Journal of Genetics Volume 62 Number 2 December 1975

EFFICIENCY OF SELECTIVE BREEDING BASED ON A PHENOTYPIC INDEX

PREM NARAIN AND A.K. MISHRA*
Institute of Agricultural Research Statistics
Library Avenue, New Delhi 110 012

Genetic improvement in a quantitative trait in a population is usually achieved by adopting selective breeding on the basis of the phenotypic values of a trait. The rate of this improvement can, however, be increased if the variation in this trait due to auxiliary traits, particularly at the environmental level, are minimised as far as possible. In such a case selection is made on the basis of an index expressed as deviation of the phenotypic value of the trait from its expected value predicted with the help of the auxiliary traits. Such an index may be called 'phenotypic index' to distinguish it from 'selection index' introduced by Hazel (1943). The genetic improvement expected on the basis of such a selection procedure depends on the number of auxiliary traits, the heritabilities of the traits and the genetic as well as phenotypic correlations between pairs of traits. However, unlike the case of selection index, a knowledge of the estimates of genetic parameters is not necessary for constructing the phenotypic index. As such it is easier to adopt this procedure. If we take the expected genetic improvement in the main trait without the use of any auxiliary traits as a standard of comparison, the efficiency of selective breeding based on the phenotypic index can be expressed as the ratio of the two expected genetic improvements in the main trait. Selection on the basis of phenotypic index is then useful whenever this ratio is expected to be greater than one. This idea of increasing the rate of genetic improvement was first initiated by Rendel (1954) who found that the efficiency of selective breeding for a trait of incomplete heritability may be increased by basing selection on an index which corrects the variation of the main trait for measurable variation introduced by other traits at the environmental level. Osborne (1957) gave a revised estimate for the efficiency of selective breeding for the case when the traits are also genetically correlated. Purser (1960) and Searle (1965) further considered this technique. However, these studies considered only one auxiliary trait. No attempt has so far been made to include more than one auxiliary trait in this method of selection. One can use the technique of partial regression for correcting the variation in the main trait due to several auxiliary traits and investigate the conditions under which this efficiency is increased. This article therefore deals with a study of a 'phenotypic index' based on several auxiliary traits.

---

* Indian Institute of Management, Ahmedabad

THEORY

Consider n auxiliary traits $x_k (k = 1, 2, \ldots, n)$ related to the main trait y. Let the phenotypic value and breeding value of x_k and y, expressed as deviations from the population means, be denoted by $P(x_k)$, $P(y)$ and $A(x_k)$, $A(y)$ respectively. Also, let the phenotypic values be standardised to have unit variances so that the heritabilities of the traits, $h^2(x_k)$ and $h^2(y)$ are the same as the respective genetic variances. Since the regression coefficient of A(y) on P(y) is $h^2(y)$, the expected genetic gain in y due to selection made on the basis of y itself is

$$\triangle G = i\, h^2(y)$$

where i is the intensity of selection.

Consider now selection, with the same intensity, made on the basis of a 'phenotypic index' given by

$$I_P = P(y) - \sum_{k=1}^{n} b_k P(x_k) \tag{2}$$

where b_k is partial regression coefficient of P(y) on $P(x_k)$. The expected genetic gain in y is now

$$\triangle G^* = i\, b_{A(y)I_P} \sigma(I_P) \tag{3}$$

where $b_{A(y)I_P}$ is the regression coefficient of A(y) on I_P and $\sigma(I_P)$ is the phenotypic standard deviation of I_P. The regression coefficient of A(y) on I_P is, in view of (2), equivalent to partial regression coefficient of A(y) on P(y) when $x_1, x_2, \ldots, x_n$ are held constant. Similarly $\sigma(I_P)$ is the standard deviation of P(y) eliminating the effects of the auxiliary traits.

In order to obtain an expression for $b_{A(y)I_P}$ we set up the relationship,

$$E[A(y)] = a_o P(y) + \sum_{k=1}^{n} a_k P(x_k) \tag{4}$$

and evaluate a_o with the help of the resulting normal equations:

$$\begin{aligned} a_o + \underline{R}'_o \, \underline{a} &= h^2(y) \\ a_o \underline{R}_o + \underline{\underline{R}} \, \underline{a} &= h^2(y) \, \underline{C} \end{aligned} \tag{5}$$

where $\underline{R}'_o = (R_{o1}, R_{o2}, \ldots, R_{on})$, R_{ok} being the phenotypic correlation coefficient between y and x_k, $\underline{\underline{R}}$ is the n x n correlation matrix of phenotypic correlation coefficients R_{kl} between the auxiliary traits and $\underline{C}' = (C_1, C_2, \ldots C_n)$, C_k being $r_{ok}\, h(x_k)/h(y)$ where r_{ok} is the genetic correlation coefficient between y and x_k. It may be noted that C_k is the relative efficiency of indirect selection based on x_k as discussed by Searle (1965). We then get

$$b_{A(y)I_P} = a_o = h^2(y)(1-\underline{R}_o'\,\underline{\underline{R}}^{-1}\,\underline{C})(1-\underline{R}_o'\,\underline{\underline{R}}^{-1}\,\underline{R}_o)^{-1} \tag{6}$$

Similarly, to obtain an expression for $\sigma(I_P)$, we set up the relationship,

$$E[P(y)] = \sum_{k=1}^{n} b_k P(x_k) \ , \tag{7}$$

and evaluate b_k 's with the help of

$$\underline{\underline{R}}\ \underline{b} = \underline{R}_o \tag{8}$$

where $\underline{b}' = (b_1, b_2, \ldots, b_n)$. This gives

$$\sigma^2(I_P) = 1 - \sum_{k=1}^{n} b_k R_{ok}$$

$$= 1 - \underline{R}_o'\,\underline{\underline{R}}^{-1}\,\underline{R}_o \tag{9}$$

Denoting the efficiency of selection by phenotypic index relative to individual selection by E_P and using (1), (3), (7) and (10), we get

$$E_P = \Delta G^*/\Delta G$$

$$= (1-\underline{R}_o'\,\underline{\underline{R}}^{-1}\,\underline{C})(1-\underline{R}_o'\,\underline{\underline{R}}^{-1}\,\underline{R}_o)^{-1/2} \tag{10}$$

However, if we consider selection, with the same intensity, made on the basis of the usual selection index of Hazel (1943) given by

$$I_S = P(y) - \sum_{k=1}^{n} w_k P(x_k) \ , \tag{11}$$

we have to choose optimum values of $\underline{w}' = (w_1, w_2, \ldots, w_n)$. For this we maximise the ratio of genetic gain in y due to selection on the basis of I_S to the gain by direct selection on y. This ratio is efficiency of selection based on I_S given by

$$E_S = \Delta G^{**}/\Delta G$$

$$= b_{A(y)I_S}\,\sigma(I_S)/h^2(y)$$

$$= (1-\underline{w}'\,\underline{C})(1-2\underline{w}'\,\underline{R}_o + \underline{w}'\,\underline{\underline{R}}\,w)^{-1/2} \tag{12}$$

It is found that E_S is maximum when

$$\underline{w} = R^{-1}(\underline{R}_o - \underline{C}K) \tag{13}$$

and the maximum value of E_S is given by

$$E_S = (1-\underline{R}'_o \underline{\underline{R}}^{-1} \underline{C} + K \underline{C}' \underline{\underline{R}}^{-1} \underline{C})(1-\underline{R}'_o \underline{\underline{R}}^{-1} \underline{R}_o + K^2 \underline{C}' R^{-1} \underline{C})^{-1/2} \qquad (14)$$

where

$$K = (1-\underline{R}'_o \underline{\underline{R}}^{-1} \underline{R}_o)(1-R'_o \underline{\underline{R}}^{-1} \underline{C})^{-1} \qquad (15)$$

If the genetic correlation coefficients r_{ok}'s are all zero, i.e. $\underline{C} = \underline{O}$, then $\underline{w}$ reduces to

$$\begin{aligned} \underline{w} &= \underline{\underline{R}}^{-1} \underline{R}_o \\ &= \underline{b} \end{aligned} \qquad (16)$$

in view of (8) and I_S reduces to I_P. This means when all the auxiliary characters are related to the main character only at the environmental level, the phenotypic index is optimal with efficiency.

$$E_S = E_P = (1-\underline{R}'_o \underline{\underline{R}}^{-1} R_o)^{-1/2} \qquad (17)$$

However, even if all the r_{ok}'s are not zero, the phenotypic index could be used though its efficiency would then be less than maximal. This is seen from the relationship between the two efficiencies, given by

$$E_P = (E_S^2 - \underline{C}' \underline{\underline{R}}^{-1} \underline{C})^{-1/2} \qquad (18)$$

CONDITIONS FOR THE USE OF PHENOTYPIC INDEX

It is apparent from the theory given above that if we choose auxiliary characters which have no genetic correlation with the main character, the use of phenotypic index for selection is optimal and is expected to result in maximum genetic improvement in the character. However, if we happen to choose auxiliary characters which are genetically related to the main character, the use of phenotypic index may still result in more genetic improvement in the character than that expected on directly selecting for it, provided certain conditions are satisfied. We therefore investigate below the conditions under which E_P is greater than one.

The relation (10) shows that E_P is always more than unity whenever the corresponding elements of the vectors $\underline{R}_o$ and $\underline{C}$ are of opposite signs. But when this is not so, E_P cannot exceed unity if $\overline{C} \geqslant \underline{e}$ where $\underline{e}$ is a vector with unit elements. However, if $\underline{C} < \underline{e}$, the efficiency may or may not be greater than unity. In such cases, in order that $E_P > 1$, we must have

$$\left(\underline{R}_o' - \frac{2\,\underline{C}'}{1+\underline{C}'\,\underline{\underline{R}}^{-1}\underline{C}}\right)\underline{\underline{R}}^{-1}\,\underline{C} > 0 \qquad (19)$$

Thus when the corresponding elements of $\underline{R}$ and $\underline{C}$ are of the same signs, E_P is greater than one provided $\underline{C} < \underline{e}$ and (19) is satisfied. For example, when $n = 1$, we have only two parameters $C_1 = C$ and $R_{o1} = R$ affecting the efficiency. It is always more than one whenever R and C are of opposite signs. However, when R and C are either both positive or both negative, the efficiency exceeds unit only when $|C| < 1$ and $|R| < 2|C|/(1+C^2)$.

It is further seen from (10) that for given values of $\underline{\underline{R}}^{-1}$ and $\underline{C}$, the efficiency is a function of $\underline{R}'$. So differentiating (10) with respect to $\underline{R}_o'$ and equating it to zero, we find that the efficiency is maximum or minimum when $\underline{R}_o = \underline{C}$. The second differential of E_P at $\underline{R}_o = \underline{C}$ is, however, positive so that the efficiency is minimum when the corresponding elements of $\underline{R}_o$ and $\underline{C}$ are equal in magnitude and possess the same sign. It becomes one when either $\underline{R}_o = \underline{o}$ or $\underline{R}_o = 2\underline{C}/(1+\underline{C}'\,\underline{\underline{R}}^{-1}\,\underline{C})$. For one auxiliary character, when R and C are equal and of the same sign, the efficiency is less than one and possesses the minimum value of $(1-R^2)^{1/2}$. It becomes one when either $R = o$ or $R = 2C/(1+C^2)$.

We now assume that the auxiliary traits are uncorrelated so that $\underline{R}$ is an identity matrix. We then have

$$\begin{aligned} E_P &= (1-\underline{R}_o'\,\underline{C})\,(1-\underline{R}_o'\,\underline{R}_o)^{-1/2} \\ &= (1-\sum_{k=1}^{n} R_{ok}C_k)\,(1-\sum_{k=1}^{n} R_{ok}^2)^{-1/2} \qquad (20) \end{aligned}$$

Now the efficiency depends on whether the sum of products of R_{ok} and C_k over all the auxiliary characters is positive or negative. It may happen that for some of the characters R_{ok} and C_k may have opposite signs and for others they may have the same sign. But if the sum of their products happen to be negative, the efficiency will be greater than unity. On the contrary, it will be less than one. The least value of E_P will be $(1-\sum_{k=1}^{n} R_{ok}^2)^{1/2}$ and less than one when $R_{ok}=C_k$ for each k. The value of E_P will be $(1+\sum_{k=1}^{n} R_{ok}^2)(1-\sum_{k=1}^{n} R_{ok}^2)^{-1/2}$ and greater than one when $R_{ok} = -C_k$ for each k. If, however, all C_k's are equal to C but R_{ok}'s are not equal, the efficiency reduces to $(1-C\sum_{k=1}^{n} R_{ok})(1-\sum_{k=1}^{n} R_{ok}^2)^{-1/2}$. Now the efficiency will be greater than one if C and the sum of R_{ok} values over the auxiliary characters are of opposite signs. But if all R_{ok}'s are equal to R but C_k's are not equal, the efficiency becomes $(1-R\sum_{k=1}^{n} C_k)(1-nR^2)^{-1/2}$ which would be greater than unity if R and the sum of C_k values over the auxiliary characters are of opposite signs. If, however, all C_k's are equal to C and all R_{ok}'s are equal to R, we have

$$E_P = (1-nRC)(1-nR^2)^{-1/2} \tag{21}$$

In this form, the effect of the number of auxiliary characters on the efficiency can be seen. The results are presented in Table 1.

Table 1. Effect of the number of traits and the phenotypic correlation on the efficiency of phenotypic index

C	R n	-0.20	-0.10	0.00	0.10	0.20
+0.2	1	1.061	1.025	1.00	0.985	0.980
	5	1.348	1.123	1.00	0.923	0.899
	10	1.809	1.264	1.00	0.843	0.775
	15	2.532	1.410	1.00	0.759	0.633
	20	4.027	1.573	1.00	0.674	0.477
-0.2	1	0.980	0.985	1.00	1.025	1.061
	5	0.899	0.923	1.00	1.123	1.348
	10	0.775	0.843	1.00	1.264	1.809
	15	0.633	0.759	1.00	1.410	2.532
	20	0.447	0.674	1.00	1.573	4.027

It is apparent that the efficiency increases with the number of auxiliary traits when R and C are having opposite signs and it decreases when R and C are having the same signs. For given n, the efficiency decreases as R increases for positive values of C. This relation is, however, reversed for negative values of C. Furthermore, we see that when n is small, the changes in the values of efficiency for different values of R keeping C as fixed is very small but when n is large there is a rapid change in the values of the efficiency. For example, when n = 20, C = +0.2, the efficiency varies from 0.447 to 4.027 but when n = 1, it varies only from 0.980 to 1.061.

APPLICATION TO DAIRY CATTLE BREEDING

In order to demonstrate the practical relevance of the technique of phenotypic index developed in this paper, breeding data on cattle of Kankrej breed collected from an organised herd at Anand (Gujarat), India were used. Records were available for the period 1945 to 1963 in respect of milk yield in first lactation (y), age at first calving (x_1) and body weight of calf at birth at first calving (x_2) for 180 daughters from 13 sires. Table 2 shows the heritabilities, phenotypic and genetic correlations used in the calculation of three different phenotypic indices. The genetic parameters were estimated on an intra-sire basis. The table also shows the three phenotypic indices as well as their estimated relative efficiencies.

Table 2. Efficiency of Phenotypic index for milk yield.

Trait	Heritability		Correlation between traits	
			Phenotypic	Genetic
Milk yield in first lactation (y)	0.415	$y-x_1$	+0.195	-0.048
Age at first calving (x_1)	0.504	$y-x_2$	-0.184	+0.213
Body weight of calf at first calving (x_2)	0.555	x_1-x_2	+0.122	-0.350

Phenotypic indices	Relative Efficiency %
$I_1 = y - 1.34\ x_1$	103
$I_2 = y + 30.56\ x_2$	106
$I_3 = y - 1.51\ x_1 + 35.00\ x_2$	111
Selection on y alone	100

It is apparent from this table that for improving the milk yield in first lactation, the use of age at first calving as an auxiliary trait results in an increase in the efficiency of selective breeding by about 3%. This increase in efficiency rises to 6% if the auxiliary trait happens to be the body weight of calf at birth in first calving. It is interesting to find that when both of these auxiliary traits are used simultaneously to correct for variations in the milk yield in the first lactation, the relative efficiency increases by as much as 11%.

In the above illustration, the estimates of heritabilities and genetic correlations were used to work out the relative efficiencies of the phenotypic index but not the index itself. However, we may use them to obtain the relative efficiencies of the selection index in the three cases. It is found that with x_1, the efficiency is 104, with x_2 it is 110 and with both x_1 and x_2, it is 113. It is apparent that use of selection index instead of phenotypic index would result in higher relative efficiency but the increase is only marginal. At the same time, unlike phenotypic index, the working out of the selection index would require the estimation of genetic parameters. In view of the cost involved in meeting this requirement, the use of phenotypic index may be preferable even though some efficiency is sacrificed, provided it is ensured from the past experience, that the phenotypic and genetic correlations for either of the two auxiliary traits are expected to have opposite signs.

SUMMARY

This article discusses the efficiency of selective breeding based on 'phenotypic index' which is defined as the deviation of the phenotypic value of the trait from its expected value predicted with the help of one or more auxiliary traits. The conditions under which the efficiency of such a procedure is greater than one have been theoretically studied. The practical relevance of this technique has also been demonstrated by applying it to breeding data on cattle.

REFERENCES

HAZEL, L.N. (1943). The genetic basis for constructing selection indexes. Genetics 28, 476-490.

OSBORNE, R. (1957). Correction for regression on a secondary trait as a method of increasing the efficiency of selective breeding. Australian Journal of Biological Sciences, 10, 365-366.

PURSER, A.F. (1960). The use of correction for regression on a second character to increase the efficiency of selection. Biometrical Genetics, Edited by O. Kempthorne, 234 pp. Pergamon Press, London.

RENDEL, J.M. (1954). The use of regression to increase heritability. Australian Journal of Biological Sciences, 7, 368-378.

SEARLE, S.R. (1965). The value of indirect selection. I. Mass Selection. Biometrics 21, 682-707.