

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Papers and Publications in Animal  
Science

Animal Science Department

---

October 1991

## COMPARISON OF METHODS OF ESTIMATING VARIANCE COMPONENTS IN PIGS

J. W. Keele

*University of Nebraska-Lincoln*

T. E. Long

*University of Nebraska-Lincoln*

R. K. Johnson

*University of Nebraska-Lincoln*, rjohnson5@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/animalscifacpub>



Part of the [Animal Sciences Commons](#)

---

Keele, J. W.; Long, T. E.; and Johnson, R. K., "COMPARISON OF METHODS OF ESTIMATING VARIANCE COMPONENTS IN PIGS" (1991). *Faculty Papers and Publications in Animal Science*. 49.  
<https://digitalcommons.unl.edu/animalscifacpub/49>

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

## COMPARISON OF METHODS OF ESTIMATING VARIANCE COMPONENTS IN PIGS<sup>1,2</sup>

J. W. Keele<sup>3</sup>, T. E. Long<sup>4,5</sup> and R. K. Johnson<sup>4</sup>

U.S. Department of Agriculture, Agricultural Research Service,  
Clay Center, NE 68933-0166 and  
University of Nebraska, Lincoln 68583-0908

### ABSTRACT

Components of variance due to average effects of genes ( $\sigma_g^2$ ), environmental effects common to littermates ( $\sigma_c^2$ ), and environmental effects peculiar to individual pigs ( $\sigma_e^2$ ) were estimated ( $\hat{\cdot}$ ) by the Pseudo Expectation Approach (PE). Data were litter size (LS), backfat (BF; centimeter) and ADG (kilograms/day) collected from the Nebraska Gene Pool swine population between 1967 and 1986. Mean square errors (MSE) for  $\hat{h}^2$  and  $\hat{c}^2$  ( $\hat{\sigma}_g^2$  and  $\hat{\sigma}_c^2$  divided by phenotypic variance) by PE and nested ANOVA and  $h^2$  estimated by offspring on parent regression (REGOP) were evaluated using simulation of 200 repetitions of the Nebraska Gene Pool population. Parameter values for  $\sigma_g^2$ ,  $\sigma_c^2$ , and  $\sigma_e^2$  used in simulations were PE estimates from the Gene Pool population. Estimates of  $h^2$  from PE were  $.18 \pm .06$  for LS,  $.56 \pm .06$  for BF, and  $.16 \pm .05$  for ADG. Estimates of  $c^2$  from PE were  $.01 \pm .03$  for LS,  $.09 \pm .02$  for BF, and  $.19 \pm .03$  for ADG. Compared with REGOP, PE yielded  $\hat{h}^2$  with smaller MSE for BF and ADG and larger MSE for LS. The MSE of PE was smaller than the MSE of the nested ANOVA estimate for all estimates and traits. These results were interpreted to suggest that considerable gains in precision in estimation of genetic parameters could be achieved by accounting for all relationships in lieu of accounting for only half- and full-sib relationships or parent-offspring relationships. Key Words: Backfat, Litter Size, Growth Rate, Computer Simulation, Monte Carlo Method, Heritability

J. Anim. Sci. 1991. 69:1428-1434

### Introduction

Estimates of components of variance ( $\sigma_x^2$ , where x is the source of variation) are needed

to develop selection strategies for genetic improvement of pigs. Estimation of  $\sigma_x^2$  requires small amounts of computer time and memory regardless of method if parents do not have observed phenotypes, dams are nested within sires, sires are nested within contemporary group, and parents are unrelated (nested model). In most data sets collected from pigs, both parents and progeny have observed phenotypes, parents have progeny in more than one contemporary group, the progeny of a boar(sow) in one contemporary group may have a different dam(sire) than his(her) progeny in another contemporary group, and there are nonzero relationships among parents. Therefore, the animal model (see Equation 1

<sup>1</sup>Published as Paper No. 9060, Journal Series, Nebraska Agric. Res. Div., Lincoln.

<sup>2</sup>Authors gratefully acknowledge reviewers for helpful suggestions and Sherry Kløver for manuscript preparation.

<sup>3</sup>Author to whom reprint requests should be made. Roman L. Hruska U.S. Meat Anim. Res. Center, P.O. Box 166, Clay Center.

<sup>4</sup>Anim. Sci. Dept., Univ. of Nebraska, Lincoln.

<sup>5</sup>Present address: Univ. of New England, Anim. Genet. and Breed. Unit, Armidale, NSW, Aust.

Received October 13, 1989.

Accepted November 5, 1990.

below; Quaas and Pollak, 1980) provides a more accurate description of most pig data than the nested model. However, assuming the animal model instead of the nested model requires more computer resources to obtain  $\hat{\sigma}_x^2$ . Does the additional precision achieved from using the animal model instead of the nested model justify the increased computational requirements? One objective was to obtain  $\hat{\sigma}_x^2$ ,  $\hat{h}^2$ , and  $\hat{c}^2$  for litter size (LS), backfat (BF), and ADG. A second objective was to compare the precision of  $\hat{\sigma}_x^2$ ,  $\hat{h}^2$ , and  $\hat{c}^2$  by three methods: ANOVA (assuming the nested model; Henderson, 1953), Pseudo Expectation Approach (assuming the animal model; Schaeffer, 1986), and regression of offspring on parent (REGOP).

#### Materials and Methods

**Animal Data.** Data were from the University of Nebraska Gene Pool population, a 14-breed synthetic population that has been closed since 1965. Data and traits investigated were described by Long et al. (1990). Data collected between 1967 and 1986 were included in the analysis. Litter size was defined as number of fully formed pigs at birth (LS). Average daily gain (kilograms/day) was from weaning at 28 d of age to 90 kg (ADG). Backfat probe (centimeter) was adjusted to 90 kg of live weight (BF). Litter size was observed for reproducing females only, whereas BF and ADG were observed for all young pigs. The Gene Pool population is composed of four lines (Figure 1). Selection practices included selection on ovulation rate, LS, age at puberty, and random selection. Data from all four lines were used for LS, but only data from the randomly bred control line were used for ADG and BF. Males and females were only used for one generation, so there was only one litter per female.

**Animal Model.** The assumed model includes average effects of genes for animals ( $g$ ), effects of generation  $\times$  sex subclasses ( $\beta$ ), environmental effects common to littermates ( $c$ ), and other environmental effects peculiar to the individual pig ( $e$ ). The model was:

$$y = X\beta + Zg + Cc + e \quad [1]$$

where  $X$ ,  $Z$ , and  $C$  are incidence matrices

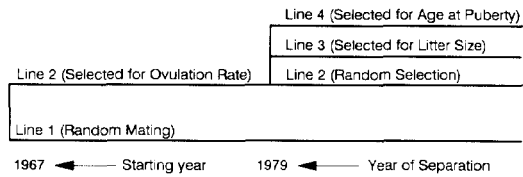


Figure 1. Evolution of Nebraska Gene Pool swine population lines.

relating  $\beta$ ,  $g$ , and  $c$  to phenotypes. Random effects were assumed to have the following distributions:  $g \sim N(O, A\sigma_g^2)$ ,  $c \sim N(O, I\sigma_c^2)$ , and  $e \sim N(O, I\sigma_e^2)$ , where  $A$  is Wright's numerator relationship matrix and  $I$  represents an identity matrix of appropriate size. The random vectors,  $g$ ,  $c$ , and  $e$ , are not correlated with one another; hence, the variance-covariance matrix among the phenotypes is  $\sigma_g^2 ZAZ' + \sigma_c^2 CC' + \sigma_e^2 I$ . Genetic maternal effects and effects due to dominance were assumed to be equal to zero. If these effects are actually not zero, they contribute to  $c$  and result in nonzero covariances among the elements of  $c$  and between the elements of  $c$  and the elements of  $g$ . Genetic maternal effects and effects due to dominance were not simulated, so [1] is correct for the simulated data but possibly not correct for the Gene Pool data. Likewise, the simulated data were normally distributed, whereas the real data may not be.

**Methods of Estimation.** Estimates of  $\sigma_x^2$  can be obtained by the following methods: least squares (analysis of variance; Henderson, 1953), maximum likelihood (ML; Hartley and Rao, 1967), restricted maximum likelihood (REML; Patterson and Thompson, 1971), minimum norm quadratic unbiased estimation (MINQUE; La Motte, 1970, 1971; Rao, 1971), symmetric differences squared (SDS; Grimes and Harvey, 1980), weighted SDS (SDS; Christian, 1980; Keele and Harvey, 1989), pseudo expectation approach (PE; Schaeffer, 1986), tilde-hat approach (TH; Van Raden and Jung, 1988), and Henderson's simple method (HSM; Hudson and Van Vleck, 1982; Henderson, 1986). Also, heritability can be estimated directly as twice the regression of offspring on parent (REGOP; Falconer, 1981). However,  $c^2$  cannot be estimated with REGOP.

Most students of variance component estimation agree that with unlimited computer resources and normally distributed variables REML is the method of choice because it accounts for df used to estimate fixed effects (ML does not), it is robust to certain types of selection bias (Rothschild et al., 1979; Meyer and Thompson, 1984; Sorenson and Kennedy, 1984), it does not depend on prior values for  $\sigma_x^2$ , and it is precise relative to other methods. On the other hand, methods that only yield estimates in the parameter space (i.e., REML, ML, and PE) are biased. This bias becomes larger relative to the parameter value as the parameter value approaches the limits of the parameter space.

Although the nested model, dams nested within sires and sires nested within contemporary group, is not the correct model for the Gene Pool data, we chose to study ANOVA with the nested model because it is computationally and conceptually simple. We chose to study an approximate REML method instead of REML because computing REML under the animal model would have been too time-consuming with the computers that we had available to us. Our choice of PE instead of TH and HSM was arbitrary. One would expect  $\sigma_x^2$ ,  $\hat{h}^2$ , and  $\hat{e}^2$  obtained by REML to be only slightly more precise than estimates by PE when mating is random (Schaeffer, 1986; Ouweltjes et al., 1988; Van Raden and Jung, 1988).

For nested ANOVA, variance components for sire within contemporary group ( $\hat{\sigma}_s^2$ ), dam within sire ( $\hat{\sigma}_d^2$ ), and within full-sib family ( $\hat{\sigma}_e^2$ ) were computed by equating the type I (sequential) mean squares to their expectations and solving for the variance components. Following Falconer (1981),  $\hat{\sigma}_g^2 = 4\hat{\sigma}_s^2$ ,  $\hat{\sigma}_c^2 = \hat{\sigma}_d^2 - \hat{\sigma}_s^2$  (assuming zero variance due to genetic maternal effects and dominance effects) and  $\hat{\sigma}_e^2 = \hat{\sigma}_e^2 - \hat{\sigma}_d^2 - \hat{\sigma}_s^2$ .

Pseudo expectation approach (Schaeffer, 1986) estimates of components of variance at the  $k + 1$  iterate were obtained by the following equations:  $\hat{\sigma}_g^2[k+1] = \hat{g}'[k]Z'My / \text{tr}(Z'MZA)$ ,  $\hat{\sigma}_c^2[k+1] = \hat{c}'[k]C'My / \text{tr}(C'MC)$ , and  $\hat{\sigma}_e^2[k+1] = [y'My - \hat{g}'[k]Z'My - \hat{c}'[k]C'My] / \text{tr}(M)$ , where  $M = I - X(X'X)^{-1}X'$ . For an

arbitrary square matrix  $B$ ,  $\text{tr}(B)$  is the sum diagonal elements of  $B$ .  $\hat{g}^{[k]}$  and  $\hat{c}^{[k]}$  were obtained by solving Henderson's (1973) mixed model equations with  $\hat{\sigma}_g^2[k]$ ,  $\hat{\sigma}_c^2[k]$ , and  $\hat{\sigma}_e^2[k]$  as the parameters.

Inbreeding was taken into account in the computation of  $A^{-1}$  following Quaas (1976). However, the inbreeding coefficient was not included as a covariate in the model. Pseudo expectation estimates of variance components were assumed to be converged when the largest of the absolute values of the deviations of the current estimates from the estimates from the previous iteration was less than  $1 \times 10^{-4}$  for LS,  $1 \times 10^{-5}$  for BF, and  $1 \times 10^{-7}$  for ADG.

Offspring on sire (for BF and ADG) or dam regressions were computed on a within generation  $\times$  sex subclasses (contemporary group) basis. This results in estimates of  $h^2$  that are translation invariant to contemporary group effects because in the Gene Pool populations all the sires or dams of the pigs for a particular contemporary group were reared in a common contemporary group. Parent phenotypes were repeated for every progeny phenotype. Heritability was then calculated as twice the regression coefficient.

*Simulated Data.* Two hundred repetitions of each data set (LS, BF, and ADG) were simulated. Simulations were done so that the pedigree structure of the simulated pigs was the same as that for the Gene Pool population. No selection was simulated even though selection on reproductive characters (LS, number of ovulations, and age at puberty; Figure 1) was practiced in the Gene Pool lines (Zimmerman and Cunningham, 1975; Cunningham et al., 1979; Johnson et al., 1984). Data were simulated following model [1]. The values for  $\sigma_x^2$  used in the simulation were estimates from PE for the Gene Pool population. Elements of  $\beta$ ,  $c$ , and  $e$  were simulated as  $\mu + \sigma_\beta \text{NID}(0,1)$ ,  $\hat{\sigma}_c \text{NID}(0,1)$ , and  $\hat{\sigma}_e \text{NID}(0,1)$ , respectively, where  $\text{NID}(0,1)$  is an independent random number drawn from a normal distribution with a mean of zero and a variance of 1 (IMSL, 1987), and  $\mu$  is the mean of the population. Even though  $\beta$  was simulated as a vector of normal random variables, it was regarded as fixed in model [1]. Simulation of  $\beta$  as a random variable was strictly for convenience and does not influence the results because all the estimates of variance components studied

TABLE 1. NUMBERS OF PIGS, RECORDS, LITTERS, SIRES, DAMS, GENERATION  $\times$  SEX, SUBCLASSES, AND NONZERO ELEMENTS IN THE HALF-STORED COEFFICIENT MATRIX

Variable	Trait		
	Litter size	Backfat	ADG
Number			
All pigs	2,838	2,161	2,240
Pigs with phenotypes	2,099	1,953	2,077
Litters	1,249	652	655
Sires	663	304	304
Dams	1,215	658	661
Generation $\times$ sex subclasses	19	36	36
Nonzero elements in coefficient matrix	16,454	12,619	13,120

are translation invariant. Elements of  $g$  were simulated recursively as follows. Let  $j$  and  $k$  be the sire and dam of the  $i^{\text{th}}$  pig from the Gene Pool population. The  $i^{\text{th}}$  element of  $g$  was  $g_i = \frac{1}{2}(g_j + g_k) + \phi_i \delta_g \text{NID}(0,1)$ , where  $\phi_i$  is equal to  $\frac{1}{2} - \frac{1}{4}F_j - \frac{1}{4}F_k$  if both  $j$  and  $k$  are identified,  $\frac{1}{4}(3 - F_j)$  if  $j$  is identified but  $k$  is not,  $\frac{1}{4}(3 - F_k)$  if  $k$  is identified but  $j$  is not, and 1 if neither  $j$  nor  $k$  is identified. The inbreeding coefficient of the  $i^{\text{th}}$  pig is  $F_i$ .  $j(k)$  and  $g_0$  were set to zero if  $j(k)$  was not identified. The random number stream for each simulation of 200 data sets was initiated with a seed of 173131 so simulations could be repeated if necessary. Values for  $\mu$  were seven pigs/litter for LS, 2 cm for BF, and .5 kg/d for ADG. Values for  $\hat{\sigma}_\beta^2$  were 5 for LS, .075 for BF, and .002 for ADG. For each of the 200 data sets and for each trait,  $\hat{\sigma}_x^2$ ,  $\hat{h}^2$ , and  $\hat{c}^2$  were obtained by each of the three methods. Means and MSE were computed for  $\hat{\sigma}_x^2$ ,  $\hat{h}^2$ , and  $\hat{c}^2$  for each of the three methods and for each trait. Mean square errors were the mean of the squared deviations of estimates from parameter values. Empirical CV for PE were computed as the square root of MSE divided by the parameter values times 100. Both MSE and CV are inflated if estimates are biased.

#### Results and Discussion

*Size of Data Set.* Numbers of all pigs, pigs with phenotypes, litters, sires, dams, generation  $\times$  sex subclasses, and nonzero elements in the coefficient matrix for LS, BF, and ADG are given in Table 1. There were approximately twice as many sires and dams for LS as

there were for ADG or BF. There were approximately twice as many dams as sires for all data sets. The large number of litters relative to the number of pigs with records indicates the small average number of pigs/litter that had phenotypes recorded on them (1.7 pigs/litter for LS, 3.0 for BF, and 3.2 for ADG). This was because postweaning performance was recorded only on pigs selected to become parents. This selection was random for ADG and BF because only data from the random mating line was used for those analyses. The LS data were collected from four lines. Selection was random for one line. For the other three lines, selection practices varied with generation but included selection based on ovulation rate, LS, and age at puberty (Zimmerman and Cunningham, 1975; Cunningham et al., 1979; Johnson et al., 1984).

Storing only nonzero elements of the coefficient matrix saved a considerable amount of computer memory when computing  $\hat{g}$  and  $\hat{c}$ . The percentages of elements on or above the diagonal of the coefficient matrix that were nonzero were only .41% for LS, .54% for BF, and .52% for ADG (Table 1).

*Estimates from Data.* Estimates of  $\hat{\sigma}_x^2$ ,  $\hat{h}^2$ , and  $\hat{c}^2$  by PE are given in Table 2 and their empirical CV are in Table 3. The value for  $\sigma_c^2$  for LS might truly be zero as indicated by the magnitude of  $\hat{\sigma}_c^2$  and its large empirical CV (Table 3) and the fact that PE only yields positive estimates.

Young et al. (1978) reviewed the literature relative to heritability estimates of production traits in swine and, after pooling results from 20 studies of LS and 19 studies for BF, reported an average  $\hat{h}^2$  of .10 for LS and .41 for BF. In this study,  $\hat{h}^2$  for LS and BF tended to be higher than the pooled estimates reported

TABLE 2. ESTIMATES OF COMPONENTS OF VARIANCE ( $\sigma_g^2$ ), HERITABILITY ( $h^2$ ), AND THE PROPORTION OF VARIANCE DUE TO ENVIRONMENTAL EFFECTS COMMON TO LITTERMATES ( $c^2$ ) BY PSEUDO EXPECTATION APPROACH OBTAINED FROM THE NEBRASKA GENE POOL POPULATIONS

Trait	$\hat{\sigma}_g^2$	$\hat{\sigma}_c^2$	$\hat{\sigma}_e^2$	$\hat{h}^2$	$\hat{c}^2$
Litter size	1.30	.04	5.75	.18	.01
Backfat (cm <sup>2</sup> )	.12	.02	.08	.56	.09
ADG, kg/d <sup>2</sup>	$6.19 \times 10^{-4}$	$7.49 \times 10^{-4}$	$25.98 \times 10^{-4}$	.16	.19

by Young et al. (1978). When the Nebraska Gene Pool population was developed, one of the motivations for producing a 14-breed synthetic line was to increase the additive genetic variance in the line above levels found in the foundation breeds. The estimates of  $h^2$  for LS and BF obtained in this study would suggest that an increase in additive genetic variance was accomplished. Estimates of  $c^2$  were low for LS and BF, which is to be expected because these traits are measured at a point temporally removed from the period during which pigs shared a common litter environment.

For ADG, Young et al. (1978), after pooling results from 31 studies measured over various periods, reported an average  $\hat{h}^2$  of .36; ADG was calculated from approximately 42 to 140 d of age. In the present study, ADG was measured from weaning at 28 d to 90 kg. By measuring ADG from an earlier age, growth would tend to be affected to a greater degree by common environmental effects incurred in the preweaning environment, such as size of the litter and nutrition provided by the sow. This would tend to reduce  $\hat{h}^2$  and increase  $\hat{c}^2$ . This can be seen in the  $\hat{h}^2$  and  $\hat{c}^2$  from the current study for ADG, in which  $\hat{c}^2$  is larger than  $\hat{h}^2$ .

Approximately half the data used to calculate  $\hat{h}^2$  and  $\hat{c}^2$  for LS was from lines undergoing selection for either ovulation rate,

litter size, or age at puberty. Van Raden and Jung (1988) have studied the effects of selection on REML, PE, TH, and HSM. Using simulated data, they reported that estimates of  $h^2$  by PE were biased downward when selection was directly on the trait for which variance components were being estimated. A similar result would be expected for a trait with a large genetic covariance with the trait selected. In these data, direct selection for litter size was only in one line in later generations and thus composed approximately 15% of the data. In the earlier generations there was selection for a correlated trait, ovulation rate, and this perhaps could have had an effect on  $\hat{h}^2$  and  $\hat{c}^2$  by PE for LS. However, these estimates would not be expected to be biased downward very much because  $\hat{h}^2$  by PE (.18) was larger than  $\hat{h}^2$  by REGOP (.14; not reported in a table). Data used for parameter estimates for BF and ADG were from the randomly bred control line and, therefore, should not be subject to any of the potential problems selection might pose to estimating variance components.

*Simulation Study.* The average, smallest, and largest number of iterations required for PE estimates to achieve convergence for the simulation study are given in Table 4. More iterations were required for LS data than for BF or ADG. Convergence criteria were met for all replicates.

TABLE 3. EMPIRICAL CV<sup>a</sup> FOR GENETIC PARAMETER ESTIMATES BY PSEUDO EXPECTATION APPROACH

Trait	$\hat{\sigma}_g^2$	$\hat{\sigma}_c^2$	$\hat{\sigma}_e^2$	$\hat{h}^2$	$\hat{c}^2$
Litter size	34.9	595.0	6.3	32.5	552.9
Backfat	15.2	27.0	13.7	11.0	25.9
ADG	32.9	15.0	5.1	31.4	14.3

<sup>a</sup>Empirical CV was computed as  $100 \times (\text{mean squared error})^{.5}$  divided by parameter value.

TABLE 4. AVERAGE, SMALLEST, AND LARGEST NUMBER OF ITERATES REQUIRED TO ACHIEVE CONVERGENCE<sup>a</sup> BY THE PSEUDO EXPECTATION METHOD

Trait	Average	Smallest	Largest
Litter size	89.3	16	215
Backfat	18.6	12	34
ADG	25.5	10	92

<sup>a</sup>Estimates of variance components were assumed converged when the largest of the absolute values of the deviations of the current estimates from the estimates from the previous iterate was less than  $1 \times 10^{-4}$  for litter size,  $1 \times 10^{-5}$  for backfat, and  $1 \times 10^{-7}$  for average daily gain.

The number of replications out of 200 that yielded negative nested ANOVA estimates of variance components or negative REGOP estimates of  $h^2$  are given in Table 5. The number of replicates yielding negative estimates increased when the value of  $\sigma_x^2$  input to the simulation was decreased, as expected. Negative estimates of ANOVA  $\sigma_e^2$  did not occur because this estimate is always positive. Negative REGOP  $h^2$  were obtained only for ADG.

The empirical CV of  $\hat{\sigma}_x^2$ ,  $\hat{h}^2$ , and  $\hat{\epsilon}^2$  by PE are given in Table 3. Empirical CV were computed from deviations from the values input to the simulation so they are a measure of dispersion plus bias. With the exception of  $\hat{\sigma}_c^2$  and  $\hat{\epsilon}^2$  for LS, the CV were less than 100%, an objective that is frequently not achieved when estimating genetic parameters.

The large empirical CV for  $\hat{\sigma}_c^2$  and  $\hat{\epsilon}^2$  for LS were artifacts because the parameter values (PE estimates from real data) were small. In general, PE estimates were relatively precise,

TABLE 5. NUMBER OF REPLICATES OUT OF 200 THAT YIELDED NEGATIVE NESTED ANOVA ESTIMATES OF VARIANCE COMPONENTS OR NEGATIVE REGRESSION OF OFFSPRING ON PARENT (REGOP) ESTIMATES OF HERITABILITY

Trait	$\hat{\sigma}_g^2$	$\hat{\sigma}_c^2$	$\hat{\sigma}_e^2$	REGOP Heritability	
				Sire	Dam
Litter size	13	101	0	—	0
Backfat	0	19	0	0	0
ADG	33	2	0	7	6

TABLE 6. MEAN SQUARE ERRORS (MSE) FOR ANOVA DIVIDED BY MSE FOR PSEUDO EXPECTATION APPROACH

Trait	$\hat{\sigma}_g^2$	$\hat{\sigma}_c^2$	$\hat{\sigma}_e^2$	$\hat{h}^2$	$\hat{\epsilon}^2$
Litter size	4.60	4.98	2.59	5.17	5.86
Backfat	4.14	7.65	3.11	6.78	8.93
ADG	7.53	6.57	5.75	8.43	7.17

indicating that the data set was large enough. The MSE of  $\hat{\sigma}_x^2$ ,  $\hat{h}^2$ , and  $\hat{\epsilon}^2$  by nested ANOVA divided by MSE of  $\hat{\sigma}_x^2$ ,  $\hat{h}^2$ , and  $\hat{\epsilon}^2$  by PE are given in Table 6. Ratios greater than 1 indicate that PE is better than nested ANOVA. Pseudo expectation approach was better than nested ANOVA for all simulations and estimates (ratios of MSE ranged from 2.59 to 8.93).

Ratios of MSE of  $\hat{h}^2$  by REGOP divided by MSE of  $\hat{h}^2$  by PE are given in Table 7. Pseudo expectation approach was worse than REGOP for LS and better than REGOP for BF and ADG.

*Discussion.* The advantage of Pseudo expectation approach over nested ANOVA was greater than its advantage over REGOP for estimating heritability. Doubling the size of the current study by augmenting the Gene Pool population with an independent replicate (with a design identical to that of the Gene Pool population) would result (on the average) in variance component estimates with MSE that are half as large as those obtained from the current study. Substituting PE for nested ANOVA would be expected to gain more information than doubling the size of the current study because the MSE of nested ANOVA estimates were always greater than twice the MSE of PE estimates. This conclusion is in conflict with the views of many experimentalists, who suppose that more can be achieved through increasing numbers of

TABLE 7. MEAN SQUARE ERRORS (MSE) FOR REGRESSION OF OFFSPRING ON PARENT DIVIDED BY MSE FOR PSEUDO EXPECTATION APPROACH ESTIMATES OF HERITABILITY

Trait	Sire	Dam
Litter size	—	.78
Backfat	1.94	1.90
ADG	2.38	2.40

animals than through selection of statistical method.

The estimates of  $h^2$  and  $c^2$  for LS and BF from this study tended to be somewhat higher than pooled estimates from the literature, and  $\hat{h}^2$  for ADG tended to be lower than literature values, but these estimates do fall within the range of values reported by several researchers. Similarly,  $\hat{h}^2$  for all three traits investigated from this population are reasonable.

Selection can bias all methods studied here. Restricted maximum likelihood is less biased and more precise than any of the methods studied here (Ouweltjes et al., 1988; Van Raden and Jung, 1988). Therefore, the advantage of PE over ANOVA or REGOP is a conservative estimate of the advantage of REML over ANOVA or REGOP. In addition, one would expect the advantage of REML or PE with animal model over nested ANOVA and REGOP to be greater than that reported here if the data included parents with progeny in several contemporary groups and dams mated to more than one sire.

#### Implications

Considerable gains in precision of estimating genetic parameters can be realized by accounting for all relationships in lieu of accounting for only half- and full-sib relationships or parent-offspring relationships.

#### Literature Cited

- Christian, L. E. 1980. Modification and simplification of symmetric differences squared procedure for estimation of genetic variances and covariances. Ph.D. Dissertation. The Ohio State Univ., Columbus.
- Cunningham, P. J., M. E. England, L. D. Young and D. R. Zimmerman. 1979. Selection for ovulation rate in swine: Correlated response in litter size and weight. *J. Anim. Sci.* 48:509.
- Falconer, D. S. 1981. *Introduction to Quantitative Genetics* (2nd Ed.). Longman House, Essex, UK.
- Grimes, L. W. and W. R. Harvey. 1980. Estimation of genetic variances and covariances using symmetric differences squared. *J. Anim. Sci.* 50:634.
- Hartley, H. O. and J.N.K. Rao. 1967. Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika* 54:93.
- Henderson, C. R. 1953. Estimation of variance and covariance components. *Biometrics*. 9:226.
- Henderson, C. R. 1973. Sire evaluation and genetic trends. In: *Proc. Anim. Breed. Genet. Symp. in Honor of Dr. Jay L. Lush*. p 10. Am. Soc. Anim. Sci., Champaign, IL.
- Henderson, C. R. 1986. Recent developments in variance and covariance estimation. *J. Anim. Sci.* 63:208.
- Hudson, G.F.S. and L. D. Van Vleck. 1982. Estimation of components of variance by method 3 and Henderson's new method. *J. Dairy Sci.* 65:2018.
- IMSL. 1987. *IMSL User's Manual: Stat/Library, Fortran Subroutines for Statistical Analysis*. IMSL Problem-solving Software Systems, Houston, TX.
- Johnson, R. K., D. R. Zimmerman and R. J. Kittok. 1984. Selection for components of reproduction in swine. *Livest. Prod. Sci.* 11:541.
- Keele, J. W. and W. R. Harvey. 1989. Estimation of components of variance and covariance by symmetric differences squared and minimum norm quadratic unbiased estimation: A comparison. *J. Anim. Sci.* 67:348.
- La Motte, L. R. 1970. A class of estimators of variance components. Tech. Rep. 10, Univ. of Kentucky, Lexington.
- La Motte, L. R. 1971. Locally best quadratic estimators of variance components. Tech. Rep. 22, Univ. of Kentucky, Lexington.
- Long, T. E., R. K. Johnson and J. W. Keele. 1990. Effects of errors in pedigree on three methods of estimating breeding value for litter size, backfat and average daily gain in swine. *J. Anim. Sci.* 68:4069.
- Meyer, K. and R. Thompson. 1984. Bias in variance and covariance component estimators due to selection in a correlated trait. *Z. Tierz. Zuechtungsbiol.* 101:33.
- Ouweltjes, W., L. R. Schaeffer and B. W. Kennedy. 1988. Sensitivity of methods of variance component estimation to culling type of selection. *J. Dairy Sci.* 71:773.
- Patterson, H. D. and R. Thompson. 1971. Recovery of interblock information when block sizes are unequal. *Biometrics*. 58:545.
- Quaas, R. L. 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics*. 32:949.
- Quaas, R. L. and E. J. Pollak. 1980. Mixed model methodology for farm and ranch beef cattle testing programs. *J. Anim. Sci.* 51:1277.
- Rao, C. R. 1971. Minimum variance quadratic unbiased estimation of variance components. *J. Mult. Anal.* 1:445.
- Rothschild, M. F., C. R. Henderson and R. L. Quaas. 1979. Effects of selection on variances and covariances of simulated first and second lactations. *J. Dairy Sci.* 62:996.
- Schaeffer, L. R. 1986. Pseudo Expectation Approach to variance component estimation. *J. Dairy Sci.* 69:2884.
- Sorensen, D. A. and B. W. Kennedy. 1984. Estimation of genetic variances from unselected and selected populations. *J. Anim. Sci.* 59:1213.
- Van Raden, P. M. and Y. C. Jung. 1988. A general purpose approximation to restricted maximum likelihood: The tilde-hat approach. *J. Dairy Sci.* 71:187.
- Young, L. D., R. A. Pumfrey, P. J. Cunningham and D. R. Zimmerman. 1978. Heritabilities and genetic and phenotypic correlations for prebreeding traits, reproductive traits and principal components. *J. Anim. Sci.* 46:937.
- Zimmerman, D. R. and P. J. Cunningham. 1975. Selection for ovulation rate in swine: Population, procedures and ovulation response. *J. Anim. Sci.* 40:61