Path Analysis, SEM and the Classical Twin Model

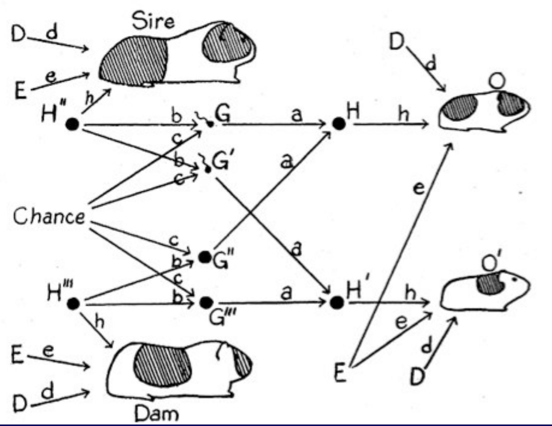
Michael Neale[1] & Frühling Rijsdijk[2]

[1] Virginia Institute for Psychiatric and Behavioral Genetics
Virginia Commonwealth University

[2] MRC SGDP Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London

Aims of Session

1. Introduction to SEM
2. Path Coefficient ACE Model
3. Variance Components ACE Model
4. ADE Model
5. RAM Algebra

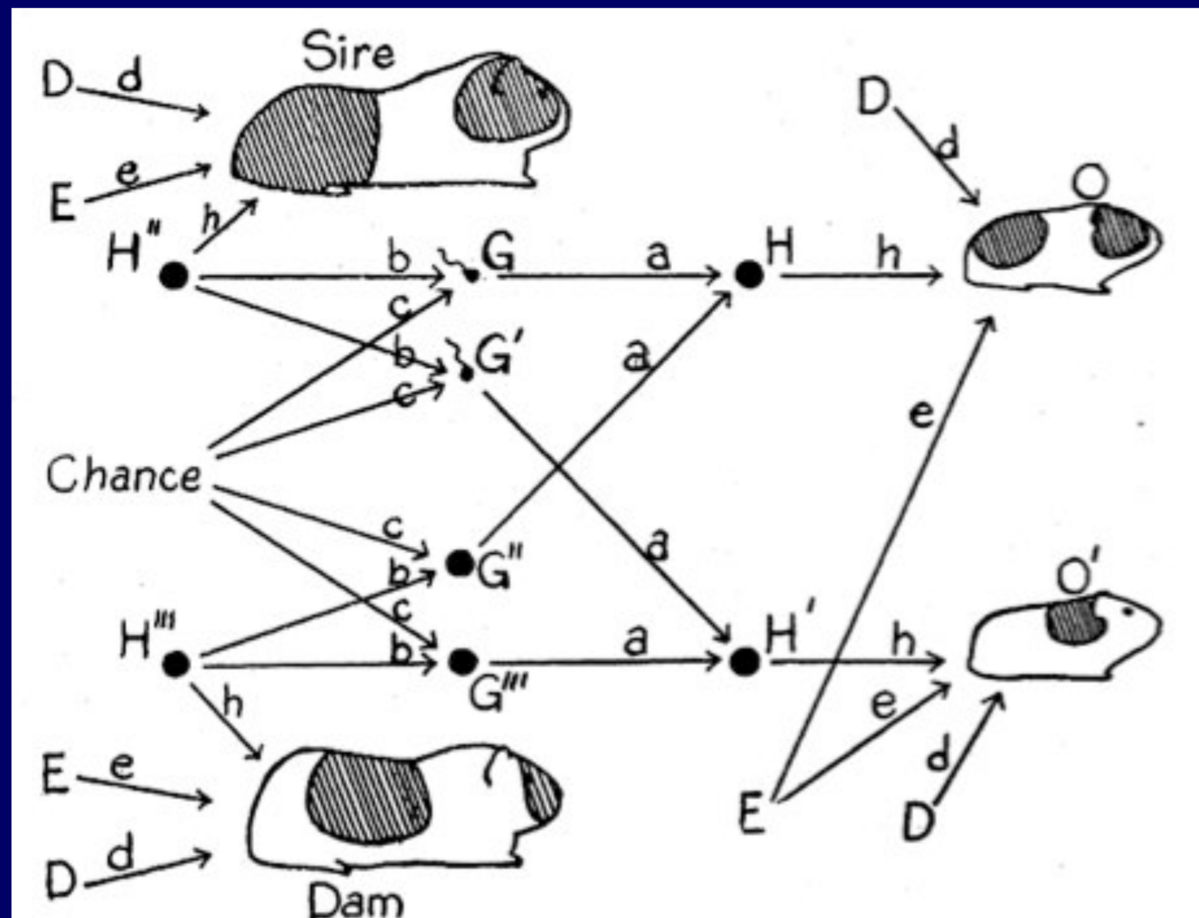


Path Analysis and SEM

- Sewall Wright (1921) PNAS, 6, 320–332
- Causal and correlational relationships between variables in path diagrams
- One-to-one mathematical equivalence with simple matrix algebra expression

- Structural equation modelling (SEM) is a unified platform for path analysis, regression, factor and variance components models

Path Analysis and SEM



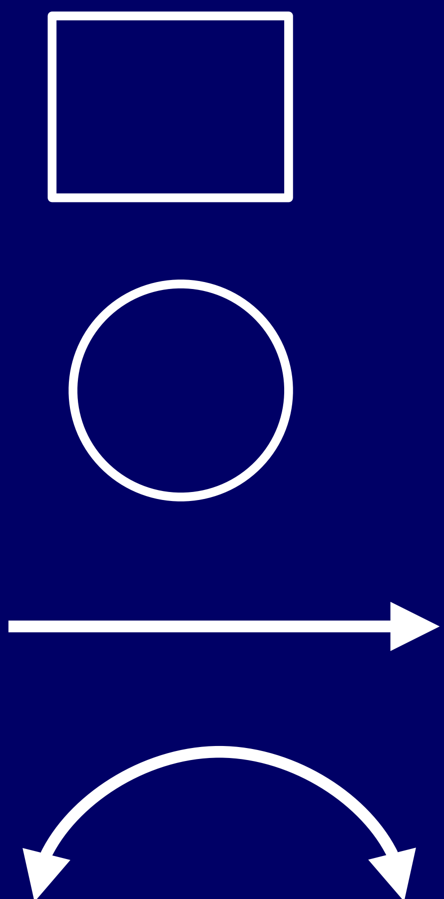
OpenMx Software

R package

Open Source

Since 1990

Path Diagram Conventions



Observed Variables

Latent Variables

Causal Paths

Covariance Paths

Tracing Rules of Path Analysis

1 Find All Distinct Chains between Variables:

a Go backwards along *zero or more* single-headed arrows

b Change direction at *one and only one* Double-headed arrow

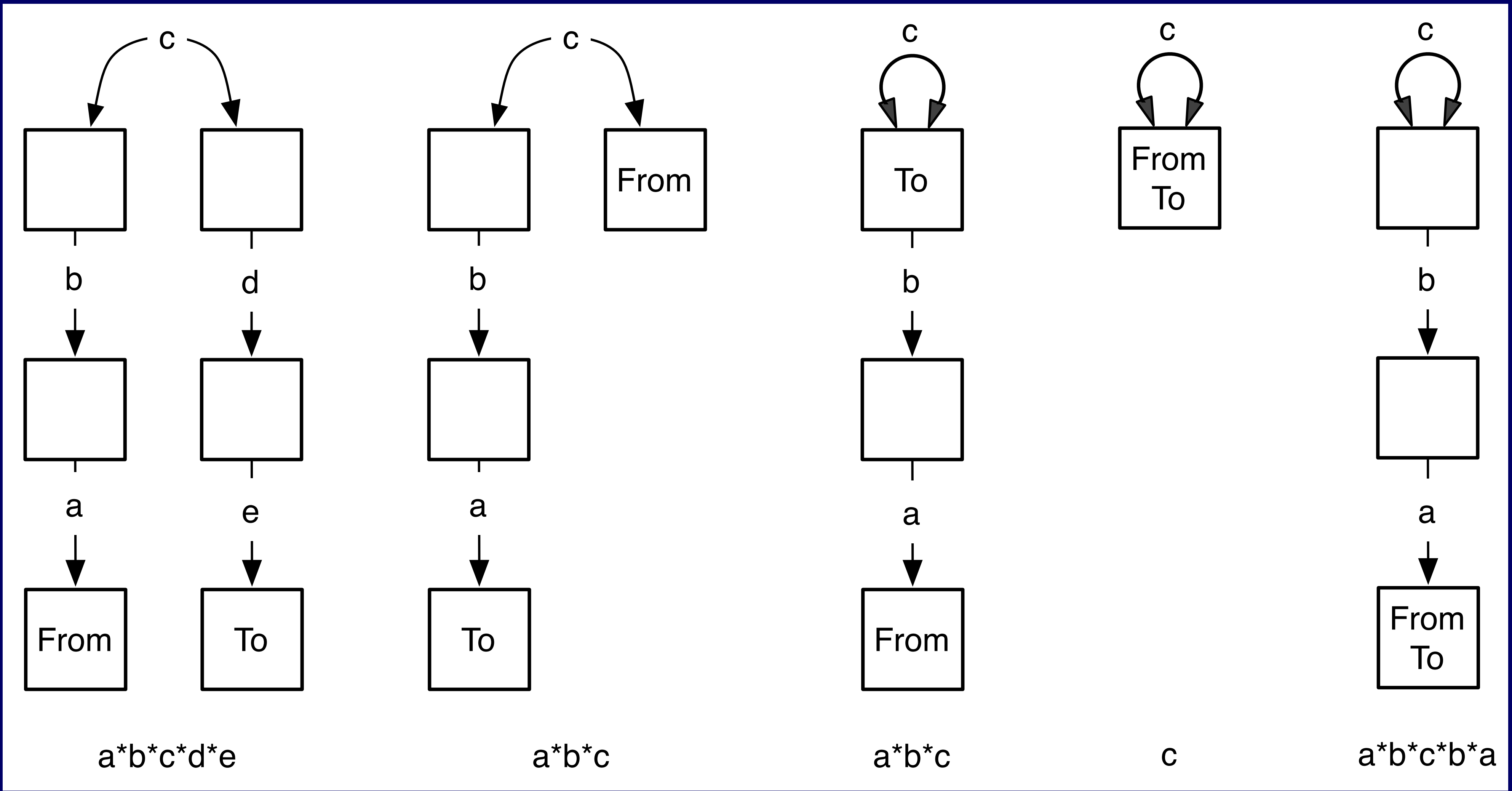
c Trace forwards along *zero or more* Single-headed arrows

2 Multiply path coefficients in a chain

3 Sum the results of step 2.

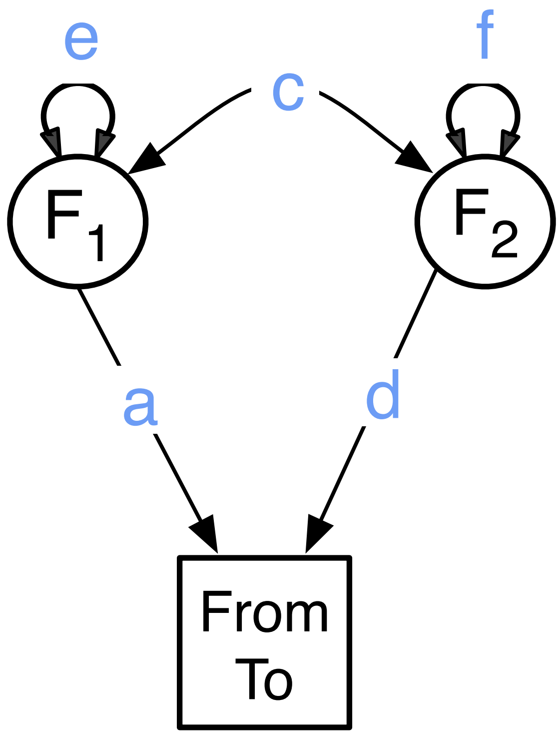
For covariance of a variable with itself (Variance), chains are distinct if they have different paths *or* a different order

Chain Examples I



Thou shalt not pass through adjacent arrowheads

Chain Examples II



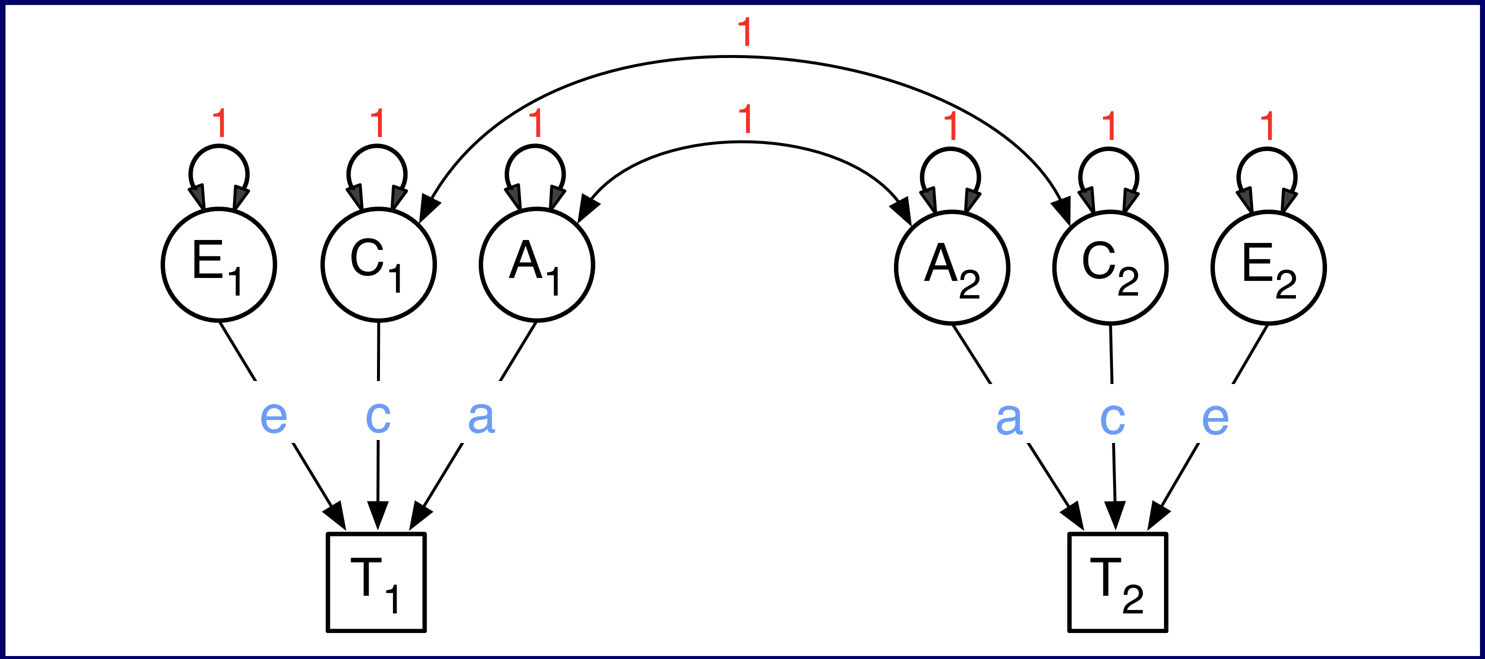
$$a*e*a + d*f*d + a*c*d + \underline{d*c*a}$$

Variance: Chains in Different Order Count

Path Diagrams for the Classical Twin Model

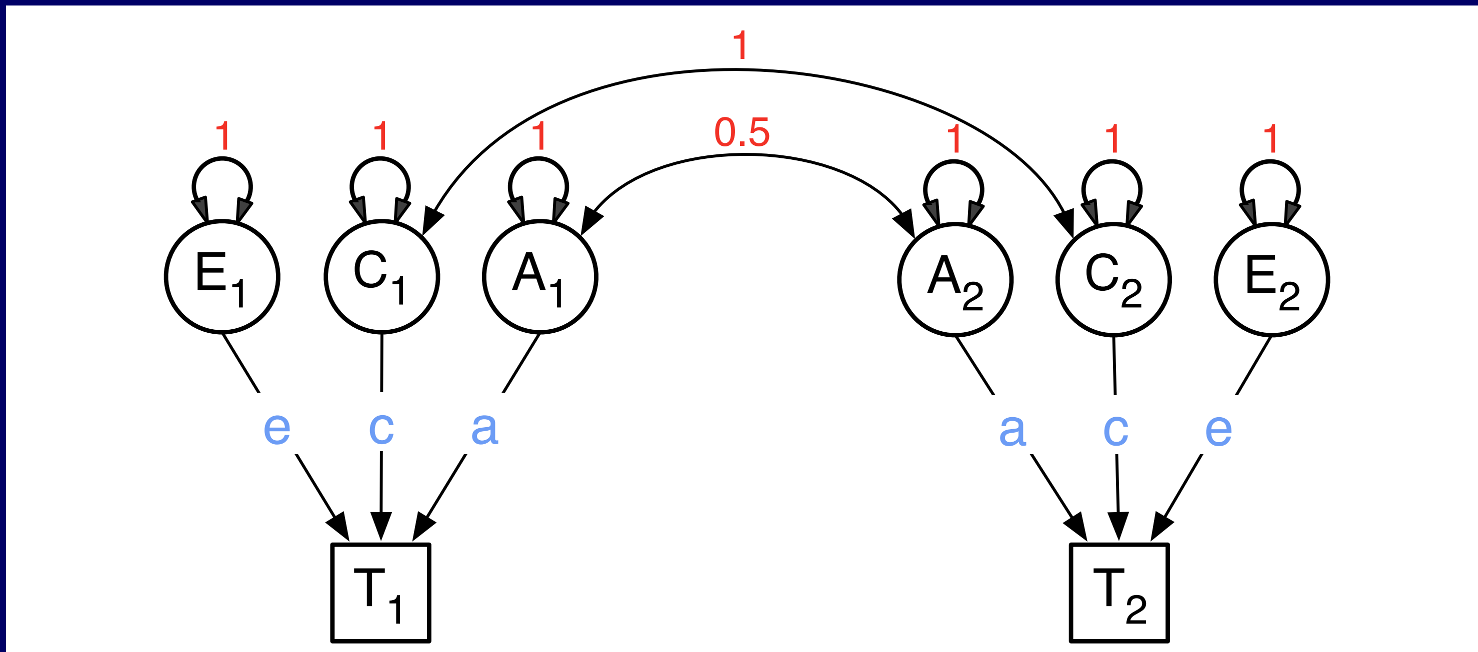
Part 1: Path Coefficients

Path Model for an MZ Pair



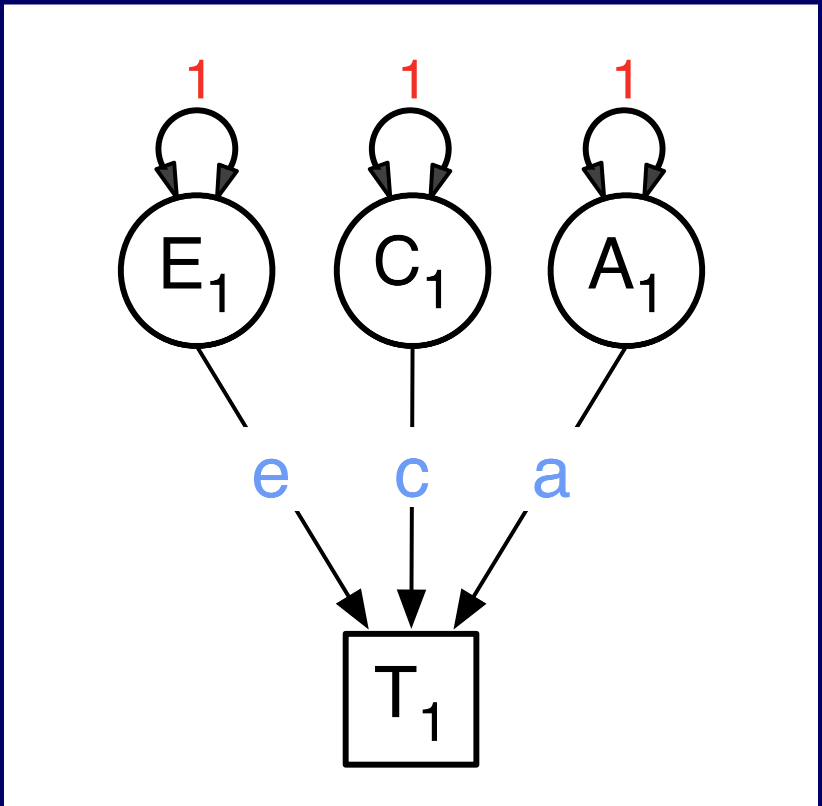
Latent variables A1 C1 and E1 have variance 1, and cause phenotype T1 via path coefficients a, c and e.
Same model for T2. $Cov(A_1,A_2)=1$

Path Model for a DZ Pair



Latent variables A1 C1 and E1 have variance 1, and cause phenotype T1 via regression paths a, c and e.
Same model for T2. Cov(A1,A2) = .5

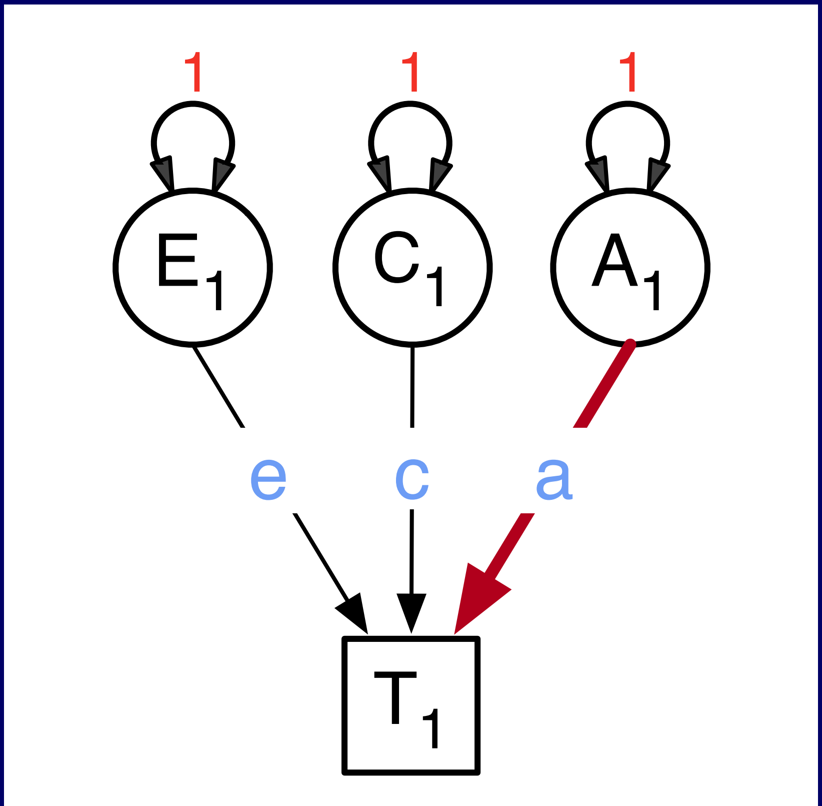
Variance of Twin 1 AND Twin 2 (for MZ and DZ pairs)



What Chains?

Total Variance = $a^2 + c^2 + e^2$

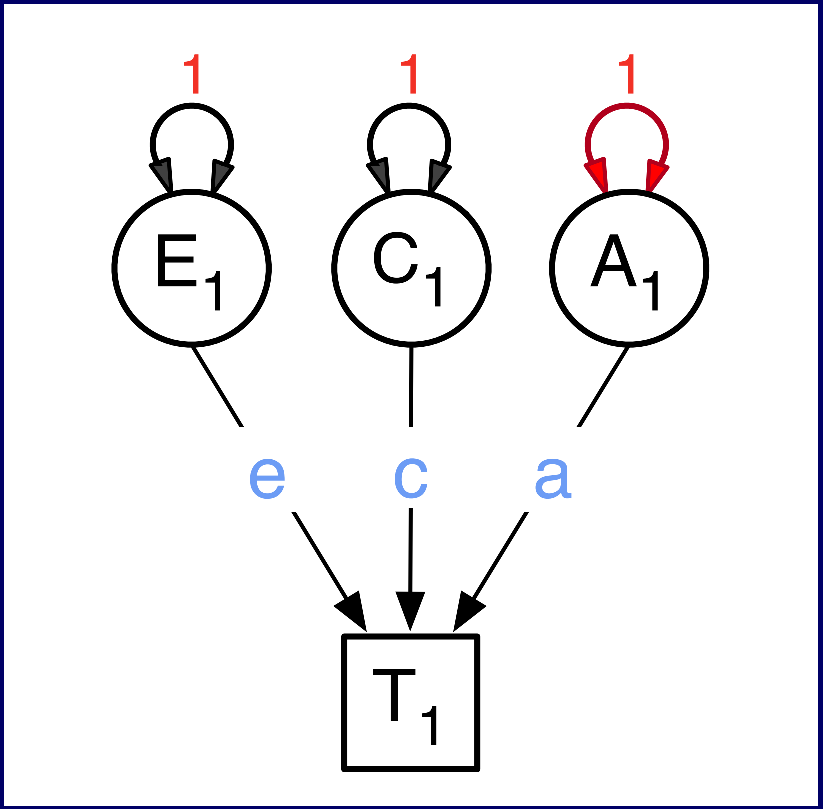
Variance of Twin 1 AND Twin 2 (for MZ and DZ pairs)



a* =

Total Variance = $a^2 + c^2 + e^2$

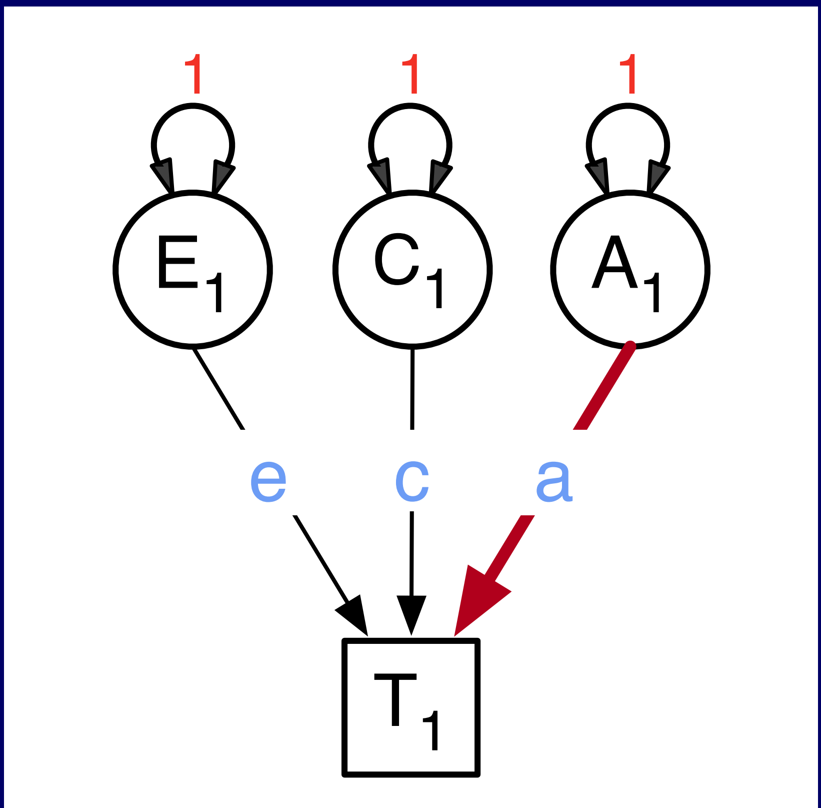
Variance of Twin 1 AND Twin 2 (for MZ and DZ pairs)



a*1

Total Variance = $a^2 + c^2 + e^2$

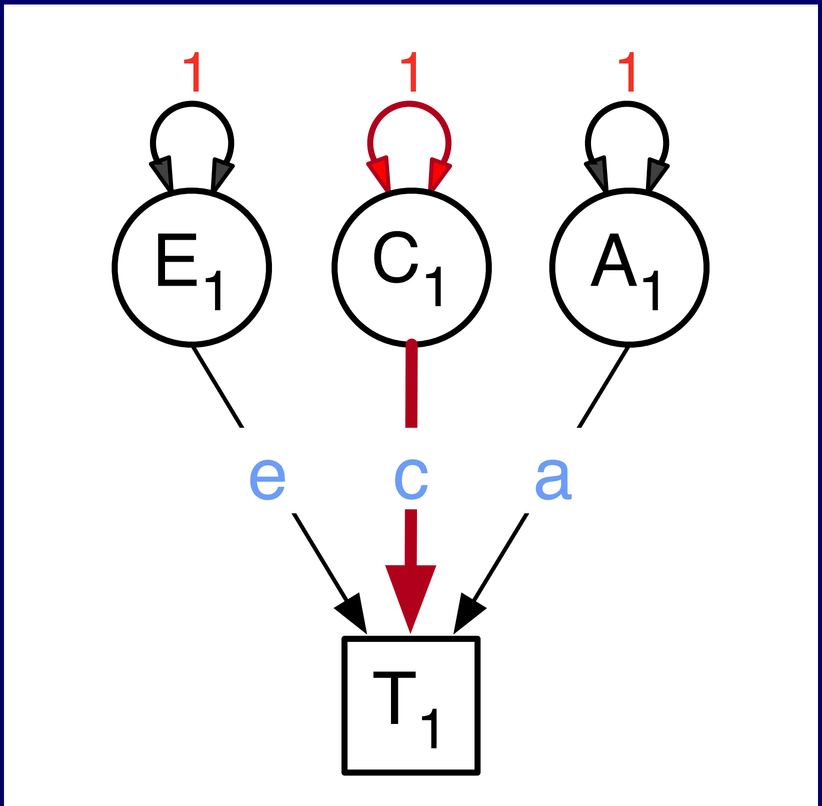
Variance of Twin 1 AND Twin 2 (for MZ and DZ pairs)



$a*1*a = a^2$

Total Variance = $a^2 + \ldots$

Variance of Twin 1 AND Twin 2 (for MZ and DZ pairs)



$a*1*a = a^2$

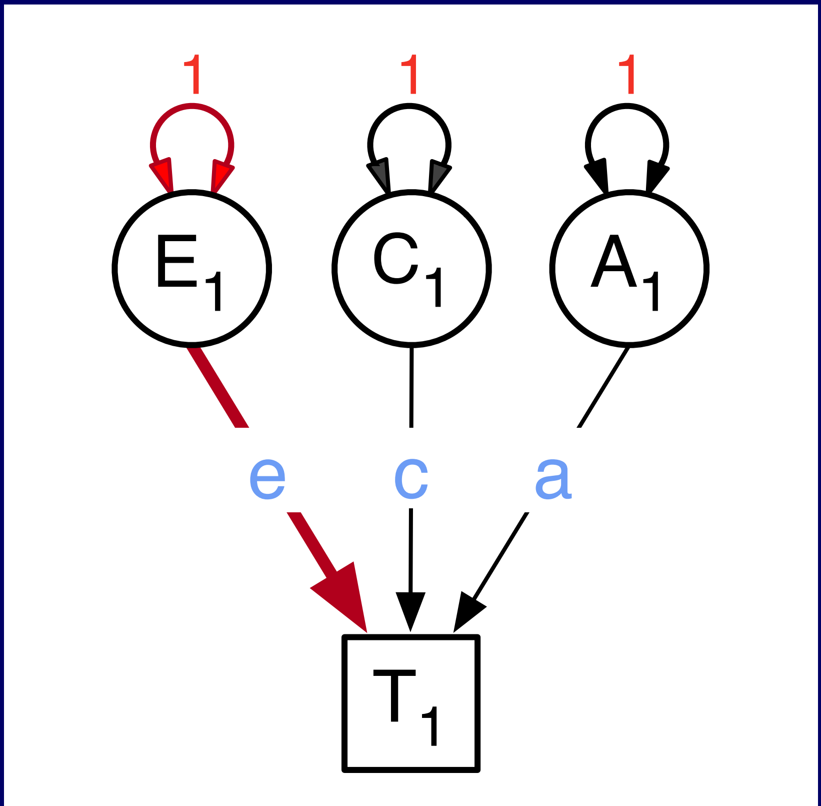
$+$

$c*1*c = c^2$

$+$

Total Variance = $a^2 + c^2 + \ldots$

Variance of Twin 1 AND Twin 2 (for MZ and DZ pairs)



$a*1*a = a^2$

$+$

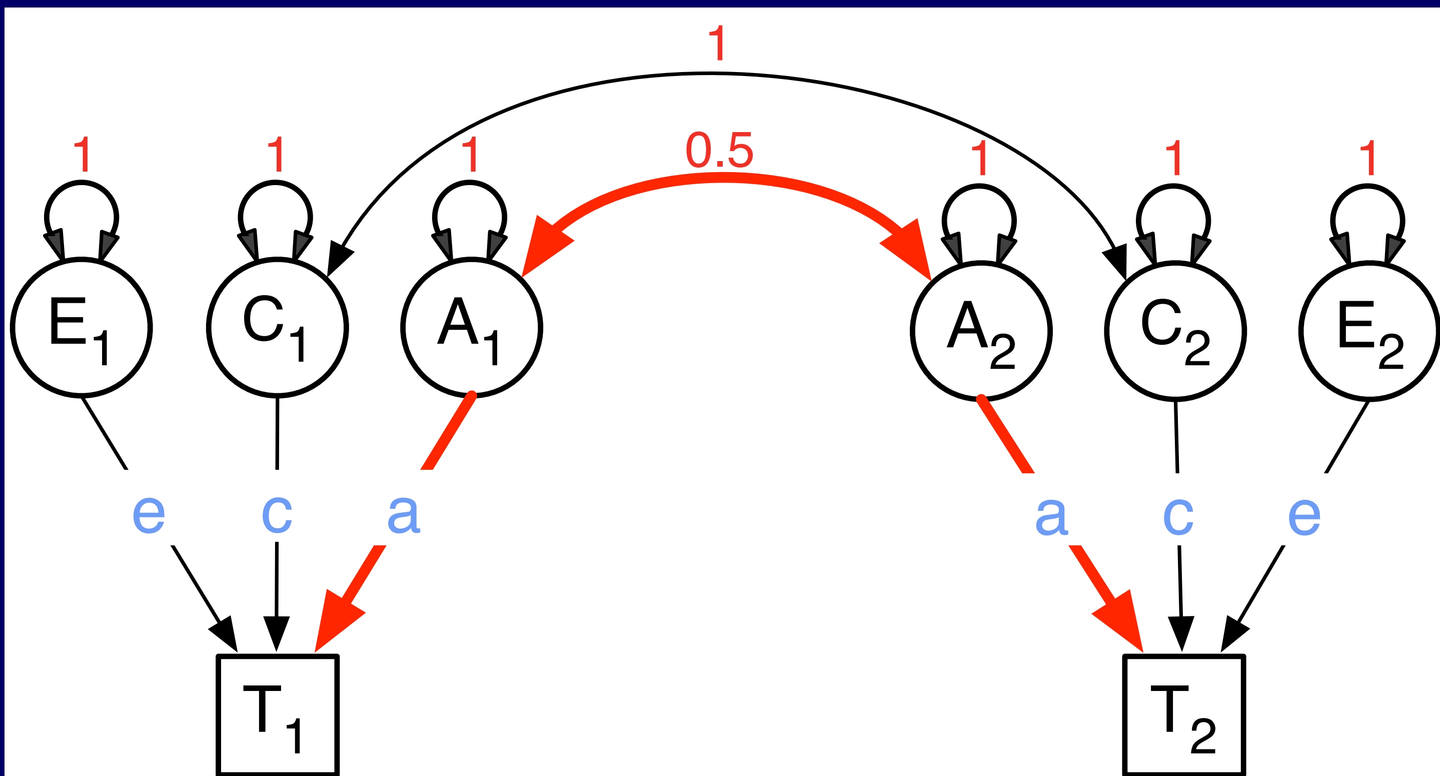
$c*1*c = c^2$

$+$

$e*1*e = e^2$

Total Variance = $a^2 + c^2 + e^2$

Covariance of Twin 1 AND Twin 2 (for DZ pairs)

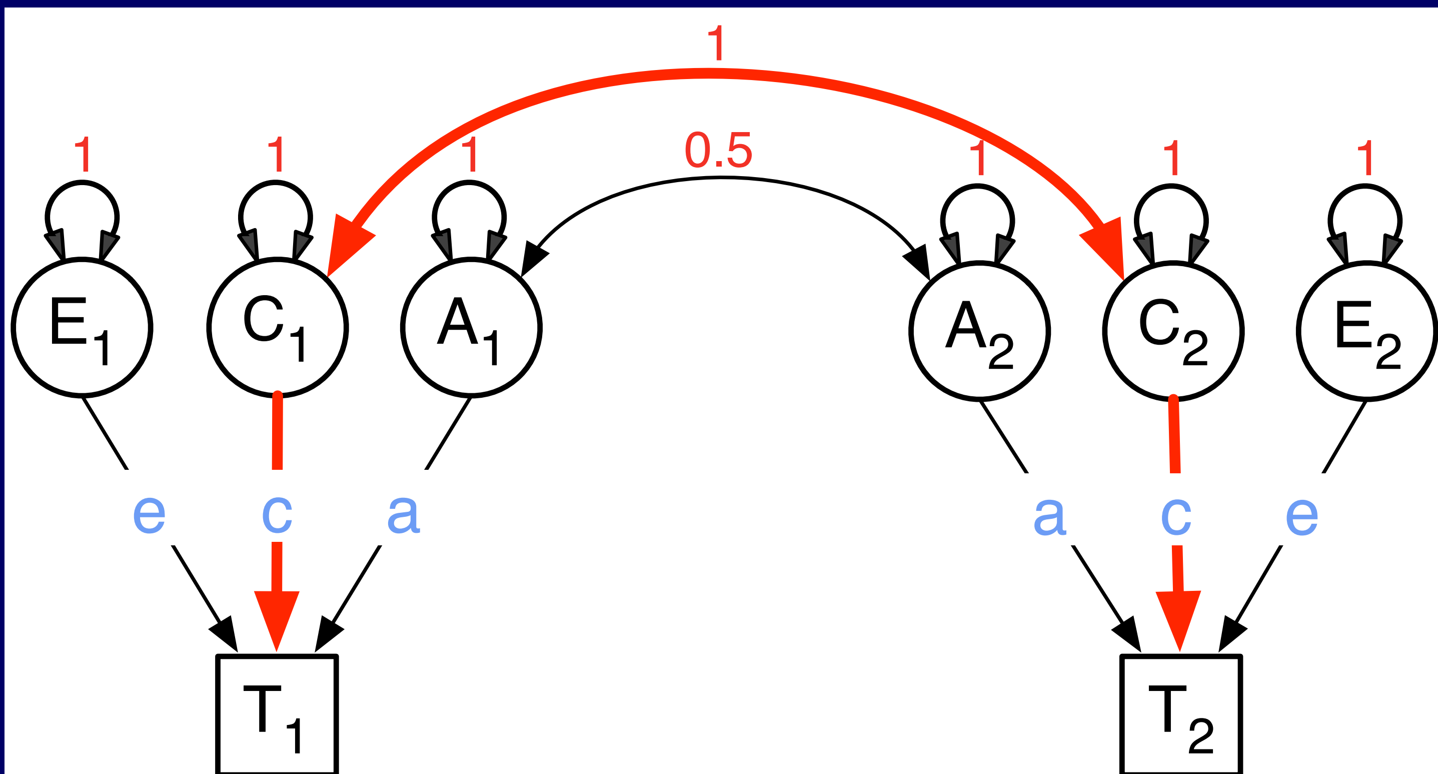


$a^*.5^*a = .5a^2$

$+$

Covariance = $.5a^2 + \ldots$

Covariance of Twin 1 AND Twin 2 (for DZ pairs)



$a^*.5^*a = .5a^2$

+

$c^*\ 1^*c = c^2$

Total Covariance = $.5a^2 + c^2$

Predicted Variance-Covariance Matrices ACE Path Model

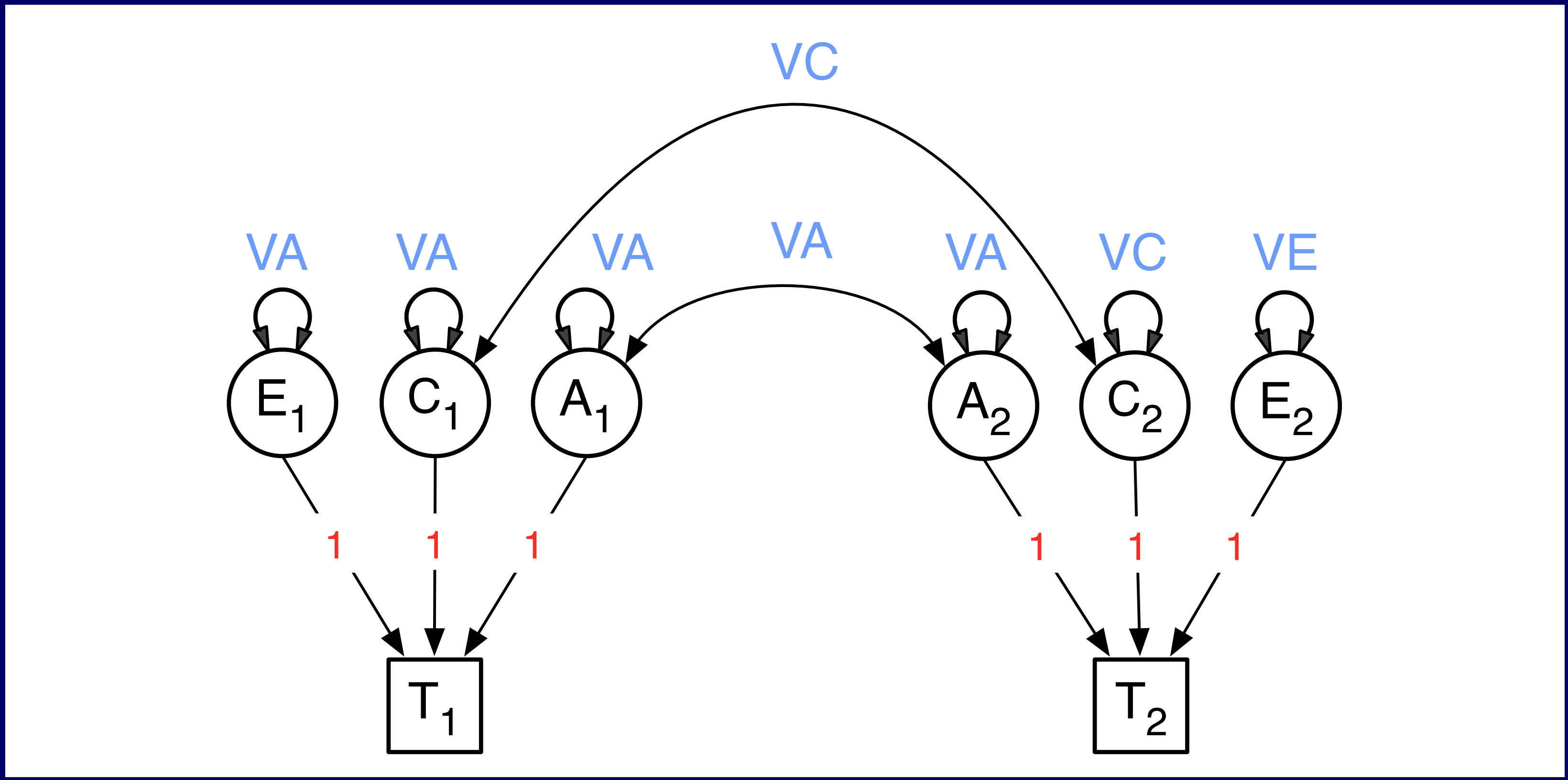
$$\text{Cov MZ} \quad \begin{array}{c|cc} & \text{Tw1} & \text{Tw2} \\ \hline \text{Tw1} & & \\ \text{Tw2} & & \end{array} \begin{pmatrix} a^2+c^2+e^2 & a^2+c^2 \\ a^2+c^2 & a^2+c^2+e^2 \end{pmatrix}$$

$$\text{Cov DZ} \quad \begin{array}{c|cc} & \text{Tw1} & \text{Tw2} \\ \hline \text{Tw1} & & \\ \text{Tw2} & & \end{array} \begin{pmatrix} a^2+c^2+e^2 & \frac{1}{2}a^2+c^2 \\ \frac{1}{2}a^2+c^2 & a^2+c^2+e^2 \end{pmatrix}$$

Path Diagrams for the Classical Twin Model

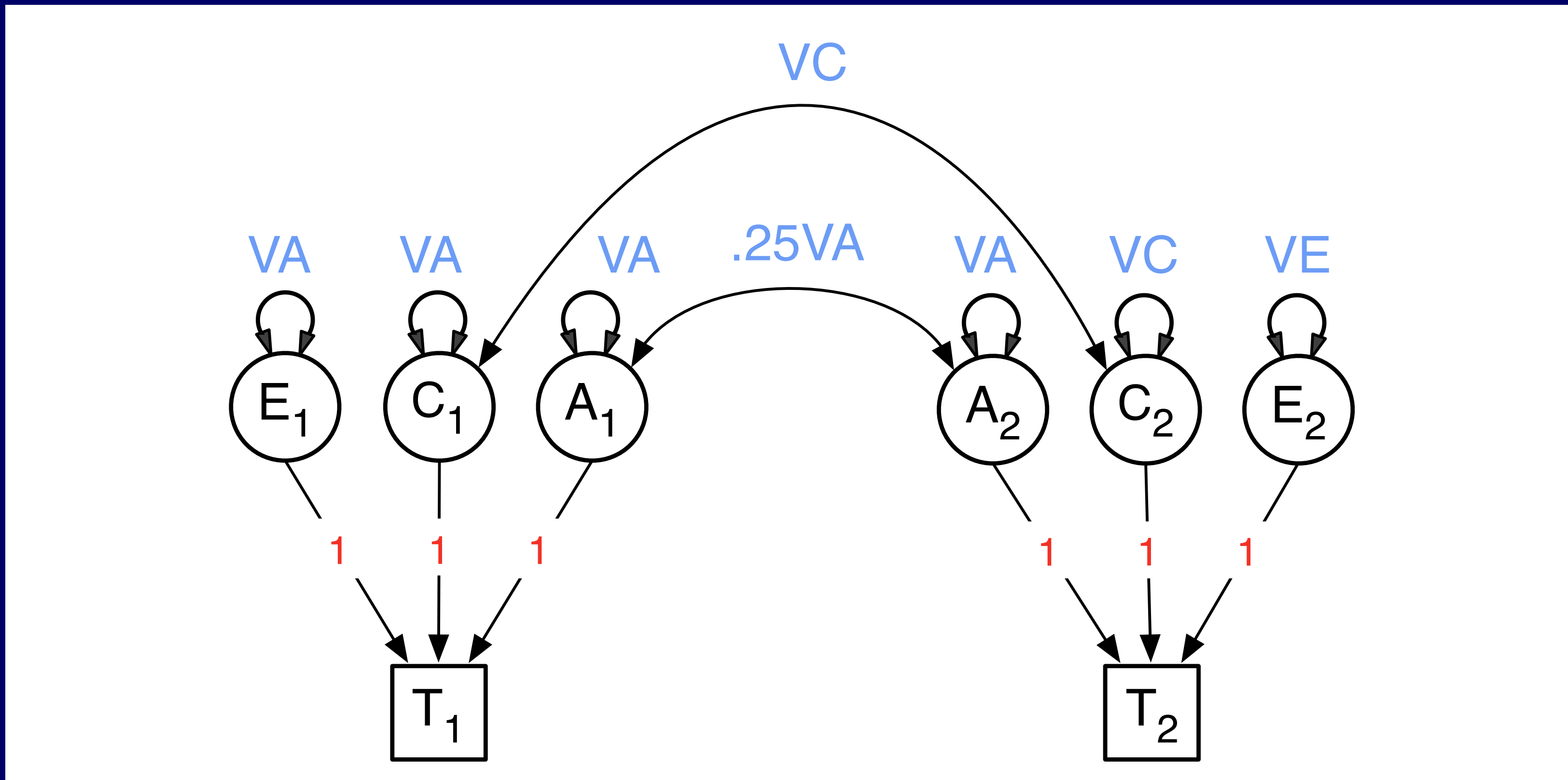
Part 2: Variance Components

Variance Component Model: MZ



Latent variables A1 C1 and E1 have variances VA, VC and VE, and cause phenotype T1 via regression paths 1. Same model for T2

Variance Component Model: DZ



Latent variables A1 C1 and E1 have variances VA, VC and VE, and cause phenotype T1 via regression paths 1. Same model for T2

Predicted Variance-Covariance Matrices ACE VC Model

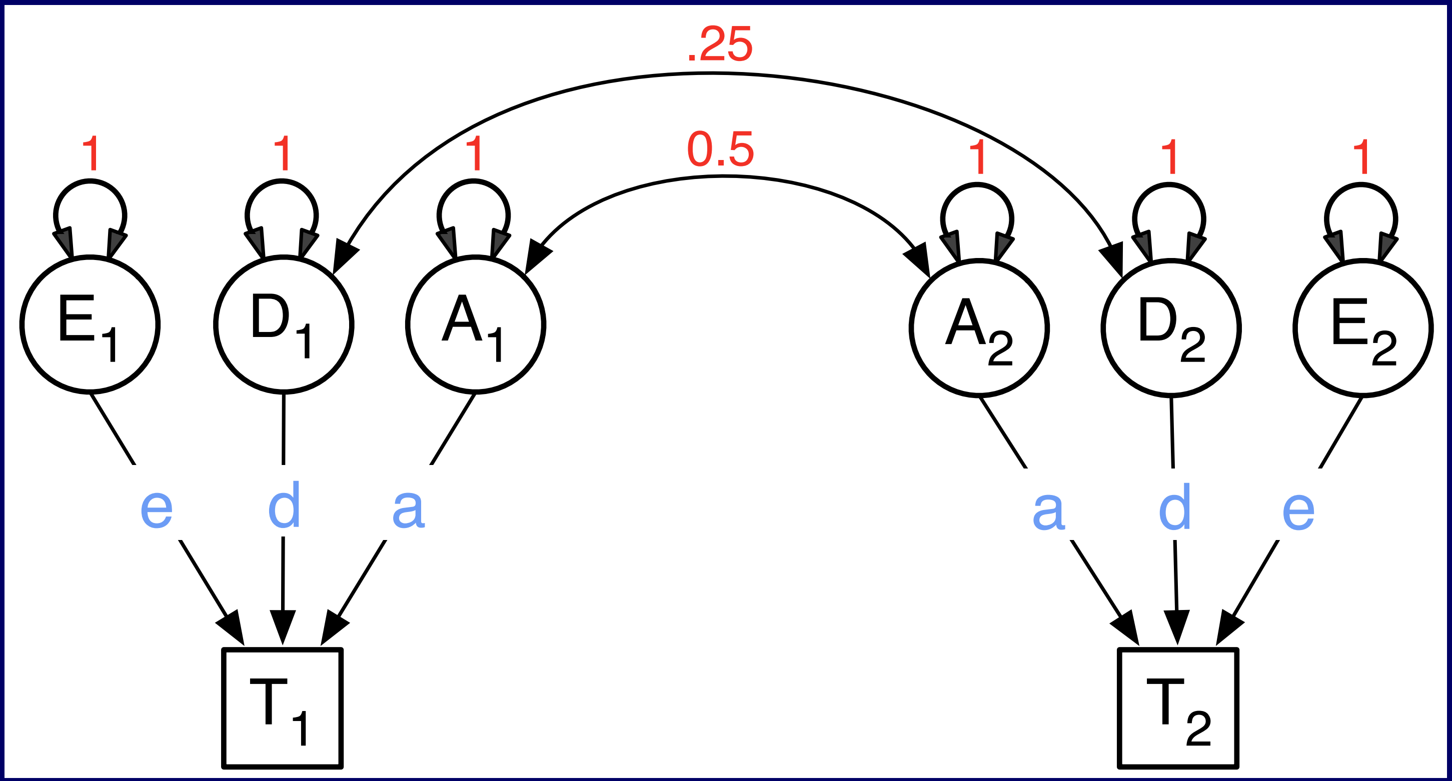
$$
\text{Cov MZ} \quad
\begin{array}{c c}
 & \begin{array}{cc} \text{Tw1} & \text{Tw2} \end{array} \\
\begin{array}{c} \text{Tw1} \\ \text{Tw2} \end{array} &
\begin{pmatrix} V_A+V_C+V_E & V_A+V_C \\ V_A+V_C & V_A+V_C+V_E \end{pmatrix}
\end{array}
$$

$$
\text{Cov DZ} \quad
\begin{array}{c c}
 & \begin{array}{cc} \text{Tw1} & \text{Tw2} \end{array} \\
\begin{array}{c} \text{Tw1} \\ \text{Tw2} \end{array} &
\begin{pmatrix} V_A+V_C+V_E & .5V_A+V_C \\ .5V_A+V_C & V_A+V_C+V_E \end{pmatrix}
\end{array}
$$

What's the Difference?

- Path: Implicit Boundary Constraint
  - Estimate a but a^2 *never* negative
- Variance Component: Unbounded
  - Estimates VA, VC and VE can be positive or negative
- Variance Component may fit better
  - No bias from implicit boundary
- Negative Variances? Model wrong?

ADE Path Coefficient Model
DZ pairs



MZ Covariance = $a^2 + d^2$

DZ Covariance = $.5a^2 + .25d^2$

Total Variance = $a^2 + d^2 + e^2$

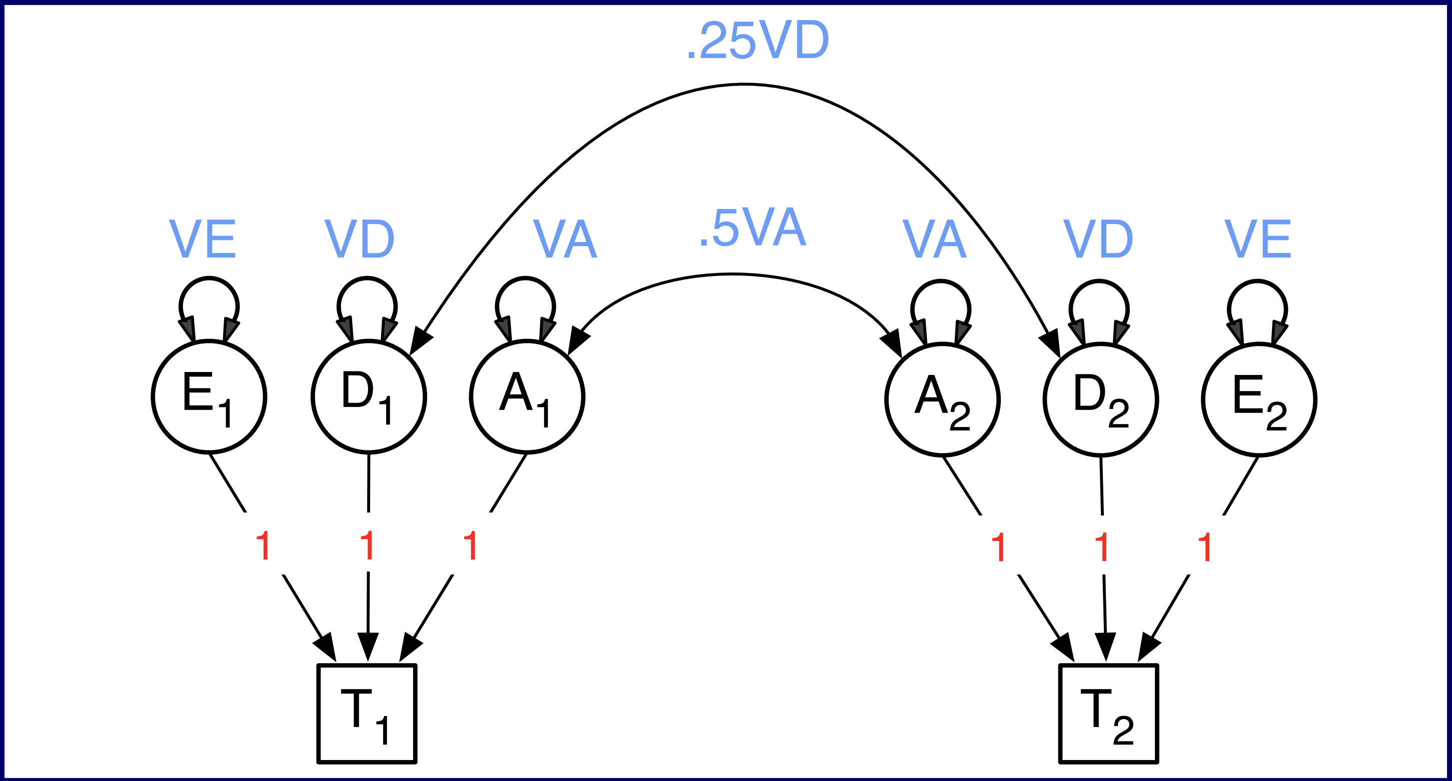
Predicted Var-Cov Matrices
ADE Model

$$\text{Cov MZ} \quad \begin{array}{c|cc} & \text{Tw1} & \text{Tw2} \\ \hline \text{Tw1} & \multicolumn{2}{c}{\multirow{2}{*}{}} \\ \text{Tw2} & & \end{array}$$

$$\text{Cov MZ} = \begin{array}{cc} & \begin{array}{cc} \text{Tw1} & \text{Tw2} \end{array} \\ \begin{array}{c} \text{Tw1} \\ \text{Tw2} \end{array} & \begin{pmatrix} a^2+d^2+e^2 & a^2+d^2 \\ a^2+d^2 & a^2+d^2+e^2 \end{pmatrix} \end{array}$$

$$\text{Cov DZ} = \begin{array}{cc} & \begin{array}{cc} \text{Tw1} & \text{Tw2} \end{array} \\ \begin{array}{c} \text{Tw1} \\ \text{Tw2} \end{array} & \begin{pmatrix} a^2+d^2+e^2 & \frac{1}{2}a^2+\frac{1}{4}d^2 \\ \frac{1}{2}a^2+\frac{1}{4}d^2 & a^2+d^2+e^2 \end{pmatrix} \end{array}$$

ADE Variance Component Model
DZ pairs



MZ Covariance = VA + VD

DZ Covariance = .5VA + .25VD

Total Variance = VA + VD + VE

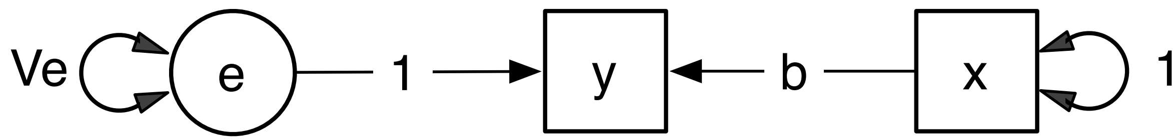
Predicted Variance-Covariance Matrices ADE VC Model

$$\text{Cov MZ} \quad \begin{array}{c c} & \begin{array}{cc} \text{Tw1} & \text{Tw2} \end{array} \\ \begin{array}{c} \text{Tw1} \\ \text{Tw2} \end{array} & \begin{pmatrix} V_A+V_D+V_E & V_A+V_D \\ V_A+V_D & V_A+V_D+V_E \end{pmatrix} \end{array}$$

$$\text{Cov DZ} \quad \begin{array}{c c} & \begin{array}{cc} \text{Tw1} & \text{Tw2} \end{array} \\ \begin{array}{c} \text{Tw1} \\ \text{Tw2} \end{array} & \begin{pmatrix} V_A+V_D+V_E & .5V_A+.25V_D \\ .5V_A+.25V_D & V_A+V_D+V_E \end{pmatrix} \end{array}$$

One-to-one Translation to Matrices

RAM Algebra: Standardized Univariate Regression



Asymmetric Arrows $\quad \mathbf{A} = \begin{array}{c} \\ x \\ y \\ e \end{array}\begin{array}{c} \begin{array}{ccc} x & y & e \end{array} \\ \begin{bmatrix} 0 & 0 & 0 \\ b & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$

Symmetric Slings $\quad \mathbf{S} = \begin{array}{c} x \\ y \\ e \end{array}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & V_e \end{bmatrix}$

Funky Filter $\quad \mathbf{F} = \begin{array}{c} x \\ y \end{array}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

One-to-one Translation to Matrices

Completely General RAM Algebra
Expected Covariance Matrix

$$F * (I - A)^{-1} * S * (I - A)^{-1'} * F'$$

Thank you:
Jack McArdle & Steve Boker
Workshop Faculty & Students & NIH
Also see: http://onyx.brandmaier.de for path model drawing software