Using pooled data to estimate genetic parameters for social interaction traits

Katrijn Peeters, Esther Ellen and Piter Bijma

WAGENINGEN UR
For quality of life

STW
Enabling new technology

NWO
Netherlands Organisation for Scientific Research

HENDRIX GENETICS

Pooling phenotypic data

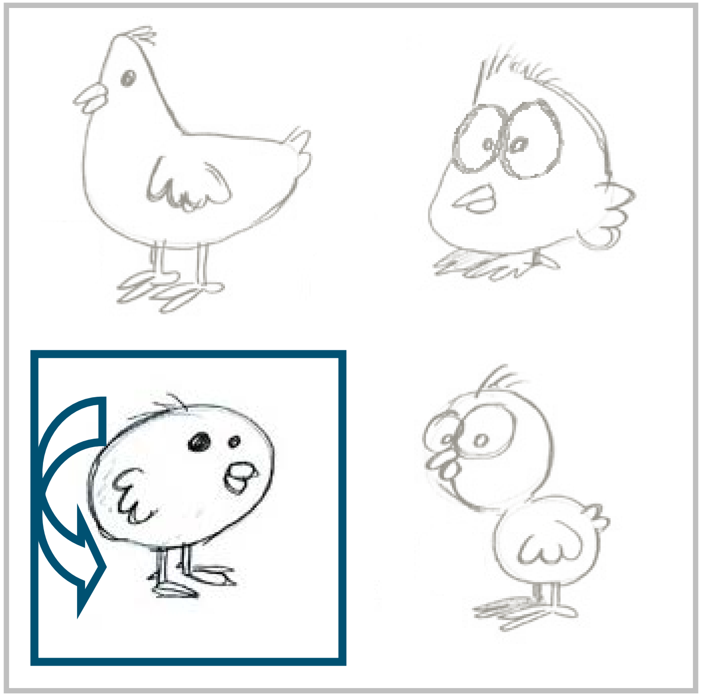
Collecting individual data
in group housing systems

Expensive *Difficult*

Pooling data

WAGENINGEN UR
For quality of life

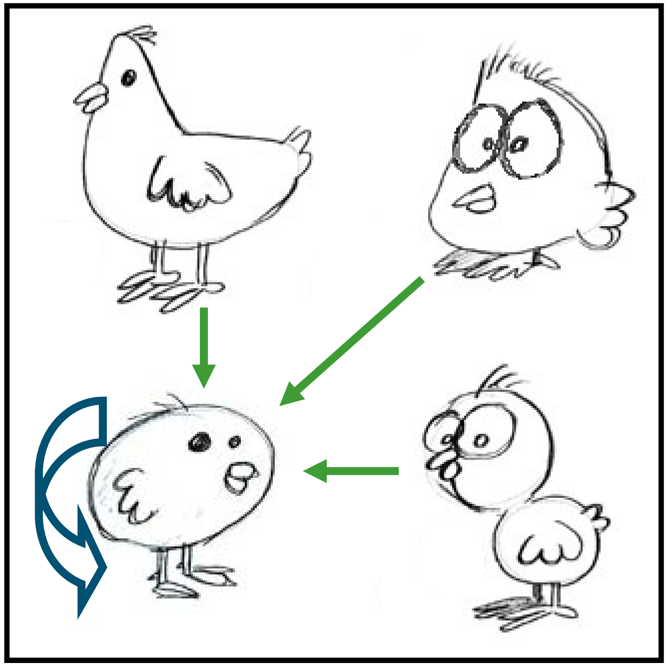
Social interaction theory



Direct

$$P_i = A_{D_i} + E_{D_i}$$

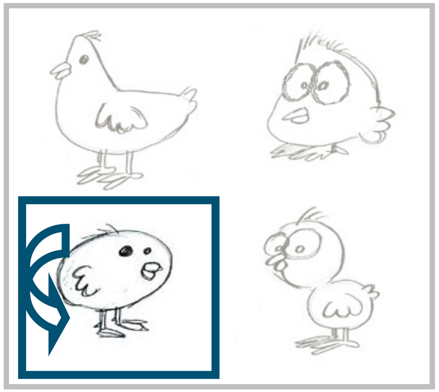
Social interactions theory



Direct | **Social**

$$P_i = A_{D_i} + E_{D_i} + \sum_{i \neq j}^{n-1} A_{S_j} + \sum_{i \neq j}^{n-1} E_{S_j}$$

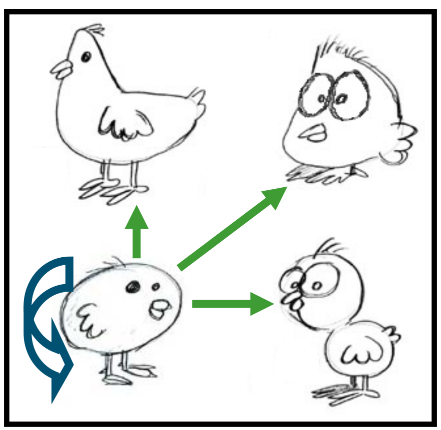
Direct model ⇔ Direct-Social model



Direct

$$A_D$$

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

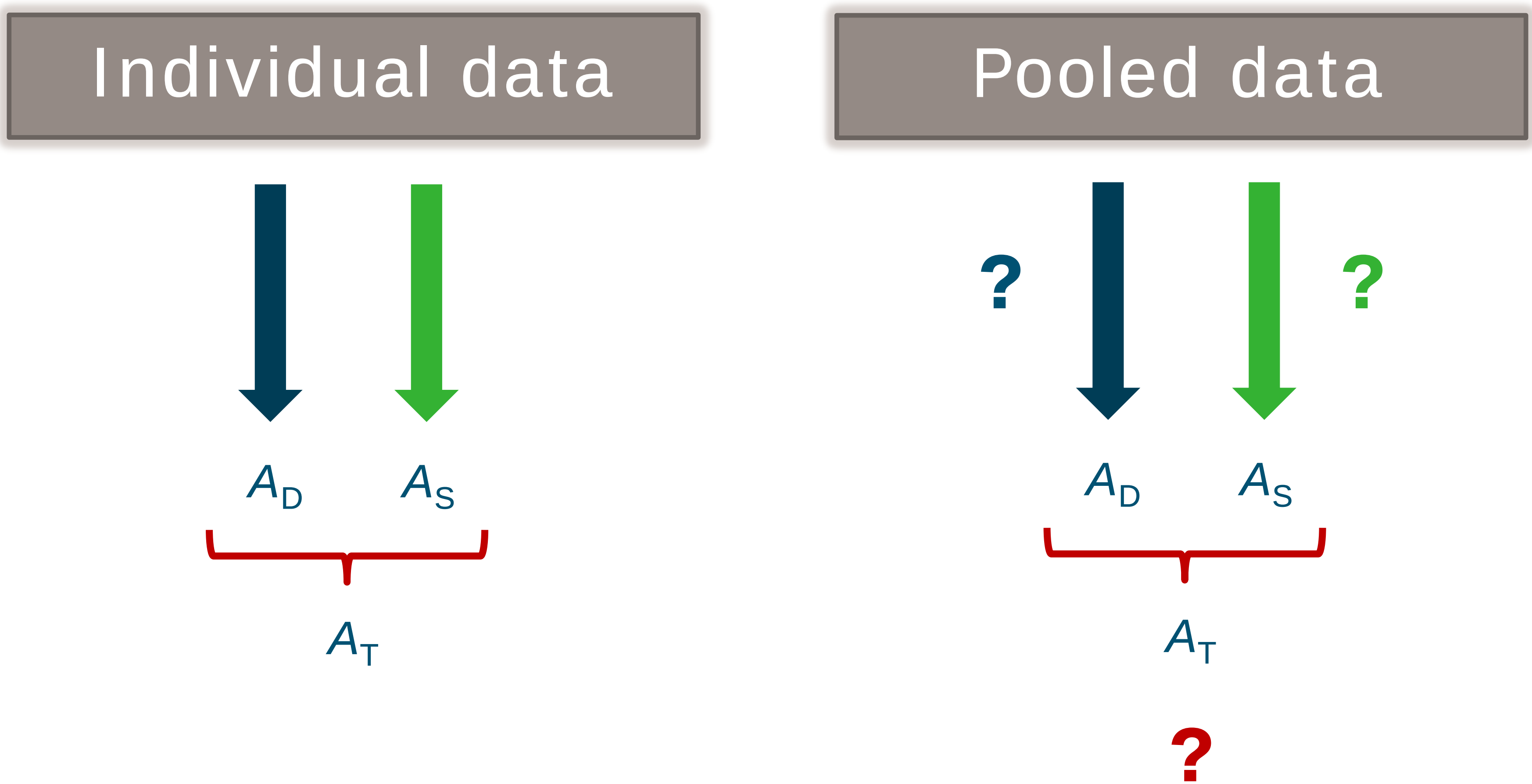


Direct-indirect

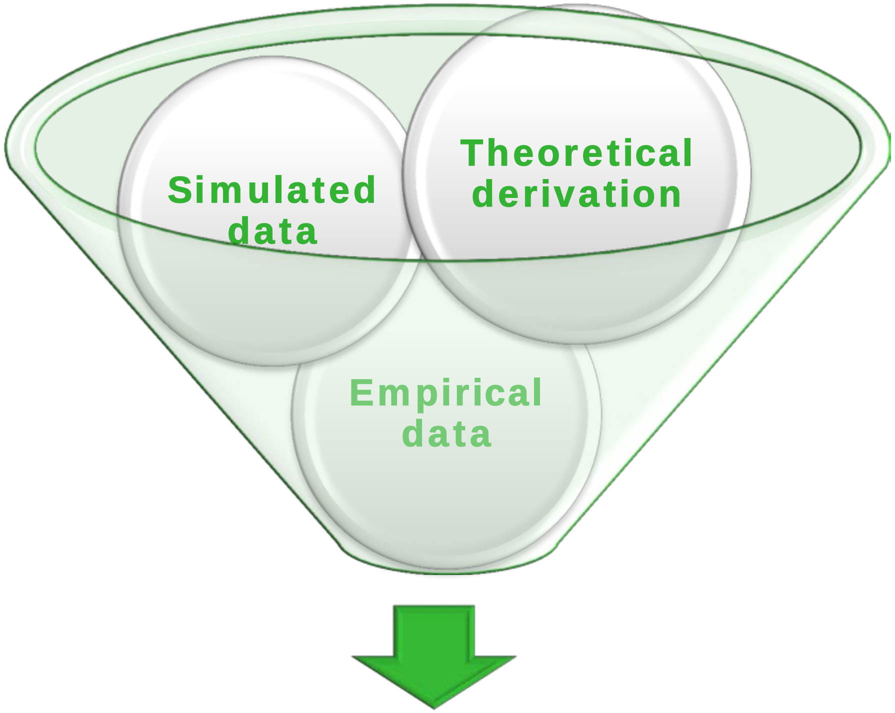
$$A_T = A_D + 3A_S$$

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

Aim: what can we estimate?



Three Methods

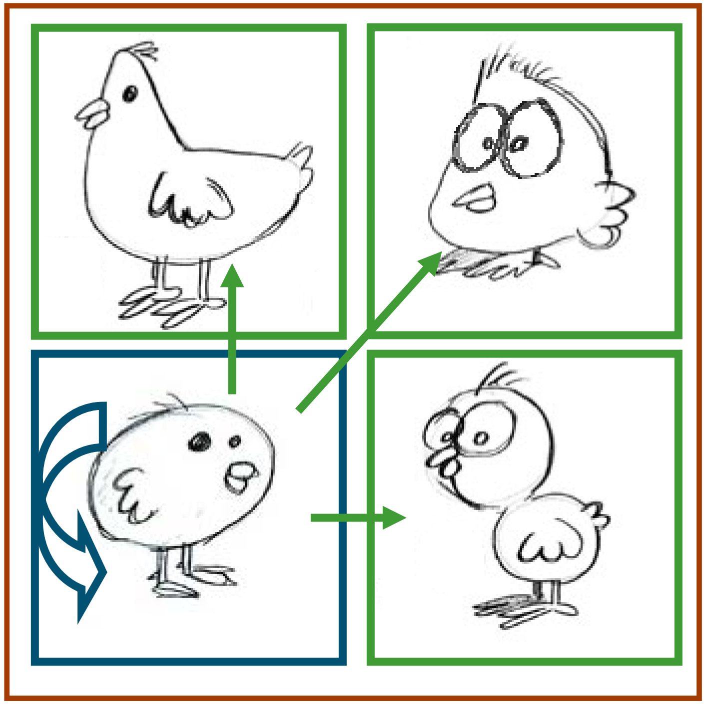


Can we estimate

direct and social genetic parameters

from pooled data?

1. Theoretical derivation: Individual data



$$P_1 = \mathbf{A_{D1}} + A_{S2} + A_{S3} + A_{S4} + E$$

$$P_2 = A_{D2} + \mathbf{A_{S1}} + A_{S3} + A_{S4} + E$$

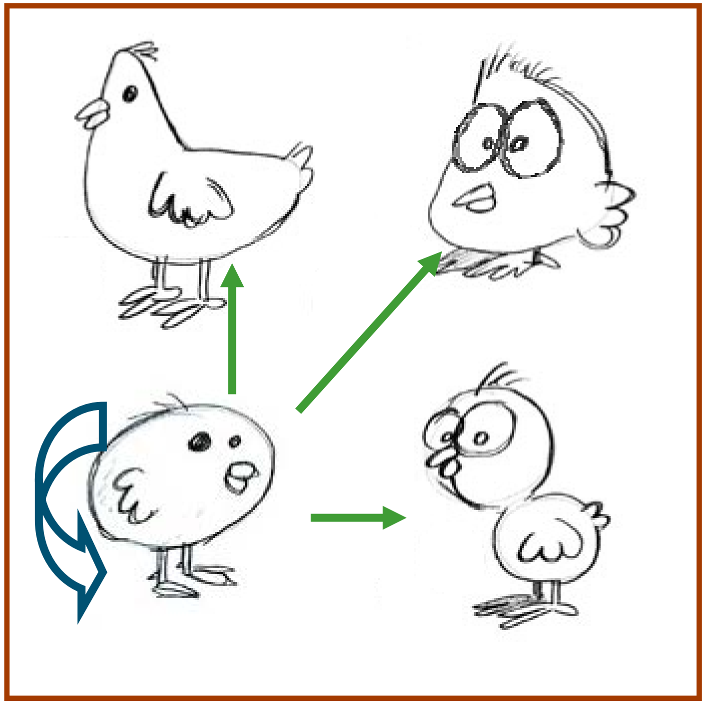
$$P_3 = A_{D3} + \mathbf{A_{S1}} + A_{S2} + A_{S4} + E$$

$$P_4 = A_{D4} + \mathbf{A_{S1}} + A_{S2} + A_{S3} + E$$

Information on A_D and A_S is obtained from different phenotypes ($P_1 \neq P_{2/3/4}$)

Both A_D and A_S are estimable

1. Theoretical derivation: Pooled data (P*)



$$P^* = \boldsymbol{A_{D1}} + A_{D2} + A_{D3} + A_{D4} + 3(\boldsymbol{A_{S1}} + A_{S2} + A_{S3} + A_{S4}) + E$$

An individual's A_D and its full A_S is always lumped into a single record (P^*)

➔ A_D and A_S are fully confounded

$$P^* = \boldsymbol{A_{T1}} + A_{T2} + A_{T3} + A_{T4} + E$$

➔ A_T can be estimated

➔ **y** = **Xb** + Σ**Za** + **e** will estimate A_T

WAGENINGEN UR
For quality of life

2. Simulated data

- 500 Sires, 500 Dams, 12 offspring per mating
- 12,000 individuals
- Direct and Social effects
- 3,000 groups of 4 individuals each
- Record = pooled phenotype per group

2. Simulated data

- Input

$\sigma^2_{A_D} = 1;\ \sigma^2_{A_S} = 1;\ \sigma_{A_{DS}} = 0.5$
$\sigma^2_{E_D} = 2;\ \sigma^2_{E_S} = 2;\ \sigma_{E_{DS}} = 0$

- Pooled data model (ASReml)

$\mathbf{y} = \mathbf{Xb} + \Sigma\mathbf{Za} + \mathbf{e}$

$\mathrm{y} \sim \mathrm{mu}\ \ \mathrm{FIXED}\ \ !\mathrm{r}\ \ A_1\ \mathrm{and}(A_2)\ \mathrm{and}(A_3)\ \mathrm{and}(A_4)$

- Expected values ⇔ obtained values (100 replicates)

$\sigma^2_{A_T} = \sigma^2_{A_D} + 6\,\sigma_{A_{DS}} + 9\,\sigma^2_{A_S} = 13 \quad \Leftrightarrow \quad \hat{\sigma}^2_A = 13.05\ (\pm 2.12)$
$\sigma^2_{E^*} = 4\left(\sigma^2_{E_D} + 6\,\sigma_{E_{DS}} + 9\,\sigma^2_{E_S}\right) = 80 \quad \Leftrightarrow \quad \hat{\sigma}^2_E = 80.32\ (\pm 7.30)$

Note: $\hat{\sigma}^2_A \neq \hat{\sigma}^2_{A_D}$ but $\hat{\sigma}^2_A = \hat{\sigma}^2_{A_T}$

3. Empirical data: cannibalistic laying hens

- 12,944 Laying hens → 6,092 W1 / 6,852 WB

- Individual survival time:

$$y \sim mu \;\; FIXED \;\; !r \;\; A_{D1} \;\; A_{I2} \; and(A_{I3}) \; and(A_{I4}) \;\; Cage$$

- Group survival time (pooled per cage):

$$y \sim mu \;\; FIXED \;\; !r \;\; A_{T1} \; and(A_{T2}) \; and(A_{T3}) \; and(A_{T4})$$

3. Empirical data

Individual data

Pooled data

	W1	WB
$\sigma^2_{A_D}$	705 (±171)	1,404 (±301)
$\sigma_{A_{DS}}$	59 (±61)	-162 (±105)
$\sigma^2_{A_S}$	104 (±41)	292 (±72)
σ^2_{Cage}	799 (±166)	1,191 (±238)
σ^2_{E}	7,980 (±210)	12,675 (±365)

$$\sigma^2_{A_D} + 6\,\sigma_{A_{DS}} + 9\,\sigma^2_{A_S} = \sigma^2_{A_T}$$

$$16\,\sigma^2_{Cage} + 4\sigma^2_E = \sigma^2_{E^*}$$

3. Empirical data

Individual data		
	W1	**WB**
$\sigma^2_{A_T}$	1,996 (±640)	2,521 (±842)
$\sigma^2_{E^*}$	44,700 (±2,526)	69,752 (±3,513)

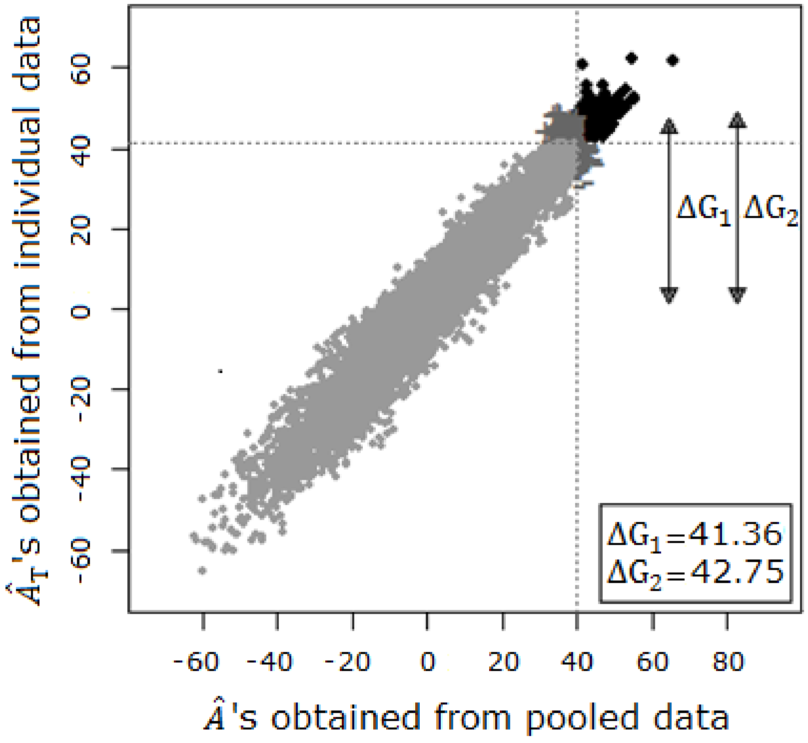
Pooled data		
	W1	**WB**
$\sigma^2_{A_T}$	1,979 (±643)	2,521 (±845)
$\sigma^2_{E^*}$	44,750 (±2,538)	69,750 (±3,519)

Indeed: analysis of pooled records yields $\hat{\sigma}^2_{A_T}$ rather than $\hat{\sigma}^2_{A_D}$

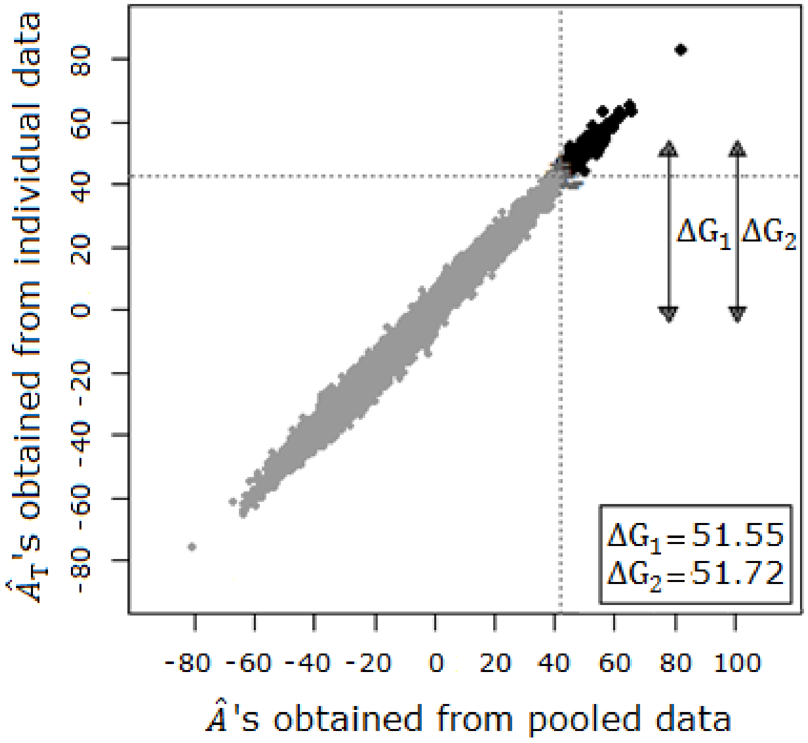
WAGENINGEN UR
For quality of life

3. Empirical data: The cost of pooling

W1

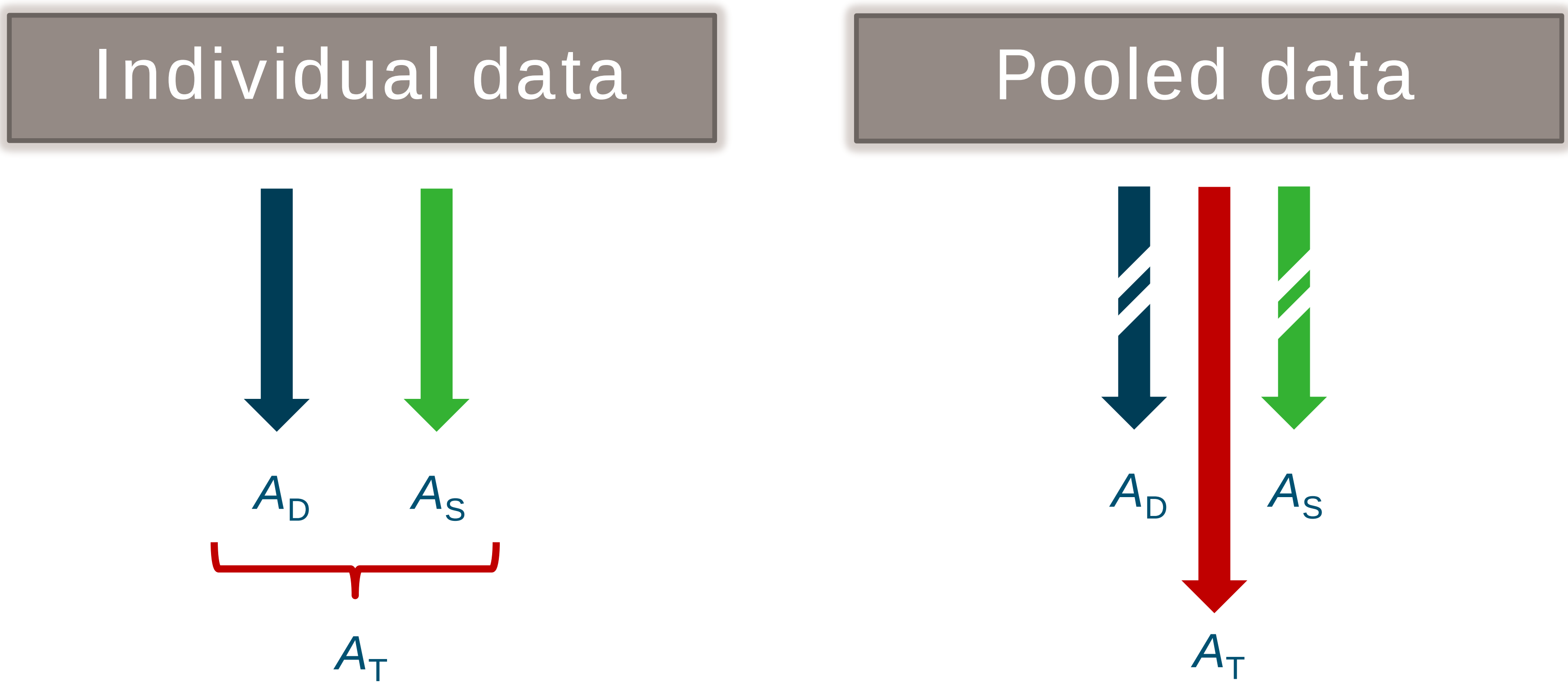


WB



Pooling causes surprisingly little reduction in response to selection for socially-affected traits.

Conclusion



With social interactions,

estimated Var(A) from pooled data will differ from ordinary Var(A).

$$\mathrm{Var}(A)_{pooled} = \mathrm{Var}(A_T)$$

Social interactions

Peeters *et al.*, Genetics Selection Evolution **45**:27

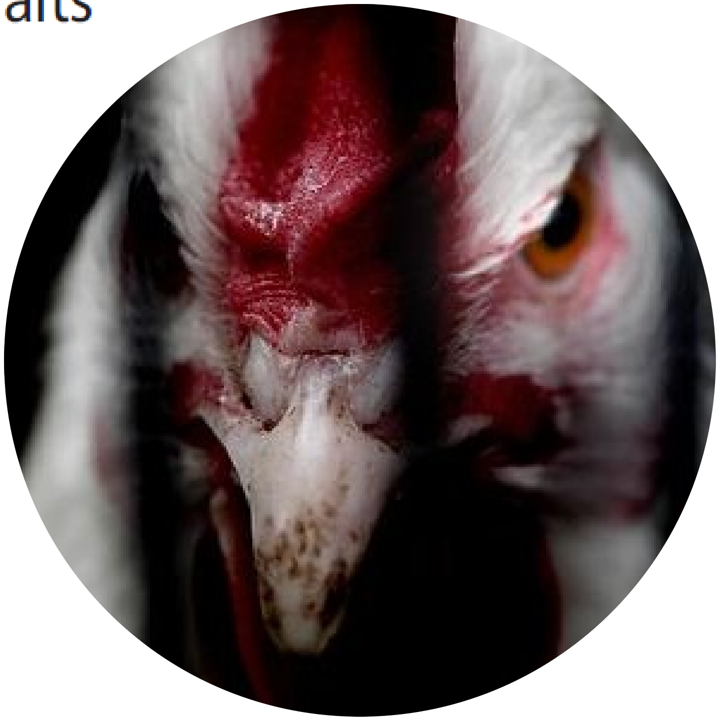
katrijn.peeters@wur.nl

Genetics Selection Evolution

RESEARCH Open Access

Using pooled data to estimate variance components and breeding values for traits affected by social interactions

Katrijn Peeters*, Esther Dorien Ellen and Piter Bijma



WAGENINGEN UR For quality of life

STW Enabling new technology

NWO Netherlands Organisation for Scientific Research

HENDRIX GENETICS

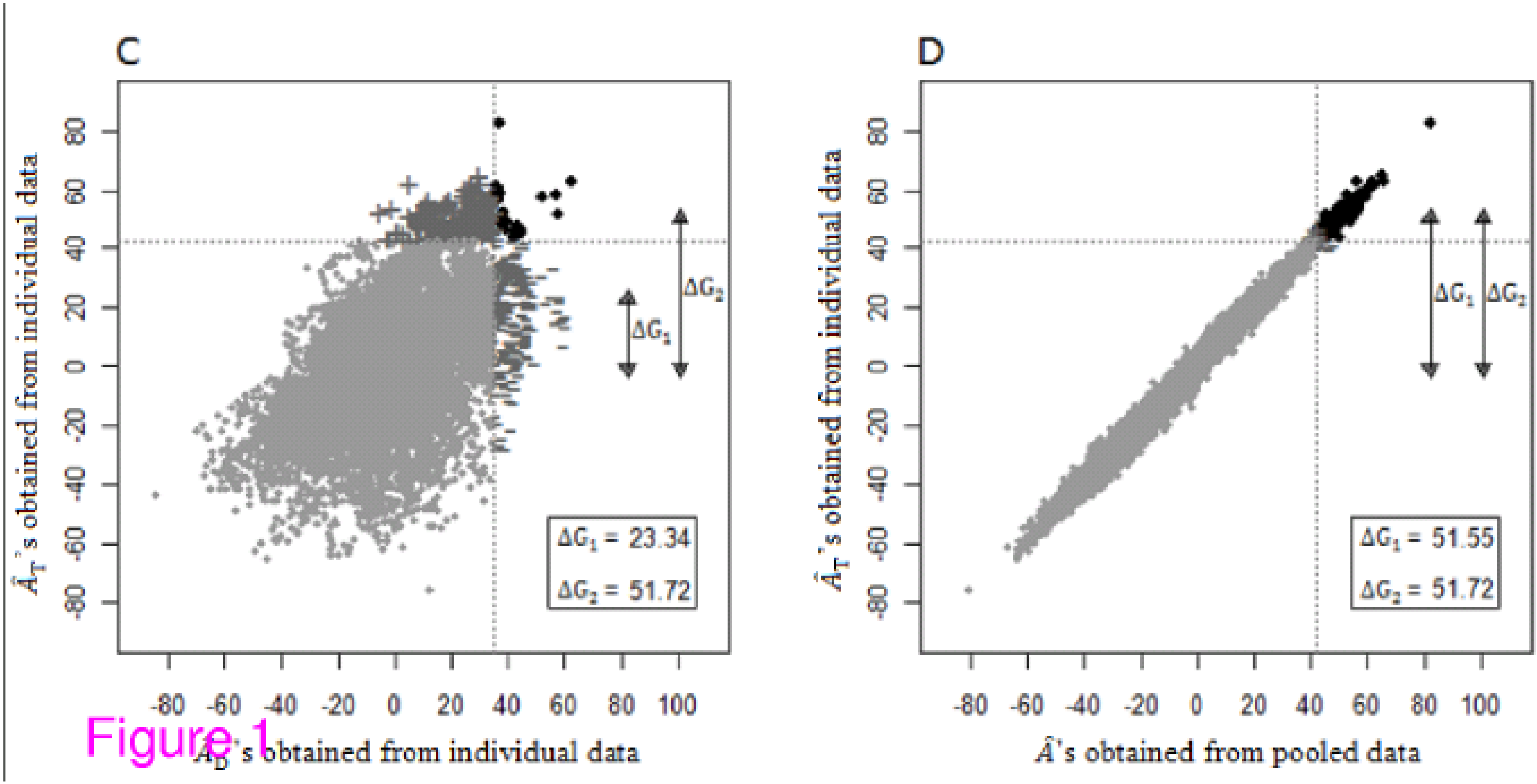

C
D
Â_T's obtained from individual data
Â_D's obtained from individual data
Â's obtained from pooled data
ΔG1
ΔG2
ΔG1 = 23.34
ΔG2 = 51.72
ΔG1 = 51.55
ΔG2 = 51.72

Figure 1