WEIGHTING INFORMATION FROM RELATIVES TO SELECT FOR MILK IN HOLSTEINS

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
Oliver Wendell Deaton
1964

This is to certify that the

thesis entitled

WEIGHTING INFORMATION FROM RELATIVES TO
SELECT FOR MILK IN HOLSTEINS

presented by

OLIVER WENDELL DEATON

has been accepted towards fulfillment of the requirements for

Ph.D. **degree in** Dairy

Lon D. McGilliard
Major professor

Date June 3, 1963

O-169

LIBRARY
Michigan State
University

WEIG

WEIGHTING INFORMATION FROM RELATIVES TO SELECT FOR MILK IN HOLSTEINS

By

Oliver Wendell Deaton

AN ABSTRACT OF A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Dairy

1964

WE

Sc

informatio

ulations of

F

average w

use in dev

were use

average u

correlati

were 0. 1

Holstein

records

ficients

should b

iority of

ABSTRACT

WEIGHTING INFORMATION FROM RELATIVES TO SELECT FOR MILK IN HOLSTEINS

by Oliver Wendell Deaton

Selection indexes for milk production in Holstein cattle using information from close relatives were developed and tested in various populations of cows recorded in Michigan DHIA.

Records of lactations measured as deviations from the annual herd average were used to choose the appropriate measure of milk production to use in developing a selection index. Linear multiple regression equations were used to predict the daughter's deviation in first lactation from herd average using various records of the cow as independent variables. The simple correlations of the cow's first record with the first record of the daughter were 0.149 for 904 Guernsey cows and their daughters, and 0.256 for 1,526 Holstein cows and their daughters. The correlations of the cow's later records were much smaller in both breeds. The partial regression coefficients indicated that nearly all of the emphasis among records of the cow should be placed on the cow's first record to predict the superiority or inferiority of the first record of the daughter. Multiple correlation coefficients

indicated tha
the cow were
cow's first r

Sel
breeding val
using 7,63
combination
heritability
sisters on t
variances)
derived fro

T
creased th
own recor
curacy of
the cow's
of informa

or offspri
0.55 for
sired by a
grandpar

indicated that averages of either the first two or the first three records of the cow were poorer predictors of the daughter's first record than was the cow's first record alone.

Selection indexes to predict with maximum accuracy the general breeding value of individual Holsteins for milk production were developed using 7,638 deviations of first lactations from herd averages in a variety of combinations of the cow, her dam, her daughters, and her half-sisters. The heritability used was 0.246 which was derived from the regression of paternal sisters on the cow. Other estimates of heritability (with larger sampling variances) ranged from 0.123, derived from intra-sire correlation, to 0.436, derived from intra-dam correlation.

The records of a cow's dam and maternal sisters only slightly increased the accuracy of estimating her genotype providing the cow had an own record. Daughters and paternal sisters added considerably to the accuracy of estimating her genotype. The multiple correlation of the index with the cow's genotype ranged from 0.50 to 0.73 depending on the kinds and amounts of information available from relatives.

Multiple correlation coefficients for individuals without an own record or offspring (heifers and young bulls) varied from 0.12 for one half-sister to 0.55 for many relatives. In estimating the genotype of a young bull that is sired by a well proven sire, the usefulness of information on the maternal grandparents is limited to the dam's paternal sisters if their numbers are

sufficient. T
no apparent v

Sele
own record i
The first rec
index and als
0.166 with th
accuracy ne

Th
progress fo
sires.

sufficient. The maternal granddam and the dam's maternal sisters are of no apparent value.

Selection by index was compared with mass selection on the cow's own record in a test population of 429 Holstein cows and their 498 daughters. The first record of an unselected daughter was correlated with the cow's index and also with the cow's own first record. The resulting correlations of 0.166 with the index and 0.140 for the cow's record indicated an increase in accuracy near 19 per cent in favor of index selection.

The index appeared to be a practical method to increase genetic progress for milk production especially to select potential dams of future sires.

WEI

WEIGHTING INFORMATION FROM RELATIVES TO SELECT FOR MILK IN HOLSTEINS

By

Oliver Wendell Deaton

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Dairy

1964

T

D. McGil

period of

greatly a

assistan

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Dr. Lon D. McGilliard for his guidance and friendly counsel throughout the entire period of study.

The advice and moral support of Dr. Clint Meadows was also greatly appreciated.

Thanks are also due to A. J. Thelan and his associates for their assistance in processing the data.

TABLE OF CONTENTS

LIST OF TABLES

Figure

1. R

2. C

3. G

LIST OF ILLUSTRATIONS

the "de
primar
an imp
allowi
This s
omic
for br
val se
dairy

be es
obser
tests
cura
time
crea

CHAPTER I

INTRODUCTION

To improve a population genetically is to increase the frequency of the "desirable" genes or gene combinations. This change is accomplished primarily by selection although in some cases the system of mating can play an important role. Voluntary selection involves ranking individuals and allowing them to reproduce at rates proportional to their genetic worths. This selection is limited by many forces such as natural selection and economic considerations, forces causing losses of individuals that would be kept for breeding purposes. A low reproductive rate and a long generation interval set a low limit on annual genetic changes possible in a population such as dairy cattle.

An animal's breeding value for traits with small heritabilities can be estimated most accurately if aids to mass selection are utilized. Repeated observations, information on ancestors and collateral relatives, and progeny tests are such aids. The breeder must compromise between additional accuracy in evaluating his animal's breeding values and a shorter interval of time between generations. Yearly genetic progress actually may be increased by using less accurate information earlier if by this the generation interval

construc
merit.
designe
an obje
to mai
breede
of an in
(c) a c

genetic
situati
non-no
param

dairy
progra
fying b
betwe
and ph
used t

interval can be shortened sufficiently.

The concept of combining information in a selection index is to construct a number to be proportional to an animal's breeding worth or net merit. Such numbers are used to rank individuals for selection and are designed to maximize genetic progress. An index helps to make selection an objective process. Merely by reducing subjective judgment and by helping to maintain consistent goals, an index can be a valuable asset to the dairy breeder. Indexes can be developed from information on (a) several traits of an individual, (b) the same trait on an individual and its relatives, or (c) a combination of the two.

Even when all voluntary selection is based on the index, maximum genetic progress is limited to how close the model of the index fits the real situation. Deviations from the linear model, non-additive gene action, and non-normality of the data as well as inaccurate estimation of the population parameters usually limit genetic progress.

Numerous selection indexes for improving the productive traits of dairy cattle have been proposed, but their application in dairy cattle breeding programs has been limited. The indexes have been constructed with simplifying but untested assumptions. The assumption that no relationship exists between the sire and his mates is frequently made. Seldom have the genetic and phenotypic correlations among the various sub-groups of the populations used to construct the index actually been calculated. The environmental correlatio

or inferre

from smal

I

small, as

tions are

other que

for selec

various

close r

formati

tions be

theory

daughte

correlations existing among the various groups of cows are frequently assumed or inferred from other studies to avoid the large sampling errors resulting from small numbers in the population available for study.

In large populations of dairy cattle where sampling errors can be small, assumptions involving the mating system and environmental correlations are the most obvious areas requiring further investigation. However, other questions involving the validity and applicability of a specific index for selection for production of milk in dairy cattle may concern:

1. What measure of milk production should be used in the index? Has sufficient account of the variation, reliability and interrelationships among the different lactations been made?
2. Is the index substantially more accurate than simpler selection methods?
3. Are the results realized in a cow population close to the theoretical predictions?

One object of this investigation was to ascertain weighting factors for various combinations of information on the milk production of a cow and her close relatives. A second objective was to compare the usefulness of information from the cow and the various types of relatives by the correlations between indexes and genotypes. A third objective was to compare theory with results accomplished by correlating cows' indexes with their daughters' production.

This

cattle breeding

the informatio

critical analys

This investigation should add to the knowledge and practice of dairy cattle breeding by suggesting improved criteria for selection, by increasing the information needed to examine breeding theory, and by stimulating more critical analyses relating to animal improvement.

selectio
has been
lined th
farm an
merit.
maximi
of an a

individ
the two
lection
ces bet
bers, a
(b), a
family

CHAPTER II

REVIEW OF LITERATURE

Previous Indexes

The increased efficiency of net merit or total score as a basis for selection as compared to independent culling levels or the tandem method has been demonstrated by Hazel and Lush (1942). Hazel (1943) clearly outlined the theory and genetic basis for developing a selection index to improve farm animals. This paper dealt with several traits which comprised net merit. Path coefficients and multiple regression techniques were used to maximize the linear correlation between the index and the breeding value of an animal.

Lush (1947) investigated the expected consequences of selecting on individuality alone, on the family average, or on an optimum combination of the two. The conditions which favored family selection over individual selection were (a) a large number of individuals per family, (b) large divergences between the environmental and genetic correlations among family members, and (c) low heritabilities of the traits under selection. In the case of (b), a large genetic correlation with a small environmental correlation among family members suggested using the family average in a positive manner to

determine th
large enviro
average neg
says an indi
credit becau
than usual d
based on a c
always to ec
ever, in dai
survival tra
ic correlati

Nu
have been p
widespread
programs a
clude:

1.

2.

3.

4.

determine the breeding value of the individual; whereas, small genetic and large environmental correlations implied the need to consider the family average negatively as an environmental correction. The latter situation says an individual from a family with high merit should be given negative credit because its performance and the family average are likely larger than usual due to favorable environmental conditions. Gains from selection based on a combination of individual and family performance were shown always to equal or to exceed gains based solely on individual selection. However, in dairy cattle low reproductive rates and inbreeding degeneration of survival traits seriously limit developing sizeable families with high generic correlations among members.

Numerous selection indexes for all major classes of farm animals have been proposed, yet their use in cattle breeding programs has not been widespread. The limited application of selection indexes in dairy breeding programs appears to be the result of several interrelated factors which include:

1. Ineffective education of breeders: lack of knowledge of the existence of indexes as well as not knowing how to use indexes.
2. Labor, expense, and records required.
3. The accuracy of selection by index has not been clearly demonstrated to be much more than simpler methods of selection.
4. Reluctance on the part of the breeder to apply indexes. That is,

bre
pro
σ''
sou
ger

Severa
purposes of dis
use informatior
which selection

Indexe
index for intra
cow and her cl
used were: he
repeatability
average of al
this was wei
real produci
assumed zer
be 0.1. The
selection sole
1.2 depending

From

> breeders seem in effect to overestimate heritability of the productive traits and are sometimes distrustful of "figuring" or "pencil pushing." Many breeders seem convinced that only sound judgment (their own) can properly evaluate the total genetic worth of a cow.

Several selection indexes for dairy cattle have been published. For purposes of discussion these are divided into two groups: (1) Indexes which use information from relatives in selection for a single trait. (2) Indexes in which selection is for multiple traits.

Indexes for a Single Trait. — Legates and Lush (1954) developed an index for intraherd selection for fat yield in Jerseys utilizing records of the cow and her close relatives, dam, daughters, and sisters. The statistics used were: heritability, 0.201 (single record basis ignoring effects of year); repeatability, 0.412; correlation between paternal sisters, 0.120. The average of all of the cow's records was the measure of production although this was weighted according to the inverse of its variance to estimate the cow's real producing ability. The genetic relationship of a sire with his mates was assumed zero, and the relationship between mates of a sire was assumed to be 0.1. The ratios of progress from selection on the index as compared to selection solely on the cow's own records were of the magnitude of 1.0 to 1.2 depending on the kinds and amounts of information available.

From Iowa Holstein data, McGilliard (1962) constructed an index for intraherd se

fat. In develop

herd average w

on the cow and

Iowa Board of

Michigan State

Skjerv

computing parti

combinations of

parents) were i

necessary were

Using

index from 18,

information abo

the genetic pot

bulls with their

as accurate for

eight or nine A

Indexe

berry (1955) de

yield alone, and

sidered as net m

for intraherd selection on hundreds of pounds milk corrected to 3.5 per cent fat. In developing this index, deviations of each lactation from the annual herd average were used to estimate the parameters required, and information on the cow and her close relatives was utilized. This index has been used in Iowa Board of Control herds for a number of years and is presently in use at Michigan State University.

Skjervold and Ødegard (1959) presented a correlation matrix for computing partial regression coefficients as index weights. Although all combinations of performance tests and ancestral information (through grandparents) were included, no actual data were used, and the assumptions necessary were numerous.

Using deviations from the herd average, Barr (1962) developed an index from 18,675 Canadian Holstein-Friesian records. This index utilized information about milk production from parents and grandparents in assessing the genetic potential of young bulls. The correlation of the indexes of 28 bulls with their A.I. proofs was 0.32. The index was estimated to be about as accurate for evaluating the bull's breeding value for milk production as eight or nine A.I. daughters.

Indexes Involving Selection for Several Traits. —Tabler and Touchberry (1955) developed several indexes for Jerseys. Milk yield alone, fat yield alone, and a combination of milk yield, fat yield, and type were considered as net merit; and various combinations of these traits in the individual were used to

estimated 10

addition to he

selection for

fat yield. In

decrease in t

were 0.25, 0

yield, fat yie

In a

found less ac

with milk yie

considerably

Jerseys of th

a better crit

yield than wa

0.24 respec

Ha

population u

the value of

equally. Pa

of informati

were used to construct the indexes. The genic value of milk yield could be estimated 10 per cent more accurately if the cow's fat yield considered in addition to her milk yield. Milk yield appeared to be a better criterion of selection for the genetic improvement of production (milk and fat) than did fat yield. Including type along with milk yield resulted in a 15 per cent decrease in the expected genetic gain of milk and fat yield. Heritabilities were 0.25, 0.20, 0.56 and 0.25, respectively, for single records of milk yield, fat yield, per cent fat, and type.

In a later study using Holstein data, Tabler and Touchberry (1959) found less advantage (5.8 per cent) in including fat yield in an index along with milk yield to select for pounds of milk and fat. The Holsteins showed considerably more genetic variability for milk and fat yield than did the Jerseys of the earlier study. Again, the conclusion was that milk yield was a better criterion for genetic improvement of a combination of milk and fat yield than was fat yield alone. Heritability estimates were 0.27, 0.57, and 0.24 respectively, for milk yield, per cent fat, and fat yield.

Harvey and Lush (1952) published two indexes derived from the same population used by Legates and Lush. The first index alloted type one-third the value of fat production, while the second weigh type and production equally. Partial regression coefficients were given for various combinations of information on the cow and her daughter.

Us

The use of

breeding w

research w

a logical m

much of th

Th

breeding v

pointed out

unequal nu

adjusted fo

and fair m

B

as compar

ducing abi

to the rel

accuracy

mation bu

should hol

from Ber

in the ave

Measures of Production

Usefulness of Single and Multiple Records of Lactation Production. — The use of the lifetime average of production records for predicting a cow's breeding worth has become widely accepted and recommended by breeders, research workers, and purebred organizations. Averaging has seemed to be a logical method to reduce the effects of the intangible factors which cause much of the variation in production records.

The theory using the highest record of a cow as an indicator of her breeding value was criticized by Berry and Lush (1939). These workers pointed out the unfairness of using the highest record to compare cows with unequal numbers of records. The average of all records, appropriately adjusted for the variation of the average, was recommended as an accurate and fair means to compare cows.

Berry (1945) studied the reliability of averages of butterfat records as compared to various single records in evaluating a cow's probable producing ability and her breeding worth. The second record added considerably to the reliability of estimating a cow's probable producing ability over the accuracy of a single first lactation, and additional records added new information but at a rapidly decreasing rate. Theoretically, the same principle should hold in estimating a cow's breeding value. Generally, the correlations from Berry's data bear out this contention, but including the second lactation in the average, as well as including those beyond the fourth, actually lowered

the valu
was ind
second
weighte
record
and fif
involv

differ
daugh
real
recor

of di
from
cattl
of a
the
The
fluc
inte
bil

the value of the average for predicting the production of the daughters. This was indicated by the correlations of the cow's records with the first and second records of her daughters. A smaller correlation was obtained from a weighted average of the first two records of the cow than from her first record alone. Also the correlations involving averages including the fourth and fifth records were smaller than the corresponding correlation coefficents involving an average of the first three records.

In a study of 169 Ayrshire sires, Putnam et al. (1943) found that differences in progeny test results using first records or all records of daughters and dams were small and not significant. They suggested that a real saving of labor with no sacrifice in accuracy could be made by using first records only.

Several workers have reported information concerning the reliability of different lactation records for selection. Based on daughter-dam correlations from milk and butterfat yields in thirteen herds of Swedish Red and White cattle, Johansson and Hansson (1940) stated: "Among the first three records of a cow the second lactation yield was found to be the poorest measure of the capacity of production and the first lactation was found to be the best." They felt that the second lactation was more sensitive to environmental fluctuations than were other lactations. The absence of a preceding calving interval or dry period was mentioned as a factor contributing to less variability and more reliability in first records as compared to other lactations.

Johansson (1955) studied the heritability of butterfat yield of the first three lactations in 4,912 daughter-dam pairs of Swedish cattle. The heritability estimates were: first lactation, 0.33; second lactation, 0.10; and third lactation, 0.24. Johansson concluded, "The first lactation record is significantly superior to the second and slightly superior to the third as an indicator of the cow's inherent capacity for yield." Again environmental factors including nutritional and management factors were imputed to cause much of the differences in reliability of the various records.

Rendel et al.(1957) also found the heritability of the first lactation yield to be higher than the heritability of the second. From regression of daughter on dam within sire and herd for six English breeds, the heritability of the second record was 0.24 as compared to 0.43 for the first. The regression of the daughter's first record on the dam's first record was 0.21, whereas the regression of the daughter's first record on the average of four records of the dam was only 0.20.

From a study involving heritabilities, repeatabilities, and correlations of various single records and averages among 8,413 Brown Swiss records from 38 Wisconsin herds, Johnson and Corley (1961) concluded that the first records appeared to be as valuable as any other single record or average of records in selecting for production traits.

Freeman (1960) using deviations from the annual herd average studied the genetic relationship among the first three lactations of Holsteins.

Herita

0.24

relat

fat; t

0.40

cate

fere

hav

on

of

we

br

th

r

Heritabilities for the first, second, and third lactations were: for milk, 0.36, 0.24 and 0.26; for fat, 0.43, 0.35, and 0.26, respectively. The genetic correlations between first and second records were 0.68 for milk and 0.80 for fat; the values for first with third and second with third were of the order of 0.40 for milk and for fat. It was suggested that this could be evidence to indicate that, to some extent, different sets of genes influence production in different lactations.

The unequal reliability and variability of the various lactation records have not been given due consideration in many of the theoretical investigations on selection indexes. If the first lactation record is a more reliable indicator of a cow's breeding value than are later lactations, simpler selection criteria would be available. The reduced time and expense required to assess the breeding values of cows and bulls could be used to considerable advantage by the breeder. If different genes affect the different lactation yields, the breeder may need to re-examine his goals and management practices as well as his selection methods in order to make optimum use of the various records.

Many breeders seem to be quite concerned about the advisability of selecting dairy cows and sires on first lactation performance. That some heifers produce well in their initial lactations and then "burn out" in later lactations seems to be an idea not at all uncommon among breeders. Likewise, the supposition seems prevalent that many of the cows which have long productive lives are those which do not yield well in their first lactations but

tend to be "s
degree, the

Abu
in first lact
amount of r
longevity.
results of p
ing unselec
interrelati
ity is scan

H
tween yiel
second la
and the in
yield and
performa
on the in
that sele
lifetime

basis of
regressi

tend to be "slow-starters." Should these claims be justified even to a moderate degree, the wisdom of selection on first lactations would be debatable.

Abundant evidence of a critical nature about the relation of production in first lactation to later records is lacking. One reason for the limited amount of research on this subject is the confounding effects of selection on longevity. Sound economics dictate that considerable selection be made on the results of performance in the first lactation. Such a situation makes obtaining unselected data impossible. Knowledge of the genetic and physiological interrelationships of early lactation yield, reproductive fitness, and longevity is scanty.

Hickman and Henderson (1955) studied the genetic relationship between yield by the heifer and the increase in production from the first to the second lactation. The genetic correlations between yield in the first lactation and the increase in yield from the first to second lactation were +0.25 for fat yield and -0.04 for milk yield. The authors concluded that sire selection on performance of the daughters in first lactation should have little or no effect on the increase in production with age of the sire's offspring. It appeared that selection on production in first lactation was compatible with increased lifetime production.

Robertson and Khishin (1958) came to similar conclusions on the basis of regressing the increase from first to second lactation (and also the regression from the second to third lactation) on heifer yield. Their sample

was th
Englis

and fa
3,579
corre
ducti
more

prod
Jers
of a
of th
yiel
with
no
old
pro
ide

inv
15

was the offspring of 1,273 sires each, with at least 35 daughters, from five English breeds.

Gaalaas and Plowman (1963) studied the relationships between milk and fat yield in first lactation and longevity in 79 Holstein herds involving 3,879 daughters of 123 sires. Small but highly significant regression and correlation coefficients were obtained between final age in the herd and production in first lactation. These workers concluded that cows producing more in the first lactation had a longer productive life in the herd.

The relationship between production in first lactation and subsequent production levels was studied by Rennie and Bremner (1961) in a Canadian Jersey population. Regression coefficients of the average of all records of a cow on her first record were between 0.47 and 0.62 for various groups of their data. When classified on the basis of mature equivalent butterfat yield in the first lactation, the group averages maintained their same rank, with less spread, throughout all subsequent lactations. These workers found no evidence to substantiate the claim that high production in the two-year-old cow injures future usefulness. On this basis the authors state, "Sire proving programs based on two-year-old records appear to properly identify those sires of superior breeding value for production at all ages."

Additional evidence to support the conclusions reached by previous investigations has been found by Parker (1962). This study involved 18,250 lactations from Ontario Holstein herds. Rather large regression

and c
yield
and t
phen
lacta
resp
to b
yiel
on t

any
lar
sho
whi
bre
the
an
iar
oth

an
of

and correlation coefficients of subsequent lactation yield on first lactation yield were obtained. The regressions were 0.60 for milk and 0.59 for fat, and the corresponding phenotypic correlations were 0.53 and 0.52. The phenotypic correlations between longevity as measured by number of lactations and first lactation yield were 0.34 and 0.33 for milk and fat respectively. These figures were derived for ratios of production records to breed class averages. The general conclusion was that first lactation yield provided a reliable indication of future performance and that selection on the first record was a sound practice.

Removal of Effects of Herd and Year-Seasons. —In data arising from any sizeable number of herds, differences between herds are one of the largest single causes of variation in dairy production records. Table 1 shows a number of estimates of the per cent of the total variation in production which has been attributed to herds. These estimates represent a number of breeds in various locations. Herds generally contribute 25-45 per cent of the total variation in populations of this nature. In the study by Johansson and Hansson only thirteen highly selected herds were studied. This peculiarity may account for their estimate being noticeably smaller than any other values reported.

McGilliard (1952) extensively reviewed and discussed the theoretical and practical aspects of using the herd average in estimating breeding values of dairy cattle. Expressing a cow's production as a deviation from some

TAB

Refe

Bar

Ber

Hick
Hen

Joha
Han

John
Cor

Leg
Lusk

Mill
McC

Plu

Spec

TABLE 1. —Per cent of total variation in production from differences between herds

References		Population Studied	Variation Due to Herds (Per Cent of Total)	
			Milk	Fat
Barr	(1962)	43,498 lactations from Ontario Holstein ROP herds	26	
Bereskin	(1962)	39,000 lactations from Iowa DHIA centrally processed herds	27	27
Hickman and Henderson	(1955)	First and second lactations of 3,912 cows in 1,094 NY DHIA herds	33	
Johansson and Hansson	(1940)	13 Swedish Red and White herds		6-7
Johnson and Corley	(1961)	8,413 HIR records from 38 Wisconsin Brown Swiss herds	25	
Legates and Lush	(1954)	23,330 lactations of 12,405 Jersey HIR cows in 293 herds		39
Miller and McGilliard	(1959)	First lactation DHIA-IBM records from . . . 4,677 Holsteins	35	39
		. . . 1,001 Guernseys	39	43
		. . . 501 Jerseys	33	35
Plum	(1935)	5,860 records of 2,316 cows in Iowa Cow Testing Associations		33
Specht	(1957)	51,656 records of 26,700 Holstein cows in Michigan DHIA	31	33

type of contemporary or herd average has become common in recent years. The use of deviations is an approach toward eliminating from the data of production the large differences between herds as well as portions of the variation between years and seasons.

Expressing a cow's record solely as a deviation from the contemporary or herd-mate average considers all differences between herds to be environmental. Henderson et al. (1954) proposed sire evaluation procedures using deviations with adjustments to account for genetic differences between herds. In correcting a bull's daughters' records for differences between herds, the average production of the daughters was reduced 0.6 of the amount by which the unweighted mean of the contemporary herd average exceeded the average of all herds in the population. This was expressed as: Corrected Daughter Average = Daughter Average - 0.6 (Contemporary Herd Average - Average of all Herds). More recently, VanVleck et al. (1961) described this procedure, but the estimate of the intra-sire regression of daughters' records on adjusted herd-mate average was revised to 0.9 instead of 0.6.

Adjusting records for environmental differences between herds should remove enough of the variation to justify the additional computations required. That is, the adjusted records should reflect enough more accurately true sire differences that the increase in accuracy justifies the computations. The need for such adjustment can be judged somewhat from Table 2, which gives some estimates of the genetic differences between herds in production traits.

Ref

Bru

Her
Ca

Lu
Str

Pi
Lu

Ro
M

R
R

g

e

r

s

t

v

TABLE 2. —Genetic differences between herd averages

Reference		Population Studied	Per Cent of Differences Genetic	
			Milk	Fat
Brumby	(1959)	450 calves involved in herd transfer		10
		40 sets of identical twins split into different herds		10
Henderson and Carter	(1957)	10,292 progeny records of 595 AI sires	18	
Lush and Straus	(1942)	2,142 dau-dam comparisons of 7,850 AI sires proven in Iowa DHIA		6-7
Pirchner and Lush	(1959)	2,903 AI Holstein dau-dam pairs	6.5	6.5
		880 Iowa AI Holstein heifers		10
		1,072 Jersey HIR cows		14
Robertson and McArthur	(1955)	Offspring of 225 British AI bulls	12	16
Robertson and Rendel	(1954)	3,152 heifers of three English breeds	10	

Most of the differences between herds appear to be due to non-genetic causes. Using deviations from herd average assumes all differences between herds are non-genetic. Refinements in adjusting production records for genetic differences between herds appear, at best, to give only slight increases in accuracy at the expense of considerable additional computations. According to Johansson (1961), "Artificial breeding from bull studs will tend to erase the genetic differences between herds, apart from those due

the a

reco

the

cow

devi

tion

Hol

sir

vol

in

twe

Ma

tio

inv

ma

va

av

ju

due to limited herd size (sampling errors)." The inaccuracies arising from the assumption that herds are genetically similar seem to be small.

The most appropriate method of constructing deviations of production records from their herd or contemporary averages is not all clear. Ideally, the deviations of records would reflect only the real differences between cows. The herd average should be constructed in such a manner that the deviated records would be relatively free of environmental sources of variation but would retain most of the genetic differences.

Tucker and Legates (1962) investigated the lactation records of 442 Holstein herds to determine effective methods of using herd mates in dairy sire evaluation. Evidence derived from a quartic regression technique involving variances due to (a) month in which the daughters freshen, (b) month in which the herd mates freshen, and (c) environmental dissimilarities between the months, suggested using two seasons, October through April and May through September. These workers also recommended that first lactations should be compared with first lactations of herd mates. The difficulty involved in obtaining comparisons with sufficient numbers of herd mates often makes these recommendations impractical.

VanVleck et al. (1961) discussed the usefulness of deviations from various contemporary averages for sire evaluation. The contemporary averages compared were (1) regressed adjusted stablemate averages, (2) adjusted stablemate averages, (3) stablemate averages, and (4) herd averages.

The
var
own
dep
in t
vari
rar
var

dev
of f
wer
sea
the
in
adj
Fo
thu
fol
tho
au
th

The desirability of having an unbiased estimator with minimum sampling variance was emphasized. Deviations from averages containing the cow's own record (only the herd average in this group) are biased in a manner which depends on the number of records involved in the average. Small differences in the theoretical and actual variance components were found among the various unbiased contemporary averages compared. Each of the contemporary averages investigated was about as effective as the others in removing variation caused by herd-year-seasons.

Bereskin and Hazel (1962) discussed the effectiveness of using deviations for evaluating sires. Five plans for removing effects of season of freshening on deviations from herd averages were considered. The plans were (1) fixed year-seasons using May-September and October-April as seasons, (2) rolling seasons consisting of five consecutive months centered on the date of freshening, (3) pooled fixed seasons using the seasons as described in (1) but pooled over three years, (4) pooled rolling seasons, and (5) two adjacent fixed seasons which actually made each "season" one year in length. For milk production, the smallest component of variance within sires, and thus the most effective plan for removing effects of season, was plan (2) followed closely by plan (1). Plan (5) was considerably less effective than those plans involving pooling data over different years, plans (3) and (4). The authors cite a previous analysis from the same data wherein 79 per cent of the effects of herd-year-season on milk production records were removed by devia

of obtai

seasons

ness of

calenda

by deviations from regressed adjusted herd-mate averages.

In the practical situation of evaluating individual cows, the necessity of obtaining sufficient numbers in contemporary comparisons favors longer seasons. A compromise between numbers of contemporaries and effectiveness of removing seasonal effects often dictates the use of seasons based on calendar years.

of herds

Comple

length r

earlies

of age f

expres

herd av

mature

of the

herd a

years

lations

separa

a sele

CHAPTER III

SOURCE OF DATA

The data came from four different populations (or sub-populations) of herds tested in Michigan DHIA. All records were 305 day-2X-M. E. Completed records shorter than 305 days were used without adjustment for length regardless of their duration. The first lactation was defined as the earliest recorded lactation which began when the cow was less than 37 months of age for Holsteins or less than 35 months for Guernseys. All records were expressed as deviations from the herd average in ten pounds of milk. The herd average with which each record was compared was the average of the mature equivalent lactations begun during a calendar year by all other cows of the same breed in the herd. A cow's own record was not included in the herd average with which her record was compared. Records from herd-years with less than three lactations were excluded.

The investigation was divided into three phases with different populations of herds and cows represented in each phase.

Phase I. —Populations of Guernsey cows and Holstein cows used separately to determine the appropriate records to consider in constructing a selection index for milk. The first lactations of 904 registered Guernsey

cows
invol
were

regis
reco
1952

the i
usec
invo

Th
les
he

of
us

cows whose dams had at least one of their first five lactations recorded were involved. Twenty-four herds were included in this group, and the records were made during the years 1947 to 1962.

The Holstein group included the records of first lactations of 1,592 registered heifers whose dams had at least one of their first three lactations recorded. These records were made in 77 different herds during the years 1952 to 1962.

Phase II. —The population from which the parameters to construct the index were derived. Only Hoslteins recorded in DHIA-IBM herds were used. Included were the first lactation records from 8,984 cows in 196 herds involving the years from 1954 to 1962. Some averages of this population were:

M. E. milk production	12,981 lbs.
M. E. fat production	473 lbs.
Deviation from herd average	-114 lbs.
Standard deviation	2,121 lbs.
Age at calving	28.5 months

The negative average deviation of milk production apparently resulted from the less selected first lactations of heifers being compared with all cows in the herd which included older cows that had survived repeated selection.

Phase III. —The test population used to judge the practical usefulness of the index. First lactations from 145 Holstein herds on DHIA-IBM were used. All records were made during the period 1954 to 1962. To be included in this

reduce

clude

in this group all cows had to have dams with first lactations. This limitation reduced the test population to 498 cows. No herds in this group were included in the population of Phase II.

dec

abi

vid

the

alt

Ex

th

av

re

to

co

fo

fo

in

CHAPTER IV

METHODS AND RESULTS

Measure of Production

The first major problem in establishing a selection index was to decide what records to use and how to use them to represent the producing ability of each cow. The number of records varying from cow to cow provides different amounts of information with different variances. To combine these necessitates care to insure equitable treatment among cows. Possible alternatives are single records or multiple records weighted in various ways. Examples of single records are the first record, a random single record, or the best or worst record. Ways of combining records might be the lifetime average, the average adjusted to equal variance, or a linear combination of records, each weighted by the information contained. Phase I was designed to explore the optimum combination of the available records of a cow, and, consequently, to determine the proper weighting for combining these records for use as the cow's phenotype.

The first record of any of a cow's daughters was used as a criterion for estimating the optimum weights for the various records of the cow. As in a practical situation, a measure was sought which would predict most

accu
cont
cien
dete
was
dep
no
que
rer

be
we
m
St
pr

fi
th
s

A
t

accurately which cows would have daughters farthest above (or below) the contemporary average. The appropriate weight, partial regression coefficient, to assign to each of a cow's records to accomplish this ranking was determined by linear multiple regression. The first record of the daughter was the dependent variable, and the various records ot the cow were the independent variables. As deviations from annual herd averages were used, no further adjustments were made for herds, years, or seasons. Consequently, genetic differences which may have existed between herds were removed by this procedure along with the environmental differences.

The records from 24 registered Guernsey herds and 77 Holstein herds were used for separate analyses. In the Guernsey herds the records were made from 1947 to 1962, and all records from the Holstein herds were made between 1952 and 1962. These herds are cooperating with Michigan State University and Michigan Artificial Breeders Cooperative in a breeding program which includes the production of young sires to be sampled in A. I.

From the Guernsey herds 904 cows with at least one of the first five lactations recorded had daughters with first records. Among these cows there was a range in numbers of records from 391 fifth lactations to 683 second lactations. Only 651 had first lactations.

In the Holstein herds 1,526 cows had daughters with first records. As including the fourth and fifth records appeared to add little information in the Guernsey data, only the first three records were included in the Holstein

grou
and
tion
mac
Als
The
ber

ter
cov
lat
Th
sin
ca
he
ca

Th
an
ab
te
re

group. Among the 1,526 cows, 853, 1,072, and 1,024 had first, second, and third records respectively. The apparent reason for this unlikely distribution of records was the short time involved. That is, many of the older cows made their first lactations previous to the time the herd started testing. Also, numerous purchased, untested cows were introduced into these herds. Therefore, the number of second and third records was larger than the number of first records.

The simple correlations of each of the cow's records with the daughter's first record are in Table 3. The correlations of the first record of a cow with the first record of a daughter are noticeably larger than the correlations of first record of the daughter with subsequent records of the cow. This situation may be caused partially by environmental and physiological similarities peculiar to first lactations such as the absence of a preceding calving interval and the infrequent occurrence of milk fever among first calf heifers. Different sets of genes affecting the various lactations could also cause more than usual similarity of lactations of the same sequence.

Table 3 also gives standard deviations of the production records. The production records are deviations from the herd average. In both Holsteins and Guernseys the variation is smallest in the cows' first records. Presumably the cow's first records represent a selected sample whereas the daughters are relatively unselected. Although the later records of the cows represent increasingly selected samples, their large variation apparently reflects

TAB

D

TABLE 3. —Correlations of cows' records with their daughters' first records

Guernseys

Cow's record	Correlation with daughter's first record	Standard deviation of milk production records in pounds deviation from herd average
1	0.149	1640
2	0.075	1730
3	0.081	1700
4	-0.016	1650
5	0.042	1760
Daughter's first		1730

Holsteins

Cow's record	Correlation with daughter's first record	Standard deviation of milk production records in pounds deviation from herd average
1	0.256	2040
2	0.145	2140
3	0.154	2130
Daughter's first		2150

feeble genetic selection and large random environmental fluctuations.

Table 4 lists the partial regression coefficients in the cases involving up to three records per cow. These are the weights to be used with the corresponding records of the cow to predict the first lactation of the daughter. That is:

$$\hat{Y} = \bar{Y} + \sum_{i=1}^{3} b_i X_i$$

where; $\bar{Y}$ = the average of all daughters' deviated records.

$\hat{Y}$ = estimated daughter's deviation from herd average.

b_i = partial regression coefficient of the daughter's first record on the ith record of the cow.

X_i = record of the cow's ith lactation as a deviation from the herd average.

i = lactation number of the cow's record.

Although the regressions and correlations indicate heritability is much larger in Holsteins than in Guernseys, many similarities between the breeds are apparent. The regression of a daughter's first record on the first record of her dam gives an estimate of one-half of heritability for this trait. These data indicate heritabilities of 0.54 for Holsteins and 0.32 for Guernseys. These estimates ignore differences between sires and between herds although by being deviations from herd average, they should be similar to intra-herd estimates. Comparing the multiple correlation coefficients shows that the addition of a second or third record, even with optimum weighting, gives little

TABLE 4. —Partial regression coefficients of daughters' first records on various records of the cow

Guernseys

Record of cow	Partial regression coefficients			Multiple correlation coefficient
	b_1	b_2	b_3	
First only	0.158			0.149
First two	0.159	-0.002		0.149
First three	0.155	0.003	0.011	0.149
Ave. first two				0.130
Ave. first three				0.123

Holsteins

Record of cow	Partial regression coefficients			Multiple correlation coefficient
	b_1	b_2	b_3	
First only	0.270			0.256
First two	0.253	0.037		0.258
First three	0.245	0.030	0.019	0.259
Ave. first two				0.229
Ave. first three				0.219

or no information about the superiority of the daughter beyond that furnished by the first record alone.

Correlations between the average of two or three records and the daughter's first record were estimated using the variances, covariances and correlations among these records. When the average of two records of a cow was used, the correlation with the daughter's first record was actually smaller than the correlation using the cow's first record alone. An even smaller correlation was obtained using the average of three records of a cow. This seems to be simply a diluting effect. A cow's first record is a good indicator of the daughter's first record, whereas the second and third records are very poor indicators. If equal variability existed among the cow's records, averaging would weight each the same. Actually the unweighted average gives slightly more emphasis to the more variable and much less reliable second and third records.

Repeatabilities were measured in these data by a weighted average of the various correlations of a cow's records with each other. Weighting was based on the number of records involved in the various correlations. Due to small numbers, the records beyond the third were omitted from this calculation. The resulting repeatability values were 0.53 for Guernseys and 0.52 for Holsteins. These high values of repeatability indicate that records beyond the first actually add little new information about the real producing ability of a cow. The heritability for the Holstein group was slightly higher

tha
sar
Th
thi
flu
it
fir

er
re
re
co
m
le
u
s

v
f
s
c

than the estimate of repeatability. Of course, this may have been due to sampling variation as the numbers involved in either estimate were not large. This could also indicate that repeatability was smaller because the second and third records included in repeatability were more sensitive to environmental fluctuations than the first but were excluded from heritability. That is, if it were possible to derive repeatability using only records as reliable as first lactations, the repeatability would be larger than heritability.

The partial regression coefficients indicate that nearly all the emphasis among records of the cow should be placed on the cow's first record to predict the superiority or inferiority of her daughter's first record. The multiple correlation coefficients of the various records of the cow with the first record of the daughter indicate that little additional information can be extracted from the second and third lactation records regardless of the manner of using them. The logical conclusion seemed to be to use the first lactation record to express a cow's phenotype for genetic studies.

Theoretical Basis for Constructing an Index

A selection index is a numerical expression constructed to predict with maximum accuracy the genic or general breeding value of individuals for milk production. The general breeding value (G_1) of an individual is synonymous with its genotype. As used here, G_1 refers to the genes which contribute in an additive manner to that individual's milk production. When the correlation between the index (I) and G_1 is maximum and I is normally distributed, truncation selection will result in maximum improvement in G per generation.

The index to be constructed is some function of the phenotypes for milk production of the individual and her close relatives. The linear prediction equation for the index is

$$I = b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$$

The b's are partial regression coefficients chosen such that the correlation between the index and the individual's genotype (G_1) is maximum within the limits of the accuracy of the estimated parameters. The sums of squares of the deviations of I from G_1 will thus be minimum. The X's are phenotypes or averages of phenotypes (first records of milk production expressed as deviations from herd averages) of the various groups of relatives with subscripts 1 for the cow or individual to which I refers, 2 for the dam of the cow designated 1, 3 for the average of the cow's daughters, 4 for the average of the paternal sisters of the cow, and 5 for the average of maternal sisters of 1.

The problem of predicting G_1 from the various X's is reflected in Figure 1.

Computing the b's in such a manner that G_1 is best predicted from the X's requires calculating the various correlations or covariances among X's and estimating the covariances between the X's and G_1. The correlations

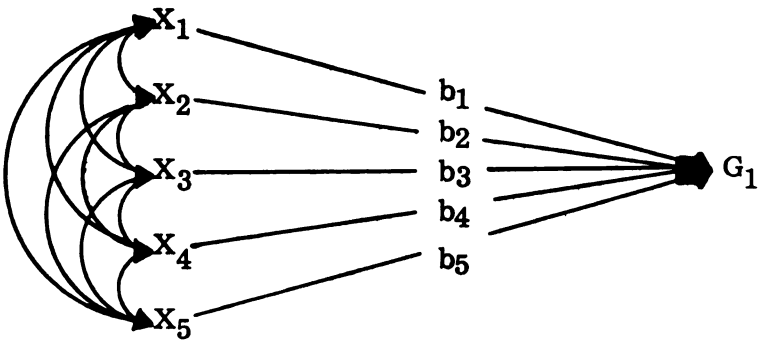


Fig. 1. —Relationships involved in predicting G_1 from the various phenotypes of the cow and her relatives

or the covariances among the X's are calculated directly from records of production, but the covariances involving G_1 must be inferred from information such as is illustrated in Figure 2, which shows the causal relationships involved.

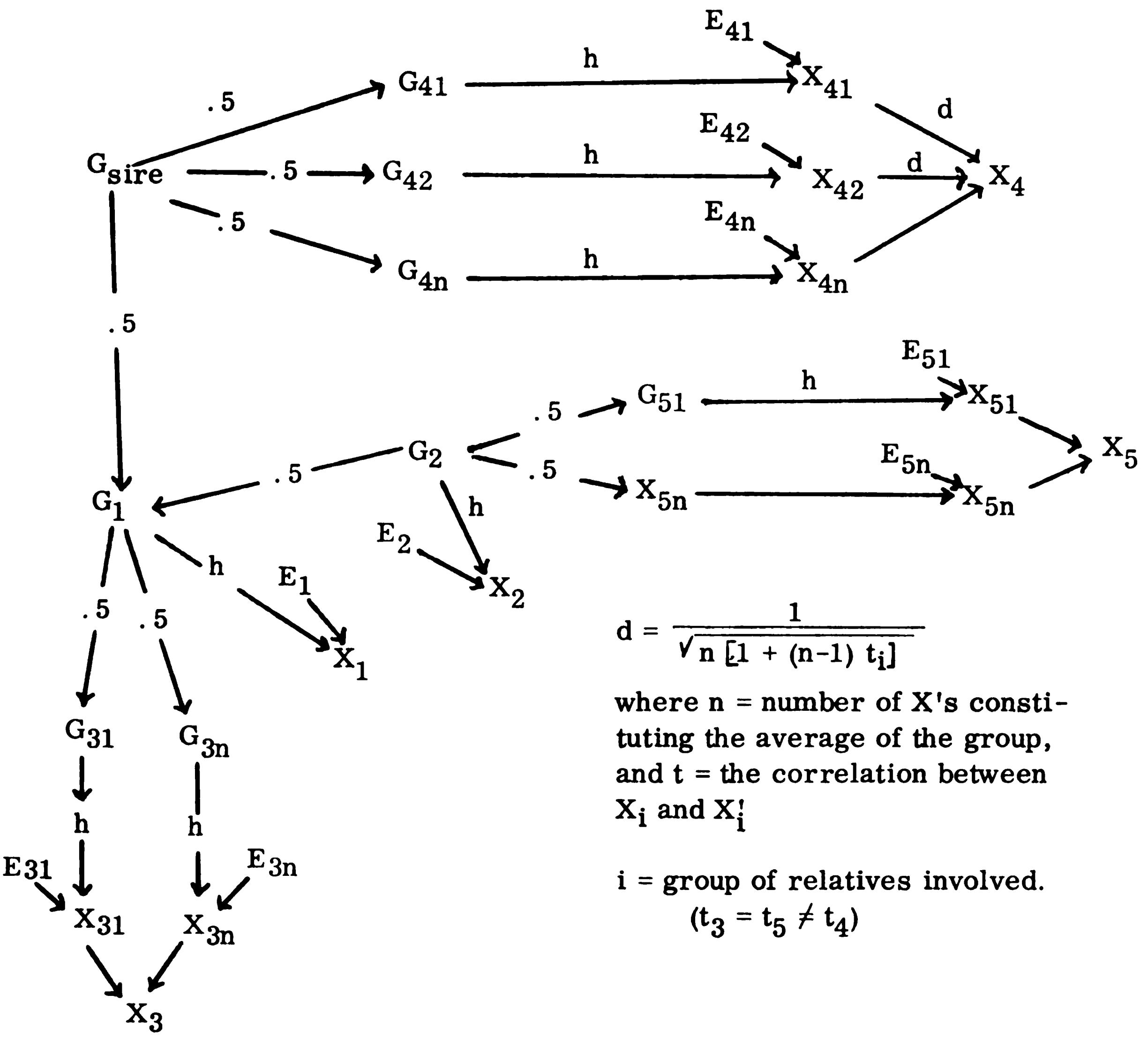
The correlation between any two items can be shown to be the sum of the products of the paths between them. The covariance between a cow's genotype and her phenotype would be:

$$\text{Cov } X_1G_1 = r_{X_1G_1}\ \sigma_{X_1}\ \sigma_{G_1}$$

$$r_{X_1G_1} = h = \frac{\sigma_{G_1}}{\sigma_{X_1}}$$

$$\text{thus Cov } X_1G_1 = \sigma^2_{G_1}$$

The covariances of the other X's with G_1 will be either one-half or one-fourth of Cov X_1G_1, depending on the genetic relationship between the



$$d = \frac{1}{\sqrt{n\,[1 + (n-1)\,t_i]}}$$

where n = number of X's constituting the average of the group, and t = the correlation between X_i and X_i'

i = group of relatives involved.
$(t_3 = t_5 \neq t_4)$

E = environment
G = genotype
X = phenotype

$h = \frac{\sigma_G}{\sigma_X} = \sqrt{\text{heritability}}$

<u>Subscripts</u>:

(first position)	(second position)
1 = cow	the individuals composing the average
2 = dam	
3 = average of daughters	
4 = average of paternal sisters	
5 = average of maternal sisters	

FIG. 2.—Causal relationships among phenotypes of the cow and her close relatives

particular group concerned and G_1.

The b's are derived by solving simultaneous equations from a variance-covariance matrix of the following form:

$$\begin{array}{lllll}
b_1 \text{ Var } X_1 & + b_2 \text{ Cov } X_1X_2 & + \ldots & + b_5 \text{ Cov } X_1X_5 & = \text{Cov } X_1G_1 \\
b_1 \text{ Cov } X_2X_1 & + b_2 \text{ Var } X_2 & + \ldots & + b_5 \text{ Cov } X_2X_5 & = \text{Cov } X_2G_1 \\
\cdot & \cdot & & \cdot & \cdot \\
\cdot & \cdot & & \cdot & \cdot \\
b_1 \text{ Cov } X_5X_1 & + b_2 \text{ Cov } X_5X_2 & + \ldots & + b_5 \text{ Var } X_5 & = \text{Cov } X_5G_1
\end{array}$$

The b's will change as either the size of the matrix (number of groups of relatives included) or the magnitude of the diagonal elements (the variances of the groups of relatives) change.

The variances of X_1 and X_2 do not change (except as the records represent a selected sample) because only single records are involved. The variances of X_3, X_4, and X_5 decrease as the number of individuals increases. The variance of averages from groups of size n can be measured from the formula:

$$\sigma^2_{\bar{X}} = \frac{\sigma^2_X + (n-1) \text{ Cov } XX'}{n}$$

where: $\sigma^2_{\bar{X}}$ = the variance of averages of groups of size n.

σ^2_X = the variance of individuals

and Cov XX' = the covariance among individuals of the group.

The multiple correlation coefficient (R) represents the correlation between I and G_1 and is a measure of the accuracy of the index for predicting the cow's genotype. R^2 is the fraction of the variance in G_1 associated with variation in I. Estimates of these expressions are obtained in the following manner:

$$R^2_{IG_1} = \frac{b_1 \text{ (Cov X } G_1) + \ldots + b_n \text{ (Cov } X_n G_1)}{\sigma^2_{G_1}}$$

$$R^2_{IG_1} = \frac{\sigma^2_I}{\sigma^2_{G_1}} \quad ; \quad R_{IG} = \frac{\sigma_I}{\sigma_{G_1}}$$

Estimates of Parameters

Phenotypic variances and covariances. —The variances and covariances needed to construct the index were obtained from the first lactation records of 8,984 Holstein cows, 7,638 of which had information available on some relative and 1,346 with an own record only. Only the 7,638 cows could be included in any of the calculations of phenotypic covariances among related groups. The variance of the 7,638 records was 44,059 and differed little from 44,971 which was the variance of the entire 8,984 records. Since the covariances had to be drawn from the 7,638 cows, the variance among them was used as the variance of the individual to form the index.

As an index for evaluating animals in different herds as well as within herds was desired, and because differences between herds had been removed by deviations from herd averages, all variances and covariances were calculated ignoring herds.

The phenotypic covariances among the groups of relatives were calculated from the sums of products. The covariances observed between single relatives were:

	X_1 Cow	X_2 Dam	X_3 Daughter	X_4 Paternal sister	X_5 Maternal sister
with:					
X_1		7,987	8,145	2,698	4,651
X_2			8,704	0	7,237
X_3				437	2,484
X_4					971

The various groups of relatives composed a number of sub-sets of data with different variances and different sample sizes. The variances of individuals in the various sub-sets were in close agreement with the variance of 44,059 from the larger population. Therefore, 44,059 was used as the variance of single records for the cow and for the individual in all groups of relatives. That is $\sigma^2_{X_1}$, $\sigma^2_{X_2}$, $\sigma^2_{X_3}$, $\sigma^2_{X_4}$, and $\sigma^2_{X_5}$ all equal 44,059 when they represent single relatives ($n = 1$). The variances of X_3, X_4, and X_5 become smaller as more individuals are included in the average. The variance of the average of the group can be obtained from the appropriate variances and covariances. As previously mentioned, the variance of an average of n individuals with like variance is:

$$\sigma^2_{\bar{X}} = \frac{\sigma^2_X + (n-1)\text{ Cov. } XX'}{n}$$

The Cov XX' is the covariance among individuals within the group constituting the average. The daughters represented by X_3 and the maternal sisters as X_5 have the same genetic relationship and usually similar environments within sets. The Cov XX' for these groups is equivalent to the covariance between the cow and her maternal sisters. To compute $\sigma^2_{\bar{X}}$ for X_3 and X_5, 4,651 was used as Cov XX'. The variance of groups of size n for daughters or maternal sisters was:

$$\sigma^2_{X_3} \text{ or } \sigma^2_{X_5} = \frac{44,059 + (n-1)\,4,651}{n}$$

Paternal sisters are often distributed in numerous herds where environmental similarities are usually small. Therefore, Cov XX' for the paternal sisters (X_4) should represent this situation of daughters of a sire scattered in many herds. The covariance between the cow and her paternal sisters was considered to estimate appropriately Cov XX' in determining the variances of the average of X_4. The variance of the average of groups of paternal sisters of size n was:

$$\sigma^2_{X_4} = \frac{44,059 + (n-1)\,2,698}{n}$$

Heritability. —Some knowledge of heritability is needed to evaluate the covariances between the various X's and G_1 that constitute the right side of the variance-covariance matrix as illustrated by the relationships in Figure 2.

Estimates of heritability from various portions of the data are given in Table 5. All of the estimates ignore sires and herds to approach more nearly the situation relating to an index to be used to compare cows from different herds. The heritabilities given were nearly the same as those indicated within herd-sire groups.

An analysis cross-classifying herds and sires was one of the methods used to estimate heritability from intra-sire components of variance. The cross-classified model is more efficient than the hierarchical classification and provides an opportunity to check for possible herd by sire interaction.

The model used was:

$$Y_{ijk} = \mu + h_i + s_j + hs_{ij} + e_{ijk}$$

where Y_{ijk} denotes the record (as a deviation from herd average) made by the k^{th} daughter of the j^{th} sire in the i^{th} herd. μ is a mean of all deviated records. h_i is the amount the i^{th} herd causes the records made in that herd to deviate from the average of all herds. s_i is the amount the j^{th} sire causes the average of his daughters to deviate from the average of all sires. hs_{ij} is the amount the particular combination of the i^{th} herd and the j^{th} sire causes the records of this combination to deviate from the additive combination of the i^{th} herd and the j^{th} sire. e_{ijk} is the amount the ijk^{th} record deviates from the average of all the records of the j^{th} sire in the i^{th} herd. It was assumed that, except for μ, all elements of the model are random, uncorrelated

TABLE 5. —Estimates of heritability

Method	Heritability	Sampling variance of heritability
Regression of cow on dam	0.376	.0018
Regression of paternal sisters on cow	0.246	.0005
Intra-dam correlation	0.436	.0057
Intra-sire correlation		
Interaction model	0.123	.0008
No interaction model	0.326	.0026

variables with zero expectation and variances σ_h^2, σ_s^2, σ_{hs}^2, and σ_e^2. These parameters are estimated by statistics correspondingly designated as H, S, HS, and E.

$$\text{Heritability} = \frac{4\,S}{S + HS + E}$$

The hierarchical model for intra-sire analysis of variance components was:

$$Y_{ijk} = \mu + h_i + s_{ij} + e_{ijk}$$

where Y_{ijk} denotes the record (as a deviation) made by the k^th daughter of the j^th sire in the i^th herd. μ is a population mean of deviated records. h_i is the amount the i^th herd causes the records made in that herd to deviate from the average of all herds. s_{ij} is the amount the j^th sire causes the average of his daughters in the i^th herd to deviate from the average of all daughters in the i^th herd. e_{ijk} is the amount the ijk^th record deviates from the average of all records of the daughters of the j^th sire in the i^th herd. It was assumed that, except for μ, all elements of the model are random, uncorrelated variables with zero expectation and variances σ_h^2, σ_s^2, and σ_e^2. The statistics which estimate these parameters are designated H, S, and E, respectively.

$$\text{Heritability} = \frac{4\,S}{S + E}$$

The hierarchical model was also used for intra-dam components of variance. The model was:

$$Y_{ijk} = \mu + h_i + d_{ij} + e_{ijk}$$

where Y_{ijk} denotes the deviated record made by the k^th daughter of the j^th dam in the i^th herd. μ, h_i and e_{ijk} are the same as defined above. d_{ij} is the amount the j^th dam causes the average of her daughters in the i^th herd to deviate from the average of all daughters in the i^th herd. It was assumed

TABLE 6.—Components of variance for milk production in first lactation*

Source	D. F.	Mean Square	Expected Mean Square
Cross Classified Model for Paternal Sisters			
Herds	195	83,210	$\sigma_e^2 + 6.4\ \sigma_{hs}^2 + 6.3\ \sigma_s^2 + 33.8\ \sigma_h^2$
Sires	676	69,315	$\sigma_e^2 + 5.9\ \sigma_{hs}^2 + 9.8\ \sigma_s^2 + 5.9\ \sigma_h^2$
Herd x Sire	1502	39,008	$\sigma_e^2 + 0.9\ \sigma_{hs}^2 - 0.8\ \sigma_s^2 + 2.6\ \sigma_h^2$
Residual	4721	39,694	σ_e^2
Components: H = 621, S = 1324, HS = 2199, E = 39694			
Hierarchical Model for Paternal Sisters			
Herds	195	83,210	$\sigma_e^2 + 6.4\ \sigma_s^2 + 33.8\ \sigma_h^2$
Sires/Herds	2178	48,414	$\sigma_e^2 + 2.5\ \sigma_s^2$
Residual	4271	39,694	σ_e^2
Components: H = 617, S = 3526, E = 39694			
Hierarchical Model for Maternal Sisters			
Herds	187	58,494	$\sigma_e^2 + 2.4\ \sigma_d^2 + 18.3\ \sigma_h^2$
Dams/Herds	1353	46,223	$\sigma_e^2 + 2.2\ \sigma_d^2$
Residual	1899	36,366	σ_e^2
Components: H = 638, D = 4449, E = 36366			

---

*Milk production was expressed in deviations of 10 pounds from the annual herd average.

that, except for μ, all elements of the model are random, uncorrelated variables with zero expectation and variances σ_h^2, σ_d^2, and σ_e^2. The statistics which estimate these parameters are designated H, D, and E, respectively. Similarly, heritability is estimated as: 4 D / D + E.

The sampling variances of the heritability estimates were calculated as described by Falconer (1960). The sampling variances for the intra-sire estimates of heritability are only rough approximations as the numbers of offspring per sire were variable.

The herd by sire interaction component in the cross-classified analysis was larger than expected on the basis of previous studies. Consequently, the heritability indicated from this model is much smaller than the other estimates from this population. A satisfactory explanation for this interaction is not available. Some conditions that could possibly cause a herd by sire interaction on deviated, first lactation records may include: (1) marked differences from herd to herd involving preferential treatment among the daughters of various sires, (2) wide differences in age structure among herds, and (3) a non-random distribution of year or season of calving among certain sires' daughters. The fact that many herds are involved makes the first two possibilities seem unlikely. Although (3) seems to be more likely to occur than either (1) or (2), the real nature of these effects remains unexplained. Some differences could arise from differences in fertility of sires among seasons; yet, this effect would seem to be small as measured in the production of the daughters' first lactations.

The estimates of heritability derived from groups of dams are noticeably larger than those estimates obtained from groups of sires. The estimates from groups of dams appear to be somewhat inflated by common environmental effects within herds. The nature of such common environmental effects that may exist is not apparent. Possible causes would include (1) a positive correlation between different daughters of the same dam caused by similar preferential treatment, (2) maternal effects, and (3) the use of standard age correction factors. The latter possibility would depend upon the existence of real genetic differences in the rate of maturity among cows.

Due to the difficulties in finding the real nature of the herd by sire interaction and because of the lack of agreement among estimates, the heritability with the smallest sampling variance was used. The value of 0.246 derived from the regression of the paternal sisters on the cow was used as heritability.

The Index

The variance-covariance matrix from which the b's were derived was:

Equation	b_1	b_2	b_3	b_4	b_5		
X_1	44059	7987	8145	2698	4651	=	10839
X_2	7987	44059	8704	0	7237	=	5420
X_3	8145	8704	44059	437	2484	=	5420
X_4	2698	0	437	44059	971	=	2710
X_5	4651	7237	2484	971	44059	=	2710

The diagonal elements are variances of the various X's when the X's are single individuals. The diagonal elements of X_3, X_4, and X_5 were changed to the variances of averages and for each change a new set of equations was solved. A large number of sets of equations was solved to arrive at b's for different combinations of kinds and amounts of information.

Sample portions of the weights from the index and the corresponding multiple correlation coefficients are shown in Tables 7 and 8. These tables give only a small portion of the comparisons that were made; yet, they should give a good picture of how the weights vary with changes in the number of relatives involved.

A more concise picture of the relative usefulness of the various kinds and amounts of information can be judged from Table 9 which gives only the multiple correlation coefficients of the cow's index with her genotype.

An R value of 0.50 may be considered as the "base" for all the combinations which include the cow's own record. This is the value obtained where only the cow's own record is used to estimate her genic value. This value of R is the square root of heritability.

The addition of the dam's record raises R only to 0.52. Eight daughters plus the cow's record gives an R of 0.62, the last few being nearly as useful as the first one or two in increasing the accuracy. The addition of maternal sisters to the cow's record makes only a small increase in R even if as many as eight are considered.

TABLE 7. —**Partial regression coefficients and multiple R's for all five groups of relatives**

Part 1, $N_5 = 1$

Daughters N_3	Paternal Sisters N_4	b_1	b_2	b_3	b_4	b_5	R
1	1	.22	.07	.07	.05	.02	.55
1	2	.21	.07	.07	.09	.02	.56
1	3	.21	.07	.07	.13	.02	.56
1	10	.20	.07	.07	.33	.02	.60
1	20	.19	.07	.07	.45	.02	.61
1	50	.18	.07	.07	.62	.01	.64
1	100	.18	.08	.07	.71	.01	.65
1	200	.17	.08	.07	.76	.01	.66
2	1	.21	.06	.13	.05	.02	.56
2	2	.20	.06	.13	.09	.02	.57
2	3	.20	.06	.13	.13	.02	.58
2	10	.19	.06	.13	.33	.02	.61
2	20	.18	.06	.13	.45	.01	.63
2	50	.17	.06	.13	.62	.01	.65
2	100	.17	.07	.13	.71	.01	.67
2	200	.16	.07	.13	.76	.01	.67
3	1	.20	.05	.19	.05	.02	.58
3	2	.20	.05	.19	.09	.02	.58
3	3	.19	.05	.19	.13	.02	.59
3	10	.18	.05	.19	.33	.02	.62
3	20	.17	.05	.19	.45	.01	.64
3	50	.16	.05	.19	.62	.01	.66
3	100	.16	.06	.19	.71	.01	.68
3	200	.16	.06	.19	.76	.01	.68
8	1	.16	.01	.41	.05	.02	.62
8	2	.16	.01	.41	.09	.02	.63
8	3	.16	.01	.41	.13	.02	.64
8	10	.15	.01	.41	.33	.01	.67
8	20	.14	.02	.41	.45	.01	.68
8	50	.13	.02	.41	.62	.01	.71
8	100	.12	.02	.41	.70	.01	.72
8	200	.12	.02	.41	.76	.01	.72

TABLE 7. --Continued

Part 2, $N_5 = 2$

N_3	N_4	b_1	b_2	b_3	b_4	b_5	R
1	1	.21	.06	.07	.05	.04	.55
1	2	.21	.06	.07	.09	.04	.56
1	3	.21	.07	.07	.13	.04	.56
1	10	.20	.07	.07	.33	.03	.60
1	20	.19	.07	.07	.45	.03	.62
1	50	.18	.07	.07	.62	.02	.64
1	100	.17	.07	.07	.70	.02	.65
1	200	.17	.08	.07	.76	.02	.66
2	1	.21	.05	.13	.05	.04	.56
2	2	.20	.05	.13	.09	.04	.57
2	3	.20	.05	.13	.13	.04	.58
2	10	.19	.06	.13	.33	.03	.61
2	20	.18	.06	.13	.45	.03	.63
2	50	.17	.06	.13	.62	.02	.65
2	100	.17	.06	.13	.70	.02	.67
2	200	.16	.07	.13	.76	.02	.67
3	1	.20	.04	.19	.05	.04	.58
3	2	.19	.04	.19	.09	.04	.58
3	3	.19	.05	.19	.13	.04	.59
3	10	.18	.05	.19	.33	.03	.62
3	20	.17	.05	.19	.45	.03	.64
3	50	.16	.05	.19	.62	.02	.66
3	100	.16	.05	.19	.70	.02	.68
3	200	.15	.06	.19	.76	.02	.68
8	1	.16	.01	.41	.05	.03	.63
8	2	.16	.01	.41	.09	.03	.63
8	3	.16	.01	.41	.13	.03	.64
8	10	.15	.01	.41	.33	.03	.67
8	20	.14	.01	.41	.45	.02	.68
8	50	.13	.02	.41	.62	.02	.71
8	100	.12	.02	.41	.70	.01	.72
8	200	.12	.02	.41	.76	.01	.73

TABLE 7. —Continued

Part 3, $N_5 = 3$

N_3	N_4	b_1	b_2	b_3	b_4	b_5	R
1	1	.21	.06	.07	.05	.06	.55
1	2	.21	.06	.07	.09	.06	.56
1	3	.21	.06	.07	.13	.06	.56
1	10	.20	.07	.07	.33	.05	.60
1	20	.19	.07	.07	.45	.04	.62
1	50	.18	.07	.07	.61	.03	.64
1	100	.17	.07	.07	.70	.03	.65
1	200	.17	.07	.07	.75	.03	.66
2	1	.20	.05	.13	.05	.06	.57
2	2	.20	.05	.13	.09	.06	.57
2	3	.20	.05	.13	.13	.05	.58
2	10	.19	.06	.13	.33	.04	.61
2	20	.18	.06	.13	.45	.04	.63
2	50	.17	.06	.13	.61	.03	.65
2	100	.16	.06	.13	.70	.03	.67
2	200	.16	.06	.13	.76	.02	.67
3	1	.20	.04	.19	.05	.06	.58
3	2	.20	.04	.19	.09	.05	.59
3	3	.19	.04	.19	.13	.05	.59
3	10	.18	.05	.19	.33	.04	.62
3	20	.17	.05	.19	.45	.04	.64
3	50	.16	.05	.19	.61	.03	.66
3	100	.16	.05	.19	.70	.03	.68
3	200	.15	.05	.19	.76	.02	.68
8	1	.16	.00	.41	.05	.05	.63
8	2	.16	.01	.41	.09	.05	.63
8	3	.16	.01	.41	.13	.04	.64
8	10	.14	.01	.41	.33	.04	.67
8	20	.14	.01	.41	.45	.03	.68
8	50	.13	.02	.41	.61	.02	.71
8	100	.12	.02	.41	.70	.02	.72
8	200	.12	.02	.41	.76	.02	.73

TABLE 7. —Continued

Part 4, $N_5 = 8$

N_3	N_4	b_1	b_2	b_3	b_4	b_5	R
1	1	.21	.05	.07	.05	.12	.56
1	2	.21	.05	.07	.12	.11	.57
1	3	.20	.05	.07	.12	.11	.57
1	10	.19	.06	.07	.32	.09	.60
1	20	.19	.06	.07	.44	.08	.62
1	50	.18	.06	.07	.61	.07	.64
1	100	.17	.07	.07	.69	.06	.65
1	200	.17	.07	.07	.75	.05	.66
2	1	.20	.04	.13	.05	.12	.57
2	2	.20	.04	.13	.09	.11	.58
2	3	.20	.04	.13	.12	.11	.58
2	10	.18	.05	.13	.32	.09	.61
2	20	.18	.05	.13	.44	.08	.63
2	50	.17	.06	.13	.61	.06	.65
2	100	.16	.06	.13	.69	.06	.67
2	200	.16	.06	.13	.75	.05	.67
3	1	.19	.03	.18	.05	.11	.58
3	2	.19	.03	.18	.09	.11	.59
3	3	.19	.04	.18	.12	.11	.59
3	10	.18	.04	.18	.32	.09	.63
3	20	.17	.04	.18	.44	.08	.64
3	50	.16	.05	.18	.61	.06	.67
3	100	.15	.05	.19	.69	.05	.68
3	200	.15	.05	.19	.75	.05	.68
8	1	.16	.00	.41	.05	.10	.63
8	2	.16	.00	.41	.09	.09	.64
8	3	.15	.00	.41	.12	.09	.64
8	10	.14	.01	.41	.33	.07	.67
8	20	.14	.01	.41	.44	.06	.68
8	50	.13	.01	.41	.61	.04	.71
8	100	.12	.01	.41	.70	.04	.72
8	200	.12	.02	.41	.75	.03	.73

TABLE 8. —Partial regression coefficients and multiple R's for all information except dams records

Part 1, N_5 = 1

Daughters N_3	Paternal Sisters N_4	b_1	b_3	b_4	b_5	R
1	1	.23	.08	.05	.03	.53
1	2	.22	.08	.09	.03	.54
1	3	.22	.08	.12	.03	.55
1	10	.21	.08	.33	.03	.58
1	20	.20	.08	.44	.03	.60
1	50	.19	.08	.61	.02	.62
1	100	.19	.08	.69	.02	.64
1	200	.19	.08	.74	.02	.64
2	1	.21	.15	.05	.03	.55
2	2	.21	.15	.09	.03	.56
2	3	.21	.15	.12	.03	.57
2	10	.20	.15	.33	.03	.60
2	20	.19	.15	.44	.02	.62
2	50	.18	.15	.61	.02	.64
2	100	.17	.15	.69	.02	.65
2	200	.17	.15	.75	.02	.66
3	1	.20	.21	.05	.03	.57
3	2	.20	.21	.09	.03	.58
3	3	.20	.21	.13	.03	.58
3	10	.18	.21	.33	.02	.61
3	20	.18	.21	.44	.02	.63
3	50	.17	.21	.61	.02	.66
3	100	.16	.21	.69	.02	.67
3	200	.16	.21	.75	.02	.68
8	1	.16	.42	.05	.02	.62
8	2	.16	.42	.09	.02	.63
8	3	.16	.42	.13	.02	.64
8	10	.15	.42	.33	.01	.67
8	20	.14	.42	.45	.01	.68
8	50	.13	.43	.61	.01	.71
8	100	.12	.43	.70	.01	.72
8	200	.12	.43	.75	.01	.72

TABLE 8. —Continued

Part 2, $N_5 = 2$

N_3	N_4	b_1	b_3	b_4	b_5	R
1	1	.22	.08	.05	.06	.54
1	2	.22	.08	.09	.06	.54
1	3	.22	.08	.12	.06	.55
1	10	.21	.08	.32	.05	.58
1	20	.20	.08	.44	.05	.60
1	50	.19	.08	.60	.04	.63
1	100	.18	.08	.69	.04	.64
1	200	.18	.08	.74	.04	.65
2	1	.21	.15	.05	.05	.56
2	2	.21	.15	.09	.05	.56
2	3	.21	.15	.12	.05	.57
2	10	.19	.15	.33	.05	.60
2	20	.19	.15	.44	.04	.62
2	50	.18	.15	.60	.04	.64
2	100	.17	.15	.69	.04	.65
2	200	.17	.15	.74	.03	.66
3	1	.20	.20	.05	.05	.57
3	2	.20	.21	.09	.05	.58
3	3	.20	.21	.12	.05	.58
3	10	.18	.21	.33	.04	.62
3	20	.18	.21	.44	.04	.63
3	50	.17	.21	.61	.03	.66
3	100	.16	.21	.69	.03	.67
3	200	.16	.21	.74	.03	.68
8	1	.16	.42	.05	.04	.63
8	2	.16	.42	.09	.03	.63
8	3	.16	.42	.13	.03	.64
8	10	.15	.42	.33	.03	.67
8	20	.14	.42	.45	.02	.68
8	50	.13	.42	.61	.02	.71
8	100	.12	.43	.70	.02	.72
8	200	.12	.43	.75	.01	.72

TABLE 8. —Continued

Part 3, $N_5 = 3$

N_3	N_4	b_1	b_3	b_4	b_5	R
1	1	.22	.08	.05	.08	.54
1	2	.22	.08	.09	.08	.55
1	3	.22	.08	.12	.08	.55
1	10	.20	.08	.32	.07	.58
1	20	.20	.08	.44	.07	.60
1	50	.19	.08	.60	.06	.63
1	100	.18	.08	.06	.06	.64
1	200	.18	.08	.74	.05	.65
2	1	.21	.14	.05	.08	.56
2	2	.21	.14	.09	.07	.56
2	3	.20	.14	.12	.07	.57
2	10	.19	.15	.32	.06	.60
2	20	.19	.15	.44	.06	.62
2	50	.18	.15	.60	.05	.64
2	100	.17	.15	.69	.05	.66
2	200	.17	.15	.74	.05	.66
3	1	.20	.20	.05	.07	.57
3	2	.20	.20	.09	.07	.58
3	3	.19	.20	.12	.07	.59
3	10	.18	.21	.32	.06	.62
3	20	.18	.21	.44	.05	.63
3	50	.17	.21	.60	.05	.66
3	100	.16	.21	.69	.04	.67
3	200	.16	.21	.74	.04	.68
8	1	.16	.41	.05	.05	.63
8	2	.16	.41	.09	.05	.63
8	3	.16	.41	.12	.05	.64
8	10	.14	.42	.33	.04	.67
8	20	.14	.42	.45	.03	.68
8	50	.13	.42	.61	.03	.71
8	100	.12	.42	.70	.02	.72
8	200	.12	.42	.75	.02	.72

TABLE 8. —Continued

Part 4, $N_5 = 8$

N_3	N_4	b_1	b_3	b_4	b_5	R
1	1	.21	.07	.04	.16	.55
1	2	.21	.07	.08	.15	.55
1	3	.21	.07	.12	.15	.56
1	10	.20	.08	.31	.14	.59
1	20	.19	.08	.43	.13	.61
1	50	.18	.08	.59	.11	.63
1	100	.18	.08	.67	.11	.64
1	200	.18	.08	.72	.10	.65
2	1	.20	.14	.04	.14	.56
2	2	.20	.14	.08	.14	.57
2	3	.20	.14	.12	.14	.58
2	10	.19	.14	.32	.12	.61
2	20	.18	.14	.43	.11	.62
2	50	.17	.14	.59	.10	.65
2	100	.17	.14	.68	.10	.66
2	200	.17	.14	.73	.09	.66
3	1	.19	.20	.04	.13	.58
3	2	.19	.20	.08	.13	.59
3	3	.19	.20	.12	.13	.59
3	10	.18	.20	.32	.11	.62
3	20	.17	.20	.43	.10	.64
3	50	.16	.20	.59	.09	.66
3	100	.16	.20	.68	.08	.67
3	200	.15	.20	.73	.08	.68
8	1	.16	.40	.05	.10	.63
8	2	.16	.41	.09	.09	.64
8	3	.15	.41	.12	.09	.64
8	10	.14	.41	.32	.07	.67
8	20	.14	.41	.44	.07	.68
8	50	.13	.42	.61	.05	.71
8	100	.12	.42	.69	.05	.72
8	200	.12	.42	.75	.04	.72

TABLE 9. —**Multiple correlation coefficients** using various combinations of **information**

Information Available				Paternal Sisters X4								
Cow X_1	Dam X_2	Dau's. X_3	Maternal Sisters X_5	0	1	2	3	10	20	50	100	200
1				.50	.50	.51	.52	.56	.57	.60	.61	.62
1	1			.52	.53	.54	.54	.58	.60	.63	.64	.65
1		1		.52	.53	.54	.54	.58	.60	.62	.63	.64
1		2		.54	.55	.56	.56	.60	.61	.64	.65	.66
1		3		.56	.57	.57	.58	.61	.63	.65	.67	.67
1		8		.62	.62	.63	.64	.67	.68	.71	.72	.72
1			1	.50	.51	.52	.52	.56	.58	.58	.59	.60
1			2	.51	.51	.52	.53	.56	.58	.61	.62	.63
1			3	.51	.52	.52	.53	.56	.58	.61	.62	.63
1			8	.52	.53	.53	.54	.57	.59	.61	.62	.63
1	1	1		.55	.55	.55	.56	.60	.61	.64	.65	.66
1	1	2		.55	.56	.57	.57	.61	.63	.65	.66	.67
1	1	3		.57	.57	.58	.59	.62	.64	.66	.68	.68
1	1	8		.62	.62	.63	.64	.67	.68	.71	.72	.72
1	1		1	.52	.53	.64	.55	.58	.60	.63	.64	.65
1	1		2	.53	.53	.54	.55	.58	.60	.63	.64	.65
1	1		3	.53	.54	.54	.55	.58	.60	.63	.64	.65
1	1		8	.53	.54	.55	.55	.59	.60	.63	.64	.65
1	1	1	1	.54	.55	.56	.56	.60	.61	.64	.65	.66
1	1	2	1	.56	.56	.57	.58	.61	.63	.65	.67	.67
1	1	3	1	.57	.58	.58	.59	.62	.64	.66	.68	.68
1	1	8	1	.62	.62	.63	.64	.67	.68	.71	.72	.72
1	1	1	2	.54	.55	.56	.56	.60	.62	.64	.65	.66
1	1	2	2	.56	.56	.57	.58	.61	.63	.65	.67	.67
1	1	3	2	.57	.58	.58	.59	.62	.64	.66	.68	.68
1	1	8	2	.62	.63	.63	.64	.67	.68	.71	.72	.73
1	1	1	3	.54	.55	.56	.56	.60	.62	.64	.65	.66
1	1	2	3	.56	.57	.57	.58	.61	.63	.65	.67	.67
1	1	3	3	.57	.58	.59	.59	.62	.64	.66	.68	.68
1	1	8	3	.62	.63	.63	.64	.67	.68	.71	.72	.73
1	1	1	8	.55	.56	.56	.57	.60	.62	.64	.65	.66
1	1	2	8	.56	.57	.58	.58	.61	.63	.65	.67	.67
1	1	3	8	.57	.58	.59	.59	.63	.64	.67	.68	.68
1	1	8	8	.62	.63	.64	.64	.67	.68	.71	.72	.73

The situation is quite different in the case of adding paternal sisters to the cow's record. The first three paternal sisters increase R the same amount as do eight maternal sisters; whereas, additional paternal sisters continue to raise R. The number of paternal sisters required to equal the information furnished by eight daughters is somewhere between 100 and 200; yet, this accuracy is closely approached when 50 paternal sisters are used.

In the cases where a third source of information is added to that from records of the cow and dam, much the same situation exists as if no record of the dam was used. When the numbers involved in the third source of information are at all large, little change in R would occur if the dam's record were omitted.

The limited value of the maternal sisters in increasing R is an obvious feature of Table 9. The highest increase in R that can be made by maximum use of the maternal sisters is 0. 02 units; yet, for most situations additional maternal sisters do not change R at all.

Substantial increases in R can be obtained by the addition of paternal sisters. This appears to be true for all combinations except those cases where very large numbers of the other groups of relatives are used.

Some justification could be made for deleting completely the dam and maternal sisters. However, the addition of a record from a dam or a maternal sister is most useful in situations where other information is limited. Such situations are common most of the time.

Table 10 lists a number of correlation coefficients applicable to heifers and young bulls. This table gives R values for several combinations of information which do not include the individual's own record or records on daughters. In such situations the dam's record alone provides only a fair indication of the individual's genotype as indicated by an R of 0. 25 Neither maternal nor paternal sisters give a high R unless large numbers are used. However, when both maternal and paternal sisters are used in sufficient numbers, quite large R's can be obtained. Obtaining some information from both sides of the pedigree adds considerably to the accuracy of estimating an individual's genotype. However, the use of both dam and maternal sister information furnish little more accuracy than either source used alone. Records of paternal sisters can be more valuable than information from the female side of the pedigree because of the larger numbers of relatives possible and also because of the independence of the information due to less covariance among paternal sisters as compared to female relatives.

The R values from Table 10 indicate that indexing a young sire, with only the information considered here, cannot become highly accurate in estimating his genotype. Such mediocre accuracy is a result of utilizing only three sources of information, two of which furnish little evidence.

If 20 or more paternal sisters are available, most of the information from the sire's side of the pedigree is utilized. However, the dam's record and records of a few maternal sisters do not render nearly so much information

TABLE 10. —Accuracies (R_{IG_1}) of information available for young animals

Information Available		Paternal sisters X_4								
Dam X_2	Maternal sisters X_5	0	1	2	3	10	20	50	100	200
-	-	--	.12	.29	.31	.32	.38	.44	.47	.48
1	-	.25	.28	.30	.32	.41	.45	.50	.53	.54
-	1	.12	.17	.21	.23	.34	.39	.45	.47	.49
-	2	.14	.20	.23	.26	.34	.40	.45	.48	.50
-	3	.18	.23	.25	.27	.37	.41	.46	.49	.50
-	8	.26	.29	.31	.32	.40	.43	.48	.50	.52
1	1	.26	.29	.31	.33	.41	.46	.51	.53	.54
1	2	.27	.30	.32	.34	.42	.46	.51	.53	.55
1	3	.28	.31	.33	.34	.42	.46	.51	.53	.55
1	8	.31	.33	.35	.36	.44	.47	.52	.54	.55

about the female side of the pedigree. The question then arises as to the possibility of including the other close relatives of the dam of the young sire. The value of information from the relatives of the son's dam (other than those which are already included in the son's index) will vary as the kind and amount of relatives change. Several sets of simultaneous equations were solved to evaluate the usefulness of information from the dam's other relatives. The R's, correlations between the young sire's index and his genotype, for various combinations of information on the cow (dam of the young sire) are listed in Table 11. The young son is considered to be sired by a sire unrelated to the dam and proven in A. I. by 50 daughters (paternal sisters to the young son).

Method A comes from Table 10 and represents the index on the son using nothing beyond the parents. Method B uses all the information included in Method A in addition to information on the maternal grandparents.

The 50 paternal sisters of the young sire alone give an R of 0.44, and the addition of the record of his dam raises the accuracy to 0.50. Method B raises the accuracy to 0.51, 0.51, and 0.52 if the dam has 10, 20, or 100 paternal sisters, respectively. If the son's dam and his 3 maternal sisters are used, both the methods use the same information and give an R of 0.51. The addition of 100 paternal sisters of the son's dam increases the accuracy of Method B to only 0.52 However, the dam's dam and the dam's maternal sisters add no accuracy to estimating the son's genotype if several of the dam's paternal sisters are available.

The R's listed in Table 11 would be somewhat altered if the information from the A. I. sire was not independent of the information on the female side of the pedigree. Inbreeding would tend to increase the R's due to the reduction of sampling variation caused by Mendelian segregation. On the other hand, the R's would be lowered by environmental similarities which increase the covariance among relatives in the individual's pedigree.

From Table 11 the accuracy of estimating a young sire's genotype can be increased slightly by considering his dam's relatives which are not already included in his index. All of the increase in accuracy comes from his dam's paternal sisters. The maternal granddam and the dam's maternal

TABLE 11. —Accuracy of indexing a young sire

Information on cow (dam of young sire)					Daughters of A. I. sire (paternal sisters of young sire)	Accuracy* (R_{IG})	
X_1	X_2	X_3	X_4	X_5		Method A	Method B
-	-	-	-	-	50	.44	.44
1	-	-	-	-	50	.50	.50
1	-	-	10	-	50	.50	.51
1	-	-	20	-	50	.50	.51
1	-	-	100	-	50	.50	.52
1	-	3	-	-	50	.51	.51
1	-	3	100	-	50	.51	.52
1	1	3	100	-	50	.51	.52
1	1	3	100	3	50	.51	.52

*Correlation of young sire's genotype with his index.

Method A uses information on the young sire's dam and sibs.

Method B uses information as in Method A plus maternal grandam and dam's sibs.

sisters appear to be of no value in estimating the genotype of a young sire. In estimating the genotype of a young bull whose sire's and dam's genotypes are relatively well known, the dam's paternal sisters can be of some use, especially if several are available.

Evaluation of the Index on a Test Population

The multiple correlation coefficient between the genotype of a cow and her index was considered not to be a completely adequate method of evaluating the index. A test of the practical usefulness in a population of cows was desired.

A comparison of selection by index with mass selection on the cow's record was made. The test population was chosen on the same basis as the population used to derive the index (Phase II). The records from 145 herds were used in this analysis. The records were made during the same years as were those in Phase II, but none of the same herds were involved.

The first lactation record (as a deviation from herd average) of an unselected daughter was used as the criterion for comparing selection of the cow by index with selection based on own phenotype. The test population was limited to cows that had an own record as well as at least one daughter with a record. This limitation reduced the test population to 429 cows and 498 daughters. In calculating the indexes of the cows, the particular daughter considered as the dependent variable was excluded to avoid injecting a part-whole relationship into the correlations.

The relationships from the test population appear in Table 12. The test population was noticeably more variable than the population used to derive the index, but the heritability was nearly the same. However, the small numbers involved leave a wide margin for sampling variation.

TABLE 12. —Relationships between phenotype, index, and daughter of 498 cows in the test population

	Daughter Y	Cow X	Index I	XY	IY
Average production (pounds deviation from herd average)	44	370	137		
Variances and Covariances*	53,297	58,916	4,568	7,825	2,582
Standard deviation (pounds deviation from herd average)	2,309	2,427	676		

$r_{XY} = 0.140$

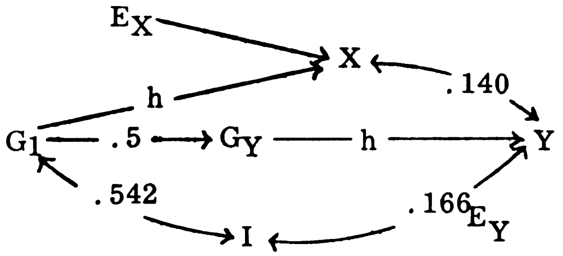
$r_{IY} = 0.166$ $\qquad \frac{r_{IY}}{r_{XY}} = 1.19$

$b_{YX} = 0.133$ $\qquad$ Heritability = 0.266

*Coded units of 100 pounds.

The correlation of the cow's index with an unselected daughter's record was 0.166 as compared to 0.140 for the correlation of the cow's own phenotype with her daughter's record. These correlations indicate that selection by index in this population would have been 19 per cent more effective than selection based on the cow's own phenotype.

An average or aggregate R for the test population was determined by using Fisher's Z transformation. This procedure consisted of (a) obtaining an R value based on the number of each type of relatives available for each cow, (b) converting to Z by use of standard statistical tables, (c) averaging the Z values for the test population, and (d) converting back to R from the Z table. The resulting R value of 0.542 was an indication of the average amount of information used in deriving the indexes. Figure 3 shows the relationships involved in the test population. These relationships were the basis for comparing the calculated correlations with those that would be expected by theoretical inferences.



I = index of cow
E = environment
G_1 = genotype of cow
G_Y = genotype of daughter
$h = \sqrt{\text{heritability}} = .516$
Y = daughter's phenotype
X = cow's phenotype

FIG. 3. —Genetic and phenotypic relationships in the test population

The correlation of 0.166 between the cow's index and her daughter's phenotype is higher than would be expected on the basis of the other relationships shown in Figure 3. Sampling variation in this small population seems to be the most likely cause of divergence between calculated and theoretical values. Other possibilities would include (a) underestimation of heritability in the test population, the index population, or both; or (b) underestimating aggregate R by a different weighting of the individual R's. In the case of (a) a heritability of either 0.378 in the test population or approximately 0.40 in the index population would be required singly to make the theoretical correlation between I and Y agree with the calculated value of 0.166. A more plausible explanation would seem to be that heritability was underestimated to a lesser degree in both populations. In reference to (b) it seems probable that aggregate R would be underestimated by the use of large numbers of cows with low R's. That is, the most abundant groups of cows were those with the least amount of information, the most variable in genic values, and yet influence aggregate R the most.

Regardless of the nature of the difference between actual and theoretical estimates, the 19 per cent advantage of index selection over mass selection appears to be sufficient evidence to justify a moderate amount of effort and expense to index cows, especially dams of sires, for selection.

The increased accuracy of selection in the test population cannot, in itself, be taken as proportional to genetic gain in a practical situation. The amount of information available from relatives in a typical herd should be in close proximity to that of the test population. Although the test population should have more than average information available from relatives as a result of being a selected sample of older cows, this accuracy should be offset by the loss of one daughter per cow omitted from consideration as the dependent variable. The amount of information which would normally be available on potential dams of sires may be considered considerably more than that of the test population.

The influence of selection for other traits and any factor reducing the selection differential or increasing the generation interval would result in less genetic progress for milk production than that indicated in the test population.

CHAPTER V

APPLICATION OF RESULTS

The index was developed to measure breeding values of cows and young animals and was not intended to substitute for or to replace current methods of evaluating sires with progeny information. The existing methods of evaluating sires have been developed to a high degree of refinement and accuracy. The index was developed for the purpose of increasing the accuracy of female selection within as well as between herds. The increase in accuracy of female selection by index over selection on own performance should be from 10 to 20 per cent depending on the amount of information available from relatives. The genetic gain which can be obtained directly by female selection with the use of the index appears to be real but not large.

The greatest usefulness of the index appears to be in selecting the cows that will become the dams of future sires. The economic and biological limitations which reduce both the number of bulls that a stud can test and the intensity with which tested bulls can be selected will result in genetic progress below the maximum possible. A.I. units, however, have wide latitude in deciding which bulls will be tested, and this index offers an effective method of selecting cows to produce these sires. The most promising and practical approach to genetic progress and breed improvement appears to be in producing numerous young bulls from well proven sires and out of intensely selected dams, testing as many of these as possible, and culling among these bulls as severely as feasible on the basis of information from progeny.

Prediction equations applicable to heifers and young sires were among those developed. The accuracy of these equations is generally low in comparison to those which utilize the individuals's own record or records of daughters. In estimating the genotype of a young bull sired by a well proven sire, the dam's paternal sisters can be of limited use if several are available; yet the dam's maternal sisters and the dam's dam are of no apparent value. The situation is to be expected since these relatives are of limited use in estimating the breeding value of the dam herself. Prediction equations could easily be developed from the parameters available wherein young sires could be indexed using all the information included in the dam's index. Actually there seems to be little occasion to use such information. The weights for the various relatives would be nearly the same as would result from an average of the sire's index and the dam's index if there were a moderate amount of information available from both sire and dam.

Although the index could be used to advantage by the breeder doing his own computations, the calculations required could become quite involved. Widespread use of the index would be expected to be in herds which obtain periodical listings of individual cow records. A computer program to index routinely the cows in cooperator herds would seem to be a worthwhile enterprise for universities or A.I. organizations concerned with young sire programs. Such a program could serve as a pilot project which could eventually be applied to all tested herds.

This investigation was intended primarily to develop index weights which could be applied in dairy breeding programs to improve milk production. The index has furnished little evidence to answer fundamental questions on breeding theory. Many perennial questions have been raised anew during the course of this investigation. These questions include:

1. What are the major sources of environmental correlations between various related groups of cows?
2. What is the nature of the apparent difference in reliability of the different lactations of a cow?
3. Do different genes affect production in different lactations?
4. How is milk yield related to longevity, fertility, and rate of maturity?
5. What genetic and management factors are responsible for herd by sire interactions?
6. Do deviated records have inherent biases or require special analytic procedures?

Other questions whose answers will lead to more useful indexes and sounder breeding practices would relate to:

1. Curvilinear analysis of production data.
2. Non-additive gene action affecting productive traits.
3. Effects of various mating systems on productive traits.
4. Genetic correlations among various traits.
5. Applicability of age correction factors for genetic studies.

CHAPTER VI

SUMMARY

Selection indexes for milk production in Holstein cattle using information from close relatives were developed and tested in various populations of cows recorded in Michigan DHIA.

Records of lactations measured as deviations from the annual herd average were used to choose the appropriate measure of milk production to use in developing a selection index. Linear multiple regression equations were used to predict the daughter's deviation in first lactation from herd average in using various records of the cow as independent variables. The simple correlations of the cow's first record with the first record of the daughter were 0.149 for 904 Guernsey cows and their daughters, and 0.256 for 1,526 Holstein cows and their daughters. The correlations of the cow's later records were much smaller in both breeds. The partial regression coefficients indicated that nearly all of the emphasis among records of the cow should be placed on the cow's first record to predict the superiority or inferiority of the first record of the daughter. Multiple correlation coefficients indicated that averages of either the first two or the first three records of the cow were poorer predictors of the daughter's first record than was the cow's first record alone.

Selection indexes to predict with maximum accuracy the general breeding value of individual Holsteins for milk production were developed using 7,638 deviations of first lactations from herd averages in a variety of combinations of the cow, her dam, her daughters, and her half-sisters. The heritability used was 0.246 which was derived from the regression of paternal sisters on the cow. Other estimates of heritability (with larger sampling variances) ranged from 0.123, derived from intra-sire correlation, to 0.436, derived from intra-dam correlation.

The records of a cow's dam and maternal sisters only slightly increased the accuracy of estimating her genotype providing the cow had an own record. Daughters and paternal sisters added considerably to the accuracy of estimating her genotype. The multiple correlation of the index with the cow's genotype ranged from 0.50 to 0.73 depending on the kinds and amounts of information available from relatives.

Multiple correlation coefficients for individuals without an own record or offspring (heifers and young bulls) varied from 0.12 for one half-sister to 0.55 for many relatives. In estimating the genotype of a young bull that is sired by a well proven sire, the usefulness of information on the maternal grandparents is limited to the dam's paternal sisters if their numbers are sufficient. The maternal granddam and the dam's maternal sisters are of no apparent value.

Selection by index was compared with mass selection on the cow's own record in a test population of 429 Holstein cows and their 498 daughters. The first record of an unselected daughter was correlated with the cow's index and also with the cow's own first record. The resulting correlations of 0.166 with the index and 0.140 for the cow's record indicated an increase in accuracy near 19 per cent in favor of index selection.

The index appeared to be a practical method to increase genetic progress for milk production especially to select potential dams of future sires.

REFERENCES

REFERENCES

BARR, G. R. Selecting Young Dairy Bulls on Differences Between Relatives and Their Contemporaries. Ph.D. Thesis, Iowa State University Library, Ames, Iowa. 1962.

BERESKIN, B. and HAZEL, L. N. The Role of Herd Averages in Dairy Sire Evaluation. Unpublished Mimeograph. Iowa State University, Ames Iowa. 1962.

BERRY, J. C. Reliability of Averages of Different Numbers of Lactation Records for Comparing Dairy Cows. J. Dairy Sci., 28: 355. 1945.

BERRY, J. C. and LUSH, J. L. High Records Contrasted with Unselected Records and with Average Records as a Basis for Selecting Cows. J. Dairy Sci. 22:607. 1939.

BRUMBY, P. J. The Relationship of Inheritance and Environment to Herd Production. Ruakura Farmers' Conference Week, Hamilton, New Zealand. Proceedings: 188. 1959.

FALCONER, D. S. Introduction to Quantative Genetics. The Ronald Press Co., New York. 1960.

FREEMAN, A. E. Genetic Relationships among the First Three Lactations of Holstein Cows. (Abstr.) J. Dairy Sci. 43:876. 1960.

GAALAAS, R. F. and PLOWMAN, R. W. Relationship between Longevity and Production in Holstein-Friesian Cattle. J. Dairy Sci. 46:27. 1963.

HARVEY, W. R. and LUSH, J. L. Genetic Correlation between Type and Production in Jersey Cattle. J. Dairy Sci. 35:199. 1952.

HAZEL, L. N. The Genetic Basis for Constructing Selection Indexes. Genetics 28:476. 1943.

HAZEL, L. N. and LUSH, J. L. The Efficiency of Three Methods of Selection. J. Heredity 33:393. 1942.

HENDERSON, C. R. and CARTER, H. W. Improvement of Progeny Tests by Adjusting for Herd, Year, and Season of Freshening. (Abstr.) J. Dairy Sci. 40:638. 1957.

HENDERSON, C. R., CARTER, H. W. and GODFREY, J. T. Use of the Contemporary Herd Average in Appraising the Progeny Tests of Dairy Bulls. (Abstr.) J. Animal Sci. 13:959. 1954.

HICKMAN, C. G. and HENDERSON, C. R. Components of the Relationship between Level of Production and Rate of Maturity in Dairy Cattle. J. Dairy Sci. 38:883. 1955.

JOHANSSON, I. The First Lactation Yield as a Basis for Selection as Compared with the Second and Third Lactations. Proc. Brit. Soc. Animal Prod. :102. 1955.

JOHANSSON, I. Genetic Aspects of Dairy Cattle Breeding. Univ. Illinois Press. Urbana, Ill. 1961.

JOHANSSON, I. and HANSSON, A. Causes of Variation in Milk and Butterfat Yield of Dairy Cows. Kungl. Lantb. Tidersk. 79, No. 6-1/2:1. 1940.

JOHNSON, L. A. and CORLEY, E. L. Heritability and Repeatability of First, Second, Third and Fourth Records of Varying Duration in Brown Swiss Cattle. J. Dairy Sci. 44:535. 1961.

LEGATES, J. E. and LUSH, J. L. A selection Index for Fat Production in Dairy Cattle Utilizing the Yields of the Cow and her Close Relatives. J. Dairy Sci. 37:744. 1954.

LUSH, J. L. Family Merit and Individual Merit as a Basis for Selection. Part 1. Amer. Naturalist 81:362. 1947.

LUSH, J. L. and STRAUS, F. S. The Heritability of Butterfat Production in Dairy Cattle. J. Dairy Sci. 25:975. 1942.

McGILLIARD, L. D. Usefulness of the Herd Average in Estimating Breeding Values of Dairy Cattle. Ph. D. Thesis. Iowa State University Library, Ames, Iowa. 1952.

McGILLIARD, L. D. Unpublished Data. Michigan State University. East Lansing, Michigan. 1962.

MILLER, R. H. and McGILLIARD, L. D. Relations between Weight at First Calving and Milk Production During the First Lactation. J. Dairy Sci. 42:1932. 1959.

PARKER, R. J. The Relationship between First Lactation Production, Size, and Type and Subsequent Performance in Holstein-Friesian Cattle. M. S. A. Thesis. University of Toronto Library, Toronto, Ontario.

PIRCHNER, F. and LUSH, J. L. Genetic and Environmental Portions of the Variation among Herds in Butterfat Production. J. Dairy Sci. 42:115. 1959.

PLUM, M. Causes of Differences in Butterfat Production of Cows in Iowa Cow Testing Associations. J. Dairy Sci. 18:811. 1935.

PUTNAM, D. N., BOWLING, G. A. and CONKLIN, C. T. The use of First Records Versus Average of All Records in Dam-Daughter Comparisons when Proving Sires. J. Dairy Sci. 26:967. 1943.

RENDEL, J. M., ROBERTSON, A., ASKER, A. A., KHISHIN, S. S. and RAGAB, M. T. The Inheritance of Milk Production Characteristics. J. Agric. Sci. 48:426. 1956.

RENNIE, J. C. and BREMNER, J. Relationship of Levels of Production between First Lactation Records and Subsequent Records Made by Jersey Cattle in Canada. Paper presented at the Fourth International Conference of the World Jersey Cattle Bureau, Cambridge, England. 1961.

ROBERTSON, A. and KHISHIN, S. S. The Effect of Selection for Heifer Yield on the Production Level of Mature Cows. J. Agric. Sci. 50:12. 1958.

ROBERTSON, A. and McARTHUR, A. T. G. Genetic Differences between Bull Breeding Herds. Proc. Brit. Soc. Animal Prod. 94. 1955.

ROBERTSON, A. and RENDEL, J. M. The Performance of Heifers Got by Artificial Insemination. J. Agric. Sci. 44:184. 1954.

SPECHT, L. W. Rates of Improvement by Progeny Testing in Herds of Various Sizes. Ph. D. Thesis. Michigan State University Library, East Lansing, Michigan.

SKJERVOLD, H. and ØDEGARD, A. K. Estimation of Breeding Value on the Basis of the Individual's Own Phenotype and Ancestor's Merits. Acta Agric. Scand. 9:341. 1959.

TABLER, K. A. and TOUCHBERRY, R. W. Selection Indices Based on Milk and Fat Yield, Fat Per Cent, and Type Classification. J. Dairy Sci. 38: 1155. 1955.

TABLER, K. A. and TOUCHBERRY, R. W. Selection Indices for Milk and Fat Yield of Holstein-Friesian Dairy Cattle. J. Dairy Sci. 42:123. 1959.

TUCKER, W. L. and LEGATES, J. E. Effective Use of Herdmates in Dairy Sire Evaluation. Paper presented at the Fifty-Fourth Annual Meeting of the Amer. Soc. Animal Prod. Chicago, Illinois. 1962.

VANVLECK, L. D., O'BLENESS, G. W. and HENDERSON, C. R. Comparison of Procedures Used for Evaluating Sires Used in Artificial Insemination. J. Dairy Sci. 44:708. 1961.

VANVLECK, L. D., HEIDHUES, T. and HENDERSON, C. R. Analysis of Deviations of Dairy Records from Different Contemporary Averages. J. Dairy Sci. 44:269. 1961.

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 03070 9616