Detection of genomic pre-selection with Mendelian sampling variance test

A.-M. Tyrisevä[1], E. A. Mäntysaari[1], J. Jakobsen[2], G. P. Aamand[3], J. Dürr[2], W. F. Fikse[4], M. H. Lidauer[1]

[1]MTT Agrifood Research Finland
[2]Interbull Centre, Sweden
[3]NAV Nordic Cattle Genetic Evaluation, Denmark
[4]Dept. Animal Breeding and Genetics, SLU, Sweden

STØTTET AF
mælkeafgiftsfonden

Naturerhverv.dk

Ministeriet for Fødevarer, Landbrug og Fiskeri

Den Europæriske Landbrugsfond for Udvikling af Landdistrikterne

Danmark og EU investerer i landdistrikterne.

Link til: European Agricultural Fund for Rural Development

Background

- After genomic pre-selection, the next generation AI bulls are not a random sample of the progeny of their parents

- In theory, also the genetic variance reduces due to pre-selection

- If not accounted for in the evaluation model, EBVs will be biased → problems in MACE

- **Aim**:
  - Is it possible to detect effect of genomic pre-selection in MS means or variances and utilize information in the validation of national evaluation models?

Study outline

- Sample populations were simulated:
  a) Control scheme
  b) GPS scheme, where from a certain year onwards genomic preselection starts:
     - All young bulls better than their parents
- To accomplish the goal:
  - Simple and fast way to simulate data for real population was tested:
    - Method retains existing pedigree and data structure
    - Generates genetic trend
    - Allows an easy way to include genomic pre-selection

Data used for simulations

- 750 000 Danish Holstein cows from 2000 herds and born during a 20-year time interval

- One observation was simulated for each cow

- Model:

$$y = herd + BV + e$$

- Simulated h^2 was 0.25

- Thus, only the herd and the pedigree structure were retained from the original data

MTT

Simple method to simulate data with genetic trend

1. Replace original **y** by a yearly increasing value carrying a desired annual trend
   - Records of animals having progenies were set missing to ensure that the average MS of parents would not be regressed towards yearly means

2. Estimate breeding values from pseudo records
   - Pseudo BVs will be in synchrony with parent and progeny averages and expected yearly means of BVs

3. Calculate MS terms of the animals from pseudo BVs

4. Use the MS terms from step 3 to simulate true breeding values

→ Genetic trend created in the first step was transmitted to the generated data through the MS terms

Considering change in genetic variance

- In a case of unselected MS, the MS term ϕ_i of animal i would be generated followingly:

$$\phi_i \mid E[\phi_i] \sim N[E[\phi_i], d_{ii}\sigma_u^2]$$

$E[\phi_i]$ = expected value of MS term
d_{ii} = diagonal of an animal i in $\mathbf{A}$

- Howerver, an $E[\phi_i]$ different from zero yields into inflated BV variance:

$$Var(\phi_i) = d_{ii}\sigma_u^2 + Var(E[\phi_i])$$

- This was avoided by carrying out a variance reduction:

$$\phi_i \mid E[\phi_i] \sim N[E[\phi_i], (1-k)d_{ii}\sigma_u^2]$$

$1\text{-}k$ = variance reduction factor

Considering change in genetic variance

- Obtaining variance reduction factor *(1-k)*:

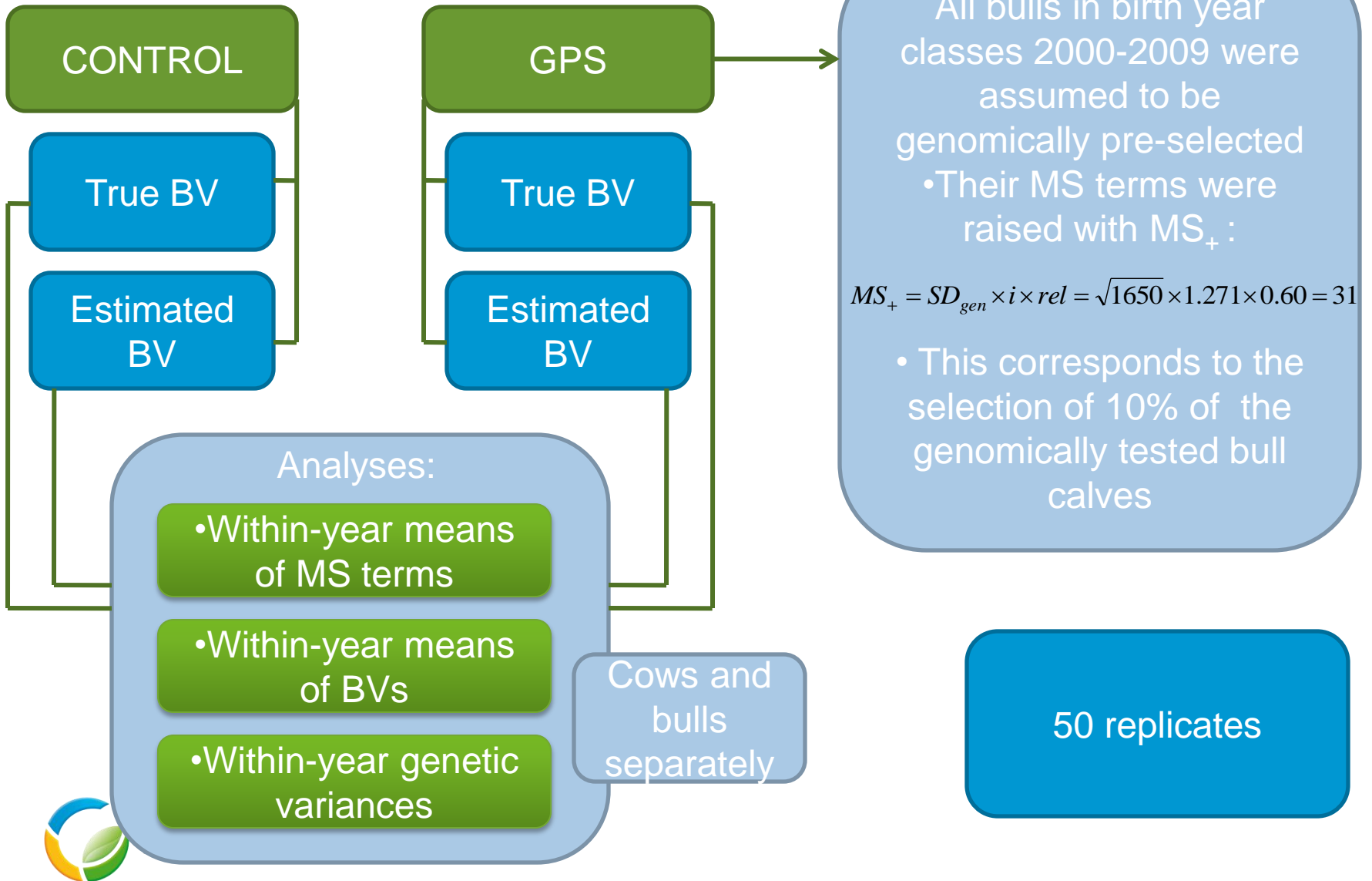$$k = i(i - x) \quad \text{and} \quad i = E[\phi]/\sigma_\phi$$

i = selection intensity
x = deviation of truncation point from mean in standard deviation units
σ_ϕ = standard deviation of MS term

  - Standard formula from Falconer et al. 1996
  - x is unknown → approximation of *(1-k)*

- *(1-k)* is an exponential function of i and a satisfactory approximation can be obtained by a simple linear fit of its logarithmic value:

$$(i - k)_i = Exp(-1.18969\,|i| + 0.10805 i^2)$$

MTT

Design of the study

CONTROL

GPS → All bulls in birth year classes 2000-2009 were assumed to be genomically pre-selected
- Their MS terms were raised with MS_+ :

$$MS_+ = SD_{gen} \times i \times rel = \sqrt{1650} \times 1.271 \times 0.60 = 31$$

- This corresponds to the selection of 10% of the genomically tested bull calves

True BV

Estimated BV

True BV

Estimated BV

Analyses:
- Within-year means of MS terms
- Within-year means of BVs
- Within-year genetic variances

Cows and bulls separately

50 replicates

Effect of genomic pre-selection on true and estimated breeding values.
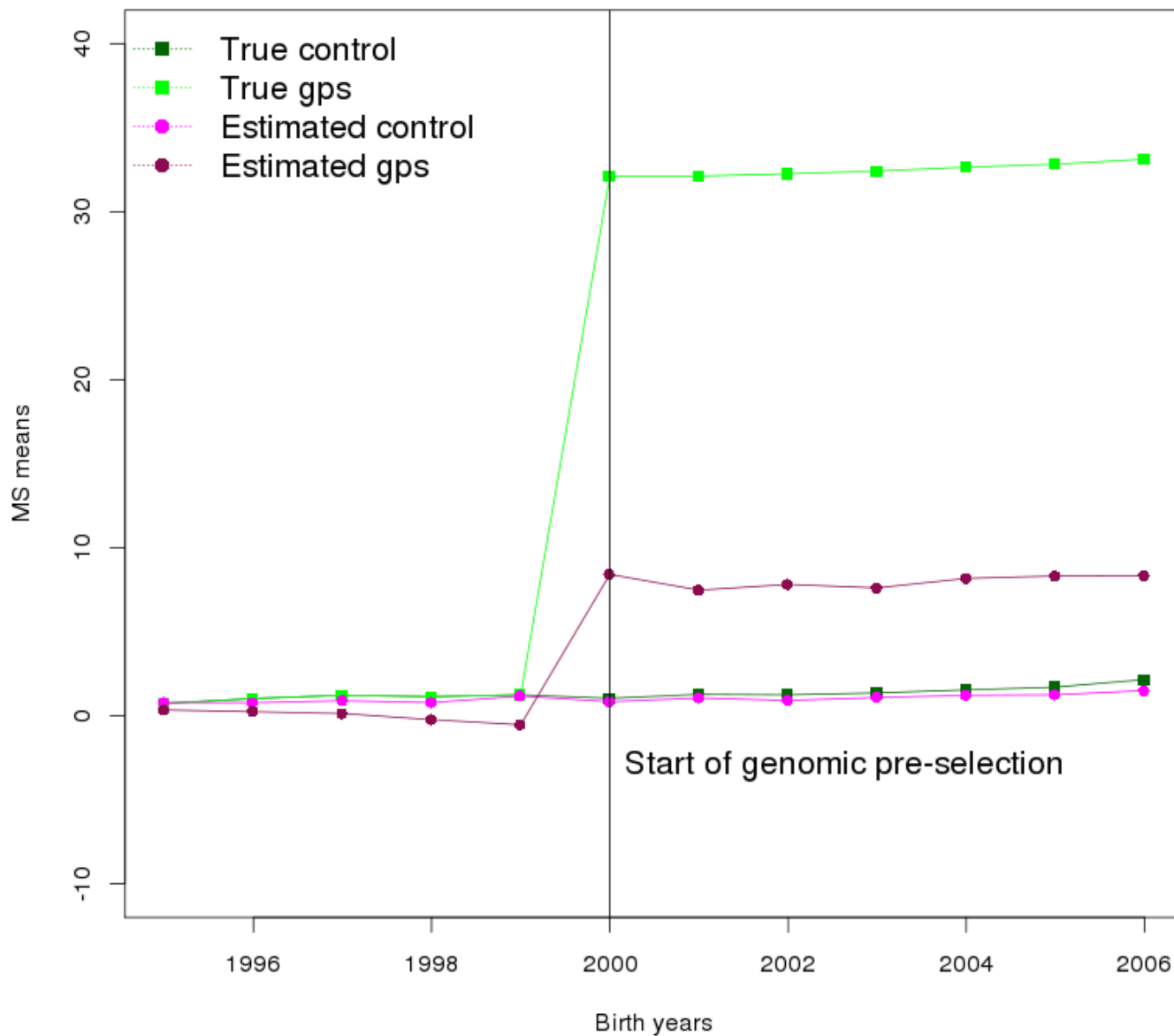Bulls averaged over 50 replicates.

Breeding values bulls

EBVs under-estimate the effect of genomic pre-selection
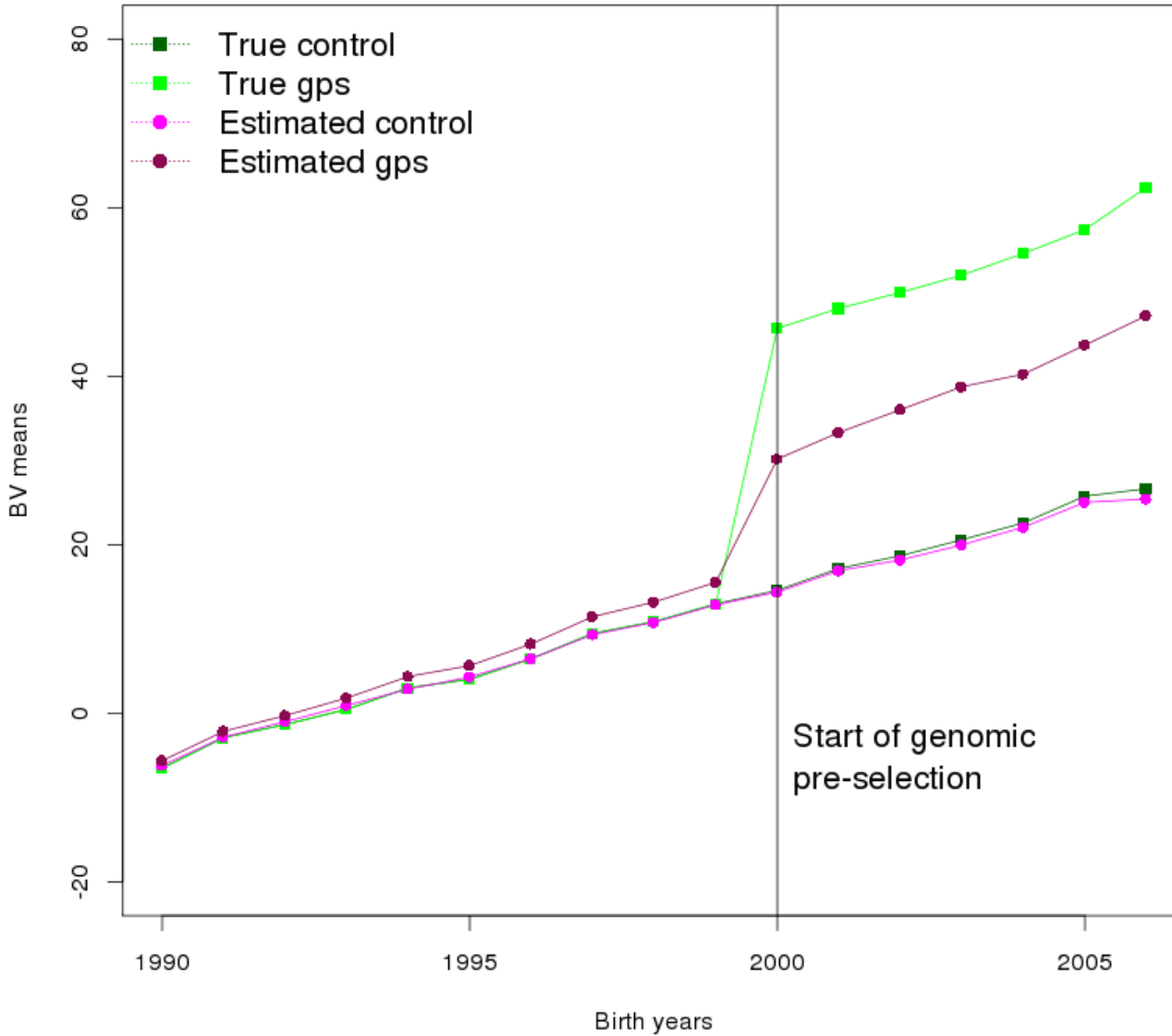
MS means bulls

Effect of genomic pre-selection on true and estimated MS means. Bulls averaged over 50 replicates.
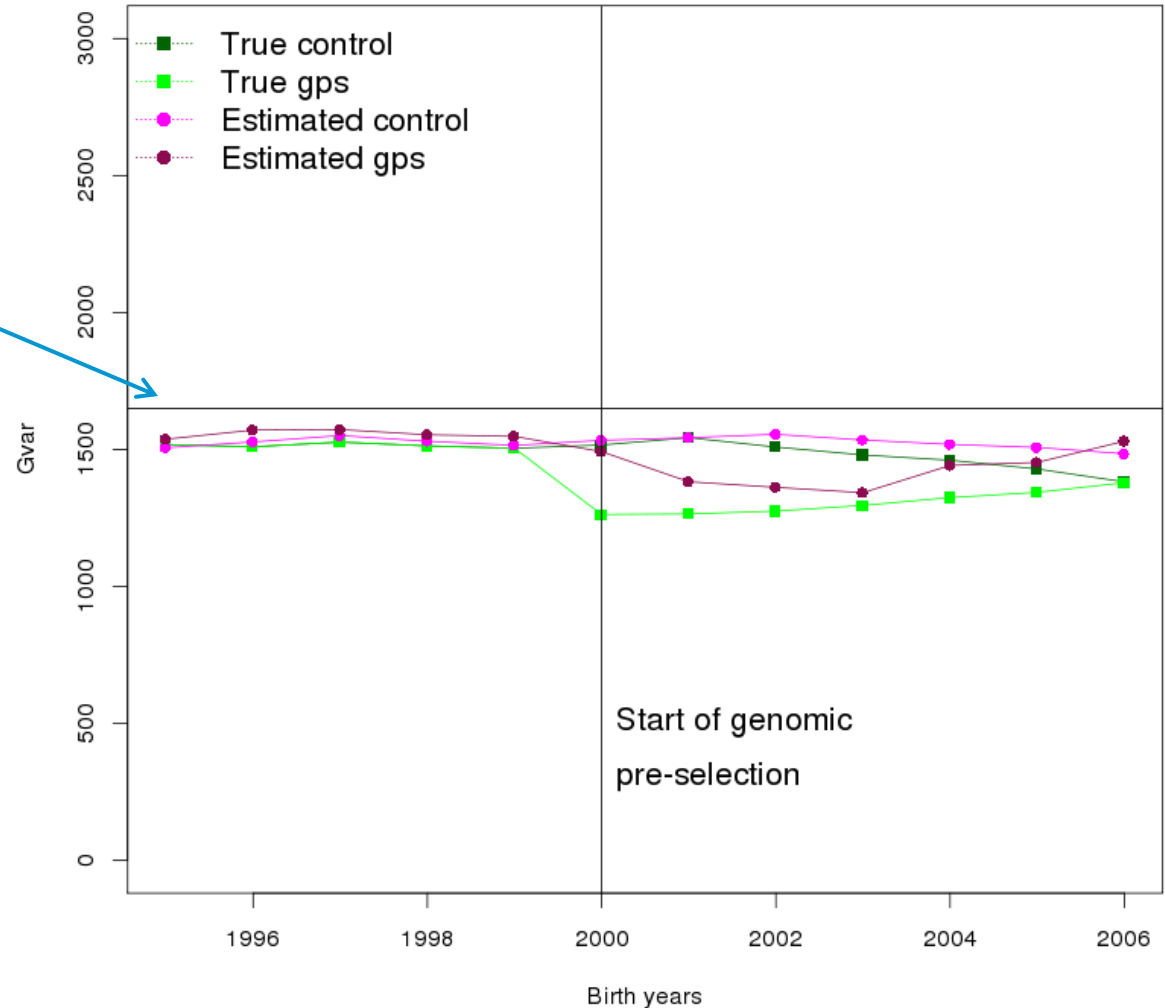
Within-year genetic variances bulls

Estimates from MS variance validation software (Tyrisevä et al. 2012 Interbull Bulletin 46)

$$\sigma_{u_i}^2 = \frac{\sum_{k=1}^{q_i} d_k \hat{m}s^2{}_k}{q_i - \sum_{k=1}^{q_i} d_k PEV(\hat{m}s_k)}$$
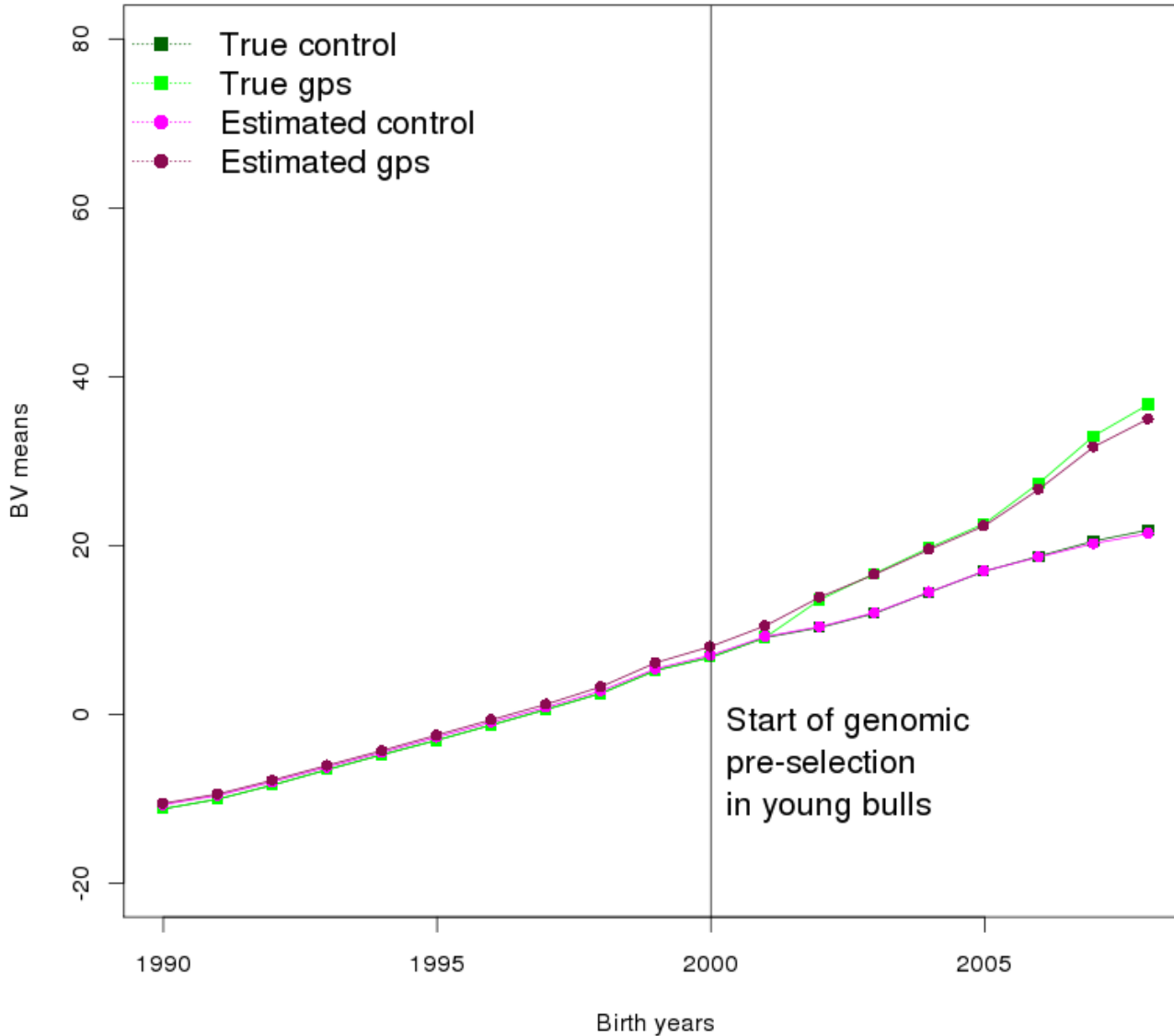
Applied genetic variance for the simulation of BVs was 1650

Effect of genomic pre-selection on genetic variances. Bulls averaged over 50 replicates.

- True control
- True gps
- Estimated control
- Estimated gps

Start of genomic pre-selection

Gvar

Birth years

Effect of genomic pre-selection on true and estimated breeding values. Cows averaged over 50 replicates.

Breeding values cows

Compared to control scheme, BVs of cows having genomically pre-selected sires were increased
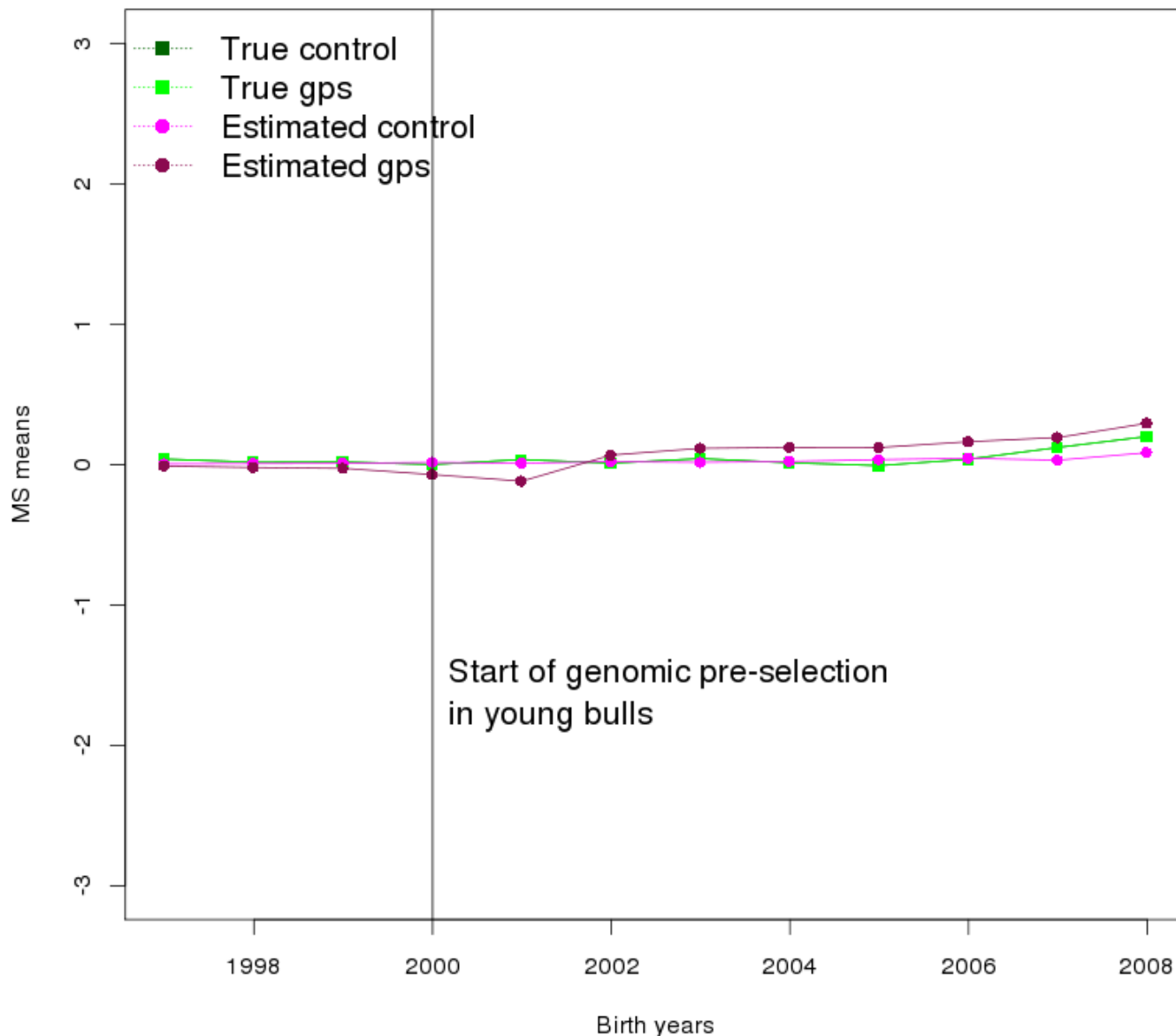
MS means cows

The overall level of MS means was zero

The MS mean deviated slightly from the expactation for daughters of genomically pre-selected bulls

**Effect of genomic pre-selection on true and estimated MS means.
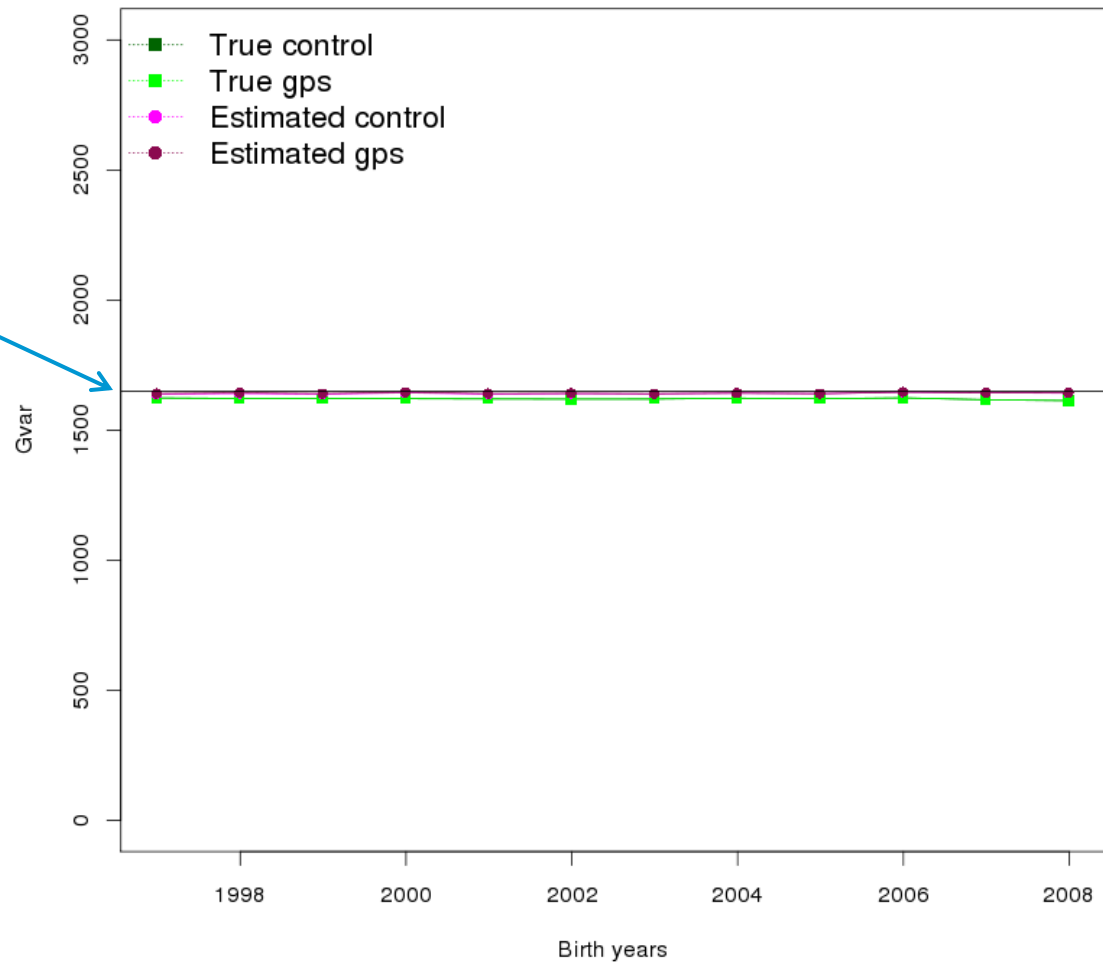Cows averaged over 50 replicates.**

Legend:
- True control
- True gps
- Estimated control
- Estimated gps

Start of genomic pre-selection in young bulls

MS means

Birth years

Within-year genetic variances cows

Estimates from MS variance validation software (Tyrisevä et al. 2012 Interbull Bulletin 46)

$$\sigma_{u_i}^2 = \frac{\sum_{k=1}^{q_i} d_k \hat{m}s^2_{\;k}}{q_i - \sum_{k=1}^{q_i} d_k PEV(\hat{m}s_k)}$$

Applied genetic variance for the simulation of BVs was 1650

Effect of genomic pre-selection on genetic variances. Cows averaged over 50 replicates.

Legend:
- True control
- True gps
- Estimated control
- Estimated gps

y-axis: Gvar

x-axis: Birth years (1998, 2000, 2002, 2004, 2006, 2008)

Conclusions

- Genomic pre-selection can be detected by calculating yearly means of MS terms from EBVs

- In bulls, MS means clearly deviated from the zero expectation
  - Easy to obtain from the MS variance validation program

- At the start of genomic pre-selection, estimated genetic variance decreased
  - and started to increase again, when genomically pre-selected bulls become bull sires

- Applied simulation method was found useful to study the effects of genomic pre-selection

MTT