

This dissertation has been 65-12,475  
microfilmed exactly as received

HILLERS, Joe Karl, 1938-  
GENETIC GAIN IN SIRE SELECTION.

Iowa State University of Science and Technology  
Ph.D., 1965  
Agriculture, animal culture

University Microfilms, Inc., Ann Arbor, Michigan

GENETIC GAIN IN SIRE SELECTION

by

Joe Karl Hillers

A Dissertation Submitted to the  
Graduate Faculty in Partial Fulfillment of  
The Requirements for the Degree of  
DOCTOR OF PHILOSOPHY

Major Subject: Animal Breeding

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University  
Of Science and Technology  
Ames, Iowa

1965

## TABLE OF CONTENTS

	Page
INTRODUCTION	1
SOURCE AND ADJUSTMENT OF DATA	6
METHODS, RESULTS, AND DISCUSSION	17
Rate of Maturity	17
Genetic Improvement	34
Breeding Value of Mates of Sires	53
SUMMARY	60
BIBLIOGRAPHY	64
ACKNOWLEDGMENTS	67

## INTRODUCTION

Selection of sires in dairy cattle breeding, as in all large animal species, is the most important path of genetic improvement in the population. Robertson and Rendel (1950) proposed that 76 percent of the total possible genetic improvement in populations of dairy cattle is contributed by the selection of sires. Flock (1964) studied selection among cows in Iowa D.H.I.A. herds. He found selection of dams of bulls actually resulted in saving bulls for breeding from the top 80 percent of all cows, when the cows were ranked on an index considering only production. He also found there was no selection at all among dams of cows. This supports the proposal of Robertson and Rendel.

Because of the importance of sire selection, much money and time have been spent in developing the methods of sire evaluation through biometrical processes to a rather precise form. Bereskin (1963) has a very thorough discussion of the evolution of present-day sire evaluation methods. Sire proofs used to be based on the daughter-dam comparison which gave the sire credit for any non-genetic changes which occurred between the time the dam and the daughter freshened. Sometimes only the best records of daughters were considered in a sire's proof under the reasoning that best records demonstrated the sire's maximum ability.

Present-day sire proofs are based on the daughter-herdmate comparison. This is the comparison of the daughter with other animals in the herd who freshened during the same period of time. Theoretically at least, this type of a comparison does not confound environmental improvement with sire effects. With the continually increasing use of artificial insemination and frozen semen, many daughters in several herds are usually available for the computation of an A.I. sire's proof. It is also generally accepted that all of the daughter's records should be included in the sire's proof.

The most commonly accepted formula for computing a sire's proof is the one used by the U.S.D.A., for example see Agricultural Research Service (1964);

Predicted daughter avg. = Breed avg.

$$+ \frac{\text{No. of daughters}}{\text{No. of daughters} + 12} (\text{Adj. daughter avg.} - \text{Breed avg.}) .$$

The adjusted daughter average is the deviation of the daughter average from the adjusted average of their herdmates. The adjusted herdmate average is the herdmate average corrected by regression techniques for error in estimating the true herdmate average and for genetic differences between herds. The predicted daughter average is then estimated by using the number of daughters of each sire to correct the

adjusted daughter average for error in estimating the true daughter average, then this is expressed relative to the breed mean.

By studying the formula given above, it is possible to distinguish some of the possible sources of error in obtaining the predicted daughter average. The first obvious place where error could enter the formula would be in the three regressions which are used in its computation. One of these regressions involves the number of daughters and one involves the number of herdmates. The effect of any sampling error in the estimation of these two regressions is reduced as the numbers of daughters and herdmates increases. Probably the largest source of error in the use of these three regressions is the general application of them to data from different parts of the country. For example, are the genetic differences between herds the same in New York, Iowa, and California? Another example would be the difference between New York and Iowa data in the regressions used to correct for number of herdmates (Bereskin, 1963). It is very probable that the formula used by the U.S.D.A. fits average D.H.I.A. data. However, does it allow accurate comparisons between bulls proven in different parts of the country, especially with small numbers of daughters?

Genetic improvement in the population is another source of possible error which the formula can not take into

account. Sire proofs are relative to the average genetic worth of the herdmate population from which they are calculated. There is no credit given for genetic differences between herdmate populations across years or in different areas of the nation, if these genetic differences exist. The rate of genetic improvement in the dairy population as a whole, probably has been and still is very slow. At the present time, comparisons are usually made between bulls who are nearly contemporary, so any improvement made in the population would have very little effect. However, with increased ease of storing semen and the rather extensive use of frozen semen from highly proven A.I. sires for planned matings to produce the next generation of A.I. bulls, the time may come when genetic improvement may bias comparisons between bulls.

Differences between sires in the rate of maturity of their daughters could bias the adjusted daughter average. Other cows who calve in the same herd-year-season as the daughter are used in the computation of the herdmate average. The factors developed by Kendrick (1955) or similar ones developed for a smaller area, are usually used to correct for differences in age. If the daughters of one sire are more mature at a given age than those of other sires in the population, using the same correction factor for all cows at this age will over-correct the daughters of the one sire and

put a positive bias in his proof.

Any non-randomness of the mates of a sire will cause a bias in his proof. This non-randomness of mates will bias the daughter-herdmate comparison in the opposite direction to its effect on the daughter-dam comparison. If the mates of a sire are  $\Delta G$  from the mean genetic worth of the population, the daughters will be expected to be  $\frac{\Delta G}{2}$  from the mean because of their dams. This will cause a bias in the daughter-herdmate comparison of  $\frac{\Delta G}{2}$  and in the daughter-dam comparison of  $\frac{-\Delta G}{2}$ . The direction and magnitude of this bias will be a function of the type of non-randomness used in picking the mates of the sires.

There are many other factors which could bias a sire's proof. One example would be a genotype x environment interaction. Inbreeding and non-additive genetic variance would be other examples. The effect of sampling in the daughters of a sire which get reported and environmental correlations among the daughters can also affect a sire's evaluation.

The purpose of this study was to investigate some of these possible biases and propose possible methods of correction. The possible sources of bias investigated in this study were genetic improvement, rate of maturity, and non-random mating of the sires.

## SOURCE AND ADJUSTMENT OF DATA

The data used in this study were made available by the American Breeders Service Inc. (A.B.S.). The data were collected in 76 large herds in the San Joaquin Valley, the Central Coastal region, and San Bernardino County in California. A.B.S. chose this area to measure the effectiveness of its sire-selection program because of the similarity of climatic conditions throughout the area and the economy of record collection because of large numbers of A.B.S. daughters per herd.

All herds in the area which were on standard D.H.I.A. test, bred at least 100 Holstein cows per year, milking at least 50 daughters of A.B.S. sires and kept accurate herd identification and parentage records were included in this study. An attempt was made to obtain records for all cows calving in the herds during the period in which these data were collected. According to Koch,<sup>1</sup> no attempt was made to select herds on the basis of production except as would occur by meeting the above requirements.

The average production of 299,624 cows on D.H.I.A. test in California in 1963 was 12,689 pounds of milk and 483

---

<sup>1</sup>Koch, H. R., American Breeders Service Inc., Palo Alto, California. Description of data. Private Communication. 1964.

pounds of butterfat. In the general regions where the 76 herds used in this study are located, the Central Coastal Region, San Bernardino County, and the San Joaquin Valley, the D.H.I.A. averages were 12,572 pounds of milk and 480 pounds of butterfat for 229,743 cows (Pelissier<sup>1</sup>). The average actual production of the 3285 cows calving in 1962 in these data was 15,125 pounds of milk and 542 pounds of butterfat. This indicates that these data are from selected herds.

A total of 110,084 records was available in 76 herds for this study. All records were corrected to a 2X milking, 305 day, mature-equivalent basis. The age-conversion factors used were those developed by Kendrick (1955). No exacting analysis was undertaken to determine how well these age-corrections fit the data. Averages for each month of age were obtained from all of the data and the fluctuations in the age corrected data appeared to be of a random nature with respect to age with no evidence of any trends in under- or over-correcting. Lactations in which the cow died or was sold while being milked were extrapolated to 305 days using the factors adopted by the Holstein-Friesian Association in 1958. No records were included in which the cow was milked for less than 150 days. This cut-off point could result in

---

<sup>1</sup>Pelissier, C. L., University of California, Davis, California. Description of California D.H.I.A. data. Private Communication. 1965.

some selection of records.

The modal age of calving in these data was 30 months while in Iowa it is 25 months. It is often the practice in these California herds to rear the heifers on rather poor pasture, which probably results in heifers being bred later than is the common practice in Iowa.

Records were available from some of the herds starting in January, 1951 with records from other herds becoming available at various times until June, 1958 when the last herd entered the program. The number of records per herd ranged from 154 to 5626. The last completed lactations available in these data were started in October, 1963. However, the number of lactations started after December, 1962 was so small that they were discarded in the analysis of these data.

A model explaining the environmental effects of herds and seasons could be written as:

$$Y_{ijk} = \mu + S_i + H_j + (SH)_{ij} + e_{ijk} \quad (1)$$

where  $Y_{ijk}$  is the production of the  $k^{\text{th}}$  record started in the  $j^{\text{th}}$  herd in the  $i^{\text{th}}$  year-season. Because of the large size of the herds involved in this study, it was decided to use seasons which were two months in length. Using seasons of this length meant losing some exactness in estimating the herd-year-season average, but some exactness in correcting

for seasonal fluctuations was gained by this procedure as compared to using six-month seasons. The six seasons per year were January-February, March-April, May-June, July-August, September-October, and November-December. With these divisions, the number of cows calving in each herd-year-season ranged from 1 to 353 with an over-all average of 25.4. At least one lactation was started in 79.1 percent of the possible herd-year-seasons which occurred from January, 1951 to December, 1962 in the 76 herds in this study.

Analyses of variance were computed using Model (1) to determine the importance of the herds and seasons as sources of variation. The mean squares were equated to their expectations and the resulting components of variance are given in Table 1. The corresponding percentage values obtained by Bereskin (1963) in Iowa D.H.I.A. data are given in Table 2.

Table 1. Analyses of herds and seasons

Source	df	Milk		Butterfat	
		Component	Percent of total variance	Component	Percent of total variance
Season	71	1,728,755	18.5	2197	17.9
Herd	75	1,467,613	15.7	2082	16.9
Season x herd	4,182	443,471	4.7	560	4.6
Within	105,755	5,721,800	61.1	7458	60.7

Table 2. Bereskin's analyses of herds and seasons

Source	Percent of total variance	
	Milk	Butterfat
Seasons	2.3	1.8
Herds	28.1	29.2
Seasons x herds	3.2	4.1
Within	66.4	64.9

Other studies (Barr, 1962; Van Vleck *et al.*, 1961d) have shown that seasons account for less than 5 percent of the total variance and herds account for more than 25 percent of the total variation. One of the reasons that seasons account for so much more variation in the California data than was found in the other studies is the definite time trend which was present in the California data. Milk production increased from 12,122 pounds in the November-December season of 1951 to 17,149 pounds in the November-December season of 1962. The relatively small herd component in these data was probably due to the requirements which were set up at the initiation of the collection of this data which tended to restrict the differences between herds.

The effect of sires was not included in Model 1 as it was in other studies of this type (Barr, 1962; Henderson, 1956). Leaving the effect of sires out of the herd-year-season

analysis requires the assumption of independence of the sire effect from these environmental effects. This assumption should hold reasonably well in these data since there is no reason to suspect that there has been any significant change in the ability of A.B.S. to pick genetically-superior bulls for use in their stud during the time period involved in these data or that better sires were used in better herds.

The applicability of the results of this study of herd and year effects to the entire California D.H.I.A. population would be limited since these herds are a select group of California herds. For example, the between-herd component obtained in this study would probably be an under-estimate of the between-herd component expected from the whole D.H.I.A. population. Also, all California herds may not have the same time trend present in these 76 herds. There will be no attempt to make inferences about the effects of seasons and herds to the entire California D.H.I.A. population. The other effects investigated later in this study should be relatively independent of herd and season effects and more applicable to regular dairy data. This would mean the interpretation of the effects other than herd and season would not have to be limited to these 76 herds.

The within herd-year-season components of variance are higher in this study than the ones found by Bereskin (1963) in Iowa data. The components are 5,721,800 compared to

5,014,000 for milk and 7458 compared to 6764 for butterfat. This difference could be attributed to the numbers which were in the herd-year-seasons of the two studies. As stated previously, the average number of cows calving in a herd-year-season in these data was 25 cows, while it was only 12 in the Iowa data. The fact that Bereskin (1963) used only two seasons per year compared to six in this study would tend to increase the within herd-year-season component in the Iowa data because of less complete season correction.

Another factor which could cause the variance to be greater in this study than was found by Bereskin was the difference in production. Average production in this study was 13,785 pounds of milk and 504 pounds of butterfat. The corresponding averages found by Bereskin were 12,278 pounds of milk and 444 pounds of butterfat. Studies generally show an increase in variance corresponding to an increase in average production for these two characters (Van Vleck, 1963).

All records used in this analysis were expressed as deviations from their regressed, adjusted herd-mate average as discussed by Van Vleck et al. (1961a). This was done to remove a large portion of the environmental variance between herd-year-seasons.

Using the previously stated model, the herd-year-season average could be written:

$$\bar{y}_{ij.} = \mu + S_i + H_j + (SH)_{ij} + \frac{\sum e_{ijk}}{n_{ij}} \quad (2)$$

The corresponding true herd-year-season average would be:

$$\bar{Y}_{ij} = \mu + S_i + H_j + (SH)_{ij} \quad (3)$$

With the usual assumptions of the analysis of variance, the least squares estimate of the true herd-year-season average could be written as:

$$\hat{\bar{Y}}_{ij.} = \mu + \frac{\hat{\sigma}_S^2 + \hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2}{\hat{\sigma}_S^2 + \hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2 + \frac{\hat{\sigma}_e^2}{n_{ij}}} (\bar{y}_{ij.} - \mu) \quad (4)$$

Equations 2 and 3 are usually written (Bereskin, 1963; Heidhues et al., 1961):

$$\bar{y}_{ij.} = \mu_i + H_j + (SH)_{ij} + \frac{\sum e_{ijk}}{n_{ij}} \quad (5)$$

and

$$\bar{Y}_{ij} = \mu_i + H_j + (SH)_{ij} \quad (6)$$

by absorbing the year-season effect into the mean, thus forming a year-season mean. This would mean that the least squares estimate of the true herd-year-season average could be written as:

$$\hat{\bar{y}}_{ij.} = \mu_1 + \frac{\hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2}{\hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2 + \frac{\hat{\sigma}_e^2}{n_{ij}}} (\bar{y}_{ij.} - \mu_1) \quad (7)$$

It is possible to show that  $\sigma_H^2 + \sigma_{SH}^2$  can be estimated by the component  $\sigma_{H/S}^2$ , the herd within year-season component, computed from an analysis using a hierarchal model:

$$y_{ijk} = \mu + S_i + H_{ij} + e_{ijk}$$

where  $y_{ijk}$  is the same in this model as in the previous model. This can be demonstrated by taking the expectation of the herd within year-season mean square with respect to Model 1.

This estimate of the regression of herd-mate average on its expected value,

$$\frac{\hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2}{\hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2 + \frac{\hat{\sigma}_e^2}{n_{ij}}},$$

is different from what Bereskin (1963) used in his study. Bereskin used as an estimate

$$\frac{\hat{\sigma}_b^2}{\hat{\sigma}_b^2 + \frac{\hat{\sigma}_e^2}{n_i}},$$

where  $\hat{\sigma}_b^2$  is the between herd-year-seasons component of

variance,  $\hat{\sigma}_e^2$  is the within herd-year-season component of variance and  $n_i$  is the number in the  $i^{\text{th}}$  herd-year-season. By taking the expectations of the between herd-year-season mean square with respect to model (1), the component  $\hat{\sigma}_b^2$  can be shown to estimate  $\sigma_{SH}^2 + K_1\sigma_H^2 + K_2\sigma_S^2$  with both  $K_1$  and  $K_2$  being less than one when unequal numbers are involved. When the effect of year-season is put into the mean as in Equations 5, 6, and 7, the component of variance for year-seasons should not be included as part of the regression estimate. The practical extent of the difference between the regression estimate in Equation 7

$$\frac{\hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2}{\hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2 + \frac{\hat{\sigma}_e^2}{n_{ij}}}$$

and Bereskin's (1963),

$$\frac{\hat{\sigma}_b^2}{\hat{\sigma}_b^2 + \frac{\hat{\sigma}_e^2}{n_i}}$$

will be a function of the magnitude of the year-season component of variance. In most studies, as stated previously, the year-season component amounts to less than five percent of the total variance. This would have a very small effect on the regression. Another way of expressing these regres-

sions is  $\frac{n}{n + A}$  where  $A = \frac{\hat{\sigma}_e^2}{\hat{\sigma}_b^2}$  in Bereskin's study or

$$A = \frac{\hat{\sigma}_e^2}{\hat{\sigma}_H^2 + \hat{\sigma}_{SH}^2}$$

in this study. If the year-season component is included in the regression estimate, the estimate of A in this study would be about 1.5 and when it is not included  $A = 2.9$ . The difference is not this large in Bereskin's (1963) data. Because of these results, the regression estimate used to predict the true herd-year-season average in this study was  $\frac{n}{n+3}$ . This corresponds to the  $\frac{n}{n+2}$  used by Bereskin (1963) and Barr (1962) and the  $\frac{n}{n+1}$  used by the Cornell group (Van Vleck et al., 1961d).

No estimate of the genetic fraction of the between-herd variance was made from these data. It was assumed that 0.9 of the differences between herds were due to environmental factors. This is the value accepted by Bereskin (1963) and others. In reality, something more than 0.1 of the herd differences in these data may be genetic, because of the small between-herd variance relative to that found in other studies. However, one of the requirements before a herd could enter the bull-proving program was that it have at least 50 daughters of A.B.S. bulls. This restriction would certainly reduce the genetic differences between herds. Because of these offsetting factors, it was decided that 0.9 would be used in this study.

## METHODS, RESULTS, AND DISCUSSION

## Rate of Maturity

One factor which could cause a sire's proof to be biased would be differential maturity rates of daughters of different bulls. This would be especially true of a bull's initial proof, which would probably be based only on first lactations. If the daughters of one sire mature slower than the daughters of another sire, it could be possible to get a change in the ranking of these two bulls if second lactations were used as compared to the ranking based on first lactations whether the average ages differ or not.

This problem has been investigated previously by Hickman and Henderson (1955) by comparing the increase from first to second lactation of different sires' progeny. The results of this type of investigation are dependent upon the type of selection which was practiced on the heifers on the basis of their first lactations. This investigation will be discussed later in this section.

Only first lactations were used in the present study in an attempt to remove the effect of selection from the results. There is some selection, even among first lactations, since there may be some family selection, but this is unavoidable in field data. Interpretation of results found in this study are limited because the ages used as first lactations, 23 to

35 months, give little information on ultimate mature production. However, real differences between sires in rates of maturity should be easiest to demonstrate at the immature ages involving the steepest part of the maturity curve.

For these reasons, the regression of actual production during first lactation on age in months was used as a measure of rate of maturity in this study. A simple linear regression of actual production on age in months was calculated for each sire from his daughter's production. All lactations started between the ages of 23 and 35 months were considered as first lactations and used in this portion of the study.

Between these ages, there is a slight curvilinearity in the rate of maturity. This would mean that the linear regressions used in this study could estimate different slopes if the average age of the daughters differed from sire to sire. The average age of the daughters of the sires used in this part of the study ranged from 28.5 to 30.8 months of age. The standard deviations of age of the daughters of these sires ranged from 2.4 to 3.4 months. From these statistics, it was concluded that the possibility of the curvilinearity of the maturity curve affecting the linear regressions was so slight that an investigation of deviations from linearity was not warranted. If curvilinearity was the only cause of the differences between the linear regressions, it could be removed by regressing

age-corrected production on age. This would correct for differences due to curvilinearity, but would not correct for maturity differences between sires. Multiplicative age factors, for example those developed by Kendrick (1955), are computed to remove this curvilinearity. The regression of age-corrected production on age should be zero if the age-correction factors used are the proper ones for the population being studied.

Another factor which could cause the regressions to differ would be the sampling error in estimating the regressions from small numbers of daughters per sire. There were 28 sires in the study who had at least 44 daughters with first lactations. Only these 28 sires were used in this portion of the analysis. The average number of daughters per sire was 337 with a range of 44 to 1287. With only 13 possible classes on the independent axis, numbers of this size should give reasonable accuracy in estimation.

The within-sire regressions of actual production on age are given in Table 3.

The average regressions, weighted by number, are +161.89 pounds of milk and +6.06 pounds of butterfat per month of age. The standard errors of the regressions given in Table 3 range from 19 to 92 for the milk regressions and from 0.7 to 3.4 for the butterfat regressions.

These standard errors are not valid estimates in the

Table 3. Within-sire regressions of production on age

Sire	Number of daughters	Regressions	
		Milk	Fat
1018	143	+215.40	+11.23
1021	323	+258.16	+10.29
1028	458	+115.31	+ 4.75
1117	54	+249.50	+ 8.89
1118	44	+103.14	+ 5.49
1122	88	+127.18	+ 6.31
1126	109	+207.68	+ 5.68
1601	1278	+147.51	+ 5.80
1602	138	+234.30	+ 8.73
1603	767	+205.52	+ 7.06
1605	429	+203.10	+ 8.07
1606	343	+161.42	+ 7.10
1607	818	+193.14	+ 6.76
1608	633	+199.49	+ 7.21
1609	558	+176.52	+ 6.31
1610	555	+172.41	+ 5.83
1611	192	+ 49.24	+ 1.38
1702	224	+132.52	+ 3.75
1717	105	+ 96.39	+ 4.47
1718	72	- 9.20	- 3.73
1720	315	+163.08	+ 6.69
1725	532	+ 45.12	+ 2.22
1728	307	+155.48	+ 5.30
1735	323	+ 94.19	+ 3.91
1738	55	+152.93	+ 5.02
1741	318	+ 96.17	+ 4.02
1746	138	+115.49	+ 4.51
1747	130	+243.55	+ 8.04

strict sense. One assumption used to estimate the standard error of a regression is that the variance of the dependent variable will be the same for each class on the independent axis. This assumption may not hold in this analysis if there is a positive correlation between the mean and the variance for these production traits in dairy cattle (Van Vleck, 1963).

This change in the variance will not affect the regression estimate itself, since the regression merely connects the means of the conditional distributions. The estimate of the standard errors obtained in this study are a weighted average of functions of the variances of the different conditional distributions involved.

To obtain an indication of the importance of the heterogeneity of the within-sire regressions of production on age, analyses of variance were computed to test the added reduction in variance obtained from fitting the individual regressions. The results are given in Table 4. Both of the tests are significant with  $P < .01$ .

These analyses indicate that there is heterogeneity between sires in the rate of maturity of their daughters, at least for this measure of it. These are not exact tests of significance because homogeneity of variance is an assumption in an analysis of variance. However, Cochran (1947) stated the over-all F test is probably the least affected by heterogeneity of variance. The stated levels of probability may not be exact but it would appear that there are real differences between sires for this measure of rate of maturity.

Another factor which could cause the within-sire regressions to be different would be the incomplete cross-classification of herds and sires, since each sire did not appear in each herd. The sire with the smallest number of

Table 4. Analyses of variance to test heterogeneity of regressions of production on age

Source	Degrees of freedom	Sums of squares	Mean squares
Milk			
Total within sire	9430	45,133,385,197	
Reduction due to fitting $b_1$	28	2,394,450,384	
Reduction due to fitting $b$	1	2,160,114,900	
Additional reduction	27	234,335,484	8,679,092
Error	9402	42,738,934,813	4,545,728
computed F = 1.91			
Butterfat			
Total within sire	9430	61,519,122	
Reduction due to fitting $b_1$	28	3,443,815	
Reduction due to fitting $b$	1	3,082,117	
Additional reduction	27	361,698	13,396
Error	9402	58,075,307	6,176
computed F = 2.16			

daughters, 1118, had daughters in 13 different herds. The sire with the next smallest number of daughters, 1117, had progeny in 15 different herds. The sire with the most daughters, 1601, appeared in all but three of the seventy-six herds.

Regressions of actual production on age were calculated within each herd and year to indicate if these environmental factors influence maturity rate. The resulting regressions are given in Tables 5 and 6.

The regressions of actual milk production on age within herd ranged from -317.66 pounds to +284.45 pounds. The corresponding range for butterfat production was from -11.45 pounds to +11.07 pounds. These indicate large differences between herds in the rate of maturity of their daughters. One possible explanation of the negative regressions of production on age which are present in some herds would be the increase in production across years coupled with a decrease in age of first calving.

The regressions of production on age were estimated in each year from 1951 to 1962 in herd 19 in an attempt to find an explanation for the negative regressions obtained in that herd. This herd was chosen because of its large size and negative regression. These results along with average production and age are given in Table 7.

The averages in Table 7 show there was an increase in

Table 5. Regression of first lactation actual production on age within herd

Herd	Number	Milk	Butterfat	Herd	Number	Milk	Butterfat
1	128	+112.14	+ 4.97	41	378	+ 48.38	+ 1.39
2	203	+149.71	+ 4.87	42	101	+252.42	+ 8.06
3	147	+189.27	+ 3.34	43	1071	+220.37	+ 7.40
4	261	+179.24	+ 9.12	44	497	+177.03	+ 6.18
5	138	+263.07	+10.49	45	292	+118.16	+ 1.46
6	380	+ 26.56	+ 2.31	46	260	+ 42.33	+ 1.90
7	417	+ 82.13	+ 3.51	47	621	+ 32.47	+ 1.12
8	341	- 50.82	- 0.73	48	281	+121.50	+ 5.63
9	356	+242.52	+11.07	49	205	+110.64	+ 2.86
10	140	+ 60.35	+ 0.94	50	190	+160.95	+ 5.66
11	317	+193.45	+ 8.54	51	117	-175.17	- 0.96
12	299	+284.45	+10.53	52	163	+ 25.22	+ 2.53
13	909	-230.05	- 8.73	53	634	-317.66	-11.45
14	511	+217.30	+ 9.40	54	225	+ 68.21	+ 3.95
15	207	+257.47	+ 9.64	55	313	- 51.82	- 2.67
16	293	+195.14	+ 6.27	56	636	+ 3.51	- 0.74
17	472	+193.12	+ 7.29	57	457	+168.06	+ 7.01
18	1340	+136.80	+ 4.83	58	240	+ 64.49	+ 2.49
19	1554	-230.32	- 6.90	59	81	-161.38	- 6.48
20	548	+ 47.41	+ 1.50	60	108	+ 73.15	+ 5.96
21	1062	+191.00	+ 6.99	61	181	+116.51	+ 5.78
22	311	+240.51	+ 8.10	62	265	+ 73.83	+ 3.31
23	712	+182.82	+ 6.27	63	499	+179.03	+ 6.57
24	329	+116.94	+ 4.31	64	256	+ 38.74	+ 0.12
25	198	+ 34.76	+ 2.78	65	450	+ 47.42	+ 0.65
26	384	+216.75	+ 8.77	66	233	-267.57	- 9.06
27	262	+ 7.70	- 1.48	67	377	+ 7.03	+ 2.42
28	234	+ 73.71	+ 2.94	68	39	+154.79	+ 2.82
29	168	-165.45	- 5.34	69	240	- 44.91	- 1.36
30	173	- 6.00	- 0.87	70	428	- 77.21	- 2.56
31	205	+243.17	+ 7.81	71	62	+130.48	+ 5.80
32	249	- 28.57	- 4.03	72	15	+196.08	+ 9.72
33	323	+145.17	+ 4.78	73	158	- 39.40	+ 2.17
34	561	+106.43	+ 3.35	74	137	+220.84	+ 4.86
35	542	+ 65.75	+ 1.87	75	403	-207.79	- 9.28
36	224	+164.71	+ 4.65	76	27	+114.74	+ 1.31
37	248	+157.65	+ 6.02				
38	283	+244.67	+ 7.49				
39	163	+176.12	+ 6.51				
40	122	+214.40	+ 8.41				

Table 6. Regression of first lactation actual production on age, average production, and age for each year

Year	Number	Regression		Averages		Age
		Milk	Butter-fat	Milk	Butter-fat	
1951	1139	- 11.85	-0.65	9301	347	30.1
1952	1246	+146.25	+5.49	9607	352	30.1
1953	1475	+152.38	+5.40	9898	365	30.4
1954	1847	+123.01	+4.83	9936	367	30.0
1955	2230	+100.21	+4.37	10409	384	30.3
1956	2535	+154.05	+5.82	10390	388	30.2
1957	2865	+113.24	+4.08	11140	416	30.1
1958	3093	+161.59	+6.07	11259	419	30.4
1959	3206	+ 44.98	+2.34	11847	437	30.6
1960	3386	+ 94.86	+4.16	12475	461	30.7
1961	2566	+ 53.19	+2.24	12619	468	30.3
1962	766	+120.91	+4.20	13148	478	30.5

Table 7. Regression of actual production on age in herd 19

Year	Number	Regressions		Averages		Age
		Milk	Butter-fat	Milk	Butter-fat	
1951	114	-612.84	-20.37	8806	344	32.1
1952	153	-105.02	- 4.53	9908	365	31.8
1953	135	-114.44	- 3.75	9928	369	31.9
1954	198	-248.40	- 9.73	10074	363	31.6
1955	211	- 4.65	- 1.10	10761	386	31.5
1956	185	- 26.38	- 1.77	11019	398	30.5
1957	86	+ 90.59	+ 2.69	11495	410	30.1
1958	102	+117.60	+ 3.80	11296	407	29.4
1959	85	+ 68.17	+ 4.90	11849	417	29.9
1960	96	+ 84.41	+ 4.30	13317	462	29.1
1961	97	+ 50.97	+ 2.91	13743	478	30.3
1962	92	+ 74.08	+ 0.31	13605	476	29.8

production accompanied by a decrease in age in herd 19. This could cause the over-all regression of production on age to be negative. There is no readily apparent reason why the regressions for the early years in herd 19 are negative. The average ages given in Table 6 for each year do not show any over-all trend in age of calving. If all of the herds with negative regressions of production on age in Table 5 had a decline in average age of calving, some herds would have had an increase in average age of calving. This would cause the regressions in these herds to be more highly positive. The results in Table 7 show the regressions of actual production on age are apparently non-homogeneous across herd-year combinations, at least in this large herd. The apparent non-homogeneity of these herd-year regressions would confuse the interpretation of any over-all regression of production on age.

In an attempt to determine the dependency of the within-sire regressions on the within-herd regressions for the herds in which the sires appeared, the average within-herd regressions were obtained for two sires. The two sires picked, 1718 and 1738, have small numbers of daughters and have regressions of actual milk production on age of -9.20 pounds and +152.93 pounds, respectively. If herds affect the within-sire regressions of production on age, it should be evident for the sires who appear in relatively few herds.

Each sire had daughters in eighteen different herds and ten of these eighteen herds had daughters of both sires. The unweighted averages of the within-herd regressions for milk production were +56.7 for sire 1718 and +49.0 for sire 1738. The average ages of the daughters of these sires at first calving were 28.5 and 28.7 months respectively. Sire 1718 was in the stud from 1956 to 1960 and sire 1738 was in the stud from 1957 to 1962. These results would indicate that it is unlikely that differences between herds and years would cause the difference between the regressions for these two bulls.

The regressions of production on age calculated for each year are given in Table 6. They are positive except for those in 1951. There may be some selection of records entering in the regressions for 1951. A.B.S. started maintaining these records in 1957 and took all available records in the herds back through 1951. If some records were destroyed or lost from 1951 to 1957, they could have been for the lower or poorer cows in production. This type of selection should have raised the average production in that year, but this is not apparent from the averages given in Table 6.

A positive correlation between production and growth rate could be a factor in causing the negative within-herd regressions of production on age. It is possible that the

management of the herds with negative regressions bred their heifers when they reached a certain weight. If this were true, and there was a positive correlation between production and growth rate, the regressions would tend to be negative. However, most herds probably have their heifers bred when they reach a certain weight, at least within age limits. It is rather hard to visualize that breeding is done on a weight constant basis in only fifteen of the herds.

In an attempt to determine the relative importance of sires and herds in determining rate of maturity, regressions were calculated for each sire-herd combination. An analysis of variance was computed using the model:

$$Y_{ij} = \mu + S_i + H_j + (SH)_{ij} + e_{ij}$$

where,  $Y_{ij}$  is the regression of actual production on age for the daughters of the  $i^{\text{th}}$  sire in the  $j^{\text{th}}$  herd. Two separate analyses were done; one involved regressions from all sire-herd combinations which had two or more daughters in each subclass ( $n > 1$ ), the other involved only those regressions which had 3 or more daughters in them ( $n > 2$ ). The resulting components of variance are given in Table 8.

In the analyses involving all the regressions, the ratio of the herd component to the sire component is 3.0 for milk and 1.0 for butterfat. In the analyses involving only those regressions with more than 2 daughters, the corresponding

Table 8. Components of variance for analyses of sire - herd regressions of actual production on age

Source	df	Milk	Butterfat
n > 1			
Sires	27	23,191,120	39,984
Herds	75	69,909,930	42,107
Error	981	4,004,093,400	5,209,445
n > 2			
Sires	27	24,506,427	28,621
Herds	75	51,613,610	15,014
Error	832	3,276,499,400	4,598,081

ratios are 2.1 and 0.5. The difference between the ratios obtained from the two analyses might be due to the relative number of observations per sire and per herd. In the analyses involving all regressions, the error in estimation,  $e_{ij}$ , has less chance of averaging out of the between-herd differences than out of the between-sire differences since there are about three times as many observations for each sire as for each herd.

With only one observation per subclass, it is impossible to estimate the relative importance of the two error

components:  $(SH)_{ij}$ , the sire x herd interaction, and  $e_{ij}$ , the error of estimation. The reasoning presented in the previous paragraph is more valid if the error mean square is mainly due to  $e_{ij}$ , the error of estimation. In the extreme case where there is no  $e_{ij}$  effect, the reasoning is invalid. The ratio of the sum of the herd and sire component to the error component is as great or greater for the analyses where  $n > 2$  or the number of daughters per subclass was 3 or more. This indicates that the error mean squares were not mainly due to  $e_{ij}$ . However, taking out the 149 regressions estimated with only two observations, probably causes little reduction in the average size of  $e_{ij}$ .

The unequal numbers in these data will force a positive correlation between the sire and herd components, since the effect of both are measured on the same animals. This will cause the sum of the two components to be larger in comparison to the error mean squares than would occur if equal numbers were present. There is no way of determining the extent of this correlation or how it affects the sire and herd components.

If the  $(SH)_{ij}$  component is important, the measurement of the maturity rate of the daughters of a sire will depend on what herds have used the bull. As discussed previously, the relative size of the error mean square to the sum of the sire and herd components for the two analyses indicates that

the  $(SH)_{ij}$  term could be important. There is no obvious reason why the error component should be so much larger than the sum of the herd and sire components for the butterfat analyses than for the milk analyses. Hickman and Henderson (1955), using the increase from first to second lactation as a measure of rate of maturity, found the  $(SH)_{ij}$  component was 3 and 130 times as large as the sire component for milk in two different analyses. This tends to confirm that maturity rate of the daughters of a sire is a function of the herds in which the sire has daughters.

Hickman and Henderson (1955) found the ratio of the herd component of variance to the sire component of variance for increase from first to second lactation to be 6.5 for milk and 6.9 for butterfat from one analysis and 0.9 for milk and 1.5 for butterfat from another analysis. Year-seasons were included as a component in the latter analyses and not in the former analyses. Apparently herds were partially confounded with year-seasons. The validity of the results obtained from using the increase from first to second lactation as a measure of rate of maturity depends upon the assumption that selection differentials for first lactations are the same for all sires. If there are differences between sires in rate of maturity, unequal selection differentials would tend to make these differences less extreme than would be expected if there were no selection at all. However, they

found no evidence of unequal selection differentials among sires. Analyses of this type only test for significance among average differences between sires and do not exclude the possibility that there are differences between individual sires in selection differential.

Hickman and Henderson (1955) computed estimates of heritability of increase from first to second lactation that ranged from 0.10 to 0.24. These heritabilities indicate some genetic control of increase from first to second lactation or at least between-sire differences. Their estimates of the heritability of milk and butterfat production were between 0.40 and 0.64 which are higher than are usually found.

Robertson and Khishin (1958) present the reasoning that if there is a lower genetic variance among second lactations than among first lactations as has been found by Molinuevo and Lush (1964) and Deaton and McGilliard (1964), there must be differences between sires in rate of maturity. A lower genetic variance in second lactations as compared to firsts would indicate the change from first to second lactation would have to differ from one sire to another. This would mean either a change in the ranking of the bulls from first to second lactation or, at least a closer grouping.

As discussed previously, it is impossible to make any inferences to mature production from the regressions given in Table 3. For instance, it is impossible to tell

if the daughters of bull 1611, whose regression of actual milk on age is +49.24 pounds, are closer to maturity at first lactation than the daughters of bull 1602, whose corresponding regression is +234.30. The difference in the regressions may indicate the daughters of 1611 approach maturity slower than daughters of 1602. An accurate method of determining the position of these regressions with respect to mature production is not readily apparent and would probably add little value in interpreting whether there are differences between sires in rate of maturity.

The regressions given in Table 3 and other analyses presented in this study would indicate that there are differences between sires in the rate of maturity of their daughters. This would mean that a sire's proof based on first lactations could be biased and the ranking of a group of sires on the basis of second lactations could differ from a ranking on first lactations. If there are differences between sires in rate of maturity, the only records that could be used in a sire's proof that would not cause a bias would be those made when the daughters are mature. This is usually between 6 and 8.5 years of age in Holsteins. It is almost certain that the bias caused by differences in maturity rate would not be large enough to justify waiting until the daughters of a bull are this old before computing a proof on the bull. Waiting for mature records would cause

the generation interval on the sire's side to be at least 1.5 times as long as would be possible from selecting the bulls on the basis of the first lactations of their daughters. The mistakes made in selecting on first lactations would have to be rather large to offset this increase in generation length.

### Genetic Improvement

One of the problems facing animal breeders is how to measure genetic improvement in a population. Several methods of estimating genetic improvement have been proposed but each has its limitations. Some of the methods will be discussed here.

One method of measuring genetic improvement was suggested by Rendel and Robertson (1950). They proposed that genetic improvement could be accomplished in four ways: selection of dams of cows, selection of dams of bulls, selection of sires of cows, and selection of sires of bulls. The evaluation of genetic improvement by this method involves estimating the average genetic superiority of the four types of parents, summing these averages, and dividing by the sum of the four possible generation intervals. This method is useful in estimating the absolute genetic improvement in the over-all population only in populations in which all records can be expressed relative to a similar genetic and

environmental base. This usually restricts the method to a single herd. The genetic statistics used in estimating the average genetic superiority of the parents must be the true population parameters for the estimate of genetic improvement to be correct. Also, selection must have been only for the trait in question. The effect of correlated responses can not readily be taken into account by this method.

Another method of estimating genetic trend has been proposed by Henderson et al. (1959). This method involves the computation of maximum likelihood estimates of genetic groups which are usually separated by time of birth. This method attempts to estimate the genetic trend free of environmental effects by utilizing repeated records on the same cow to estimate year-to-year differences. The exactness of the estimates of genetic trend by this method depends on the assumption that the assumed age-correction factors are correct for the population in question and repeatability is known without error and constant for all records. In dairy data covering a long time period, there has to be an incomplete cross-classification of cows and years. This situation can only complicate the estimation of the environmental and genetic trends from the data.

Another method of estimating genetic improvement in a population involves the comparison of daughters of A.I. sires with those of non-A.I. sires. This method has been

discussed by Van Vleck and Henderson (1961c). The method requires dividing the population into cows sired by artificial insemination and cows sired by non-A.I. sires. The genetic trend in the non-A.I. population and the genetic superiority of the A.I. population relative to the non-A.I. population are estimated by the solution of least squares equations. The over-all genetic improvement of the population is not obtained. It is possible to estimate the change in magnitude of the differences between the two parts of the population across time, so the primary usefulness of the method appears to be in comparing the A.I. population to the non-A.I. population. This method requires reasonable numbers of both A.I. and non-A.I. daughters that are contemporary for any accuracy of estimation. One assumption required in an analysis of this type is that sires are mated to random dams. If the best dams were mated to A.I. sires, the calculated genetic superiority of the A.I. sires would be inflated.

Still another method of estimating genetic improvement in a population was discussed by Smith (1962). This method measures the change in the relative position of the daughters of a bull with respect to their herdmates over time. The regression of the deviation of the herdmate's average from the daughter's average on time would measure one-half of the genetic improvement in the population. This method of

estimating genetic improvement also requires the assumption that sires are mated to a random sample of dams. Accuracy of estimates of genetic improvement computed by this method should be improved as the time a sire is used in the population is lengthened. This method of estimating genetic improvement is easily applied to daughters of A.I. sires and was used in this study.

Each record was expressed as a deviation from its herd-mate average as discussed in an earlier section. The regression of these deviation records on time within each sire would be an estimate of one-half the genetic progress in the population. If the genetic change is positive, the regression will be negative. For this analysis, time was expressed as coded months ranging from 1 for January, 1951 to 144 for December, 1962. Only first lactations were used in this portion of the analysis in an attempt to remove as much of the effect of selection as possible. The resulting regressions of deviation production on time in months for each sire are given in Table 8 along with the range in months of freshening of his daughter's first lactations.

Standard errors of the regressions presented in Table 9 range from 3 to 65 pounds for the regressions involving milk and from 0.1 to 2.3 pounds for the regressions involving butterfat.

The average of the 28 regressions shown in Table 9

Table 9. Within-sire regressions of production on time

Sire	Number of daughters	Regressions		Range in months
		Milk	Butterfat	
1018	144	- 11.49	-0.02	56
1021	327	- 11.73	-0.05	80
1028	464	+ 13.38	+0.54	92
1117	54	+ 64.08	+1.28	36
1118	45	+ 15.70	-0.40	24
1122	88	+ 93.48	+4.77	21
1126	113	+ 8.63	-0.05	18
1601	1299	+ 4.92	+0.13	115
1602	139	+ 17.33	+0.48	92
1603	778	- 4.62	-0.17	86
1605	433	+ 2.88	+0.10	95
1606	344	+ 2.63	+0.49	94
1607	824	+ 2.68	+0.01	53
1608	640	+ 13.69	+0.36	62
1609	569	+ 5.38	+0.25	63
1610	556	- 0.70	-0.00	50
1611	194	- 34.17	-1.30	84
1702	229	+ 6.68	+0.06	77
1717	105	+ 13.64	+1.86	29
1718	74	-104.65	-5.66	17
1720	315	+ 8.36	+0.23	39
1725	536	- 17.98	-0.76	81
1728	309	- 55.07	-1.17	32
1735	325	- 2.61	-0.32	74
1738	56	- 12.00	-1.04	23
1741	322	- 21.88	-0.66	52
1746	140	- 34.35	-1.41	80
1747	132	+ 15.66	+0.25	33

weighted by number of daughters, gives an estimated monthly genetic improvement of 2.5 pounds of milk and 0.06 pounds of butterfat. The estimated yearly genetic improvement would be 30 pounds of milk and 0.72 pounds of butterfat.

The genetic improvement measured by this method is the improvement actually accomplished in the female population

if the assumption of sires being mated to a random sample of dams holds in the population. This genetic improvement could be from either direct selection or a correlated response. The calculated rate of genetic improvement is an average improvement for the period in which the data were collected and may or may not be the present rate of genetic improvement in the population.

Analyses of variance were computed to test the homogeneity of the regressions given in Table 9. The results are given in Table 10. The F ratios are both significant at the 0.01 level of probability.

One type of selection among sires could cause a bias in the genetic improvement estimated by the method used here. If sires were selected on the basis of the records of their first few daughters in the data, a bias could exist in the calculated genetic improvement. Any positive errors involved in the production of these first daughters would inflate the estimate of the rate of genetic improvement in the population based on these sires. If this type of selection were practiced, the daughters used in making the selections should be removed from the data used in calculating the rate of genetic improvement. The bulls selected by A.B.S. for use in their stud were picked on the basis of several traits, one of which was a proof based on a few daughters by natural service, usually in one herd. Little, if any, selection was practiced

Table 10. Analyses of variance to test homogeneity of regressions of production on time

Source	Degrees freedom	Sums of squares	Mean squares	F ratio
Milk				
Total within sire	9535	51,299,570,821		
Reduction due to fitting $b_1$	28	305,974,305		
Reduction due to fitting $b$	1	1,288,785		
Additional reduction	27	304,745,520	11,286,871	2.10
Error	9507	50,993,596,516	5,363,794	
Butterfat				
Total within sire	9535	67,261,980		
Reduction due to fitting $b_1$	28	461,057		
Reduction due to fitting $b$	1	3,218		
Additional reduction	27	457,839	16,957	2.41
Error	9507	66,800,923	7,026	

among the bulls on the basis of the production of their progeny in these data. Any factor, such as selecting on natural proofs, which affects the genetic worth of the

individuals in the population would be contributing to the genetic improvement in the population and should be measured. It is the non-genetic error affecting the records which can bias the estimated rate of genetic improvement. So even though these 28 sires are a selected group, the method by which they were selected should not affect the estimate of genetic improvement.

Another factor which could cause the regression of daughter's deviation production on time to be biased would be a change in the average breeding value of the mates of the bull. If a bull is mated to genetically better cows as he is progressively used over time in the population, the deviation of his daughters will tend to increase in a positive direction. This would be a form of non-random mating. A.B.S. distributed liquid semen to its technicians until September, 1958 when it started using frozen semen. Selection of bulls to breed specific cows was rather difficult while liquid semen was being used because a technician did not carry semen from all the bulls all the time. The regression of the deviation production of the mates of a sire on time would indicate if there has been a change in the average merit of the mates of the sire.

The relation between the regression of daughter's deviation production on time and the regression of dam's deviation production on time could be used to partially

correct for non-random mating of the sires. This relation, which can be expressed as a regression, would correct for any non-random mating common to all sires. An example of this might be a tendency to mate all new sires in the stud to poorer cows. By letting:

$D_{ij}$  = the deviation of the first lactation of the  $j^{\text{th}}$  daughter of the  $i^{\text{th}}$  sire from her herdmate average,

$T_{ij}$  = the coded month of freshening of the  $ij^{\text{th}}$  daughter, and

$D'_{ij}$  and  $T'_{ij}$  are the same for the dam of the  $ij^{\text{th}}$  daughter,

the within-sire regression of daughter's deviation production on time would be:

$$b_{DT} = \frac{\sum_j (D_{ij} - \bar{D}_i)(T_{ij} - \bar{T}_i)}{\sum_j (T_{ij} - \bar{T}_i)^2}$$

and the regression of the dam's deviation production on time would be:

$$b'_{DT} = \frac{\sum_j (D'_{ij} - \bar{D}'_i)(T'_{ij} - \bar{T}'_i)}{\sum_j (T'_{ij} - \bar{T}'_i)^2}$$

Under the assumption that there are no covariances other than daughter-dam, the covariance of the two regressions would be:

$$\text{Cov}(b, b') = \frac{\sum_j (D_{ij} - \bar{D}_i)(T_{ij} - \bar{T}_i)(D'_{ij} - \bar{D}'_i)(T'_{ij} - \bar{T}'_i)}{[\sum_j (T_{ij} - \bar{T}_i)^2][\sum_j (T'_{ij} - \bar{T}'_i)^2]} .$$

Assuming the dam is known for every daughter and  $T_{ij} - \bar{T}_i$  is random for both daughter and dam with respect to both  $D_{ij} - \bar{D}_i$  and  $D'_{ij} - \bar{D}'_i$ , the covariance of the two regressions will reduce to:

$$\text{Cov}(b, b') = \frac{\sum_j (T_{ij} - \bar{T}_i)(T'_{ij} - \bar{T}'_i) \text{Cov}(\text{daughter}, \text{dam})}{[\sum_j (T_{ij} - \bar{T}_i)^2][\sum_j (T'_{ij} - \bar{T}'_i)^2]} .$$

Considering  $T'_{ij} - \bar{T}'_i$  constant for each  $j$ , the variance of the regression of dam's deviation production on time would be:

$$\text{Var}(b') = \frac{\sum_j (T'_{ij} - \bar{T}'_i)^2 \text{Var}(\text{dams})}{[\sum_j (T'_{ij} - \bar{T}'_i)^2]^2} .$$

By assuming an equal calving age for all dams and equal age at first calving for all cows, the regression,  $b_{bb'}$ , would be:

$$b_{bb'} = \frac{\frac{\sum_j (T_{ij} - \bar{T}_i)^2 \text{Cov}(\text{daughter}, \text{dam})}{\left[ \sum_j (T_{ij} - \bar{T}_i)^2 \right]^2}}{\frac{\sum_j (T_{ij} - \bar{T}_i)^2 \text{Var}(\text{dam})}{\left[ \sum_j (T_{ij} - \bar{T}_i)^2 \right]^2}}$$

which reduces to:

$$b_{bb'} = \frac{\text{Cov}(\text{daughter}, \text{dam})}{\text{Var}(\text{dam})}.$$

In the above derivation it was assumed that there was no correlation among  $T_{ij} - \bar{T}_i$ ,  $T'_{ij} - \bar{T}'_i$ ,  $D_{ij} - \bar{D}_i$ , and  $D'_{ij} - \bar{D}'_i$  for all  $j$ . This would not hold true for  $T_{ij} - \bar{T}_i$  and  $D'_{ij} - \bar{D}'_i$ . As the dam's deviation production for first lactation,  $D'_{ij} - \bar{D}'_i$ , increases there should be an increased chance that she would have a heifer calf at some later calving. This will cause a positive correlation between  $T_{ij} - \bar{T}_i$  and  $D'_{ij} - \bar{D}'_i$ . This would cause the true  $\text{Cov}(b, b')$  to be larger than it would be if this correlation was not present. Any selection among cows for production would force this correlation to be positive.

Other assumptions utilized in the derivation were equal age at first calving for all dams and daughters and equal calving age for all dams. These assumptions were necessary

in order to get the  $(T_{ij} - \bar{T}_i)^2$  term in the numerator of the Cov (b,b') and the  $[\sum_j (T_{ij} - \bar{T}_i)^2]^2$  term in the denominator. Any deviation of the real data from these assumptions will cause more reduction in the denominator of the Cov (b,b') than in the numerator. This will also cause the true Cov(b,b') to be larger than the Cov (b,b') would be under these assumptions.

Evaluation of the  $b_{bb'}$  from the formula derived here would be difficult, if not impossible, unless these assumptions are made. The magnitude of the regression will vary from sire to sire and for each sample of daughters of each sire. As discussed above, the regression will be somewhat larger than the regression of daughter on dam. The simplest way to evaluate the regression would be to calculate the regression using the within-sire regressions as observations. This method of evaluation would not be precise in this study since there are only 28 of these pairs of regressions available. Therefore, no attempt to estimate this regression was made, but the correlation of  $T_{ij} - \bar{T}_i$  and  $D'_{ij} - \bar{D}'_i$  was evaluated for each sire. If enough sires were available for this type of evaluation, an indication of the importance of the above assumptions could be obtained by comparing the computed  $b_{bb'}$  to the regression of daughter on dam computed from the same data.

The within-sire regressions of production of the dam on

time and the within-sire correlations of deviation production of the dam with month of first calving of the daughter are given in Table 11.

The standard errors of the regressions of dam's milk production on time range from 4 pounds to 383 pounds for those involving milk and from 0.17 pounds to 14.31 pounds for those involving butterfat. The estimates of the standard errors were larger than the computed regressions in 43 of the 56 cases. The average regression of dam's production on age, weighted by number, was +0.0047 pounds of milk and -0.0004 pounds of butterfat. These small average regressions would indicate there were no non-random mating systems common to all sires.

Correlations between the within-sire regression of daughter's production on time and regression of dam's production on time were calculated using the values given in Tables 9 and 11. These correlations were +0.02 for the regressions involving milk and +0.09 for the regressions involving butterfat. These very small correlations indicate that the regression of dam's production on time has little effect on the regression of daughter's production on time. Therefore, this possible cause of bias in the estimation of genetic improvement will receive no further consideration.

The correlations given in Table 11 of dam's first lactation production with the month of calving of her

Table 11. Correlations and regressions involving dam's production

Sire	Number	Regression of dam's production on time		Correlation of dam's production with month of calving of the daughter	
		Milk	Butterfat	Milk	Butterfat
1018	56	-11.79	-0.56	+0.21	+0.14
1021	135	- 8.53	-0.31	-0.11	-0.11
1028	267	- 3.66	-0.33	+0.05	+0.02
1117	11	-83.17	-0.93	+0.43	+0.13
1118	9	+69.75	-8.03	+0.16	+0.08
1122	49	+ 7.84	+0.17	-0.07	+0.06
1126	67	+ 1.94	-0.30	+0.26	+0.17
1601	523	- 1.79	+0.02	+0.06	+0.08
1602	36	+ 4.47	+0.07	+0.12	+0.22
1603	296	- 6.11	-0.42	+0.00	-0.05
1605	165	+23.65	+1.00	+0.13	+0.19
1606	133	- 4.18	+0.21	+0.15	+0.15
1607	406	- 4.52	-0.15	-0.01	-0.02
1608	330	+10.35	+0.27	+0.04	+0.04
1609	277	- 0.46	-0.07	+0.01	-0.03
1610	314	+13.96	+0.23	+0.09	+0.07
1611	97	+ 8.61	+0.42	+0.12	+0.17
1702	111	+11.55	+0.23	+0.01	-0.02
1717	60	- 1.62	-0.06	+0.16	+0.31
1718	44	-37.39	-0.70	+0.11	+0.06
1720	174	+ 6.74	+0.26	-0.03	-0.07
1725	316	- 9.46	-0.52	-0.00	-0.03
1728	189	+ 0.42	+0.05	-0.04	-0.01
1735	200	- 5.43	-0.01	-0.06	-0.03
1738	37	-12.88	-0.61	+0.23	+0.26
1741	210	+ 0.65	-0.10	+0.01	+0.03
1746	82	+13.48	+0.29	+0.06	-0.01
1747	76	- 1.61	-0.07	-0.03	+0.08

daughter are all small and most are positive. The average of these within-sire correlations weighted by number, was +0.034 for milk production and +0.032 for butterfat production. These tend to support the positive correlation

previously theorized in this section.

Results presented in this section indicate that there has been about 30 pounds genetic improvement per year in milk production and 1 pound genetic improvement per year in butterfat production. Expressed as a percent of the mean production, as given in Table 6, the genetic improvement made per year was 0.26% of the mean for milk production and 0.24% of the mean for butterfat production. These figures are not as large as an animal breeder would hope, but they do show that there has been some genetic improvement in the population. Most bulls in the A.B.S. stud were selected on the basis of a natural service proof and other traits including type. This rate of genetic improvement should be increased if A.B.S. starts selecting its bulls from a young sire sampling program rather than on the basis of a natural proof.

Most of this genetic improvement must have been due to the selection of bulls since culling cows creates at best a small selection differential. The average deviation production of the daughters of the 28 sires involved in these analyses was +346 pounds of milk and +19 pounds of butterfat. The minimum age a sire could be purchased on the basis of a natural proof would be about 5 years. Another year would elapse before the first daughters were born in these data. The 28 sires in Table 9 had daughters freshening over an average range of 72 months. Assuming the frequencies of

freshenings were symmetrical across this range and the average age at first calving was 30 months, an estimate of the generation interval for these sires would be 9 years. This would mean the yearly contribution from the selection of sires would be about 38 pounds of milk and 2 pounds of butterfat. This would tend to support the hypothesis that the estimated genetic improvement came from the selection of sires. However, the similarity of these results is partly automatic since the same daughter's records are used in both analyses.

Arave et al. (1964) used both the maximum likelihood method of Henderson (1959) and the method suggested by Smith (1962) to estimate genetic improvement. The maximum likelihood estimate of genetic improvement in fat-corrected milk was 0.7 percent of the mean per year. The pooled intra-sire regression estimate of genetic improvement in the largest herd in the study was 56 pounds of fat-corrected milk per year. This is in close agreement with the maximum likelihood estimate from this same herd of 60 pounds per year. This study involved 12 Jersey herds which were used to progeny test bulls from the inbred Jersey lines of the University of California herd. The method by which these bulls were selected and the fact that they were not Holsteins may affect the comparison of these results with those found in the present study.

Walton (1961) estimated genetic improvement by the method

suggested by Rendel and Robertson (1950) and the method suggested by Henderson (1959). The estimated yearly genetic improvement using Rendel and Robertson's method was 0.59 percent of the mean for milk and 0.58 percent of the mean for butterfat. The corresponding maximum likelihood estimates were 1.36 percent of the mean for butterfat and 1.30 percent of the mean for milk production. Data used in this study came from the Iowa State University Holstein herd in which a program of selecting for production was followed as closely as practically possible. Under these conditions, the genetic improvement should be expected to be larger than would be found in field data. Assuming different values for repeatability and age-correction factors caused much variation in the estimates of genetic improvement from the maximum likelihood method. This herd was closed to outside breeding about 1930 and made little use of progeny testing. Therefore, the estimates obtained by Rendel and Robertson's method are probably more accurate than the estimates obtained from the maximum likelihood method.

Van Vleck and Henderson (1961c) estimated the annual genetic improvement in New York D.H.I.A. herds to be about 2 pounds of butterfat and 45 pounds of milk among natural service progeny. The corresponding rates of improvement among A.I. progeny were about 3 pounds of butterfat and 55 pounds of milk. Both of these rates of genetic improvement

are higher than the ones found in the present study. The reasons for the differences between these results and the ones found in this study are not obvious but may be due to the difference in methods.

Van Vleck and Henderson (1961b) used the estimated genetic improvement they had found in New York D.H.I.A. data (Van Vleck and Henderson 1961c) to study the effect of genetic improvement on the ranking of bulls. They concluded that genetic trend has little effect in the evaluation of contemporary sires. They did conclude that genetic trend probably should be considered if the sires to be compared were used in widely different time periods.

The maximum rate of genetic improvement expected with progeny testing and artificial insemination has been reported to be about 1.7 percent of the mean per year by Specht and McGilliard (1961). The genetic improvement found in this study was well below the maximum possible improvement as would be expected where selection of sires was based on a natural service proof and was lower than the rate of improvement found in the three studies cited here. Two studies reviewed here (Arave et al., 1964; Walton, 1961) involved herds which were under University supervision while the herds in this study were operated as practical dairy farms. Herds under supervision by a University probably place more emphasis on production in making selections than would be

expected under field conditions.

With the use of frozen semen and the ease of storing frozen semen, it is possible to visualize a time when even 30 pounds of genetic improvement per year could affect selection among bulls. To correctly select among bulls separated by ten to fifteen years, some credit must be given for genetic improvement. The method of estimating genetic improvement used in this study should measure only the additive genetic improvement made in the population. If there has been 30 pounds additive genetic improvement per year in the population, the proof of the sires in the early years would not be comparable to the proof of the later bulls. If the true rate of genetic improvement were known, it should be added to the proofs of the later bulls to make them comparable to the proofs which could have been expected for these bulls if they had been proven in the original population. However, the use of estimates which are inaccurate, to correct for this genetic improvement may introduce more error than making no correction at all. Only estimates of genetic improvement which are relatively accurate should be used to correct for the genetic improvement. At the present time, there is little need to correct for genetic improvement when comparing between bulls, since there is usually not a long period separating the bulls.

### Breeding Value of Mates of Sires

One factor which can bias a bull's proof is the average genetic merit of the cows to which he is mated. If a sire is mated to above-average cows, the progeny would be expected to be above-average on the merit of their dams alone. This bias could influence progeny tests of any size, but would usually be expected to exert more influence when the number in the progeny test is small. The effect of dams is one reason single herd proofs are not accepted as being as accurate as multi-herd proofs.

Analyses of variance were computed using the model:

$$Y_{ij} = \mu + S_i + e_{ij}$$

where  $Y_{ij}$  is the deviated production of the  $j^{\text{th}}$  dam mated to the  $i^{\text{th}}$  sire. Only production during first lactation was used in these analyses. All records were expressed as deviations from their herdmate's average, as discussed earlier. Separate analyses were run for the dams of the first ten, twenty and forty daughters of a sire, ordered by calving date of the daughter, who had first lactation records available for their dams. If a dam had more than one daughter, her record was repeated.

The resulting analyses of variance for milk and butterfat production are given in Table 12. In the analyses involving  $n = 10$ , the F ratio for butterfat is significant

with  $P < .01$  and the F ratio for milk is significant with  $P < .05$ . All other F ratios are non-significant.

The average production of the three groups of mates of each sire are given in Table 13.

The non-significant difference between sires in the analyses in Table 12 does not mean that 20 or 40 dams will not affect the average production of the daughters of a sire. The regression of daughter's production on dam's production is the same regardless of number. Even at  $n = 40$ , the range

Table 12. Analyses of variance of mates of sires

Source	Degrees freedom	Milk		Butterfat	
		Mean square	F ratio	Mean square	F ratio
n = 10					
Total	239				
Sires	23	8,441,995	1.75	14,205	2.19
Error	216	4,801,736		6,484	
n = 20					
Total	479				
Sires	23	5,257,230	<1.0	8,025	1.16
Error	456	5,595,017		6,889	
n = 40					
Total	959				
Sires	23	5,594,850	1.05	7,308	1.06
Error	936	5,317,379		6,885	

Table 13. Average production of the mates of sires

Sire	Milk			Butterfat		
	n = 10	n = 20	n = 40	n = 10	n = 20	n = 40
1018	-1043	-389	+ 93	-26	- 7	+ 3
1021	- 298	+ 89	+325	- 1	+ 3	+13
1028	+ 739	+513	- 58	+29	+20	+ 4
1122	+ 922	+541	+771	+17	+ 8	+22
1126	- 588	-614	-264	- 1	- 7	+ 3
1601	-2015	-864	-519	-76	-38	-16
1603	- 205	+ 90	-198	- 3	+10	+ 2
1605	- 20	+157	- 31	-10	+ 1	- 8
1606	-1347	-984	-561	-47	-27	-15
1607	+ 286	+547	- 35	+40	+36	+ 2
1608	- 802	+327	-293	-21	- 2	- 7
1609	+ 889	-102	-152	+44	+10	+ 9
1610	+ 505	+385	+146	+35	+31	+22
1611	- 502	+117	+127	-32	- 8	- 2
1702	+1048	+ 88	- 41	+42	+ 5	- 4
1717	-1073	-203	-182	-32	-18	- 7
1718	+ 388	-206	+498	+28	- 3	+25
1720	- 73	-262	- 65	+10	+ 2	+ 9
1725	+1891	+958	+917	+74	+33	+38
1728	+1148	+428	+526	+64	+36	+25
1735	- 927	-733	- 85	-32	-21	+ 2
1741	- 538	-133	-127	0	- 2	+ 2
1746	+ 314	-809	-400	+42	0	+ 6
1747	- 330	-492	+297	-22	-23	+ 8

in the average production of the mates of these 24 sires is 1478 pounds of milk and 54 pounds of butterfat. As stated previously, the analysis of variance tests whether the average differences between sires are significant and gives little indication of differences between individual sires.

Bereskin (1963) in his study of Iowa D.H.I.A. records computed the genetic gain expected from including dams in a

sire's proof, relative to not including them, to be 1.007 for milk production and 1.025 for butterfat production. These analyses involved 351 sires with an average of 8.9 mates per sire. These values indicate that, on the average, little would be gained by including the effect of the dams in the proof of a sire.

Lush (1944) found that including dams with their daughters in the selection of sires should cause from 1.12 to 1.20 times as much progress as choosing sires on the average of their daughters alone. As Lush stated, some of this 1.12 to 1.20 was due to partially correcting for between-herd differences. The use of records expressed as deviations from herd-mate averages will lower this gain of improvement.

Van Vleck, et al. (1962) computed an index for ranking sires on the basis of both their daughter's average and the average of their dam's records. The index could be written:

$$I = b_1 (\bar{Y}_o - \bar{X}_o) + b_2 (\bar{Y}_d - \bar{X}_d)$$

where  $\bar{Y}_o$  is the average of the daughters,  $\bar{X}_o$  is the average of their herdmates,  $\bar{Y}_d$  is the average of the dams of the daughters, and  $\bar{X}_d$  is the average of the herdmates of the dams. The value the authors obtained for  $b_1$  was  $\frac{n}{n + 14.76}$  which was very similar to the  $\frac{n}{n + 15}$  obtained when only the daughter's average was considered. The value obtained for  $b_2$  was  $-\frac{n}{8n + 118.08}$  which approaches the regression of

daughter on dam,  $\frac{1}{8}$ , when  $n$  approaches  $\infty$ . When  $n$  is small, dams contribute very little to the index. The similarity of  $b_1 = \frac{n}{n + 14.76}$  and the coefficient  $\frac{n}{n + 15}$  used when only the daughter's records are considered, indicates the small contribution of the correction for dams to the variance of the daughter's deviation corrected for dam's production.

Using  $b_2 = -\frac{n}{8n + 118.08}$ , the difference between the averages of the extreme groups of 40 dams from these 24 sires could be expected to contribute about 130 pounds of milk to the differences in the proofs of these two bulls. If a stud was selecting one of these two bulls to keep in service, this contribution could lead to the selection of the bull with the poorer genetic worth.

In practical situations, one factor which will inhibit or at least reduce the effectiveness of correcting for dams, is incomplete knowledge of the dams of the daughters. In this study the 28 sires used in the genetic improvement analyses had 9563 daughters with first lactations. Only 4670 of these daughters had dams with first lactations in the data. If these data had been collected over a longer period of time, there probably would have been a higher percentage of dam's records available. However, the normal turnover of herds in D.H.I.A. testing programs would limit the possible gain from the inclusion of dam's records in most situations.

The correlation between the production of the dams of the first ten daughters of a sire and the production of the dams of his second ten daughters was  $-0.25$  for milk production and  $-0.13$  for butterfat production. There is no obvious biological reason why these correlations should be negative but they would mean that for the average of these sires the correction for the dams of the first 10 daughters would be in the opposite direction to the correction for the dams of the second 10 daughters. This indicates that some caution should be exercised in making corrections for dams when complete dam information is not available. Theoretically, a correction should be made for dam's records, but in most practical situations, such a correction either shouldn't or couldn't be used.

There are other types of non-random mating which can bias a sire's proof. For example, what proportion of herdmate averages are computed from daughters of one sire? The same type of bias in a sire's proof could result from non-randomness of dams of herdmates as from dams of daughters. Either of these two possible sources could bias a sire's proof. Because of the relative numbers involved, non-randomness of herdmates would be expected to have less effect on the sire's proof than non-randomness of dams. A thorough investigation of the parentage of herdmates would require very complete records collected over a long period of time

since parental identification of herdmates would be more of a problem than parental identification of daughters. An investigation of the parentage of herdmates would be rather difficult and probably of little value, particularly for the evaluation of A.I. sires.

## SUMMARY

The effects of three factors which could bias a sire's proof were investigated in this study. These factors were: (1) differences between sires in the rate of maturity of their daughters, (2) genetic improvement in the population, and (3) differences between sires in the worth of their mates.

The within sire regression of actual production during first lactation on age in months was used as an indicator of rate of maturity for a sire's daughters. Only cows calving between the ages of 23 and 35 months of age and sires with more than 44 daughters were considered in these analyses. The regression of actual production on age for each of 28 sires ranged from -9.20 to +258.16 pounds of milk per month of age and from -3.73 to +11.23 pounds of butterfat per month of age. There were significant differences between sires in the regression of actual production on age in months for both milk and butterfat. There were also large differences between herds and years for this measure of rate of maturity when the regressions were calculated within herds and within years disregarding sires. The effect of herds on rate of maturity was estimated to be from two to three times as important as the effect of sires for milk production. Because of incomplete cross classification of sires, herds, and years, some confounding could exist which would make

interpretations questionable. It would appear, however, that there are some differences between sires in the maturity of their daughters. If there are differences between sires in the rate of maturity of their daughters, the only records which would not bias the proof of a sire would be those made when the daughters are mature. However, waiting for mature records would result in the lengthening of the generation interval of sires by a factor of at least 1.5. It is doubtful that the increase in accuracy by selecting on mature records rather than first records would offset this increase in generation interval and result in an increase in genetic improvement per year.

The estimate of the genetic improvement per year in these data was 30 pounds of milk and 1 pound of butterfat. The method used to estimate the rate of genetic improvement was discussed by Smith (1962). This is far less genetic gain than appears to be possible when selection is for a single trait but it is in the positive direction. This rate of improvement expressed as a percent of the mean production was 0.26% for milk production and 0.24% for butterfat production. Sire proofs are usually based on daughter-herdmate comparisons which do not correct for any genetic improvement in the population over time. If selections are to be made between bulls on the basis of daughter-herdmate comparisons, some correction should be made for any genetic improvement

which has been made in the time interval separating the two bulls. However, comparisons are not usually made between bulls separated by long periods of time, so genetic improvement of the magnitude found in this study would have very little effect.

Analyses of variance were computed using the production of the dams of the first ten, twenty and forty daughters of a sire to determine if there were differences between sires in the genetic worth of their mates. All records used in these analyses were expressed as deviations from herdmate averages. There was a significant difference between sires in the production of their first ten mates for both milk and butterfat. The differences between sires were not significant for the first twenty or forty mates. The difference in milk production between the two extreme groups of 40 mates would have been expected to contribute about 130 pounds of milk to the difference in the proofs of these two bulls. This would indicate it should be possible to make more accurate selections between bulls if their proofs are corrected for their mate's production, even with 40 daughters available. However, production information of the dams of all daughters is seldom available. The fraction of daughters which should have dam's records available before a correction for dams would be expected to give reasonable results is not known. With less than 50 daughters, probably between two-thirds and

three-fourths of them should have production of their dams available before a correction should be made. As the number of daughters increases, the fraction needed to make a reasonable correction should decrease, along with the necessity of making any correction for dam's production at all. It is probably safe to conclude that in most practical situations no attempt should be made to correct for mates of sires in computing the proofs of A.I. sires.

In conclusion, each of the sources of possible bias in sire proofs investigated in this study appear to be real. The need and/or practicality of making corrections for any of these effects does not seem warranted under present conditions of proving bulls used in artificial insemination. The general acceptance of the ability of the daughter-herdmate comparison to detect genetic differences between bulls would indicate the magnitude of possible biases in sire proofs must be relatively small in comparison to the gross genetic differences between bulls. However, in the future the error caused by the three sources investigated in this study, may be relatively large enough that corrections should be made for them.

## BIBLIOGRAPHY

- Agricultural Research Service  
 1964 D.H.I.A. sire summary list. United States  
 Department of Agriculture Publication ARS-44-149.
- Arave, C. W., R. C. Laben and S. W. Mead  
 1964 Measurement of genetic change in twelve Cali-  
 fornia dairy herds. Journal of Dairy Science  
 47:278-283.
- Barr, G. R.  
 1962 Selecting young dairy bulls on differences  
 between relatives and their contemporaries.  
 Unpublished Ph.D. thesis. Ames, Iowa, Library,  
 Iowa State University of Science and Technology.
- Bereskin, B.  
 1963 Effects of genetic and environmental variance on  
 dairy sire evaluation. Unpublished Ph.D. thesis.  
 Ames, Iowa, Library, Iowa State University of  
 Science and Technology.
- Cochran, W. G.  
 1947 Some consequences when the assumptions for the  
 analysis of variance are not satisfied. Bio-  
 metrics 3:22-38.
- Deaton, O. W., and L. D. McGilliard  
 1964 First, second, and third records of a cow to  
 estimate superiority of her daughters. Journal  
 of Dairy Science 47:1004-1006.
- Flock, D. K.  
 1964 Selection among Holstein-Friesian cows. Unpub-  
 lished Ph.D. thesis. Ames, Iowa, Library, Iowa  
 State University of Science and Technology.
- Heidhues, T., L. D. Van Vleck and C. R. Henderson  
 1961 Actual and expected accuracy of sire proofs  
 under the New York system of sampling bulls.  
 Zeitschrift für Tierzucht und Zuchtungsbiologie  
 75:323-330.

- Henderson, C. R.  
1956 Cornell research on methods of selecting dairy sires. New Zealand Society of Animal Production Proceedings 16:69-74.
- Henderson, C. R., O. Kempthorne, S. R. Searle and C. M. Von Krosigk  
1959 The estimation of environmental and genetic trends from records subject to culling. Biometrics 15:192-218.
- Hickman, C. G. and C. R. Henderson  
1955 Components of the relationship between level of production and rate of maturity in dairy cattle. Journal of Dairy Science 38:883-890.
- Kendrick, J. F.  
1955 Standardizing dairy herd improvement association records in proving sires. U. S. Department of Agriculture Agricultural Research Service Publication 52-1.
- Lush, J. L.  
1944 The optimum emphasis on dams' records when proving dairy sires. Journal of Dairy Science 27:937-951.
- Molinuevo, H. A. and J. L. Lush  
1964 Reliability of first, second, and third records for estimating the breeding value of dairy cows. Journal of Dairy Science 47:890-893.
- Rendel, J. M. and A. Robertson  
1950 Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. Journal of Genetics 50:1-8.
- Robertson, A. and S. S. Khishin  
1958 The effect of selection for heifer milk yield on the production level of mature cows. Journal of Agricultural Science 50:12-16.
- Robertson, A. and J. M. Rendel  
1950 The use of progeny testing with artificial insemination in dairy cattle. Journal of Genetics 50:21-31.
- Smith, C.  
1962 Estimation of genetic change in farm livestock using field records. Animal Production 4:239-251.

- Specht, L. W. and L. D. McGilliard  
1960 Rates of improvement by progeny testing in dairy herds of various sizes. *Journal of Dairy Science* 43:63-75.
- Van Vleck, L. D.  
1963 Genotype and environment in sire evaluation. *Journal of Dairy Science* 46:983-987.
- Van Vleck, L. D., T. Heidhues and C. R. Henderson  
1961a Analysis of deviations of dairy records from different contemporary averages. *Journal of Dairy Science* 44:269-281.
- Van Vleck, L. D. and C. R. Henderson  
1961b Effect of genetic trend on sire evaluation. *Journal of Dairy Science* 44:1877-1880.
- Van Vleck, L. D. and C. R. Henderson  
1961c Measurement of genetic trend. *Journal of Dairy Science* 44:1705-1710.
- Van Vleck, L. D., C. R. Henderson and H. W. Carter  
1962 Evaluation of sires available through planned mating. *Journal of Animal Science* 21:30-32.
- Van Vleck, L. D., L. H. Wadell and C. R. Henderson  
1961d Components of variance associated with milk and fat records of artificially sired Holstein daughters. *Journal of Animal Science* 20:812-816.
- Walton, R. E.  
1961 Results of selection for production in a Holstein herd. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.
- Williams, E. J.  
1959. Regression analysis. 1st ed. New York, Wiley and Sons.

## ACKNOWLEDGMENTS

I wish to extend my sincere appreciation to Dr. A. E. Freeman for his counsel and assistance throughout my graduate study and in the preparation of this manuscript. Thanks are also due to the Animal Breeding graduate students for many helpful hints and ideas presented during the course of this study.

I also want to thank Mr. Harlan R. Koch and Dr. Robert E. Walton of American Breeder's Service, Inc. for making the data available for use in this study and for giving much sound advice on the description and interpretation of the data.