Tests of single-step GBLUP for production traits in US Holsteins

Y. Masuda[1], I. Misztal[1], P. M. VanRaden[2], and T. J. Lawlor[3]

1 University of Georgia, USA
2 AGIL, USDA, USA
3 Holstein Association USA, Inc., USA

Purposes

- To show computational stability
  - Calculate inbreeding in $\mathbf{A}^{-1}$ to be consistent with $\mathbf{A}_{22}^{-1}$ and $\mathbf{G}^{-1}$
  - Include unknown parent groups (UPGs) for genotyped animals in $\mathbf{H}^{-1}$
- To compare genetic trends between the traditional BLUP-EBV and ssGBLUP-GEBV with all available genotypes
  - One way to detect genomic pre-selection effect
  - Milk, fat, and protein yield for US Holsteins
- To discuss possible effects of trend differences on MACE

Data

Data	Description	Number of records
Phenotypes	Milk, fat, and protein yield from US Holsteins; from 1990 to 2015	50,970,954
Pedigree	3 generations back from phenotyped cows or genotyped animals; 300 UPGs	29,651,623
Genotypes	Both male and female; including young bulls and heifers (#SNPs = 60671)	764,029

Model

- Three-trait repeatability model

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{ZQg} + \mathbf{Wp} + \mathbf{Hs} + \mathbf{e}$$

- Relationship matrix

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}_{APY}^{-1} - \omega \mathbf{A}_{22}^{-1} \end{bmatrix}$$

  - ω: a parameter to compensate for missing pedigree can be replaced by UPGs
  - $\mathbf{G}_{APY}^{-1}$: 18,359 core animals randomly selected

Inbreeding and UPGs

- QP-transformation for $\mathbf{A}^{-1}$ (Westell et al., 1988; Quaas 1988)

$$\mathbf{A}^{*}=\begin{bmatrix}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{Q}\\ -\mathbf{Q}'\mathbf{A}^{-1} & \mathbf{Q}'\mathbf{A}^{-1}\mathbf{Q}\end{bmatrix}$$: Henderson’s rule with inbreeding

- QP-transformation for $\mathbf{H}^{-1}$ (Misztal et al., 2013)

$$\mathbf{H}^{*}=\underbrace{\mathbf{A}^{*}+\begin{bmatrix}0 & 0 & 0\\ 0 & \mathbf{G}^{-1}-\mathbf{A}_{22}^{-1} & 0\\ 0 & 0 & 0\end{bmatrix}}_{\text{Already considered}}+\underbrace{\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & -(\mathbf{G}^{-1}-\mathbf{A}_{22}^{-1})\mathbf{Q}_{2}\\ 0 & -\mathbf{Q}_{2}'(\mathbf{G}^{-1}-\mathbf{A}_{22}^{-1}) & \mathbf{Q}_{2}'(\mathbf{G}^{-1}-\mathbf{A}_{22}^{-1})\mathbf{Q}_{2}\end{bmatrix}}_{\text{Extra terms (Matilainen et al. 2016)}}$$

Iteration in PCG

Model	**Trait**	**Inb. in $\mathbf{A}^{-1}$**	**UPG in $\mathbf{H}^{-1}$**	$\boldsymbol{\omega}$	**# of iterations***
ssGBLUP	Single	No	No	Any	N/A
	Single	Yes	No	0.90	452
	Three	Yes	No	0.90	1,274
	Three	Yes	Yes	1.00	464
BLUP	Three	Yes			402

* Computing with 6 cores; Convergence criterion: $Cr = 10^{-15}$.

Timing

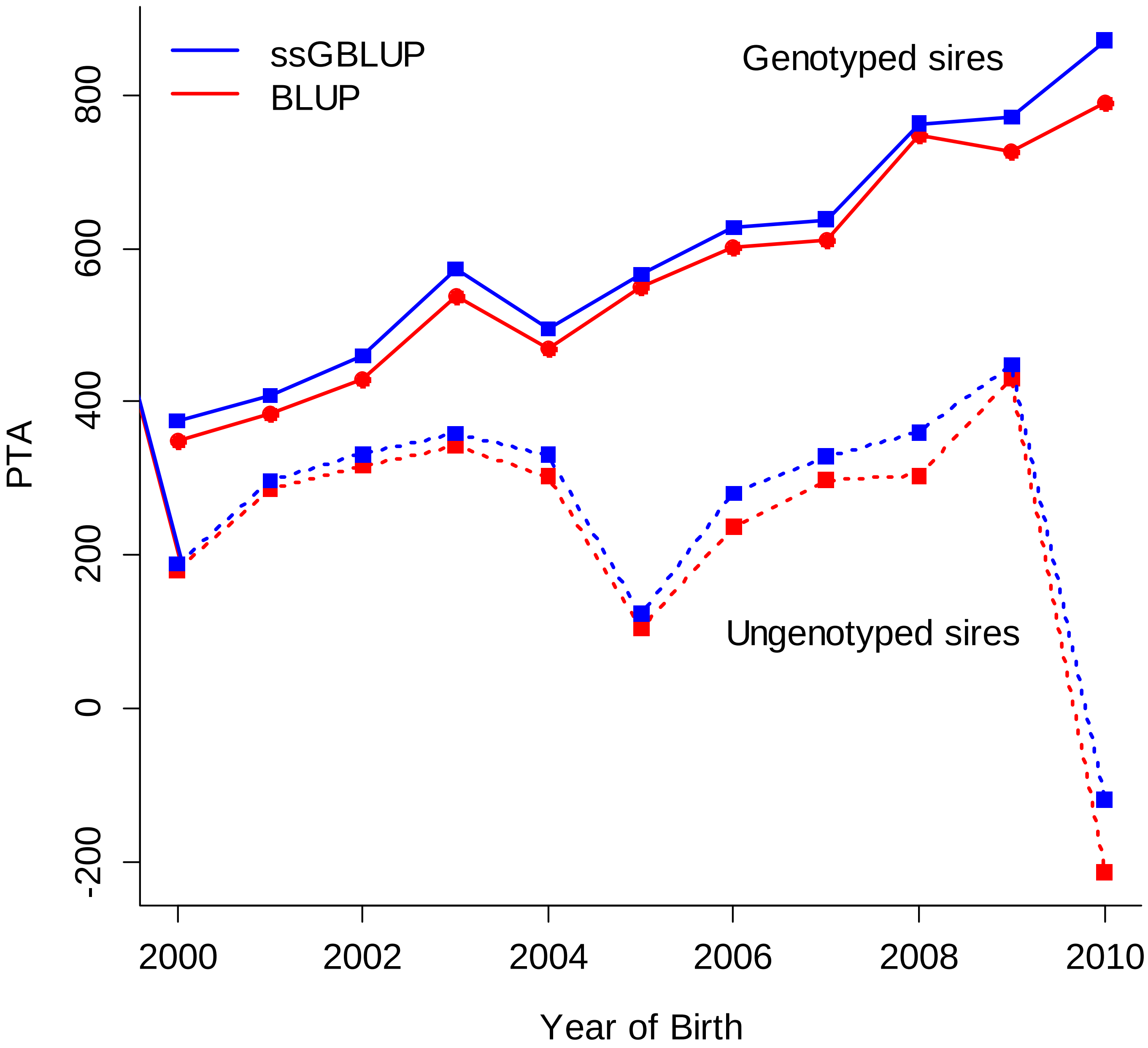
Model	Trait	Inb. in $\mathbf{A}^{-1}$	UPG in $\mathbf{H}^{-1}$	ω	Timing per iteration	Total Time*
ssGBLUP	Single	No	No	Any	N/A	N/A
	Single	Yes	No	0.90	40 sec.	5 h 10 min
	Three	Yes	No	0.90	80 sec.	28 h 30 min
	Three	Yes	Yes	1.00	90 sec.	12 h 58 min.
BLUP	Three	Yes			51 sec.	6 h 31 min.

* Affected by background jobs; Computing with 6 cores; Excluding computations for G_{APY}^{-1} etc.

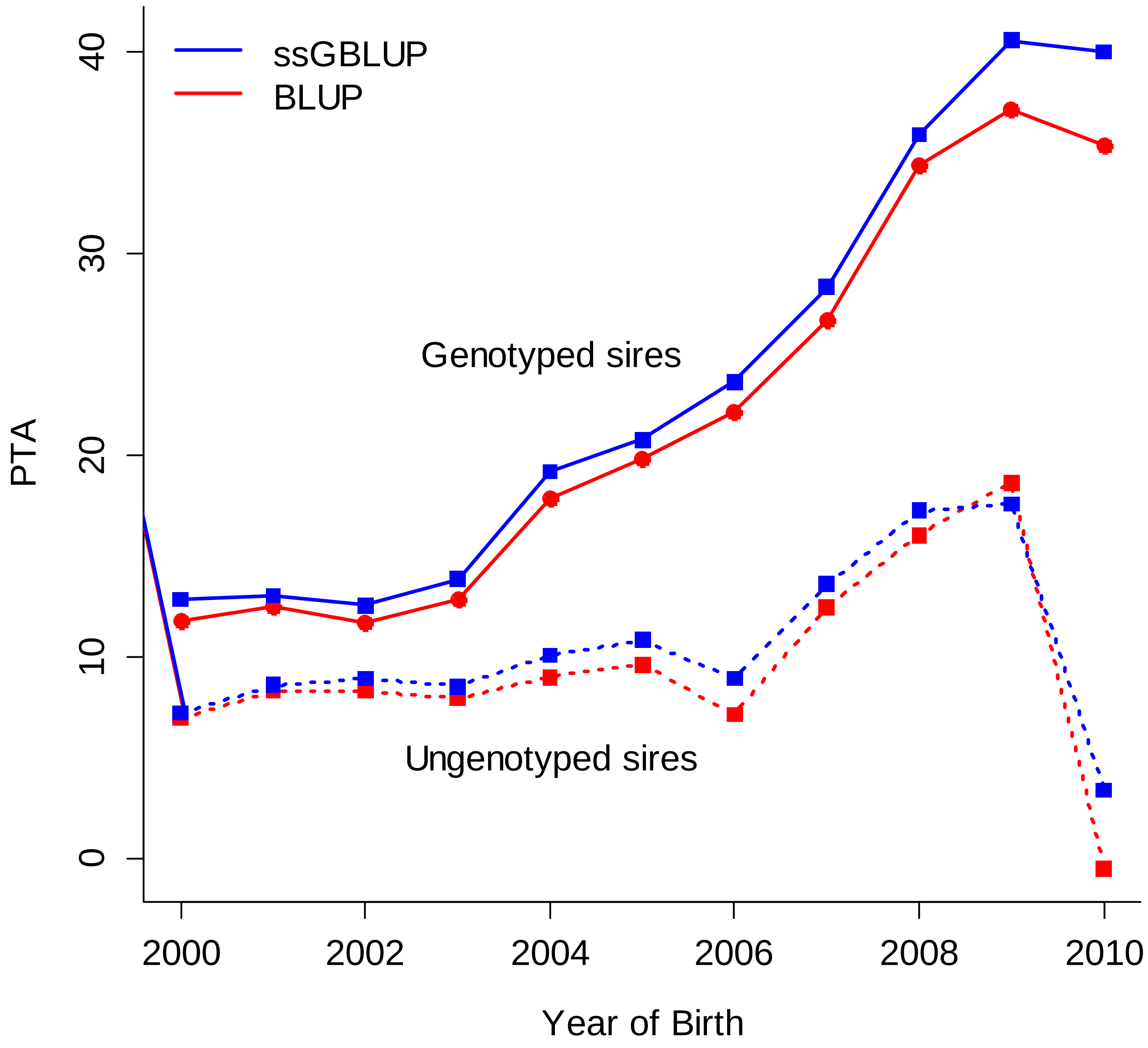
Genetic trends

- Expectations
  - Higher genetic trends from ssGBLUP for genotyped animals.
    = The traditional BLUP is biased down.
  - Because … they are not getting the full credit for how much their Mendelian sampling is above PA.
- Implication for MACE
  - MACE is biased down.
  - The current MACE will evaluate bulls being lower than their domestic within-country ssGBLUP results.

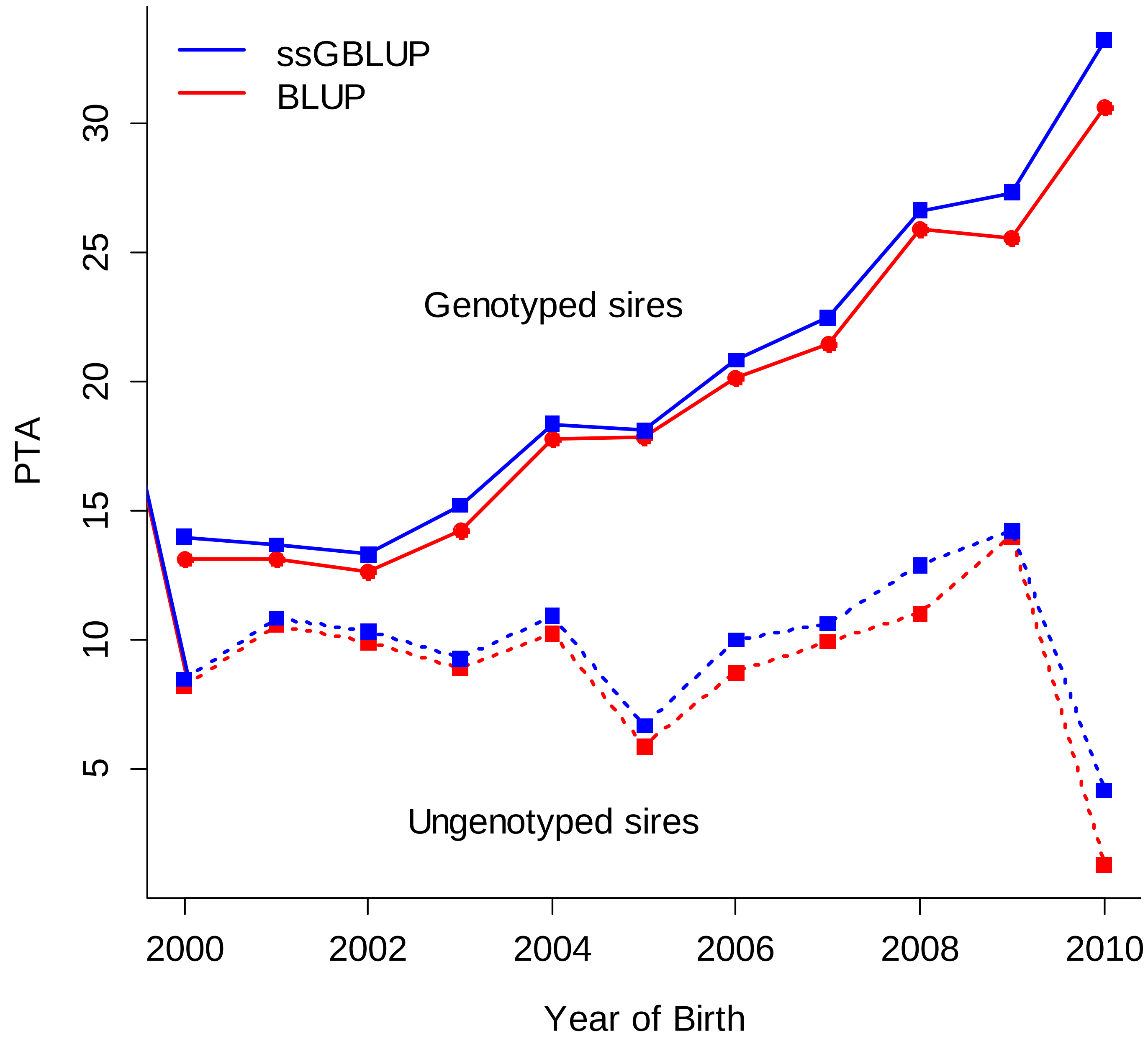
Sires with at least 50 daughters with record(s): milk



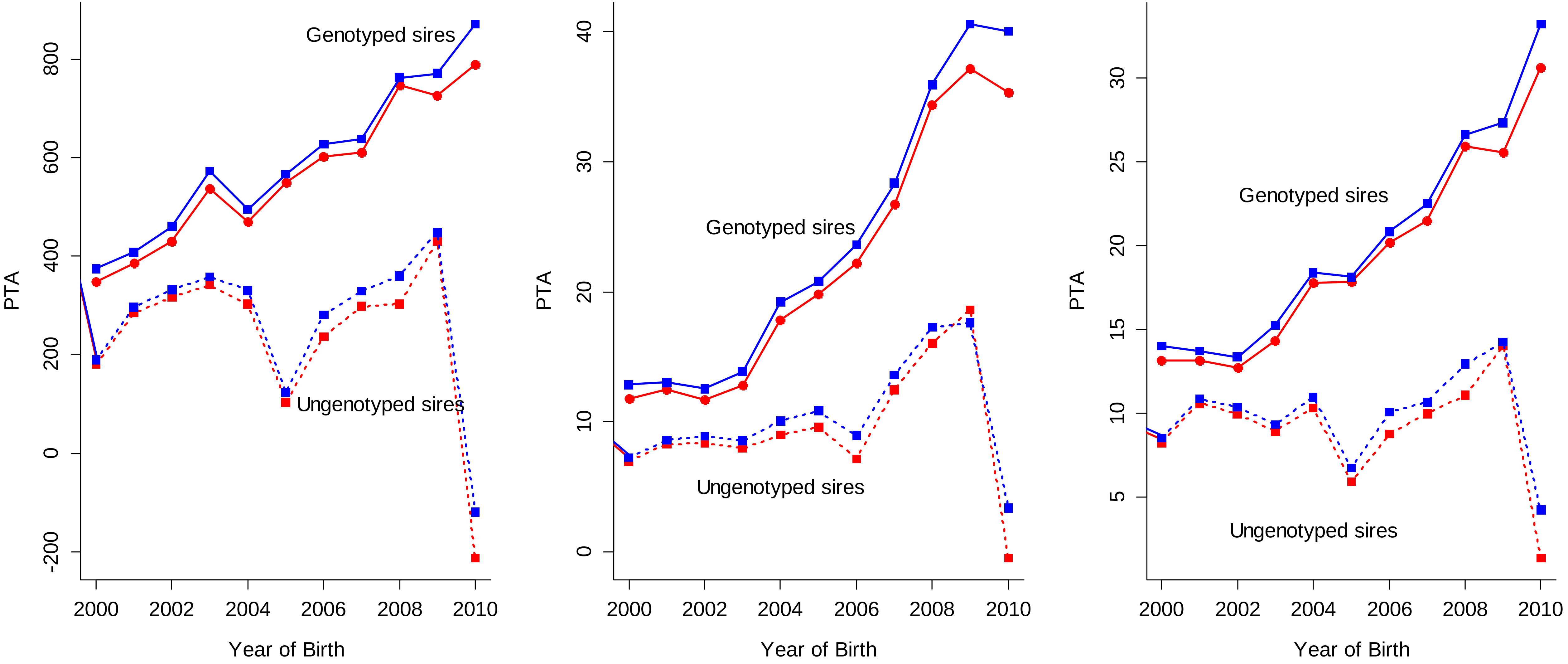
Sires with at least 50 daughters with record(s): fat

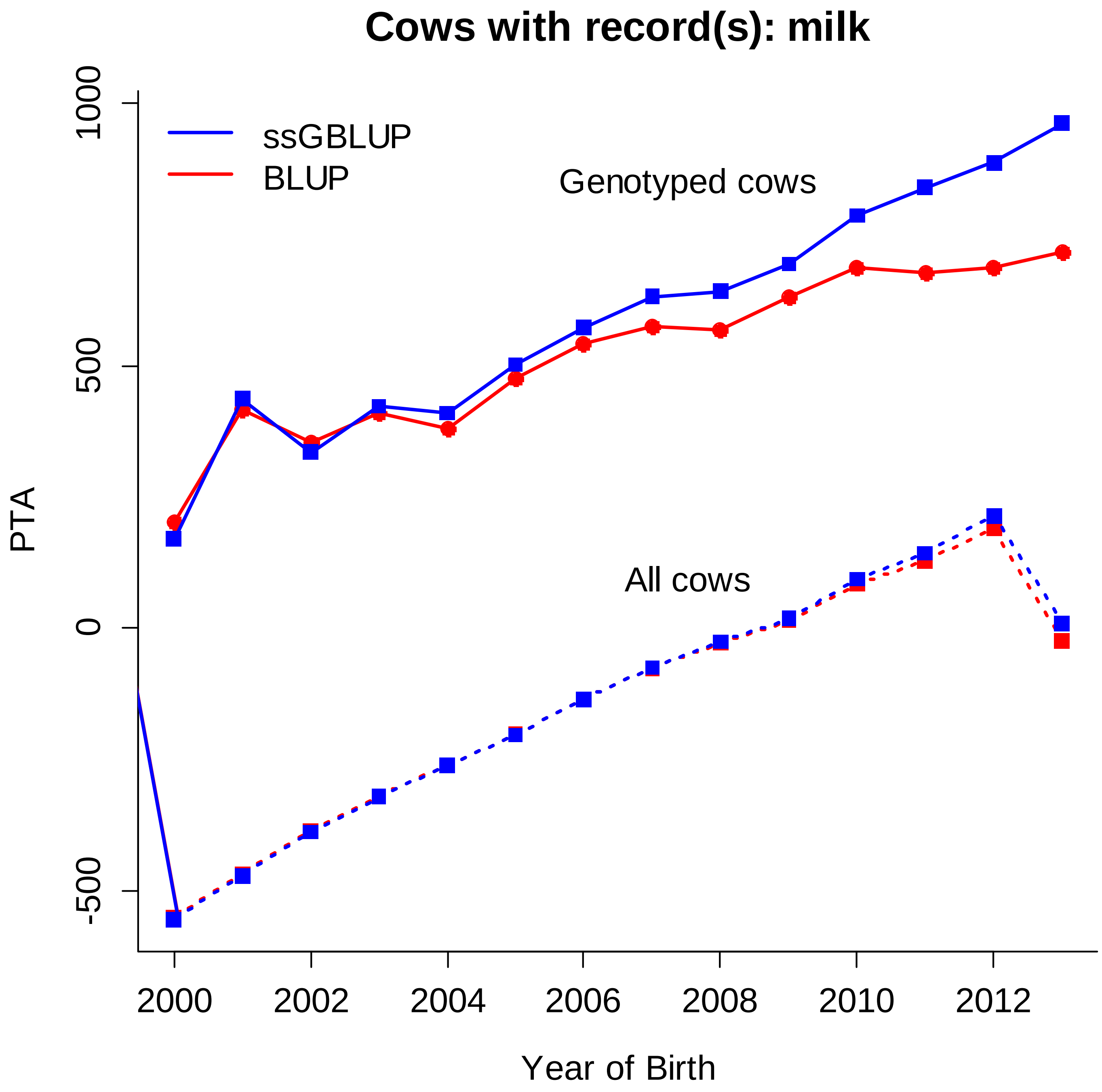


Sires with at least 50 daughters with record(s): protein

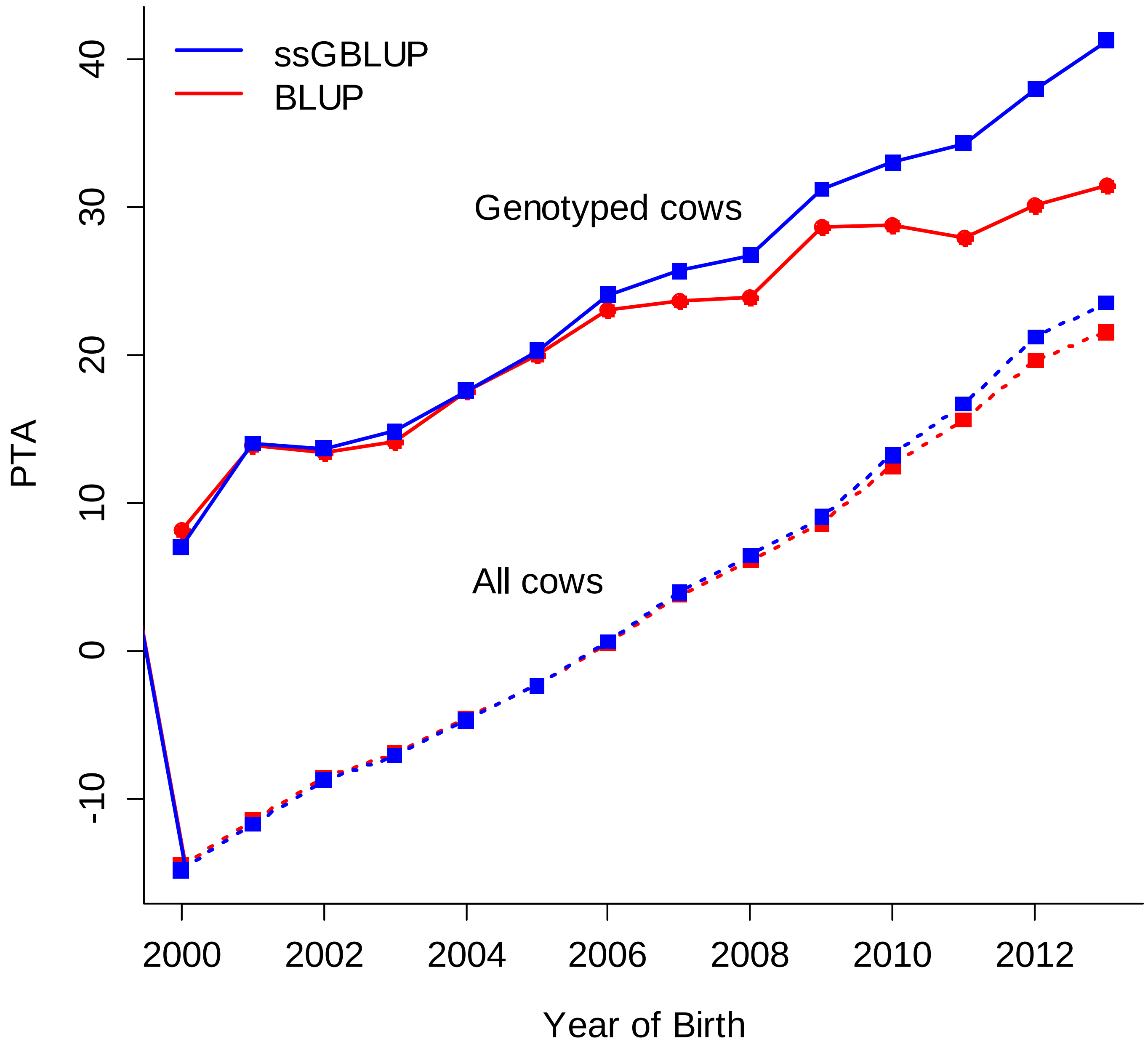


Sires with at least 50 daughters with record(s): milk,fat, and protein

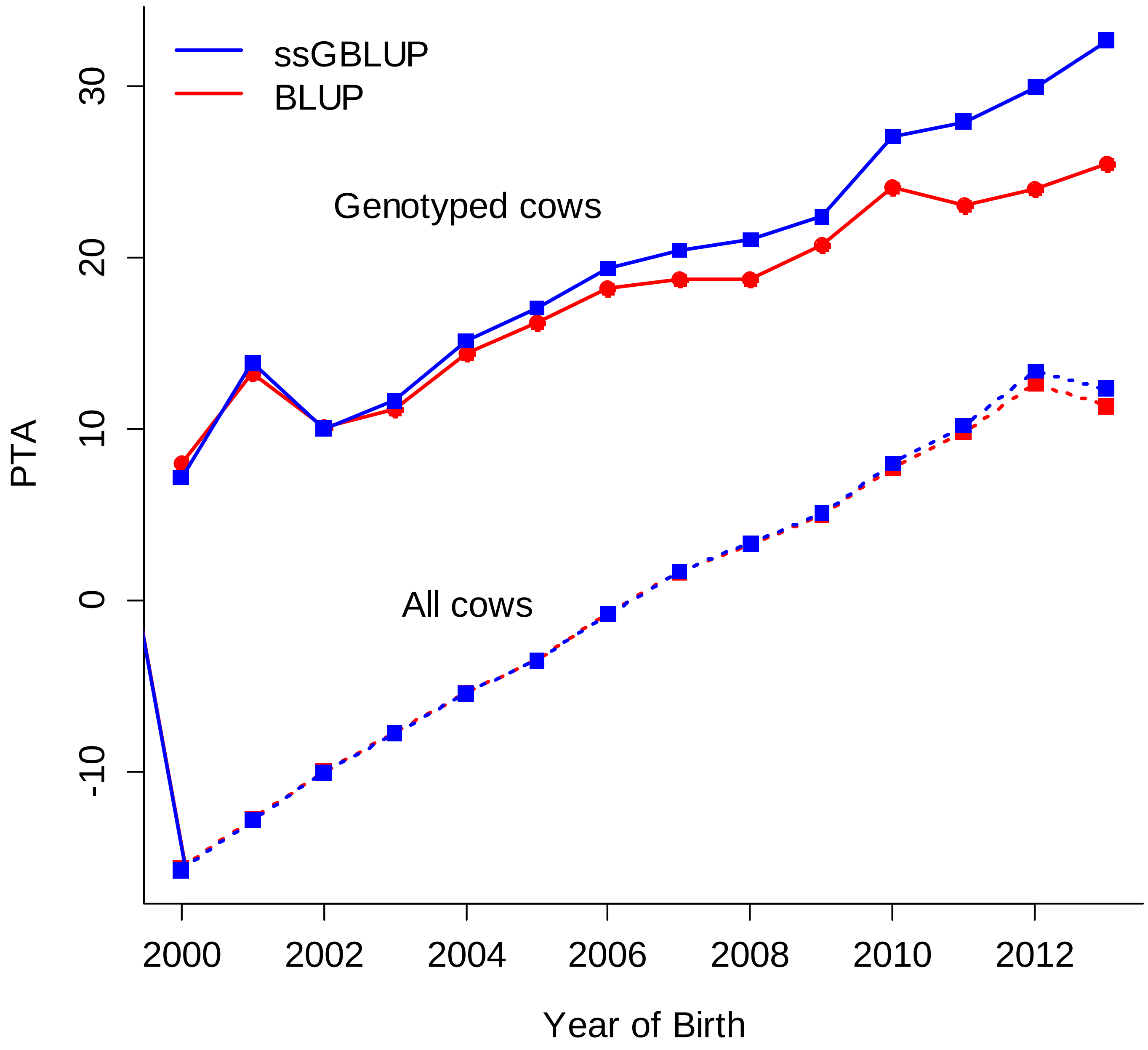



Cows with record(s): milk
ssGBLUP
BLUP
Genotyped cows
All cows
PTA
1000
500
0
-500
2000
2002
2004
2006
2008
2010
2012
Year of Birth

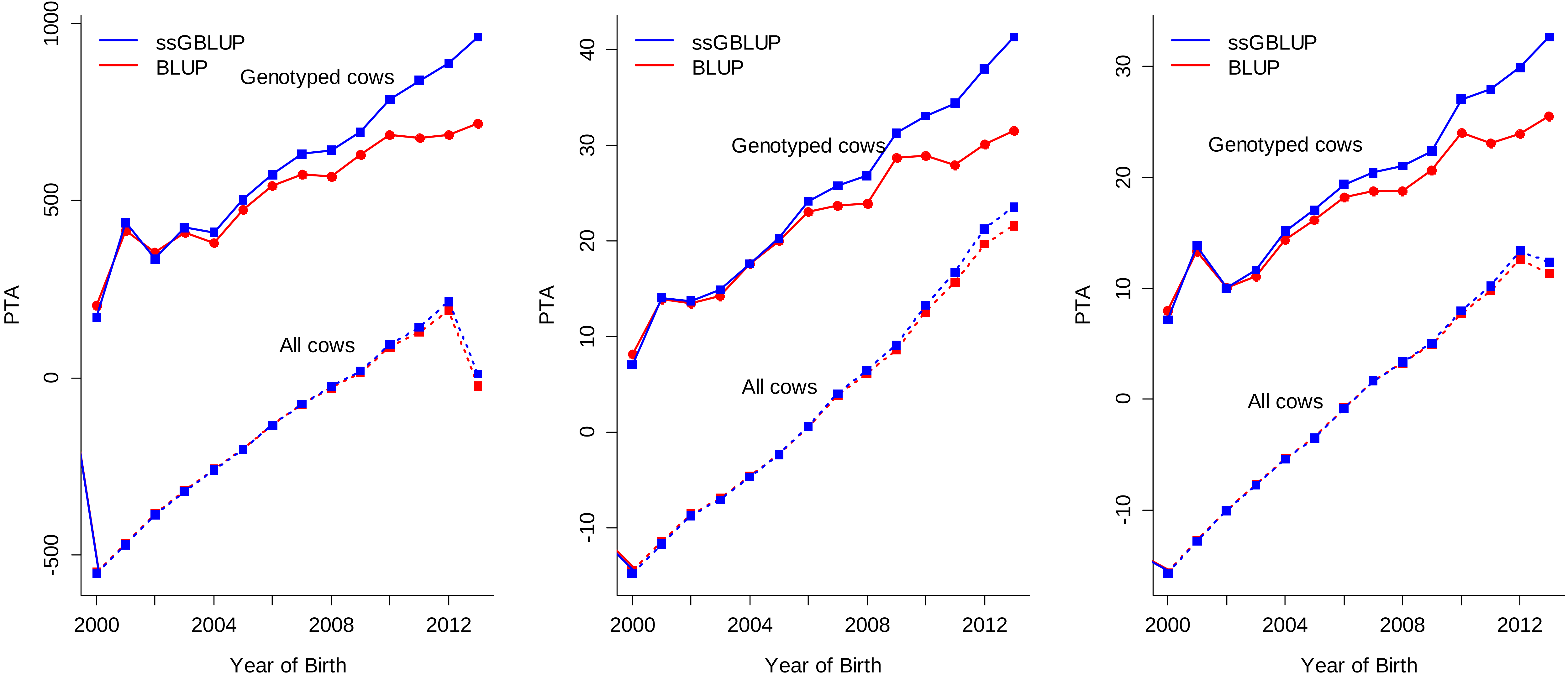
Cows with record(s): fat



Cows with record(s): protein



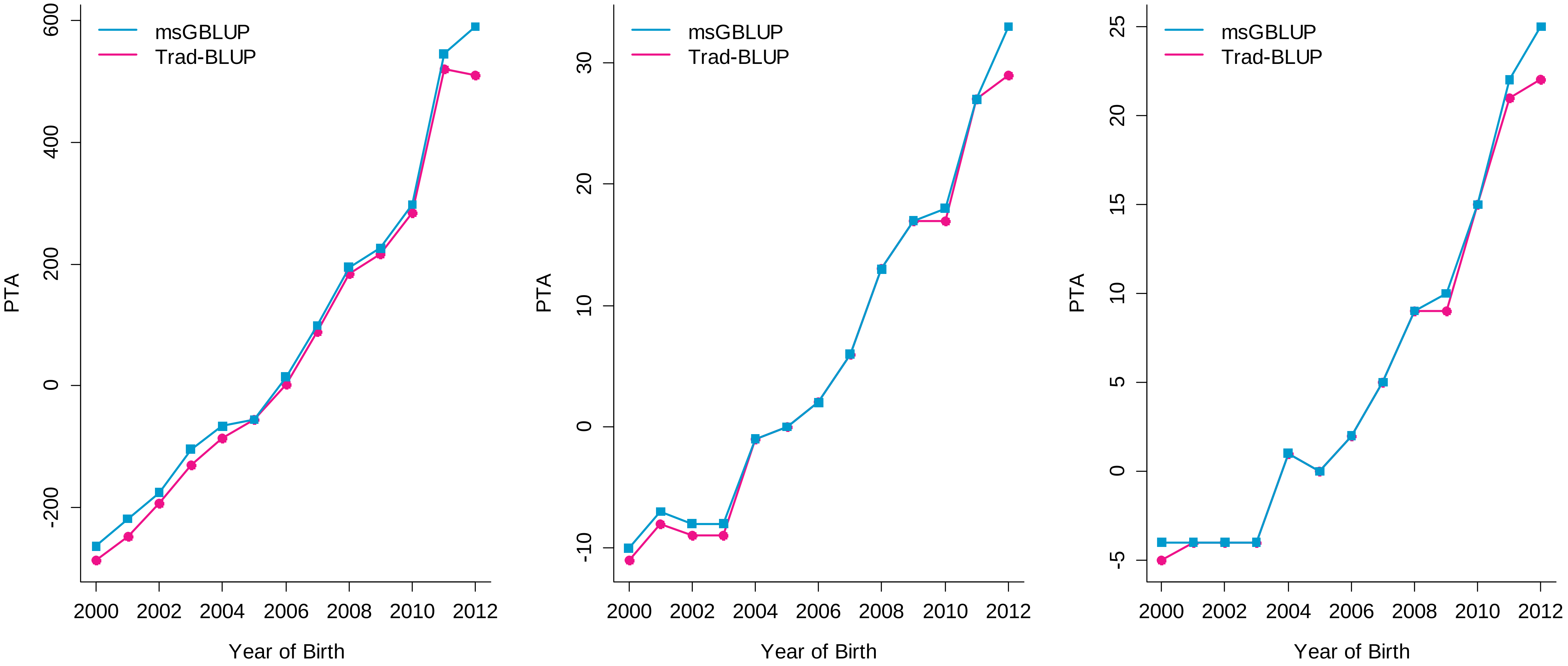
Cows with record(s): milk, fat, and protein



The US official evaluation

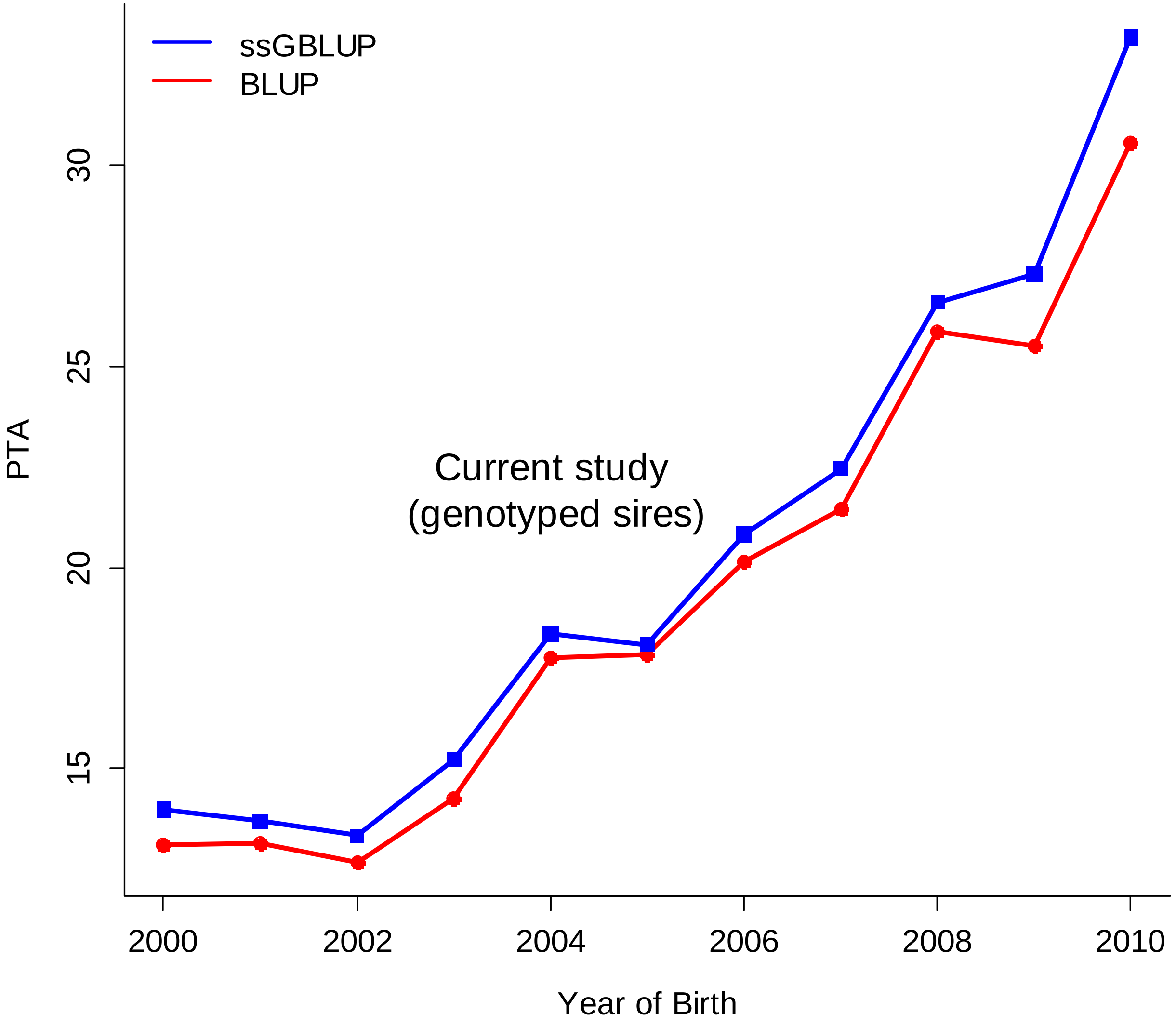
- Official multi-step method (msGBLUP)
  - One of the goals: makes it similar to BLUP i.e. minimized difference between PTA and GPTA.
  - Includes foreign information (MACE and foreign dams).
  - Includes an adjustment to reduce cow bias.
  - Single-step transfers genomic information from progeny to parents. Multi-step does not.
- Trends
  - Sires with at least 10 daughters with record(s)
  - Cows with record(s)

US Official GPTA for proven sires: milk, fat, and protein

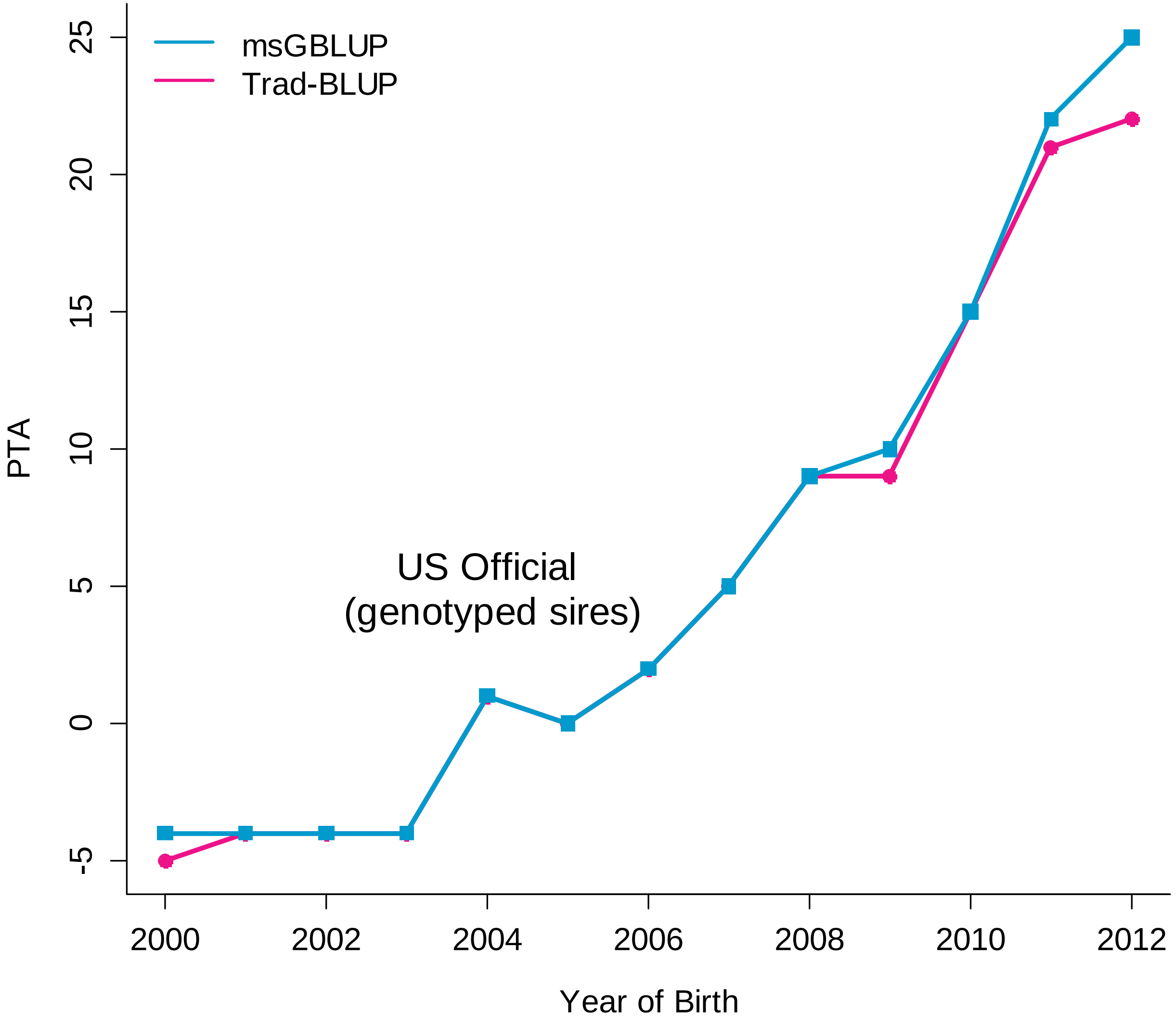


* At least 10 daughters with record(s)

Comparison of genetic trends in protein yield for proven sires

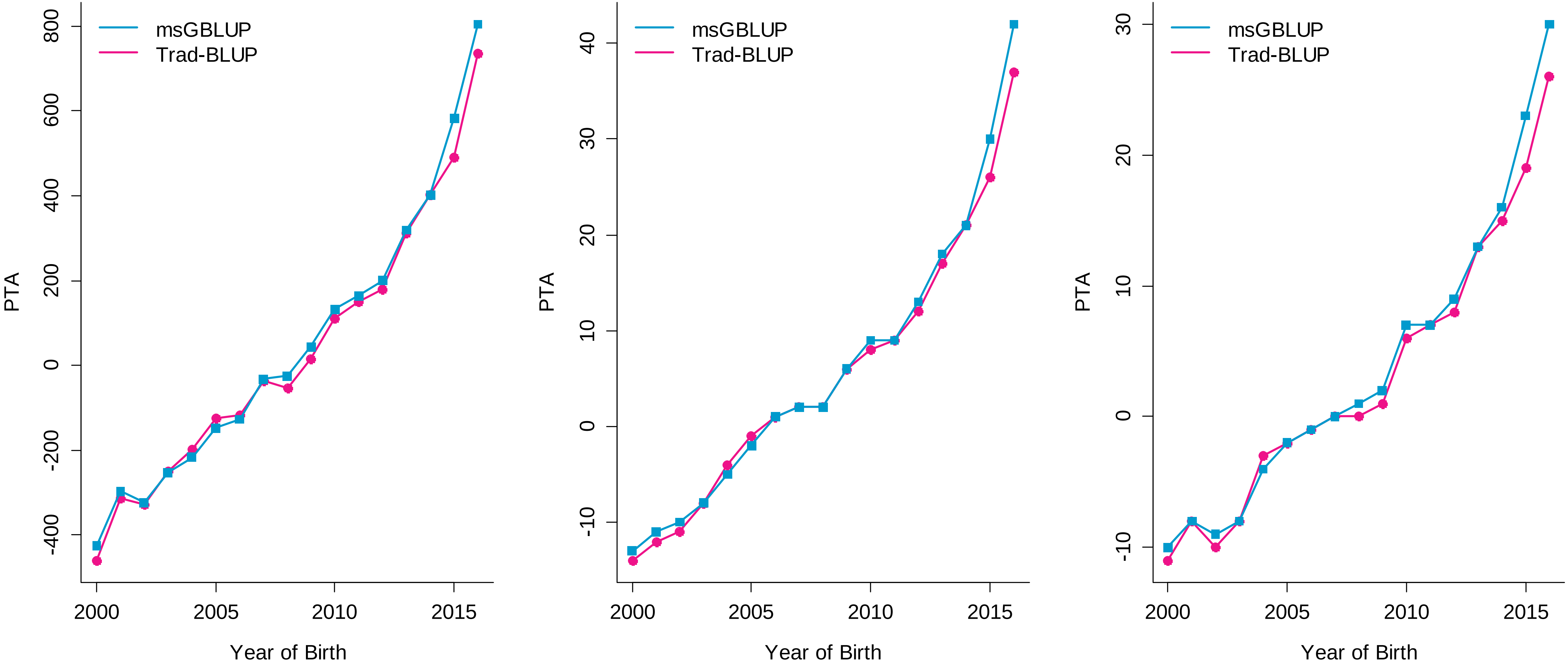


* At least 50 daughters with record(s)

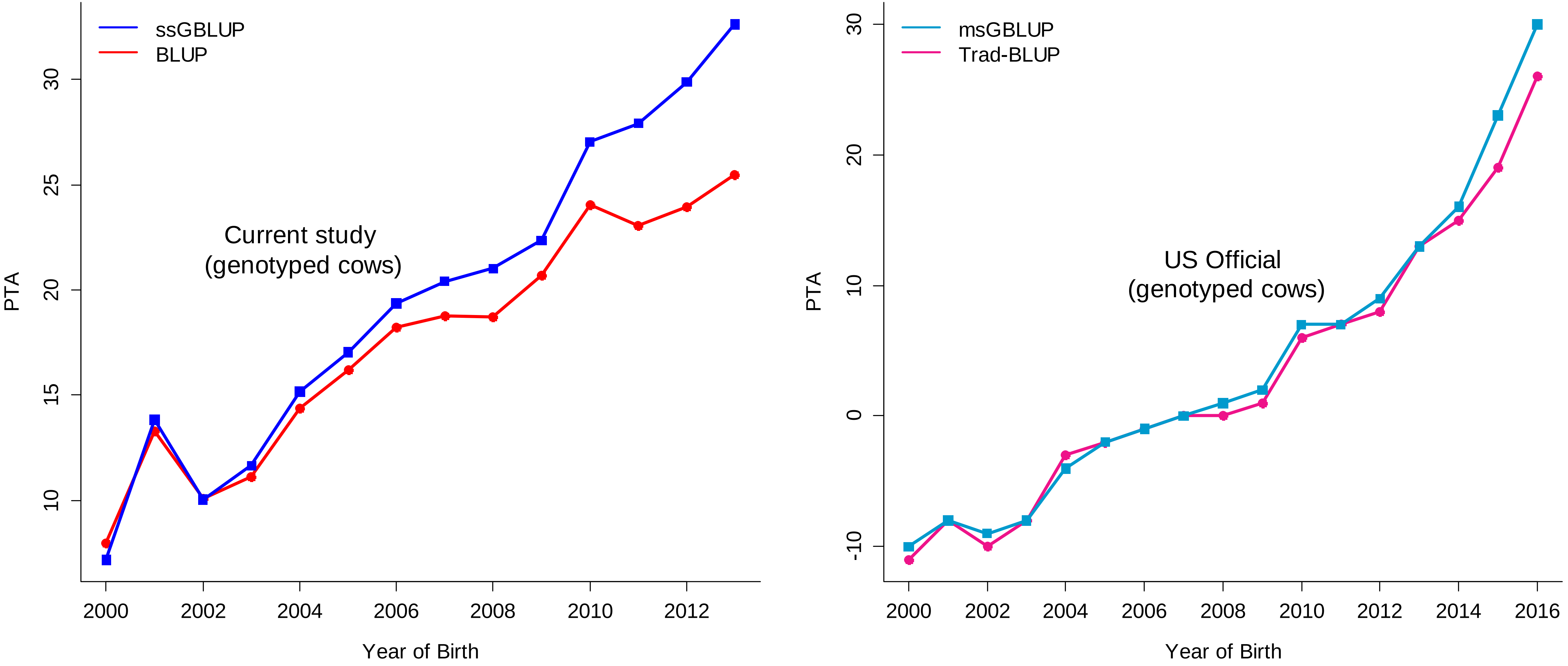


* At least 10 daughters with record(s)

US Official GPTA for cows with record(s): milk, fat, and protein



Comparison of genetic trends in protein yield for cows with record(s)



Questions

- Is ssGBLUP too high in recent years, or are msGBLUP and BLUP too low because genomic pre-selection is not accounted for?
- Is recent progress underestimated officially for genotyped cows and proven bulls (domestic and foreign)?
- If countries publish ssGBLUP, but send BLUP for MACE, will only foreign bulls be underestimated?
- Could other statistical methods in MACE provide unbiased foreign and domestic EBVs?

Summary

- Single-step GBLUP gets a stable convergence.
- Single-step GBLUP provides very similar genetic trends to the traditional evaluation except for the last few years.
- Bull trends are more similar to the USDA official trends.
- Genetic trend differences indicate that traditional BLUP evaluations appear to be showing the effect of genomic pre-selection bias.
- Trend differences in ssGBLUP vs. BLUP will cause domestic vs. MACE bull differences.
- Foreign bulls evaluated under MACE will, for the most part, be biased down when compared to domestic bulls evaluated with ssGBLUP.

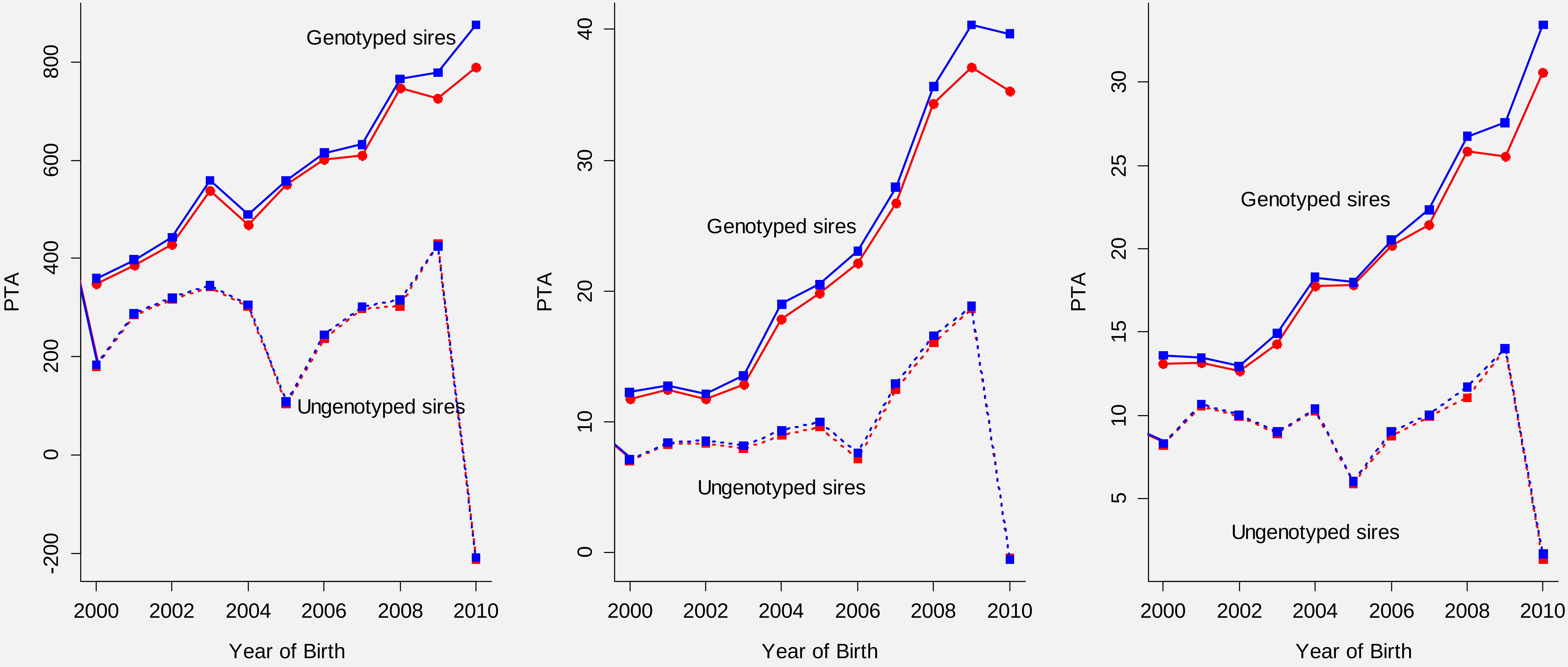
Acknowledgement

- USDA NIFA for partial financial support.
- Council of Dairy Cattle Breeding for phenotype, genotype, and pedigree data.
- John Cole and Melvin Tooker (USDA-AGIL) for preparing initial data sets and a computing environment.

Appendix

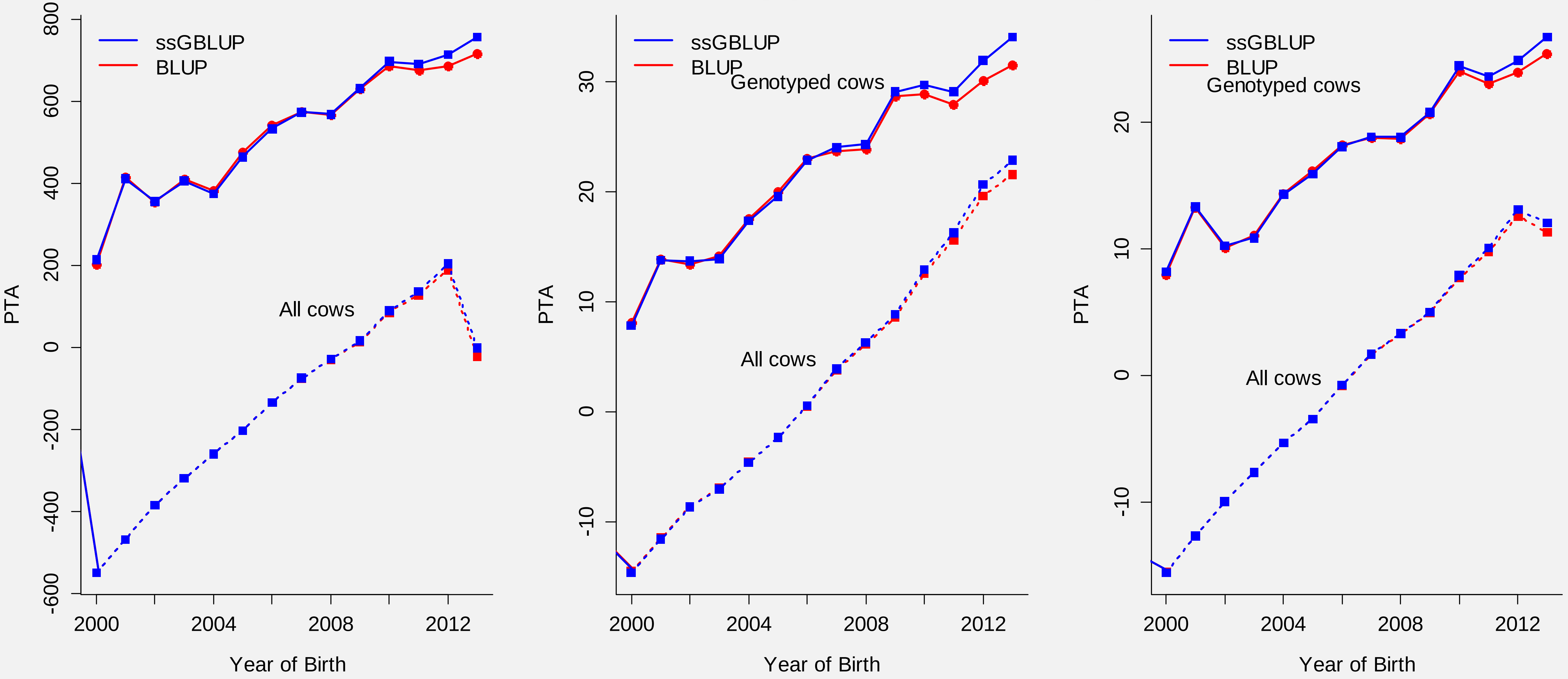
Single-step GBLUP with only bull genotypes

Sires with at least 50 daughters with record(s): milk,fat, and protein



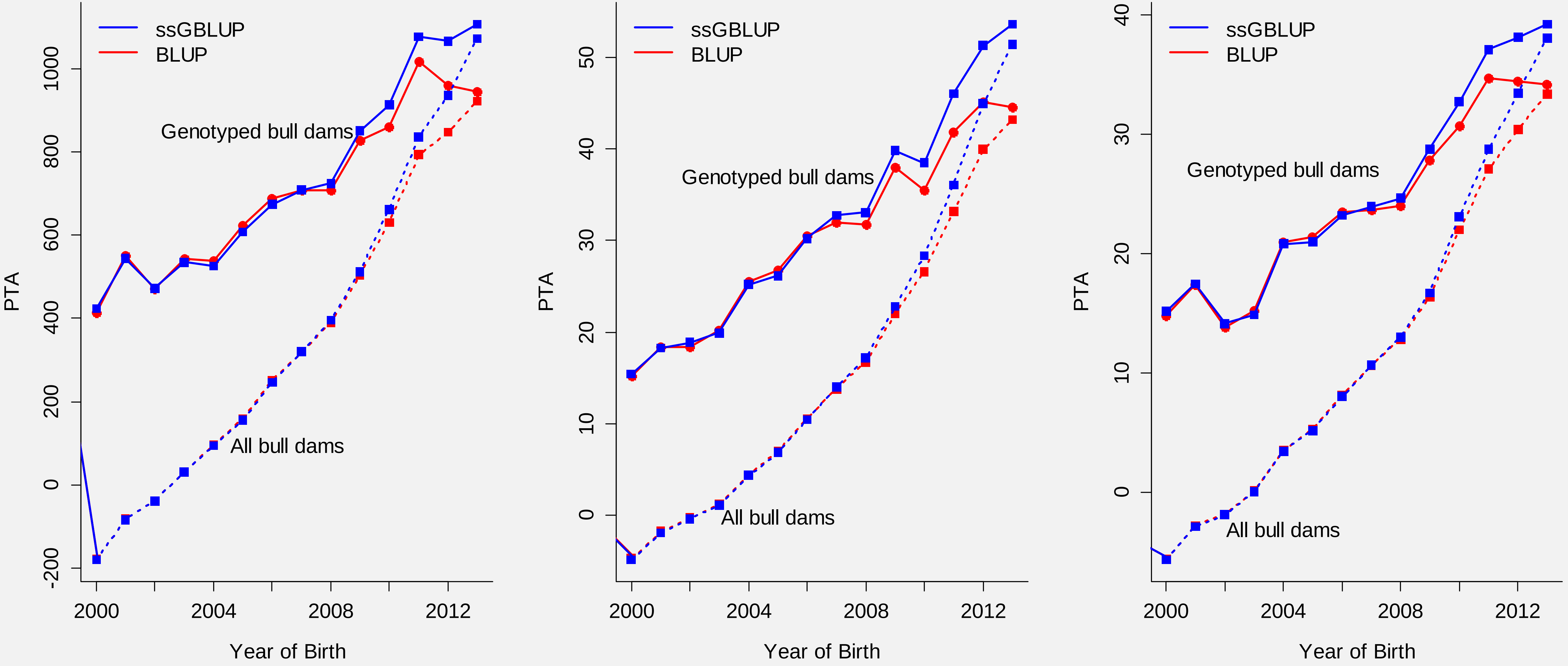
From analyses with bull genotypes only

Cows with record(s): milk, fat, and protein



From analyses with bull genotypes only

Bull dams with record(s): milk, fat, and protein



From analyses with bull genotypes only