Factor Analysis Models as Approximations

Robert C. MacCallum
University of North Carolina – Chapel Hill

Presented at conference on Factor Analysis at 100
Chapel Hill, NC – May, 2004

General Principle:

A factor analysis model is not an exact representation of real-world phenomena.

Always wrong to some degree, even in population.

- Nonlinearity
- Effects of minor factors
- Many other sources of error

At best, model is an approximation of real world.

This is the nature of scientific models in general.

“ . . . no model is completely faithful to the behavior under study. Models usually are formalizations of processes that are extremely complex. It is a mistake to ignore either their limitations or their artificiality. The best one can hope for is that some aspect of a model may be useful for description, prediction, or synthesis.”

Cudeck & Henly (1991)

View is now widely understood and accepted.

Focus on two aspects of this principle:

1) Evolution of this idea in the context of FA.

- Examine how view of relationship between model and real world has changed.
- Start with Spearman (1904).
- Consider work by Thomson, Thurstone, Tucker, and others.

2) Implications for parameter estimation.

- Consequences for performance of different estimation methods.
- Will demonstrate and explain potentially important effects.

Part I: Historical Perspective

Principle seems self-evident now.
Was not part of mindset in 1904.

Formal expression of models, statistical theory came much later.
Recognition of approximate nature of models.

Consider older writings from modern perspective.

Examine how early investigators viewed relationship between model and real world.

How this view has evolved since Spearman (1904).

Two Disclaimers

- A selective review, not exhaustive.
  Focus on major contributors, classic papers.

- Review is interpretive in some places.
  Authors not explicit.
  Reader can make inferences.
  Alternative interpretations may be reasonable.

Spearman (1904)

Did not specify a formal FA model.

Trying to support a substantive theory.

Observation of positive correlations among tests taken as supporting evidence for *g*.

Performance on a mental test determined by two factors: *g* and *s*

Sought to show *g* alone accounted for correlations among various measures of ability.

Analysis of earlier studies showed inconsistent patterns of correlations, no consistent support for theory.

Spearman attributed failures to weak methods:

- Design
- Sample
- Measures
- Analyses

Believed that if studies were done properly, results would support his theory.

Spearman conducted series of studies.
Great care with respect to design, sample, measures.

Statistical analyses included use of partial correlations, correction for attenuation.

Results showed corrected correlations consistently around 1.00.

Spearman viewed this as strong support for theory.

Also observed hierarchical pattern of correlations in correlation matrix: Evidence for *g*, differential saturation of tests by *g*.

Some strongly stated conclusions:

"On the whole then, we reach the profoundly important conclusion that there really exists a something that we may provisionally term "General Sensory Discrimination" and similarly a "Genral Intelligence," and further that the functional correspondence between these two is not appreciably less than absolute."

p. 272

“Whenever branches of intellectual activity are at all dissimilar, then their correlations with one another appear wholly due to their being all variously saturated with some common fundamental Function (or group of Functions).”

p. 273

A clear theme in Spearman (1904):

- There is a "truth" with regard to the structure of intelligence: Two-Factor Theory.

- It is accessible via sufficiently careful design, methods, analyses.

- Available evidence supports two-factor theory as being true.

Next 20-30 years: Refinement of criteria for verification of two-factor theory:

- Hierarchical structure of R.

- Intra-columnar correlation:

  Two-factor theory $\Rightarrow$ ICC = 1.0.

- Tetrad differences:

  Two factor theory $\Rightarrow$ tetrad differences = 0.

Numerous papers attempting to provide further support for Spearman’s theory using these criteria.

Hart & Spearman (1912)

Considered competing theories:

- Two-factor theory.
- Complex array of elementary functions.
- Group factors.

Implications of each model for pattern of elements in R: Different predictions about ICC.

A problem in model comparison / selection.

Obtained data from 14 empirical studies.

Applied inter-columnar correlation criterion:

Eliminated variables that were "too similar."

Corrected for bias due to sampling error.

Results: Mean ICC for each study ≈ 1.0.

Hart & Spearman conclusion:

The theory has now been "proved true." (p. 60)

Hart and Spearman finding seemed definitive.

Dodd (1928): "It seemed to be the most striking quantitative fact in the history of psychology."

But concerns and questions arose.
Empirical and methodological studies.

Godfrey Thomson – series of papers beginning 1916.

- Other theories produce same structure.
- Spearman's analytical methods suspect.

Thomson (1920): Spearman criteria can be satisfied as a result of a different underlying process and structure.

Thomson’s Sampling Theory of mental abilities.

Satisfaction of ICC and tetrad criteria not sufficient for validation of 2-factor model.

“The proof of the Theory of Two Factors which is based on the presence of hierarchical order therefore falls to the ground.”

Spearman’s selection and deletion of variables may invalidate support of theory.

Numerous supportive studies deleted tests that caused violations of criteria (Hart & Spearman, 1912; Brown & Stephenson, 1933)

Thomson (1920): “The Hart and Spearman criterion for the degree of perfection is erroneous and creates the extreme perfection it purports to detect.”

Spearman (1920): “ . . . as regards the fundamental theory, I venture to maintain that this has now been demonstrated with finality.”

Early years:

Debate about truth/falsity of two-factor theory.

Theory implies what correlations should look like (after corrections, etc.).

Spearman school: If characteristics of data can be shown to hold exactly, theory is proven.

Thomson school: Evaluation of criteria is flawed; theories not sufficiently differentiated by criteria.

Hazy distinction between theory and model.

Focus gradually shifted away from establishing truth/falsity of 2-factor theory.

Spearman gave ground as evidence accumulated.
Acknowledged need for group factors.
Only after *g*, and as few as possible.

Alternative theories proposed; imply different factor models.

Thomson wrote explicitly about finding solution that provides best approximation of data.

- Different types and degrees of approximation.
- Question of how many factors are required.

Thurstone on the model as an approximation:

"Our assumptions oversimplify the phenomena we are trying to comprehend."

" . . . performance can be expressed, in first approximation, as a linear function of the primaries."

"Perhaps the equation (1) should be very much more complex."

"While working with this simplifying assumption, we can expect to find the principal landmarks or dimensions of mind."

Thurstone:

- Was very explicit about model as approximation.
- Extended formal approach to expressing model.

Substantive theory:

- Two factor model usually inadequate.
- Need more factors; may include *g*.

Evaluating the model:

- Tetrad criterion too strong.
- Focus on residual correlations:

  No expectation of fully explaining data.

  Recognize many sources of lack of fit.

Lawley & Maxwell (1963)

Full representation of common factor model as a statistical model:

- Distribution theory
- Parameters with true population values
- Parameter estimates
- Statistical inference, model testing

Provided a basis for understanding how the model may be incorrect:

- Distributional assumptions
- Linearity
- Minor factors
- Individual differences in loadings

Bartholomew:
Critical advance provided by Lawley & Maxwell.
On Spearman's contributions: "Factor analysis was born before its time."

Notion of model as approximation gradually became part of mindset.

Reflected in numerous classic papers:

Jöreskog (1969): Comments on likelihood ratio test of model fit.

Tucker & Lewis (1973): Development of TLI based on differentiating sampling error, model error.

Tucker, Koopman, & Linn (1969): Simulation of correlation matrices incorporating effects of minor factors.

More extensive study and representation of approximate nature of models:

MacCallum & Tucker (1991):
Analysis of sources of error in common factor model; implications for theory and practice.

Cudeck & Henly (1991):
Framework for sources of error in covariance structure models; implications for model comparison.

The way we now think about models in general.

Box (1979): "Models, of course, are never true, but fortunately it is only necessary that they be useful. For this it is usually needful only that they not be grossly wrong."

Tukey (1961): "In a single sentence, the moral is: Admit that complexity always increases, first from the model you fit to the data, thence to the model you use to think and plan about the experiment and its analysis, and thence to the true situation."

Factor analysis models are of this nature.

- Approximations to real-world phenomena.
- Always wrong to some degree, in variety of ways.
- May still be useful if not grossly wrong.

This perspective has evolved over the past 100 years.

Part II: Implications for theory and practice.

There are many.

Focus on one:
Implications regarding parameter estimation.

Will show how the presence and nature of error in the model has direct and strong implications for performance of different estimation methods.

Central issue: The nature of error.

Data contains error of various types:

- Sampling error (random, nonrandom).
- Model error from various sources.

Estimation methods make assumptions about error.

- ML: All error is normal theory random sampling error; no model error.

When using a particular estimation method, consider issue of correspondence between assumptions about error and nature of error.

Does it matter?

Example

Illustration presented by MacCallum & Tucker (1991).

Using a modified version of methods used by Tucker, Koopman and Linn (1969), a population correlation matrix was generated that included model error.

- 12 MVs, 3 common factors in major domain.
- Influences of 50 minor factors simulated to represent model error.

Note: No sampling error is present. All error is model error.

Data Generating Parameters:

	Major Domain Factors			Uniqueness	Minor Variance
	1	2	3		
1	**.95**	0	0	.000	.098
2	**.95**	0	0	.000	.098
3	**.95**	0	0	.000	.098
4	**.95**	0	0	.000	.098
5	**.95**	0	0	.000	.098
6	0	**.70**	0	.413	.098
7	0	**.70**	0	.413	.098
8	0	**.70**	0	.413	.098
9	0	**.70**	0	.413	.098
10	0	0	**.50**	.653	.098
11	0	0	**.50**	.653	.098
12	0	0	**.50**	.653	.098

Simulated Population Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1	**1.00**											
2	**.94**	**1.00**										
3	**.87**	**.88**	**1.00**									
4	**.89**	**.90**	**.95**	**1.00**								
5	**.96**	**.94**	**.89**	**.86**	**1.00**							
6	-.01	-.01	.06	.08	-.04	**1.00**						
7	-.06	-.06	.06	.03	-.06	**.53**	**1.00**					
8	.00	-.06	-.04	.02	-.06	**.49**	**.52**	**1.00**				
9	-.01	.05	-.02	-.06	.04	**.45**	**.45**	**.42**	**1.00**			
10	.06	.07	-.02	.02	.05	.02	-.06	-.04	.02	**1.00**		
11	.04	.05	-.05	-.06	.07	-.05	-.05	-.04	.05	**.29**	**1.00**	
12	.01	-.06	-.02	-.07	.04	-.06	.00	.02	.00	**.21**	**.27**	**1.00**

Population correlation matrix was factor analyzed by OLS and ML methods, retaining 3 factors.

	OLS Solution				**ML Solution**			
	Factor	Factor	Factor		Factor	Factor	Factor	
Variable	1	2	3	Uniqueness	1	2	3	Uniqueness
1	**.96**	-.01	.06	.07	**.96**	-.02	.10	.06
2	**.96**	-.02	.04	.07	**.96**	-.01	.06	.08
3	**.94**	.05	-.10	.11	**.95**	.05	-.13	.08
4	**.94**	.06	-.12	.09	**.96**	.06	-.28	.00
5	**.96**	-.06	.10	.06	**.97**	-.03	.25	.00
6	.01	**.72**	.00	.48	.01	**.72**	-.12	.47
7	-.03	**.75**	-.01	.44	-.02	**.74**	-.06	.44
8	-.04	**.69**	.03	.53	-.03	**.67**	-.03	.54
9	-.01	**.61**	.14	.61	-.01	**.64**	.30	.50
10	.04	-.04	**.45**	.79	.04	-.03	**.06**	.99
11	.01	-.07	**.63**	.60	.01	.04	**.23**	.94
12	-.02	-.04	**.42**	.82	-.01	-.02	**.19**	.96

Both OLS and ML recover the first two major domain factors. However, ML does not recover the weak third factor.

Reason?

Note the nature of error in the data, relative to assumptions:

Data: Model error, no sampling error.

ML: All error is normal theory random sampling error

$\Rightarrow$ more error in smaller correlations.

OLS: No assumption about sampling vs. model error

$\Rightarrow$ error unrelated to size of correlations.

Did this issue give OLS an advantage?

Further study of this phenomenon:

MacCallum, Tucker, & Briggs (2001)
Briggs & MacCallum (2003)
Recent work with help from Michael Browne and Li Cai.

Extended to other structures, numbers of strong vs. weak factors, amount and type of error, other estimation methods.

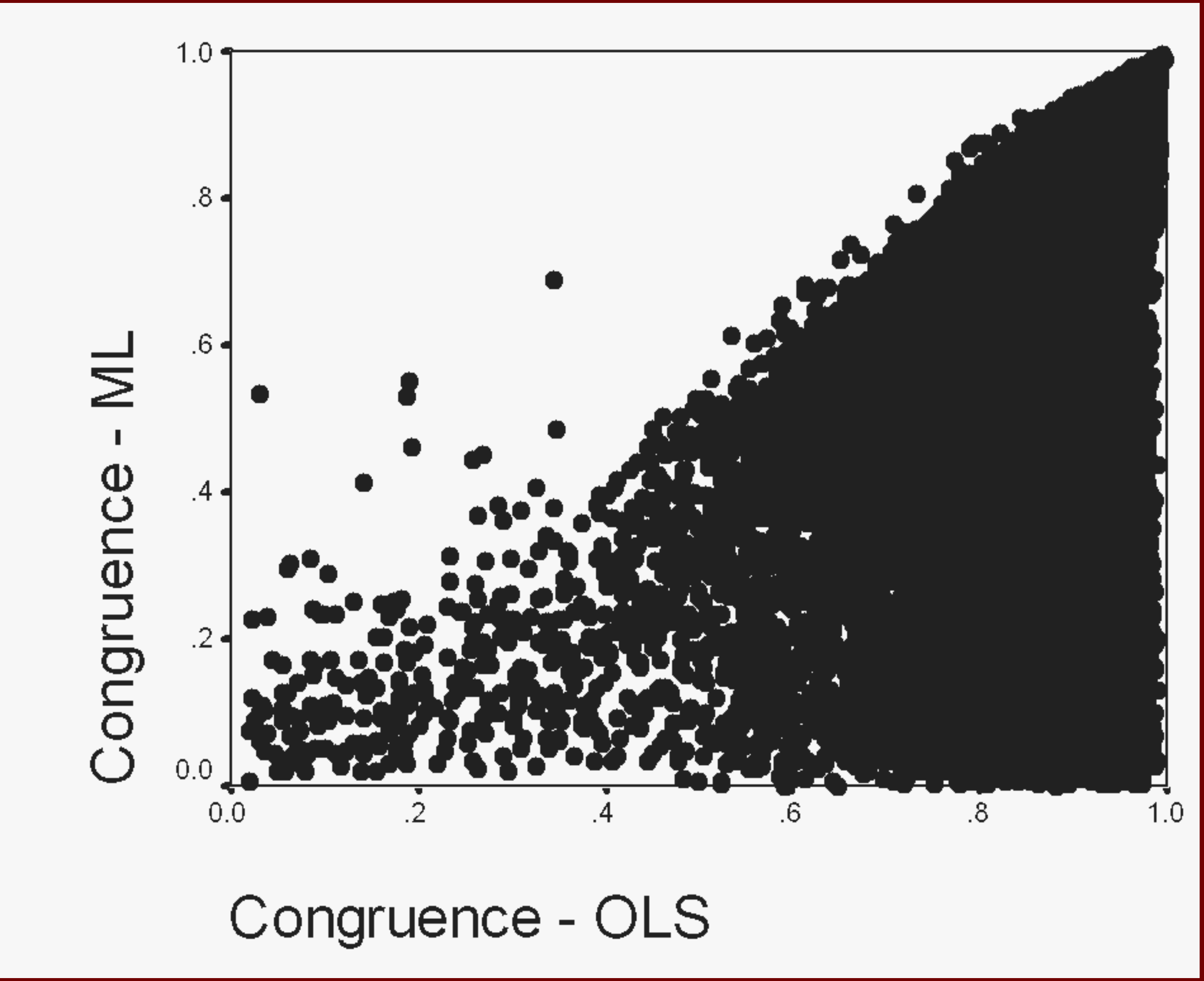
Consider results from one selected condition:

- 16 measured variables
- 4 factors (3 strong, 1 weak)
- Model error introduced via minor factors
- Sampling error, N = 100

- Generate 500 population correlation matrices.
- For each population, generate 200 sample correlation matrices.

- Analyze each matrix using ML and OLS.
- Examine recovery of 4th factor.

Congruence coefficients from 100,000 data sets



Means: ML: .78 OLS: .90

SDs: ML: .21 OLS: .08

Explanation for advantage of OLS?

Consider discrepancy functions:

ML: $$F_{ML} = \ln\left|\Sigma\right| + tr(S\Sigma^{-1}) - \ln\left|S\right| - p$$

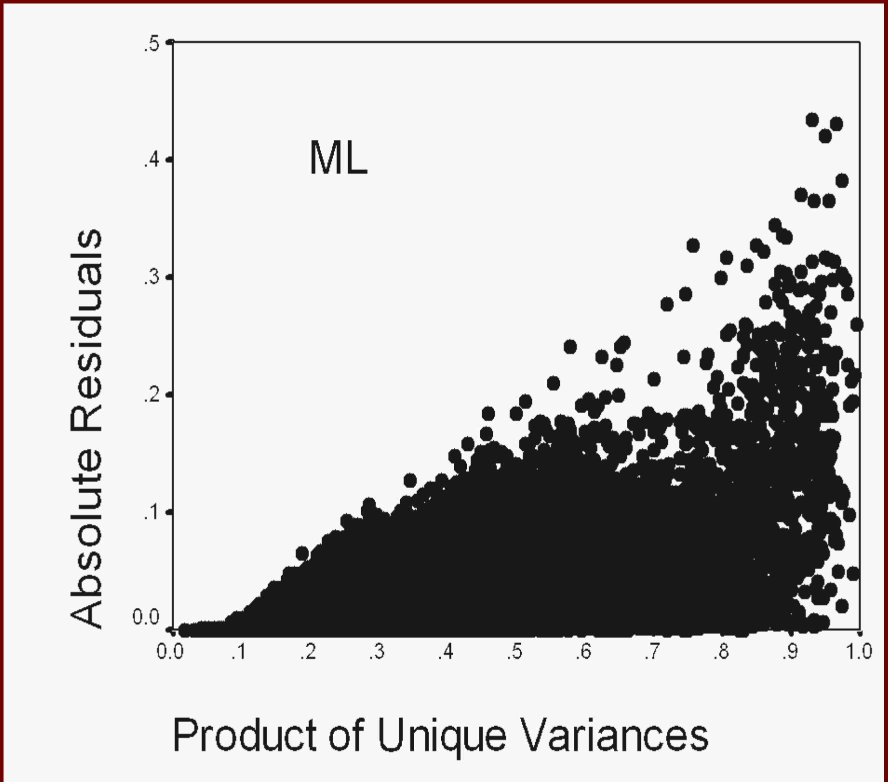
$$F_{ML} \cong \sum_j \sum_k \left[{(s_{jk} - \hat{\sigma}_{jk})^2} \middle/ {u_j^2 u_k^2} \right]$$

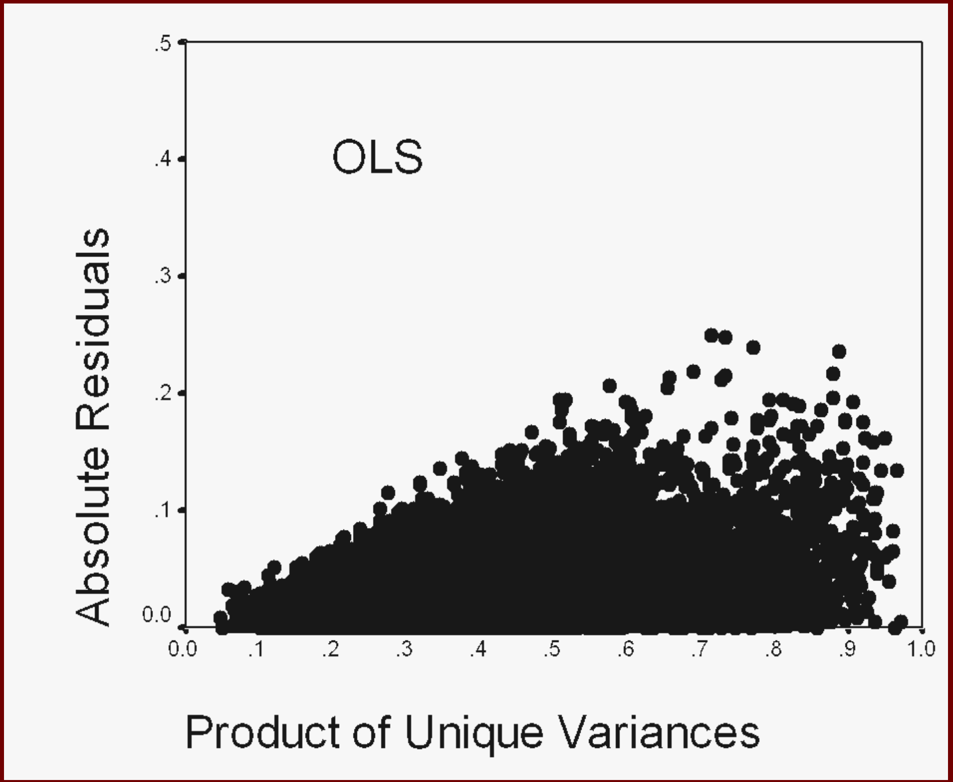
Squared residuals associated with larger product of unique variances receive less weight.

OLS: $$F_{OLS} = \sum_j \sum_k (s_{jk} - \hat{\sigma}_{jk})^2$$

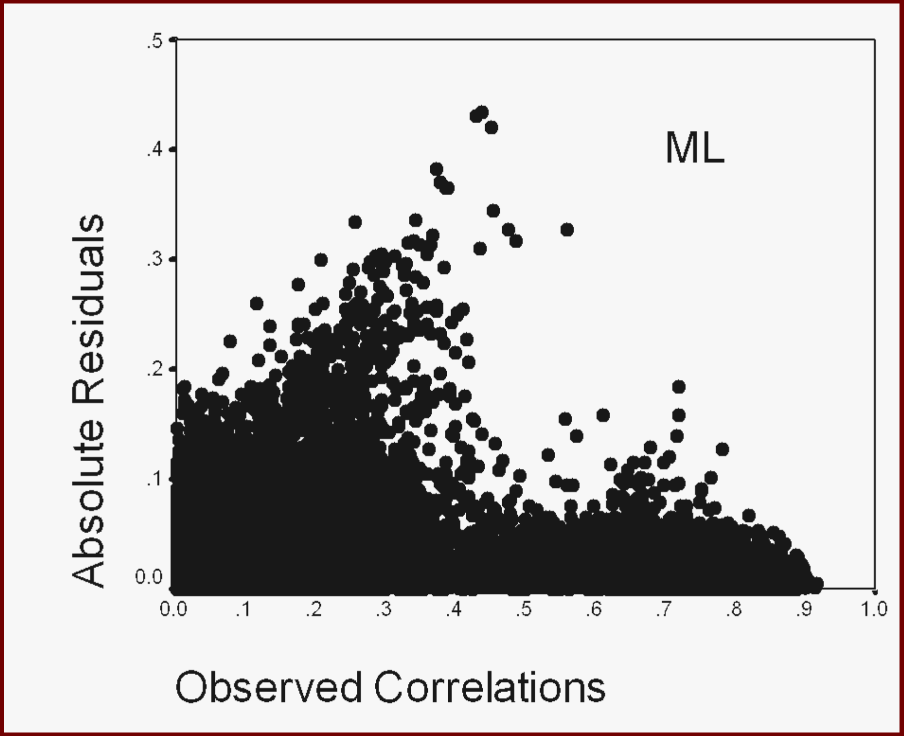
All squared residuals weighted equally.

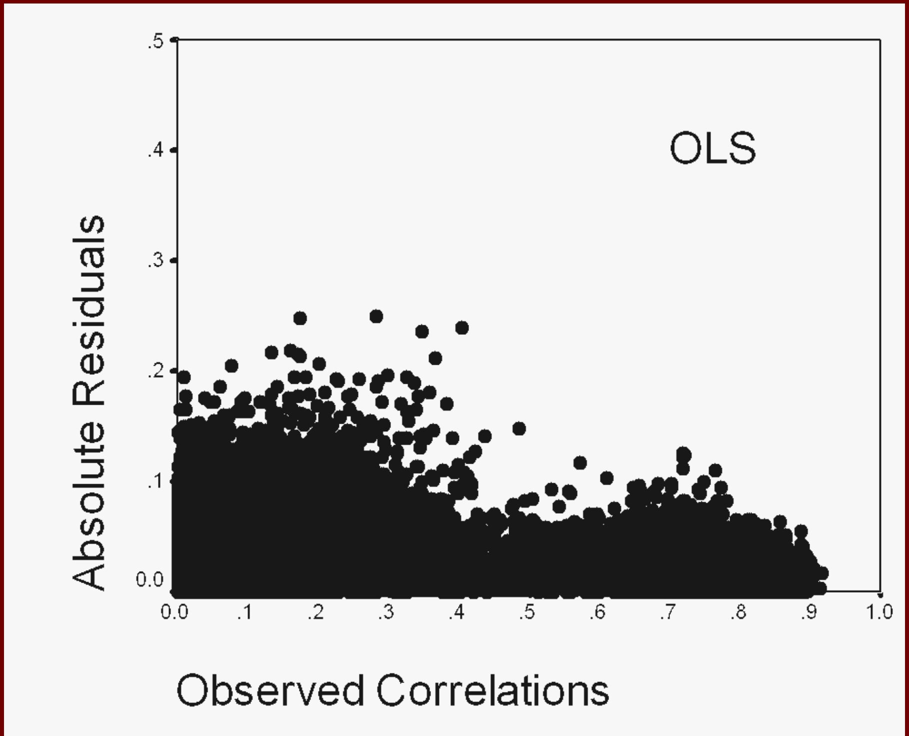
Relationship of residuals to unique variance products (From 400 data sets: 20 populations, 20 samples each; 48,000 points in each plot)





Relationship between residuals and correlations





Proposed explanation of these findings:

Different estimation methods make different assumptions about nature of error

→ Differential weighting of residuals associated with small correlations

→ Differential recovery of weak factors.

Consider another estimation method:
Alpha Factor Analysis (Kaiser & Caffrey, 1965)

A different view about error:
- No sampling error with respect to individuals.
- Consider observed set of variables to be selected from a universe of variables.
- Error in data is "psychometric error" arising from not having the complete universe of variables.

Objective: Determine common factors that have maximum correlation with corresponding factors in universe of variables.

Approximate discrepancy function for Alpha FA:

$$F_{ALPHA} \cong \sum_j \sum_k \left[(s_{jk} - \hat{\sigma}_{jk})^2 \big/ (1 - u_j^2)(1 - u_k^2) \right]$$

Compare to ML:

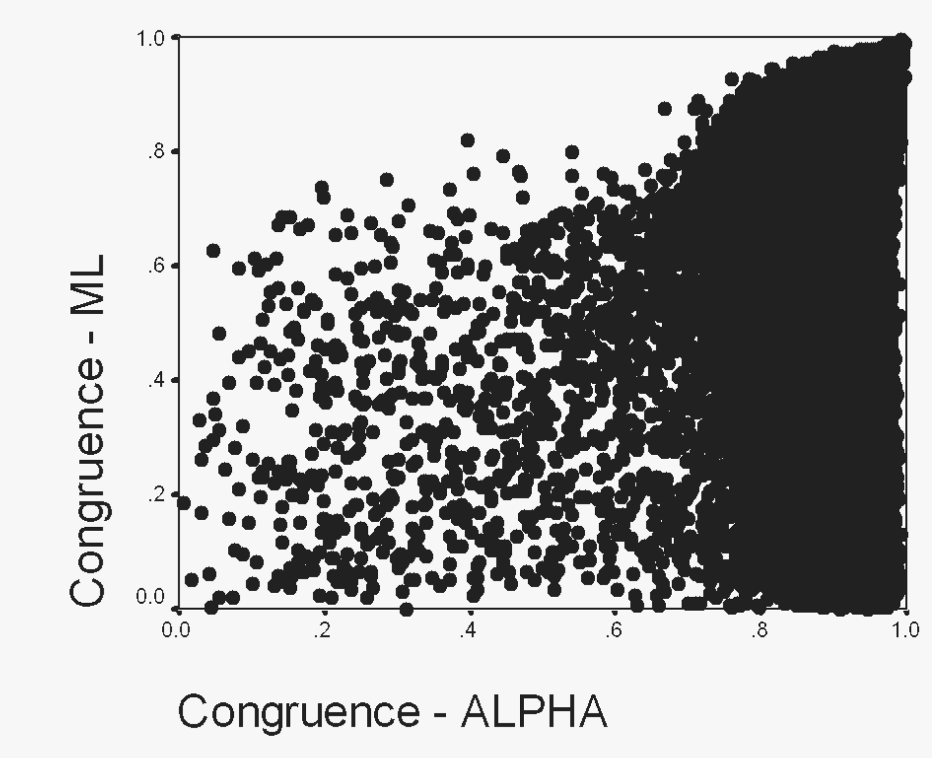
$$F_{ML} \cong \sum_j \sum_k \left[(s_{jk} - \hat{\sigma}_{jk})^2 \big/ u_j^2 u_k^2 \right]$$

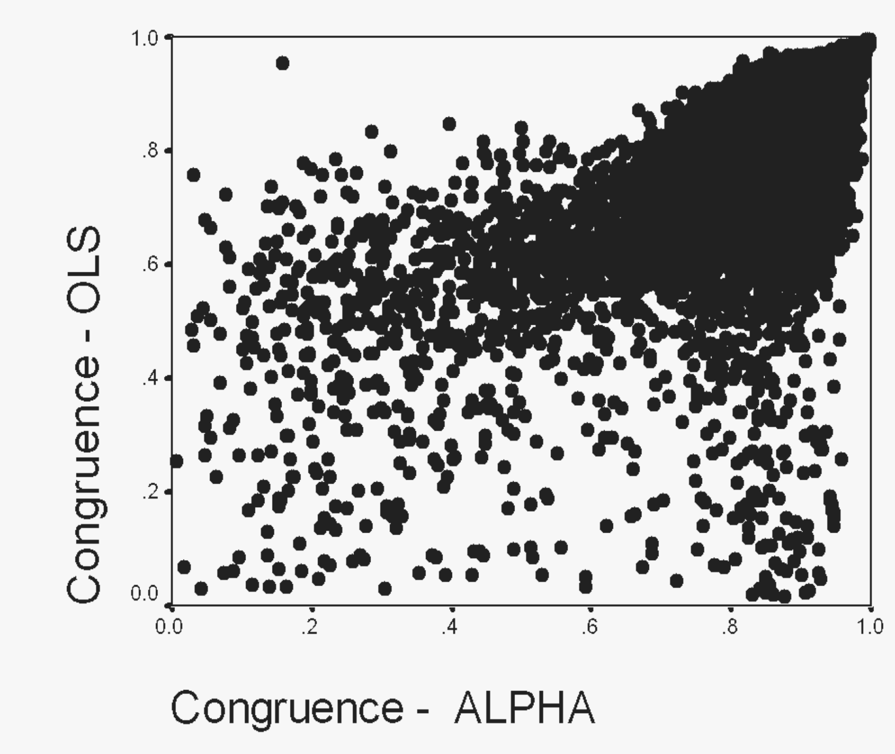
Note inverse weighting of squared residuals.

Alpha assumes less error in correlations associated with large products of unique variances.

Consider ability of Alpha to recover weak factors.

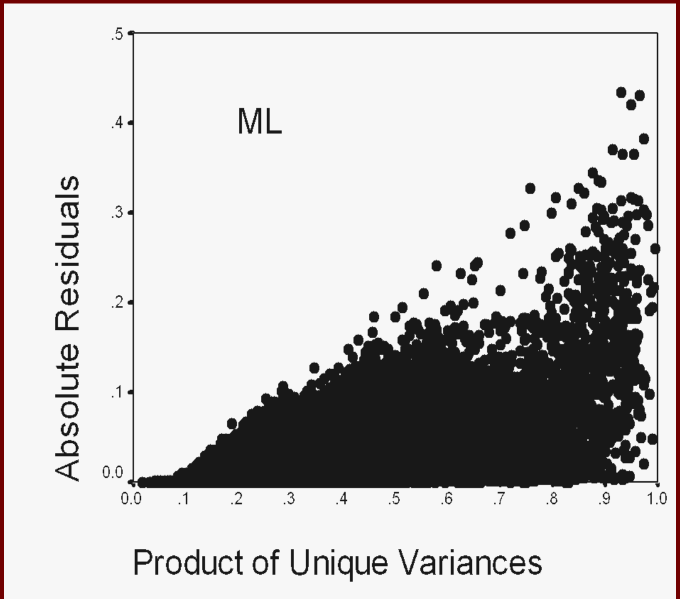
Recovery of weak factor by three methods

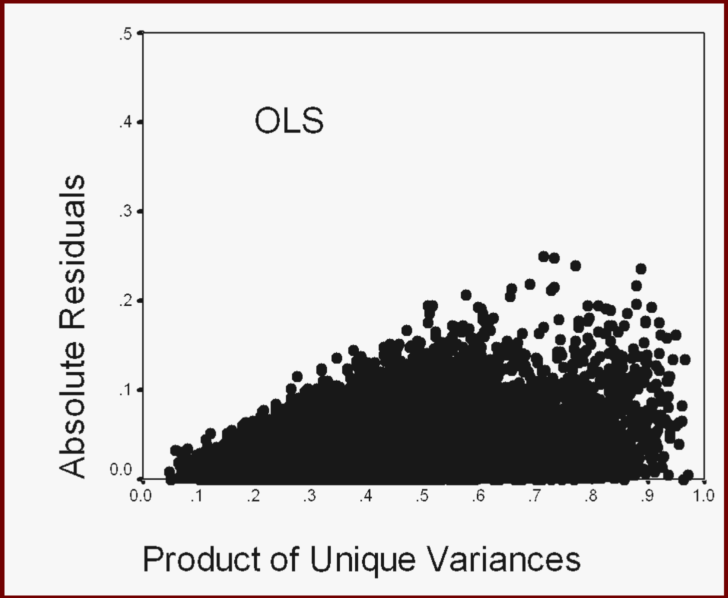


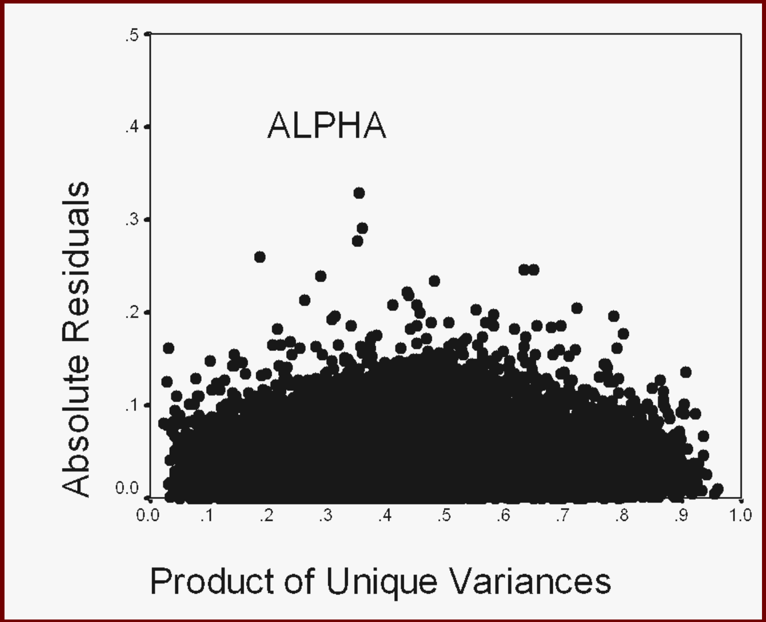


Means:	Alpha:	.91	SDs:	Alpha:	.08
	OLS:	.90		OLS:	.08
	ML:	.78		ML:	.21

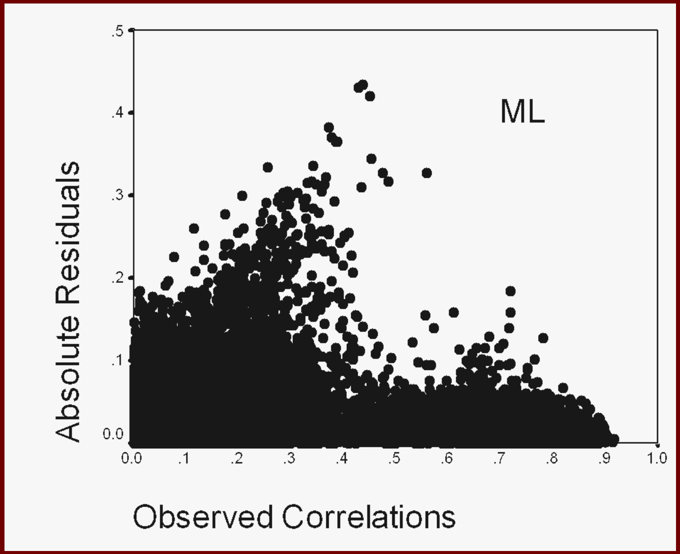
Relationship of residuals to unique variance products

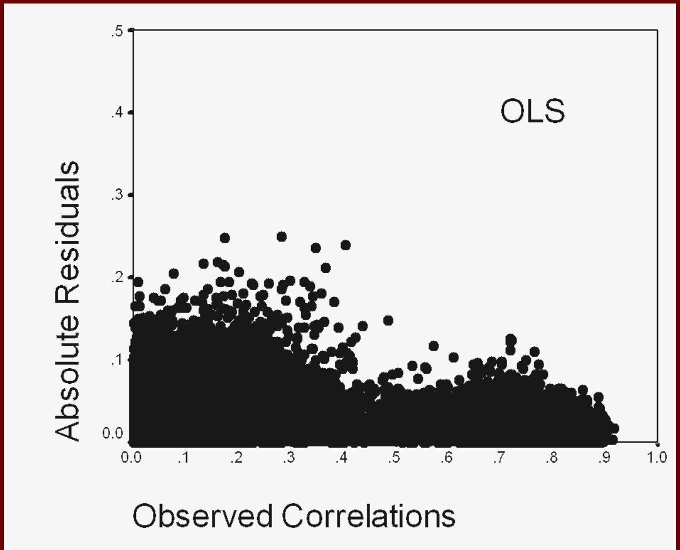


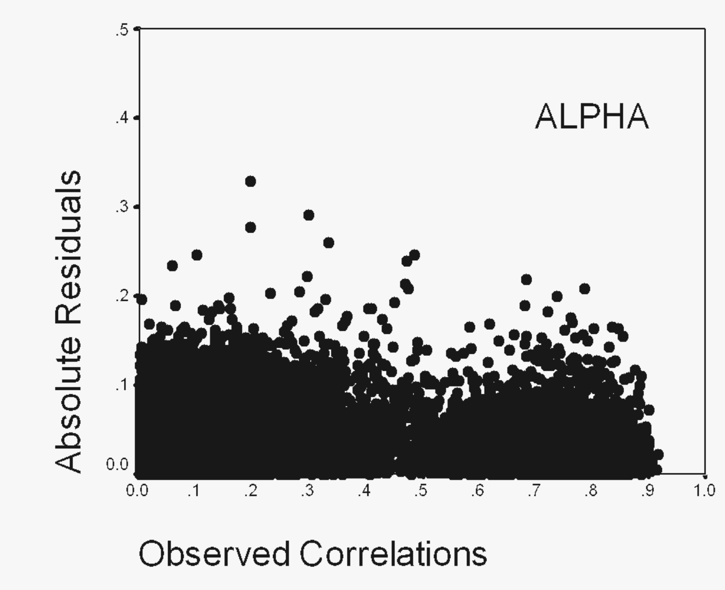




Relationship of residuals to correlations







Summary: Part II

Recognition that FA models are approximations implies that not all error is sampling error.

Different estimation methods make different assumptions about nature of error.

Correspondence between assumptions and nature of error in data has real consequences for parameter estimates.

Consequences have implications for how we might best do factor analysis in practice.

Conclusion

• Like all other scientific models, factor analysis models are approximations of the real world.

• Some early controversy due in part to fact that this principle was not yet developed or understood.

• This principle must be kept in mind when we do FA and interpret our findings.

• This principle may have important implications for how we do factor analysis.