How the study of the number of starts and the starting status can inform about selection bias when using earnings for breeding evaluations in racehorses

59th Annual Meeting of the EAAP, Session H36-4, Vilnius, Lithuania, August 27th 2008.

Langlois B. Blouin C. INRA-SGQA 78350 Jouy-en-Josas (France) bertrand.langlois@jouy.inra.fr

Abstract:
In France, horses earning some money represent 30% of the horses born for French Trotters, 41% of the Thoroughbred horses in flat races and only 23% in jumping races. The question of a selection bias, when using earnings for breeding value estimation is therefore raised. To analyse this matter we looked at the number of starts of 2-5 year-old French Trotters and Thoroughbreds born between 1996 and 2000 (58 841 and 19 957 resp.). Two variables were studied:-1 a none or all variable (starter/non starter) -2 number of starts excluding zero in trotting, flat and jumping races. The first variable reflects the phenomenon being prepared to participate in races. It appeared "highly heritable" according to very high sire components (14%, 14%, 21% resp.) and also very high dam components (11%, 13%, 13% resp.). The second variable seemed however, mainly environmentally or trainer dependant. The sire component was very low in trotting and jumping races (1%, 4% resp.) and average (10%) in flat races. A noticeable dam effect was observed in all three cases as shown by the greater dam component (3%, 11%, 14% resp.).The distribution of the number of starts for horses earning money compared to those not earning anything allows with some hypotheses to estimate the number of horses prepared to race but that never started. It leads to the correction of the apparent selection rate from 30% to 86% in trotting, from 23% to 61% in jumping races and from 41% to 61% in flat races. It can be concluded that earning horses can be assumed as representative in trotting but that they are relatively pre-selected in galloping races. The real selection rates seem, however, much greater than the observed ones because a lot of horses born that do not appear in races are in fact not tested.

Key words:
Horse # French Trotter # Thoroughbred # selection bias # breeding value estimation

Introduction

Breeding value estimation for race horses allows planning a better selection policy than mass selection on phenotype as it is currently done. However, in the two cases, the selection criterion is based on public racing results. One can note that horses earning money have to start at least in one race and that they represent only 30% of the horses born for the French Trotter, 41% of the Thoroughbred horses in flat races and 23% in jumping races. The question of a selection bias, when using earnings for breeding value estimation is therefore often raised.

In this paper we will first estimate the genetic parameters of different criteria used to describe this selection process. We will secondly analyse more deeply the question of the number of starts to try to estimate the real selection basis which is probably not the number of horses born but the number of horses prepared to race.

Materials and methods

Data:

We considered all French Trotters and Thoroughbreds born between 1996 and 2000 in order to get five promotions older than 5 years in 2006. That was in total 58 841 trotters from 976 sires and 23 913 dams and 19 957 Thoroughbreds from 887 sires and 9 571 dams. The racing results were obtained from the "Société d'Encouragement à l'Elevage du Cheval Français" (SECF), (http://www.cheval-français.com) and France Galop (http://www.france-galop.com). Pedigree data were provided by the "Système d'Identification Répertoriant les Equidés (SIRE), the national computer system for Stud-books (http://www.haras-nationaux.fr).

Variables:

1- The starting status. It is an all or none variable which takes the value 0 for non starter and 1 for starter. It will be analysed under a probit scale.
2- The number of starts, zeros included.
3- The number of starts zeros excluded.
4- The number of starts of earning horses.

Models:

Sire, Sire and dam and animal models were applied to the data using the ASREML software (Gilmour et al. 2002).

The fixed effects considered were the following:

1- Year (5 levels from 1996 to 2000)
2- Sex (2 levels, males and females)
3- Breeding area (3 levels)
   a- Normandy (76, 27, 28, 61, 14, 50)
   b- Greater west with Brittany with north and Ile de France-centre (29, 22, 56, 35, 44, 85, 17, 79, 49, 53,,72)+(62, 59, 02, 80, 60)+(77, 78,91, 93,94,95, 45, 41,37, 36, 86,87,16, 23, 63, 15, 43, 19)
   c- Rest of France (southwest, southeast, east and greater Lyon)
4- Breeder's category (4 levels according to the number of horses produced over the period: 1-2; 3-5; 6-8; 9+)

Studies of the distributions:

The distribution of the number of starts can be cut into two parts: that of horses earning money and that of horses not earning anything. This latter one can be modelled by $p_n=p_1^n$ where p_n is the probability of zero earning after n starts and p_1 is the probability of being not placed in one race. One can remark that p_1 can be chosen to get the best fit with the observed values. We chose to fit for the two classes with maximum and minimum number (one) by a log linear function.

Otherwise, having the number Nn of placed and Nn' of non placed horses for each number of starts n, it is also possible to simply infer for each number of starts the total number N0 of horses prepared to race:

$$Nn' = \frac{Nn'}{(Nn+Nn')} \times N0$$

$$N0 = Nn+Nn'$$

Because for zero start no horse earns any money N_0=1.00 x N0. The distribution of the number of starts therefore gives the distribution of N_0 the number of horses

prepared to race but which never started. For maximizing we chose to estimate it through the mode of the distribution of the number of starts.
N_0 is then estimated in two different ways.

Results

Heritability estimations are recapitulated in tables 1, 2, 3 and 4 for each four variables in the three kinds of races.

Table1. Estimations of the heritabilities of the Starting Status

	Model	2y.-old	3y.-old	4y.-old	5y.-old	2-5y.-old
Flat races	Sire	0.73	0.76	0;.37	0.31	0.62
	Sire	**0.68**	**0.70**	**0.34**	**0.29**	**0.57**
	Dam	0.43	0.56	0.33	0.32	0.54
Jumping races	Sire	-	0.71	0.74	0.57	0.91
	Sire	-	**0.43**	**0.70**	**0.85**	**0.85**
	Dam	-	0.38	0.40	0.51	0.51
Trotting races	Sire	0.16	0.46	0.52	0.45	0.58
	Sire	**0.05**	**0.45**	**0.50**	**0.44**	**0.56**
	Dam	0.15	0.37	0.39	0.30	0.43

Table2. Estimations of the heritabilities of the number of starts, zero starts. included

	Model	2y.-old	3y.-old	4y.-old	5y.-old	2-5y.-old
Flat races	Sire	0.22	0.41	0.22	0.13	0.39
	Sire	**0.22**	**0.41**	**0.22**	**0.13**	**0.40**
	Dam	0.52	0.64	0.44	0.39	0.70
	Animal	0.33	0.58	0.28	0.15	0.61
Jumping races	Sire	-	0.15	0.24	0.17	0.31
	Sire	-	**0.15**	**0.23**	**0.17**	**0.31**
	Dam	-	0.45	0.40	0.38	0.59
	Animal	-	0.23	0.30	0.24	0.49
Trotting races	Sire	0.02	0.13	0.16	0.13	0.18
	Sire	**0.02**	**0.13**	**0.17**	**0.13**	**0.18**
	Dam	0.21	0.37	0.33	0.23	0.35
	Animal	0.02	0.17	0.21	0.14	0.23

Table 3: Estimations of the heritabilities of the number of starts, zero starts excluded

	Model	2y.-old	3y.-old	4y.-old	5y.-old	2-5y.-old
Flat races	Sire	0.24	0.32	0.26	0.14	0.37
	Sire	**0.25**	**0.33**	**0.27**	**0.15**	**0.38**
	Dam	0.46	0.41	0.52	0.37	0.57
	Animal	0.24	0.39	0.34	0.20	0.49
Jumping races	Sire	-	0.04	0.07	0.09	0.15
	Sire	-	**0.03**	**0.08**	**0.10**	**0.16**
	Dam	-	0.42	0.18	0.35	0.44
	Animal	-	0.09	0.08	0.10	0.20
Trotting races	Sire	0.01	0.08	0.04	0.03	0.06
	Sire	**0.00**	**0.09**	**0.04**	**0.03**	**0.06**
	Dam	1.03	0.52	0.31	0.32	0.29
	Animal	0.05	0.09	0.04	0.03	0.06

Table4. Estimations of the heritabilities of the number of starts for earning horses

	Model	2y.-old	3y.-old	4y.-old	5y.-old	2-5y.-old
Flat races	Sire	0.33	0.33	0.19	0.07	0.30
	Sire	**0.36**	**0.34**	**0.19**	**0.07**	**0.31**
	Dam	0.50	0.43	0.47	0.34	0.49
	Animal	0.35	0.43	0.32	0.12	0.35
Jumping races	Sire	-	0.05	0.10	0.04	0.12
	Sire	-	**0.05**	**0.10**	**0.05**	**0.12**
	Dam	-	0.65	0.21	0.51	0.44
	Animal	-	0.11	0.11	0.03	0.13
Trotting races	Sire	0.00	0.12	0.06	0.04	0.07
	Sire	**0.00**	**0.13**	**0.07**	**0.05**	**0.07**
	Dam	0.81	0.68	0.54	0.39	0.35
	Animal	0.09	0.14	0.06	0.05	0.06

From these tables one can observe the very high estimations, maybe over estimations, of the heritabilities of the starting status: 0.54< h^2 <0.73 in flat races, 0.51< h^2 <0.91 in jumping races and 0.37< h^2 <0.52 in trotting races when the disputable 2 year-old's results are excluded.

The same estimations for the number of starts when zeros are excluded are average (0.14< h^2 <0.49) for flat races, but low in the other cases: 0.03< h^2 <0.20 in jumping races and 0.00< h^2 <0.09 in trotting races.

These values increase slightly for the number of starts when zeros are included because this criterion is a mixing with the very "heritable" starting status. They do not change very much when considering only earning horses. All these estimations do not consider the maternal components because for all the criteria based on the number of starts a maternal effect clearly appears. This is not the case for the starting status, the Dam component of the variance being in that case not as high as the Sire component. A differential exploitation in the number of starts of the offspring between mares clearly appears

One can also note that the maximum of the heritability estimations are found for 3 year-olds in flat races, 4 year-olds in jumping races and 3-4 year-olds in trotting races. We would also recommend the Sire component of the variance in the Sire-Dam model (in bold in the tables) multiplied by 4, to give the best estimations of the heritabilities according to the age of performance and art of races.

The analysis of the statistical significance of the fixed effects show that sex is significant in most cases as is the year with the exception of jumping races. The region and the breeder's category are currently significant for starting status but this effect mostly disappears for the number of starts, zeros starts excluded, or when restricted to earning horses. We will therefore limit the presentation of the results to the effect of sex for racing status and number of starts, zero starts excluded (Table5) and to the effect of the region and breeder's category on the starting status (Table6). The effect of year depending on the policy of the racing organisations is not interesting here.

From table 5 we can conclude that females are more precocious than males in flat races. They have a greater chance of being a starter at 2years of age (advantage of 1/10 of the standard deviation of the underlying probit function). This advantage decreases slowly, and at 5 years of age the situation is reversed (disadvantage of 2/10 of the upper cited standard deviation). Finally it results that over the 2-5 years career, males and females have an equal chance to appear in races. However, females have in mean, a half of a start less than males because their higher number of starts with 2 and 3 years of age is not sufficient to counterbalance their disadvantage with 4 and 5 years of age.

In jumping races females have a lower probability to appear in races than males and with age this decreases even more. Their number of starts, similar to those of 3year old males, is lesser after, leading to a mean difference of near 0.7 starts disadvantage over the 3-5year-olds career.

In Trotting races, females have a lesser probability than males to appear in races whatever their age. When they appear, their number of starts is similar to those of 2 and 3 year old males but lower after leading over the 2-5year-old career to a 2.1 starts handicap.

Table5: Estimated fixed effects of sex

Racing art	Level	2y.-old	3y.-old	4y.-old	5y.-old	2-5y.-old
Starting Status						
Flat races	Males Females	0.00 0.11**	0.00 0.04*	0.00 -0.02.	0.00 -0.24**	0.00 0.01
Jumping races	Males Females		0.00 -0.07**	0.00 -0.20**	0.00 -0.29**	0.00 -0.22**
Trotting races	Males Females	0.00 -0.02	0.00 -0.09**	0.00 -0.12**	0.00 -0.22**	0.00 -0.11**
Number of starts (zeros excluded)						
Flat races	Males Females	0.00 0.36**	0.00 0.27**	0.00 -0.30**	0.00 -0.37*	0.00 -0.52*
Jumping Races	Males Females		0.00 0.03	0.00 -0.19*	0.00 -0.17	0.00 -0.66**
Trotting races	Males Females	0.00 0.07	0.00 0.06	0.00 -0.21**	0.00 -0.69**	0.00 -2.10**

Table 6: Effects of the region and breeder's category on the starting status

Factor	Level	2y.-old	3y.-old	4y.-old	5y.-old	2-5y.-old
Flat races						
Region	(a) (b) (c)	**0.00** **-0.21** **0.03**	**0.00** ***0.00*** **-0.13*****-0.58*** **-0.12*****-0.07***	0.00 -0.03 -0.16	0.00 -0.02 -0.14	0.00 -0.06 -0.13
Breeder's category	[1-2] [3-5] [6-8 [9+]	**0.00** **0.19** **0.24** **0.41**	**0.00** ***0.00*** **0.20** ***0.16*** **0.31** ***0.26*** **0.45** ***0.55***	**0.00** ***0.00*** **0.15** ***0.33*** **0.23** ***0.44*** **0.27** ***0.84***	**0.00** *0.00* **0.14** *-0.11* **0.14** *-0.03* **0.18** *0.55*	**0.00** ***0.00*** **0.19** ***0.46*** **0.30** ***0.81*** **0.43** ***1.48***
Jumping races						
Region	(a) (b) (c)		0.00 0.06 0.01	0.00 0.18 0.05	**0.00** **0.22** **0.05**	0.00 0.19 0.04
Breeder's category	[1-2] [3-5] [6-8 [9+]		**0.00** **0.12** **0.19** **0.22**	**0.00** **0.10** **0.13** **0.19**	0.00 0.08 0.10 0.08	**0.00** **0.13** **0.20** **0.23**
Trotting races						
Region	(a) (b) (c)	0.00 -0.14 -0.16	**0.00** **-0.08** **-0.14**	**0.00** **-0.10** **-0.16**	**0.00** **-0.09** **-0.13**	**0.00** **-0.10** **-0.16**
Breeder's category	[1-2] [3-5] [6-8 [9+]	0.00 0.06 -0.00 0.11	**0.00** **0.09** **0.10** **0.18**	**0.00** **0.06** **0.08** **0.16**	**0.00** **0.06** **0.07** **0.15**	**0.00** **0.07** **0.08** **0.16**

Legend of table 6:
(a)Normandy; (b) Greater west + north + Ile de France + centre; (c) southwest + southeast + east and greater Lyon
***Bold**: significant at the 5% level; Italics: corresponding effect on number of starts (zeros excluded)*

Table 6 shows that the effects of the breeding region, as those of the breeder's category are mostly not significant on the number of starts (zeros excluded or limited to the earning horses). Only slight significant effects were found for flat races where an increase in number of starts was observed with the increasing number of foals produced by the breeder over the period (in italics on table 6). Most of the significant effects were observed on the starting status. The Normandy breeding region offers a greater chance to their foals to become starters in flat and trotting races. This is not the case for jumping races where there is a slight tendency in favour of region b (great west etc…) expresses the fact that specialised breeding for jumping races is not located in Normandy which is more specialised in flat races. In the three racing arts, the chance of a foal to become a starter increases with the number of foals produced by the breeder over the period, or with his/her professionalism.

Figures 1, 2 and 3 give the distributions of the number of starts for horses starting and for horses not earning anything from 2 to 5 years of age.

FIG 1: Distribution of starts in flat races from 2 to 5 years of age

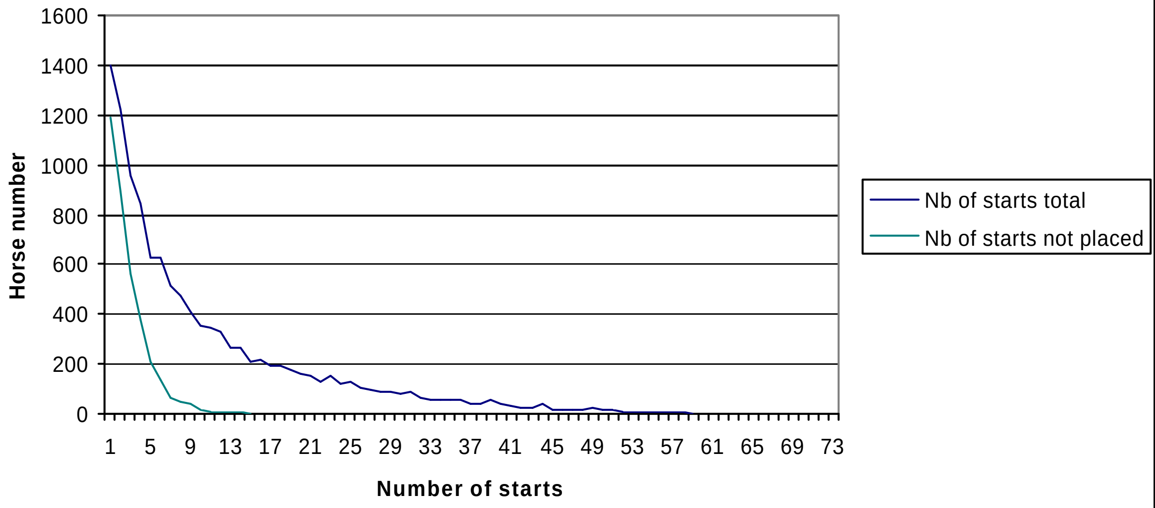


FIG 2: Distribution of starts in jumping races from 3 to 5 years of age

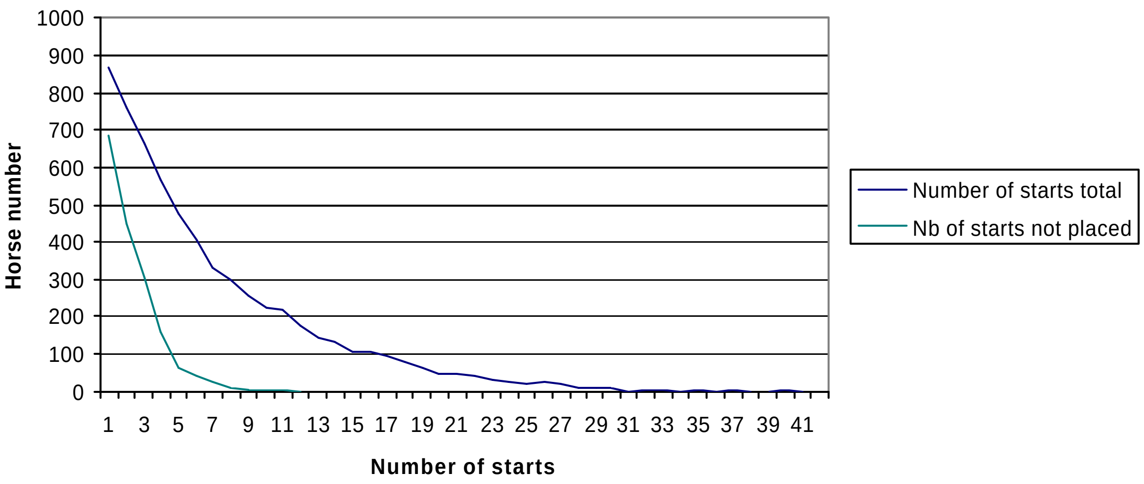
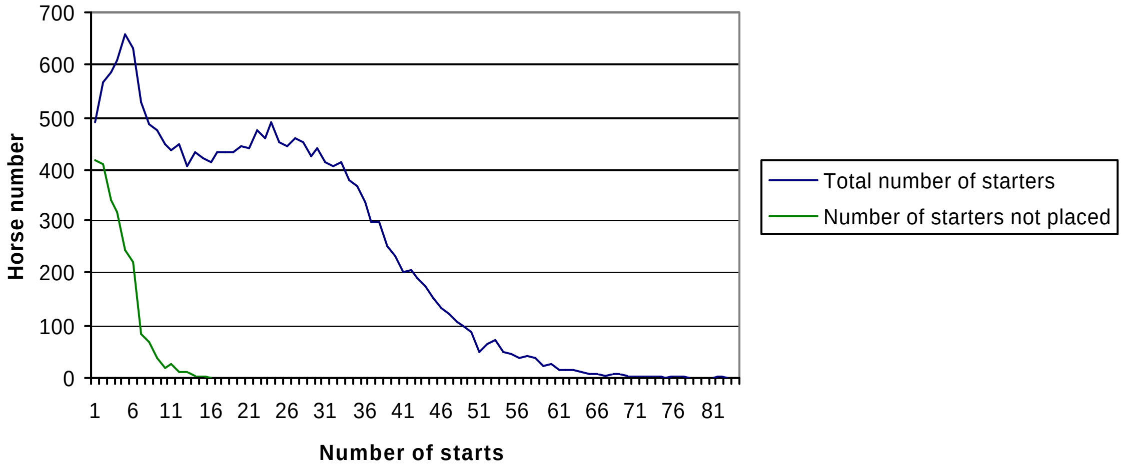


FIG 3: Distribution of the number of starts in trotting races from 2 to 5 years of age



The same kind of distribution is found for every age and allows the estimation of N_0 (mode of the distribution for starters, 1401, 866, 660, here for flat jumping and trotting races resp.) and N_0' (exponential model, 1807, 1133, 610 here resp.) estimating the number of horses prepared to race but which never started. Table 7 summarises the numbers obtained which are converted in Table 8 for the estimation of different selection rates.

Table 7: Observed and calculated numbers.

N_0: number of prepared horses which never started (mode of N_s). N_0': idem (Exponential adjustment). N_s: number of starters. N_e: number of horses earning money

Age	N_0	N_0'	N_s	N_e	N_0+N_s	$N_0'+N_s$
			Flat races			
2y. old	1013	1711	3542	2375	4555	5253
3y. old	1503	2493	9753	6680	11256	12246
4y. old	1502	2070	7196	4425	8698	9266
5y. old	749	1030	3798	2483	4547	4828
2-5y. old	1401	1807	11821	8250	13222	13628
			Jumping races			
3y. old	821	1306	2521	1674	3342	3827
4y. old	891	1264	4698	3249	5589	5962
5y. old	656	816	3319	2303	3975	4135
3-5y.old	866	1133	6304	4545	7170	7437
			Trotting races			
2y. old	486	332	1159	838	1645	1491
3y. old	1636	1297	15063	12686	16699	16360
4y. old	1198	1240	18211	15306	19409	19451
5y. old	1011	1463	14069	11420	15080	15532
2-5 y. old	660	610	19782	17573	20441	20390

Table 8: corresponding selection rates (%)

Age	N_e/born horses	$N_e/(N_0'+N_s)$	$N_e/(N_0+N_s)$	N_e/N_s
Flat races				
2y. old	12	45	52	67
3y. old	33	55	59	68
4y. old	22	48	51	61
5y. old	12	51	55	65
2-5 y. old	**41**	**61**	**62**	**70**
Jumping races				
3y. old	8	44	50	66
4y. old	16	54	58	69
5y. old	12	56	58	69
3-5y.old	**23**	**61**	**63**	**72**
Trotting races				
2y. old	1	56	51	72
3y. old	22	78	76	84
4y. old	26	79	79	84
5y. old	19	74	76	81
2-5 y. old	**30**	**86**	**86**	**89**

The estimated amount of horses prepared to race that did not race is important. It is mostly greater than half of the number of horses starting but which did not earn anything in the two hypotheses tested. One can also notice that the two hypotheses led to relatively similar estimations of these numbers. These estimations did not impact very much on the estimated selection rates of the earning horses shown in table 8 also due to the very great percentage of earning horses (respectively 70, 72 and 89%of the starters in flat, jumping and trotting races). When considering the 2 to 5 year old career 41, 23 and 30% of the horses born earn some money and they represent in fact a selection rate of 61-62, 61-63 and 86%. If one would consider the starting horses as the only selection basis, these rates would increase to 70, 72 and 89%.
The rates calculated by Langlois and Blouin (2004) by comparison of BLUP breeding value estimations of earning and non earning horses were 90% for 35% of born thoroughbreds in flat races, 86% for 16% in jumping races and 85% for 32% in trotting races. With the present estimation of N_0 and N_0', we are therefore in good agreement with previous results on trotters. However previous results on thoroughbreds may have underestimated the selection bias in flat and jumping races.

Discussion

The heritabilities obtained here for starting status are very high, up to 0.70 in flat races, to 0.85 in jumping races and to 0.56 in trotting races. This latter result recently presented and discussed by Rose et al. (2007) is very similar to those found by Langlois and Vrijenhoek (2004) by the mother-offspring regression. It is also quite high: $0.30<h^2<0.56$ here and $0.32<h^2<0.43$ estimated before and 0.65 for the qualification. To be a starter appears therefore highly sire and mare dependant in the three racing arts studied but particularly in jumping and flat races. Some pre-selection according to the pedigree can therefore not be avoided even if the level of this pedigree does not necessarily play a role. This means that some pedigrees have a better chance of being starters than others. It could be the effect of fashion or of the market. Some breeders and trainers aim to test horses produced by some studs (or mares) and some stallions lead to the discard of the others. Indeed, two phenomena should be distinguished:

-One is to be prepared to participate in races or not be prepared. It seems highly heritable. However, it could be sire and mare dependent but not really heritable leading to a kind of artefact.

-The second phenomenon is when prepared to race, to take part in a different number of races. Except for flat races, it seems highly environmental or trainer dependant. Indeed, the heritabilities of the number of starts (zeros excluded or for earning horses) are moderate in flat races and very low in jumping and trotting races Therefore, as recommended before for trotters (Langlois and Vrijenhoek, 2004), we confirm that this criteria should be used as a correction factor (co variable) for annual or career traits but not as a selection trait. Similar results were found by Arnason (1999) and Svobodova et al. (2005).

Concerning the fixed effects, the results are self explanatory for sex (Table 5) and there is no need for adding much comment. In general females have a slight lesser chance of being starters and when they start they have a lesser number of starts. However, in the very early career, particularly in flat races at 2 and 3 years of age females appear more precocious than males.

The preliminary study by Rose et al. 2007 used six regions of birth. We considered here that three were sufficient. Normandy clearly appears as the leading region in terms of starting status for flat and trotting races, followed by other regions north of the Loire (great west, north and centre) which are leading in jumping races. The rest of France (east, southeast, southwest) give lesser chance of being starters in the three types of races. These regions of birth effects are mostly not significant for the number of starts, zero starts excluded, except for the 3y.-olds in flat races (Table 6).

The effect of breeder's category is mostly significant on the variable starting status. A clear advantage can be observed with the increasing number of horses produced which is an indicator of a higher professional level of the breeder. This effect is still observed only for flat races for the number of starts, excluding zero starts. This can be connected with the fact that in flat races, number of starts appears to be moderately heritable. This is not the case for jumping and trotting races. In flat races the number of starts remains sire dependant as is the starting status. It may explain that the effect of breeder's category on starting status and on the number of starts is more closely linked. This is not observed for jumping nor trotting races because the dependence of the number of starts toward stallions is too loose in those cases.

The estimation of the number of horses prepared for racing but which never raced by the mode of the distribution of the number of starts N_0 or by an exponential adjustment of the distribution of the number of starts of non placed horses N_0' are approximate. They suppose that the total potential of starters is fully expressed for the maximum of the distribution of starters in the first case (N_0), or in the second case (N_0') that the distribution law proposed to model the distribution of the number of starts for non placed horses fit reality well (this is hardly the case). However, the selection rates obtained (table8) are in relatively good agreement.

It can be deduced from this study that horses earning some money between 2 and 5 years of age are representative of respectively 61-62; 61-63; 86 percent of the best horses born for thoroughbreds in flat and jumping races and for French Trotters. The corresponding observed proportions are far below respectively 41; 23 and 30 per cent of the horses born. From the comparisons of Blup estimations of horses earning or not some money, Langlois and Blouin (2004) previously obtained the corresponding selection rate for placed horses: 90; 86 and 85 for respectively 35; 16 and 32 per cent of the horses born.

We are therefore in very good agreement for trotters where estimations of breeding values are not biased by selecting only earning horses. This is probably due to the fact that 89% of starting horses are earning some money.

On the contrary, in the preceding study we probably underestimated the bias in the case of Thoroughbreds. According to the present study, horses having earnings are evaluated in the best 61-63% born horses and they represent only 70-72% of the starting horses. In this case, horses with earnings appear therefore selected among the best and estimation of breeding values may under-evaluate real genetic differences between horses. This may have led in the preceding study to considerably underestimate the real bias.

For trotters, 86 % as selection rate will hardly bias breeding value estimations made on earnings. An 89 % as selection rate when considering the selection basis formed by the starting horses will also not introduce much bias.

Conclusion

Two different conclusions should be made at the end of this study, one for the French trotter and one for the Thoroughbred:

Because for trotters the selection bias when taking into account only placed horses is seemingly not so big as it was thought a priori, the recommendations for breeding value estimation made before (Langlois and Vrijenhoek, 2004) can be attenuated.

It was recommended to use two principal traits, the qualification and the career earnings, adding optionally the best time for facilitating the comparison between countries. Number of starts was proposed as an important correction factor for earnings (and best time) but not as a selection criterion. Here we confirm these propositions, replacing the qualification by being a starter or not which is a heritable criterion. Further studies would lead to the estimation of the genetic parameters of the variable earning money or not which is probably of the same nature as being a starter, but does not need any additional information as the number of starts here and being qualified before, to be treated. However, the correction of the selection bias expected from the introduction of such none or all variables in the estimation of breeding value of trotters is probably not as important as it was thought primarily.

For thoroughbreds it was estimated that earning horses belong to the best 60-65% of the horses born. The bias is then much greater even if not so great when inferred directly from the proportions of horses born earning money (respectively 41 and 23 per cent in flat and jumping races). In this case, the lower percentage of horses earning money among the starters (70-72 %) indicate that using performance criteria like ranks in races which can take into account horses without earnings is recommended because it will reduce the bias in estimations however not completely. In this case to use a censored variable conditioning the observation of the performance in a tobit model (Tobin, 1958) should be a solution to investigate. The number of starts could act as the censored variable but particular attention should be given to the threshold (best 60-65% not best 41 or 23%).

Some further thought is also needed about the meaning of the high heritability of the variable being a starter or not. What is the biological or animal husbandry background of this observation? A clear answer to this question is still failing. To enlighten this question the estimation of the genetic correlation of this variable with earnings and best time should be done properly. It is important because supposing that being a starter only depends on the horse quality, one could propose to run breeding value estimations only on this criteria. It is indeed available for every horse born, is very heritable and would therefore lead to unbiased estimations with very high accuracy. But this would be the estimation of the horse quality only in the case of a very high genetic correlation with performance measurements.

References

Arnason, Th. 1999. Genetic evaluation of Swedish standard-bred trotters for racing performance traits and racing status. Journal of Animal Breeding and Genetics 116, 387-398.

Gilmour A.R., Gogel B.J., Cullis B.R., Wehlam S.J., Thompson R., 2002. ASReml User guide release 1.0 VSN International Ltd, Hemel Hemstead, Hp1 1ES, UK

Langlois B., Blouin C., 2004. Practical efficiency of breeding value estimations based on annual earnings of horses for jumping, trotting and galloping races in France. Livest. Prod. Sci. 87, 99-107.

Langlois B., Vrijenhoek T. 2004. Qualification status and estimation of breeding value in French Trotters. Livest. Prod. Sci. 89, 187-194.

Langlois B., Blouin C. 2007. Annual, career or single race records for breeding value estimation in race horse. Livest. Sci. 106, 132-141.

Rose A., Blouin C., Langlois B. 2007. How number of starts inform about the selection bias when using performers for breeding evaluation: the example of French trotters. 58th Annual Meeting of the eaap, Session 17, 13p. Dublin, Ireland, August 26th-29th 2007.

Svobodova S., Blouin C., Langlois B. 2005. Estimation of genetic parameters of thoroughbred racing performance in the Czech Republic. Anim. Res., 54, 499-509.

Tobin J. 1958. Estimation of relationships for limited dependant variables. Econometrica, 26, 24-36.