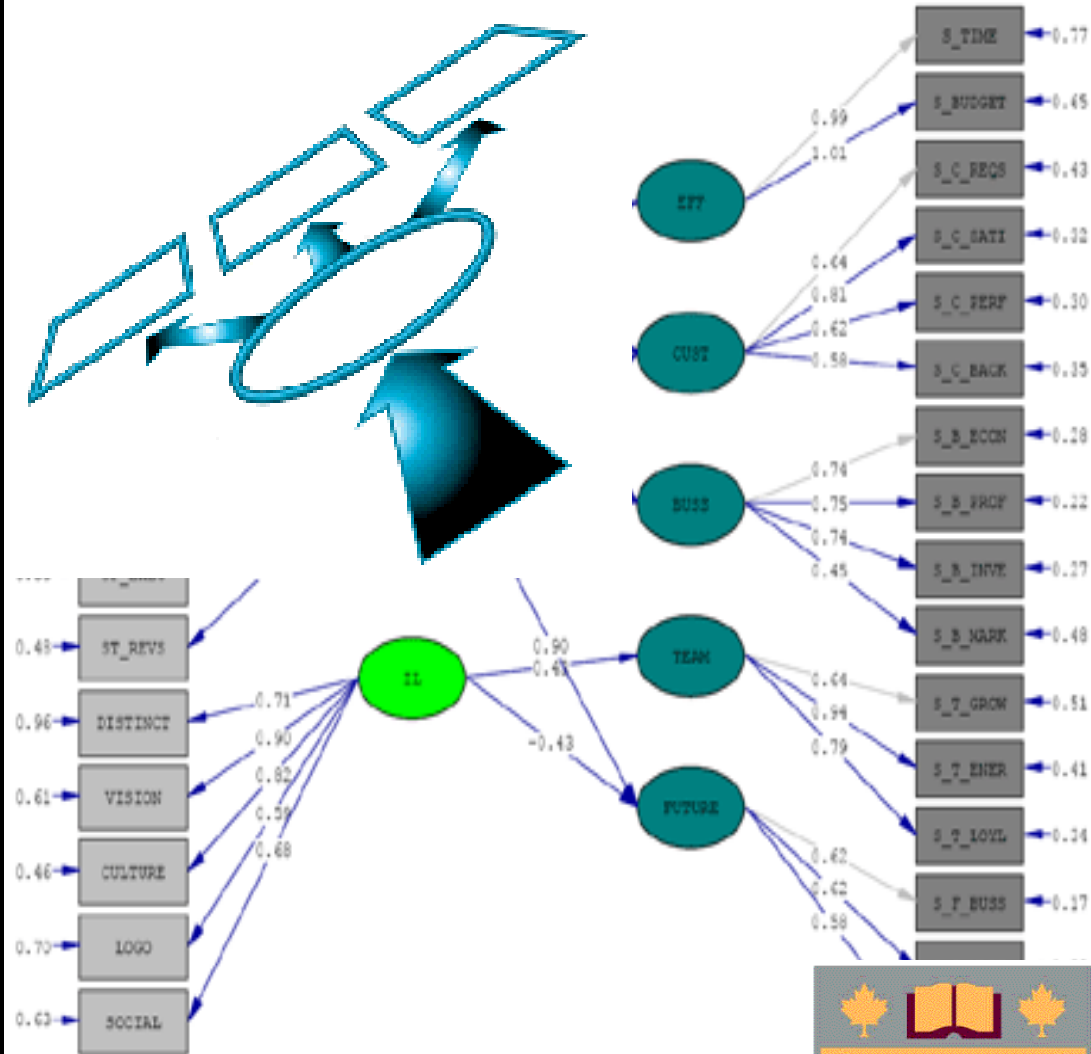


TSSEM

TWO-STAGE STRUCTURAL EQUATION MODELING

A META-ANALYTIC SEM TECHNIQUE



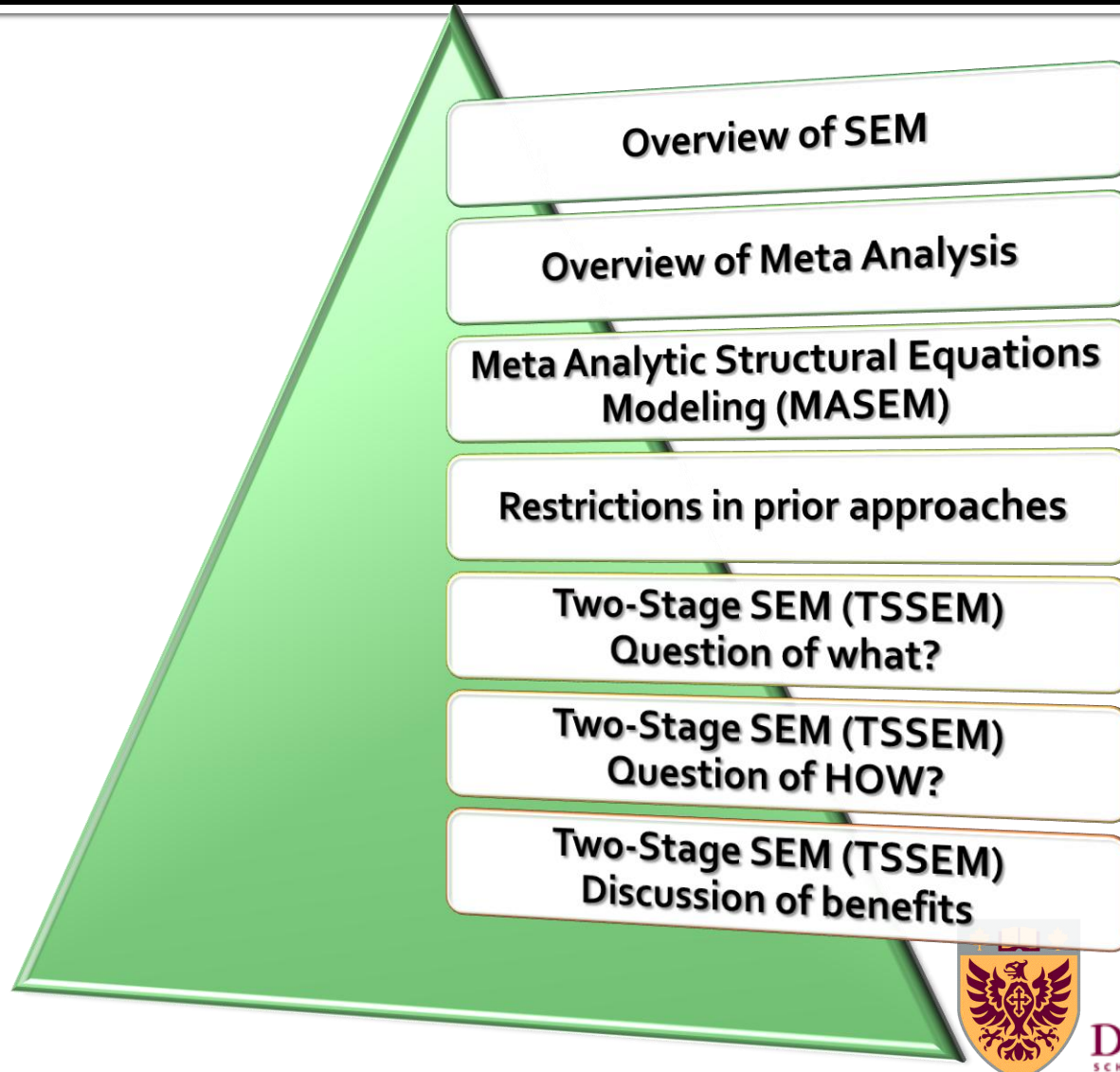
HAMED QAHHRI-SAREMI

DDSA Lunch & Learn Program
February 2010.



DeGroote
SCHOOL OF BUSINESS

Agenda

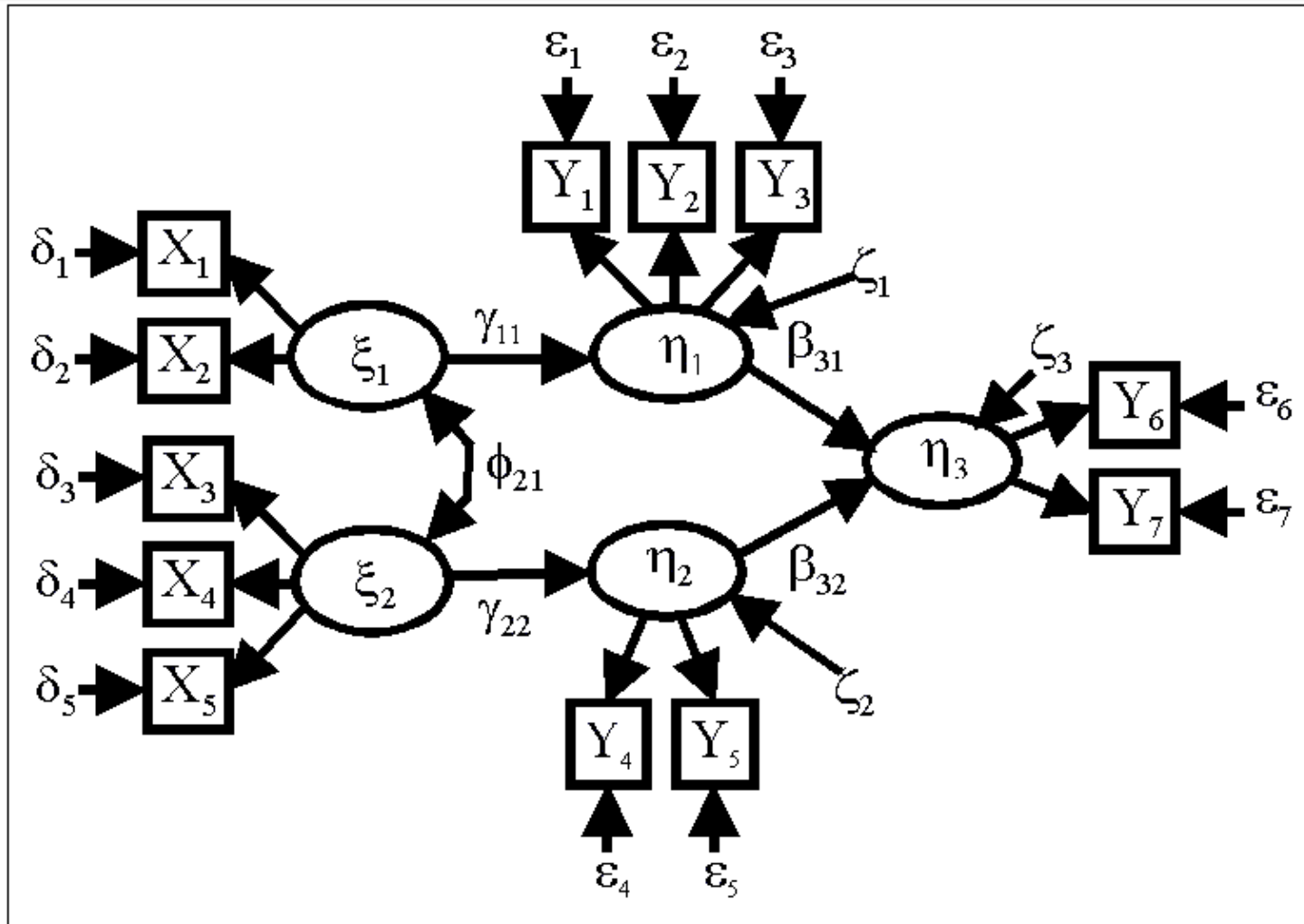


Overview of SEM

(Cheung & Chan, 2005)

- A frequently used multivariate technique for testing hypothetical models (common in behavioral sciences).
- Is a confirmatory technique
 - Tests models that are conceptually derived, a priori
 - Tests if the theory fits the data
- Combination of factor analysis and multiple regression
 - Can simultaneously test measurement and structural relationships
- A simple review of the literature shows its increase in popularity as a research tool

Overview of SEM



Overview of Meta Analysis

- Coined by **Glass (1976)**:
 - “the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings”.
- It is widely used as a statistical technique to synthesize research findings in social sciences.

Meta Analytic Structural Equations Modeling (MASEM)

(Viswesvaran & Ones, 1995)

- Combination of Meta Analysis and SEM techniques in synthesizing the prior studies findings in testing structural models.
- Involving the techniques of synthesizing matrices of effect sizes (e.g. correlation matrices) and fitting SEM

Meta Analytic Structural Equations Modeling (MASEM)

(Cheung & Chan, 2005)

- Stage 1:
 - Test of composite hypothesis about the homogeneity of correlation matrices across k studies
 - In case of heterogeneity:
 - Categorical Moderators
 - Random-effect assumption and averaging the correlations
 - Calculating pooled correlation matrix
 - List-wise / pair-wise deletion



Prior Approaches to Meta Analytic SEM (MASEM) – Stage 1

(Hunter & Schmidt , 1990, 2004)

- Univariate Approach
 - Based on the assumption of independency in elements of pooled correlation matrices.

(Hedges & Olkin, 1985)

- Multivariate Approach
 - Using multivariate approaches to estimate the pooled effect size (e.g. pooled correlation matrix) and the weight matrix from independent studies at the first stage.

Meta Analytic Structural Equations Modeling (MASEM)

(Cheung & Chan, 2005)

- Stage 2:
 - Using the “pooled Correlation matrix” as the “observed covariance matrix” in fitting SEM.
 - Using CFA or OVPA models to fit the model to the pooled correlation matrix.

Restrictions in prior approaches

(Cheung & Chan, 2005)

- Appropriate sample size for fitting SEM?
 - When pair-wise deletion is involved
 - Using ad-hoc solutions like: harmonic mean, median, the total, and etc.
- **BUT**, type-I error of Chi-square test statistics, goodness-of-fit indices, the statistical power, and the standard errors of parameters estimate are dependent on the **SAMPLE SIZE** used for fitting SEM.

Restrictions in prior approaches

- Possibility of Non-positive definite matrix as the input matrix because of pair-wise deletion
 - Non-positive definite= negative values for matrix determinant
(Cheung & Chan, 2005)
 - The determinant of a covariance or correlation matrix is a measure of "generalized variance."
(Wothke, 1993)
 - Since negative variances are undefined, and since zero variances apply only to constants, it is troubling when a covariance or correlation matrix fails to have a positive determinant.

Restrictions in prior approaches

(Cheung & Chan, 2005)

- Ignoring between-studies variations (second-order sampling error)
 - Sampling variation is reflected in the standard errors of pooled correlations
 - Sampling variation is well addressed in the first stage (i.e. different ns), BUT totally missed in the second stage (i.e. ignoring the effect of k).

Restrictions in prior approaches

(Cheung & Chan, 2005)

- The assumption of independency of correlations does NOT hold in many cases
 - Need to take into account the covariances/correlations of correlations (i.e. covariation among correlations).
 - Totally missed in univariate approaches

Restrictions in prior approaches

- Analyzing a correlation matrix instead of a covariance matrix
(Cheung & Chan, 2005)
 - It's generally incorrect to analyze correlation matrix instead of covariance matrix in SEM.
 - Lead to incorrect chi-square statistics and the standard errors of parameters estimates.

(Cudeck, 1989)

Two-Stage SEM (TSSEM)

Question of what?

- TSSEM is a two-stage approach for conducting MASEM on correlations/covariance matrices (Cheung & Chan, 2005; Cheung & Chan, 2009)
 - To incorporate meta-analytic techniques and SEM into a unified framework.
 - Stage 1 (in SEM context):
 - Homogeneity test of correlation/covariance matrices across studies
 - Calculating pooled correlation/covariance matrix
 - Calculating an asymptotic covariance matrix (ACM) of pooled correlation/covariance matrix
 - Stage 2:
 - Fit the model to the pooled matrices using SEM

Two-Stage SEM (TSSEM)

Question of what?

(Cheung & Chan, 2005, 2008, 2009)

- Stage 1 of TSSEM :
 - Using multiple-group CFA approach to test the homogeneity of correlation matrices across studies
 - Testing the chi-square difference of unconstrained vs. constrained models
 - Calculate the pooled correlation/covariance matrices
 - But in the multivariate sense, with no assumption of independency of correlations
 - Output is a (pxp) pooled correlation/covariance matrix, where p is the number of variables

Two-Stage SEM (TSSEM)

Question of what?

- Stage 1 of TSSEM : (Cheung & Chan, 2005, 2008, 2009)
 - Calculating an asymptotic covariance matrix (ACM) as a weight matrix for pooled correlation matrix in the second stage
 - ACM is a $\frac{p(p+1)}{2} \times \frac{p(p+1)}{2}$ matrix of variances (diagonal elements) and covariances (off-diagonal elements) of pooled correlations/covariances.
 - ACM reflects:
 - Inter-dependency of pooled correlations/covariances
 - Second-order (between studies) variations
 - Second-order effect of sample size in stage 2

Two-Stage SEM (TSSEM)

Question of what?

- Stage 1 of TSSEM : (Cheung & Chan, 2005, 2008, 2009)
 - TSSEM use LISREL to do the homogeneity test (i.e. a back & forth process between TSSEM software and LISREL)
 - We have fit-indices in addition to chi-square index for the result of homogeneity test
 - Chi-square gets significant in cases with very large N (i.e. the case in many MAs)
 - If the homogeneity hypothesis of correlations/covariances is NOT rejected, we can proceed to second stage
 - In case of rejection, like post-hoc analysis, the output of LISREL will give the fit indices for every study in the MA

Two-Stage SEM (TSSEM)

Question of what?

- Stage 2 of TSSEM: (Cheung & Chan, 2005, 2008, 2009)
 - Pooled correlation matrix and its asymptotic covariance matrix (i.e. as the weight matrix) are used as the input to SEM
 - The total sample size of all studies is used as the input sample size for the SEM
 - The only SEM package currently work with TSSEM is LISREL.

Two-Stage SEM (TSSEM)

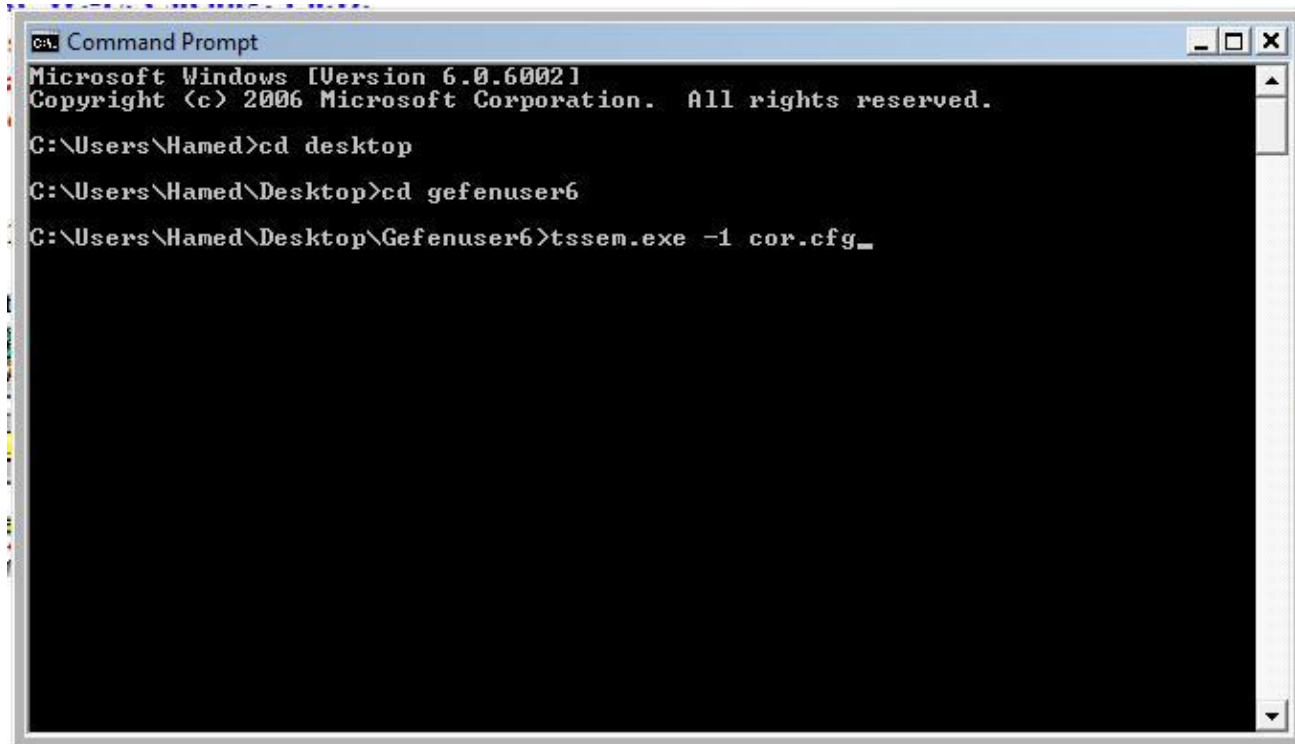
Question of HOW?

(Cheung, 2009)

- Download the free TSSEM software from:

<http://courses.nus.edu.sg/course/psycwlm/internet/tssem.zip>

- Don't be surprised it is a DOS-driven software.



```
Command Prompt
Microsoft Windows [Version 6.0.6002]
Copyright (c) 2006 Microsoft Corporation. All rights reserved.

C:\Users\Hamed>cd desktop
C:\Users\Hamed\Desktop>cd gefenuser6
C:\Users\Hamed\Desktop\Gefenuser6>tssem.exe -1 cor.cfg_
```

Two-Stage SEM (TSSEM)

Question of HOW?

- No installation, just copy TSSEM.EXE and two other files (i.e. cor.cfg & cor.dat) to your folder.
- TSSEM.EXE generates required LISREL syntax files and processes the intermediate data files in the TSSEM procedure.

(Cheung, 2009)

Two-Stage SEM (TSSEM)

Question of HOW?

(Cheung, 2009)

- Steps of analysis:
 1. Prepare your data (i.e. consistent correlation matrices, sample size in each study, and missing variables in each study)

Preparing data in spreadsheet

[illegible]

Two-Stage SEM (TSSEM)

Question of HOW?

- Steps of analysis:

1. Configure cor.dat and cor.cfg (this file is the main file which is used till the end of the process) based on prepared data
 1. Copy correlation matrices to cor.dat
 2. Copy all other information (Study number, sample size, and missing variables of the studies in the same order with correlation matrices) to cor.cfg
 3. Complete other required syntax for the first stage in cor.cfg based on TSSEM guideline (just copy and paste)
 4. Complete the bottom part of the syntax in cor.cfg for the second stage based on guidelines in TSSEM package (these are LISREL syntax for model and data specification and will be used when you reach the second stage)

Preparing cor.dat data file

```

1.0000
0.7370 1.0000
0.5970 0.6940 1.0000
0.4868 0.6598 0.4973 1.0000
0.4390 0.5180 0.4580 0.4159 1.0000
0.5729 0.8273 0.4988 0.6200 0.1153 1.0000
0.4830 0.5580 0.5530 0.3723 0.5600 0.5456 1.0000
0.4460 0.5160 0.4990 0.3956 0.2290 0.3160 0.4500 1.0000
0.4510 0.6699 0.5973 0.4100 0.0810 0.5139 0.4876 0.2844 1.0000
0.4000 0.7000 0.1700 0.6000 0.1579 0.3880 0.5050 0.4913 0.6110 1.0000
0.1100 0.4273 0.2784 0.6190 0.0757 0.3053 0.3529 0.2570 0.4170 0.3660 1.0000
1.0000
0.5000 1.0000
0.0000 0.0000 1.0000
0.5000 0.7000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
0.2000 0.4000 0.0000 0.5000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
0.4000 0.7000 0.0000 0.6000 0.0000 0.0000 0.0000 0.5000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
1.0000
0.5300 1.0000
0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000
1.0000

```

```
# This is the Stage 1 specification
```

```
AN=1
```

```
# AN=1: Analysis of correlation matrices
```

```
# AN=2: Analysis of covariance matrices
```

```
# AN=3: Testing the homogeneity of variances.
```

```
NG=41
```

```
NI=11
```

```
CM SY FI=cor.dat
```

```
|
```

```
GP=1
```

```
NO=523
```

```
GP=2
```

```
NO=1004
```

```
MV=3 5 6 7 9 11
```

```
GP=3
```

```
NO=122
```

```
MV=3 4 5 6 7 8 9 10 11
```

```
.
```

```
.
```

```
.
```

```
GP=41
```

```
NO=86
```

```
MV=2 3 4 6 7 8 9 11
```

```
# This is the Stage 2 model specification
```

```
% LA
```

```
% 'Intent' 'PU' 'PEOU' 'Trust' 'SN' 'Compat' 'Effic' 'Secur' 'Conven' 'Involv' 'Tech'
```

```
%MO NX=7 NY=4 PH=SY,FR PS=DI,FR BE=FU,FI GA=FU,FI
```

```
%fr be(1,2) be(1,3) be(1,4) be(2,3) be(2,4) be(4,3)
```

```
%fr ga(3,2) ga(3,3) ga(3,5) ga(3,6) ga(4,1) ga(4,2) ga(4,4) ga(4,5) ga(4,6) ga(4,7)
```

```
# End of input syntax
```

Preparing cor.cfg syntax file

Two-Stage SEM (TSSEM)

Question of HOW?

- Stage 1: (Cheung & Chan, 2007)
 - Run "TSSEM.EXE -1 COR.CFG" in your DOS command window. This produces cor1.ls8 file which is a LISREL syntax for test of heterogeneity and calculation of pooled correlation/covariance matrix as well as ACM.

! Generated by TSSEM, Mike Cheung (2007)

! Analysis of correlation matrices

TI TSSEM Stage 1: Group 1

DA NG=41 NI=11 NO=523

CM SY FI=cor.dat

MO NX=11 NK=11 LX=DI,FR TD=ZE PH=ST,FR

OU

TI TSSEM Stage 1: Group 2

DA NI=11 NO=1004

CM SY FI=cor.dat

MO NX=11 NK=11 LX=DI,FR TD=ZE PH=ST,FR

OU

TI TSSEM Stage 1: Group 3

DA NI=11 NO=122

CM SY FI=cor.dat

MO NX=11 NK=11 LX=DI,FR TD=ZE PH=ST,FR

OU

TI TSSEM Stage 1: Group 4

DA NI=11 NO=240

CM SY FI=cor.dat

MO NX=11 NK=11 LX=DI,FR TD=ZE PH=ST,FR

OU

TI TSSEM Stage 1: Group 5

**View of cor1.ls8 syntax file
created for LISREL**



Two-Stage SEM (TSSEM)

Question of HOW?

- Then run the LISREL on cor1.ls8 and see the goodness of fit results in cor1.out for the homogeneity test.

Global Goodness of Fit Statistics

Degrees of Freedom = 216

Minimum Fit Function Chi-Square = 2533.60 (P = 0.0)

Normal Theory Weighted Least Squares Chi-Square = 2206.94 (P = 0.0)

Estimated Non-centrality Parameter (NCP) = 1990.94

90 Percent Confidence Interval for NCP = (1843.87 ; 2145.41)

Minimum Fit Function Value = 0.19

Population Discrepancy Function Value (F0) = 0.15

90 Percent Confidence Interval for F0 = (0.14 ; 0.16)

Root Mean Square Error of Approximation (RMSEA) = 0.17

90 Percent Confidence Interval for RMSEA = (0.16 ; 0.17)

P-Value for Test of Close Fit (RMSEA < 0.05) = 0.00

Expected Cross-Validation Index (ECVI) = 0.54

90 Percent Confidence Interval for ECVI = (0.53 ; 0.55)

ECVI for Saturated Model = 0.0099

ECVI for Independence Model = 1.26

Chi-Square for Independence Model with 2255 Degrees of Freedom = 16795.11

Independence AIC = 17697.11

Model AIC = 7186.94

Saturated AIC = 5412.00

Independence CAIC = 21531.56

Model CAIC = 28357.21

Saturated CAIC = 28418.72

Normed Fit Index (NFI) = 0.85

Non-Normed Fit Index (NNFI) = -0.66

Parsimony Normed Fit Index (PNFI) = 0.081

Comparative Fit Index (CFI) = 0.84

Incremental Fit Index (IFI) = 0.86

Relative Fit Index (RFI) = -0.57

**View of cor1.out output file
for the stage-1
(i.e. homogeneity test of
correlations)**

Two-Stage SEM (TSSEM)

Question of HOW?

- You results are:
 - Homogeneous, proceed to stage 2
 - $RMSEA \leq 0.08$,
 - $NFI \geq 0.9$,
 - $NNFI \geq 0.9$,
 - $CFI \geq 0.9$
 - Heterogeneous, stay and address the heterogeneity through:
 - Moderator testing (e.g. cluster analysis)
 - Justifying the assumption of random effect

Two-Stage SEM (TSSEM)

Question of HOW?

(Cheung & Chan, 2007)

- Stage 2:
 - Run "TSSEM.EXE -2 COR.CFG" in your DOS command window. This produces cor2.ls8 file which is a LISREL syntax for test of SEM based on pooled correlation/covariance matrix, ACM, and N.

Two-Stage SEM (TSSEM)

Question of HOW?

cor2.ls8 - Notepad

File Edit Format View Help

! Generated by TSSEM, Mike Cheung (2007)

! Analysis of correlation matrices

TI TSSEM Stage 2

DA NI=11 NO=13388 MA=KM

KM=cor2.cor SY

AC=cor2.ack SY

LA

'Intent' 'PU' 'PEOU' 'Trust' 'SN' 'Compat' 'Effic' 'Secur' 'Conven' 'Involv' 'Tech'

MO NX=7 NY=4 PH=SY,FR PS=DI,FR BE=FU,FI GA=FU,FI

fr be(1,2) be(1,3) be(1,4) be(2,3) be(2,4) be(4,3)

fr ga(3,2) ga(3,3) ga(3,5) ga(3,6) ga(4,1) ga(4,2) ga(4,4) ga(4,5) ga(4,6) ga(4,7)

OU ME=WL

**View of cor2.ls8 syntax file
for the stage-2 analysis in LISREL**

Two-Stage SEM (TSSEM)

Question of HOW?

- Then run the LISREL on cor2.ls8 and see the goodness of fit results in cor2.out for the SEM analysis.

Goodness of Fit Statistics

Degrees of Freedom = 18

Minimum Fit Function Chi-Square = 603.56 ($P = 0.0$)

Estimated Non-centrality Parameter (NCP) = 585.56

90 Percent Confidence Interval for NCP = (509.01 ; 669.52)

Minimum Fit Function Value = 0.045

Population Discrepancy Function Value ($F0$) = 0.044

90 Percent Confidence Interval for $F0$ = (0.038 ; 0.050)

Root Mean Square Error of Approximation (RMSEA) = 0.049

90 Percent Confidence Interval for RMSEA = (0.046 ; 0.053)

P-Value for Test of Close Fit ($RMSEA < 0.05$) = 0.62

Expected Cross-Validation Index (ECVI) = 0.052

90 Percent Confidence Interval for ECVI = (0.047 ; 0.059)

ECVI for Saturated Model = 0.0099

ECVI for Independence Model = 1.28

Chi-Square for Independence Model with 55 Degrees of Freedom = 17102.93

Independence AIC = 17124.93

Model AIC = 699.56

Saturated AIC = 132.00

Independence CAIC = 17218.45

Model CAIC = 1107.66

Saturated CAIC = 693.14

Normed Fit Index (NFI) = 0.96

Non-Normed Fit Index (NNFI) = 0.90

Parsimony Normed Fit Index (PNFI) = 0.32

Comparative Fit Index (CFI) = 0.97

Incremental Fit Index (IFI) = 0.97

Relative Fit Index (RFI) = 0.89

**View of cor2.out output file
for the stage-2
(i.e. SEM analysis)**

Two-Stage SEM (TSSEM)

Discussion of benefits

- TSSEM incorporates the precision of estimation in stage 1 to the stage 2
 - Through weighting the pooled correlations/covariances by their variance and covariances which are calculated based on their sample size (both k and n parameters involved).
 - Therefore, when using N as sum of all n s in SEM, no inflation in Chi-square and fit-indices happen.

(Cheung & Chan, 2005, 2008, 2009)

Two-Stage SEM (TSSEM)

Discussion of benefits

(Cheung & Chan, 2005, 2008, 2009)

- TSSEM provides the information about second-order (between-studies) variation in the pooled correlation/covariance matrix through ACM as its weight matrix
 - If there's much (little) sampling variation in the pooled correlation/covariance, the asymptotic variance of the pooled correlation/covariance that is a diagonal element in ACM will also be larger(smaller).
 - When the ACM as the weight matrix is inverted, more (less) weight is given to the pooled correlation/covariance with less (more) variation

Two-Stage SEM (TSSEM)

Discussion of benefits

- TSSEM utilizes more information in analyzing the fit for pooled correlation/covariance matrices through utilizing their covariances as their weights.

(Cheung & Chan, 2005, 2008, 2009)

Two-Stage SEM (TSSEM)

Discussion of benefits

- TSSEM finally uses a covariance matrix in the stage 2 (i.e. through multiplying pooled correlation/covariance matrix by ACM as the weight matrix)
 - This responds to the concern of inputting pooled correlation instead of an observed covariance matrix in SEM
 - No inaccurate chi-square or fit-indices produced

(Cheung & Chan, 2005, 2008, 2009)

References:

- Cheung, M. W. L. (2009). TSSEM: A LISREL syntax generator for two-stage structural equation modeling (Version 1.11) [Computer software and manual].
- Cheung, M. W. L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods, 10*(1), 40-64.
- Cheung, M. W. L., & Chan, W. (2009). A two-stage approach to synthesizing covariance matrices in meta-analytic structural equation modeling. *Structural Equation Modeling, 16*(1), 28-53.
- Cheung, M. W. L. (2008). A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling. *Psychological Methods, 13*(3), 182-202.
- Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105*, 317-327.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3-8.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*: Sage Publications.
- Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining META-ANALYTIC STRUCTURAL EQUATION MODELING 63 psychometric meta-analysis and structural equations modeling. *Personnel Psychology, 48*, 865-885.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for metaanalysis*. Orlando, FL: Academic Press.
- Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 256-93). Newbury Park, CA: Sage.

