July 1968

Relative Selection Efficiency in Retrospect of Selected Young Sires

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

Relative Selection Efficiency in Retrospect of Selected Young Sires

L. D. VAN VLECK
Animal Science Department, Cornell University, Ithaca, New York 14850

Abstract

The indexes used in retrospect in estimating one-half genetic value (daughter superiority) of progeny of 541 Holstein matings contracted by Eastern Artificial Insemination Cooperative were determined. A set of information on relatives to be complete would include records of paternal sisters, the dam, paternal sisters of the dam, maternal sisters, maternal grandam, and maternal sisters of the dam. The average index of daughter superiority for the 541 progeny was between 406 and 455 kg of milk, depending on which combination of relatives was used. Paternal sisters accounted for 45 to 50% of the total daughter superiority. Records of maternal relatives, particularly records of the grandam and the dam, generally received too much emphasis relative to theoretical weights. The overemphasis on the dam's records may be due to the greater opportunity for selection of dams than for selection of sires. Efficiency of the indexes actually used in retrospect was computed as the ratio of the average correlation between the retrospective index and the true daughter superiority to average correlation between the index, using theoretical weights and true daughter superiority. The most efficient index in retrospect used records of paternal sisters, the dam, and paternal sisters of the dam, but was only 84% as efficient as the theoretically best index.

Allaire and Henderson (1, 2, 3) have described the amount and efficiency of selection practiced by New York dairymen for production in several lactations and for many type traits included in the New York type appraisal program. The efficiency of selection of matings to produce young bulls for sampling by artificial insemination organizations in New York has not previously been examined.

The Eastern Artificial Insemination Cooperative (formerly New York Artificial Breeders Cooperative) completes, at the time of deciding to make a mating, the records for a six-point pedigree of the propective young calf. They have available tables prepared by Dr. C. R. Henderson in 1953 for indexing the daughter superiority of the future calf.

The purpose of this study was to determine in retrospect what weights were actually being used as compared to the theoretically best weights determined from selection index theory, assuming that selection was for milk yield alone. The relative selection efficiency of the index used in retrospect relative to the theoretically best index was then determined.

Received for publication July 23, 1968.

Methods

Information on up to six kinds of relatives of prospective young bulls was available. A complete set of information with all milk records (2×, 305-day, mature equivalent) expressed as deviations from herdmate average would be:

X_1 = the average of p_1 artificially sired paternal sisters—this average was treated as if each daughter had only one record, although this was not always true,

X_2 = the average of n_2 records of the dam,

X_3 = the average of p_3 AI or natural-service paternal sisters of the dam,

X_4 = the average of p_4 maternal sisters with n_{4j} records for the j^{th} sister,

X_5 = the average of n_5 records on the maternal grandam, and

X_6 = the average of p_6 maternal sisters of the dam with n_{6j} records for the j^{th} maternal sister of the dam.

Figure 1 shows the complete set of relatives used in the index.

If information for some relatives was missing, the available records were used in applicable combinations.

The usual equations to find the weighting factors for estimating the daughter superiority (one-half the genetic value, $\frac{1}{2} G_\alpha$) of the individual, a, resulting from the mating are in matrix notation:

$$Pb = g.$$

P is the phenotypic variance-covariance matrix of the available information; b is the vector of

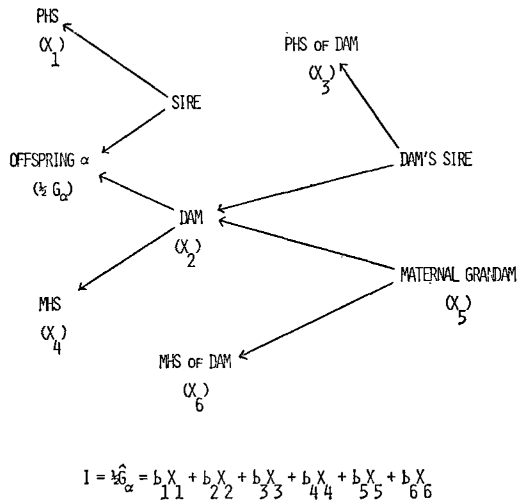

FIG. 1. Diagram of records of relatives used in estimating daughter's superiority (one-half genetic value) for offspring (α) of planned matings.

weighting factors to be used in the index, I, which estimates one-half the genetic value of the progeny, α; and g is the vector of covariances between the X_i and $\frac{1}{2}\,G_\alpha$. Note that the estimate of one-half the genetic value, $\frac{1}{2}\,\hat{G}_\alpha = I = \Sigma\; b_iX_i$ where the X_i are the averages described previously. More explicitly, when all information is available:

$$P = \sigma_p^2 \begin{bmatrix} d_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & d_2 & h/4 & h/2 & h/2 & h/4 \\ 0 & h/4 & d_3 & h/8 & 0 & 0 \\ 0 & h/2 & h/8 & d_4 & h/4 & h/8 \\ 0 & h/2 & 0 & h/4 & d_5 & h/2 \\ 0 & h/4 & 0 & h/8 & h/2 & d_6 \end{bmatrix}$$

and $g' = h\sigma_p^2$ (1/8 1/4 1/16 1/8 1/8 1/16) where σ_p^2 is the phenotypic variance, which was assumed to be the same for all kinds of relatives and h is heritability, assumed to be .25.

$d_1 = (p_1 + 15)/16\,p_1,$
$d_2 = (1 + n_2)/2\,n_2,$
$d_3 = (p_3 + 15)/16\,p_3$ if AI proof,
$d_3 = (p_3 + 7)/8\,p_3$ if NS proof,
$d_4 = 1/\Sigma[16\,n_{4j}/(7\,n_{4j} + 8)] + 1/16,$
$d_5 = (1 + n_5)/2\,n_5,$ and
$d_6 = 1/\Sigma[16\,n_{6j}/(7\,n_{6j} + 8)] + 1/16.$

These diagonal coefficients correspond to heritability = .25, repeatability = .50, and environmental correlation among natural-service daughters of .0625. These equations were set up assuming that only additive genetic effects contribute to covariance among relatives, except for the environmental covariance between natural-service paternal sisters, and that the relationships among relatives were as indicated in the matrix P.

Dickerson et al. (4) showed that the weights used in an index can be found in retrospect by substituting for g, the vector, s, of deviations of the averages of the relatives of the selected animals from the averages of all animals that could have been selected. Harvey and Bearden (5) utilized this procedure in preparing tables of correlated response. Allaire and Henderson (2) expanded and described the method of Dickerson et al. in matrix notation. Thus, the actual weights, w, which in retrospect were used in an index, can be found as:

$w = P^{-1}s$, whereas the theoretically best weights (6) from selection index theory are:

$b = P^{-1}g$. The actual index is $I' = \Sigma\, w_iX_i$ and the theoretical index is $I = \Sigma\, b_iX_i$. The index actually used as determined in retrospect will be termed the actual index.

What is necessary to compare the actual index with the theoretically best index is the relative selection efficiency (RSE) which can be defined, as usual, as RSE $= \Delta\, G'/\Delta\, G$ where $\Delta\, G'$ is the expected genetic progress per generation using the actual index and $\Delta\, G$ is the expected genetic progress per generation using the theoretical index. If the X's and G_α follow a multivariate normal distribution, then $\Delta\, I' = \sigma_{I'}D$ and $\Delta\, I = \sigma_I D$ where $\Delta\, I'$ and $\Delta\, I$ are the selection differentials of the actual and theoretical indexes, $\sigma_{I'}$ and σ_I are the standard deviations of the indexes, and D is the height of the ordinate in the normal distribution of I or I' at the truncation point for selection divided by the fraction selected. A reasonable assumption is that the intensity of selection is the same for both I and I'. Then,

$\Delta\, G' = \beta_{GI'}\,\Delta\, I' = \sigma_{GI'}\,D/\sigma_{I'}$ and
$\Delta\, G = \beta_{GI}\,\Delta\, I = \sigma_{GI}D/\sigma_I$ where

β is the regression of genetic value on the index and σ_{GI} and $\sigma_{GI'}$ are the covariances between genetic value and the indexes I and I'. Thus, RSE $= (\sigma_{GI'}\sigma_I)/(\sigma_{GI}\sigma_{I'})$ as is usually given. In matrix notation this becomes

RSE $= [(w'g)(b'Pb)^{1/2}]\;/\;[(b'g)(w'Pw)^{1/2}]$.

The b's, w's, and RSE were computed for each prospective progeny. RSE (a) was the average of the RSE for all prospective matings. RSE was also computed in two other ways:

$$\text{RSE(b)} = [(\bar{\bar{w}}'g)(\bar{\bar{b}}'P_A\bar{\bar{b}})^{1/2}]\;/\;[(\bar{\bar{b}}'g)(\bar{\bar{w}}'P_A\bar{\bar{w}})^{1/2}] \text{ and}$$

$$\text{RSE(c)} = [(\overline{w'g})(\overline{b'P_Ab})^{1/2}]\;/\;[(\overline{b'g})(\overline{w'P_Aw})^{1/2}] \text{ where}$$

P_A is the unweighted average variance-covariance matrix (actually P with average diagonal coefficients),

$\bar{w}$ is the vector of unweighted averages of the actual weights,

$\bar{b}$ is the vector of unweighted averages of the theoretical weights,

$\tilde{w} = P_A^{-1}\tilde{s}$,

$\tilde{b} = P_A^{-1}g$, and

$\bar{s}$ is the vector of unweighted averages of the selection differentials. In all cases the w were standardized back to the same relative scale as the b by multiplying w by b_1 divided by the absolute value of w_1 where b_1 is the first weight in the vector b and w_1 is the first weight in the vector w—the weights associated with X_1, the average of the paternal half-sisters. Thus, when standardized $w_1 = b_1$ or $-w_1 = b_1$. All three procedures attempt to combine situations with different amounts of information into an average approximation.

Since all relatives of all prospective progeny did not have known records, the above procedure was followed for seven combinations of relatives. These combinations and the number of each are given in Table 1, together with the average selection differentials. Note that many prospective progeny may be included in several of the combinations. The average diagonal coefficients of P for the various combinations of relatives are shown in Table 2.

Results and Discussion

The average theoretical index for the progeny was between 406 and 455 kg of milk for pedigree information, including the dam and her relatives as well as the proof of the sire. The average contribution to the expected daughter superiority of the progeny by the proof of the sire was 45 to 50% of the total.

The average relative selection efficiency by Method (a), the average correlation between the theoretical index value and true daughter superiority, and the average actual and theoretical weights are given in Table 3. As expected, the correlation, r_{IG}, between true daughter superiority and the theoretical index of daughter superiority declined as fewer kinds of relatives were used in the index. The reduction in the average theoretical correlation was relatively

TABLE 1. Combinations of relatives and number of each used in calculating relative selection efficiencies.

Paternal sisters	Dam	Paternal sisters of dam	Maternal sisters	Maternal grandam	Maternal sisters of dam	Average index	No. of matings
	Average selection differential (kg)						
439	1,869	343	829	714	494	447	131
440	1,901	386	768	699		449	189
437	1,923	419	792			445	211
447	2,100	411				442	539
447	2,100					406	539
446						209	541
445	2,073	395		864		445	480

TABLE 2. Average diagonal coefficients for finding actual and theoretical selection index weights.

Relatives in index	Paternal sisters	Dam	Paternal sisters of dam	Maternal sisters	Maternal grandam	Maternal sisters of dam
6	.0678	.6090	.0738	.6200	.6508	.468
5	.0674	.6133	.0721	.6305	.6683	
4	.0674	.6160	.0733	.6664		
3	.0671	.6781	.0687			
2	.0671	.6781				
1	.0671					
4	.0671	.6792	.0580		.6357	

small if the dam's records were included—.573 with sire proof and dam records, as compared to .614 with records of all six relatives.

The average relative selection efficiency was greatest when fewer relatives were included. Relative selection efficiency, however, compares efficiency only where records of the same relatives are included in both the actual and theoretical index. Net efficiencies of the actual indexes for different combinations of relatives can be compared, using the correlation between true daughter superiority and the actual indexes. This correlation is the product of RSE and r_{IG}. On this basis, net efficiency of selection was greatest when the three-point pedigree (sire proof, dam's records, and proof of the dam's sire) was used. Use of the three-point pedigree had only slightly greater net efficiency than using only the sire's proof and the dam's records (.516 versus .513). Use of the sire proof alone was not much less efficient than the three-point index (.479 versus .516). The smallest net efficiency was for use of records of all six relatives. What should be noted is that the best of the actual indexes was only 84% as efficient as the theoretically best index (.614 versus .516).

The reason RSE(a) for the sire proof alone is not unity is that two of the 541 sire proofs were negative, which implied negative selection for those two progeny.

The number of matings associated with different combinations of records should also be considered. These numbers provide an indication of the amount of selection which can be practiced. If records on the maternal sisters and maternal sisters of the dam are required before a mating is made, then the amount of possible selection is markedly reduced, since only 131 of 541 matings were backed by records on all six relatives and only 211 of 541 were backed by records on four types of relatives, including maternal sisters. Nearly all matings included the three most important points in the pedigree, although in general practice this might not be true. The Eastern Artificial Insemination Cooperative apparently required those three points before contracting the mating.

The Artificial Insemination stud would decrease selection efficiency by requiring more than records of paternal sisters, the dam, and paternal sisters of the dam. The optimum procedure, considering both accuracy of evaluation and selection intensity, would be to use the records of the three-point pedigree weighted according to the theoretical weights. A slightly more efficient set of theoretical weights would result if the numbers of records of each paternal sister and paternal sister of the dam were considered.

Comparison of the average actual and theoretical weights provides some indication for the reduced efficiency of the actual indexes. The two pedigrees with negative sire proofs

TABLE 3. Average actual (w) and theoretical (b) selection index weights, relative selection efficiencies, relative selection efficiency (a), correlation between theoretical index and true daughter superiority (r_{IG}), and relative genetic gain from actual versus theoretical index for different combinations of relatives in the actual selection index.

	Averaged weights								
	Paternal sisters	Dam	Paternal sisters of dam	Maternal sisters	Maternal grandam	Maternal sisters of dam	Relative selection efficiency	r_{IG}	Relative genetic gain
w	.461	.212	.333	.004	.261	.081	.729	.614	.448
b	.466	.073	.148	.031	.029	.018			
w	.465	.334	.494	.010	.191		.769	.613	.471
b	.468	.074	.150	.029	.032				
w	.465	.356	.470	.028			.801	.605	.485
b	.467	.081	.142	.033					
w	.464	.336	.419				.867	.595	.516
b	.468	.079	.163						
w	.464	.375					.895	.573	.513
b	.468	.095							
w	.464						.989	.484	.479
b	.468								
w	.463	.320	.450		.144		.829	.605	.502
b	.469	.071	.171		.036				

may have caused some inflation of the average actual weights for records on the maternal side of the pedigree. In general, however, too much emphasis was given to the performance of the maternal relatives, except for maternal sisters. The greatest overemphasis appears to have been on records of the maternal grandam, but from a practical basis the important overemphasis was on the dams' records. Alternatively, the real possibility is that the sire's proof was underemphasized because of more opportunity for selection on the dam than on the sire, or because factors other than daughter milk production were considered in choosing sires to produce young bulls.

Tables 4 and 5 present information similar to that in Table 3 for the two other methods of calculating relative selection efficiency. Both methods make the actual index appear better than does Method (a), especially Method (b) which uses average diagonal coefficients for the phenotypic variance-covariance matrix and average selection differentials. Method (b) does not put much emphasis on serious errors in individual indexes. Both alternative methods of calculating relative selection efficiency do, however, also indicate either an overemphasis on records of the maternal side of the pedigree, particularly of records of the dam, or an underemphasis on the sire.

The efficiency of selection of dams was examined by repeating the procedures after excluding the paternal sisters of the offspring. Results are summarized in Table 6. Calculation of relative selection efficiency by Methods (b) and (c) suggests that nearly optimum emphasis was placed on maternal relatives. The relative selection efficiencies computed by Method (a), however, suggest a slight overemphasis on records of the maternal grandam and maternal sisters of the dam and an underemphasis on records of paternal sisters of the dam. These results suggest that there has been a tendency to put too little selection emphasis on sires for mating to produce young selected bulls. The dams of the potential young bulls

TABLE 4. Actual ($\tilde{w}$) and theoretical ($\tilde{b}$) selection index weights, relative selection efficiencies, relative selection efficiency (b), and correlation between theoretical index and true daughter superiority (r_{IG}) for different combinations of relatives using average diagonal coefficients and selection differentials in the computations.

	Weights from average diagonals and selection differentials							
	Paternal sisters	Dam	Paternal sisters of dam	Maternal sisters	Maternal grandam	Maternal sisters of dam	Relative selection efficiency	r_{IG}
$\tilde{w}$	.461	.183	.158	.043	.033	.039	.926	.609
$\tilde{b}$	.461	.076	.137	.023	.030	.014		
$\tilde{w}$	.464	.184	.206	.034	.039		.931	.609
$\tilde{b}$	.464	.076	.141	.023	.032			
$\tilde{w}$	.463	.192	.228	.038			.927	.600
$\tilde{b}$	.463	.083	.132	.025				
$\tilde{w}$	.466	.194	.242				.929	.591
$\tilde{b}$	.466	.078	.157					
$\tilde{w}$	.466	.217					.926	.570
$\tilde{b}$	.466	.092						
$\tilde{w}$	.466						1.000	.482
$\tilde{b}$	.466							
$\tilde{w}$	.466	.181	.242		.060		.931	.602
$\tilde{b}$	.466	.070	.165		.035			

TABLE 5. Actual $(\overline{w})$ and theoretical $(\overline{b})$ selection index weights, relative selection efficiencies, relative selection efficiency (e), and correlation between theoretical index and true daughter superiority (r_{IG}) for different kinds of relatives using average actual weights standardized to $\overline{b}_1 = \overline{w}_1$.

	Standardized average weights							
	Paternal sisters	Dam	Paternal sisters of dam	Maternal sisters	Maternal grandam	Maternal sisters of dam	Relative selection efficiency	r_{IG}
$\overline{w}$	.466	.214	.336	.004	.263	.082	.762	.615
$\overline{b}$	.466	.073	.148	.031	.029	.018		
$\overline{w}$	.467	.336	.497	.010	.193		.798	.614
$\overline{b}$	.467	.074	.150	.029	.032			
$\overline{w}$	.469	.324	.455	.146			.818	.605
$\overline{b}$	.469	.071	.171	.036				
$\overline{w}$	.468	.339	.423				.829	.595
$\overline{b}$	.468	.079	.163					
$\overline{w}$	.468	.379					.817	.573
$\overline{b}$	.468	.094						
$\overline{w}$	.468						1.000	.468
$\overline{b}$	.468							
$\overline{w}$	.467	.358	.473		.028		.821	.605
$\overline{b}$	.467	.081	.141		.032			

appear to have received more emphasis than sires, although the emphasis on each part of the maternal pedigree has been nearly optimum.

Conclusions

If selection is for milk yield alone, there has been underemphasis on records of paternal sisters in predicting the daughter superiority of sons of planned matings. The underemphasis on records of the sire's daughters accounted for most of the reduction in selection efficiency. If opportunity for selection is considered, selection should be based primarily on the records of three points in the pedigree—sire proof, dam's records, and proof of the sire of the dam. This combination was actually the most effective combination, but was only about 84% as efficient as the theoretically best index. Part of the reduction in efficiency may also be due to emphasis on other traits, such as milk fat percentage and type. Further study will be necessary to determine what weight has been given to milk fat percentage and type relative to milk yield.

References

(1) Allaire, F. R., and C. R. Henderson. 1966. Selection practiced among dairy cows. I. Single lactation traits. J. Dairy Sci., 49: 1426.
(2) Allaire, F. R., and C. R. Henderson. 1966. Selection practiced among dairy cows. II. Total production over a sequence of lactations. J. Dairy Sci., 49: 1435.
(3) Allaire, F. R., and C. R. Henderson. 1967. Selection practiced among dairy cows. III. Type appraisal and lactation traits. J. Dairy Sci., 50: 194.
(4) Dickerson, G. E., C. T. Blunn, A. B. Chapman, R. M. Kottman, J. L. Krider, E. J. Warwick, J. A. Whatley, Jr., M. L. Baker, J. L. Lush, and L. M. Winters. 1954. Evaluation of developing inbred lines of swine. Missouri Agr. Exp. Sta., Res. Bull. 551.
(5) Harvey, W. R., and G. D. Bearden. 1962. Tables of expected genetic progress in each of two traits. Agr. Res. Serv., USDA, 20: 12.
(6) Henderson, C. R. 1963. Selection index and expected genetic advance. Statistical genetics and plant breeding. Natl. Acad. Sci., Natl. Res. Council. Publ. 982: 141.

TABLE 6. Actual (w) and theoretical (b) selection index weights, relative selection efficiencies (RSE), and correlation between theoretical index and true daughter superiority (r_{IG}) by Methods (a), (b), and (c) for different combinations of maternal relatives.

	Average weights					Method (a)		Method (b)		Method (c)	
	Dam	Paternal sisters of dam	Maternal sisters	Maternal grandam	Maternal sisters of dam	RSE	r_{IG}	RSE	r_{IG}	RSE	r_{IG}
w	.076	.126	.031	.046	.027	.716	.381	.977	.375	.980	.381
b	.073	.148	.031	.029	.018						
w	.071	.144	.044	.048		.774	.378	.988	.373	.966	.378
b	.074	.150	.029	.072							
w	.078	.111	.042			.827	.364	.997	.358	.971	.364
b	.081	.142	.033								
w	.077	.173				.928	.346	.991	.342	.999	.346
b	.079	.163									
w	.069	.169		.059		.863	.363	.992	.360	.970	.363
b	.071	.171		.035							