Universität Rostock Traditio et Innovatio

The use of a Tobit-like-classification in genetic evaluation of German trotters

A.-E. Bugislaus[1], E. Stamer[2], N. Reinsch[3]

[1]University of Rostock, Germany
[2]TiDa, Germany
[3]Leibniz Institute for Farm Animal Biology FBN, Germany

Actual breeding value evaluation in German trotters

- **Model:** BLUP animal model utilizing individual race results from all starting trotters
- **Traits:** Earnings per race
  Rank at finish
  Racing time per km

- **Question:** Do only trotters with earnings show their real racing potential whereas trotters without earnings are not driven to their limit?

Censored race results could be responsible for a potentially severe bias in actual genetic evaluation in German trotters

Objectives of the study

- Definition of censoring in race results of German trotters
- Tobit-like-Threshold-Model of racing performances
- Use of a real Tobit-Model for the censored trait racing time per km
- Comparison with linear model that treated all individual racing times per km as uncensored

Data for genetic estimations

- Total data set consisted of 105,981 race performances from 6,504 trotters *(mean of racing time per km = 79.7 s/km)*
- Data set involved 14,148 races with 7.5 participants in average
- Starting method was in all races the auto start
- Pedigree back to the fourth generation was included (20,703 animals)

Censoring in race results of German trotters and description of different used genetic models

Trait	Uncensored racing time per km (y)	Tobit-like-Threshold-Model for placing status	Censored racing time per km for Tobit-Model (y*)
Exemplary results of one race sorted by ranks at finish	y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8 y_9 y_{10}	5 4 3 2 1 0 0 0 0 0	y_1 y_2 y_3 y_4 y_5 y_6^* y_7^* y_8^* y_9^* y_{10}^*
Model	Linear-Model: Bayesian analysis	Threshold-Model: Bayesian analysis	Tobit-Model: Bayesian analysis with data augmentation
Program	LMMG (REINSCH, 1996)	LMMG_TH (REINSCH, 1996)	LMMG_TOB (REINSCH, 2011)

Tobit-Model for censored trait racing time per km:
Data augmentation

- For each $y^* > y_5$ the threshold is determined as standardized value:

$$t = \frac{y_5 - x_i'\beta - z_i'u}{\sigma_e}$$

- A random variable $a_i < t$ is drawn from a truncated standard normal distribution and is subsequently transformed to the original scale:

$$y_i^* = a_i * \sigma_e + x_i'\beta + z_i'u$$

- For one iteration y_i^* is treated as observation, in the following round y_i^* is again determined

Univariate genetic-statistical model

$$y = Xb + Z_1a + Z_2pe + e$$

y: vector of observations containing either the uncensored or censored trait racing time per km or the threshold trait placing status of each trotter in each individual race

b: fixed effects

a: random animal effect

pe: random permanent environmental effect

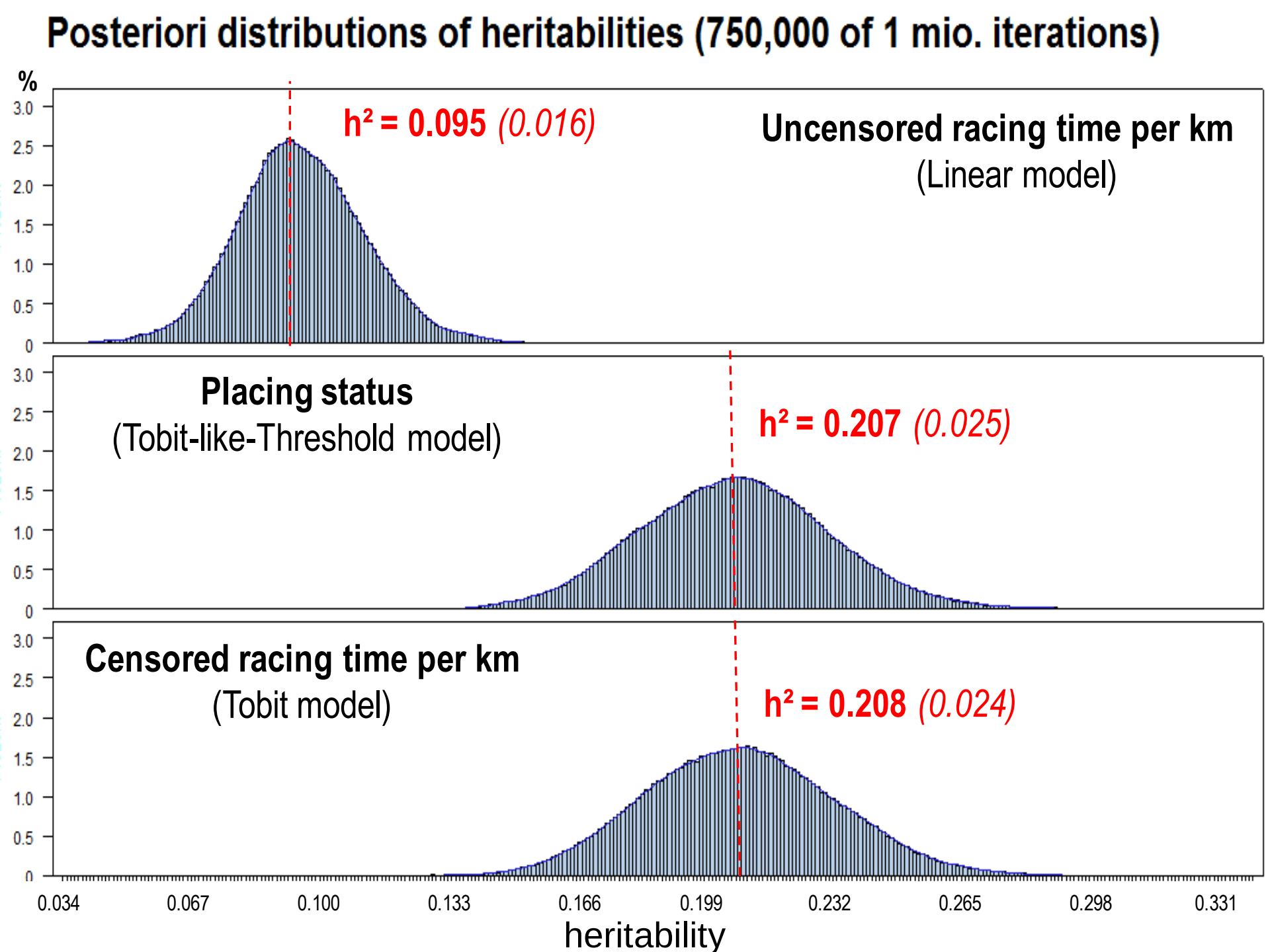
e: residual effect

X, Z_1, Z_2: incidence matrices

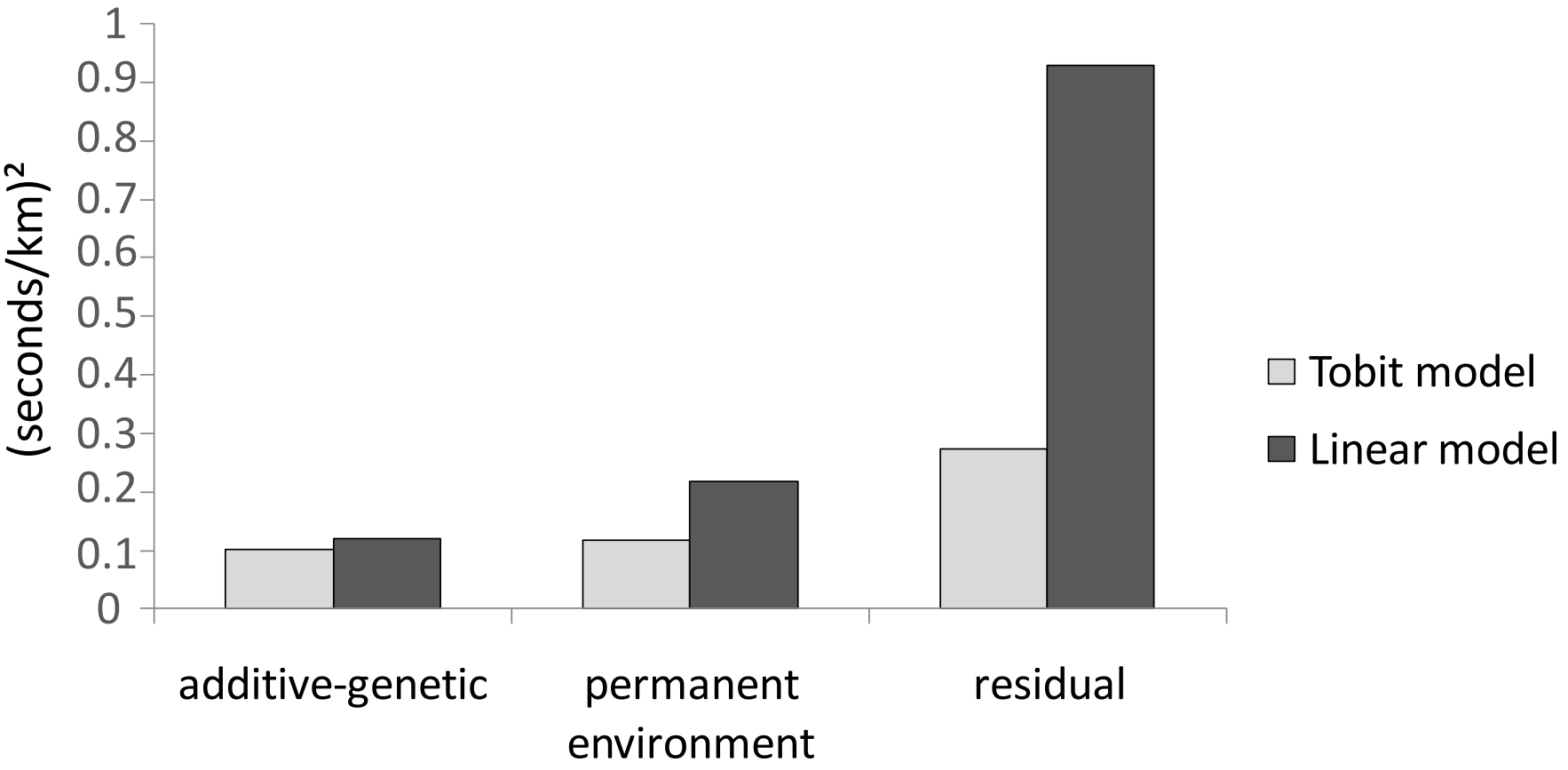
- 1 million cycles were generated (Gibbs sampling algorithm)
- As burn-in period 250,000 rounds were considered

Fixed effects

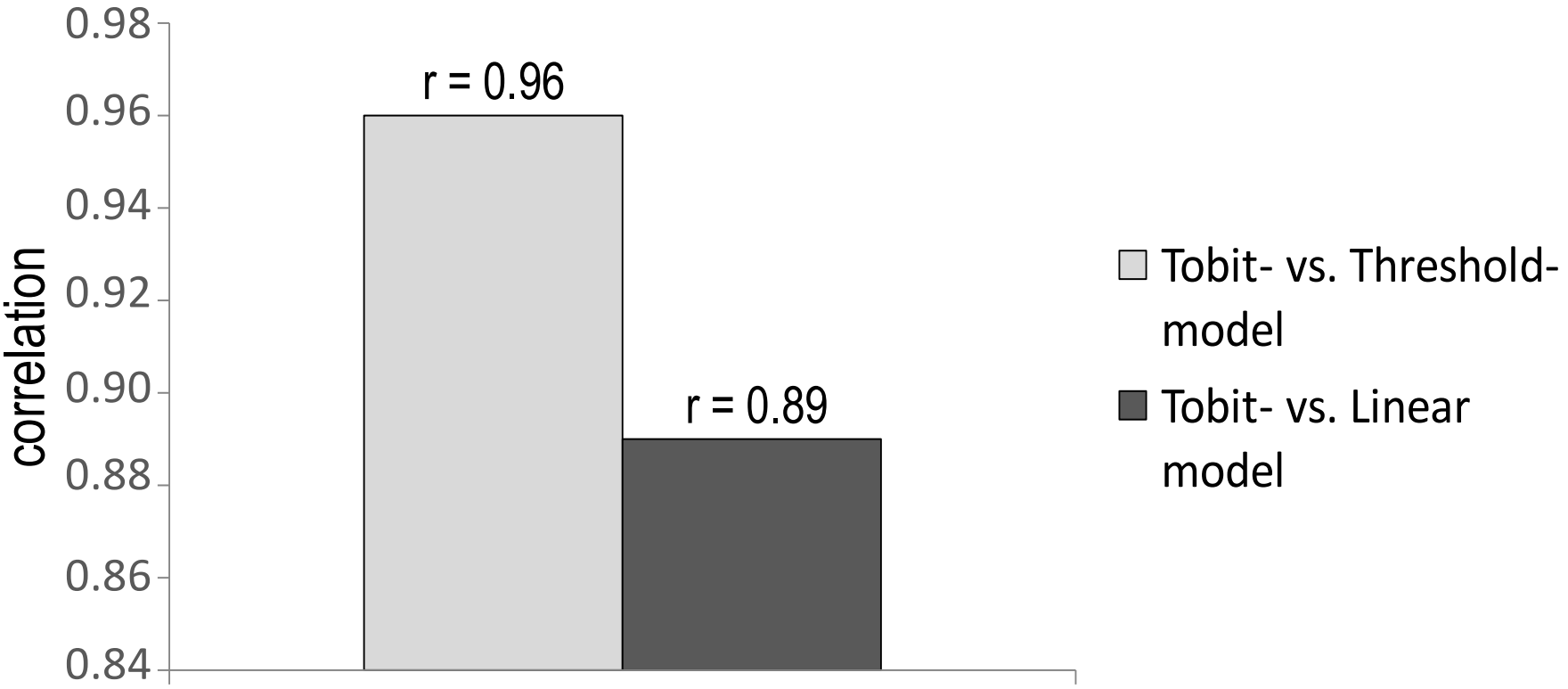
- sex (stallion, mare, gelding)
- age of trotter (12 classes)
- year-season of race (three months are one season)
- condition of race track (fast, good, medium, heavy, muddy)
- distance of race (10 distance classes)
- driver (1, …, 1572)
- each individual race (1, …, 14148)


Posteriori distributions of heritabilities (750,000 of 1 mio. iterations)
%
3.0
2.5
2.0
1.5
1.0
0.5
0
h² = 0.095 (0.016)
Uncensored racing time per km
(Linear model)
Placing status
(Tobit-like-Threshold model)
h² = 0.207 (0.025)
Censored racing time per km
(Tobit model)
h² = 0.208 (0.024)
0.034
0.067
0.100
0.133
0.166
0.199
0.232
0.265
0.298
0.331
heritability

Variance components estimated for racing time per km with two different models



Rank correlations between breeding values estimated with different genetic models



Conclusion

- Trotters without earnings didn‘t show their real racing potential and should be regarded as censored observations.
- Tobit-(like-Threshold)-Models with censored race results represented good suitability for genetic evaluation.
- Heritability estimates for the threshold trait placing status and the censored trait racing time per km were very similar.
- Also the high rank correlation between the breeding values of placing status and the breeding values of censored racing time per km showed great agreement.

Thank you for your attention!