Bivariate Analysis of Individual and Pooled Data on Social Interaction Traits

Katrijn Peeters, Jeroen Visscher, Piter Bijma

Data collection

Cheap/Easy

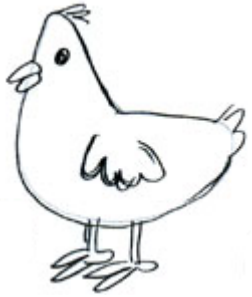
Expensive/Difficult

Individual data

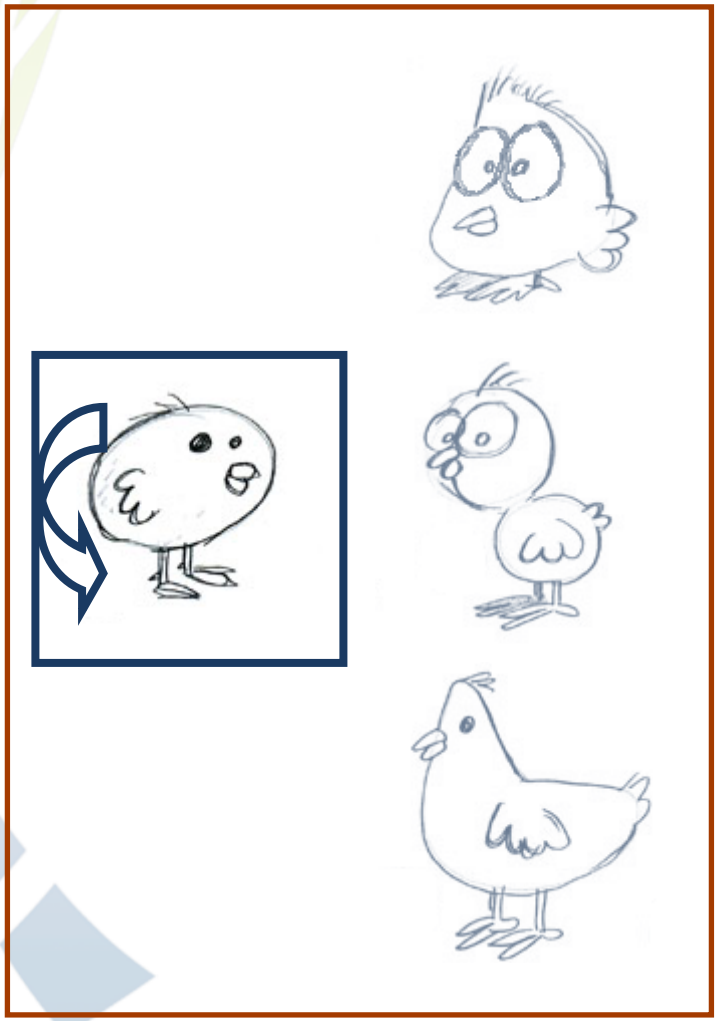
Pooled data

Social interactions

PECK ME
PECK ME
…



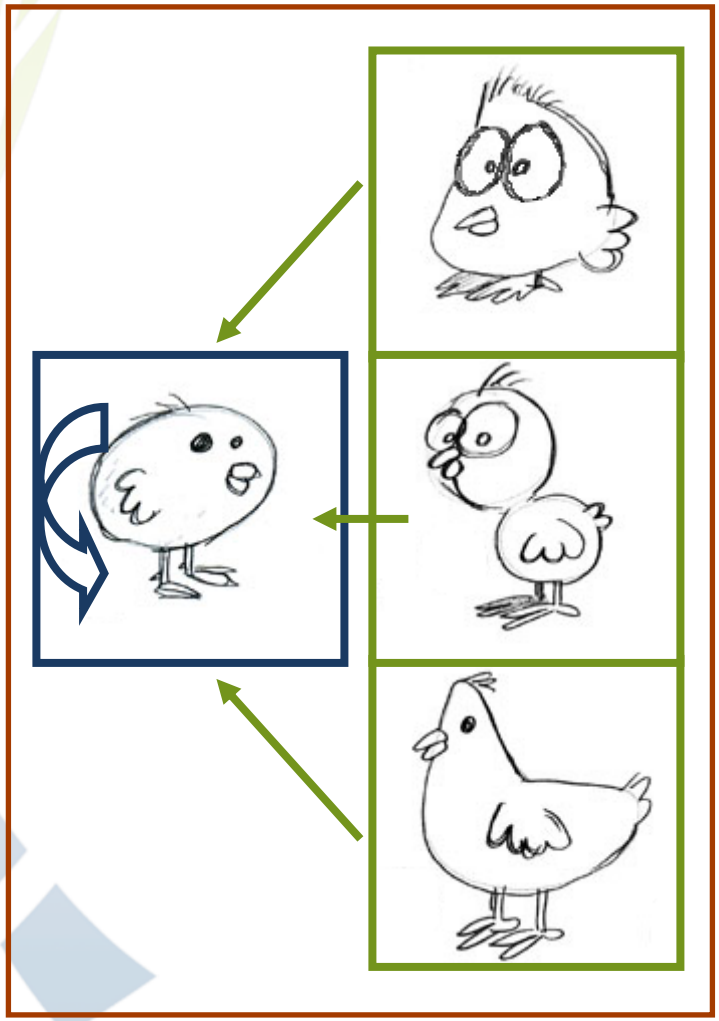
Social interactions theory



Direct

$$P_i = A_{\mathrm{D}_i} + E_{\mathrm{D}_i}$$

Social interactions theory

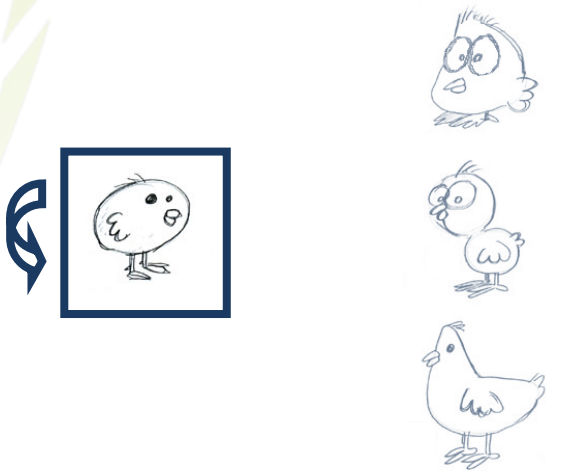


Direct

Indirect

$$P_i = A_{\mathrm{D}_i} + E_{\mathrm{D}_i} + \sum_{i \neq j}^{n-1} A_{\mathrm{I}_j} + \sum_{i \neq j}^{n-1} E_{\mathrm{I}_j}$$

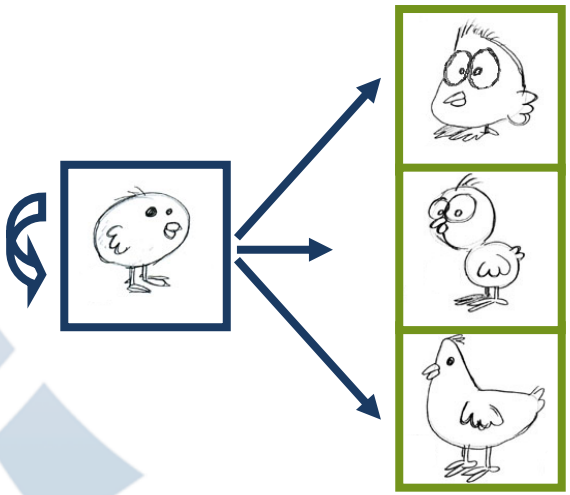
Direct model⇔ Direct-indirect model



Direct

$$A_D$$

$$\sigma^2_{A_D}$$



Direct-indirect

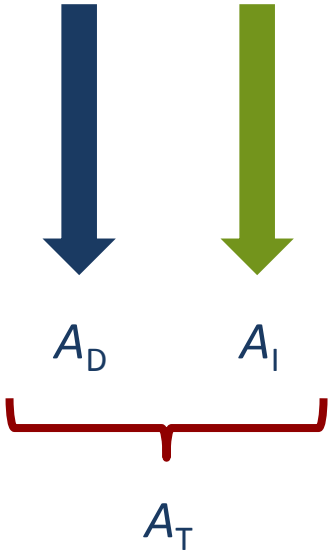
$$A_T = A_D + 3A_I$$

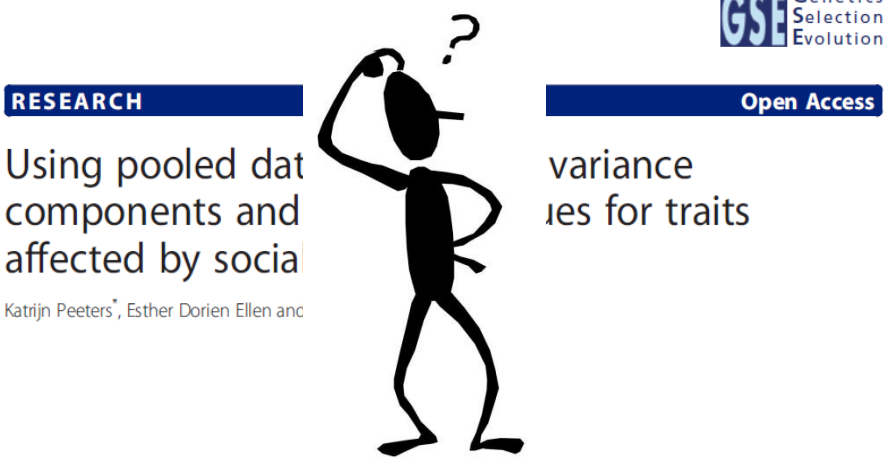
$$\sigma^2_{A_T} = \sigma^2_{A_D} + 6\sigma_{A_{DI}} + 9\sigma^2_{A_I}$$

Direct, indirect and total BV ~ Data

Individual data

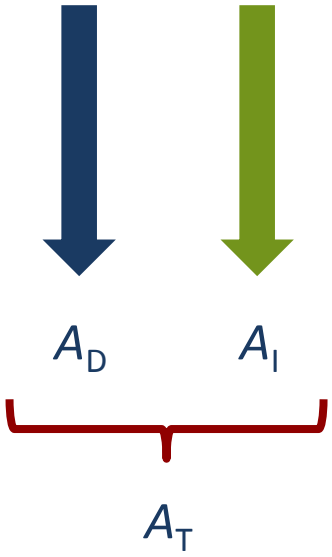
Pooled data



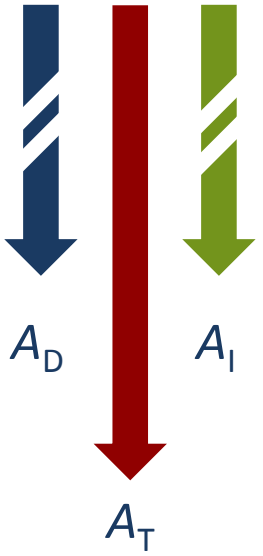


Direct, indirect and total BV ~ Data

Individual data

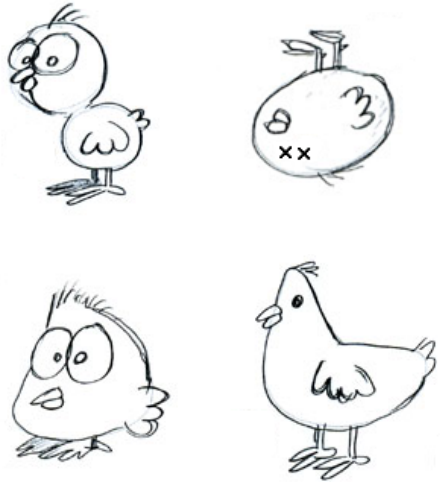


Pooled data

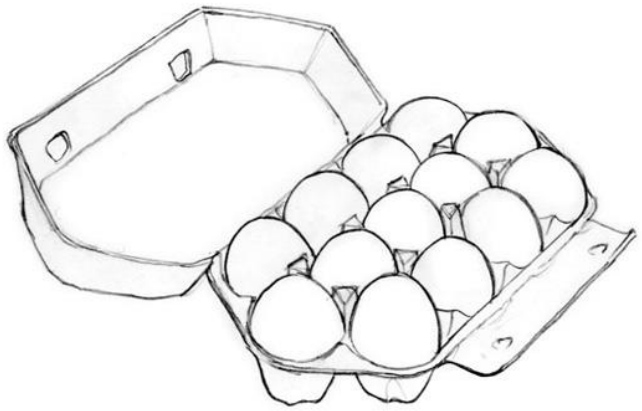


Research question

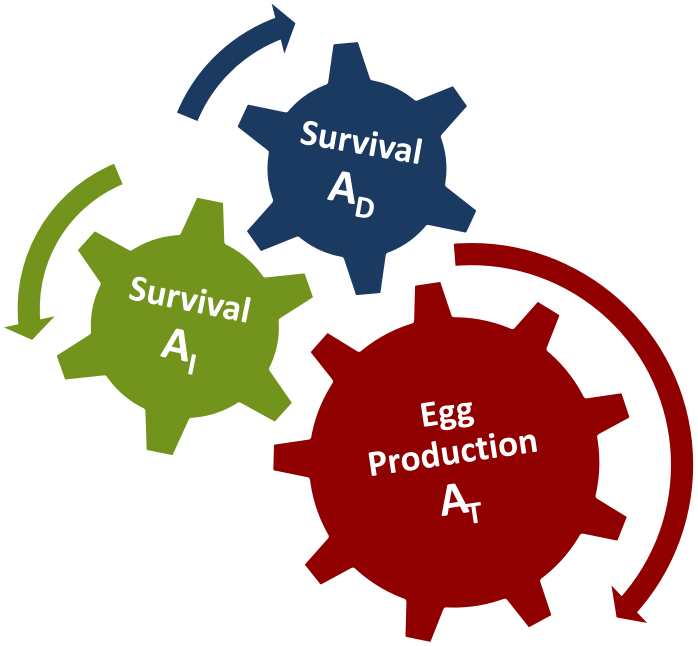
Individual data



Pooled data



Research question



Univariate model

Individual data

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_{\mathrm{D}}\mathbf{a}_{\mathrm{D}} + \mathbf{Z}_{\mathrm{I}}\mathbf{a}_{\mathrm{I}} + \mathbf{Vcage} + \mathbf{e}$$

Pooled data

$$\mathbf{y}^{*} = \mathbf{X}^{*}\mathbf{b}^{*} + \mathbf{Z}_{\mathrm{T}}^{*}\mathbf{a}_{\mathrm{T}} + \mathbf{e}^{*}$$

Univariate model (ASReml example)

Individual data

$$y = FE + A_1 + A_2 \text{ and}(A_3) \text{ and}(A_4) + \text{cage}$$

Pooled data

$$y = FE + A_1 \text{ and}(A_2) \text{ and}(A_3) \text{ and}(A_4)$$

Univariate model: Genetic

Individual data

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_I\mathbf{a}_I + \mathbf{Vcage} + \mathbf{e}$$

$$\begin{bmatrix} \sigma^2_{a_D} & \sigma_{a_{DI}} \\ \sigma_{a_{DI}} & \sigma^2_{a_I} \end{bmatrix}$$

Pooled data

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}^*_T\mathbf{a}_T + \mathbf{e}^*$$

$$\left[\sigma^2_{a_T}\right]$$

Bivariate model: Genetic

Individual data

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_I\mathbf{a}_I + \mathbf{Vcage} + \mathbf{e}$$

Pooled data

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}_T^*\mathbf{a}_T + \mathbf{e}^*$$

$$\begin{bmatrix} \sigma^2_{a1_D} & \sigma_{a1_{DI}} & \sigma_{a1_{D^2T}} \\ \sigma_{a1_{DI}} & \sigma^2_{a1_I} & \sigma_{a1_{I^2T}} \\ \sigma_{a1_{D^2T}} & \sigma_{a1_{I^2T}} & \sigma^2_{a2_T} \end{bmatrix}$$

Univariate model: Environmental

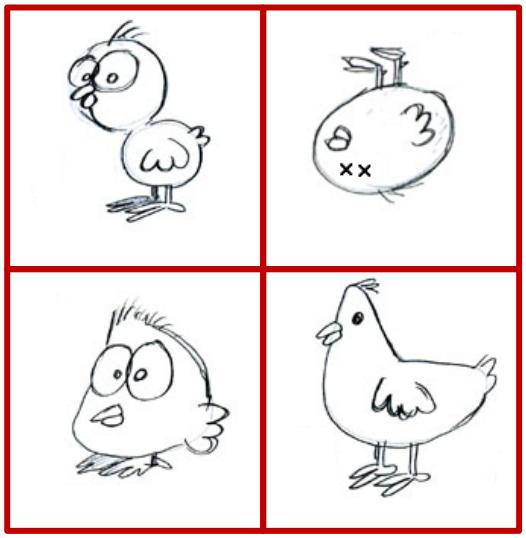
Individual data

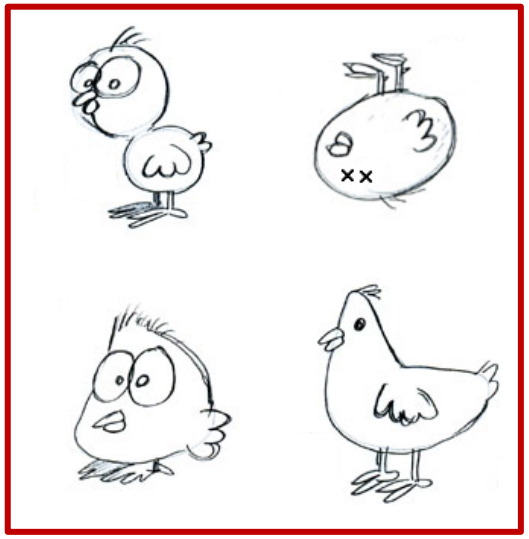
$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_I\mathbf{a}_I + \mathbf{Vcage} + \mathbf{e} \qquad [\sigma^2_{cage}] \qquad [\sigma^2_e]$$

Pooled data

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}^*_T\mathbf{a}_T + \mathbf{e}^* \qquad [\sigma^2_{e^*}]$$

Individual ➔ Pooled

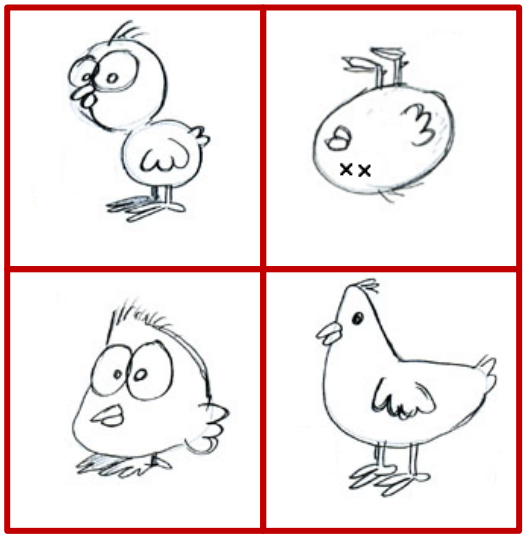


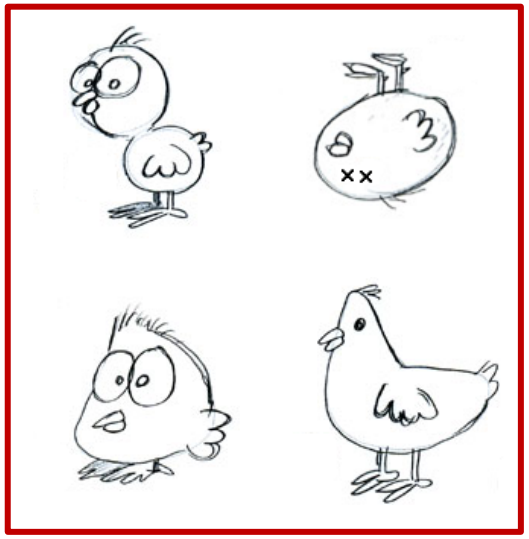


$$cage + e_1$$
$$cage + e_2$$
$$cage + e_3$$
$$cage + e_4$$

$$4 * cage + \sum e$$

Individual ➔ Pooled



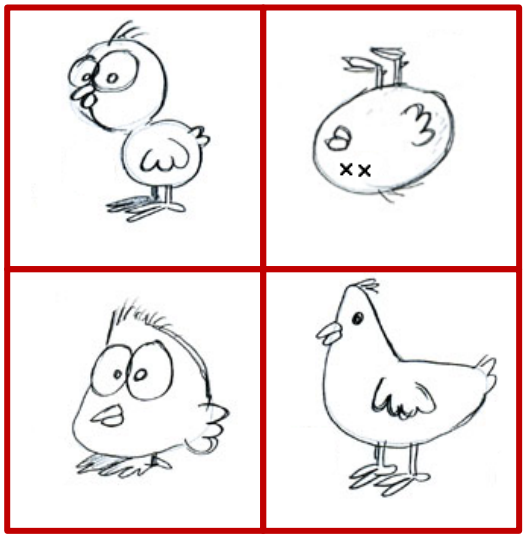


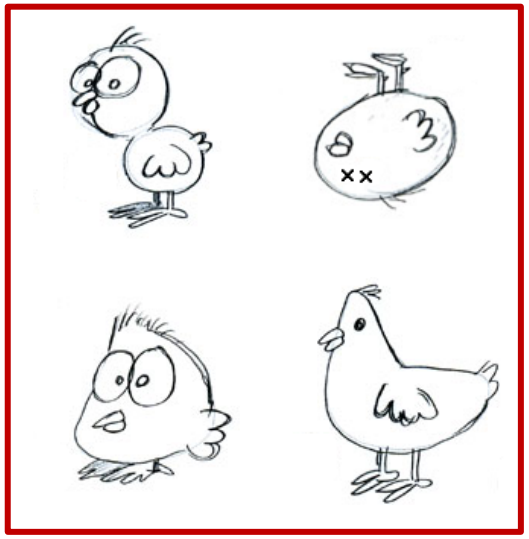
σ^2_{cage}

σ^2_e

$4 * cage + \sum e$

Individual ➔ Pooled





σ^2_{cage}

σ^2_e

$$\sigma^2_{e^*} = 4^2 * \sigma^2_{cage} + 4 * \sigma^2_e$$

Univariate model: Environmental

Individual data

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_I\mathbf{a}_I + \mathbf{Vcage} + \mathbf{e} \qquad [\sigma^2_{cage}] \qquad [\sigma^2_e]$$

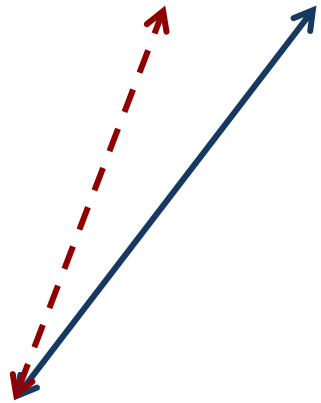
Pooled data

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}^*_T\mathbf{a}_T + \mathbf{e}^* \qquad [\sigma^2_{e^*}]$$

Bivariate model: Environmental

Individual data

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_I\mathbf{a}_I + \mathbf{Vcage} + \mathbf{e}$$



$$\left[\sigma^2_{cage_1}\right]$$

$$\begin{bmatrix} \sigma^2_{e_1} & \sigma_{e^{\times}_{12}} \\ \sigma_{e^{\times}_{12}} & \sigma^2_{e^*_2} \end{bmatrix}$$

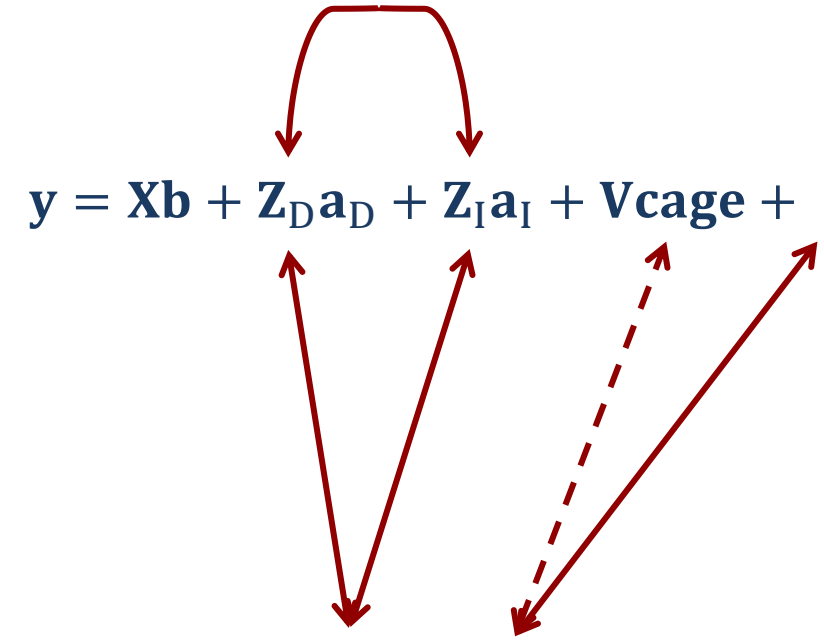
Pooled data

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}^*_T\mathbf{a}_T + \mathbf{e}^*$$

Bivariate model: BIAS!!

Individual data

Pooled data



$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_D a_D} + \mathbf{Z_I a_I} + \mathbf{Vcage} + \mathbf{e}$$

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}_T^*\mathbf{a}_T + \mathbf{e}^*$$

$$\begin{bmatrix} \sigma^2_{a1_D} & \sigma_{a1_{DI}} & \sigma_{a1_{D^2T}} \\ \sigma_{a1_{DI}} & \sigma^2_{a1_I} & \sigma_{a1_{I^2T}} \\ \sigma_{a1_{D^2T}} & \sigma_{a1_{I^2T}} & \sigma^2_{a2_T} \end{bmatrix}$$

$$\left[\sigma^2_{cage_1}\right]$$

$$\begin{bmatrix} \sigma^2_{e_1} & \sigma_{e^{\times}_{12}} \\ \sigma_{e^{\times}_{12}} & \sigma^2_{e^*_2} \end{bmatrix}$$

Simulated data: BIAS!!

	True parameters	Estimated parameters
$\sigma^2_{a1_D}$	1.00	1.37 (± 0.21)
$\sigma^2_{a1_I}$	1.00	1.45 (± 0.16)
$\sigma^2_{a2_T}$	11.50	17.70 (± 2.26)
$\sigma_{a1_{DI}}$	0.25	0.66 (± 0.15)
$\sigma_{a1_{D}2_T}$	1.00	2.99 (± 0.49)
$\sigma_{a1_{I}2_T}$	1.00	3.24 (± 0.34)

Variances trait 1

Variance trait 2

$$\begin{bmatrix} \sigma^2_{a1_D} & \sigma_{a1_{DI}} & \sigma_{a1_{D}2_T} \\ \sigma_{a1_{DI}} & \sigma^2_{a1_I} & \sigma_{a1_{I}2_T} \\ \sigma_{a1_{D}2_T} & \sigma_{a1_{I}2_T} & \sigma^2_{a2_T} \end{bmatrix}$$

Simulated data: BIAS!!

	True parameters	Estimated parameters
$\sigma^2_{a1_D}$	1.00	1.37 (± 0.21)
$\sigma^2_{a1_I}$	1.00	1.45 (± 0.16)
$\sigma^2_{a2_T}$	11.50	17.70 (± 2.26)
$\sigma_{a1_{DI}}$	0.25	0.66 (± 0.15)
$\sigma_{a1_D{}^2T}$	1.00	2.99 (± 0.49)
$\sigma_{a1_I{}^2T}$	1.00	3.24 (± 0.34)

Covariances

$$\begin{bmatrix} \sigma^2_{a1_D} & \sigma_{a1_{DI}} & \sigma_{a1_D{}^2T} \\ \sigma_{a1_{DI}} & \sigma^2_{a1_I} & \sigma_{a1_I{}^2T} \\ \sigma_{a1_D{}^2T} & \sigma_{a1_I{}^2T} & \sigma^2_{a2_T} \end{bmatrix}$$

Bivariate model: Environmental

Individual data

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_I\mathbf{a}_I + \mathbf{Vcage} + \mathbf{e}$$

$$\begin{bmatrix} \sigma^2_{cage_1} & \sigma_{cage_{12}} \\ \sigma_{cage_{12}} & \sigma^2_{cage_2} \end{bmatrix}$$

Pooled data

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}^*_T\mathbf{a}_T + \mathbf{Vcage} + \mathbf{e}^*$$

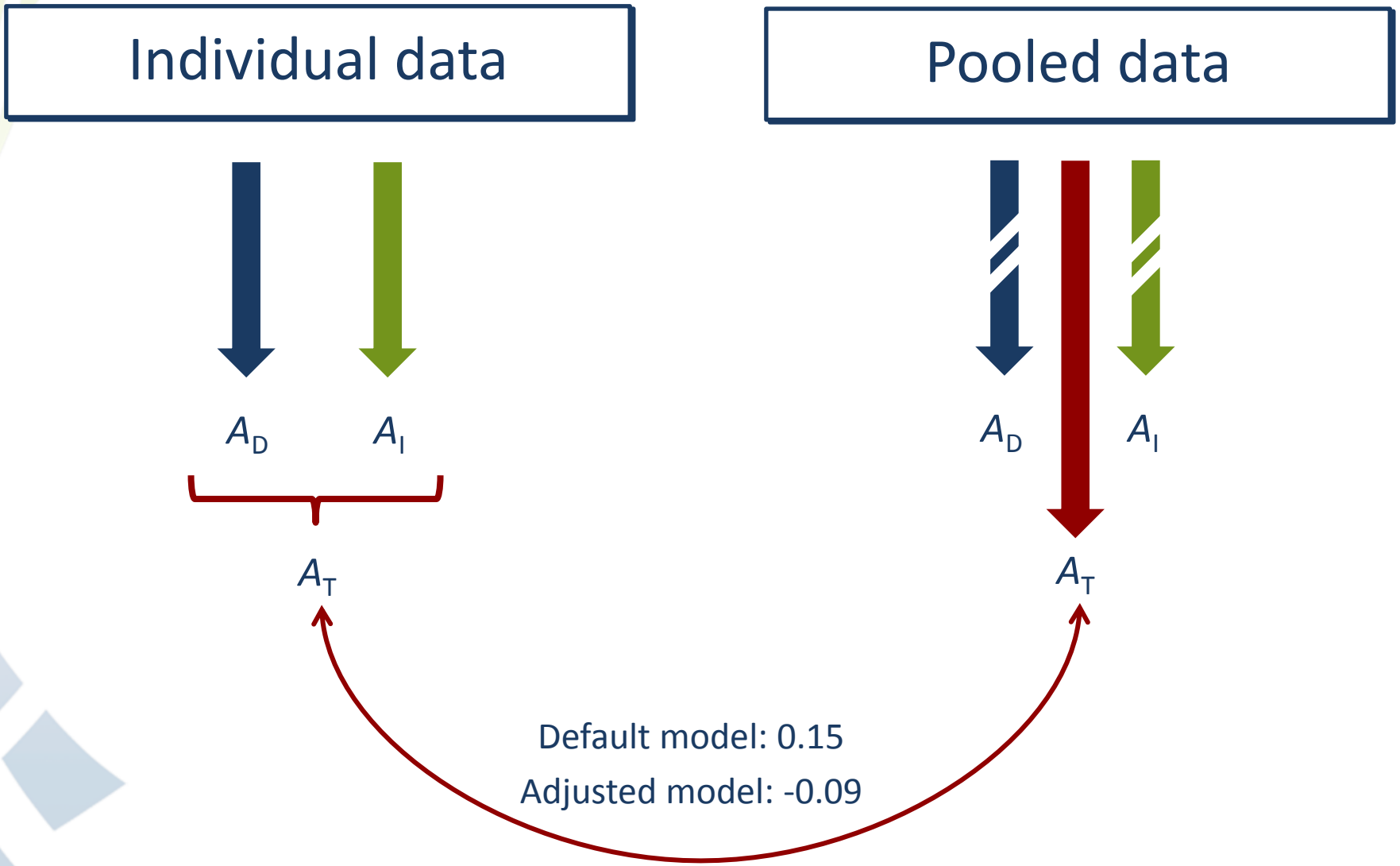
$$\begin{bmatrix} \sigma^2_{e_1} & \sigma_{e^\times_{12}} \\ \sigma_{e^\times_{12}} & \sigma^2_{e^*_2} \end{bmatrix}$$

Fix

Simulated data: Biased → Unbiased

	True parameters	Estimated parameters Default model	Estimated parameters Adjusted model
$\sigma^2_{a1_D}$	1.00	1.37 (± 0.21)	1.01 (± 0.17)
$\sigma^2_{a1_I}$	1.00	1.45 (± 0.16)	1.00 (± 0.15)
$\sigma^2_{a2_T}$	11.50	17.70 (± 2.26)	11.64 (± 1.91)
$\sigma_{A1_{DI}}$	0.25	0.66 (± 0.15)	0.26 (± 0.13)
$\sigma_{A1_D 2_T}$	1.00	2.99 (± 0.49)	1.02 (± 0.46)
$\sigma_{A1_I 2_T}$	1.00	3.24 (± 0.34)	1.02 (± 0.44)

Correlation
Survival time ~ Early egg production



Problem:

Incorrect variance components
due to a non fitted environmental correlation

Solution:

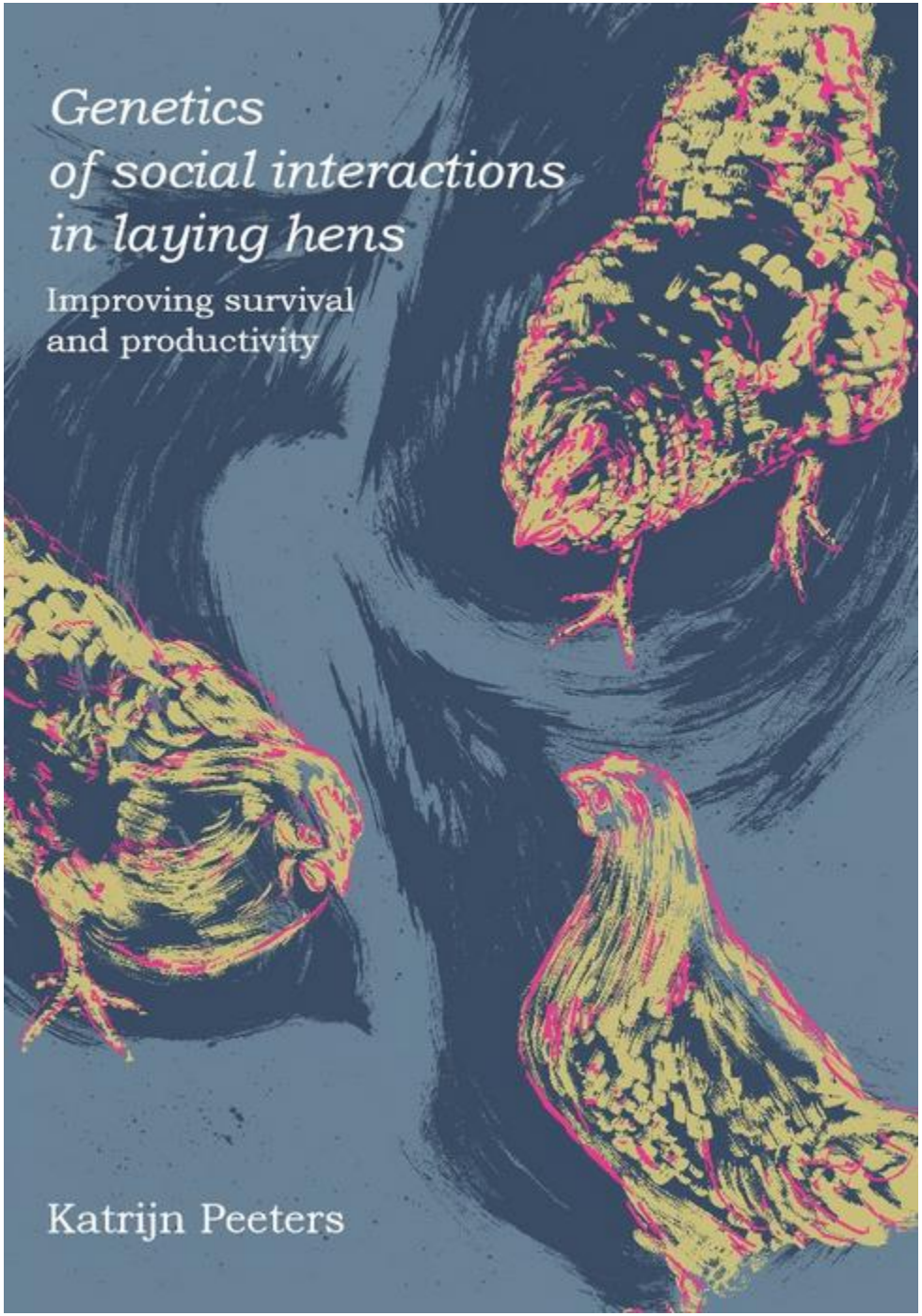
1. Add a pooled cage effect

+

2. Fit a correlation between the individual cage effect and pooled cage effect

+

3. Fix the pooled cage effect to avoid over-parameterisation


Genetics
of social interactions
in laying hens
Improving survival
and productivity
Katrijn Peeters