MAESTRÍA EN INVESTIGACIÓN OPERATIVA Y ESTADÍSTICA

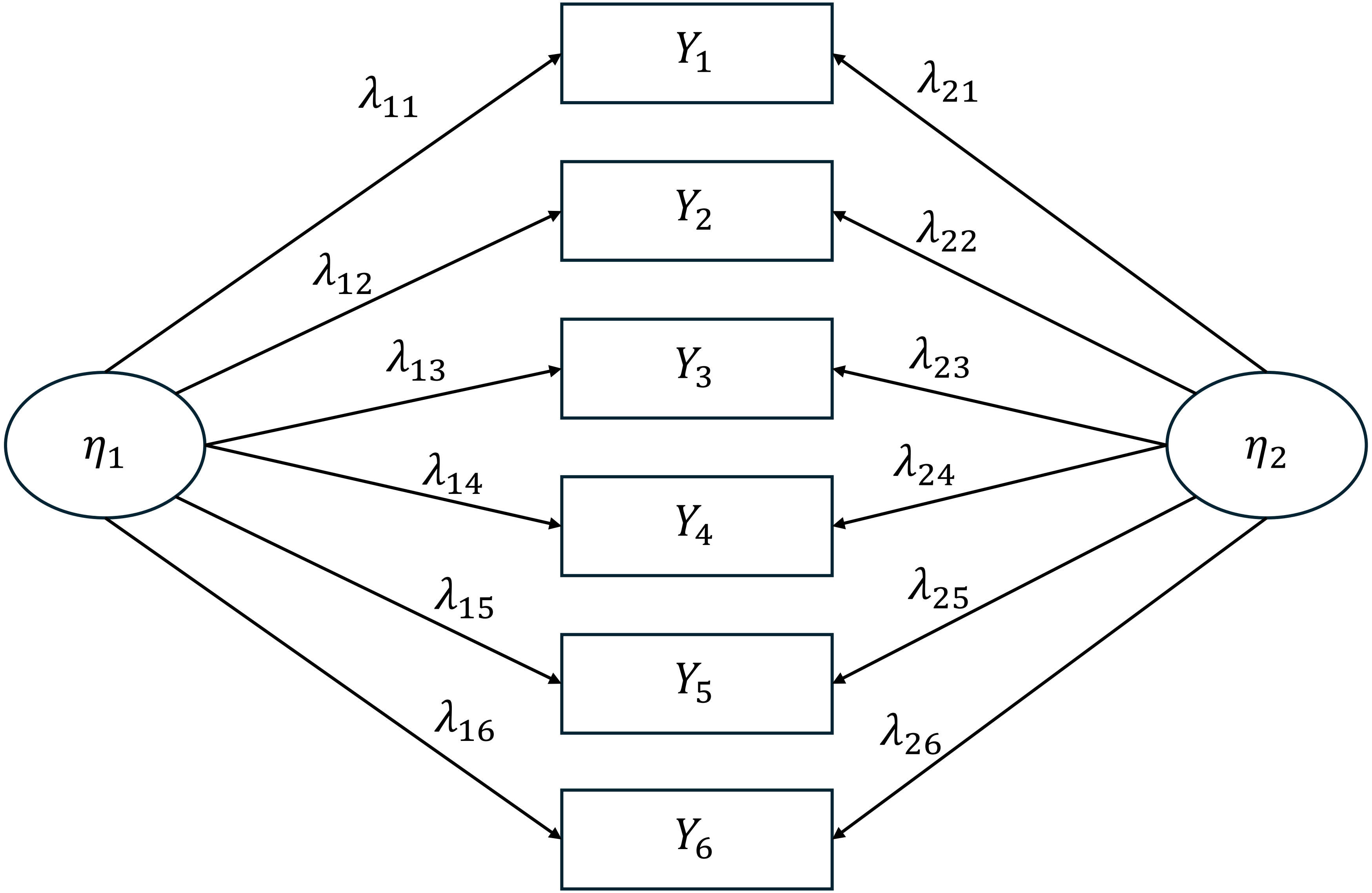
Universidad Tecnológica de Pereira

Multivariate statistics

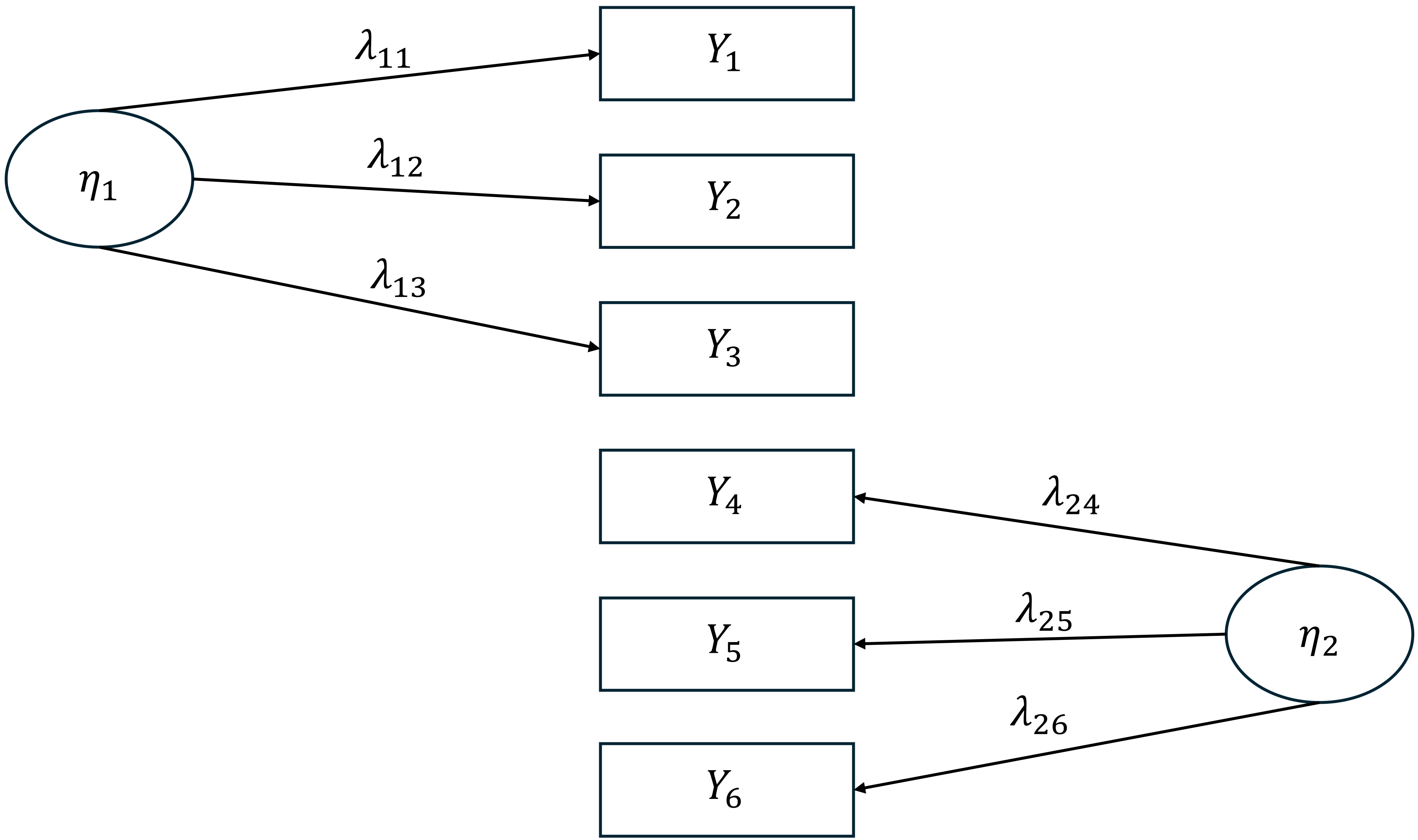
Confirmatory Factor Analysis

May 31, 2024

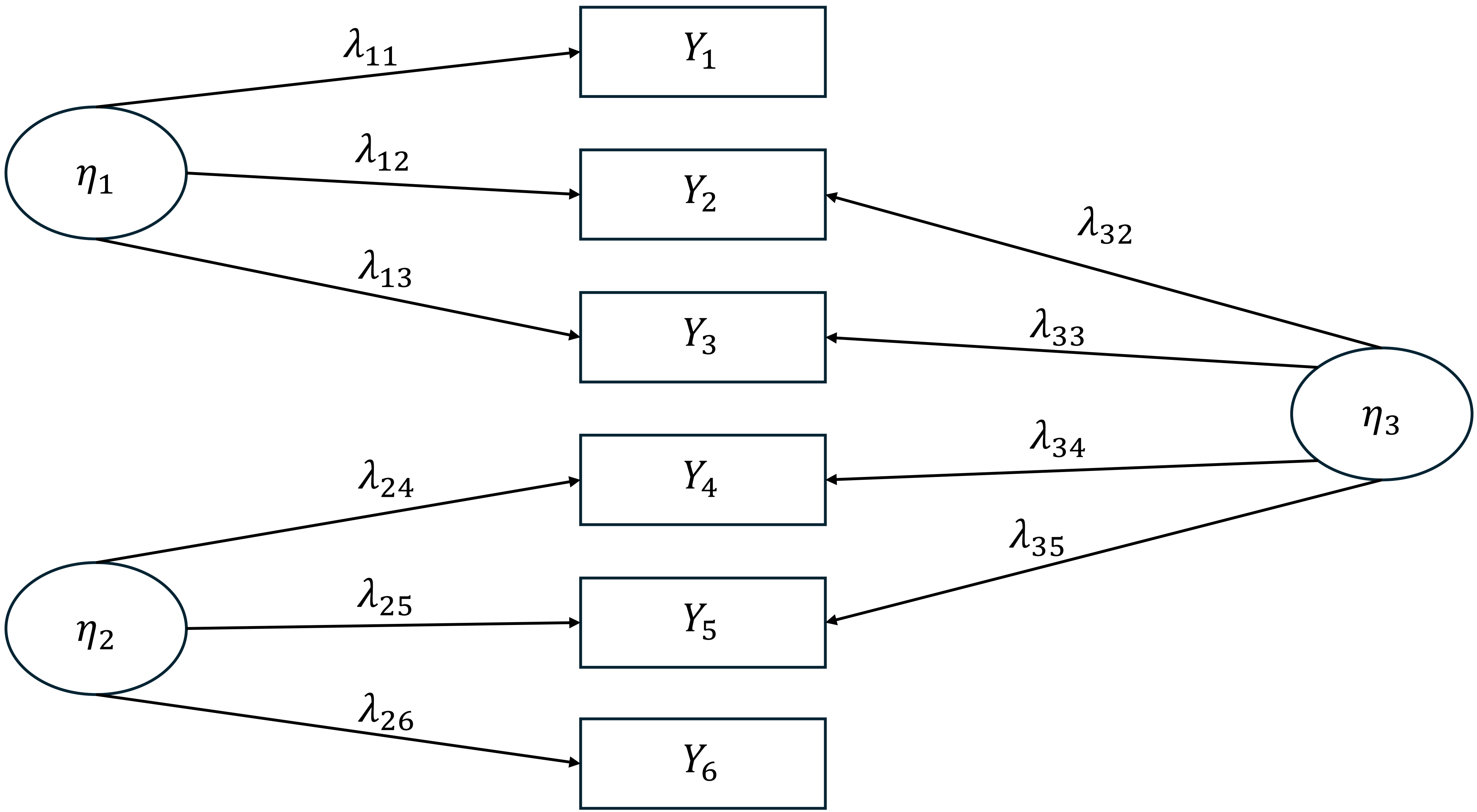
Orlando Sabogal-Cardona
@Antonio Sabogal
orlando.sabogal.20@ucl.ac.uk



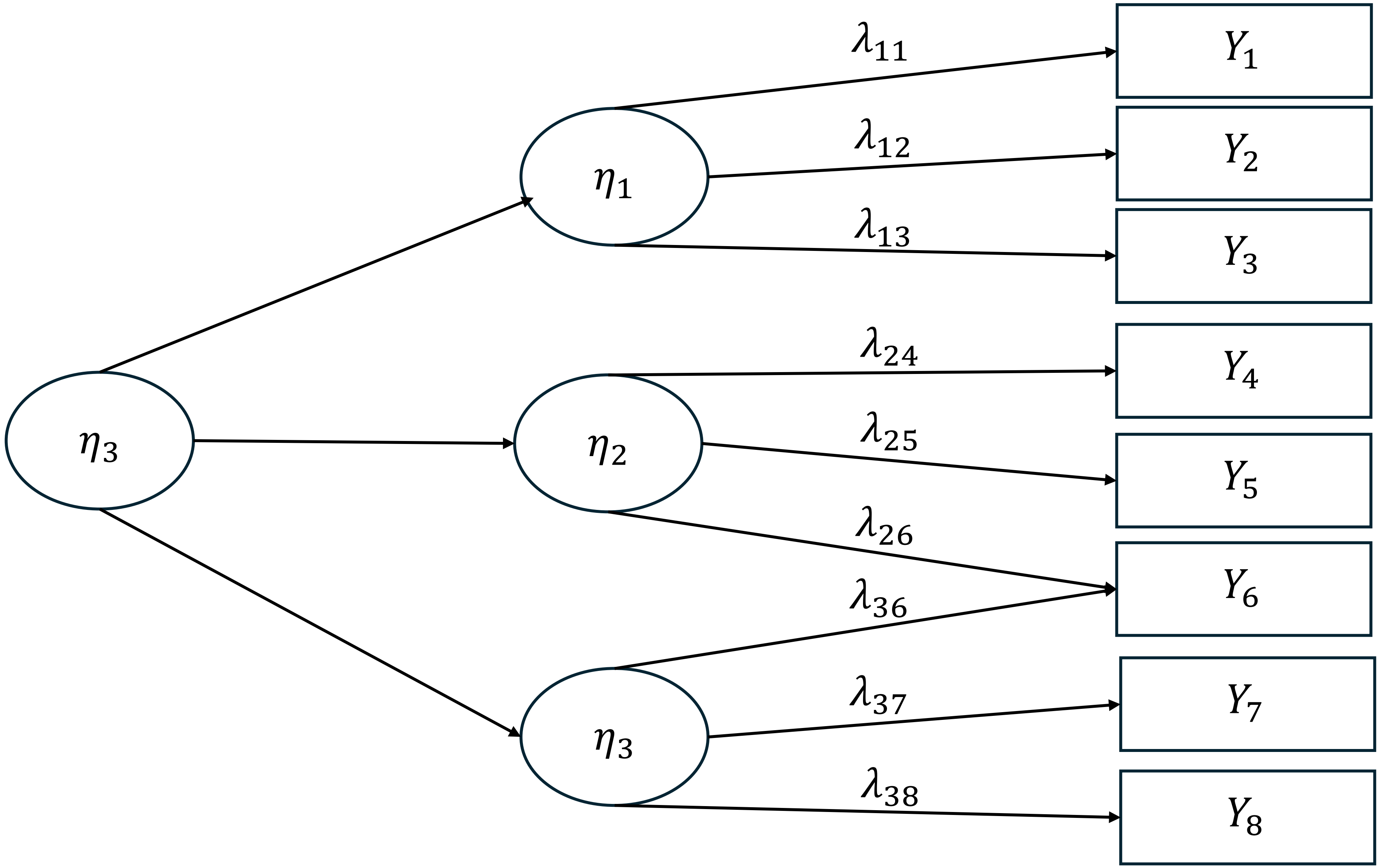
This is an EFA


η_1
λ_{11}
λ_{12}
λ_{13}
Y_1
Y_2
Y_3
Y_4
Y_5
Y_6
λ_{24}
λ_{25}
λ_{26}
η_2

This is a CFA



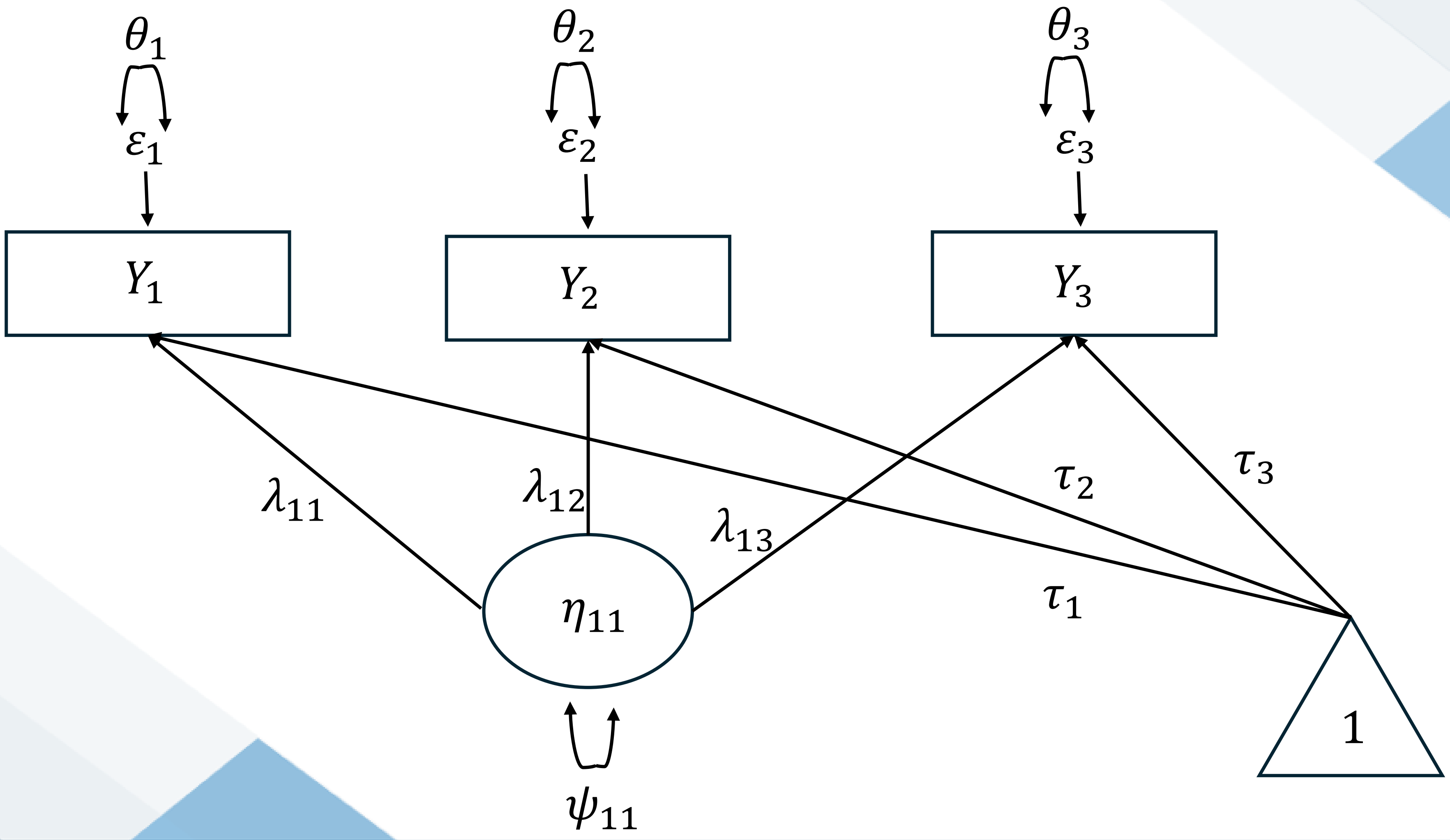
This is also a CFA



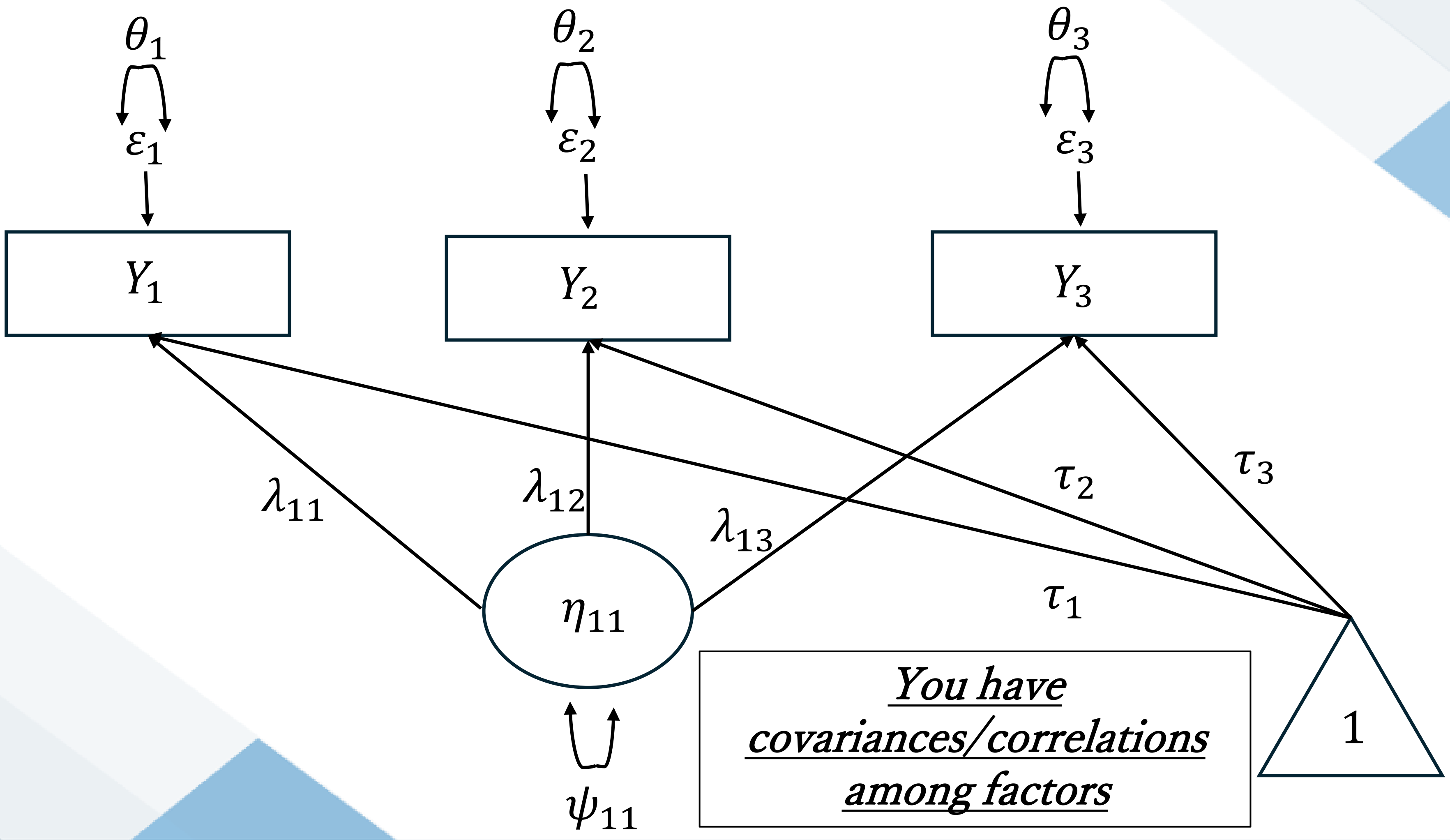
This is also a CFA

MAESTRÍA EN INVESTIGACIÓN OPERATIVA Y ESTADÍSTICA

Remember…



Remember…



You have covariances/correlations among factors

CFA (you already know this)

The common factor model

$$Y = \Lambda\xi + \varepsilon$$

Used to predict the covariance/correlation matrix

$$\boxed{\Sigma = \Lambda\Psi\Lambda' + \Theta}$$

Y: *Matrix of observed indicator variables*

ξ: *Matrix of factors*

Λ: *Factor loading matrix*

ε: *Matrix of unique factors* (*source in variance not associated with* ε)

n: *number of observations*
p: *number of observed variables*
$Y: n * p$
$\Lambda: p * m$
$\xi: n * m$
$\varepsilon: n * p$

S: *Observed covariance or correlation matrix*

Σ: *Model correlation or covariance matrix*

Ψ: *Correlation matrix of the factors*

Θ: *Diagonal matrix of unique error variances*

m: *number of factors*
$S: p * p$
$\Sigma: p * p$
$\Psi: m * m$
$\Theta: p * p$

Make "Σ" as similar to "S" as possible

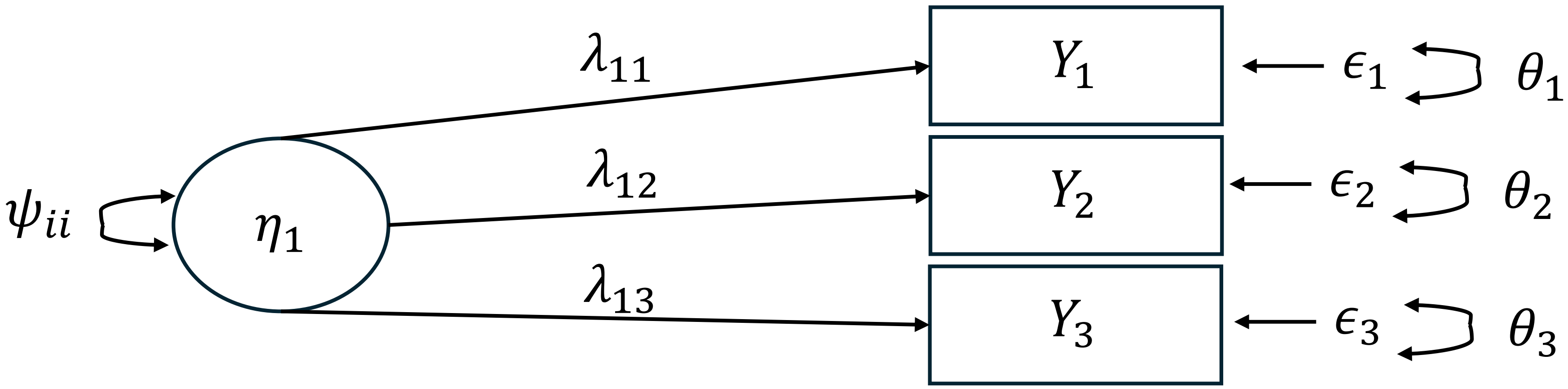
$$\min(\Sigma - S)$$

$$\Sigma = \Lambda\Psi\Lambda' + \Theta$$

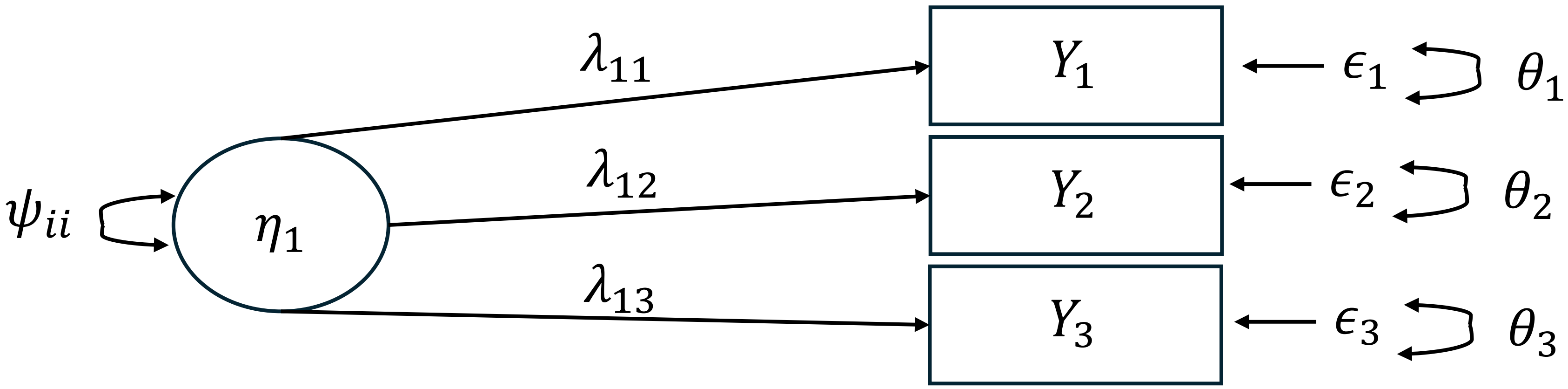
$$\min(\Sigma - S)$$

$$\Sigma = \Lambda\Psi\Lambda' + \Theta$$

Does this make sense to you?



$$Y_1 = \lambda_{11}\eta_1 + \epsilon_1$$



$$Y_1 = \lambda_{11}\eta_1 + \epsilon_1$$

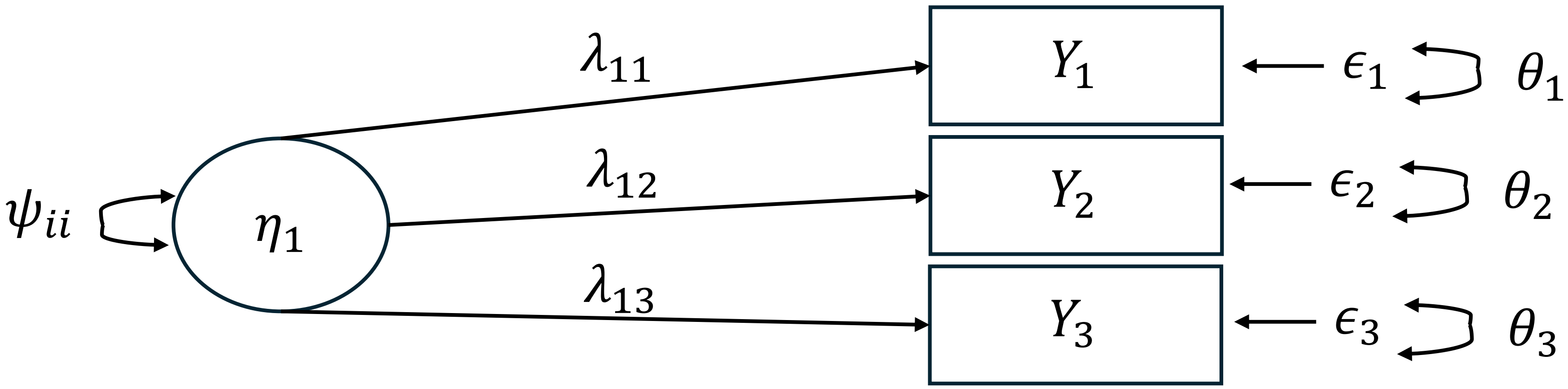
$$var(\epsilon_i) = \theta_i$$

$$var(\eta_i) = \psi_{ii}$$

$$cov(\eta_i, \eta_j) = \psi_{ij}$$

$$var(aX + bY) = a^2 var(X) + b^2 var(Y)$$

See:
https://statproofbook.github.io/P/var-lincomb.html#:~:text=Theorem%3A%20The%20variance%20of%20the,v(X%2CY).



$$Y_1 = \lambda_{11}\eta_1 + \epsilon_1$$

$$var(Y_1) = var(\lambda_{11}\eta_1 + \epsilon_1)$$

$$var(Y_1) = \lambda_{11}^2 var(\eta_1) + var(\epsilon_1)$$

$$\boxed{var(Y_1) = S_{11}^2 = \lambda_{11}^2\psi_{11} + \theta_1}$$

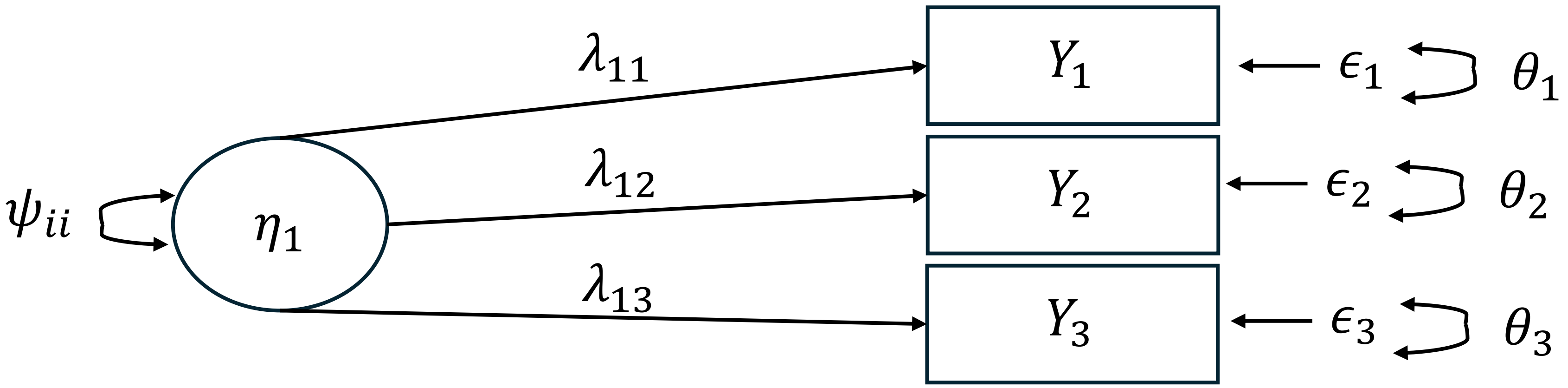
$$var(\epsilon_i) = \theta_i$$

$$var(\eta_i) = \psi_{ii}$$

$$cov(\eta_i, \eta_j) = \psi_{ij}$$

$$var(aX + bY) = a^2 var(X) + b^2 var(Y)$$

See:
https://statproofbook.github.io/P/var-lincomb.html#:~:text=Theorem%3A%20The%20variance%20of%20the,v(X%2CY).

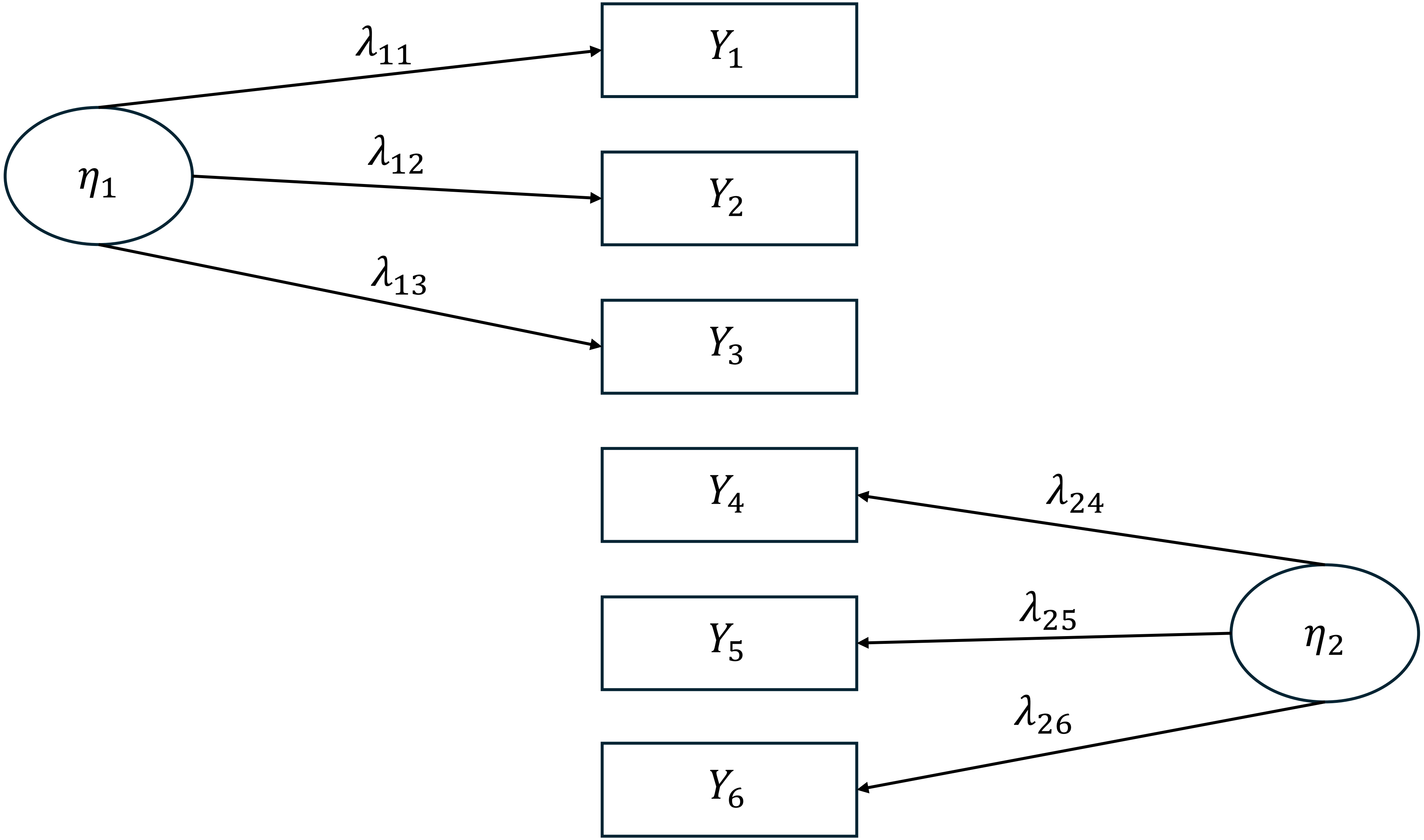


$$Y_1 = \lambda_{11}\eta_1 + \epsilon_1 \qquad\qquad Y_2 = \lambda_{21}\eta_1 + \epsilon_2$$

$$cov(aX + bY, cX + dY) = acCov(X, X) + adCov(X, Y) + bcCov(Y, X) + bdCov(Y, Y)$$

$$cov(Y_1, Y_2) = cov(\lambda_{11}\eta_1 + \epsilon_1, \lambda_{21}\eta_1 + \epsilon_1)$$

$$\boxed{cov(Y_1, Y_2) = \lambda_{11}\lambda_{12}\psi_{11}}$$



λ_{11}
Y_1
η_1
λ_{12}
Y_2
λ_{13}
Y_3
Y_4
λ_{24}
Y_5
λ_{25}
η_2
λ_{26}
Y_6

$$\Lambda = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{12} & 0 \\ \lambda_{13} & 0 \\ 0 & \lambda_{24} \\ 0 & \lambda_{25} \\ 0 & \lambda_{26} \end{bmatrix}$$

$$\Psi = \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix}$$

$$\Theta = \begin{bmatrix} \theta_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_6 \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{12} & 0 \\ \lambda_{13} & 0 \\ 0 & \lambda_{24} \\ 0 & \lambda_{25} \\ 0 & \lambda_{26} \end{bmatrix} \qquad \Psi = \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix} \qquad \Theta = \begin{bmatrix} \theta_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_6 \end{bmatrix}$$

$$\Lambda\Psi = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{12} & 0 \\ \lambda_{13} & 0 \\ 0 & \lambda_{24} \\ 0 & \lambda_{25} \\ 0 & \lambda_{26} \end{bmatrix} \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix} = \begin{bmatrix} \lambda_{11}\psi_{11} & \lambda_{11}\psi_{12} \\ \lambda_{12}\psi_{11} & \lambda_{12}\psi_{12} \\ \lambda_{13}\psi_{11} & \lambda_{13}\psi_{12} \\ \lambda_{24}\psi_{21} & \lambda_{24}\psi_{22} \\ \lambda_{25}\psi_{21} & \lambda_{25}\psi_{22} \\ \lambda_{26}\psi_{21} & \lambda_{26}\psi_{22} \end{bmatrix}$$

$$\Lambda\Psi\Lambda' = \begin{bmatrix} \lambda_{11}\psi_{11} & \lambda_{11}\psi_{12} \\ \lambda_{12}\psi_{11} & \lambda_{12}\psi_{12} \\ \lambda_{13}\psi_{11} & \lambda_{13}\psi_{12} \\ \lambda_{24}\psi_{21} & \lambda_{24}\psi_{22} \\ \lambda_{25}\psi_{21} & \lambda_{25}\psi_{22} \\ \lambda_{26}\psi_{21} & \lambda_{26}\psi_{22} \end{bmatrix} \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{24} & \lambda_{25} & \lambda_{26} \end{bmatrix}$$

$$\Lambda\Psi\Lambda' = \begin{pmatrix} \lambda_{11}^2\psi_{11} & \lambda_{11}\lambda_{21}\psi_{11} & \lambda_{11}\lambda_{31}\psi_{11} & \lambda_{11}\lambda_{42}\psi_{12} & \lambda_{11}\lambda_{52}\psi_{12} & \lambda_{11}\lambda_{62}\psi_{12} \\ \lambda_{21}\lambda_{11}\psi_{11} & \lambda_{21}^2\psi_{11} & \lambda_{21}\lambda_{31}\psi_{11} & \lambda_{21}\lambda_{42}\psi_{12} & \lambda_{21}\lambda_{52}\psi_{12} & \lambda_{21}\lambda_{62}\psi_{12} \\ \lambda_{31}\lambda_{11}\psi_{11} & \lambda_{31}\lambda_{21}\psi_{11} & \lambda_{31}^2\psi_{11} & \lambda_{31}\lambda_{42}\psi_{12} & \lambda_{31}\lambda_{52}\psi_{12} & \lambda_{31}\lambda_{62}\psi_{12} \\ \lambda_{42}\lambda_{11}\psi_{21} & \lambda_{42}\lambda_{21}\psi_{21} & \lambda_{42}\lambda_{31}\psi_{21} & \lambda_{42}^2\psi_{22} & \lambda_{42}\lambda_{52}\psi_{22} & \lambda_{42}\lambda_{62}\psi_{22} \\ \lambda_{52}\lambda_{11}\psi_{21} & \lambda_{52}\lambda_{21}\psi_{21} & \lambda_{52}\lambda_{31}\psi_{21} & \lambda_{52}\lambda_{42}\psi_{22} & \lambda_{52}^2\psi_{22} & \lambda_{52}\lambda_{62}\psi_{22} \\ \lambda_{62}\lambda_{11}\psi_{21} & \lambda_{62}\lambda_{21}\psi_{21} & \lambda_{62}\lambda_{31}\psi_{21} & \lambda_{62}\lambda_{42}\psi_{22} & \lambda_{62}\lambda_{52}\psi_{22} & \lambda_{62}^2\psi_{22} \end{pmatrix}$$

$$\Sigma = \Lambda\Psi\Lambda' + \Theta$$

$$\Sigma = \begin{pmatrix} \lambda_{11}^2\psi_{11} + \theta_1 & \lambda_{11}\lambda_{21}\psi_{11} & \lambda_{11}\lambda_{31}\psi_{11} & \lambda_{11}\lambda_{42}\psi_{12} & \lambda_{11}\lambda_{52}\psi_{12} & \lambda_{11}\lambda_{62}\psi_{12} \\ \lambda_{21}\lambda_{11}\psi_{11} & \lambda_{21}^2\psi_{11} + \theta_2 & \lambda_{21}\lambda_{31}\psi_{11} & \lambda_{21}\lambda_{42}\psi_{12} & \lambda_{21}\lambda_{52}\psi_{12} & \lambda_{21}\lambda_{62}\psi_{12} \\ \lambda_{31}\lambda_{11}\psi_{11} & \lambda_{31}\lambda_{21}\psi_{11} & \lambda_{31}^2\psi_{11} + \theta_3 & \lambda_{31}\lambda_{42}\psi_{12} & \lambda_{31}\lambda_{52}\psi_{12} & \lambda_{31}\lambda_{62}\psi_{12} \\ \lambda_{42}\lambda_{11}\psi_{21} & \lambda_{42}\lambda_{21}\psi_{21} & \lambda_{42}\lambda_{31}\psi_{21} & \lambda_{42}^2\psi_{22} + \theta_4 & \lambda_{42}\lambda_{52}\psi_{22} & \lambda_{42}\lambda_{62}\psi_{22} \\ \lambda_{52}\lambda_{11}\psi_{21} & \lambda_{52}\lambda_{21}\psi_{21} & \lambda_{52}\lambda_{31}\psi_{21} & \lambda_{52}\lambda_{42}\psi_{22} & \lambda_{52}^2\psi_{22} + \theta_5 & \lambda_{52}\lambda_{62}\psi_{22} \\ \lambda_{62}\lambda_{11}\psi_{21} & \lambda_{62}\lambda_{21}\psi_{21} & \lambda_{62}\lambda_{31}\psi_{21} & \lambda_{62}\lambda_{42}\psi_{22} & \lambda_{62}\lambda_{52}\psi_{22} & \lambda_{62}^2\psi_{22} + \theta_6 \end{pmatrix}$$

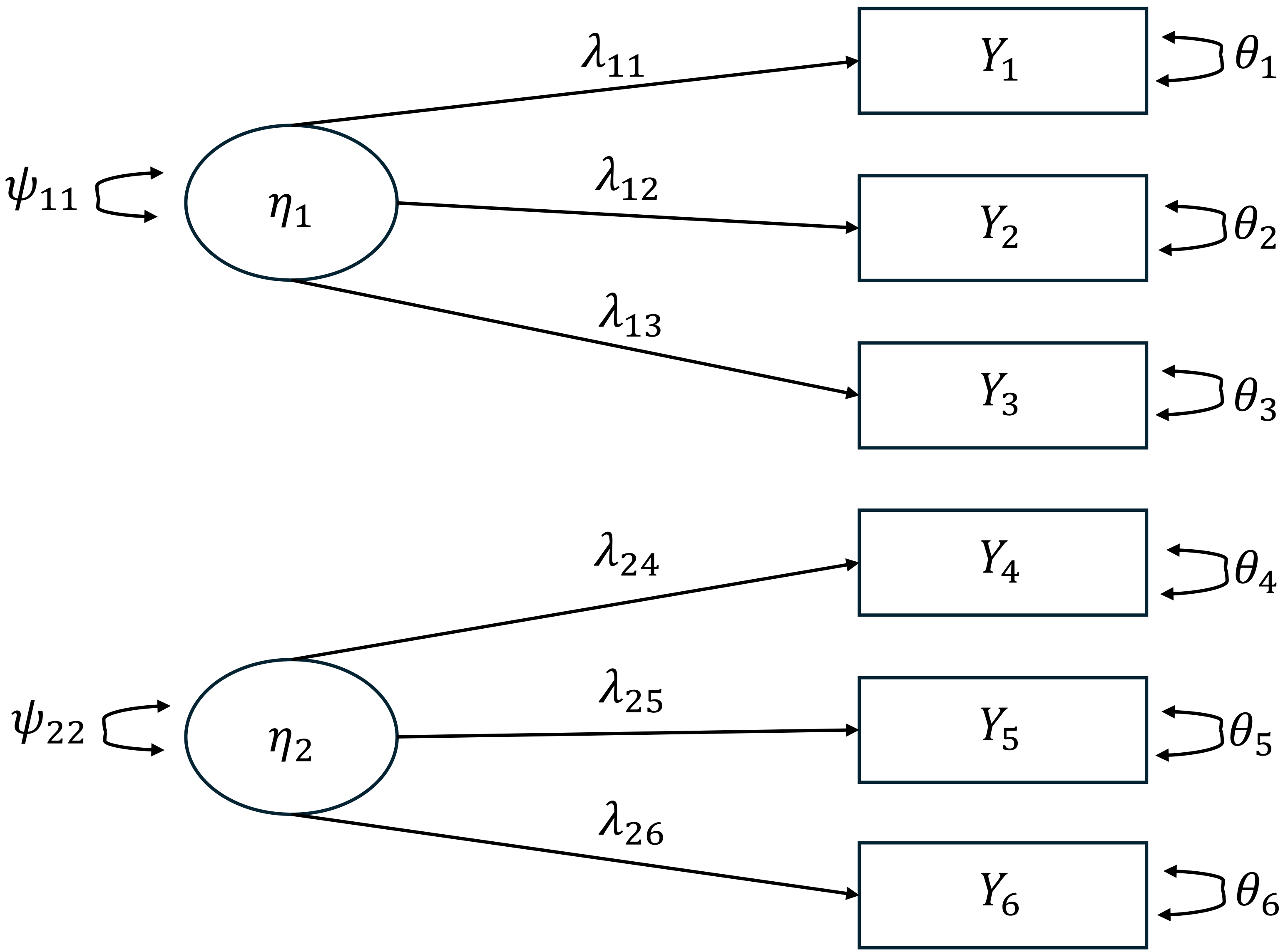
$$\min(\Sigma - S)$$

$$\Sigma = \Lambda\Psi\Lambda' + \Theta$$

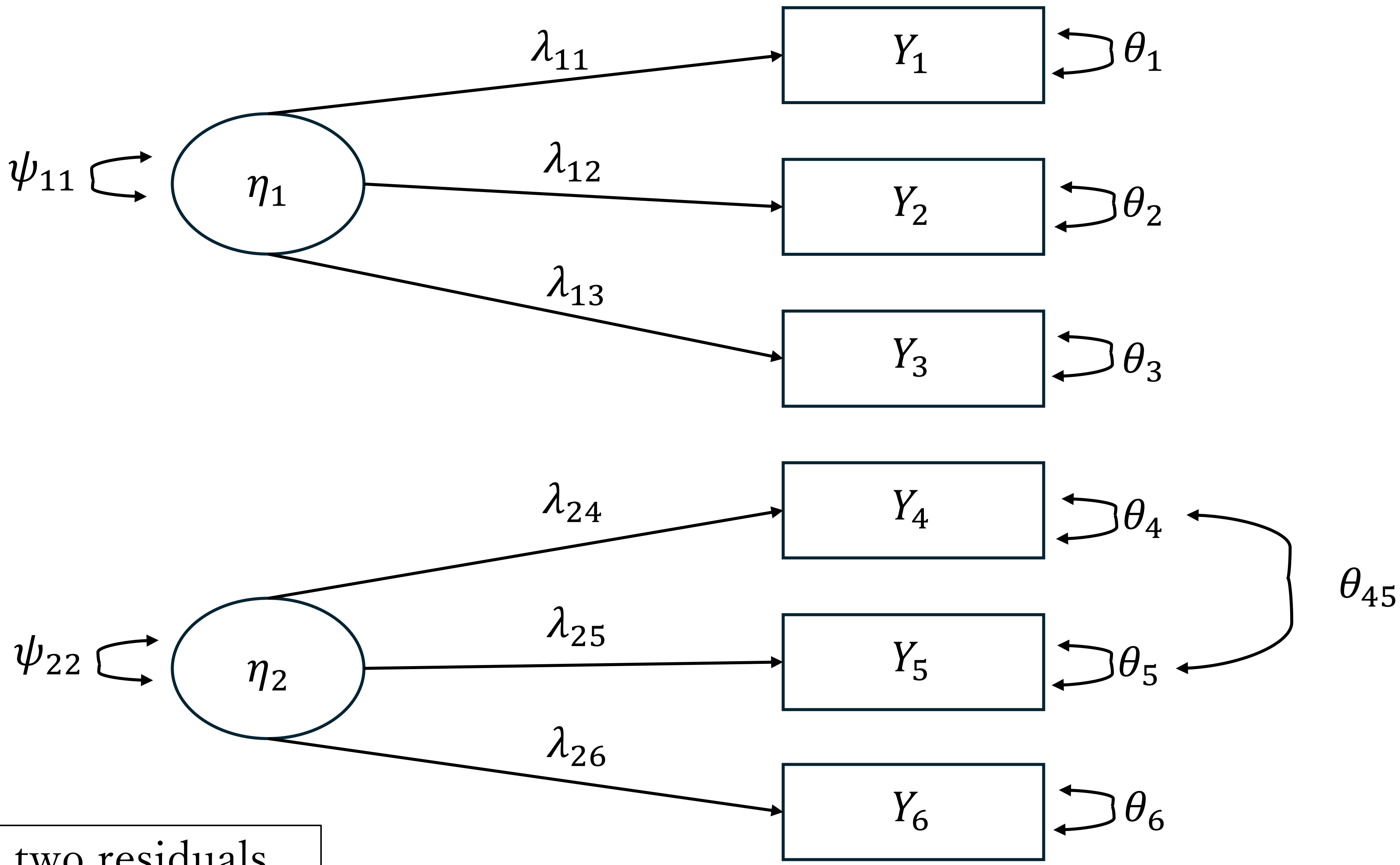
Model parameters

Quantities that the model aims to estimate from the data:

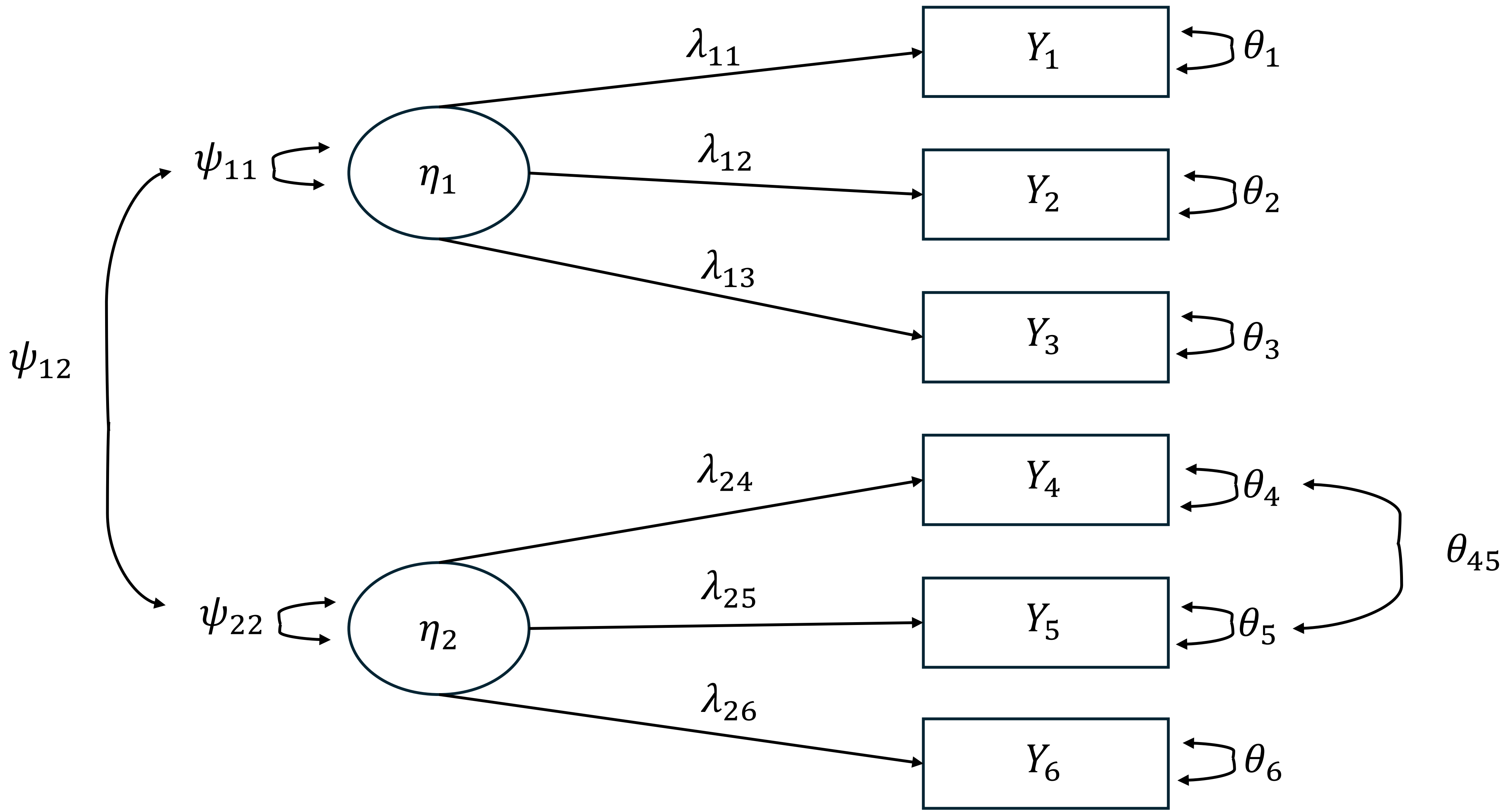
- Factor loadings
- Factor variances
- Factor Covariances
- Measurement error
- Residual covariances (error covariances)
- Intercepts



λ_{11}
Y_1
θ_1
ψ_{11}
η_1
λ_{12}
Y_2
θ_2
λ_{13}
Y_3
θ_3
λ_{24}
Y_4
θ_4
ψ_{22}
η_2
λ_{25}
Y_5
θ_5
λ_{26}
Y_6
θ_6



Why would two residuals correlate?


λ_{11}
λ_{12}
λ_{13}
λ_{24}
λ_{25}
λ_{26}
ψ_{11}
ψ_{12}
ψ_{22}
η_1
η_2
Y_1
Y_2
Y_3
Y_4
Y_5
Y_6
θ_1
θ_2
θ_3
θ_4
θ_5
θ_6
θ_{45}

Model parameters

How many pieces of information do you have?

Model parameters

How many pieces of information do you have?

HINT:

$$\Sigma = \begin{pmatrix} \lambda_{11}^2\psi_{11}+\theta_1 & \lambda_{11}\lambda_{21}\psi_{11} & \lambda_{11}\lambda_{31}\psi_{11} & \lambda_{11}\lambda_{42}\psi_{12} & \lambda_{11}\lambda_{52}\psi_{12} & \lambda_{11}\lambda_{62}\psi_{12} \\ \lambda_{21}\lambda_{11}\psi_{11} & \lambda_{21}^2\psi_{11}+\theta_2 & \lambda_{21}\lambda_{31}\psi_{11} & \lambda_{21}\lambda_{42}\psi_{12} & \lambda_{21}\lambda_{52}\psi_{12} & \lambda_{21}\lambda_{62}\psi_{12} \\ \lambda_{31}\lambda_{11}\psi_{11} & \lambda_{31}\lambda_{21}\psi_{11} & \lambda_{31}^2\psi_{11}+\theta_3 & \lambda_{31}\lambda_{42}\psi_{12} & \lambda_{31}\lambda_{52}\psi_{12} & \lambda_{31}\lambda_{62}\psi_{12} \\ \lambda_{42}\lambda_{11}\psi_{21} & \lambda_{42}\lambda_{21}\psi_{21} & \lambda_{42}\lambda_{31}\psi_{21} & \lambda_{42}^2\psi_{22}+\theta_4 & \lambda_{42}\lambda_{52}\psi_{22} & \lambda_{42}\lambda_{62}\psi_{22} \\ \lambda_{52}\lambda_{11}\psi_{21} & \lambda_{52}\lambda_{21}\psi_{21} & \lambda_{52}\lambda_{31}\psi_{21} & \lambda_{52}\lambda_{42}\psi_{22} & \lambda_{52}^2\psi_{22}+\theta_5 & \lambda_{52}\lambda_{62}\psi_{22} \\ \lambda_{62}\lambda_{11}\psi_{21} & \lambda_{62}\lambda_{21}\psi_{21} & \lambda_{62}\lambda_{31}\psi_{21} & \lambda_{62}\lambda_{42}\psi_{22} & \lambda_{62}\lambda_{52}\psi_{22} & \lambda_{62}^2\psi_{22}+\theta_6 \end{pmatrix}$$

Model parameters

How many pieces of information do you have?

$$\frac{p(p+1)}{2}$$

p: *number of observed variables*

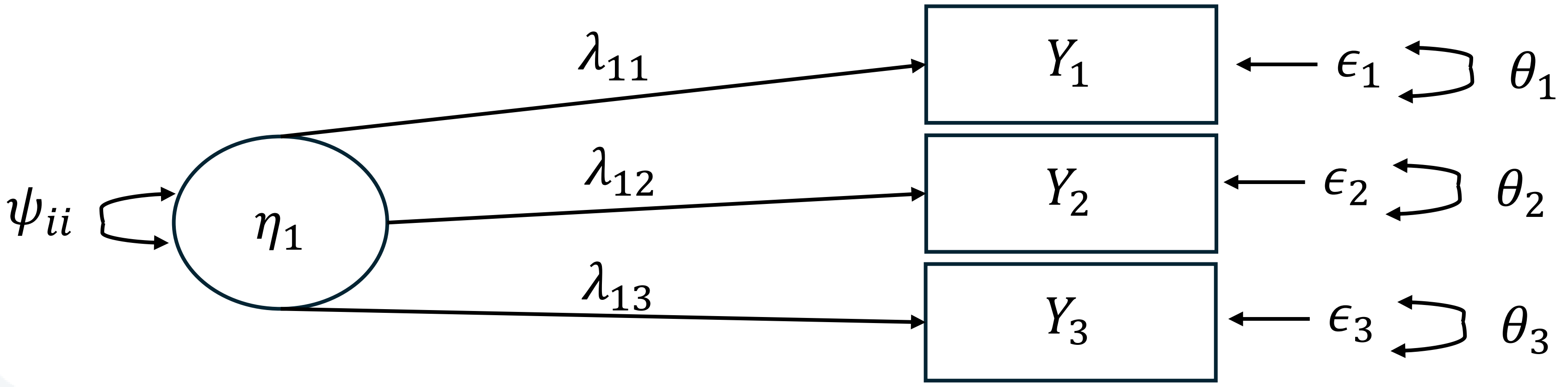
Degrees of freedom

In the context of CFA degrees of freedom refer to the number of independent pieces of information available to estimate the model parameters minus the number of parameters to be estimated. Degrees of freedom are crucial for determining model identification and for conducting statistical tests to assess model fit.

Degrees of freedom

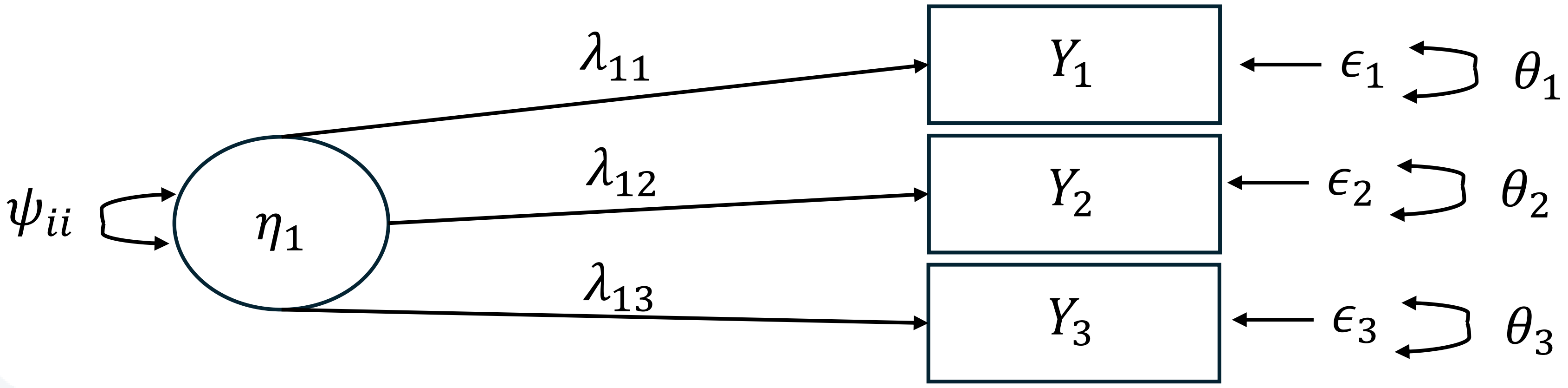
Positive Degrees of Freedom (Over-identified Model): Such models are testable, and statistical tests can be performed to assess model fit.
Zero Degrees of Freedom (Just-identified Model): These models fit the data perfectly by definition but are not testable in terms of fit.
Negative Degrees of Freedom (Under-identified Model): Such models cannot be estimated uniquely.

How many parameters?



$$\Sigma = \begin{bmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \end{bmatrix} \psi_{11} \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \end{bmatrix} + \begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}$$

How many parameters?



$$\Sigma = \begin{bmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \end{bmatrix} \psi_{11} \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \end{bmatrix} + \begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}$$

7 parameters and 6 pieces of information

Scaling the latent variable

- Since latent variables are not directly observed, their scales are not inherently defined.
- The model would be "under-identified" (conceptually) because there would be an infinite number of solutions that fit the data equally well.
- Scaling the latent variable refers to the process of fixing their scale or metric
- Scaling makes the model identifiable and interpretable.

Scaling the latent variable

There are two primary methods:

- Fixing a factor loading to 1: Sets the scale of the latent variable to be the same as that of the selected indicator variable
- Fixing the Variance of the Latent Variable to 1:This standardizes the latent variable, meaning it has a mean of 0 and a variance of 1.

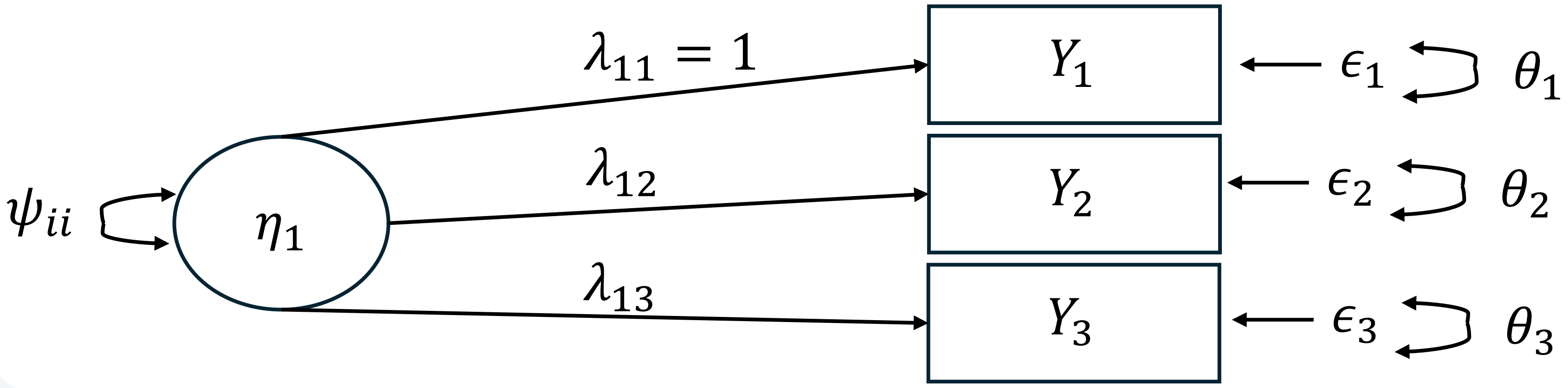
Scaling the latent variable

There are two primary methods:

- Fixing a factor loading to 1 (marker method): Sets the scale of the latent variable to be the same as that of the selected indicator variable
- Fixing the Variance of the Latent Variable to 1:This standardizes the latent variable, meaning it has a mean of 0 and a variance of 1.

…Actually, you do not have to choose...

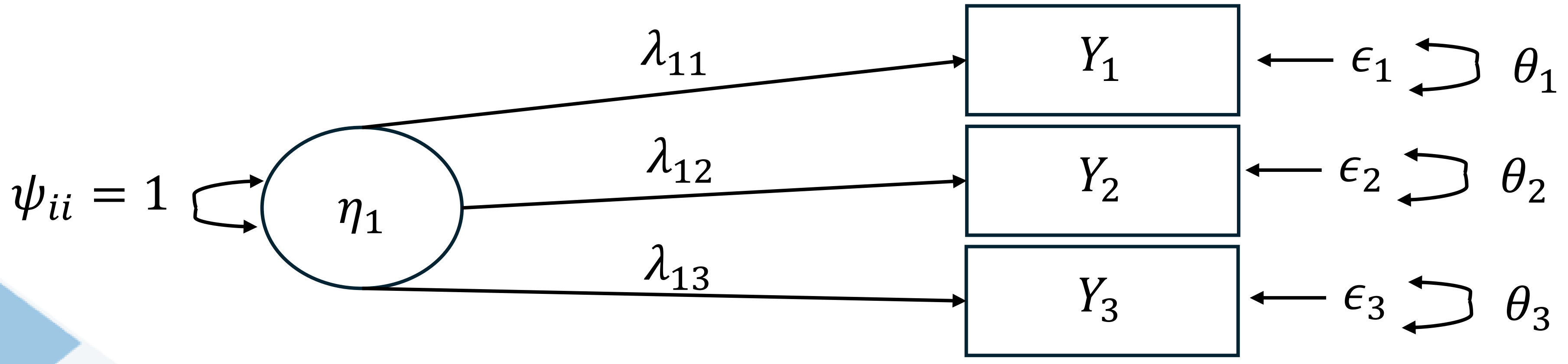
Marker method



$$\Sigma = \begin{bmatrix} 1 \\ \lambda_{12} \\ \lambda_{13} \end{bmatrix} \psi_{11} \begin{bmatrix} 1 & \lambda_{12} & \lambda_{13} \end{bmatrix} + \begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}$$

6 parameters and 6 pieces of information

Variance standardization



$$\Sigma = \begin{bmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \end{bmatrix} 1 \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \end{bmatrix} + \begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}$$

6 parameters and 6 pieces of information

Note:

Technically, you can fix any parameter to any value you want.

Goodness of fit measures (model fit statistics)

Goodness of fit measures (model fit statistics)

We need to assess how well the proposed model matches the observed data

$$\min(\Sigma - S)$$

First recommendation: check the residuals

There are plenty of GoF!

Model Chi-square test

$$\textit{Null Hypothesis}:\ \Sigma(\theta) = \Sigma$$

$$\textit{Alternative Hypothesis}:\ \Sigma(\theta) = \Sigma$$

- This test assesses the discrepancy between the observed covariance matrix and the model-implied covariance matrix.
- A non-significant chi-square indicates a good fit, meaning there is little difference between the observed and predicted covariances.
- A significant chi-square ($p < 0.05$) suggests a poor fit.
- Sensitive to sample size; **large samples may lead to significant chi-square** even for models that fit well.
- Requires multivariate normality.

SRMR

SRMR: Standardized Root Mean Square Residual

Evaluates the "reasonability" of: $(\Sigma - S)$

- Based on the residual correlation matrix (the variance-covariance could be difficult to interpret due to scale).
- Square root of the average difference of elements in the residual correlation matrix.

$$SRMR = \sqrt{\frac{\sum_{i \leq j} r_{ij}^2}{p(p+1)/2}}$$

It ranges from 0 (perfect fit) to 1 (worst fit), though it can go slightly above 1

RMSEA

RMSEA: Root Mean Square Error of Approximation

It considers complexity (penalty for model parsimony)

It ranges from 0 (perfect fit) to 1 (worst fit), though it can go slightly above 1

CFI

CFI: Comparative Fit Index

Takes advantage of a baseline model: the null (independence) model where covariances are zero.

TLI

TLI: Tucker-Lewis Index

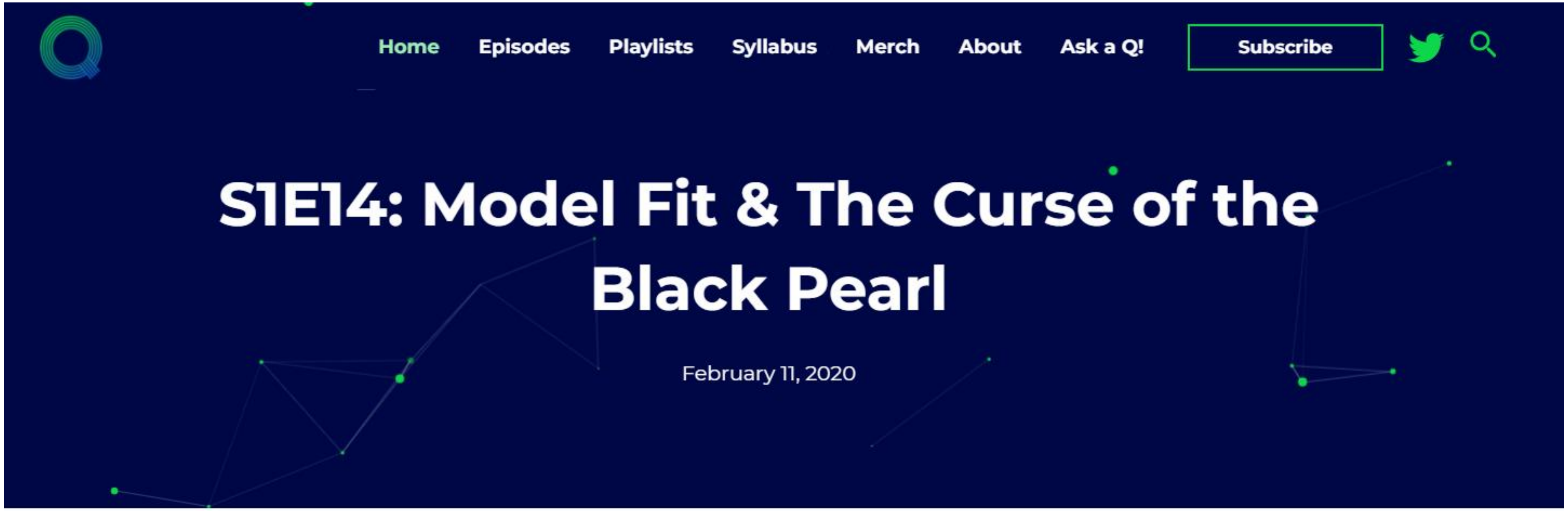
It considers complexity (penalty for model parsimony)

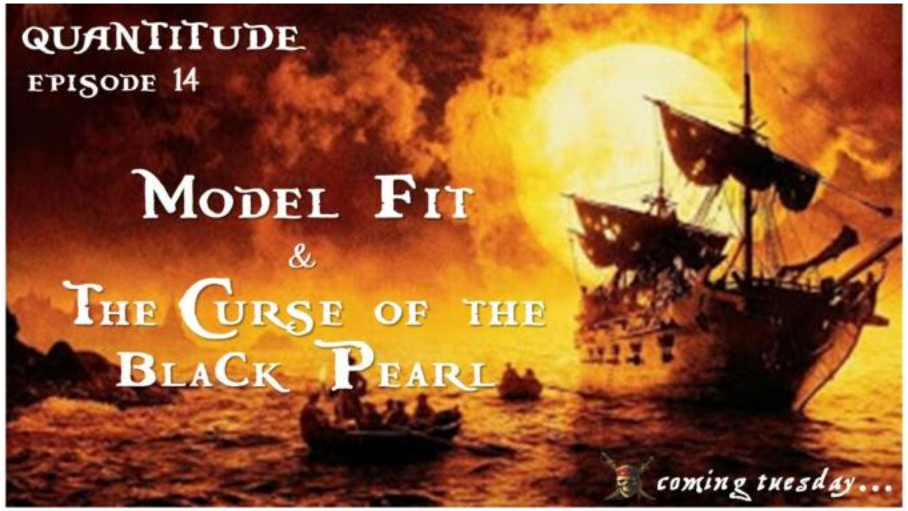
CFI and TLI range from 0 (worst fit) to 1 (perfect fit).

Fit type	Index	Interpretation for guidance
Absolute	RMR/SRMR	≤ 0.08 = good fit
	WRMR	≤ 1.00 = good fit
Parsimonious	PRATIO	Between 0.00 (saturated model) and 1.00 (parsimonious model)
	RMSEA	≤ 0.05 = very good fit ≤ 0.06 and ≤ 0.08 = good fit
	AIC	Comparative index: the lower value of this index, the better the fit
	BIC	Comparative index: the lower value of this index, the better the fit
Incremental	CFI	≥ 0.90 and ≤ 0.94 = good fit ≥ 0.95 = very good fit
	TLI	≥ 0.90 and ≤ 0.94 = good fit ≥ 0.95 = very good fit

Table 1.11. *Some goodness-of-fit indices available in lavaan*

Source: Structural Equation Modeling with lavaan (Kamel Gana & Guillaume Broc, 2019) – Page 43





https://quantitudepod.org/episode-14-model-fit-the-curse-of-the-black-pearl/

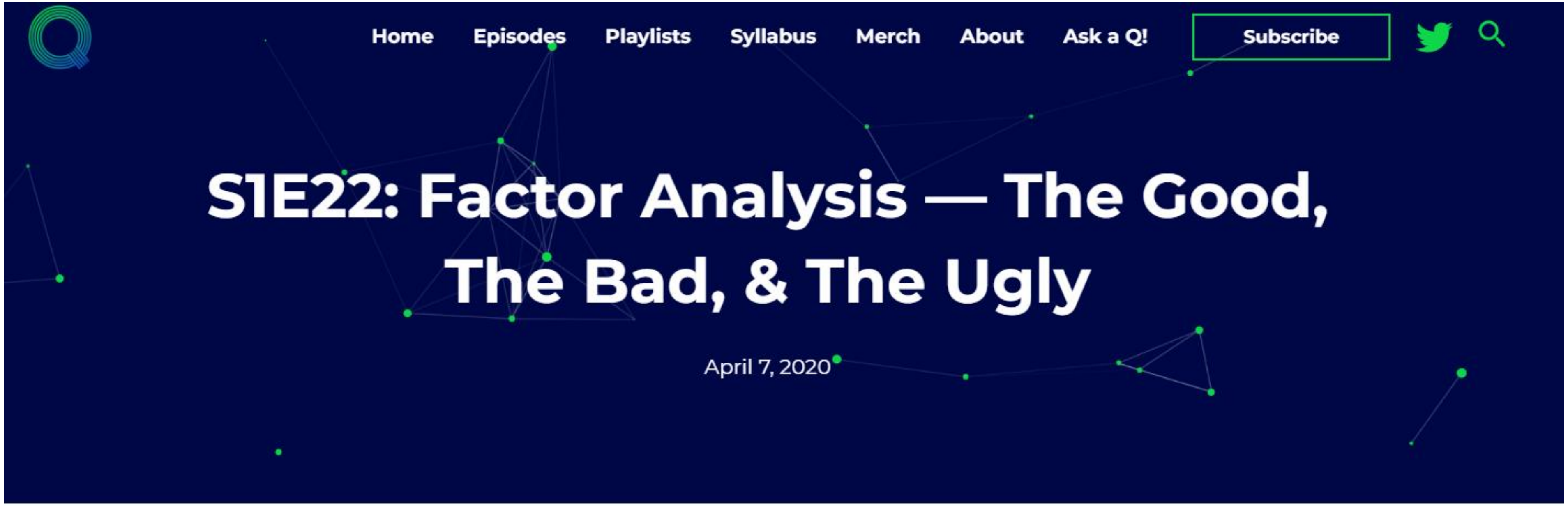
EFA vs CFA

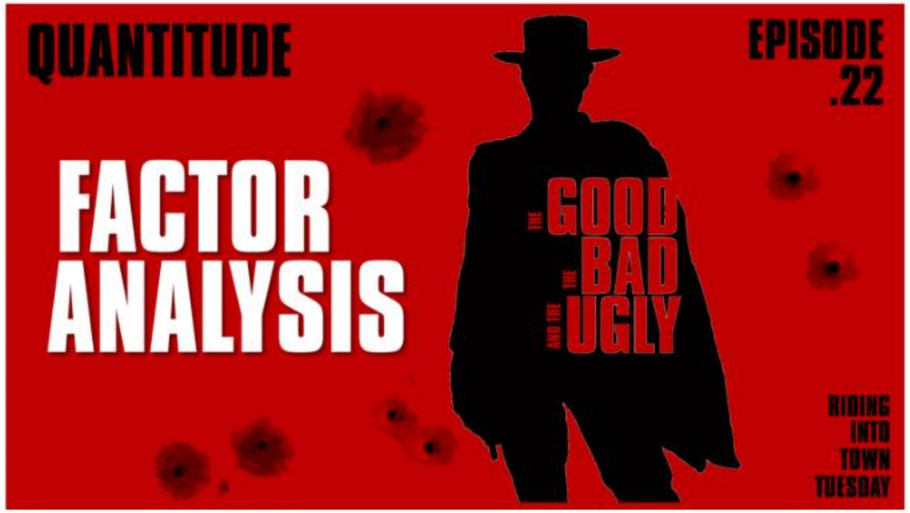
Is EFA a data driven approach?
Is EFA exploratory?
Is CFA confirmatory?
Is EF for scale development and construct validation?

EFA: no prior restrictions
CFA: you must specify several aspects

Danger Zone

Modification indices
Error correlations
Estimation methods





https://quantitudepod.org/episode-22-factor-analysis-the-good-the-bad-the-ugly/

Go to the tutorial!

MAESTRÍA EN INVESTIGACIÓN OPERATIVA Y ESTADÍSTICA

Universidad Tecnológica de Pereira

Thank you!

Orlando Sabogal-Cardona

@Antonio Sabogal

orlando.sabogal.20@ucl.ac.uk