January 1979

Fluctuations in Sire Evaluations

D. C. Bolgiano
Cornell University

L. Dale Van Vleck
University of Nebraska-Lincoln, dvan-vleck1@unl.edu

R. W. Everett
Cornell University

Fluctuations in Sire Evaluations

D. C. BOLGIANO, L. D. VAN VLECK,
and R. W. EVERETT
Department of Animal Science
Cornell University
Ithaca, NY 14853

ABSTRACT

Large fluctuations in an artificial insemination sire's evaluations concern both dairymen and artificial insemination studs. Characteristics of three bulls with high initial values which dropped on later evaluations were compared with those of 17 "normal" bulls. Characteristics included genetic yield, management, genetic trend, environmental trend, and size of herds where daughters freshened, genetic merit of mates, and average age at freshening of daughters in first lactation. Evidence suggests that fluctuations are from preferential treatment of daughters of some sires as well as randomness inherent in methods of sampling and evaluating sires.

INTRODUCTION

Over the last few decades accurate evaluation of dairy sires has been emphasized. Development of mixed model techniques for handling unbalanced data (1, 2, 3) led to the Northeast Artificial Insemination Sire Comparison (NEAISC) (4, 5, 6) used in New England and New York State. Recently, for several bulls from a stud in this region superior first evaluations dropped sharply on subsequent evaluations. Such changes in sire evaluations may be due to the random nature of sire sampling procedures or unaccounted sources of bias in evaluation procedures. In either case, such phenomena cause dairymen to have less confidence in methods of evaluation. This study attempted to determine what caused the phenomena.

METHODS

Methods to purchase and sample unproven sires were examined first. Bulls in the stud with high NEAISC ratings are mated to cows with high estimated transmitting abilities (ETA's). Male calves which result from these planned matings are brought into the stud, and at about a year of age, semen collection begins. An initial sample of 800 breeding units for each young sire is distributed evenly among the districts. Each AI (artificial insemination) technician usually receives six breeding units from a sire to use at his discretion in herds of consenting dairymen. Semen from these unproven sires is least expensive.

These sires are evaluated when complete records of milk in first lactations are available on at least 20 of their daughters. The first evaluation is based on complete records and other records in progress. A decision whether to keep a bull is subject to his first evaluation.

Those bulls returned to service will produce a second crop of daughters and be reevaluated. The second-crop evaluation is based on all new records of first lactations and those that went into the first evaluation. Usually the number of second-crop daughters is much greater than the number in the first crop (Table 1), so the second evaluation is weighted heavily by new records.

When NEAISC evaluations are computed, a confidence range (CR) is calculated for each bull. This figure is an approximate 68% confidence interval for the transmitting ability of the bull where daughter milk yields and sire transmitting abilities have multivariate normal distribution. In other words, the interval of NEAISC evaluation ± CR should contain the transmitting ability 68% of the time. A bull's transmitting ability is never known, but successive unbiased predictions with increasing amounts of information should approach it. Therefore, about 68% of later predictions would be expected to lie in the confidence range of the first evaluation. As in Table 1, most do lie in this interval.

In the NEAISC procedure the model for a cow's record of milk is

$$y_{ijkl} = h_i + g_j + s_{jk} + e_{ijkl}$$

Received August 14, 1978.

TABLE 1. NEAISC sire evaluations (kg/milk).

	1st proof			2nd proof		
Sire reg. no.	Evaluation	No. of daughters	Confidence range (kg)	Evaluation	No. of daughters	Evaluation change (2nd−1st)
1466325	149	40	178	184	1308	35
1466511	413	43	172	345	4268	-68
1471074	121	26	200	265	277	144
1471438	60	39	178	15	198	-45
1473693	303	40	178	166	184	-137
1476679[a]	518	78	195	367	4331	-151
1492023	534	21	213	417	2387	-117
1492322	455	19	218	343	761	-112
1495709	303	33	186	376	929	73
1498036	153	23	204	35	378	-118
1499133	564	27	200	742	971	178
1507569	323	40	178	214	1151	-109
1511167	451	30	190	504	1198	53
1516360	203	37	181	620	751	417
1517540	226	26	200	5	419	-221
1520161[a]	485	41	172	305	1697	-180
1520162	425	26	200	487	436	62
1526102	275	30	190	283	868	8
1536450	612	52	159	752	310	140
1537060[a]	468	39	178	53	558	-415

[a]Problem bull.

where h is the herd-year-season of freshening; g is the sire group, grouping by stud and year entered service; s is the particular sire within group; and e is the random error term associated with the record. The herd and group effects are fixed while sires are random. In matrix notation, the mixed model is

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where y is the vector of records of first lactation of the daughters of all bulls to be evaluated, X and Z are known design matrices, β is a vector of unknown fixed herd and group effects, u is a vector of unknown random sire effects, and e is a vector of random error terms. The model includes $E(\mathbf{y}) = \mathbf{X}\beta$, $E(\mathbf{u}) = 0$, $E(\mathbf{e}) = 0$, $Var(\mathbf{u}) = \mathbf{G} = \mathbf{A}(h^2/4)\sigma^2$, $Var(\mathbf{e}) = \mathbf{R}\sigma^2$, and $Cov(\mathbf{u},\mathbf{e}) = 0$, where A is the matrix of additive relationships among sires, σ^2 is the phenotypic variance of milk, and h^2 is heritability of milk yield. Estimates of fixed and random effects are obtained by solving the mixed model equations

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{Y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Y} \end{bmatrix}$$

This brief examination of the sire sampling procedure and evaluation method suggests a few possible sources of bias. The average merit of herds associated with sires may vary, but by computing as if herds are fixed, this bias is removed (7). A cow's record is classified by herd, year, and season of freshening (herd-year-season). Presently, the year is divided into two seasons, December through April and May through November. The distributions of freshening dates of first lactations of daughters were examined to see if any bulls had a large percent of daughters freshening over a short interval in a particular season which could bias comparisons if seasonal adjustments were incorrect. The distribution of freshenings was not disproportionate (11).

The model includes an effect for each herd-year-season which should account for environmental differences between cows freshening in different herds or in the same herds but in different year-seasons. The model does not account for different treatment of cows within the same herd-year-season. If some cows in a herd are treated preferentially, their records will reflect that treatment, and their sires' evaluations may be biased.

Another potential problem concerns the genetic merit of mates of unproven sires. The average merit of mates of different bulls needs to be equal. If some bulls were mated to superior cows while others had genetically inferior mates, then daughter records would not reflect clearly differences between their sires' genetic values.

Records in sire evaluation are corrected for fixed effects including number of times milked per day and age and month at freshening. Incorrect adjustment factors might cause a bull to be evaluated unfairly; e.g. if the average age at freshening of first-crop daughters of two bulls differed significantly and incorrect factors were used to adjust the records, the resulting evaluations would be inaccurate.

Most of the suggested explanations for changes in sire evaluations can be checked by analyzing production records, but the problem of detecting preferential treatment within a herd is more difficult. Our approach was to compare characteristics of herds in which daughters freshened. If daughters of a particular sire were given special treatment in many different herds, the herds might have common features. Information was compiled on all herds where 20 sires had their first- and second-crop daughters. Data were collected for each of three "problem" bulls and five to seven other bulls

TABLE 2. Average sizes of daughters' herds.

Bull	1st crop	2nd crop	Change (2nd–1st)
1466325	109.70	93.14	-16.56
1466511	118.00	96.27	-21.73
1471074	100.60	89.20	-11.40
1471438	84.80	80.29	-4.51
1473693	99.72	86.50	-13.22
1476679[a]	88.07	102.50	14.43
1492023	94.27	96.07	1.80
1492322	107.00	92.20	-14.80
1495709	95.18	92.54	-2.64
1498036	81.32	95.06	13.74
1499133	95.23	104.50	9.27
1507569	100.70	85.96	-14.74
1511167	105.10	104.10	-1.00
1516360	94.98	96.92	1.94
1517540	75.74	101.90	26.16
1520161[a]	107.40	93.79	-13.61
1520162	98.94	89.27	-9.67
1526102	97.69	91.33	-6.36
1536450	84.84	90.30	5.46
1537060[a]	93.83	104.60	10.77
Average			
All bulls	96.66	94.32	-2.34 (2.83)[b]
Problem bulls ($\bar{x}$)	96.43	100.30	3.87
Others ($\bar{y}$)	96.69	93.27	-3.42
$S_{\bar{x}-\bar{y}}$	6.71		

[a]Problem bull.

[b]Standard error of mean.

which entered service during the same year. Characteristics compared were herd size, genetic merit of the herd, management, genetic trend, and environmental trend. The averages of each of the first three characteristics were calculated for every bull over all herds where these 20 sires had their first- and second-crop daughters. Genetic and environmental trends were calculated by regressing the differences between herd average as computed by the Northeast cow ETA (NEA^{-1} ETA) procedure and breed average on years 1964 through 1977 (coded 1 to 14). For each herd, genetic and environmental trends were coded as 1 for a trend greater than or equal to zero and 0 for a trend less than zero. The percent of daughters in each crop appearing in herds with trends greater than or equal to zero was tabulated for each sire.

Differences between first-crop means and second-crop means of herd size, genetic merit and management of herds, dam ETA's, and average age at freshening were tested for significance by a "t" statistic for each characteristic. The same procedure was used to test for differences between first-crop means of "problem" bulls and "normal" bulls. In this case a pooled estimate of variance for all first-crop means of each characteristic was computed for each test as if the variances were common. Genetic and environmental trend data were tested for equality of medians between the "problem" and "normal" groups for both first and second crops.

RESULTS AND DISCUSSION

As a cow seems more likely to receive preferential treatment in a small herd than in a

TABLE 3. Average genetic merits of daughters' herds (kg/milk).

Bull	1st crop	2nd crop	Change (2nd–1st)
1466325	134.2	96.5	-37.7
1466511	68.8	118.1	49.3
1471074	161.7	86.5	-75.2
1471438	114.5	62.4	-52.1
1473693	133.3	98.4	-34.9
1476679[a]	69.7	81.6	11.9
1492023	212.5	135.5	-77.0
1492322	133.7	119.9	-13.8
1495709	106.1	141.0	34.9
1498036	116.2	91.9	-24.3
1499133	52.8	159.0	106.2
1507569	52.8	108.4	55.6
1511167	111.4	138.4	27.0
1513630	92.6	119.2	26.6
1517540	9.7	138.0	128.3
1520161[a]	-36.4	109.4	145.8
1520162	74.6	167.8	93.2
1526102	84.9	137.0	52.1
1536450	-65.2	191.6	256.8
1537060[a]	108.4	80.9	-27.5
Average			
All bulls	86.8	119.1	32.3 (18.75)[b]
Problem bulls ($\bar{x}$)	47.2	90.6	43.4
Others ($\bar{y}$)	93.8	124.1	30.3
$S_{\bar{x}-\bar{y}}$	39.9		

[a]Problem bull.

[b]Standard error of mean.

large one, herd size was examined for differences between the first- and second-crop averages of all bulls (Table 2). The first-crop average for "problem" bulls also was compared with the first-crop average of other bulls. No differences were significant ($P>.05$).

Genetic and management levels of herds were examined for similarities among problem bulls. Average genetic merit of herds where problem bulls had their first-crop daughters did not vary significantly ($P>.05$) from those where other bulls were used (Table 3). In general, second-crop daughters appeared in herds with higher genetic averages than the first crop as would be expected with upward genetic trend in the population (9, 10).

The first-crop daughters of problem bulls were in better managed herds ($P<.05$) than the other bulls (Table 4), but since herd-year-season effects are considered fixed in the model, this should not bias evaluations. Superior management in these herds may indicate a greater likelihood of preferential treatment, but this is only speculative.

Most of the daughters used to evaluate these bulls were in herds with positive genetic trend (Table 5). This was true for both first- and second-crop daughters of problem and normal bulls. In contrast, only about half of the daughters were in herds with positive environmental trend (Table 6). On a median test (8), we concluded ($P>.05$) that the distribution of percent of cows in herds with positive genetic and environmental trends did not differ significantly among problem bulls and others in both first and second crops.

TABLE 4. Average management of daughters' herds[a] (kg/milk).

Bull	1st crop	2nd crop	Change (2nd−1st)
1466325	6622.4	6690.5	68.1
1466511	6763.1	6826.6	63.5
1471074	6726.8	6722.2	-4.6
1471438	6908.2	6568.0	-340.2
1473693	6844.7	6749.4	-95.2
1476679[b]	6930.9	6989.9	59.0
1492023	6808.4	6817.5	9.1
1492322	6649.7	6699.6	49.9
1495709	6695.0	6822.0	127.0
1498036	6767.6	6690.5	-77.1
1499133	6604.3	6812.0	208.6
1507569	6912.8	6654.2	-258.5
1511167	6540.8	6785.7	244.9
1513630	6930.9	6881.0	-49.9
1517540	6699.6	6704.1	4.5
1520161[b]	7003.5	7080.6	77.1
1520162	6908.2	6822.0	-86.1
1526102	6681.4	6840.2	158.8
1536450	6722.2	6826.6	104.3
1537060[b]	6899.1	7144.1	244.9
Average			
All bulls	6781.0	6806.3	25.3 (33.7)[c]
Problem bulls ($\bar{x}$)	6944.5	7071.5	127.0
Others ($\bar{y}$)	6752.1	6759.5	7.4
$S_{\bar{x}-\bar{y}}$	70.9		

[a]Management level of a herd is obtained as a solution for a particular herd-year-season effect in the Northeast A^{-1} ETA procedure.

[b]Problem bull.

[c]Standard error of mean.

Average ETA's of dams of first- and second-crop daughters for each bull were calculated from available records. Not all daughters had dams with ETA's. The average ETA's of dams of second-crop daughters were substantially higher ($P<.05$) than ETA's of dams of first-crop daughters (Table 7). The average difference of 63 kg may be explained partly by positive genetic trend in the population although it seems unlikely that dairymen or technicians would mate best cows to unproved sires. The average ETA of dams of first-crop daughters of problem bulls was higher than the average for regular bulls, but the difference was not significant ($P>.05$). Nevertheless, the superiority of 53 kg suggests that problem bulls generally were mated to better cows than regular bulls and may indicate a potential for preferential treatment.

Finally, averages for age at freshening of daughters of these bulls were examined (Table 8). The average age at freshening of first-crop daughters was greater than the average of second-crop daughters. The difference of almost a month was significant ($P>.05$), but it was across all bulls so it should not bias evaluations of particular bulls. There was not a substantial difference ($P>.05$) in average age at freshening between first-crop daughters of problem bulls and regular bulls.

CONCLUSIONS

There appear to be several potential causes for fluctuations, including preferential treatment of daughters, nonrandom mating of unproven sires to cows, and inherent randomness of the sampling procedures. Some combination of these three factors probably causes evaluations to change. Preferential treatment can be identified only after the fact; therefore, it cannot be accounted for in the model. To remove its effects on evaluation, preferential treatment must be eliminated.

There has been some tendency to mate certain unproven bulls to genetically-superior

TABLE 5. Average coded genetic trends of daughters' herds.[a]

Bull	1st crop	2nd crop
1466325	1.0000	.8219
1466511	.7826	.8521
1471074	.8000	.8018
1471438	.9000	.7633
1473693	.7222	.7778
1476679[b]	.6667	.8180
1492023	.8667	.8406
1492322	.9048	.8113
1495709	.8929	.8343
1498036	.8214	.7702
1499133	.8372	.8503
1507569	.8828	.8026
1511167	.9474	.8583
1513630	.8163	.8425
1517540	.7778	.8405
1520161[b]	.7797	.8232
1520162	.9063	.8313
1526102	.8462	.8083
1536450	.6491	.8367
1537060[b]	.9310	.8034
Median	.8400	.8226

[a]Coded 1 for trend greater than or equal to zero and 0 for negative trend.

[b]Problem bull.

TABLE 6. Average coded environmental trends of daughters' herds.[a]

Bull	1st crop	2nd crop
1466325	.5909	.5719
1466511	.4783	.5695
1471074	.7333	.5680
1471438	.5500	.6213
1473693	.6667	.7014
1476679[b]	.4000	.5712
1492023	.7333	.6219
1492322	.5238	.5758
1495709	.5000	.6067
1498036	.5714	.6311
1499133	.3721	.5850
1507569	.5156	.5738
1511167	.5263	.5842
1513630	.6531	.6103
1517540	.5556	.5429
1520161[b]	.6271	.6146
1520162	.5313	.5639
1526102	.5641	.5798
1536450	.5614	.5578
1537060[b]	.5517	.6050
Median	.5535	.5820

[a]Coded 1 for trend greater than or equal to zero and 0 for negative trend.

[b]Problem bull.

TABLE 7. Average ETA's of daughters' dams (kg/milk).

Bull	1st crop	2nd crop	Change (2nd–1st)
1466325	20.1	110.8	90.4
1466511	26.3	125.4	99.1
1471074	-44.8	106.9	151.7
1471438	31.4	96.7	65.3
1473693	149.9	108.6	-27.6
1476679[a]	146.1	140.7	-5.4
1492023	99.4	131.5	32.1
1492322	85.7	119.5	33.8
1495709	67.5	164.8	97.3
1498036	9.9	123.8	113.9
1499133	42.4	118.5	76.1
1507569	47.3	110.7	63.4
1511167	114.4	119.8	5.4
1513630	42.6	131.2	88.6
1517540	112.0	143.7	31.8
1520161[a]	62.3	161.2	98.9
1520162	99.8	157.0	57.2
1526102	105.4	159.5	54.2
1536450	72.1	181.0	109.0
1537060[a]	143.0	168.1	25.1
Average			
All bulls	71.6	134.0	62.4 (10.1)[b]
Problem bulls ($\overline{x}$)	117.1	156.7	39.6
Others ($\overline{y}$)	63.6	130.0	66.4
$S_{\overline{x}-\overline{y}}$	30.2		

[a]Problem bull.

[b]Standard error of mean.

cows. This might be considered preferential treatment of the bulls. Both this problem and the problem of preferential treatment of daughters might be avoided by withholding identities of unproven sires when their semen is distributed originally. All pertinent information except pedigree could be given to the dairymen, estimated transmitting ability, type information on parents, etc. This would assure dairymen that they are using good semen but would keep their personal preferences from biasing evaluations. Such a system certainly would not be acceptable to purebred breeders but might have some appeal to commercial dairymen where selection of unproven sires is mainly by technicians.

The most probable explanation is that fluctuations are inherent in the random sampling procedures. When fluctuations of all 20 sires are examined (Table 1), not only problem bulls show large changes. Problem bulls are bulls that had high initial evaluations and dropped later. Other bulls showed changes of similar magnitude, but either had lower first evaluations and then declined or the change was upward, pointing out that the acquisition of problem status is arbitrary.

There appears to be a need for a sire sampling system that removes preferential treatment. A system such as suggested here seems to warrant consideration. It also must be stressed both to those who sell semen and to dairymen who use it that the method of sire evaluation is based on variability inherent in the population and the random nature of Mendelian inheritance. Therefore, a certain amount of variation is expected between subsequent evaluations of a sire.

TABLE 8. Average age of daughters at freshening (mo).

Bull	1st crop	2nd crop	Change (2nd–1st)
1466325	26.80	28.23	1.43
1466511	27.52	28.23	.71
1471074	32.90	27.94	-4.97
1471438	28.09	28.43	.34
1473693	27.81	27.64	-.17
1476679[a]	30.26	29.91	-.35
1492023	28.23	28.23	.00
1492322	29.19	28.19	-1.00
1495709	28.42	27.48	-.94
1498036	30.33	28.80	-1.53
1499133	27.72	28.12	.40
1507569	28.32	28.46	.14
1511167	28.60	28.41	-.19
1513630	26.97	28.09	1.12
1517540	27.54	27.34	-.20
1520161[a]	28.54	27.00	-1.54
1520162	27.18	27.04	-.14
1526102	29.34	26.96	-2.38
1536450	28.89	25.35	-3.54
1537060[a]	28.88	25.51	-3.37
Average			
All bulls	28.58	27.77	-.81 (.37)[b]
Problem bulls ($\bar{x}$)	29.23	27.47	-1.76
Others ($\bar{y}$)	28.46	27.82	-.64
$S_{\bar{x}-\bar{y}}$	.78		

[a]Problem bull.

[b]Standard error of mean.

ACKNOWLEDGMENTS

This research was supported in part by a grant from Eastern Artificial Insemination Cooperative, Inc., Ithaca, NY.

REFERENCES

1 Henderson, C. R. 1949. Estimation of changes in herd environment. J. Dairy Sci. 32:706. (Abstr.)

2 Henderson, C. R. 1950. Estimation of genetic parameters. Ann. Math. Stat. 21:309.

3 Henderson, C. R. 1963. Selection index and expected genetic advance. NAS-NRC:982.

4 Henderson, C. R. 1966. A sire evaluation method which accounts for unknown genetic and environmental trends, herd differences, season, age effects, and differential culling. Proc. Symp. on Estimating Breeding Values of Dairy Sires and Cows. Washington, DC.

5 Henderson, C. R. 1973. Sire evaluation and genetic trends. *In* Proc. Anim. Breeding and Genetics Symp. in Honor of Dr. J. L. Lush. ASAS and ADSA, Champaign, IL.

6 Henderson, C. R. 1974. General flexibility of linear model techniques for sire evaluation. J. Dairy Sci. 57:963.

7 Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423.

8 Hogg, R. V., and A. T. Craig. 1970. Introduction to mathematical statistics. 3rd ed. MacMillan Publishing Co., Inc., New York.

9 Hintz, R. L., R. W. Everett, and L. D. Van Vleck. 1977. Estimation of genetic trends from cow and sire evaluation. J. Dairy Sci. 61:607.

10 Van Vleck, L. D. 1976. Theoretical and actual genetic progress in dairy cattle. *In* Proc. Int. Conf. Quantitative Genetics, Iowa State University Press, Ames.

11 Van Vleck, L. D., C. R. Henderson, G. R. Wiggans, R. D. Anderson, J. J. Bakker, D. C. Bolgiano, R. A. Cady, T. R. Famula, and M. F. Rothschild. 1978. Seasonal patterns of adding new daughters to sire proofs. Genetics research. Page 9 *in* 1977-78 Rep. Eastern Artificial Insemination Cooperative, Inc.