This is to certify that the

thesis entitled

Contributions of Herd Information to Selection of Bull-Dams and to Prediction of Young Bull Transmitting Ability

presented by

Ann Elizabeth Hayes

has been accepted towards fulfillment
of the requirements for

M.Sc. degree in Animal Science

Major professor

Date June 18, 1987

CONTRIBUTIONS OF HERD INFORMATION TO SELECTION OF BULL-DAMS AND TO PREDICTION OF YOUNG BULL TRANSMITTING ABILITY

By

Ann Elizabeth Hayes

A Thesis

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

MASTER OF SCIENCE

Department of Animal Science

1987

ABSTRACT

Contributions of Herd Information to Selection of Bull-dams and to Prediction of Young Bull Transmitting Ability

By

Ann Elizabeth Hayes

Herd characteristics were examined for potential use in young bull selection. Michigan herds involved in young bull selection and sampling were compared to all herds on test in Michigan. Young bulls had been selected from herds with greater average milk production, greater milk variation, and lower genetic merit than other herds. Herd information was examined for use in prediction of young bull transmitting ability. Milk standard deviation was significant in prediction of young bull transmitting ability, while other herd characteristics were not. Young bulls were grouped by milk variation of herd of birth. Young bulls from low-variance herds were genetically superior to young bulls from high-variance herds. Predictors of young bull transmitting ability were accurate in low-variance herds, while biased in high-variance herds. Young bulls should not be selected from herds with high intraherd milk variation.

ACKNOWLEDGMENTS

I would like to thank the following people for their guidance and assistance in the development of this thesis:

Dr. I.L. Mao
Dr. T.A. Ferris
Dr. W.T. Magee
Dr. C.E. Meadows.

I would also like to thank Mr. Roger Hoyt of Select Sires Inc. and Mr. Terry Speer of the Michigan Dairy Herd Improvement Association for their cooperation in obtaining data for this research.

I especially thank my husband, Stephen, my family, and my friends. Their encouragement and love kept me going.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Genetic progress in desirable traits through the planned selection and breeding of animals, is slow for species with long generation intervals such as dairy cattle. Unlike environmental improvements, genetic improvements made through shifts in population gene frequencies are permanent. Tremendous advances have been made in the genetic improvement of dairy cattle for milk yield, and the North American Holstein is now recognized internationally for its ability to produce milk.

Despite these genetic gains and the recognized superiority of the North American Holstein, actual genetic gain has not matched the genetic gain which is theoretically possible. Decreased generation intervals and increased accuracy and intensity of animal selection will improve rate of genetic gain. These improvements take place through four paths of selection; sires of sires, sires of dams, dams of sires, and dams of dams.

Through widespread use of artificial insemination, one bull can inseminate thousands of females. With many daughters distributed in many herds, the bull's genetic merit can be very accurately estimated. Therefore, a high degree of selection intensity and accuracy exists in the sires of sires path and in the sires of dams path.

In a commercial dairy herd, nearly all heifer calves are raised as replacements and there is frequently little opportunity for voluntary culling and selection of females.

Rate of involuntary culling can be reduced through improved management. However, until embryo transfer becomes economically feasible for the average farmer, little improvement can be made to selection accuracy and intensity in the dams of dams path.

Improvements in the rate of genetic gain may be accomplished through improvement in selection of young bulls for sampling, particularly through improved selection of bull-dams.

Young bull transmitting ability is frequently predicted by estimated transmitting ability (ETA), which is ½ sire's transmitting ability and ½ dam's transmitting ability, or by pedigree index (PI), which is ½ sire's transmitting ability and ¼ maternal grandsire's transmitting ability. A young bull's transmitting ability may be estimated from the records of his daughters. This estimate is called the young bull's predicted difference (PD). Due to the random nature of Mendelian sampling, an individual's genetic merit can not be predicted very precisely.

The formula for the standard deviation of prediction error is: $\sigma_e = \sqrt{1-r_{gg}^2}\ \sigma_g$, where r_{gg}^2 is the squared coefficient of correlation between the animal's predicted breeding value and the true breeding value and σ_g is the genetic standard deviation for milk production.

Assuming a genetic standard deviation for milk production of 1,125 pounds and maximum theoretical correlations between predictions of young bull transmitting ability and young bull PD, the standard deviation of prediction error for ETA is $\sqrt{1-(.71)^2(1,125)}$ = 792 pounds of milk. The standard deviation of prediction error for PI is $\sqrt{1-(.56)^2(1,125)}$ = 932 pounds of milk.

This means that in predicting young bull PD from ETA, we will be in error by more than 792 pounds in one third of the bulls. If we use PI to predict PD, then we will be in error by more than 932 pounds in one third of the young bulls. As stated by Henderson,

> "This may seem discouraging, but there is nothing we can do about it unless we discover some accurate method of evaluation involving a measurement or set of measurements on the young sire himself. However, the picture appears much brighter if we look at the expected mean of breeding values of a group of young sires."[1]

The expected mean of breeding values for a group of young bulls equals the mean of their predicted breeding values, if predictions are unbiased and we are only confronted with Mendelian sampling. Therefore, when matings are contracted and young bulls selected for progeny testing, emphasis must be placed on accurate selection of the group rather than on individual selection.

Current practices in young bull selection place much emphasis on matings contracted to produce young bulls. Most bull studs contract matings of superior cows and bulls, relying on estimates of individual breeding values. Information on bull-dam herds is not taken into

---

[1] C.R. Henderson, "Selecting the Young Sire to Sample in Artificial Insemination," (<u>J. Dairy Sci.</u> 47:439-441, 1964), p. 439.

consideration when matings are contracted. The objective of this research was to examine ways in which herd information might improve young sire selection programs.

The initial hypothesis was that a measure of herd genetic merit might improve prediction of a young bull's transmitting ability. If proved true, this measurement could better direct sire selection committees to herds in which contract matings should be arranged, aiding in selection of genetically superior young bulls and increasing annual genetic gain.

The research was divided into three phases of investigation. The first investigation compared Michigan herd groups involved in young sire selection and sampling programs to all herds on test in Michigan. This was done to examine the success of current selection programs in locating genetically superior herds and bull-dams and to establish whether differences among herd groups existed.

The second investigation examined whether herd information might improve prediction of a young bull's transmitting ability, beyond what can be predicted from pedigree information alone. Despite the primary interest in herd genetic merit, other herd characteristics were also considered. As stated by John W. Tukey,

> Restricting oneself to the planned analysis-failing to accompany it with exploration-loses sight of the most interesting results too frequently to be comfortable. As all detective stories remind us, many of the circumstances surrounding a crime are accidental or misleading. Equally, many of the indications to be discerned in bodies of

> data are accidental or misleading. To accept all appearances as conclusive would be destructively foolish.... To fail to collect all appearances...would, however, be gross misfeasance.[2]

While failing to support the initial hypothesis, results from the second investigation revealed that intraherd standard deviation for milk production was a significant factor in prediction of young bull transmitting ability. Therefore, the third investigation grouped young bulls by intraherd milk variance of the herd from which they had been selected. Comparisons were made among variance groups on merit of contracted matings and accuracy of young bull predictions.

---

[2] John W. Tukey, Exploratory Data Analysis. (Reading, Massachusetts: Addison-Wesley Publishing Company, 1977), p.3.

2. LITERATURE REVIEW

2.1 INTRODUCTION: Man has long recognized the variation in biological characteristics which exists among animals in a species. By selecting and breeding animals superior in traits important to his own welfare, man increased the usefulness of that species. This discourse has its roots then in what has always been a primary concern of animal husbandry: the selection and breeding of superior animals. The scope of discussion will be limited to dairy cattle. However, a brief history and review of genetic theory will be given as background for the research to be presented.

The foundation for genetic theory was laid by Gregor Mendel. His research on peas was presented to the Brünn Natural Science Society in 1865 and was published in their proceedings in 1866 (Dunn, 1965). Mendel was the first to understand and describe the manner in which genetic material passes from parent to offspring. However, his work went unnoticed and its importance was unrecognized until its rediscovery in 1900.

Other scientists of the period were also engaged in genetical research. Francis Galton successfully described the statistical relationships among relatives through his work on inheritance of stature in humans. His findings were published in 1889 and his work was continued by Karl Pearson (Mather, 1949). Although these investigations failed to elucidate the laws of inheritance, they were important in the development of biometry, the application of statistical methods to biological problems.

The rediscovery of Mendel's work in 1900 led to a rivalry between the Biometricians and the Mendelians.

> The original discordance seems to have arisen because neither side understood the full implications of Mendel's fundamental separation of determinant and effect, of genotype and phenotype. The biometricians seem to have regarded continuous somatic variation as implying continuous genetic variation, and the Mendelians seem to have considered discontinuous genetic variation as incompatible with anything but obviously discontinuous somatic variation....The understanding of continuous variation awaited a fusion of the two methods of approach, the genetical and the biometrical, for each supplied what the other lacked. The one gave us the principles on which the analysis must be based; the other showed the way in which to handle continuous variation, the way of representing it in a form which made fruitful analysis possible.[3]

Work by several scientists paved the road for the eventual merging of these two schools. In 1909, Johannsen published findings which showed both heritable and non-heritable factors were responsible for variation in weight of bean seeds. Work by Nilsson-Ehle, also published in 1909, showed that several factors with small, similar effects acted cumulatively to determine grain color in wheat (Mather, 1949).

These findings were further developed, largely by East and by Fisher, into the Multiple Factor Theory (Mather, 1949). East and his collaborators showed that the

---

[3] K. Mather, Biometrical Genetics, The Study of Continuous Variation. (New York: Dover Publications, Inc., 1949), p.2.

continuous variation of several traits in both tobacco and maize could be explained by the cumulative effect of many factors. In his paper published in 1918, entitled "The Correlation Between Relatives on the Supposition of Mendelian Inheritance", Fisher gave the final evidence necessary for the merging of the two schools, and the science of Biometrical Genetics was born (Moran, 1966).

Tremendous advances have since been made in the study of Biometrical Genetics. Statistical techniques have made possible the study of traits much more complex than those which Mendel studied. However, Mendelian principles explain the transmission of genetic material from parent to offspring and necessarily form the basis of all genetic studies.

2.2 MENDELIAN GENETICS: Through his work on garden peas, Mendel formulated the theory of particulate inheritance. This theory states that units of inheritance are discrete and pass from parent to offspring. These units of inheritance, or genes as they were later termed, retain their integrity as they pass from one generation to the next. In this way, traits not expressed and seemingly lost from the population may reappear in later generations. As stated by J.L. Lush, "The idea that inheritance is unitary or particulate is as fundamental to genetics as is the concept of the atom to chemistry."[4]

---

[4] J.L. Lush, "The Genetics of Populations". (Ames, Iowa: Animal Breeding Class Notes, 1948), p. 23.

The theory of inheritance prevalent at Mendel's time was that of blending. Scientists thought material which passed from parent to offspring was fluid and that "characters of the parents blend into an intermediate level in the offspring with no apparent segregation in later generations."[5] Mendel, however, stated that genes occur in pairs, with half of the pair coming from each parent. Gene pairs segregate during the formation of gametes, so that a gamete contains only half of the parental genetic material. The full gene complement is restored by union of male and female gametes. This halving of genetic material is Mendel's law of segregation.

Mendel's second principle, the law of independent assortment, states that gene pairs segregate independently. Therefore, genes contained in a gamete are a random half of the parental gene pairs. This law was later altered to accommodate the discovery of chromosomes and linkage.

Mendel was fortunate in that the traits with which he chose to work are simply inherited. Each trait is controlled by a single pair of genes in which one allele is dominant to the other. Under these conditions, there are three possible genotypes and only two possible phenotypes for the trait. The phenotypic variation is discontinuous, as observations classify readily into one phenotypic group or the other. Mendel performed certain crosses of pea lines which were homozygous for the trait of interest, and by

---

[5] Robert C. King, A Dictionary of Genetics. (New York: Oxford University Press, 1968), p. 29.

counting the frequency with which alternate forms of the trait were expressed, he mathematically deduced the laws of inheritance.

Traits which are simply inherited, like those that Mendel studied, are called qualitative traits. Most qualitative traits are not of economic importance. There are few alternative forms which the trait may express, and ratios of their relative frequencies are predictable. Each gene has a major effect, and substitution of an allele can change the phenotype.

2.3 QUANTITATIVE GENETICS: Most traits of economic importance, like milk yield and other production traits, are quantitative traits. Quantitative traits are controlled by many genes, often referred to as polygenic systems. The basic assumptions of the Multiple Factor Theory, or polygenic systems, are that the effect of any single gene is small; gene effects are similar to each other; these effects act cumulatively; and each gene pair behaves according to Mendel's laws.

As more genes are involved, the number of possible gene combinations increases dramatically. There are 3^n different genotypes possible when n genes are involved. However, for a quantitative trait, gene effects are small and similar in effect, so there are $(2 \times n) + 1$ phenotypic classes possible (Table 1). Unlike qualitative traits, quantitative traits exhibit a continuous variation in phenotype.

Table 1. Number of possible phenotypes and genotypes for a given number of gene pairs.

Genes	Phenotypes	Genotypes
1	3	3
2	5	9
3	7	27
.	.	.
.	.	.
.	.	.
10	21	59,049
.	.	.
.	.	.
.	.	.
n	(2xn)+1	3^n

Another important characteristic of polygenic systems is the influence of the environment on phenotypic expression. The phenotype of an individual results from both its genetic composition and the environment, so that both heritable and non-heritable agents are responsible for the continuous variation in phenotype.

It must be emphasized, however, that the Multiple Factor Theory is based on assumptions. Mendelian methods can not prove that genes controlling quantitative traits are located on chromosomes in the cell nucleus and follow the laws of Mendelian inheritance. However, other types of tests have been used and there is good evidence that these assumptions are correct (Mather, 1949).

More recently, scientists have questioned whether all genetic material is contained in the cell nucleus and is inherited according to Mendelian principles. Research has been done on the occurrence of cytoplasmic inheritance. While this topic of research is still fairly new, there is evidence that cytoplasmic inheritance occurs. Bell et. al. (1985) reported that 2% of the variation in milk yield and 3.5% of the variation in fat percentage is due to cytoplasmic effects, and that cytoplasmic origin is a significant source of variation in production traits of dairy cattle. However, the Multiple Factor Theory explains inheritance of quantitative traits sufficiently well, and until more evidence is obtained in support of cytoplasmic inheritance, there seems little reason to alter existing theory.

2.4 BIOMETRICAL GENETICS: Differences between qualitative traits and quantitative traits necessitate different methods of analysis.

> [The Mendelian Method] depends for its success on the ability to assign the individuals to classes whose clear phenotypic distinctions reveal the underlying genetic differences. A certain amount of overlap of the phenotypic classes can be accomodated by the use of genetical devices; but where the variation in phenotype is fully continuous in its frequency distribution, so that no such classes can be defined, the method cannot be used. A different approach is required, one based on the use of measurement rather than frequency.[6]

This biometrical approach describes a quantitative trait by parameters of its frequency distribution. The phenotypic variance of the trait may be partitioned into portions due to heritable and non-heritable causes. Since most traits of economic importance are quantitative traits,

---

[6] K. Mather, Biometrical Genetics, The Study of Continuous Variation. (New York: Dover Publications, Inc., 1949), p.vii.

the biometrical approach is very important in animal breeding.

Prior to the development of Biometrical Genetics, selection of animals was based solely on the animals phenotype. To achieve permanent changes in a species, however, gene frequencies of the population must change. This requires an ability to distinguish among animal genotypes, selecting animals genetically superior in the trait of interest. Genetic differences among animals can be measured by biometrical techniques, allowing genetically superior animals to be identified.

Extension of Biometrical Genetics to the study of animal populations was pioneered by Sewall Wright. In his 1921 paper, entitled "Systems of Mating", Wright described the biometrical relations between parent and offspring; effects of inbreeding on the genetic composition of a population; assortative mating based on somatic resemblance; effects of selection; and general effects of mating systems on population composition (Wright, 1958).

While much of the foundation for Animal Breeding was laid by Wright, J. L. Lush was truly the father of modern Animal Breeding. Much of his research was devoted to the problem of estimating an animal's genetic merit and, therefore, its use in a breeding program. Although Lush was a prolific researcher and was author or coauthor to numerous papers, his greatest influence was perhaps as a teacher. As stated by A.E. Freeman at a symposium honoring Dr. Lush,

> Graduate teaching began a new era in Dr. Lush's life where his influence began to radiate through his graduate students. This influence continued at an increasing rate until the 1960's and was magnified by his academic grandchildren and great-grandchildren. It would be difficult to overemphasize the influence of his students on United States and world agriculture.[7]

From the initial work of Lush on prediction of an animal's genetic merit, C.R. Henderson and A. Robertson further developed genetic evaluation procedures. The computational capacity and speed of computers allowed development of advanced evaluation procedures. By their efforts, computer techniques for the evaluation of dairy bulls were developed. Much of the recent progress in milk production may be attributed to these men (Van Vleck et. al. 1987).

The introduction of artificial insemination (AI) programs made estimation of an animal's genetic merit tremendously important. Through use of AI, one male may inseminate thousands of females. It is, therefore, very important to accurately evaluate the male's genetic merit so that only superior animals are selected for use in AI programs.

In cattle, AI has been successfully coupled to the use of frozen semen, while "frozen semen has not been used as successfully in other species."[8] This combination makes a

---

[7] A.E. Freeman, "Genetical Statistics in Animal Breeding," Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. J.L. Lush. (Champaigne, Illinois: American Society of Animal Science, 1973), p. 5.

[8] R.H. Foote, "Cryopreservation of Spermatozoa and Artificial Insemination: Past, Present, and Future," (J. Androl. 3:85-100, 1982), p. 93.

powerful tool for rapid genetic improvement. However, this success has been largely limited to the dairy industry. Successful use of AI requires good heat detection. In the beef industry, many cow operations are on open range making heat detection and use of AI difficult. Another important key to a successful AI program is a good record system. Records must be kept so that superior animals may be identified.

2.5 GENETIC EVALUATION: The primary difficulty in genetic evaluation of an animal is the removal of environmental influences. Environment can affect expression of a quantitative trait either posititvely or negatively. While individual envoronmental effects can not be removed, effects which are common to a group of animals, such as a herd or season effect, may be standardized. Adjustment for group effects allows animals to be compared on a more equal basis.

The objective in genetic evaluation is to determine the animal's worth as a parent. It is clear from Table 1 that for a given number of genes many more genotypes than phenotypes are possible, and that many genotypes may express the same phenotype.

In a polygenic system, each gene has a small effect and these gene effects act cumulatively. The sum of these effects is called the animal's additive genetic value. However, genes at the same locus may interact, and one gene may be dominant to another. The sum of paired gene effects is called the animal's dominance genetic value. Additionally, genes across loci may interact so that genes at one loci affect expression of genes at another loci. This phenomenon is called epistasis, and the sum of epistatic effects is called the animal's epistatic genetic value. The sum of additive, dominance, and epistatic effects is the animal's total genetic value. Since an animal only passes to its offspring one gene from each gene pair and since gene pairs segregate independently, only additive genetic effects are heritable. Therefore, it is necessary to distinguish between an animal's total genetic value and the portion of it which is additive.

The additive genetic value is also called the animal's breeding value. However, a parent transmits to its offspring a random half of its gene pairs and not its full breeding value. The random half transmitted has an expected value equal to one half of the animal's breeding value and is called the animal's transmitting ability. Transmitting ability is the additive genetic value of an average gamete produced by the animal.

Many different procedures are used throughout the world for genetic evaluation of dairy animals. In the United States, the national evaluation of bulls and cows is done by the Department of Agriculture (USDA). Procedures for estimating an animal's transmitting ability generally include information on the individual's own performance. This is impossible in evaluation of dairy bulls for milk production, since it is a sex limited trait. The transmitting ability of a bull for sex limited traits must be estimated from the performance of his female relatives.

A bull's transmitting ability is called his predicted difference (PD) by the USDA, and the formula for estimating PD is (Norman, 1986):

$$PD82 = R(D - MCA + SMC) + (1-R)AM$$

where: **PD82** is the predicted difference of a bull relative to the 1982 genetic base. The average genetic merit of sires of 2-year-old cows calving for the first time in 1982 was set to zero, and all bulls are compared to this base (Wiggans, 1984).

(D - MCA + SMC) is the average daughter superiority of the bull. This entire term is also called the modified contemporary deviation (**MCD**). All milk records are adjusted for lactation length, age at calving, and month of calving. Daughter records (**D**) are deviated from the average of their modified contemporaries (**MCA**). Modified contemporaries are cows of similar age, calving in the same herd and season. Sire merit of contemporaries (**SMC**) must be added to the deviation to account for genetic level of contemporaries. Dam merit of contemporaries is not added to the deviation because this effect is assumed to cancel with merit of the daughter's dam (Powell, 1984).

R is the repeatability of the bull's estimated **MCD**. Repeatability is the squared correlation between estimated genetic value and true genetic value.

AM is the ancestor merit of the bull. This term provides both pedigree information and adjustment for genetic trend. Ancestors specifically considered are sire and maternal grandsire.

(1-R) is the complement of the bull's repeatability. As a bull has more daughters, his repeatability increases. Eventually very little emphasis is given to a bull's **AM**, and his PD is estimated almost entirely from progeny performance, as measured by the **MCD**.

Under USDA procedures, the transmitting ability of a cow is called her cow index (CI). The formula for estimating CI is (Powell, 1985):

$$CI82 = .5[w(\overline{MCD}_{cow} + ADC_{byc}) + (1-w)(PD82_{sire} + CI82_{dam})]$$

where: CI82 is the cow index of a cow relative to the 1982 genetic base. It is the same genetic base as defined for bull evaluation.

$(\overline{MCD}_{cow} + ADC_{byc})$ is the cow's superiority. $\overline{MCD}_{cow}$ is obtained by the same procedure used to estimate MCD in bull evaluation, except a cow's own record rather than daughter records are deviated from contemporaries, and MCD's from all lactations are averaged. Unlike bull evaluations, cow's dam merit is not expected to cancel dam merit of contemporaries, so an adjustment for average dam merit of contemporaries according birth year of the cow (ADC_{byc}) is added.

w is a weighting factor given to the dam's estimated superiority. It is a function of the cow's number of lactations and the summed repeatabilites on her sire and dam. As a cow has more records, the weighting factor increases. An average cow with one lactation record and 60% summed repeatability on her parents has a weighting factor of .16, while a cow with five lactation records and the same summed repeatability has a weighting factor of .29.[9]

$(PD82_{sire} + CI82_{dam})$ provides information on the genetic merit of the cow's sire and dam and estimates the cow's breeding value from pedigree information.

---

[9] Rex L. Powell, "Cow evaluation procedures," Fact Sheet H-2. (Agricultural Research Service, National Cooperative Dairy Herd Improvement Program, 1985), p. 2.

(1-w) is the complement of w. As the summed repeatability of the parents increases, w decreases. An average cow with five lactation records and 60% summed repeatability has a weighting factor of .29, while a cow with five lactation records and 150% summed repeatability has a weighting factor of .23.[10] w never becomes very large, so CI will always be estimated from both the cow's own performance and the genetic merit of her parents.

.5 is multiplied by the entire formula to give the estimated transmitting ability, or CI, of the cow.

Genetic evaluations allow animals to be compared and ranked by their transmitting abilities. Accurate ranking is necessary so that superior animals can be selected to produce the next generation. The primary goal of selection among dairy animals is genetic improvement of the population for milk production. As seen in Table 2, milk production has increased by more than 4,000 kilograms in the last 80 years (Voelker, 1985). However, increased milk production may not be entirely attributed to successful genetic selection. A portion of this improvement is due to better management practices. The effectiveness of a breeding program is measured by the annual rate of genetic gain in the population.

---

[10] Ibid.

Table 2. Average production of all U.S. cows since 1906 (Voelker, 1985).

Year	Milk (kg)
1906	1,640
1910	1,709
1920	1,822
1930	2,053
1940	2,101
1950	2,415
1960	3,195
1970	4,430
1980	5,404
1985	5,904

2.6 THEORETICAL GENETIC GAIN: Early work on the theoretical estimation of genetic gain was first done by Dickerson and Hazel (1944). At that time both artificial insemination and progeny testing were new practices, and there was enthusiasm over the improved accuracy in genetic selection provided by these tools. However, Dickerson and Hazel recognized that additional factors must be included when considering genetic progress. They developed an equation which determines annual genetic gain from both the genetic superiority of animals selected to become parents and the generation interval. Genetic superiority was defined as the selection differential multiplied by the accuracy of selection.

Rendel and Robertson (1950) further refined this formula as shown in [1]. The formula consists of four pathways through the bulls and cows selected to produce the next generation (Table 3).

Table 3. Four sources of genetic gain in dairy cattle (Rendel and Robertson, 1950).

Potential Parents	Actual Parents	Genetic Superiority of Selected Parents	Average Generation Interval	Potential Parents for next Generation
Bulls	Sires to produce bulls	I_{BB}	L_{BB}	→ Bulls
	Sires to produce cows	I_{BC}	L_{BC}	→ Cows
Cows	Dams to produce bulls	I_{CB}	L_{CB}	→ Bulls
	Dams to produce cows	I_{CC}	L_{CC}	→ Cows

$$[1] \qquad \frac{\text{Genetic Gain}}{\text{Year}} = \frac{(I_{BB} + I_{BC} + I_{CB} + I_{CC})}{(L_{BB} + L_{BC} + L_{CB} + L_{CC})}$$

Assuming natural service bulls and minimum generation intervals, Rendel and Robertson (1950) predicted a maximum rate of genetic gain of 1% annually in a closed herd. Realizing that most herds would not achieve the assumed optimum conditions, they predicted an annual rate of gain of .6% for an average herd.

In another paper, Robertson and Rendel (1950) expanded their predictions from rate of gain possible within a herd to rate of gain possible in a population. The use of artificial insemination rather than natural service was also incorporated. Despite the increased generation interval incurred by progeny testing, the increased accuracy of selection allowed greater genetic progress. They estimated an annual genetic gain of 1.69% for a 2000 cow population and 2.05% for a 10,000 cow population. They also reported the percent of annual genetic gain which could be attributed to each of the four paths: I_{CC}, 6%; I_{BC}, 18%; I_{CB}, 33%; I_{BB}, 43%. In conclusion, they stated that progeny testing combined with use of artificial insemination provides more rapid genetic gain than possible through other breeding schemes.

In a review of theoretical and actual genetic progress, Van Vleck (1977) stated that most predictions of theoretical genetic gain fall in the range of 1.5%-2.0%. He suggested, however, a slight shift in the proportions of genetic gain attributed to the various paths: I_{CC}, 2%; I_{BC}, 26%; I_{CB}, 32%, I_{BB}, 39%. This change is due to the larger number of inseminations per bull which is now possible. It is important to note that 71-76% of genetic gain is accomplished through the selection of young bulls entering the population, while only 24-29% of genetic gain is accomplished through selection of cows entering the population.

2.7 **ACTUAL GENETIC GAIN**: The amount of genetic gain which has actually been attained can be measured retrospectively from historical data. A method for estimating actual genetic trend was first developed by C. Smith in 1962 (1962). From the continuity of genotypes provided by sires which have progeny born in several years, genetic change over time can be measured.

Smith's method requires a regression coefficient of pooled intra-sire regressions of progeny performance on time. This regression coefficient is subtracted from the regression coefficient of population performance regressed on time. The difference between these regression coefficients measures the genetic trend due to sires. Assuming identical rates of change in the two sexes, the genetic trend of the sires is doubled to obtain genetic trend of the population. Although Smith applied this method to estimation of genetic change in swine, it has also been widely used in dairy cattle.

As more accurate procedures for genetic evaluation were developed, methods for estimating genetic gain improved. A more recent approach for measuring genetic gain is to predict average genetic merit of animals in the population and regress these averages on year. The main weakness of this approach is the relative difficulty in measuring genetic merit of the cow population as compared to the sire population. Hintz et. al. (1978) reported a genetic trend in Holstein sires of 35.8 kg/year by using this method. The validity of doubling sire trend to estimate overall genetic trend was questioned, since genetic trends measured in the cow population were only 26.1 kg/year.

Hintz et. al. (1978) also summarized previously reported estimates of genetic trend. These estimates were highly variable, ranging from 18 to 55 kg. of milk per year in the AI population. However there is difficulty in comparing estimates of genetic trend due to differences in methods used and differences in time periods and populations studied.

Recently, Lee and Freeman (1985) reported an annual genetic change of 49 kg. of milk in the sire population and 40.9 kg. in the cow population. Van Vleck *et*. *al*. (1986) reported similar findings of 47 kg./year in the sire population, 39.5 kg./year in registered cows and 38.1 kg./year in non-registered cows.

2.8 IMPROVEMENTS TO RATE OF ACTUAL GENETIC GAIN:

Although estimates of actual genetic trend have varied, none have equalled the gain which is theoretically possible. Questions as to why theoretical gain has not been achieved must be addressed within the dairy industry. While many of the factors which hinder genetic gain are known, practical methods by which these problems can be corrected are not as obvious.

2.8.1 GENERATION INTERVAL: Decreased generation intervals would increase rate of genetic gain. Pearson (1984) reported a total generation interval of 25 to 27.5 years, with the four paths contributing as follows: $L_{BB}=9$ years, $L_{BC}=6$ to 8 years, $L_{CB}=$ 5.5 years, $L_{CC}=4.5$ to 5 years. Burnside and Kuersten (1985) found that in Canadian bull studs from 1976-1977, L_{CB} was 7.5 years and L_{BB} was 11.7 years. Although Canadian studs succeeded in decreasing L_{BB}

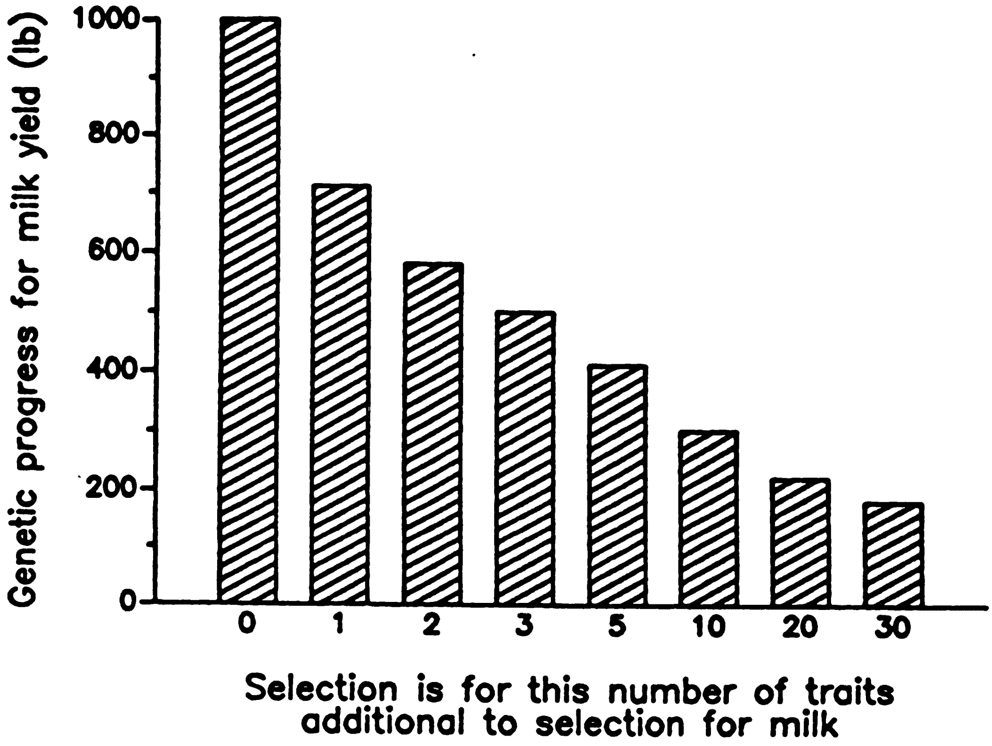
to 9 years by 1984, they urged them to further reduce L_{BB} to 7 years and to reduce L_{CB} to 6 years. In a study on U.S. AI Holsteins, Westell (1984) reported a total generation interval of 30 years, with L_{BB}=9.7 years, L_{BC}=8.5 years, L_{CB}=6.9 years and L_{CC}=4.9 years.

Van Vleck (1986a) stated that generation intervals are longer than necessary, particularly the dam of bull and sire of bull paths. A realistic goal would be to reduce these two paths to 5 years and 7 years respectively and to reduce total generation interval to 24-25 years. Van Vleck further stated that reducing generation interval from 30 years to 24 years would increase the annual rate of genetic gain by 20%.

2.8.2 INTENSITY OF SELECTION: Selection emphasis on traits other than milk production has also hindered the genetic gain theoretically possible for milk production. Dairy farmers frequently are interested in improving other traits in addition to milk production. Some farmers have an interest in type traits because of the income and prestige generated by sale of breeding stock. Others may sell milk based on milk components and have an interest in the fat or protein content of milk. Most farmers also are interested in management traits, such as milking speed, calving ease, disposition, and functional soundness. No single sire is genetically superior in all traits of interest. Therefore, AI studs offer the farmer a large array of sires from which to choose.

However, when selection is made on more than one trait, rate of improvement for any one trait is decreased. Assuming that traits are uncorrelated, the genetic progress made in one trait is proportional to $1/\sqrt{n}$, where n is the total number of traits for which selection is being made (Figure 1).

Figure 1. Relative progress in milk yield if selection emphasis is equal for several traits.
(Schmidt and Van Vleck, 1974)



Another reason why AI studs have many bulls in service simultaneously is that the top bulls can not produce enough semen to service the population. Van Vleck (1986b) stated that too many bulls are in AI use. Further dilution of semen would allow more inseminations per bull and would increase selection intensity. However, the gain achieved in selection intensity must be weighed against the cost of reduced conception rate per service caused by dilution of semen.

2.8.3 ACCUARACY OF SELECTION; Evaluation methods have developed over time to accomodate changes in the population, increased AI usage, and availability of more complete records. Evaluation procedures have also incorporated advances in computer technology, which have allowed for more accurate, less biased evaluation procedures (Van Vleck and Pollak, 1984). The major advancements in evaluation procedures are shown in Table 4.

Table 4. Summary of Sire Evaluation Methods (Van Vleck and Pollak, 1984).

Method Based on Averages	Source
Equal Parent Index	Hansson (1913)
Intermediate Index	Woodward (1922)
American Index	Yapp (1924)
Modified Mount Hope Index	Goodale (1927)
Daughter-Herdmate Difference	Peters (1913)
Percentage of Contemporaries	von Patow (1925)
Daughter-Dam Difference	Denmark (1902) Graves (1926) U.S.D.A. (1935)
Methods Based on Selection Index	
Contemporary Comparison	Johansson and Robertson
(1952)	
Herdmate Comparison	Henderson *et*. *al*. (1954)
Relative Breeding Value	U.S.D.A. (1965)
Modified Contemporary Comparison	U.S.D.A. (1974)
Methods Based on Mixed Model Equations	
Sire Comparison	Henderson (1949, 1966, 1970, 1972)
Sire Comparison with Relationships	Kennedy and Moxley (1975)
Sire Comparisons with Relationships and Adjustments for Sire of Dam	Quass *et*. *al*. (1979) Everett *et*. *al*. (1979)

Development of evaluation procedures has been an evolutionary process, and changes will continue to be made as weaknesses in existing procedures are exposed and computer technologies advance.

<u>2.8.3.1 Fluctuations in Estimates of Individual Animal Breeding Values</u>: Current sire evaluation procedures are highly sophisticated, and considerable confidence may be had in PD estimates of well proven sires. However, there are some problems in existing evaluation procedures. Van Vleck (1986a) stated that while some variaton in PD is expected as a sire has more daughters, one bull in six will have a change in PD greater than one standard error of the initial prediction. Both farmers and sire analysts are concerned when a bull's PD changes markedly from one evaluation period to the next.

Bolgiano <u>et</u>. <u>al</u>. (1979) examined this problem using several bulls whose PD dropped significantly from first to second evaluation. They attributed these changes to several causes: preferential treatment of daughters, non-random mating of unproven sires to cows, and inherent randomness of sampling procedures. They found no evidence that these fluctuations were due to differences in herd size, genetic level of herds, management level of herds, or differences in age of daughters at freshening. However, bulls with significant decreases in PD were considered a problem, while bulls with a significant increase in PD were considered normal. This may have caused difficulties in the study. Van Vleck (1986a) speculated that changes in PD are caused by disproportionate use of bulls in herds with different breeding goals.

2.8.3.2 Evaluations of Cow Breeding Value From First and All Lactations: Cow evaluations are primarily used in selection of bull dams. Problems in cow evaluation procedures will hinder selection of superior bull dams and reduce rate of genetic gain. There is evidence that cow evaluations from first lactation records are more accurate in prediction of son's PD than are evaluations which use additional lactation records.

Rothschild et. al. (1981) found this to be true in prediction of son's PD milk. However, inclusion of additional lactation records in cow index estimation was found to be useful in prediction of son's PD fat percent.

Vinson and White (1982) found that in dams of AI bulls, both milk production and estimated transmitting ability (ETA) increased in later lactations. They suggested this is due to preferential treatment of dams following a superior first lactation. They also found that regressions of son PD on dam ETA from first lactations were consistently higher than regressions on dam ETA from later or all lactations.

These conclusions were again confirmed by Murphy et. al. (1982) who concluded that selection of dams to produce sons for AI testing should be by dam ETA's from first lactations. However, USDA cow evaluations use all available lactation records and not just first lactation records. Since many studs select bull dams from cows within the top 2% on national CI ranking (elite cows), this could cause problems in selection of superior bull dams.

2.8.3.3 Problems Caused by Heterogeneous Intraherd Variances: Most evaluation techniques assume homogeneous intraherd variances. However, heterogeneity of intraherd variances have been well documented. Failure of procedures to account for this may bias genetic evaluations, decreasing selection accuracy and impeding genetic gain.

Hill et. al. (1983) split herds into two groups according to three different criteria. When herds were divided into low and high production groups, variance within sires, variance between sires, and heritability estimates increased from the low production group to the high production group. When herds were split by within-herd production variance, greater variance both within sires, between sires, and higher heritabilities were observed in the high variance group. Heritabilities in the low production group and low variance group were similar, as were heritiabilites in the high production group and high variance group. When herds were split by coefficient of variation, variance within sires and variance between sires both were greater in the high coefficient of variation group. Heritability estimates were also greater in the high coefficient group than in the low coefficient group for yield traits, but not for composition traits.

Mirande and Van Vleck (1985) also studied heritabilites of milk production, fat, and fat test at different levels of herd production. Herds were divided into four production groups. No clear association between heritability of fat test and herd production level was seen. Heritability of both fat and milk yield was smallest at the low production level. Heritability of fat was similar in the two medium production groups and the high production group, while heritability of milk yield was greater in the two medium production groups than in the high production group.

DeVeer and Van Vleck (1986) estimated heritabilites of milk yield and fat test at three production levels by daughter-dam regression. They found heritability estimates for both traits increased with increased production level.

Researchers are particulary concerned with complications that heterogeneous intraherd variances may cause in cow evaluations. Heterogeneity of herd variances may bias cow evaluations, causing selection accuracy of bull dams to be less than desired. Everett et. al. (1982) showed that in herds of equal genetic merit, herds with greater residual variance are more likely to have elite cows than herds with smaller residual variance.

Powell et. al. (1983) also documented this problem. In herds of similar sire merit, herds with higher production contained a larger number of elite cows than herds with lower production. This was attributed to the positive relationship between variance of cow MCD and herd production level, and suggests that cows in high production herds are overevaluated. They stated that, although greater emphasis is placed on MCD in high production herds than in low production herds, cow indexes in these herds should be computed with a larger heritability, and these two factors would act to cancel each other. Whether these factors cancel each other is questionable. Greater heritabilites should increase the weight placed on a cow's modified contemporary deviation. Greater heritabilities and greater empahsis on MCD of cows in high-variance herds seem to act in tandem, not opposed to each other.

2.8.3.4 Proposals to Adjust For Heterogeous Intraherd Variances: Several alternatives have been proposed to handle the problem of heterogeneous intraherd variances in bull and cow evaluations. However, none of these proposals have been accepted as fully satisfactory solutions. Frequently, a scale transformation of the original data is used to equalize heterogeneous variances. Mirande and Van Vleck (1985) examined effects of logarithmic and square root transformations. They reported that with a square root transformation, estimates of residual standard deviations by year and production level follow the same pattern as estimates on the untransformed scale. Greater residual standard deviation was still seen at higher production levels, although the square root scale greatly reduced differences between production groups. The logarithmic scale, however, reversed the pattern seen on the untransformed scale. Hence, they stated that use of a logarithmic transformation would overevaluate cows in low producing herds and underevaluate cows in high producing herds.

Van Vleck et. al. (1985) carried out a similar study, but included an additional transformation to the .4 power. The power transformation gave results similar to those of the square root transformation. Both the .4 and square root transformations were more successful in equalizing residual variation across management groups than the log transformation, but the log transformation was more succesful in equalizing variation over time. Van Vleck concluded that more research is needed to find an appropriate transformation which would equalize residual variation and at the same time properly weight records according to heritability for that environment.

Brotherstone and Hill (1986) also disputed the appropriateness of a log transformation. In additon to heterogeneity among herds in variance of milk production, they found heterogenity among herds in the coefficient of variation for milk production. Given heterogeneity in coefficient of variation, a log transformation would be insufficient. They maintained that data must also be scaled to equalize coefficients of variation.

Lofgren et. al. (1985) proposed alteration of the cow index formula rather than transformation of the data to account for differences among herd intraherd variances. Eight different methods for calculating CI were used. The difficulty with this approach was that no one CI formula was found to be uniformly best across different dairy breeds. Within breeds, cow index formulas also differed considerably when used for prediction of performance of different offspring groups. Prediction of daughter performance required a different cow index formula than prediction of son performance.

The need for genetic evaluations to account for heterogeneous genetic and environmental intraherd variances is clear. Failure to account for these differences has caused problems in genetic evaluation of dairy animals, particularly cow evaluation. Van Vleck (1986a) suggested that improvements may lie in extension of the animal model to multiple trait analysis, with milk production at different levels of management considered as separate traits. However, there are many compuational and practical difficulties in this procedure, and a real solution to the problem is still in the future.

2.9 IMPROVEMENTS TO BULL DAM SELECTION: The dam to bull path in dairy cattle selection comprises 32% of the annual genetic gain. However, the accuracy of selection in bull dams has not been as great as desired. Since this path makes such a large contribution to genetic gain, it is important to capitalize on its potential. Additionally, bull studs operate on limited resources and can not afford to test bulls from dams with unreliable evaluations. While statistical solutions are being sought to adjust for unequal intraherd variances in evaluation procedures, other means of improving bull dam selection must also be explored.

In a symposium on young sire selection, C.R. Henderson (1964) suggested an entirely different approach to the problem of bull dam selection. The alternative proposed is to first select a group of herds with desirable characteristics concerning breeding program, production level, record keeping, and management level, and then to select bull dams from within these herds.

More recently, another suggestion regarding selection of young bulls for AI sampling has been made. A young bull's pedigree index (PI) is often used to predict his transmitting ability before he is progeny tested. PI weights information on a young bull's sire PD and maternal grandsire PD, completely ignoring the dam's cow index. In arranging matings to produce young bulls for sampling, many studs set minimum standards on PI. This is done to guard against biased cow evaluations and the subsequent selection of inferior bull dams. However, cows with accurate evaluations are also disregarded, causing potentially valuable matings to be missed. Van Vleck (1986a) stated that it would be more effective to develop methods of identifying cows with potentially biased evaluations. AI studs could then identify cows with suspect evaluations and reject them from consideration as bull dams, rather than disregard all cow evaluations.

Improvements in bull dam selection could lie in alternative selection schemes, as proposed by Henderson; in methods of identifying cows with biased evaluations, as suggested by Van Vleck; or in some combination of both these suggestions.

Potential contributions of herd information to selection of bull dams and to prediction of young bull transmitting ability remained as a topic to be explored and thus became the subject of my research. Genetic and environmental herd characteristics were examined for ways in which this information might improve young sire sampling programs.

3. CHAPTER 1 EXAMINATION OF CURRENT PRACTICES IN SELECTION OF BULL-DAMS

Typically, bull studs contract matings of outstanding cows with well proven bulls to produce young bulls for sampling. Bull dams are usually selected from cows ranking among the top 2% of cows ranked nationally by cow index. An alternative method of bull dam selection would be to first select a superior group of herds. Cows in these herds would then be the pool from which bull dams are selected. This method has the advantage of concentrating candidates for selection into a manageable number of herds about which something is known (Henderson, 1964).

Two studies on current selection procedures were carried out to examine: (1) if bull dams have been selected from genetically superior herds, and (2) if cows selected as bull dams ranked among the top cows in their herds. If current selection schemes fail to locate genetically superior herds or fail to identify superior cows within a herd, then it seems very probable that the best cows are not being chosen as bull dams. Potential for improvement would then exist in the alternative scheme proposed for bull dam selection.

3.1 COMPARISON OF MICHIGAN HERDS INVOLVED IN YOUNG SIRE SELECTION AND SAMPLING TO ALL HERDS ON TEST IN MICHIGAN: Objective of the first study was to determine whether bull dams are currently being selected from genetically superior herds.

Data: The Records-in-Progress file of the Michigan Dairy Herd Improvement Association (DHIA) was obtained in October of 1985. The Records-in-Progress file captures a "snap-shot" of Michigan cows on test at the time. Cows contained in this file had either a record in progress or had completed a lactation, but had not intiated the subsequent lactation. Actual 305 day lactation records or extended 305-day records were reported for 202,835 cows in 2,510 herds. Both actual and extended records were also adjusted to a mature equivalent (ME) basis, and these were the records used in this study. Data was edited for fields lacking proper coding, such as a 0 or 1 for the lactation complete switch, leaving 200,418 records in 2,508 herds on test in Michigan.

Files from Select Sires Inc. were also obtained. These files identified Michigan herds from which young sires had been sampled 1974-1984, and herds which had participated in their program for genetic advancement by using young sire semen during the same period. Michigan herds were then coded into three herd groups.

There were 39 herds from which a young bull was selected by Select Sires for sampling. These sire contract herds (SCH) had a total of 3,359 lactation records in the current DHIA file.

There were 180 herds which had participated in Select Sires Program for Genetic Advancement for less than ten years. These progeny testing herds (PTH) had 21,139 records in the current DHIA file.

There were 93 herds which have participated in Select Sires Program for Genetic Advancement for ten or more years. These progeney testing herds (PTH+) had 12,316 records identified in the current DHIA file. Four herds were identified as both SCH and PTH+ and were included in both herd groups.

Methods: Five herd characteristics were examined. They were: (1) 305-ME milk; (2) Sire PD, which is the sire PD of cows in the herd; (3) Cow Index; (4) Calf's Sire PD, which is the PD of the cow's most recent calf; and (5) Service Sire PD, which is the PD of the bull which the cow has been most recently bred to.

The mean and variance of each characteristic was computed on a within herd basis and then pooled to estimate mean and variance for the herd group. These estimates and the percentage of herds and cow records that reported information are given in Table 5.

Table 5. Comparison of Michigan herd groups.

Characteristic	Herd Group			
	All DHI Herds	Sire Contract Herds	Progeny Testing Herds ≥10 Years	Progeny Testing Herds ≤10 Year
Number of Herds	2,508	39	93	180
Number of Records	200,418	3,359	12,316	21,139
305-ME Milk				
% Herds reporting	100%	100%%	100%	100%
%Records reporting	100%	100%	100%	100%
Mean	17,331	20,247	18,814	18,500
Standard deviation	3,101	3,414	2,881	2,996
Sire PD				
% Herds reporting	79%	100%	100%	99%
% Records reporting	47%	87%	84%	68%
Mean	265	279	373	333
Standard deviation	632	681	596	562
Cow Index				
% Herds reporting	64%	100%	100%	98%
% Records reporting	29%	62%	59%	47%
Mean	-168	-95	-44	-102
Standard deviation	407	511	433	393
Calf's Sire PD				
% Herds reporting	70%	80%	98%	95%
% Records reporting	25%	31%	32%	36%
Mean	764	751	910	821
Standard deviation	484	553	517	487
Service Sire PD				
% Herds reporting	36%	31%	36%	39%
% Records reporting	14%	9%	13%	17%
Mean	904	813	1117	921
Standard deviation	399	362	402	369

Results: The first characteristic studied was 305-ME milk average. The greatest average was seen in the SCH group, followed by the PTH+ group and then the PTH group. All three herd types participating in sire sampling programs had greater average production than an average herd on test in Michigan. The SCH group also had greatest variation for milk, while the second greatest variation was observed for all herds on test in Michigan. The PTH group ranked third in variance of milk production, and PTH+ had the smallest variation for milk.

The second characteristic studied was average sire PD. Greatest average sire PD was observed in the PTH+ group, with 100% of the herds and 84% of the cows having sire PD reported. The PTH group followed closely, with 99% of the herds and 68% of the cows having sire PD reported. The SCH group ranked third for average sire PD, with 100% of the herds and 87% of the cows having sire PD reported. Average sire PD of all herds on test was only slightly less than that observed in the SCH group. However, only 79% of all herds on test and 47% of all cows on test had sire PD reported. The greatest variation for sire PD was observed in the SCH group, followed by all herds on test, PTH+ and then PTH.

The third characteristic studied was average cow index. The highest average CI was observed in the PTH+ group, with 100% of the herds and 59% of the cows having cow index reported. The SCH group followed, with 100% of the herds and 62% of the cows having CI reported. Average CI in the SCH group was only slightly greater than that of the PTH group, which had 98% of the herds and 47% of the cows having CI reported. The lowest average CI was observed for all herds on test in Michigan, with only 64% of the herds and 29% of the cows having CI reported. The greatest variation for CI was observed in the SCH group and was much greater than the variation of CI observed in the other three herd groups. The second greatest variation for CI was observed in the PTH+ group, followed by all herds on test and then the PTH group.

The next characteristic studied was average calf's sire PD. This is the average PD of sires of the cows' most recently born calves. The PTH+ group had greatest average calf's sire PD, with 98% of the herds and 32% of the cows having calf's sire PD reported. The average seen in the PTH+ group was much greater than the average of the other groups. Second greatest average calf's sire PD was observed in the PTH group, with 95% of the herds and 36% of the cows having calf's sire PD reported. Next followed all herds on test in Michigan, with 70% of the herds and 25% of the cows having this information reported. Lowest average calf's sire PD was observed in the SCH group, with 80% of the herds and 31% of the cows having calf's sire PD reported. Greatest variation for calf's sire PD was seen in the SCH group followed by PTH+, PTH and then all herds on test. Differences between herd groups in variation of calf's sire PD were small.

The last characteristic studied was average service sire PD. This is the average sire PD of bulls to which the cows have been most recently bred. Very little information was reported for this characteristic. The greatest average service sire PD was observed in the PTH+ group, with 36% of the herds and 13% of the cows having service sire PD reported. Average of the PTH+ group was much greater than averages of the other groups. The PTH group had second greatest average service sire PD, with 39% of the herds and 17% of the cows having this information reported. The PTH group was followed closely by all herds on test, which had 36% of the herds and 14% of the cows with service sire PD reported. Lowest average service sire PD was seen in the SCH group, with only 31% of the herds and 9% of the cows having this reported. Greatest variation for service sire PD was seen in the PTH+ group, followed by all herds on test, PTH and then SCH. However, differences between herd groups in variation for service sire PD were again small.

Discussion: Although sire contract herds had greatest average milk production, they did not have greatest average cow index. Additionally, cows in the SCH group did not have greatest average sire PD, nor were they bred to superior sires. By all indications, the SCH group was not genetically superior to other herd groups in Michigan. The greater average production of the SCH group must, therefore, be attributed to better management.

Average cow index of the SCH group was greater than average CI of all herds on test. However, average sire PD of the two groups were nearly equal. The portion of CI attributed to pedigree information should be nearly equal for cows in both herd groups, assuming that average dam merit is equal in the two groups. This means that the higher cow indexes observed in the SCH group must be due to greater modified contemporary deviations.

The SCH group also had the greatest variation in all characteristics considered, except service sire PD. Several researchers (Everett et. al., 1982 and Powell et. al., 1983) have expressed concern over the bias heterogeneous intraherd variances and production levels have on cow evaluations. This study has added support to their concerns. Cows in sire contract herds had higher cow indexes than all cows on test, while they did not seem to have genetic superiority to these cows, as was discussed in the preceding paragraph. Additionally, sire contract herds had both higher production levels and greater intraherd variances than other herd groups.

Work by Everett (1984) found that most registered cows are located in herds with low genetic level, while herds with high genetic level are composed primarily of grade cows. He stated that rate of genetic gain for milk could be improved by allowing grade cows to be considered as bull dams. His study also suggests that by first considering a different pool of herds from which to select bull dams, genetic gain for milk production might be improved.

The herd group which consistently showed a superior level of genetic merit was the PTH+ group. This group had greatest average cow index; cows in these herds had greatest average sire PD; and they were bred to superior bulls. Several points for speculation may be made from these observations. In terms of genetic merit, the PTH+ group was superior to other Michigan herd groups. However, should young bulls be selected from the same population of herds in which they are also being sampled? What effect does the genetic superiority of the PTH herds have on initial estimation of young bull PD? Are bulls which are returned to service based on these initial PD estimates going to perform as well when used in the rest of the population? While these points were not investigated within this project, they are questions which deserve attention.

3.2 Rank of Bull Dams Within Herd: Objective of the second study was to examine whether cows which have been selected as bull dams were the best cows in their respective herds.

Data: Dams of 31 young bulls, born in Michigan herds from 1980-1981 and sampled by Select Sires Inc., were identified. Cow index and production data were obtained for all cows in the dam's herd for year of bull's birth.

Methods: Rank of the bull dam within her herd by both CI and 305-ME milk was determined. Four classes of rank were defined. The first class included dams that were either ranked first in the herd, or were within the top 1% of the herd. The second class included dams that were not in the top 1% of the herd, but were in the top 5% of the herd. The third class included dams that were not in the top 5% of the herd, but were in the top 15% of the herd. The fourth class included all dams that were not ranked within the top 15% of the herd. The number of observations in each of these classes is given in Table 6.

Table 6. Number of dams per class when ranked by cow index (CI) and 305-ME milk production.

Within herd rank	Number of dams by CI	Number of dams by production
Top 1%	8	8
<Top 1% and ≥Top 5%	9	9
<Top 5% and ≥Top 15%	7	5
<Top 15%	7	9
Total	31	31

Results and Discussion: When ranked by both CI and by milk production, 17 of the 31 bull dams were within the top 5% percent of their herds. However, 14 of the 31 bull dams were not in the top five percent of the herd when ranked by CI or by milk production. The correlation between CI rank class and milk production rank class was .52.

In this data set, a large percentage of cows selected as bull dams were not the top cows in their herds. Two of the 31 bull dams were below the top 40% in the herd when ranked by CI, and four bull dams were below the top 40% when ranked by milk production. Although this is a small data set covering only two years of sampling, this observation raises some important questions. Why aren't the best cows in a herd selected as bull dams? What are existing selection criteria that allow bull dams to be selected without regard to their rank in the herd?

Results of studies in sections 2.1.1 and 2.1.2 indicate that current selection schemes may not be choosing bull dams from genetically superior herds, nor do they seem to select the best cows within a herd. Improvements in selection of bull dams could boost the rate of genetic gain. An alternative selection scheme, which seems to have potential, is to first select a group of superior herds and then select the best cows within these herds. To examine this potential further, the second investigation focused on sire contract herds. Studies were carried out to examine whether herd characteristics are helpful in prediction of young bull PD and/or in identification of dams with biased evaluations.

4. CHAPTER 2 PREDICTION OF YOUNG BULL TRANSMITTING ABILITY

Progeny testing a young bull takes a large investment of time and money. Most young bulls which enter an AI stud come from the contracted mating of a high ranking, well proven bull with an outstanding cow. It takes more than six years from the time a bull is born until his first crop of daughters have completed a lactation and the bull is evaluated form progeny information. At this time, the bull stud decides which bulls are returned to service and which bulls are culled.

Bull studs always strive to achieve the highest selection intensity possible within their economic limits. Therefore, they want to sample only superior bulls and return to service the top percentage of bulls sampled. A bull stud can not afford to waste its resources sampling poor bulls.

To ensure that only superior bulls enter a stud for progeny testing, it is important to accurately predict the transmitting ability (PD) of a young bull. Four studies were undertaken in the second investigation to examine which factors are useful in prediction of young bull transmitting ability. In the first study, pedigree information was examined for its usefulness in prediction of young bull PD. In the second study, influence of genetic trend on prediction of PD was also examined. The third study examined whether young bull's herd of birth contributes to the predictability of his PD. The fourth study involved the reconstruction of sire contract herds for year of young bull's birth. Two sub-studies were then done to examine differences among herds in their sampling succes rates and to determine whether specific herd characteristics help to predict young bull PD.

4.1 PREDICTION OF YOUNG BULL TRANSMITTING ABILITY FROM PEDIGREE INFORMATION: An goal of this research was to examine whether herd information can improve predictability of a young bull's transmitting ability. Before this question could be examined, however, it was first necessary to determine how well a young bull's PD can be predicted from pedigree information alone. Usually prediction of a young bull's transmitting ability is based on genetic evaluations of his relatives. Objective of the first study was to examine prediction of young bull PD from pedigree information and determine which pedigree information or indices contribute most to the prediction.

Data: Ninety three young bulls selected from Michigan herds by Select Sires Inc. for sampling during the years 1975-1982 were identified. Sire PD, dam CI, and maternal grandsire PD at the time of the young bull's first semen allocation, the time when the bull's semen is first relased for sampling, were obtained. This is the information from which sire analysts must base their final decision of whether or not to sample the young bull.

This pedigree information was also combined into two indices which are commonly used to predict a young bull's transmitting ability:

1) **Estimated Transmitting Ability (ETA) = ½ sire PD + ½ dam CI**
2) **Pedigree Index (PI) = ½ sire PD + ¼ maternal grandsire PD.**

The initial PD estimate of the young bull was also obtained. This is his first estimate of PD which is based on daughter information. This same data set was used in sections 4.2 and 4.3.

Methods: Young bull intitial PD estimate was the dependent variable in five linear regression models. Independent variables used in the five models were:

1) young bull PI
2) young bull ETA
3) sire PD
4) dam CI
5) maternal grandsire PD

The general form of the fixed model is:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

where $\mathbf{y}$ is the vector of observations;

$\mathbf{X}$ is the matrix of observed values corresponding to the parameters to be estimated in $\mathbf{b}$;

$\mathbf{b}$ is the vector of unknowns to be estimated;

$\mathbf{e}$ is the vector of residual errors corresponding to each observation.

The five models were run under two sets of assumptions. Analyses were first done using ordinary least squares procedures, **e** ~ $N(0,I\sigma^2)$ and **y** ~ $N(Xb,I\sigma^2)$. The same five models were run again using weighted least squares procedures, **e** ~ $N(0,A)$ and **y** ~ $N(Xb,A)$, where **A** is the matrix of additive genetic relationships among young bulls. An element of this matrix is denoted as a_{ij} and is the additive genetic relationship between the ith and jth young bulls.

The assumption $V(y)=I\sigma^2$ of the ordinary least squares procedures was known to be in violation, since a high degree of relationship existed among a large number of young bulls. Weighted least squares analyses were done to determine whether prediction of young bull PD is improved when covariance among young bulls is considered. Given the pedigree information available, six types of relationships were recognized in the relationship matrix:

1) Full sibs: $a_{ij} = .5$
2) Maternal half-sibs: $a_{ij} = .25$
3) Paternal half-sibs: $a_{ij} = .25$
4) Identical maternal grandsire: $a_{ij} = .0625$
5) Sire of one young bull is maternal grandsire of another: $a_{ij} = .125$
6) Identical maternal grandsire and identical sire: $a_{ij} = .3125$

Diagonals of the relationship matrix were one, assuming that no inbreeding occurred. Results of analyses both with and without the matrix of additive genetic relationships are given in Table 7.

Table 7. Regressions of young bulls' first proof on pedigree information.

Model	Without Additive Genetic Relationship	signif. level	With Additive Genetic Relationship	signif. level
$PD_i = b_0 + b_1 (PI)_i + e_i$	$b_0 = -22.56$	.05	$b_0 = -21.68$	.05
	$b_1 = .57$	.05	$b_1 = .57$	.25
	$\sigma_e^2 = 297,469$		$\sigma_e^2 = 363,018$	
	adj. $R^2 = .048$		adj. $R^2 = .015$	
$PD_i = b_0 + b_1 (GI)_i + e_i$	$b_0 = -59.27$	.05	$b_0 = -15.96$	.05
	$b_1 = .54$	.005	$b_1 = .54$	.05
	$\sigma_e^2 = 289,219$		$\sigma_e^2 = 354,317$	
	adj. $R^2 = .074$		adj. $R^2 = .038$	
$PD_i = b_0 + b_1 (SirePD)_i + e_i$	$b_0 = -13.23$	.05	$b_0 = 27.47$	.05
	$b_1 = .28$	.10	$b_1 = .26$	.50
	$\sigma_e^2 = 305,062$		$\sigma_e^2 = 370,267$	
	adj. $R^2 = .024$		adj. $R^2 = .000$	
$PD_i = b_0 + b_1 (DamCI)_i + e_i$	$b_0 = 48.53$	.05	$b_0 = 84.09$	.05
	$b_1 = .42$	.01	$b_1 = .39$	.05
	$\sigma_e^2 = 290,603$		$\sigma_e^2 = 356,283$	
	adj. $R^2 = .070$		adj. $R^2 = .033$	
$PD_i = b_0 + b_1 (MGS\ PD)_i + e_i$	$b_0 = 115.01$	.05	$b_0 = 144.61$	.05
	$b_1 = .22$	.10	$b_1 = .21$	.50
	$\sigma_e^2 = 303,275$		$\sigma_e^2 = 367,313$	
	adj. $R^2 = .029$		adj. $R^2 = .003$	

Results and Discussion: In the ordinary least squares analyses, in which additive genetic relationships among young bulls were not considered, the regression weight of young bull PD on PI was slightly greater than the regression weight of PD on ETA. The coefficient of multiple determination, which was adjusted for small sample size (adj.R^2), was greater for ETA than for PI.

PI assumes the young bull's maternal granddam is of average genetic value and the young bull's dam is an average daughter of her sire. Available information on both the dam and the maternal granddam is ignored. When arranging contract matings, bull studs select first by cow index and then by PI to make final decisions on matings contracted. They rely on PI rather than ETA at this second stage of selection to avoid problems caused by biased cow evaluations. However, this practice disregards information which may be useful in predicting the young bull's transmitting ability. Results of these two models showed that bull dam information does contribute to the prediction of young bull PD, and ETA was a better predictor of young bull PD than was PI.

Theoretical weights for regression of son on sire and dam are both .5. In practice, however, regression coefficients depend on the number and type of records used in the son's evaluation and the sire's and dam's evaluations (Van Vleck, 1981). If the son has many daughters, regression weights for sire and dam may nearly equal .5, but generally are less than .5 (Van Vleck, 1981). Equality of these weights depends on both the number of records and the types of records used in sire and dam evaluations. The tendency is for sire regression coefficients to remain near .5, though they may be greater than .5. While regression coefficients on the dam approach .5, they are always less than .5 (Van Vleck, 1981).

In this study, sire PD was a less effective predictor of son's PD than was dam CI. The observed weight for regression of son on sire was smaller than expected and was not within the range of theoretical values. When sire PD was the independent variable, the adjusted R^2 value was only .024. Why sire PD explained such a small proportion of the variation in young bull PD is not clear, and may be due to the small data set.

The regression coefficient of son on dam was within the range of theoretical values. Although the number of records in both son's evaluation and dam's evaluation are not known, the weight of .42 seems slightly greater than expected, considering it was the son's first evaluation and not many daughters would be included in his evaluation. Van Vleck (1981) estimated that a young bull with 60 daughters, his sire having 1000 daughters, and his maternal grandsire having 1000 daughters would have a weight for regression on dam between .353 and .375, depending on the number of lactation records on his dam. When dam CI was the independent variable, the adjusted R^2 value was .070, over twice the value acheived in the sire model.

Theoretically, the regression coefficient of son on maternal grandsire is .25. In this study, the weight for regression of son on maternal grandsire PD was .22, so was within an expected range. The adjusted R^2 value of .029 was slightly greater than that obtained when sire PD was the independent variable in the model.

When the relationship matrix was considered, intercepts in all models increased, while regression coefficients remained equal or nearly equal. Why incorporation of the relationship matrix had this effect is unclear. Inclusion of the relationship matrix increased estimates of residual variance in all five models. The relationship matrix accounts for a portion of the variance among observations, and in effect states that young bulls in this data set are more alike than young bulls selected randomly from the population. Therefore, it is reasonable that estimates of residual variance increased. Increased residual variance caused significance of variables to decrease. It also caused adjusted R^2 values to decrease by more than half in all models, except the model in which ETA was the independent variable and the adjusted R^2 value was reduced by not quite half.

Accounting for covariance among young bulls did not improve predictability of PD from pedigree information. Since a goal of this research was to increase predictability of young bull PD, genetic covariance among young bulls was not considered further.

Of the five pedigree models considered, the ETA model in the ordinary least squares analysis explained the greatest proportion of variation in young bull PD. Thus, ETA would be used as the pedigree predictor of young bull PD for further work. Additional studies examined whether consideration of herd factors increases predictability of young bull PD beyond what is predicted by ETA alone.

4.2 CONSIDERATION OF GENETIC TREND IN PREDICTION OF YOUNG BULL TRANSMITTING ABILITY: Genetic trend in the population could make year of birth an important factor in predicting young bull PD. If year of birth does contributed to prediction of young bull PD, it would have to be included as a nuisance factor in later studies. Objective of this study was to examine the influence of genetic trend on young bull PD and the necessity of including year as a nuisance factor in further work with this data set.

Data: The same data set described in section 4.1 was used in this study. Young bulls in the data set were from eight years of sampling, 1975-1982. Corresponding birth years of bulls sampled were 1974-1981.

Methods: A best, linear, unbiased estimator (BLUE) was used to anlayze the following covariate model:

$$\mathbf{PD_{ij} = \mu + Y_i + b_1(ETA)_j + e_{ij}}$$

where $\mathbf{PD_{ij}}$ is the initial PD estimate of the jth young bull born in the ith year;

μ is the fixed constant common to all observations;

Y_i is the fixed effect of the ith year of birth of the young bull, with i=1,2,...8;

b is the regression coefficient of young bull PD on ETA;

ETA_j is the estimated transmitting ability (½ sire PD and ½ dam cow index) of the jth young bull;

e_{ij} is the random residual associated with PD_{ij}.

Results and Discussion: Results of the analysis are shown in Table 8. Including the fixed effect of birth year increased the adjusted R^2 value by .036 from the regression of young bull PD on ETA alone. However, tests of partial significance indicated that year was not a significant source of variation. Therefore, year was not included in further models.

Table 8. Year differences in young bulls.

Variable	Partial f-value	Significance level
μ	5.41	.05
Y_i	1.52	.19
b_1	2.78	.10

adj.R^2=.110

4.3 CONSIDERATION OF HERD OF ORIGIN IN PREDICTION OF YOUNG BULL TRANSMITTING ABILITY: A major goal of this research was to examine sire contract herds for herd characteristics which might be useful in prediction of young bull transmitting ability. Objective of this study was to examine young bull herd of origin as a fixed classification factor to determine whether it is a significant source of variation in predicition of young bull's PD value.

Data: The same data set descibed in section 4.1 was used in this study. Young bulls in the data set were from 32 different sire contract herds.

A best, linear, unbiased estimator was used to analyze the following covariate model:

$$PD_{ij} = \mu + H_i + b_1(ETA)_j + e_{ij}$$

where PD_{ij} is the first proof of the jth young bull born in the ith herd;

μ is the fixed constant common to all observations;

H_i is the fixed effect of the ith herd of birth, with i=1,2,...32;

b is the regression coefficient of young bull PD on ETA;

ETA_j is the estimated transmitting ability (½ sire PD + ½ dam CI) of the jth young bull;

e_{ij} is the random residual associated with PD_{ij}.

Results and Discussion: Results of the analysis are shown in Table 9. Including young bull's herd of birth as a fixed factor in the model more than doubled the adjusted R^2 value from the value obtained by regression on ETA alone. Tests of partial significance indicated that herd of birth had a moderate degree of significance and, thus, may be a significant source of variation in young bull PD. Herd information might be useful in selection of young sires for sampling. Therefore, further studies were done to examine what particular herd characteristics might prove useful in prediction of young bull PD.

Table 9. Consideration of herd effect in prediction of young bull PD.

Variable	Partial f-value	Significance level
μ	6.02	.05
H_i	1.46	.11
b_1	9.30	.005

adj.R^2=.200

4.4 COMPARISON OF HERDS GROUPED BY SAMPLING SUCCESS RATE: Objective of this study was to compare sire contract herds of different sampling success rates. This was done to determine whether differences existed between herds able to return to service an acceptable proportion of young bulls sampled and herds unable to return to service an acceptable proportion of bulls sampled.

To accomplish this objective, it was necessary to study sire contract herds at the time young bulls were selected for sampling.

Data: Sire contract herds were reconstructed for year of young bull's birth. Michigan herds, from which 93 young bulls had been selected by Select Sires Inc. for sampling in years 1975-1982, were identified. These bulls came from 32 sire contract herds in 61 herd/years. Information for herd/years was reconstructed with data from Michigan DHIA annual summary tapes. These tapes contained records of cows that had completed or terminated lactations in a given DHIA year, which begins on October 1 and ends on September 30. Herd characteristics compiled were:

1) size of milking herd
2) average 305 day mature equivalent (ME) milk
3) standard deviation of 305-ME milk
4) coefficient of variation for 305-ME milk
5) average transmitting ability (PD) of sires of cows in the herd
6) standard deviation of PD of sires of cows in the herd.

These herd characteristics were also used in study 4.5 to examine what particular herd characteristics contribute to selection of bull dams and to prediction of young bull PD.

The following information was available for herds in which 31 young bulls were born in 1980 and 1981.

7) average cow index
8) standard deviation of cow index
9) bull dam's rank within the herd by cow index (CI) and by 305-ME milk.

Information on rank of dam within herd was used to examine whether selected bull dams were the best cows in their respective herds. Results of this study were reported in section 3.2.

Methods: Sampling success rate was defined as the percent of young bulls sampled which were returned to service after progeny testing. Sire contract herds were divided into three groups; successful, unsuccessful, and unclassified. Characteristics of the first two classification groups were compared.

Of 93 young bulls sampled from Michigan herds 1975-1982, 18 were returned to service, 72 were rejected and 3 were still undecided as of January 1987. The overall success rate for Michigan sire contract herds was calculated as $(^{18}/_{90})$ x 100% = 20%. One bull was returned to service out of every five sampled. This success rate was the basis for catagorizing herds.

Seventeen herds sampled less than five bulls and returned none to service. These herds, involving 25 young bulls sampled, could not be classified as either successful or unsuccessful. Nine herds sampled less than five bulls and returned to service at least one bull. These nine herds, with a total of 12 bulls sampled, were classified as successful. Three herds sampled more than five bulls and achieved a success rate greater than 20%. These herds, with 16 young bulls sampled, were also classified as successful. Three herds sampled more than five bulls, but did not achieve a success rate of 20% or greater. These herds, with a total of 37 bulls sampled, were classified as unsuccessful.

A difficulty encountered in the unsuccessful group was that one of the three herds was very large. This herd averaged more than 1200 cows in the milking herd, while the other two herds averaged less than 60 milking cows. This herd dominated in calculation of group means so was, therefore, excluded from the study.

It is interesting to note that the most successful herd within the successful group sampled six young bulls and returned four to service, while the least successful herd within the unsuccessful group sampled 20 young bulls and returned only one to service. These two extreme herds are reported with a summary of the successful and unsuccessful herd groups in Table 10.

Table 10. Summary of successful and unsuccessful herd groups and of extreme herds within each classification.

	Number Herds	Number Herd/Years	Number Bulls Sampled	Number Bulls Returned	Overall Success Rate
Successful	12	24	28	15	53.6%
Extreme Successful	1	4	6	4	66.7%
Unsuccessful	2	12	26	1	3.9%
Extreme Unsuccessful	1	7	20	1	5.0%

Average young bull PI, ETA, and PD were computed for the two groups and for the two extreme herds. These means and standard errors are shown in Table 11.

Table 11. Pedigree and progeny testing results in pounds of milk for successful group, unsuccessful group, and extreme herds.

	Avg. PI	Stand. Error	Avg. ETA	Stand. Error	Avg. PD	Stand. Error
Successful	244	±45	334	±46	359	±110
Extreme Successful	166	±105	369	±64	466	±189
Unsuccessful	151	±41	211	±65	-160	±104
Extreme Unsuccessful	172	±69	168	±50	-97	±120

Weighted means and corresponding standard errors were computed for herd characteristics of the two groupings and for the two extreme herds. Pooled estimates of variance were also computed, and confidence intervals on standard deviations were constructed. These results are shown in Table 12.

Table 12. Herd characteristics of successful group, unsuccessful group, and extreme herds.

	Average 305-ME milk	95% conf. interval on S.D. milk	Average sire PD	95% conf. interval on S.D. sire PD
Successful	17,890 ±71	2,918<3,014≤3,117	-510 ±13	535<553≤572
Extreme Successful	17,117 ±139	2,766<2,918≤3,091	-569 ±25	501<529≤560
Unsuccessful	22,696 ±167	3,871<4,089≤4,337	-291 ±21	490<518≤549
Extreme Unsuccessful	22,538 ±201	3,842<4,061≤4,312	-130 ±20	372<393≤417

Results: In the successful group, average PI under estimated young bull average progeny testing results (PD), while average ETA was a very accurate predictor of average young bull PD. The interval on ETA, defined by the mean plus or minus the standard error of the mean, was entirely contained within the PD interval. However, the interval on PI barely overlapped the PD interval.

In the unsuccessful group, both average PI and average ETA greatly over estimated average PD. Additionally, intervals on both PI and ETA did not overlap the interval on PD at all.

In the extreme successful herd, both PI and ETA under estimated PD. However, PI under estimated PD by 300 pounds while ETA under estimated PD by only 100 pounds. The interval on PI did not overlap the interval on PD at all, while the interval on ETA was entirely contained within the interval on PD.

In the extreme unsuccessful herd, both PI and ETA over estimated PD. However, PI and ETA were nearly equal in their degree of over estimation. Intervals on both PI and ETA did not overlap the interval on PD.

Successful herds had lower average production and smaller production variance than did unsuccessful herds. The successful herds also had lower average sire PD than did the unsuccessful group. Variation in sire PD was slightly greater in the successful group than in the unsuccessful group, though confidence intervals on standard deviations of the two groups overlapped by a fair amount.

Average milk production was more than 5,000 pounds less in the extreme successful herd than in the extreme unsuccessful herd. Standard deviation of milk production was also much smaller in the extreme successful herd. Average sire PD was much lower in the extreme successful herd than in the extreme unsuccessful herd, while variation of sire PD was much smaller in the extreme unsuccessful herd than in the extreme successful herd.

Discussion: Due to Mendelian sampling, a young bull's transmitting ability can not be perfectly predicted. The random nature of gamete formation and combination prevents exact prediction of individual genetic value. However in prediction of a group of individuals, effects of Mendelian sampling should be reduced by the central limit theorem, so that average group prediction is more likely to equal average group breeding values estimated from progeny results, depending on the size of the group.

In successful herds, average prediction by ETA very nearly equalled average PD. In these herds, there was no obvious bias in young bull prediction by ETA. However, PI was not a very accurate predictor in these herds and under estimated average PD. However, PI should underestimate PD if superior bull dams have actually been selected. If the bull dam is an above average daughter of the maternal grandsire, rather than an average daughter as PI assumes, then PI should underestimate PD. This observation adds strength to Van Vleck's statement (1986a) that rather than ignore all cow information because some cow evaluations are inaccurate, it is better to identify and ignore only cows with biased evaluations. In unsuccessful herds it seems we are confronted with problems beyond normal Mendelian sampling. All predictions were biased, such that average predictions were much greater than average PD.

Both average milk production and average genetic level, as measured by average sire PD, appear to be greater in the unsuccessful herds than in successful herds. However, average young bull PD was much greater in the successful herds and was much more accurately predicted.

There were several difficulties in this study which must be pointed out. First of all, the criteria used for grouping herds as successful or unsuccessful may not have been completely appropriate. The randomness of Mendelian sampling is an important consideration. It may have been more appropriate to group herds based on a weighting by probability.

Another difficulty might have occurred in comparison of group characteristics. Calculation of weighted averages and pooled estimates of variance assumed that herd/years were independent observations. However, a correlation is likely to exist between characteristics of a herd in one year and characteristics of that same herd a year or two later, since many of the cows will be the same. It is not known what effect ignoring this correlation had on estimation of group characteristics.

A last consideration is that occasionally young bulls with outstanding PD's may not be returned to service due to infertility or poor functional soundness. It may be questioned whether these bulls should still be considered as unsuccessful samplings. While results of this study are in no way conclusive, they have raised important questions for further consideration.

Is the accuracy or inaccuracy of predictors what makes the difference between a successful herd and an unsuccessful herd? If so, what causes predictors to be more accurate in some herds than in others? Are cow and bull evaluations, from which index predictors are derived, influenced by herd production level, or level of herd variance? Although differences in herd characteristics were seen between the groups, it is unclear whether these differences were responsible for the different degrees of sampling success observed.

4.5 CONSIDERATION OF HERD CHARACTERISTICS IN PREDICTION OF YOUNG BULL TRANSMITTING ABILITY: It was seen in section 4.3 that a young bulls' herd of birth was significant in prediction of young bull PD. It was also observed in section 4.4 that different herds have varying degrees of success in sampling young bulls. While differences between successful and unsuccessful herds were observed, it was not clear whether these differences were responsible for differences in sampling success rates. Objective of this study was to examine particular herd characteristics to see whether they aid in prediction of young bull PD.

Data: Data described in sections 4.1 and 4.4 were also used in this study. These were the 93 young bulls chosen from Michigan herds by Select Sires Inc. during the period 1975-1982 and the corresponding reconstructed herd information of sire contract herds for time of young bulls' birth.

Methods: Ordinary least squares procedures were used to analyze five multiple linear regression models. In each model, young bull PD was regressed on ETA plus a herd characteristic. The following five herd characteristics were considered as independent variables in each of the five models.

1) average 305-ME milk production
2) standard deviation of milk production
3) coefficient of variation for milk production
4) average sire PD
5) standard deviation of sire PD.

Additional models were analysed using the 31 young bulls for which herd average cow index and herd standard deviation of cow index were availiable. Results are reported in Table 13.

Table 13. Contribution of characteristics of bull-dams' herd to prediction of young bulls' transmitting ability.

Model	Results	Significance Level
$PD_i = b_0 + b_1(ETA)_i + b_2(\text{Avg. Milk})_i + e_i$	$b_0 = 464.68$	.05
	$b_1 = .56$	.005
	$b_2 = -.03$	.23
	adj.R^2 =.081	
$PD_i = b_0 + b_1(ETA) + b_2(\text{S.D. Milk})_i + e_i$	$b_0 = 459.37$	.05
	$b_1 = .59$	.005
	$b_2 = -.15$	.09
	adj.R^2=.096	
$PD_i = b_0 + b_1(ETA)_i + b_2(\text{C.V. Milk})_i + e_i$	$b_0 = 317.35$	.05
	$b_1 = .57$	.005
	$b_2 = -2228.73$	.33
	adj.R^2=.075	
$PD_i = b_0 + b_1(ETA)_i + b_2(\text{Avg. Sire PD})_i + e_i$	$b_0 = -29.63$	.05
	$b_1 = .53$	.01
	$b_2 = .06$	N.S.
	adj.R^2=.065	
$PD_i = b_0 + b_1(ETA)_i + b_2(\text{S.D. Sire PD})_i + e_i$	$b_0 = 148.57$	.05
	$b_1 = .60$	.005
	$b_2 = -.43$	.40
	adj.R^2=.073	
$PD_i = b_0 + b_1(ETA)_i + b_2(\text{Avg.Cow Index})_i + e_i$	$b_0 = 113.02$	.001
	$b_1 = .31$	.50
	$b_2 = -.04$	N.S.
	adj.R^2=.000	
$PD_i = b_0 + b_1(ETA) + b_2(\text{S.D. Cow Index})_i + e_i$	$b_0 = -204.79$	.001
	$b_1 = .18$	N.S.
	$b_2 = .97$	.46
	adj.R^2=.000	

Results and Discussion: Partial regression on ETA was highly significant in all models analysed, except the two models which contained herd average cow index and herd standard deviation of cow index as independent variables. Of herd characteristics examined, only intraherd standard deviation for milk was significant in prediction of young bull PD. Other herd characteristics were not significant sources of variation in young bull PD. It is interesting to note that regression coefficients of young bull PD on herd milk and herd standard deviation of milk are both negative. The negative coefficients indicate that as herd milk production and intraherd milk variation increase, young bull PD was likely to decrease. This concurs with what was observed in section 4.4: that unsuccessful herds had higher milk production and greater intraherd milk variation than did successful herds.

Herd average sire PD was not at all significant in explaining variation in young bull PD. This result was slightly surprising since the original hypothesis proposed that herd genetic merit might be useful in predicting young bull PD and in bull dam selection.

Herd average cow index and herd standard deviation for cow index were not at all significant in predicting young bull PD. However, young bull ETA was also not significant in predicting young bull PD in these two models. There was too little information available for studies involving herd cow index to be valid.

Results of this investigation indicate herd differences do exist and can be useful in prediction of young bull PD and in selection of bull dams. Of herd characteristics considered, intraherd variation for milk production provided the most information in predicting young bull PD. The next phase of investigation focused on intraherd milk variation and how this herd measurement might improve sire sampling programs and the subsequent rate of genetic gain.

c
e
v
(
V
i
a

va
va
in
po
va
se
he
PD

bu
stu
com
were
In b
were

5. CHAPTER 3 EFFECTS OF INTRAHERD MILK VARIANCE ON YOUNG BULL SELECTION PROGRAMS

Problems caused by heterogeneous intraherd variances were discussed in the literature review. While bull and cow evaluation procedures assume homogeneous intraherd variances, heterogeneity of variance has been documented. (Hill et. al., 1983; Mirande and Van Vleck, 1985; DeVeer and Van Vleck, 1986) Research has indicated heterogeneity of intraherd variance may bias cow evaluations. (Everett et. al., 1982; Powell et. al., 1982)

Differences among Michigan herd groups in both genetic variance, measured by variance of sire PD, and environmental variance, measured by variance of milk production, were seen in section 3.1. In section 4.4 it was seen that herds with poor sampling success rates had greater environmental variance than herds with higher sampling success rates. As seen in section 4.5, intraherd milk variation of bull dam herds was a significant factor in prediction of young bull PD.

The third investigation involved the grouping of young bulls based on herd milk variance levels. In the first study, variance groups were characterized by the herds comprising each group. In the second study, variance groups were characterized by the young bulls comprising each group. In both these studies, correlations among characteristics were examined.

The third study examined variance groups for bias in predicted transmitting ability of young bulls. The fourth study examined variance groups for bias in the distribution of young bulls.

5.1 CHARACTERIZATION OF HERDS GROUPED BY SAMPLING SUCCESS RATE: Production and genetic characteristics of herds comprising each group were examined. Objective of this study was to compare characteristics of herds contained in each of the variance groups.

Data: Data used in this study were the 93 young bulls chosen from Michigan herds by Select Sires Inc. during the period 1975-1982, as were described in section 4.1. Young bulls were ranked by intraherd milk variation of the herds in which they were born. Three variance groups of 31 observations each were formed.

Methods: Weighted means, standard errors of weighted means, and pooled estimates of variance were computed for each variance group for 305-ME milk and sire PD. Confidence intervals on pooled estimates of standard deviations were constructed. These results are reported in Table 14. Correlations between these herd characteristics were computed for the total data set and within each variance group. Results are reported in Table 15.

Table 14. Production and genetic characteristics of herds stratified by milk variance.

Herd Charcteristic	Intraherd Milk Variance Group		
	Low	Mid	High
305-ME milk	18,090 ±63	18,339 ±32	19,718 ±53
95% confidence interval on S.D. milk	$2704<2825\le 2916$	$3423<3467\le 3512$	$4038<4110\le 4186$
Sire PD	-609 ±12	-692 ±5	-427 ±7
95% confidence interval on Sire PD	$528<544\le 561$	$537<544\le 551$	$544<554\le 564$

<u>Results and Discussion</u>: The high-variance group had the greatest average milk production, followed by the mid-variance group and then the low-variance group. Intervals on average milk, defined by the mean plus or minus standard error of the mean, did not overlap between variance groups. Levels of milk variance were significantly different in the three groups at $\alpha = .05$.

The high-variance group also had greatest average sire PD. The low-variance group had second highest average sire PD, while the middle variance group ranked last. Intervals on average sire PD did not overlap between variance groups. Average sire PD was nearly equal in mid and low-variance groups, while it was much greater in the high-variance group. Variance of sire PD was nearly equal in the three groups.

Table 15. Correlations between herd characteristics in herds stratified by milk variance.

	Intraherd Milk Variance Group			
Herd Characteristics	Low	Mid	High	Total
Avg milk/SD milk	.14	.24	.51	.57
Avg milk/avg. sire PD	.35	.54	.57	.60
Avg milk/SD sire PD	.43	.07	-.03	.08
Avg sire PD/SD milk	-.37	.03	.47	.45
SD milk/SD sire PD	-.02	.18	-.20	-.11

A moderate, positive correlation was seen between average 305-ME milk and standard deviation for milk in the total data set and was greater than the correlation within any individual variance group. Correlation between average milk and standard deviation for milk was quite small in the low-variance group and increased in the higher variance groups. This pattern was the same for correlations between average 305-ME milk and average sire PD, although correlations were generally higher.

Correlation between average 305-ME milk and standard deviation of sire PD was small but positive for the total data set. It was moderate in the low-variance group and decreased in the higher variance groups, becoming negative in the high-variance group.

Correlation between average sire PD and standard deviation of milk was moderate overall. In the low-variance group, this correlation was moderate, but negative. It was near zero in the mid-variance group and was moderate and positive in the high-variance group.

Correlation between standard deviation of milk and standard deviation of sire PD was small and negative in the total data set. It was nearly zero, but negative in the low-variance group, moderate and positive in the mid-variance group, and moderate but negative in the high-variance group.

Interpretation of these correlations is difficult, since most correlations were quite different from the high to low-variance groups. However, this could be an important clue as to why heterogeneous herd variance must be accounted for in evaluation procedures and why herd variance level influences prediction accuracy in young sire sampling.

5.2 CHARACTERIZATION OF YOUNG BULLS GROUPED BY INTRAHERD MILK VARIANCE: Pedigree information, index predictions and progeny test results of young bulls in each variance group were examined. Objective of this study was to compare characteristics of young bulls in each variance group.

Data: The same variance groups were used in this study as were descirbed in section 5.1.

Methods: Averages for sire predicted difference (PD), dam cow index (CI), maternal grandsire PD, pedigree index (PI), estimated transmitting ability (ETA), and initial PD estimate were computed in each variance group based on the 31 young bulls comprising each group. These figures along with standard errors are reported in Table 16. Correlations between young bull PD and young bull pedigree information were computed for the total data set and within each variance group. Results are reported in Table 17.

Table 16. Averages for young bull pedigree evaluations and progeny test results by intraherd milk variance group.

	Intraherd Milk Variance Group		
Average	Low-Variance	Mid-Variance	High-Variance
Dam CI	141 ±57	169 ±66	249 ±75
Sire PD	404 ±67	500 ±66	593 ±59
MGS PD	-6 ±97	153 ±92	21 ±84
PI	197 ±43	287 ±46	301 ±35
ETA	272 ±44	335 ±57	421 ±54
PD	284 ±98	63 ±99	35 ±97

Results: Average dam CI was much greater in the high-variance group than in mid and low-variance groups, where it was nearly equal. However, intervals on average dam CI, defined by the mean plus or minus standard error of the mean, overlapped between variance groups.

Average sire PD was also greatest in the high-variance group, followed by the mid-variance group and then low-variance group. Intervals on average sire PD also overlapped between variance groups.

Average maternal grandsire PD was greater in the mid-variance group than in the low and high-variance groups, which were nearly equal. However, intervals overlapped between variance groups.

Index predictions of young bull PD by both PI and ETA were greatest in the high-variance group, followed by the mid and then the low-variance groups. In all three variance groups, average PI predictions were lower than average ETA predicitons. It is not surprising that average PI is less than average ETA overall. PI assumes that the bull-dam is an average daughter of the maternal grandsire. However, if we are choosing bull-dams that are above average daughters of the maternal grandsire, then PI does not give enough credit to the bull-dam and ETA should be greater than PI.

All pedigree information indicated a genetic superiority of herds with greater intraherd milk variance. However, results of progeny tests, estimated by young bull initial PD, showed the opposite trend. Greatest average young bull PD was seen in the low-variance group. Low-variance group was followed by mid and then high-variance groups, which were nearly equal in average PD. The interval on PD in the low-variance group did not overlap PD intervals of either mid or high-variance groups. However, intervals on PD in mid and high-variance groups overlapped very much. Young bulls selected from herds with low-variance had greater average genetic merit than bulls selected from herds with higher variance.

In the low-variance group, average PI under estimated average PD, while average ETA was more accurate. The interval on ETA was entirely contained within the interval on PD. The interval on PI overlapped the interval on PD, but was not entirly contained within it.

In mid and high-variance groups, both average PI and average ETA over estimated average PD significantly. Intervals on both ETA and PI did not overlap the interval on PD. The degree of over prediction was much worse in the high-variance group than in the mid-variance group. As herd variance increased, accuracy of predictions decreased.

Table 17. Correlations between young bull pedigree information and young bull predicted difference within variance group.

	Intraherd Milk Variance Group			
Pedigree Information	Low	Mid	High	Total
Dam CI	.15	.48	.29	.28
Sire PD	-.15	.35	.54	.19
MGS PD	. 24	.23	.19	.20
PI	-.02	.37	.57	.24
ETA	-.02	.48	.49	.29

Maximum theoretical correlations between various pieces of pedigree information and young bull PD are discussed. However, these maximum theoretical correlations are based on the assumption that genetic values are known perfectly and that we are dealing with an unselected population. In a selected population, as we have in this group of young bulls, is is unlikely that the maxixmum theoretical correlation will be observed. Correlations in a selected group of animals are expected to be lower than the theoretical maximum.

In theory, the maximum correlation between dam CI and young bull PD is .50. Observed correlation between dam cow index and young bull PD was only moderate in the total data set. Correlation was greatest in the mid-variance group and was very nearly equal to the theoretical value. Correlation was moderate in the high-variance group followed by the low-variance group, in which the correlation was low.

Theoretical maximum correlation between sire PD and young bull PD is also .50. In the total data set, this correlation was low. It was negative in the low-variance group, became positive and moderate in the mid-variance group, and was close to the theoretical maximum correlation in the high-variance group.

Theoretical maximum correlation between maternal grandsire PD and young bull PD is .25. Correlations in the total data set and in all variance groups were nearly equal to the theoretical correlation.

Theoretical maximum correlation between young bull pedigree index (PI) and PD is .56. In the total data set, this correlation was less than half of the theoretical value. In the low-variance group the correlation was negligible. Correlation was moderate in the mid-variance group and was close to the theoretical value in the high-variance group.

Theoretical maximum correlation between ETA and PD is .71. This correlation was very low in the total data set and was again negligible in the low-variance group.

Correlation was approximately 70% of the theoretical correlation in both mid and high-variance groups.

As discussed in the literature review, heritability for milk increases with increased levels of herd milk production and herd variance. Differences in correlations between high and low-variance groups reported in this study may be related to different heritiabilities, or may be due to the small number of observations in each variance group, or may be due to a combination of both.

Discussion: Grouping young bulls by intraherd variance of the herd from which they were selected revealed some interesting patterns. Young bulls selected from low-variance herds can, as a group, be predicted with a great deal more accuaracy than young bulls selected from high-variance herds. However, correlations between pedigree information and young bull PD were greater in the high-variance group than the low-variance group, indicating our ability to rank young bulls is better in the high-variance group than in the low-variance group. Two studies were carried out to examine these trends in more detail.

5.3 BIAS IN PREDICTED TRANSMITTING ABILITY OF YOUNG BULLS GROUPED BY INTRAHERD MILK VARIANCE: The objective of this study was to examine mean differences among the three variance groups. Characteristics examined were: merit of contracted matings, as measured by ETA; young bull transmitting ability, as measured by PD; and young bull deviations of predicted transmitting ability from real transmitting ability, as measured by (PD-ETA).

Data: This study was accomplished using the same variance groupings defined in section 5.1.

Methods: Ordinary least squares procedures were used to analyze the following three fixed classification models:

1) $ETA_{ij} = \mu + V_i + e_{ij}$
2) $PD_{ij} = \mu + V_i + e_{ij}$
3) $(PD-ETA)_{ij} = \mu + V_i + e_{ij}$

where ETA_{ij} is the genetic index of the jth young bull in the ith variance group;

PD_{ij} is the initial PD estimate of the jth young bull in the ith variance group;

$PD-ETA_{ij}$ is the deviation of predicted transmitting ability from measured transmitting ability for the jth young bull in the ith variance group;

μ is a constant common to all observations;

V_i is the fixed effect of the ith variance group, i=1,2,3;

e_{ij} is the random residual associated with ETA_{ij}, PD_{ij}, and $(PD-ETA)_{ij}$ respectively.

The primary interest in this analysis was to compare means of the three variance groups. Orthogonal contrasts, $(V_1+V_3)-2V_2$ and V_1-V_3, were used for each of the three models. The first contrast tested whether means of high and low-variance groups were significantly different from the mean of the mid-variance group. The second contrast tested whether means in high and low-variance groups were significantly different from each other. Results are repoted in Table 18.

Mean deviation (PD-ETA) was also examined for significant difference from zero within each variance group. If no bias has been introduced into calculations of PD or ETA, then deviations (PD-ETA) are random. Under estimated PD's would cancel with over estimated PD's, and the average deviation (PD-ETA) is zero. Mean and standard deviation for (PD-ETA) were computed in each variance group. Confidence intervals on means were constructed. If the confidence interval failed to include zero, then a deviation different than zero was assumed with a significance level of α. Results are reported in Table 19.

Table 18. Significance levels for factors and contrasts in classification models involving herds grouped by level of intraherd milk variance.

Model Factors and Contrasts	Dependent Variable		
	ETA	PD	PD-ETA
Intercept	.05	.001	.001
Variance Group	.18	.17	.02
adj. R^2	.019	.020	.073
$(V_1+V_3)-2V_2$	.50	N.S.	.50
V_1-V_3	.10	.10	.01

Results and Discussion: Variance group was not a significant source of variation in predicting young bull ETA or PD. However, it was a significant source of variation in predicting PD-ETA. Results of this model showed significant differences among variance groups in accuracy of PD prediction.

High and low-variance groups were not significantly different from the middle variance group in any of the three models. However, high and low-variance groups were significantly different from each other at α=.10 or greater, in all three models. The most highly significant difference between high and low groups was in the deviation model.

Results of the first model indicated that the average mating contracted in the low-variance group was significantly different from the average mating contracted in the high-variance group. Results of the second model indicated that transmitting ability of an average young bull obtained from the low-variance group was significantly different from the transmitting ability of an average young bull obtained from the high-variance group. Results of the third model indicated that average deviation (PD-ETA) of an average young bull in the low-variance group is significantly different from the average deviation of a young bull in the high-variance group.

Table 19. **Mean, standard deviation, and significance level of difference from zero for (PD-ETA) in intraherd variance groups.**

Variance Group	Mean	Standard Deviation	Significance Level
Low	12	601.27	N.S.
Mid	-272	486.06	.01
High	-386	472.19	.001

The average deviation (PD-ETA) was not significantly different from zero in the low-variance group, while it was significantly different from zero in the mid and high-variance groups. The amount by which ETA over predicted PD increased with variance level. ETA was an unbiased predictor in the low-variance group, while it was a biased predictor in the mid and high-variance groups. Biasedness increased with variance level.

5.4 BIAS IN DISTRIBUTION OF YOUNG BULLS GROUPED BY INTRAHERD MILK VARIANCE: The objective of this study was to examine differences among the three variance groups in distributions of ETA, PD, and deviation (PD-ETA).

Data: Variance groups used in this study are teh same groups that were defined in section 5.1.

Methods: Contingency tables were constructed to test independence of variance group and distribution of ETA, PD, and (PD-ETA). ETA and PD were classified into two classes; less than or equal to the mean, or greater than the mean. The deviation (PD-ETA) was classified into classes of less than or equal to zero, or greater than zero. This classification is based on the assumption that for a group of young bulls, deviations of predicted transmitting ability from real transmitting ability will average zero. Tests of binomial proportion were done for the total distribution across all variance groups and for independent distributions within each variance group. Results are shown in Tables 20, 21, and 22 for ETA, PD, and (PD-ETA) respectively.

Table 20. Test for independence of ETA distributions in intraherd variance groups.

	Contingency Table[1]			Binomial Proportion	
Variance Group	ETA$\le\bar{X}$[2]	ETA>$\bar{X}$	Total	Test Statistic	Significance Level
Low	21	10	31	3.226	.08
Mid	13	18	31	.516	N.S.
High	13	18	31	.516	N.S.
Total	47	46	93	0	N.S.

[1] test statistic=5.506 (α=.07)

[2] $\bar{X}$=343.6

Results: Variance group was not independent of ETA distribution. Tests of binomial proportion indicated that matings in the low-variance group were significantly different than the 50/50 split expected and only 38% of the matings were above average. In the mid and high-variance groups 58% of the matings were above average. This was not not significantly different from the 50/50 split expected.

Table 21. Test for independence of PD distributions in intraherd variance groups.

	Contingency Table[1]			Binomial Proportion	
Variance Group	PD$\le\bar{X}$[2]	PD$>\bar{X}$	Total	Test Statistic	Significance Level
Low	11	20	31	2.065	.17
Mid	16	15	31	0	N.S.
High	15	16	31	0	N.S.
Total	42	51	93	.688	N.S.

[1] test statistic=1.824 (α=.30)

[2] $\bar{X}$=127.2

Variance group was independent of PD distribution. Frequencies of above and below average PD split very evenly in mid and high-variance groups. However, in the low-variance group, 65% of young bulls had above average PD's. Although the test of binomial proportion indicated this was not significantly different from the 50/50 split expected, it is especially noteworthy since only 32% of the young bulls were predicted to be above average.

Table 22. Test for independence of deviation (PD-ETA) distributions in intraherd variance groups.

	Contingency Table[1]			Binomial Proportion	
Variance Group	Dev.≤0	Dev.>0	Total	Test Statistic	Significance Level
Low	15	16	31	0	N.S.
Mid	20	11	31	2.065	.17
High	22	9	31	5.633	.02
Total	57	36	93	4.301	.05

[1] test statistic=3.535 (α=.18)

Variance group was independent of distribution of deviations (PD-ETA). In the low-variance group, the distribution was not significantly different from the 50/50 split expected. While the distribution in the mid-variance group was also not significantly different than a 50/50 split, 65% of the young bulls had negative deviations. In the high-variance group, however, there were significantly more negative deviations than positive deviations. The overall test for binomial proportion indicated that in the total data set there was not a 50/50 split, but significantly more negative deviations than positive deviations.

The assumption that average deviation equalled zero, while theoretically correct, was not upheld within this data set. For this reason, all ETA's were adjusted. The average deviation (PD-ETA) was added to all ETA's so that, for the total data set, average PD equalled average ETA and the average deviation was zero. Tests for independence and binomial proportion were done again using adjusted ETA's and results are shown in Table 23.

Table 23. Test for independence of deviation (PD- Adj.ETA) distributions in intraherd variance groups.

	Contingency Table[1]			Binomial Proportion	
Variance Group	Dev.≤0	Dev.>0	Total	Test Statistic	Significance Level
Low	11	20	31	2.065	.17
Mid	16	15	31	0	N.S.
High	20	11	31	2.065	.17
Total	47	46	93	0	N.S.

[1] test statistic=5.248 (α=.08)

After adjusting all ETA's, variance group was not independent of the deviation distribution. Adjusting ETA's moved the 50/50 split from the low-variance group to the mid variance group. It also revealed a reversing pattern of skewness in high and low-variance groups. In the low-variance group, 65% of the young bull PDs were underestimated by ETA, while in the high-variance group 65% were overestimated. However, tests of binomial proportion showed these percentages were not significantly different from the 50/50 split expected.

Discussion: That variance group was not independent of ETA indicated that matings were not arranged randomly. Matings contracted in different herd variance groups were not of equal predicted merit. Prediction of young bull PD was very accurate in low-variance herds while it was very inaccurate in high-variance herds. The test of binomial proportion for the deviation (PD-ETA) showed that ETA is unbiased in low-variance herds while it was significanlty biased in high-variance herds. Adjusting ETA's removed the bias so that no variance group was significantly different from having half the deviations above zero and half below zero. However, the test for independence on the deviation (PD-adj.ETA) revealed that the distribution of deviation is not independent of herd variance group.

6. SUMMARY AND CONCLUSIONS

The primary objective in dairy cattle breeding is to increase milk production through the selection and breeding of genetically superior animals. While much progress has been achieved, the rate of actual genetic gain for milk production has not equalled the rate of gain which is theoretically possible.

Selection of young bulls entering the population accounts for 71% of the annual genetic gain (Van Vleck, 1986a). Therefore, efforts to improve rate of actual genetic gain have focused much attention on ways in which selection programs might be improved.

The objective of this research was to examine ways in which herd information might improve young sire selection programs. Three investigations, each composed of several studies, were completed.

In the first investigation, it was observed that Michigan herd groups differed in level of milk production, variance of milk production, and genetic merit. Herds from which young sires had been selected had greater average milk production and greater variance for milk production than other herd groups. However, these herds were not genetically superior to other herd groups. Additionally, many cows selected as bull dams were not within the top 5% of their respective herds when ranked by cow index or milk production.

In the second investigation it was seen that herds which were able to return to service an acceptable proportion of young bulls sampled had lower average milk production and smaller variance of milk production than herds which were unable to return to service an acceptable proportion of young bulls sampled. It was also seen that intraherd variance contributed significantly to prediction of young bull transmitting ability.

In the third investigation, it was observed that significant differences in means and in distributions of PD, ETA, and (PD-ETA) existed between herds grouped by level of intraherd milk variance. Matings contracted in low-variance herds were not as of great a predicted genetic merit as matings contracted in herds of higher variance. However, young bulls from low-variance herds had greater average transmitting abilites than young bulls selected from herds with greater variance. Additionally, prediction of young bull PD was very accurate in herds of low-variance, while it was very biased in herds with greater variance. In low-variance herds, ETA was an extremely accurate predictor of young bull PD.

Heterogeneous intraherd variance appears to be biasing young bull predictors. As level of herd variance increases, predicted values, particularly ETA, become inflated and over estimate young bull transmitting abilities. Failure of evaluation procedures, particularly cow evaluations, to account for heterogeneous herd variance is the most likely channel through which this bias is introduced.

The goal in young sire selection is to achieve accurate group selection. Correlations of predicted genetic merit and true genetic merit are unimportant if the average transmitting ability of the selected group misses the predicted target. Committees which arrange matings to produce young bulls for AI sampling must be aware of the bias existing in young bull predictors. Accurate group selection can be achieved if young bulls are not chosen from herds of high intraherd variance. Intraherd variance can, therefore, be used to identify herds which should be avoided when contracting matings to produce young bulls. Additionally, these committees should note that in low-variance herds, ETA is a more accurate predictor than PI. Setting minimum standards on young bull PI is frequently done to avoid biasesd cow evaluations. If matings are contracted in low-variance herds, there seems little reason to maintain thi practice, and greater emphasis should be placed on ETA.

While intraherd variance shows tremendous potential for improving young sire sampling programs, this research involved only 93 young bulls and 32 herds from one state in the country. Before these recommendations are implemented, it may be wise to confirm them with a larger data set and one that incorporates more regions of the country.

7. REFERENCES

Bell, B.R., B.T. McDaniel, and O.W. Robison. "Effects of Cytoplasmic Inheritance on Production Traits of Dairy Cattle," J. Dairy Sci. 68:2038-2051, 1985.

Bolgiano, D.C., L.D. Van Vleck, and R.W. Everett. "Fluctuations in Sire Evaluations," J. Dairy Sci.. 62:760-767, 1979.

Brotherstone, Susan, and W.G. Hill. "Heterogeneity of Variance Amongst Herds for Milk Production," Anim. Prod.. 42:297-303, 1986.

Burnside, E.B., and K. Kuersten. "Trends in Generation Interval for AI Sires," Paper presented to CAAB Sire Analysts Conference, cited by L.D. Van Vleck in "Evaluation of Dairy Cattle Breeding Programs: Specialized Milk Production," Proc. 3rd World Congress on Genetics Applied to Livestock. Vol. XII: 88-95, 1986.

DeVeer , J.C., and L.D. Van Vleck. "Comparison of Heritability Estimates from Daughter on Dam Regression with Three Models to Account for Production Level of Dam," J. Dairy Sci.. 69:2890-2896, 1986.

Dickerson, G.E., and L.N. Hazel. "Effectiveness of Selection on Progeny Performance as a Supplement to Earlier Culling in Livestock," J. of Agricultural Research. 69:459-477.

Dunn, L.C.. A Short History of Genetics. New York: McGraw-Hill Book Company, 1965.

Everett, R.W.. "Genetic Potential of Grade AI Holsteins," Animal Science Mimeograph Series, No. 78. Ithaca, New York: New York Cooperative Extension Service, 1984.

Everett, R.W., J.F. Keown, and J.F. Taylor. "The Problem of Heterogeneous Within Herd Error Variances When Identifying Elite Cows," J. Dairy Sci.. 65(suppl.):100, 1982.

Henderson, C.R.. "Selecting the Young Sire to Sample in Artificial Insemination," J. Dairy Sci.. 47:439-441, 1964.

Hill, W.G., M.R. Edwards, M.K.A. Ahmed, and R. Thomas. "Heritability of Milk Yield and Composition at Different Levels of Variability of Production," Anim. Prod.. 36:59-68, 1983.

Hintz, R.L., R.W. Everett, and L.D. Van Vleck. "Estiamtion of Genetic Trends From Cow and Sire Evaluations," _J. Dairy Sci._. 61:607-613, 1978.

Lee, K.L., and A.E. Freeman. "Estimation of Genetic Change in the Registered Holstein Cattle Population," _J. Dairy Sci._. 68:2629-2638, 1985.

Lofgren, L.D., W.E. Vinson, and R.E. Pearson. "Adjustments to Cow Indexes for Milk Yield for Effects of Herd Mean and Standard Deviation," _J. Dairy Sci._. 68:3301-3311, 1985.

Mather, K.. _Biometrical Genetics, The Study of Continuous Variation_. New York: Dover Publications, Inc., 1949.

Mirande, S.L., and L.D. Van Vleck. "Trends in Genetic and Phentypic Variances for Milk Production," _J. Dairy Sci._. 68:2278-2286, 1985.

Moran, P.A.P., and C.A.B. Smith. Commentary on R.A. Fisher's Paper on "The Correlation Between Relatives on the Supposition of Mendelian Inheritance." New York: Cambridge University Press, 1966.

Murphy, P.A., R.W. Everett, and L.D. Van Vleck. "Comparison of First Lactations and All Lactations of Dams to Predict Son's Milk Evaluations," _J. Dairy Sci._. 65:1999-2005, 1982.

Norman, H.D.. "Sire Evaluation Procedures for Yield Traits," Fact Sheet H-1. Beltsville, Maryland: Agricultural Research Sevice, National Cooperative Dairy Herd Improvement Program, 1986.

Pearson, R.E.. "Response to Selection in Dairy Cattle," _Proceedings of National Invitational Workshop on Genetic Improvement of Dairy Cattle_. New York: Cornell University, 1984.

Powell, Rex L.. "Genetic Base for Cow Evaluation," _J. Dairy Sci._ 67:1359-1363, 1984.

Powell, Rex L.. "Cow Evaluation Procedures," Fact Sheet H-2. Beltsville, Maryland: Agricultural Research Service, National Cooperative Dairy Herd Improvement Program, 1985.

Powell, R.L., H.D. Norman, and B.T. Weinland. "Cow Evaluation at Different Milk Yields of Herds," _J. Dairy Sci._. 66:148-154, 1983.

Rendel, J.M., and Alan Robertson. "Estimation of Genetic Gain in Milk Yield by Selection in a Closed Herd of Dairy Cattle," J. Genetics. 50:1-8, 1950.

Robertson, Alan, and J.M. Rendel. "The Use of Progeny Testing with Artificial Insemination in Dairy Cattle," J. Genetics. 50:21-31, 1950.

Rothschild, M.F., L.W. Douglass, and R.L. Powell. "Prediction of Son's Modified Contemporary Comparison From Pedigreee Information," J. Dairy Sci.. 64:331-341, 1981.

Smith, C.. "Estimation of Genetic Change in Farm Livestock Using Field Records," Anim. Prod.. 4:239-251, 1962.

Van Vleck, L.D.. "Theoretical Weights for Regression of Son's Genetoc Evaluation on His Sire's and His Dam's Genetic Evaluations," J. Dairy Sci.. 65:164-167, 1981.

Van Vleck, L.D.. "Evaluation of Dairy Cattle Breeding Programs: Specialized Milk Production," Proceedings of 3rd World Congress on Genetics Applied to Livestock. Vol. IX:141-152, 1986a.

Van Vleck, L.D.. "Technology and Animal Breeding: Applications and Challenges for Dairy Cattle Breeding," Proceedings of 3rd World Congress on Genetics Applied to Livestock. Vol. XII:88-95, 1986b.

Van Vleck, L.D., L.R. Crow, and S.L. Mirande. "Heritability Estimates of Milk Production From Daughter on Dam Regression by Year and Management Level," J. Dairy Sci.. 68:2964-2969, 1985.

Van Vleck L.D., and John Pollak. "Sire Evaluation Methods: Past, Present, Future," Proceedings of the Ntional Invitational Workshop on Genetic Improvement of Dairy Cattle. Ithaca, New York: Cornell University, 1984.

Van Vleck, L.D., E.J. Pollak, and E.A.Branford Oltenacu. Genetics for the Animal Sciences. New York: W.H. Freeman and Company, 1987.

Van Vleck, L.D., R.A. Westell, and J.C. Schneider. "Genetic Change in Milk Yiled Estimated From Simultaneous Genetic Evaluation of Bulls and Cows," J. Dairy Sci.. 69:2963-2965, 1986.

Van Vleck, L.D.. "Theoretical and Actual Gentic Progress in Dairy Cattle," Proceedings of International Conference on Quantitative Genetics. Ames, Iowa: Iowa State University Press, 1977.

Vinson, W.E., amd J.M. White. "Prediction of Modified Contemporary Comparisons from Pedigree Information for Artificial Insemination and Natural Service Bulls," _J. Dairy Sci._. 65:1578-1587, 1982.

Voelker, D.E.. "History of Dairy Recordkeeping," Fact Sheet A-2. Beltsville, Maryland: Agricultural Research Service, National Cooperative Dairy Herd Improvement Program, 1985.

Westell, R.A.. Simultaneous Genetic Evaluation of Sires and Cows for a Large Population of Dairy Cattle," Cornell University: PhD thesis, 1984. cited by L.D. Van Vleck in "Evaluation of Dairy Cattle Breeding Programs: Specialized Milk Production," _Proc. 3rd World Congress on Genetics Applied to Livestock_. Vol XII: 88-95, 1986.

Wiggans, George R., H. Duane Norman, and Rex L. Powell. "Changes in Genetic Evaluation Procedures for January 1984," Dairy Herd Improvement Letter. Beltsville, Maryland: Animal Improvement Programs Laboratory, Animal Science Institute, Beltsville Agricultural Research Center, 1984.

Wright, Sewall. "Systems of Mating and Other Papers." Ames: Iowa State College Press, 1958.