

**Accuracy of Predicting Genetic Merit of A.I. Sampled Bulls for Final Score
from Pedigree Information.**

by

James Rohl

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State
University in partial fulfillment of the requirements of the degree of

Master of Science

in

Dairy Science


APPROVED:




Dr. R.E. Pearson, Chairman



Dr. B.G. Cassell



Dr. D.R. Notter



Dr. W.E. Vinson, Department Head

August 1994

Blacksburg, Virginia

C.2

LD
5655
V855
1994
R.645
C.2

**ACCURACY OF PREDICTING GENETIC MERIT OF A.I. SAMPLED
BULLS FOR FINAL SCORE FROM PEDIGREE INFORMATION**

By

James Rohl

Ronald E. Pearson, Chairman

Dairy Science

(ABSTRACT)

A total of 1926 A.I. sampled Holstein bulls born from 1984 to 1988 and with first proofs from Summer 1991 to Summer 1993 were used to determine the accuracy of predicting PTAT and DTD from different sources of pedigree information obtained before the bull had daughter information. Pedigree sources used were PA, PI, PTAT_{SIRE}, and PTAT_{DAM}. Simple linear regression was used to determine which pedigree source predicted PTAT or DTD with the highest accuracy (highest R^2). R^2 was higher for PA than had the other pedigree sources. R^2 s for PA to predict initial PTAT and DTD with daughter information were .59 and .18 respectively. Higher weights and R^2 s for PTAT than DTD resulted from the part whole relationship between PA and PTAT. Accuracy of prediction varied depending on when a bull received his first proof. R^2 values for PA to predict initial PTAT ranged from .35 to .69, and increased as the time of the pedigree estimate approached the date of the initial proof. R^2 values for PA to predict initial DTD ranged from .16 to .21 and increased as the time of the pedigree estimate approached the date of the initial proof.

The impact of the within herd variance correction which was implemented in the Summer 1993 summary was also evaluated. Correlations between PA, PI, PTAT_{SIRE}, and PTAT_{DAM} from the Winter 93 and Summer 93 evaluations were .98, .99, .99, and .96 respectively. Regression of the change in DTD estimated from previous PA minus actual DTD on PTAT_{DAM} S 93 - PTAT_{DAM} W 93 for bulls grouped by date of initial proof gave R^2 s from

.00 to .06. It was concluded that the variance correction had little impact on the dam's of bulls in this study.

The impact of the addition of granddaughters (son's daughters) on the PTA of the bull dam was evaluated. The mean change in $PTAT_{DAM}$ with the addition of first granddaughters was .016, indicating that the PTAT of the bull dam was slightly underestimated. R^2 s for the regression of the change in bull dam's PTAT on DTD, DTD- PA, and PTAT - PA were .39, .54, and .56 respectively.

Little evidence was found to indicate a bias based on the testing population used to prove the bull. R^2 s for the regression of PA and PI on PTAT from the bull's initial proof with daughter information ranged from .38 to .69, and .26 to .58 respectively. When PA and PI were used to estimate PTAT of a bull's second proof both within and across NAAB codes, values agreed closely.

Acknowledgements

Funding for this research was provided by 21st Century Genetics, Shiwano, Wisconsin. The data used was provided by the Holstein-Friesian Association of America.

I will be eternally grateful for Dr. Pearson's unfailing patience, advice, and generosity toward me while serving as my advisor. I also owe a large debt to the other members of my committee, Drs. D. R. Notter, B. G. Cassell, and W. E. Vinson. I have learned a great deal from each of these men, both in and out of the classroom. Another faculty member who was always willing to give advice and encouragement, and who added greatly to the quality of my time as a grad student was Dr. Dan Weigel.

Alan Pasquino, Cindy Cassady, and Erica Jones all deserve my deepest gratitude. Their patience in dealing with my computer questions made this degree possible.

My fellow grad students have done much to make my time at Virginia Tech such a positive experience. I have enjoyed knowing all of them. My long suffering office mates deserve special mention, as does Amin Ahmadzadeh, who was always ready to help in any way possible.

My deepest thanks go to my wife, Marilyn, whose constant support in every way has made all of this possible for me.

Table of Contents

Introduction	1
Review of Literature	3
Final Score: Its composition and heritability	3
Genetic Evaluations	6
Predicting Future Transmitting Abilities from Pedigree Information	9
Materials and Methods	14
Source of data	14
Data edits	14
Calculation of PA and PI	14
Definition of Initial Proof with daughter contribution	16
Accuracy of predicting bull's genetic merit from different pedigree sources	16
Assessment of effect of within herd variance correction	19
Impact of granddaughter information on Dam's PTAT	20
Assessment of impact of testing population	22
Results and Discussion	25
Predicting bull's initial proof with daughter information for PTAT and DTD from different pedigree sources.	25
Examination of PTAT by PA Decile.	43
Impact of the Variance correction.	47
Impact of Granddaughter information on Dam's PTAT	51
Assessment of impact of testing population	53
Conclusions	61
Bibliography	62
Vita	66

List of Tables

Table 1. Estimates of Heritability of Final Score . . . 5

Table 2. Number of bulls removed with each edit from bulls with birth years from 1984 to 1986 15

Table 3. Means and Standard Deviations for genetic values at time of first proof with progeny contribution greater than .5, and pedigree sources from immediately prior proof 26

Table 4. Regression of PTAT (N = 1926) and DTD (N = 150) for bull's initial proof 26

Table 5. Regression of PTAT (N = 1706) for bull's second proof on pedigree information from proof immediately prior to addition of progeny information 28

Table 6. Regression of DTD S 93 (N = 444) for bulls with initial proof in W 93 on S 92 pedigree information 28

Table 7. Means and SDs for genetic parameters of Final Score of bulls grouped by date of initial proof . . 30

Table 8. Means and SDs for Progeny Contribution (PC) for PTAT of bulls grouped by date of first proof with greater than 1 daughter 32

Table 9. Sires with greater than 20 sons in the data set.33

Table 10. Means and SDs for MGS PTAT by date of initial proof 34

Table 11. Regression of PTAT of bull's initial proof on PA from previous evaluations. 37

Table 12. Regression of PTAT of bull's initial proof on PI from previous evaluations. 37

Table 13. Regression of PTAT_{SIRE} of initial proof on PI from

previous evaluations.	38
Table 14. Regression of $PTAT_{DAM}$ of bull's initial proof on PI from previous evaluations.	38
Table 15. Regression of DTD S 93 on PA from all previous evaluations.	41
Table 16. Regression of DTD S 93 on PI from all previous evaluations.	41
Table 17. Regression of DTD S 93 on $PTAT_{SIRE}$ from all previous evaluations.	42
Table 18. Regression of DTD S 93 on $PTAT_{DAM}$ from all previous evaluations	42
Table 19. Means for PA previous to progeny information, $PTAT$ for initial proof, $PTAT$ for S93, and DTD S93 .	44
Table 20. Correlations between W 93 and S 93 proofs for PA, PI, $PTAT_{SIRE}$, and $PTAT_{DAM}$ by date of first proof group, and for all bulls	48
Table 21. R^2 s, regression coefficients, and SEs for the regression of $(DTD_{ESTIMATED} - DTD)$ on $(PTAT_{DAM} S 93 - PTAT_{DAM} W 93)$ by date of proof	50
Table 22. Means and SDs for $(PTAT_{DAM} S 93 - PTAT_{DAM} W 93)$ by date of proof	50
Table 23. Means and SDs for $\Delta PTAT_{DAM}$ by date of addition of granddaughter information	52
Table 24. R^2 s, regression coefficients, and SEs for the regression of $\Delta PTAT_{DAM}$ the son's DTD S 93 and DTD S 93 - $PA_{SON} W 93$	52
Table 25. Means and SDs for PI and PA by NAAB code . . .	56
Table 26. Means and SD for genetic values for FS by NAAB code	56
Table 27. Regression Coefficients, R^2 s, and intercepts for regression of $PTAT$ on PA and PI for different NAAB codes	57

Table 28. Means and SDs for PTAT at the time of the bull's
second proof estimated from PA and PI of the proof
immediately prior to his initial proof by NAAB
code 58

Introduction

When a young bull is selected to enter a progeny testing program, only pedigree information is available to estimate his genetic merit. The accuracy of pedigree information is directly related to the accuracy of prediction of the bull's daughter performance. Because the selection of young bulls to be progeny tested can account for much of the genetic gain possible in dairy cattle (Van Vleck, 1977), and because of the high cost involved in proving bulls (Henderson, 1964 Vinson, 1972), it is important that the pedigree information be accurate.

Over the past 5 years, genetic evaluations for both type and production have changed from sire model approaches to animal model evaluations. These evaluations are calculated for both males and females simultaneously and allow all relatives to impact the evaluation of an individual. The estimated transmitting abilities for yield and type traits of both sexes are called predicted transmitting abilities (PTA) and include information contributed by the animal's pedigree, own performance, and progeny.

The PTA may be depicted as:

$$PTA = w_1 PA + w_2 (YD/2) + w_3(2PTA_{prog} - PTA_{mate})$$

where PA = the average PTA of the sire and dam,
 YD = the yield (or type performance) deviation
 from the management group mean,
 PTA_{PROG}, PTA_{MATE} = PTA of the offspring and
 the other parent of the offspring.

The values for the w's depend upon the amount of information from each source and always sum to one. Initially



$w_1 = 1$ but as progeny information becomes available w_1 decreases and w_3 increases. Estimates of the PTAs are obtained by iteration and are dependent upon information from all relatives.

The change in method of genetic evaluation has been shown to affect the value of different measures of pedigree information in predicting the PTA of a bull for production traits (Samuelson, 1992). The relationship between pedigree estimates and PTA of a sire for type (PTAT) have not been investigated in part because of the shorter time frame since the animal model evaluation for type was initiated.

The objectives of this study were:

- 1) To determine the accuracy of predicting the genetic merit for Final Score of a bull entering a progeny test program from four different sources of pedigree information.
- 2) To assess the impact of the within herd variance correction implemented in the Summer, 1993 evaluations.
- 3) To assess the impact of granddaughter information on $PTAT_{DAM}$.
- 4) To assess the impact of the testing population as reflected by A.I. stud proving the bull on the mean and range of the resulting PTAT of a group of bulls.

Review of Literature

Final Score: Its composition and heritability

Short and Lawlor (1992) discussed the history and format of the Holstein-Friesian Association of America's type classification program. The classification program was started in 1926 for registered cattle only. Cows were given a numerical score for four major categories and overall conformation (final score; FS). In 1967, descriptive categories were recorded for eleven traits. The Sire Evaluation for Type (SET) program was initiated in 1976, which allowed unregistered daughters to contribute information to a bull's type proof. Linear classification, to score traits on a continuous scale was begun in 1983. Approximately 600,000 cows were scored in 1990, about 20% of which were grades. Response to selection for FS was approximately .13 points per year. The mean and standard deviation (SD) for FS in the registered, grade, and combined populations were 80.8 and 3.9, 78.0 and 4.7, and 79.8 and 4.2 respectively. Selection for higher FS was felt to be more effective in increasing herd life than direct selection (Short and Lawlor, 1992).

Wilcox et al. (1957) performed an experiment in which three professional classifiers scored Holstein cows every 6 months at the same time but independently of each other. Correlation among the classifiers was .69 for FS. For the same cow, repeatability of FS was .44.

Atkeson et al. (1969) analyzed final score of 12,890 Holstein cows. He found a mean FS of 81.23 with a SD of 3.58. Correlations between FS and general appearance, mammary system, rear udder, fore udder, body capacity, feet and legs, dairy character, and rump were .86, .78, .74, .67, .65, .57,

.57, and .63 respectively. Body capacity and dairy character were underemphasized and general appearance and mammary system overemphasized relative to weights on the Dairy Cow Unified Score Card.

Berland (1993) discussed the relative changes in the components that make up FS in Holstein cows. At the time the cows in this study were scored, FS was made up of 30% general appearance, 20% dairy character, 20% body capacity, and 30% mammary system. In May 1993, weights were changed to 15% frame, 20% dairy character, 10% body capacity, 15% feet and legs, and 40% udder.

Several estimates of the heritability of FS have been calculated (see Table 1). Cassell et al. (1973) analyzed data on 336,253 Holstein cows, and calculated a heritability of .31 for FS. The mean and standard deviation of FS in this data set were 80.2 and 3.7 respectively.

Vinson et al. (1975) calculated a mean and SD of 80.1 and 3.4 for FS on 78,151 Holstein cows in 2117 herds. Herds, classifiers, and herd x classifier interaction accounted for 13.8, 2.6 and 5.9 per cent of the total variance, respectively. Regression coefficients for general appearance, dairy character, body capacity, and mammary system were .435, .171, .182, and .463, with R^2 of .826.

VanRaden et al. (1990) evaluated data on 779,391 Holstein daughters of 871 sires in the January 1988 sire evaluation, and obtained estimates of heritability of final score of .27 and .29, depending on whether or not adjustment for the merit of mate's (dam's PTA) was considered. Rank correlations between adjusted and unadjusted evaluations was .996.

Short and Lawlor (1992) estimated heritabilities for FS from first lactation records of 128,601 Holstein cows. Heritability of FS was calculated to be .30 for registered

Table 1. Estimates of Heritability of Final Score

Source	Estimate
Cassell et al. (1973)	.31
Smothers et al. (1993)	.21
VanRaden et al. (1990)	
without merit of mates	.27
with merit of mates	.29
Short and Lawlor (1992)	
registered	.30
grade	.16
combined	.26

cows, .16 for grades, and .26 for the combined population.

Smother's et al. (1993) analyzed first parity classification scores for Holstein cows in 7410 herds. Heritability estimate for FS was .21. FS environmental and genetic variance decreased as the mean FS for the group increased. The same situation was found for most of the individual traits scored. Heritability was found to not change as herd mean increased. Smother's concluded that variances across herds for type traits were heterogeneous, and that within herd variances for type traits decreased with herd mean in contrast to production traits, with herds having higher mean scores having lower variances than herds with lower mean scores. Three possible reasons given were 1) closed upper end of the scale, 2) unequal category range, and 3) selection and culling.

Genetic Evaluations

In July, 1989, USDA changed the genetic evaluation procedure from the Modified Contemporary Comparison (MCC) to Best Linear Unbiased Prediction (BLUP) under an Animal Model (VanRaden and Wiggans, 1989). In the Animal Model, information from all relatives is used through the process of iteration to calculate PTAs. PA in the animal model replaces PI in the MCC to account for the ancestors' contributions. In theory, the animal model provides a more accurate prediction of genetic merit due to the inclusion of all relatives.

In the animal model, the formula used for PTA is:

$$PTA = w_1 PA + w_2 (YD/2) + w_3 (2PTA_{prog} - PTA_{mate})$$

where:

$$PA = .5(PTA_{SIRE}) + .5(PTA_{DAM})$$

YD= Average yield deviation of animal from herd management group mates (adjusted for permanent environment and sire x herd interaction).

PTA_{PROG}= the PTA of progeny of the animal evaluated.

PTA_{MATE}= the PTA of the other parent of the progeny

The weights w_1 , w_2 , and w_3 are based on the amount of information from each source and sum to one. Because bulls have no production data and their scores are not included in type evaluations, w_2 is zero for bulls. As an animal gains more progeny or individual records, more weight is placed on these and less on PA. PTA_{MATE} is subtracted from the final term to account for the genetic merit of the other parent of any progeny (i.e to adjust for unequal merit of mates).

The PTAs of the progeny and parent are interdependent. Because of this, $2PTA_{PROG} - PTA_{MATE}$ is not independent of PA. A more independent measure of progeny performance is daughter yield deviation (DYD) (VanRaden and Wiggans, 1991). DYD is calculated using the formula:

$$DYD = \frac{\sum q_{prog} w_{2\ prog} (YD_{prog} - PTA_{mate})}{\sum q_{prog} w_{2\ prog}}$$

where q_{PROG} equals 1 if mate is known and 2/3 if mate is unknown.

For type traits, the chronology for genetic evaluations is as follows (Dr. Tom Lawlor, personal communication): In 1980, BLUP sire evaluations were initiated; in 1982 gross correction for heterogeneous variances was implemented; in 1986 data used were switched from last score to first score; in 1988 relationships and unknown parent groups were included in the BLUP analysis; in 1991 an animal model using all records was initiated. In the July 1993 genetic evaluation, further statistical techniques to compensate for heterogeneous variance were implemented (Weigel and Lawlor, 1994).

PTATs were calculated using the deviation of a cow's age adjusted score from the average age adjusted scores of other cows in the same herd scored by the same scorer on the same day. The Animal Model used employs the same model and approach as the one used for production traits. Similarly, starting in 1993 a Daughter Type Deviation (DTD) was calculated in the same manner as DYT for production traits with the type deviation replacing the YD in the formula above.

In order to standardize within management group variation, the following procedure was implemented with the S 93 evaluation (Weigel and Lawlor, 1994). For each herd-year-month of classification management group, an estimate of within management group FS standard deviation was calculated from a model containing known management characteristics. Prior estimates and phenotypic variance estimates were combined using Bayesian methods. These observations were used to standardize variation across herds by using the mean of the posterior density of the phenotypic standard deviation within subclass. The average standard deviation for 30-69 cow registered herds classified since 1990 was 3.95 (Weigel and Lawlor, 1994). Age adjusted scores for individual cows are expressed as standardized deviations from the average score within their respective management groups. Thus the greatest

impact will be on cows in herds with extremely high or extremely low average FS. In high scoring herds (low variance) deviations will be expanded and PTATs will be more variable (i.e. low scoring cows will have a more negative PTAT and high scoring cows will have a more positive PTAT). In low scoring herds (high variance) the opposite is true and the deviations will be contracted and PTATs will be less variable.

In the animal model, accuracy for a PTA is defined as reliability (REL). REL is the theoretical squared correlation of the animal's predicted and true transmitting abilities. For an animal with no records or progeny information, REL is one fourth the sum of the parent REL's, which is also the REL of PA. REL is used for both cows and bulls and is based on Daughter Equivalents (DE) of information from PA and progeny information. A DE is the amount of information contributed to a parent by a daughter with one record, many herd mates, and high REL for the opposite parent (VanRaden and Wiggans, 1991). The formula for REL is:

$$REL = DE / (DE + 14)$$

The DE for PTA is the sum of DE's from PA, the animal's own record for the trait, and the progeny records adjusted for mates.

Predicting Future Transmitting Abilities from Pedigree Information

Van Vleck (1977) reviewed estimates of the sources of genetic gain in dairy cattle and found that selection of the parents of young bulls to test in the next generation yielded up to 76% of the total gain possible. Bulls chosen for proving only have pedigree information on which to base selection. Henderson (1964) demonstrated that due to Mendelian sampling the theoretical upper limit of the

correlation between potential merit based on pedigree and actual breeding value was .71. When the sire has many daughters and the dam has many records and a large family, the correlation could reach .67 in practice. However, few cows selected to be bull dams have this amount of information available at the time of selection.

Van Vleck (1982) stated that the theoretical weights for sire and dam genetic evaluations used to predict a young bull's PTA are equal when the evaluation for the sire is only from his daughters and the dam evaluation is only from her records. Theoretical weight for each is .5. Due to the fact that the bull's transmitting ability is not perfectly estimated, prediction weights on the sire and dam are often less than .5, but may be equal. Including the MGS in the dam's evaluation changed the weight of regressing son on dam very little. Because MGS is included in the dam's evaluation, the theoretical weight of MGS is zero when a young bull's transmitting ability is regressed on sire's, dam's, and MGS's transmitting abilities. The expected values for PA and PI to predict the actual transmitting ability of a young bull are both 1.

Vinson and Freeman (1972) collected data from seven AI studs to examine the intensity of selection of bulls. They found that as a group the studs returned to active service approximately 25% of the bulls sampled, and that the returned bulls exceeded the mean of the group by 2.2 points for FS. Regressions of PD's on bulls pedigree estimates based on sire alone, dam alone, and PA gave coefficients of .4, .43 and .43 respectively for milk yield, and .41, .39, and .34 for fat yield. Regressions were not reported for type data. Correlations were .2, .11, and .22 respectively, for milk, and .19, .11, and .17 for fat.

Butcher and Legates (1976) investigated estimating

breeding values for milk yield from pedigree information using the dam's first three lactations, sire proofs based on daughter's first lactations, and MGS proof based on daughter's first lactation. They found correlations of .47 between son and son's pedigree index, .43 between son and sire, .24 between son and MGS, and .21, .16, .16, .08, and .08 between son and dam's lactations one through three. They concluded that later lactations should not be used or should be given less weight than earlier ones.

Rothschild et al. (1981) examined the value of pedigree information in predicting bull's Modified Contemporary Comparison (MCC). For predicting milk and fat, MGS was more useful than dam's index. However, for fat alone, the dam's index was more accurate than the MGS PD.

Funk and Hansen (1988) calculated pedigree indexes for milk, fat, and type on 380 Holstein bulls with progeny tests.

Only PDs of male ancestors were included in this index. They concluded that including a third generation in the index had some predictive value, but that a fourth did not.

Samuelson (1992) reviewed the literature and investigated the accuracy of predicting genetic merit for milk, fat, and protein, using records from 1,644 A.I. sampled Holstein bulls born from 1984 to 1986. Pedigree sources evaluated were PA, PI, PTA_{SIRE} and PTA_{DAM} . Weighted regression techniques were used to determine which source predicted daughter yield deviation of PTA with the highest accuracy based on R^2 . For all traits, PA had a higher R^2 for daughter yield deviation (DYD) and PTA than PI. R^2 values for PI to predict first PTA milk, fat and protein were .17, .20 and .18 respectively. R^2 for PA to predict first PTA milk, fat and protein were .47, .54 and .49 respectively. Adding PTA_{DAM} to the model with PTA_{SIRE} resulted in a higher R^2 than using PTA_{SIRE} alone. R^2

values for PA and the model with PTA_{SIRE} and PTA_{DAM} were similar. Weights and R^2 s were less than expected, but reasonable accuracy was obtained in estimating a bull's DYD and PTA from pedigree estimates. Accuracy of prediction varied depending on when the bull received his first proof. R^2 values of different groups of bulls based on the date of first DYD and PTA ranged from .06 to .20, .08 to .15 and .05 to .12 for predicting first DYD from PA for milk, fat and protein, respectively. When PTA_{DAM} was added to PTA_{SIRE} , R^2 increased to the level for PA, as it must unless PTA_{DAM} and $PTAT_{SIRE}$ are correlated.

In spite of the advanced statistical tools used for calculating type evaluations, few studies looking at different pedigree indexes for FS have been done. McNeill et al. (1975) took first available age adjusted classification scores for 302,921 Holstein cows and used them to compute pedigree evaluations and type proofs for bulls in A.I. The intensity was equal to that expected from truncation selection for the top 15% of cows and 2% of bulls in normal populations. Correlations between sons type proof and dams was .28, approximately 60% greater than expected, while sire and MGS correlations were .31 and .21 respectively. McNeil concluded that selection of young bulls for progeny testing based on pedigree would be effective, and that dam's final score would be an unexpectedly accurate predictor of her son's type proof.

Stewart et al. (1976) computed pedigree indexes for 216 Holstein bulls that entered Canadian AI units from 1962 to 1970. Progeny average trait score was regressed on pedigree index ($.5 \text{ PD sire} + .04 \text{ dam's score} + .2 \text{ PD MGS} + .02 \text{ MGD score}$) for trait score, sire's trait score proof, and sire and MGS proof at the time the bull entered the stud. Regression coefficients and standard deviations for the three indexes

respectively were .34 and .07, .17 and .14, and .32 and .07. It was concluded that pedigree index was most successful in predicting breeding value for type, followed by the sire score and MGS score.

Bell et al.(1982) analyzed official classification records of 755,744 Holstein cows in order to determine the effect of genetic merit of herdmates on sire evaluations for type. Progeny were divided into three groups based on the average final score for the herd, defined as low (< 79.75), middle (79.75 to 81.75), and high (> 81.75). PDs for type were calculated by group for all sires with more than 20 progeny per group. Mean type PD's were .12, -.28, and -.45, respectively, showing that non-random mating affected herdmate comparisons for type. The bias was due to the non-random use of sires and differential culling rates for type among herds. A positive difference for sires proven in lower scoring herds, and a negative difference on those used in high scoring herds would be expected. Failure to account for the genetic quality of herdmates creates a positive difference in favor of sires proven primarily in lower scoring herds and has a negative effect on PTAT of bulls used in high averaging herds. An alternative to the herdmate comparison method then in use was suggested.

Materials and Methods

Source of data

Pedigree information, estimates of transmitting abilities, and accuracies of estimates were obtained from the Holstein Friesian Association of America. The sources of the data were the six animal model sire evaluations from Winter, 1991 to Summer, 1993. Bulls with birth years from 1984 to 1988 and with no official PTAT before Summer, 1991 were used. This resulted in a data set of 4198 bulls.

Data edits

Table 2 contains a list of edits and the number of bulls removed from the data for each edit. Initially, 4198 bulls were available. In the first edit, 2180 bulls were eliminated due to having a proof with a daughter contribution greater than .5 before Summer 91 or later than W 93. Next, 92 additional bulls were eliminated for having an NAAB stud code not equal to 1, 7, 8, 9, 11, 14, 16, 21, 23, or 29 in Winter 91. This resulted in a data set of 1926 bulls for this study.

Calculation of PA and PI

Parent averages (PA) were calculated using the formula:

$$PA = .5 (PTAT_{SIRE}) + .5 (PTAT_{DAM})$$

Pedigree index was calculated using the formula:

$$PI = .5 (PTA_{SIRE}) + .25 (PTA_{MGS})$$

Table 2. Number of bulls removed with each edit from bulls with birth years from 1984 to 1986.

Number Removed	Edit
2180	Bulls with first proof before S 91 or without first proof by S 93.
83	NAAB Code > 30
8	NAAB Code = 18
1	NAAB code = 28

Definition of Initial Proof with daughter contribution

The number of daughters was not included in the data set received. However, a progeny contribution was included for each bull. This was the sum of progeny weights where progeny from known mates had a weight of .5 and those from unknown mates had a weight of .3. No indication was provided as to when an official proof was released on a bull (> 10 daughters). For this study a bull's initial proof was defined as the first summary with a progeny contribution greater than .5 (> 1 daughter). All other proofs were numbered sequentially.

Accuracy of predicting bull's genetic merit from different pedigree sources (Objective 1)

Four sources of pedigree information considered were: Parent Average, Pedigree Index, $PTAT_{SIRE}$ and $PTAT_{DAM}$. Pedigree estimates from the Winter 1991 through Summer 1993 sire summaries were used. Ideally we would like to use the most accurate first crop proof on a bull in evaluating the accuracy of pedigree prediction (selection). However, because of the need to use only bulls with animal model pedigree information prior to scored daughters being available, we had to focus on the early proofs of each bull because of the limited time since the first animal model type evaluation.

The regression model used to determine the accuracy of predicting the bull's genetic merit from single sources was:

$$y_i = a + B_1 (\text{Pedigree Source}_i) + e_i$$

where:

$$y_i = \text{PTAT or DTD of the } i^{\text{th}} \text{ bull,}$$

$$a = y \text{ intercept,}$$

B_1 = regression coefficient,
Pedigree Source = PA, PI, PTAT_{SIRE}, or PTAT_{DAM}
 e_i = error for the i^{th} bull.

A second model was also used, where separate regression coefficients were estimated for PTAT_{SIRE} and PTAT_{DAM} simultaneously, using the model:

$$y_i = a + B_1 (\text{PTAT}_{\text{SIRE}}) + B_2 (\text{PTAT}_{\text{DAM}}) + e_i$$

where:

y_i = PTAT or DTD of the i^{th} bull,
 a = y intercept,
 B_1, B_2 = regression coefficients,
 e_i = error for the i^{th} bull.

R^2 s from the regression model were used to determine the accuracy of the different pedigree sources. Four different data sets were used to approach objective one.

The data set for the first analysis included the first PTAT and DTD with a daughter contribution of greater than .5 on each bull and the 4 pedigree sources from the summary immediately prior (Summer 91 PTAT, Winter 91 Pedigree Estimate... Summer 93 PTAT, Winter 93 Pedigree Estimate). A single regression was run for the regression of PTAT and DTD on each pedigree source. Because DTD was available only for Summer 93, only the bulls that obtained a first proof in Summer 93 were included in the DTD analysis. By using the pedigree estimate from the summary immediately prior to the one in which the bull received his first proof, the most recent pedigree information not influenced by the bull's progeny contribution was obtained. PTATs were regressed on pedigree estimates and R^2 s were computed to determine how accurately the first proof on a young bull can be predicted and which pedigree source is the best predictor.

The second data set included the PTAT from a bulls second

proof with a daughter contribution greater than .5 on each bull and the pedigree sources from the summary immediately prior to the first PTAT with a daughter contribution greater than .5 (Winter 92 PTAT, Winter 91 Pedigree Estimates...Summer 93 PTAT, Summer 92 Pedigree Estimates). Bulls which had their first daughter contribution greater than .5 in Summer 93 were not included in this data set. Because DTD was available only for Summer 93, only the bulls that obtained a first proof in Winter 93 were included in the DTD analysis. Using the information from the second proof after the initial daughter information was added was felt to correspond more closely with a bull's initial official proof. PTATs were regressed on pedigree estimates and R^2 s were computed to determine how well the second proof on a young bull can be predicted and which pedigree source is the best predictor. The first official proof (> 10 daughters) on each bull would usually occur in the PTATs classified as first or second in this study.

The third data set included all initial PTATs with a daughter contribution greater than .5, and pedigree estimates from all previous summaries (Summer 91 PTAT, Winter 91 Pedigree Estimate...Summer 93 PTAT, Winter 91, Summer 91, Winter 92, Summer 92, and Winter 93 Pedigree Estimates) and was used to determine if the date of the first proof affects the relationship between first PTAT and the different sources of pedigree information. PTATs were regressed on pedigree estimates and R^2 s were computed to determine how well the initial PTAT of a young bull can be predicted from different pedigree sources at different times prior to his initial proof.

The fourth data set included DTDs for all bulls with initial proofs in the Summer 93 summary. DTD Summer 93 was regressed on all sources of pedigree information from previous

summaries (DTD Summer 93, Winter 91, Summer 91, Winter 92, Summer 92 and Winter 93 pedigree Estimates). DTDs were regressed on pedigree estimates and R^2 s were computed to determine how well the initial DTD of a young bull can be predicted from different pedigree sources at different times prior to his initial proof. DTD was used since it is a measure of the actual differences seen in the bull's daughters, and avoids the part whole relationship with pedigree information that exists for PTAT.

In the initial studies, PA was shown to be the pedigree source with the highest accuracy of prediction based on R^2 . Thus bulls were divided into deciles based on their PA prior to their initial proof. Averages (means), maximums, minimums, and standard deviations for the PTAT of bulls in each decile were calculated and graphed against the average pedigree estimate, in order to examine the trend for bulls within deciles.

Assessment of effect of within herd variance correction (Objective 2).

In order to examine the impact of the within herd variance correction which was implemented in the Summer 93 evaluation, correlations between the pedigree sources PA, PI, $PTAT_{DAM}$, and $PTAT_{SIRE}$ for Winter 93 and Summer 93 evaluations for all bulls in the data set were calculated, based on the date of the initial proof.

Next, all bulls in the data set were divided into groups based on the date of their initial proof. A predicted DTD for each bull was calculated, based on the PA from the proof immediately prior to the bull's initial proof and using the predict option of SAS^R. The actual DTD from the Winter 93

proof was then subtracted from the predicted DTD to obtain a variable called $(DTD_{\text{PREDICTED}} - DTD)$, the difference between predicted and actual DTD values. The variable $(DTD_{\text{PREDICTED}} - DTD)$ was then regressed on the variable $(PTAT_{\text{DAM S 93}} - PTAT_{\text{DAM W 93}})$ for all bulls grouped by the date of the initial proof. Subtracting $PTAT_{\text{DAM S 93}} - PTAT_{\text{DAM W 93}}$ allows us to look at the magnitude of the change due to the variance correction. The addition of granddaughter's to the dam's proof may also impact the change from W 93 to S 93, however. The model used for the regressions was:

$$y_i = a + B_1 (X_i) + e_i$$

where:

y_i = $(DTD_{\text{PREDICTED}} - DTD)$ of the i^{th} bull,

a = y intercept,

B_1 = regression coefficient,

X_i = $PTAT_{\text{DAM S 93}} - PTAT_{\text{DAM W 93}}$

e_i = error for the i^{th} bull.

Impact of granddaughter information on Dam's PTAT (Objective 3).

Information from all relatives impacts the genetic evaluation of an individual in an animal model system. The impact of the daughters of a young bull on his dam's evaluation were examined in this study. The possibility of preferential treatment of potential bull dam's was suspected. The preferential treatment has the potential to inflate her PTAT above her true transmitting ability for type. In this case the dam's PTAT would overestimate her son's PTAT. Lower than expected daughter performance would then tend to reduce the PA. The higher average reliability of sires of sons than

of dams of sons results in the fact that the dam's PTAT would be influenced more by the addition of granddaughter information than the sire's PTAT. Many granddaughters with records randomly distributed in different herds would adjust her PTAT toward her true transmitting ability. The degree of influence of granddaughters on a bull dam's genetic evaluation is a function of the amount of information in the son's proof, the current amount of information in the dam's PTAT and the amount that the granddaughter deviation exceeds or falls below the bull's PA.

Other possible sources of information adding to a bull dam's genetic evaluation are daughters, her own records, and parents. Daughter's are a source of information which can cause changes in the PTAT of a bull dam, since she may have daughter's getting records prior to or at the same time that her granddaughter's are scored. Because a cow's FS becomes permanent when she is five years old, it is very doubtful that bull dam's own records would be adding information to her genetic evaluation by the time her granddaughter's are being evaluated. The parent information is expected to have very little effect on changing the bull dam's evaluation by the time she has granddaughter information. Therefore, while the entire change in a bull dam's PTAT cannot be accounted for by the addition of granddaughters it is the most likely source of change in the PTAT of the bull dam.

For the first regression, the $PTAT_{DAM}$ for the W 93 and S 93 evaluations was calculated for all bulls which had an initial proof with daughter information in S 93. Only bulls whose dams did not have a son in a previous evaluation were included. Of the 150 bulls in the data set which received an initial proof with daughter information in S 93, 81 had dams which did not have a son in a previous evaluation. The

PTAT_{DAM} from W 93 contained no granddaughter information from the bulls in this data set, while the S 93 evaluation did contain granddaughter information. Subtracting PTAT_{DAM} W 93 from PTAT_{DAM} S 93 gave a variable defined as $\Delta PTAT_{DAM}$. The regression of $\Delta PTAT_{DAM}$ on the son's DTD from the S 93 evaluation was then computed. Bulls with DTD above their PA are expected to increase their dam's PTAT and consequently to increase the PA while bulls with DTD below their PA are expected to have the reverse effect. Second, the regression of $\Delta PTAT_{DAM}$ on the son's DTD S 93 - PA_{SON} W 93 was run. For the third regression $\Delta PTAT_{DAM}$ was regressed on the son's PTAT S 93 - PA_{SON} W 93.

The model used for these three regressions was:

$$y_i = a + B_1 (X) + e_i$$

where:

y_i = $\Delta PTAT_{DAM}$ of the i^{th} bull,

a = y intercept,

B_1 = regression coefficient,

X = DTD S 93 for the i^{th} bull;

or DTD S 93 - PA W 93 for the i^{th} bull;

or PTAT S 93 - PA W 93 for the i^{th} bull;

e_i = error for the i^{th} bull.

In regressions 2 and 3, PA was subtracted from DTD and PTAT_{SON} because it was the best estimate of PTAT before daughter records are available. By defining the variables as differences (DTD-PA_{SON} and PTAT-PA_{SON}) we were able to examine the impact of the change in PTAT_{DAM} on the difference between the bull's actual and estimated genetic merit.

Assessment of impact of testing population (Objective 4).

In the United States, most A.I. studs sample young bulls

in special herds designated sampling herds. Semen is distributed to these herds with the goal of having a random sample of daughters, with few daughters in any one herd and a total number of daughters between 40 and 60. Because of this, most bull's are sampled in a limited number of herds. After the bull's daughters mature and enter production, the bulls to be returned to the active lineup are selected. Different A.I. studs conduct progeny tests in different herds, and in different areas of the United States, where herd size and management may differ widely, thus the potential exists that a bias in favor of some testing populations may exist.

In this data set, bulls were divided into groups based on their NAAB code. Numerical stud codes were replaced by alphabetical ones for this data set.

Means for $PTAT_{SIRE}$ and $PTAT_{DAM}$ by NAAB code were calculated, and a simple t test was used to determine if there were significant differences between them by NAAB code. Next, a bull's PA from the evaluation before his initial proof was subtracted from the PTAT of his initial proof, and a t test was used to determine if there were statistical differences between NAAB codes. A similar procedure for bulls with an initial proof in S 93 was done using DTD from the S 93 evaluation and PTAT from the W 93 one.

Two regressions were run on each group. The first was the regression of PTAT from the bull's second proof on PA from the proof immediately prior to the initial proof, in order to determine if different testing populations had an impact on the amount of variation seen in the PTAT which was explained by PA. Use of PTAT from a proof one after the initial proof was felt to give a value more typical of a bull's initial official proof.

The second regression was similar, but used PI rather

than PA.

R^2 s and regression coefficients were used to interpret the effect of testing population on PTAT. Also, intercepts and regression coefficients for each stud were compared.

Third, the predict option in SAS^R was used to estimate PTAT for the first proof after a bull's initial proof for all bulls from each stud in the data set, based on PA and PI from the proof immediately prior to the bull's initial proof. This operation was performed in two different ways. First, bulls were grouped by stud and PTAT was estimated within NAAB code. Second, all bulls in the data set were grouped together and PTAT was estimated. Then bulls were sorted by NAAB code and the results were compared to the first estimates, in order to see if any differences by NAAB code were apparent.

Results and Discussion

Predicting bull's initial proof with daughter information for PTAT and DTD from different pedigree sources. (Objective 1)

Means and SDs for first proofs with daughter information and pedigree information immediately prior to first proof are in Table 3. PA had a mean of .84 and SD of .50 while PI had a mean of .43 and SD of .41. A large part of the difference was due to the fact that PA considered the source of 100% of a bull's genes, while PI only accounted for 75%. This also accounted for part of the larger SD of PA than PI. The mean $PTAT_{DAM}$ was higher than the mean $PTAT_{SIRE}$, which may indicate that dams were more highly selected than sires or an upward bias in $PTAT_{DAM}$. The additional selection pressure on dams could be the result of the phenotypic cut off applied to bull dam selection by many studs (Funk, 1993). The higher value for DTD than PTAT was due to including only bulls with an initial proof in Summer 93. Thus these bulls were genetically superior to the average of all bulls included in the study.

Regression coefficients and R^2 s for regressing PTAT from the bull's initial proof on prior pedigree estimate and DTD S 93 are in Table 4. PA was a better predictor (higher R^2) than PI for both PTAT and DTD. The higher R^2 for PTAT than DTD reflects the fact that pedigree information is used in computing PTAT, and the part whole relationship between them. The R^2 s for $PTAT_{SIRE}$ and $PTAT_{DAM}$ summed to nearly the values of PA as expected. Dam and sire coefficients were nearly identical for predicting PTAT, possibly reflecting the part whole relationship of PA and PTAT. However, the sire coefficient was considerably larger than the one for dam in predicting DTD. Also PA had only a slightly larger R^2 than PI

Table 3. Means and Standard Deviations for genetic values at time of first proof with progeny contribution greater than .5, and pedigree sources from immediately prior proof (N = 1926).

<u>Prior Pedigree</u>	<u>Mean</u>	<u>SD</u>
PA	.84	.50
PI	.43	.41
PTAT _{SIRE}	.55	.78
PTAT _{DAM}	1.14	.64
<u>First Proof</u>		
PTAT	.82	.64
DTD ^a	.96	.97
Reliability	.64	.10

^a DTD only available for bulls with first proof in S 93 (n = 150).

Table 4. Regression of PTAT (N = 1926) and DTD (N = 150) for bull's initial proof on previous pedigree sources.

Pedigree				
Source	<u>PTAT</u>		<u>DTD</u>	
	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>
PA	.93	.59	.94	.20
PI	.92	.42	1.07	.16
PTAT _{SIRE}	.47	.37	.50	.12
PTAT _{DAM}	.43	.21	.36	.07
PTAT _{SIRE} +	.47		.54	
PTAT _{DAM}	.45	.60	.40	.21

for predicting DTD. The R^2 s did indicate a reasonable accuracy was obtained in estimating a bull's PTAT and DTD from pedigree information. PA was a better estimator than the other sources investigated and would be the preferred method for selecting bulls to enter A.I. proving programs.

These results agree with the conclusions of Samuelson (1992) who found that PA was a better indicator of a bull's actual genetic merit for milk, fat, and protein than PI. Samuelson found R^2 s of .27, .33, and .27 for the regression of initial milk, fat, and protein PTA on PI respectively, with regression coefficients of .85, .84, and .86, while he found R^2 s of .47, .54, and .49 and regression coefficients of .84, .81, and .84 for the regression of on the initial PTA of the same three traits on PA. The higher R^2 found in this study may be a function of how initial proof was defined. By specifying a progeny contribution of greater than .5 as the value for initial proof, pedigree made a larger contribution to initial proof PTAs in this study than in Samuelson's.

Regression coefficients and R^2 s for regressing second PTAT on pedigree estimates from two evaluations prior are in Table 5. For the this regression using PTAT, R^2 s and regression coefficients are very similar to those in the previous model, indicating that the addition of more progeny did not greatly impact the predictive value of the pedigree sources.

Table 6 gives the regression coefficients and R^2 s for the regression of DTD from the S 93 evaluation on pedigree estimates from S 92 for bulls with an initial proof in W 93. This allowed us to look at the impact of pedigree information on DTD of a bulls second proof. R^2 s did not change except in the case of PI, which increased from 1.07 to 1.10. Changes were less than when comparing first and second PTAT. In this

Table 5. Regression of PTAT (N = 1706) for bull's second proof on pedigree information from proof immediately prior to addition of progeny information.

Pedigree

<u>Source</u>	<u>B₁</u>	<u>R²</u>
PA	.94	.58
PI	.94	.40
PTAT _{SIRE}	.47	.35
PTAT _{DAM}	.43	.21
PTAT _{SIRE} +	.47	
PTAT _{DAM}	.40	.58

Table 6. Regression of DTD S 93 (N = 444) for bulls with initial proof in W 93 on S 92 pedigree information.

Pedigree

<u>Source</u>	<u>B₁</u>	<u>R²</u>
PA	.94	.21
PI	1.10	.17
PTAT _{SIRE}	.51	.12
PTAT _{DAM}	.36	.07
PTAT _{SIRE} +	.54	
PTAT _{DAM}	.40	.21

case the comparison represents two totally different groups of bulls. Thus any changes are confounded with bull group.

The second proof which was used in these regressions may correspond more closely with most bull's first official proof. Samuelson (1992) performed similar regressions for yield traits. His average R^2 s for PA and PI of .5 and .3 were intermediate for the two PA estimates in this study and near the estimate of PI for the second proof. R^2 for $PTAT_{DAM}$ was approximately half those found by Samuelson (1992). This comparison indicated that dam information was less useful in predicting son's type proof than his production proof. This may reflect the higher selection intensity of dams for type than for production.

Table 7 contains the means and SDs for PA and PI for all evaluations. From W91 to W 93 there was a general trend for $PTAT_{SIRE}$ and $PTAT_{DAM}$ to decrease slowly for each group of bulls. However, between W 93 and S 93 sire PTATs continued to drop while dam PTATs increased slowly (.04 to .07). This corresponds with the change in the evaluation procedure to account for heterogeneous variance. It is logical that a bull's daughters may be in slightly lower than average (higher variance) herds while bull dams may be in higher scoring (lower variance) herds. Thus the small changes in opposite directions seem reasonable. The fact that changes after daughter information became available were only slightly more negative than those before and that they impact on both sires and dams suggested the dam PTATs are not in general markedly biased upward.

Table 8 gives the means and SDs for the progeny contribution for all bulls in the data set grouped by date of first proof with daughter information. A measurement of .3 prior to the initial proof indicates that a bull had one

Table 7. Means and SDs for genetic parameters of Final Score of bulls grouped by date of initial proof.

Date of	Date of Initial Proof									
	<u>S 91</u> n=443		<u>W 92</u> n=446		<u>S 92</u> n=443		<u>W 93</u> n=444		<u>S 93</u> n=150	
<u>Parameter</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
W 91										
PA	.79	.50	.86	.52	.92	.47	.92	.44	.99	.43
PI	.39	.45	.46	.43	.48	.40	.47	.35	.54	.35
PTAT _{SIRE}	.46	.84	.58	.81	.66	.74	.59	.63	.74	.66
PTAT _{DAM}	1.11	.61	1.12	.61	1.17	.61	1.26	.62	1.26	.62
S 91		a								
PA	.79	.49	.85	.52	.92	.47	.92	.43	.99	.42
PI	.39	.45	.46	.44	.47	.40	.46	.34	.54	.33
PTAT _{SIRE}	.45	.84	.58	.82	.66	.74	.57	.62	.73	.62
PTAT _{DAM}	1.11	.62	1.12	.61	1.18	.62	1.26	.62	1.25	.63
W 92				a						
PA	.75	.50	.83	.52	.91	.48	.90	.45	.98	.44
PI	.37	.45	.45	.44	.46	.41	.45	.35	.53	.34
PTAT _{SIRE}	.43	.84	.56	.84	.64	.76	.56	.65	.74	.64
PTAT _{DAM}	1.07	.64	1.08	.61	1.17	.63	1.25	.64	1.24	.65
S 92						a				
PA	.70	.50	.77	.52	.88	.48	.86	.46	.94	.44
PI	.34	.45	.42	.44	.44	.41	.42	.36	.49	.34
PTAT _{SIRE}	.38	.84	.52	.82	.64	.74	.52	.64	.66	.64
PTAT _{DAM}	1.01	.64	1.02	.62	1.14	.64	1.21	.65	1.21	.67
W 93								a		
PA	.69	.50	.76	.53	.86	.49	.85	.47	.93	.45
PI	.34	.45	.41	.45	.43	.41	.42	.36	.48	.34
PTAT _{SIRE}	.37	.84	.51	.84	.60	.75	.51	.65	.65	.64
PTAT _{DAM}	.99	.65	1.01	.63	1.11	.65	1.20	.67	1.22	.68
S 93										a
PA	.69	.53	.76	.56	.83	.51	.82	.49	.92	.47
PI	.30	.47	.37	.46	.39	.43	.36	.38	.43	.36
PTAT _{SIRE}	.31	.87	.44	.87	.52	.79	.40	.68	.55	.66
PTAT _{DAM}	1.06	.70	1.06	.67	1.15	.69	1.24	.69	1.27	.71

a Any proof below the a in a column may be affected by the performance of the bull's daughters.

classified daughter with an unidentified mate. Of the 1926 bulls in the data set for this study, 223 had one scored daughter before their initial proof.

Progeny contribution for the first proof with progeny contribution greater than .3 ranged from 2.85 to 6.8. For the subsequent proof, progeny contribution ranged from 8.65 to 10.6. For many bulls, this would correspond to their first official Holstein Association type proof. The general tendency was for bull's progeny contribution to increase dramatically between the first proof with progeny contribution greater than .5 and the immediately subsequent proof, and then to increase more slowly.

Sires with more than 20 sons in the data set are listed in Table 9. Sires are listed in descending order by the number of sons in the data set. These sires account for 82% of the bulls in the data set (1686 out of 2053 total bulls). Bulls with greater than 50 sons in the data set accounted for 72% of the total (1479 out of 2053 total). Samuelson (1992) had found a similar concentration of sires for a slightly earlier period. He suggested that the use of the limited number of sire's of sons reduced the predictive value of $PTAT_{SIRE}$ for production traits.

Table 10 gives the means and SDs for $PTAT_{MGS}$ by date of initial proof. Values tended to be higher than for $PTAT_{SIRE}$ particularly for the most recent bulls. These means further substantiate the higher selection standards for type for dams than for sires.

Tables 11 through 14 contain the regression coefficients and R^2 s for regressing PTAT from the bull's initial proof on PA, PI, $PTAT_{SIRE}$, and $PTAT_{DAM}$ from all previous evaluations. As the date of the initial proof changed from S 91 to S 93, the number of previous proofs with pedigree information increased

Table 8. Means and SDs for Progeny Contribution (PC) for PTAT of bulls grouped by date of first proof with greater than 1 daughter¹.

Date of initial proof.										
<u>Proof</u>	S 91		W 92		S 92		W 93		S 93	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
W 91	.3									
S 91	3.5	2.9	.3							
W 92	8.6	4.9	2.8	2.2	.3					
S 92	11.5	5.9	9.1	5.2	3.5	3.0	.3			
W 93	12.6	6.5	12.9	8.0	10.6	5.2	4.0	3.4	.3	
S 93	13.0	7.0	14.0	9.9	13.5	6.4	11.9	.4	6.8	.3

1 Units are in daughter contributions, .5 for a daughter with an identified mate, .3 for a daughter with an unidentified mate.

Table 9. Sires with greater than 20 sons in the data set.

Reg #	S <u>91</u>	W <u>92</u>	S <u>92</u>	W <u>93</u>	S <u>93</u>	<u>Total</u>	<u>Name</u>
1697572	109	98	83	29	8	328	Rotate
1806201	126	57	32	18	2	235	Ned Boy
343514	6	17	53	104	22	204	Enhancer
1773417	69	36	47	21	7	181	Mark
1821208	10	56	33	19	6	124	Royalty
1842389	0	2	27	69	16	114	Jesse
1879085	0	0	6	56	45	110	Cleitus
1856904	2	2	11	35	17	69	Secret
1829401	0	5	40	13	2	60	Steady
1667366	20	16	16	1	1	54	Bell
1882797	0	0	0	15	16	33	Bell Troy
1819119	8	12	5	1	0	26	Big Al
1829881	0	0	10	8	4	23	Klark
1721509	10	5	3	3	1	22	Jetson
1839194	1	3	11	6	1	22	Bear
1840363	0	5	11	4	1	21	Bubba
352790	4	7	8	0	1	20	Starbuck
1791371	8	9	3	0	0	20	Kirk Boy
1836126	<u>0</u>	<u>0</u>	<u>18</u>	<u>2</u>	<u>0</u>	<u>20</u>	Count
Total	371	330	417	404	138	1686	
Other							
Bulls	<u>72</u>	<u>116</u>	<u>26</u>	<u>40</u>	<u>12</u>	<u>367</u>	
Total	443	446	443	444	150	1926	

Table 10. Means and SDs for PTAT_{MGS} from proof prior to initial proof by date of initial proof.

Date of Initial <u>Proof</u>	<u>n</u>	<u>Mean</u>	<u>SD</u>
S 91	48	.65	.65
W 92	53	.68	.61
S 92	52	.57	.66
W 93	57	.65	.69
S 93	34	.64	.71

from 1 to 5.

Table 11 contains the results for the regression of PTAT on PA. R^2 s ranged from .35 to .69, regression coefficients ranged from .88 to 1.02. Table 12 contains the regression coefficients and R^2 s for regressing PTAT from the bull's initial proof with daughter information on PI from all previous evaluations. R^2 s ranged from .25 to .48, regression coefficients ranged from .83 to 1.08. Table 13 contains the regression coefficients and R^2 s for regressing PTAT from the bull's initial proof on $PTAT_{SIRE}$ from all previous evaluations. R^2 s ranged from .19 to .45, regression coefficients ranged from .42 to .51. Table 14 contains the regression coefficients and R^2 s for regressing PTAT from the bull's initial proof on $PTAT_{DAM}$ from all previous evaluations. R^2 s ranged from .13 to .24, regression coefficients ranged from .38 to .49.

Values on the diagonals represent the pedigree information available the run prior to the bull receiving his initial proof with more than one daughter. The general downward trend in the R^2 s for these pedigree values from S 91 to W 93 coincides with a decrease in the SD for $PTAT_{SIRE}$ (Table 7.) However, it is not clear that it is the only or even the primary cause of this decrease. Each row represents a group of bulls receiving an initial proof at the same time. Moving right from W 91 pedigree information represents an increase in accuracy of the pedigree information and is accompanied by a slight increase in R^2 primarily from the addition information on the dam.

For the model using PI, R^2 s were smaller than for the model using PA, which may be explained by the fact that PI accounts for only 75% of a bulls genes, while PA accounts for 100%.

Regression coefficients for the model using PA ranged from .88 to 1.02, those for PI from .83 to 1.08. These values closely approached the expected value of 1 (Van Vleck, 1982). Regression coefficients for the models using $PTAT_{SIRE}$ and $PTAT_{DAM}$ ranged from .42 to .51, and .38 to .49, respectively. Values for the model with sire approached the expected value of .5 and those for dam were slightly smaller (Van Vleck, 1982).

The decrease in the R^2 s with each succeeding group of bulls starting with those that received their first proof in W 92 may be due to the reduced predictive value of sire PTATs. While a similar trend was seen by Samuelson (1992), it is even more consistent in this study.

Both R^2 s and regression coefficients for the model using $PTAT_{DAM}$ were lower than for the model using $PTAT_{SIRE}$, which is a reflection of the lower average reliability of dams than sires and the reduced SD's. While the general trends found in this study are similar to those found by Samuelson (1992), R^2 s and regression coefficients were consistently higher, probably reflecting the higher part whole relationship in this data set and the fact that dams were less useful in predicting son's proof than for production traits.

Table 11. Regression of PTAT of bull's initial proof on PA from previous evaluations.

Date of		<u>Date of PA</u>									
Initial	<u>W 91</u>	<u>S 91</u>		<u>W 92</u>		<u>S 92</u>		<u>W 93</u>		<u>S E</u>	
<u>Proof</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	
S 91	.60	.88									.03
W 92	.68	.97	.69	.98							.03
S 92	.52	.91	.55	.92	.57	.92					.04
W 93	.47	.93	.51	.99	.53	.95	.54	.96			.04
S 93	.35	.91	.42	1.02	.42	.98	.42	.98	.44	.98	.09

Table 12. Regression of PTAT of bull's initial proof on PI from previous evaluations.

Date of		<u>Date of PI</u>									
Initial	<u>W 91</u>	<u>S 91</u>		<u>W 92</u>		<u>S 92</u>		<u>W 93</u>			
<u>Proof</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>SE</u>
S 91	.41	.83									.05
W 92	.48	.99	.48	.98							.05
S 92	.42	.95	.43	.95	.40	.94					.05
W 93	.31	.96	.33	1.00	.34	.98	.34	.97			.07
S 93	.25	.93	.30	1.08	.29	1.04	.30	1.05	.30	1.06	.13

Table 13. Regression of $PTAT_{SIRE}$ of initial proof on PI from previous evaluations.

<u>Date of $PTAT_{SIRE}$</u>											
Date of											
Initial	<u>W 91</u>		<u>S 91</u>		<u>W 92</u>		<u>S 92</u>		<u>W 93</u>		
<u>Proof</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>SE</u>
S 91	.35	.42									.03
W 92	.45	.51	.45	.51							.03
S 92	.34	.47	.32	.47	.35	.46					.03
W 93	.26	.48	.28	.51	.28	.49	.28	.49			.04
S 93	.19	.43	.23	.51	.23	.49	.23	.49	.23	.50	.08

Table 14. Regression of $PTAT_{DAM}$ of bull's initial proof on PI from previous evaluations.

<u>Date of $PTAT_{DAM}$</u>											
Date of											
Initial	<u>W 91</u>		<u>S 91</u>		<u>W 92</u>		<u>S 92</u>		<u>W 93</u>		
<u>Proof</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>SE</u>
S 91	.18	.39									.04
W 92	.24	.49	.23	.48							.04
S 92	.15	.38	.17	.39	.18	.40					.04
W 93	.19	.43	.22	.45	.23	.45	.24	.45			.04
S 93	.13	.39	.15	.42	.16	.41	.16	.39	.17	.41	.07

Tables 15 through 18 contain the regression coefficients and R^2 s for regressing DTD from the bull's S 93 proof on PA, PI, $PTAT_{SIRE}$, and $PTAT_{DAM}$ from all previous evaluations.

Tables 15 contains the regression coefficients and R^2 s for regressing DTD from the bull's S 93 proof on PA. R^2 s ranged from .16 to .26, and regression coefficients ranged from .70 to .94.

Table 16 contains the regression coefficients and R^2 s for regressing DTD from the bull's S 93 proof on PI from all previous evaluations. R^2 s ranged from .12 to .22, and regression coefficients ranged from .69 to 1.1.

Table 17 contains the regression coefficients and R^2 s for regressing DTD from the bull's S 93 proof on $PTAT_{SIRE}$ from all previous evaluations. R^2 s ranged from .10 to .19, and regression coefficients ranged from .35 to .55.

Table 18 contains the regression coefficients and R^2 s for regressing DTD from the bull's S 93 proof on $PTAT_{DAM}$ from all previous evaluations. R^2 s ranged from .04 to .08, and regression coefficients ranged from .28 to .40.

Two opposite effects impacted the R^2 s and regression coefficients in all of the above regressions. Progressing down the table for date of initial proof from S 91 to S 93, less daughter information was available in this data set for each subsequent group of bulls, and R^2 s tended to decrease, while regression coefficients tended to increase slightly. However, moving horizontally across the table within date of first proof, the distance between the time of the pedigree information and the time of daughter's evaluation decreases, and R^2 s tended to increase slightly. Both R^2 s and regression coefficients for DTD were lower than those for PTAT. The lower R^2 s for the regression of DTD than PTAT reflected the part whole relationship between PTAT and PA which is not

present for DTD and PA.

Samuelson (1992) performed a similar study examining PTAs and DYDs for milk, fat and protein. However, in his study, bulls were identified by first official proof, which in many cases would correspond with the proof immediately after that which was defined as the initial proof in this study. Because of this, the R^2 values he found for regressions using PTAs were lower than those found in this study. The higher values in this study are a reflection of the fact that the initial proof as defined here still has a substantial pedigree contribution. Regression coefficients for analyses using PTAT were slightly higher than those found by Samuelson (1992). An exception to these trends was the model regressing PTAT on $PTAT_{SIRE}$ which had similar values for R^2 s and regression coefficients to those Samuelson found. This may reflect the high reliability of sire's of bulls in both studies.

In Samuelson's study, R^2 s and regression coefficients for models using DYD on PA were highest for DYD fat, next for milk and lowest for protein. Values for DYD fat approached those found in this study for the regression of DTD S 93 on PA. For the models DTD on PI and DTD on $PTAT_{SIRE}$, values for R^2 and regression coefficients were higher than those found by Samuelson (1992). However, dam information was less useful in predicting DTD than had been the case for production traits. This may be partially due to reduced variation in $PTAT_{DAM}$ compared with $PTAT_{SIRE}$. Possibly this is the result of the high phenotypic cut off used in selecting dam's of sires to be tested.

Table 15. Regression of DTD S 93 on previous PA.

Date of PA

Date of

Initial	<u>W 91</u>		<u>S 91</u>		<u>W 92</u>		<u>S 92</u>		<u>W 93</u>		
<u>Proof</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>SE</u>
S 91	.16	.70									.08
W 92	.25	.83	.26	.85							.07
S 92	.22	.84	.26	.84	.24	.85					.08
W 93	.17	.86	.17	.88	.18	.85	.18	.84			.09
S 93	.17	.89	.20	1.00	.20	.94	.19	.93	.21	.94	.16

Table 16. Regression of DTD S 93 on PI from all previous evaluations.

Date of PI

Date of

Initial	<u>W 91</u>		<u>S 91</u>		<u>W 92</u>		<u>S 92</u>		<u>W 93</u>		
<u>Proof</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>SE</u>
S 91	.12	.69									.08
W 92	.18	.86	.19	.86							.08
S 92	.21	.96	.22	.97	.22	.96					.08
W 93	.14	1.00	.15	1.03	.15	1.01	.15	.99			.11
S 93	.13	.95	.16	1.10	.15	1.05	.15	1.05	.16	1.07	.20

Table 17. Regression of DTD S 93 on PTAT_{SIRE} from all previous evaluations.

<u>Date of PTAT_{SIRE}</u>											
Date of											
Initial	<u>W 91</u>		<u>S 91</u>		<u>W 92</u>		<u>S 92</u>		<u>W 93</u>		
<u>Proof</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>SE</u>
S 91	.11	.35									.05
W 92	.18	.45	.18	.45							.04
S 92	.18	.48	.18	.48	.19	.48					.05
W 93	.13	.53	.14	.55	.14	.52	.14	.52			.06
S 93	.10	.44	.12	.51	.11	.48	.11	.49	.12	.50	.11

Table 18. Regression of DTD S 93 on PTAT_{DAM} from all previous evaluations.

<u>Date of PTAT_{DAM}</u>											
Date of											
Initial	<u>W 91</u>		<u>S 91</u>		<u>W 92</u>		<u>S 92</u>		<u>W 93</u>		
<u>Proof</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>R²</u>	<u>B₁</u>	<u>SE</u>
S 91	.04	.29									.07
W 92	.08	.40	.08	.39							.06
S 92	.04	.28	.04	.28	.05	.29					.06
W 93	.04	.31	.04	.30	.05	.31	.05	.31			.06
S 93	.05	.36	.06	.38	.06	.37	.06	.35	.07	.36	.11

Examination of PTAT by PA Decile.

Because PA showed the greatest accuracy (highest R^2) of the various pedigree sources examined, it was used as the basis of the final step in examining the accuracy of various pedigree estimates. The results of dividing the bulls into deciles based on their PAs is shown in Table 19, where the means for PA, initial PTAT, final PTAT, and DTD for bulls in each decile are shown. PA's were slightly larger than PTAT in every decile, indicating that PA was slightly inflated. The trend was for first and final PTAT to increase as decile did. However, while the rank of first and final PTAT were the same, values for final PTAT were lower than those for first PTAT. For DTD the results were less clear. DTD increased from decile 1 to 2, but results for deciles 2 through 7 showed little trend. There was also slippage at decile ten for DTD.

In Figure 1 the mean PTATs by decile are graphed. Deciles are graphed on the x axis. The y axis is the range of PTATs for the bulls in the decile. The range is shown by the vertical line, the mean by the X. It should be noted that the relationship between the deciles is linear, deciles 1 and 10 are exceptions to the trend. Bulls in decile 1 have PAs lower than expected, while bulls in decile 10 have PAs higher than expected.

In Figure 2 the mean PA and PTAT by decile are graphed. Mean PTAT corresponds very closely with PA, although there is a slight amount of slippage for deciles 9 and 10. For bulls in this data set, it appears that PA is a good predictor of PTAT ($R^2 = .59$), and that bulls with a high PA are more likely to have a high PTAT. However, this is based on PTATs with a significant amount of PA included.

Table 19. Means for PA immediately prior to progeny information, PTAT for initial proof, PTAT for S 93, and DTD S 93.

<u>Decile</u>	<u>PA</u>	1st <u>PTAT</u>	Final <u>PTAT</u>	<u>DTD</u>
1	-.01	-.03	-.12	.02
2	.39	.38	.18	.89
3	.57	.48	.38	.92
4	.70	.65	.57	.86
5	.82	.69	.60	.57
6	.93	.88	.77	1.16
7	1.06	.93	.85	.76
8	1.21	1.11	1.05	1.57
9	1.43	1.30	1.22	1.68
10	1.86	1.62	1.49	1.40

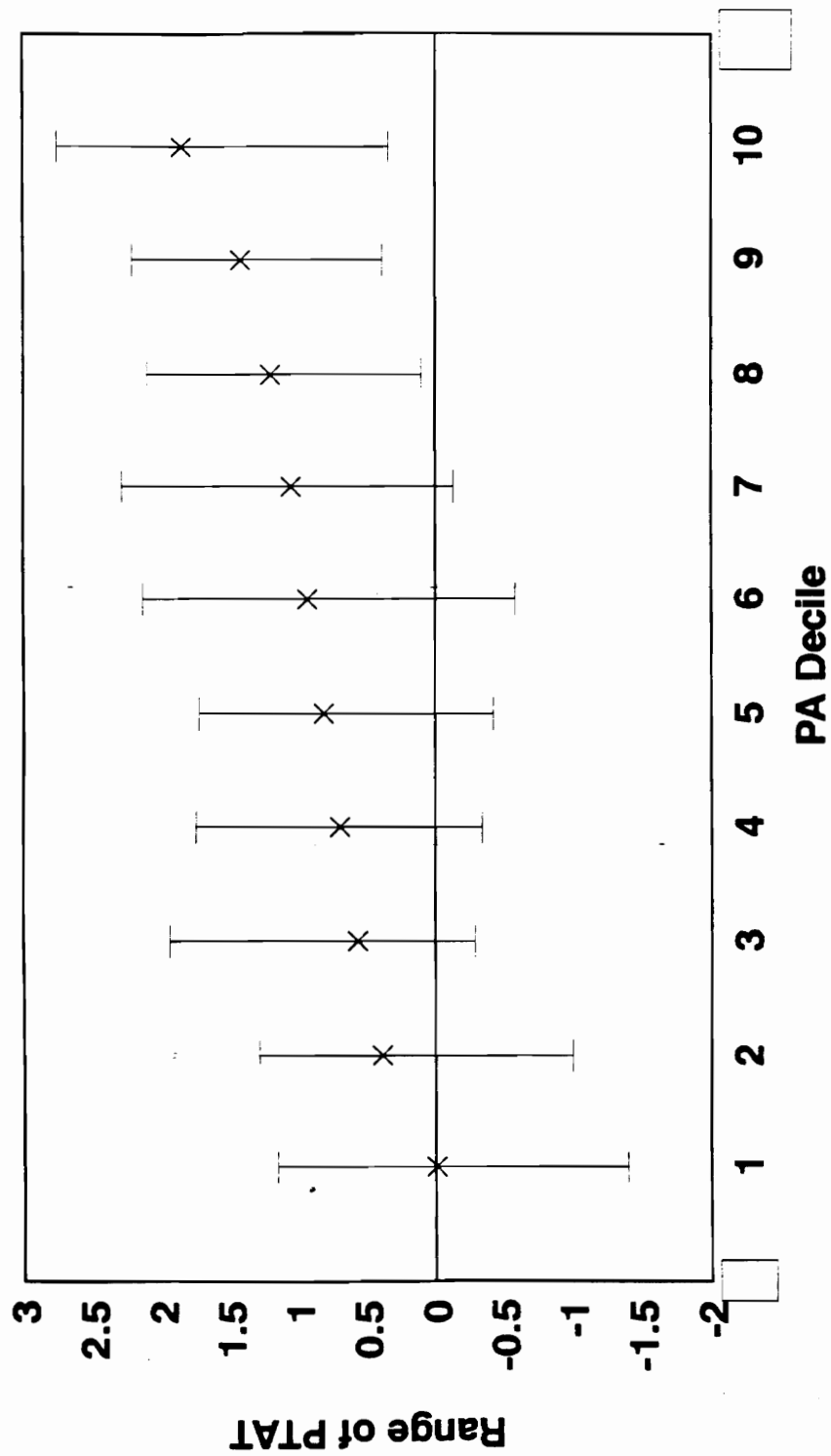


Figure 1. Mean (x) and Range (I) in PTAT for Bulls grouped by Decile of PA.

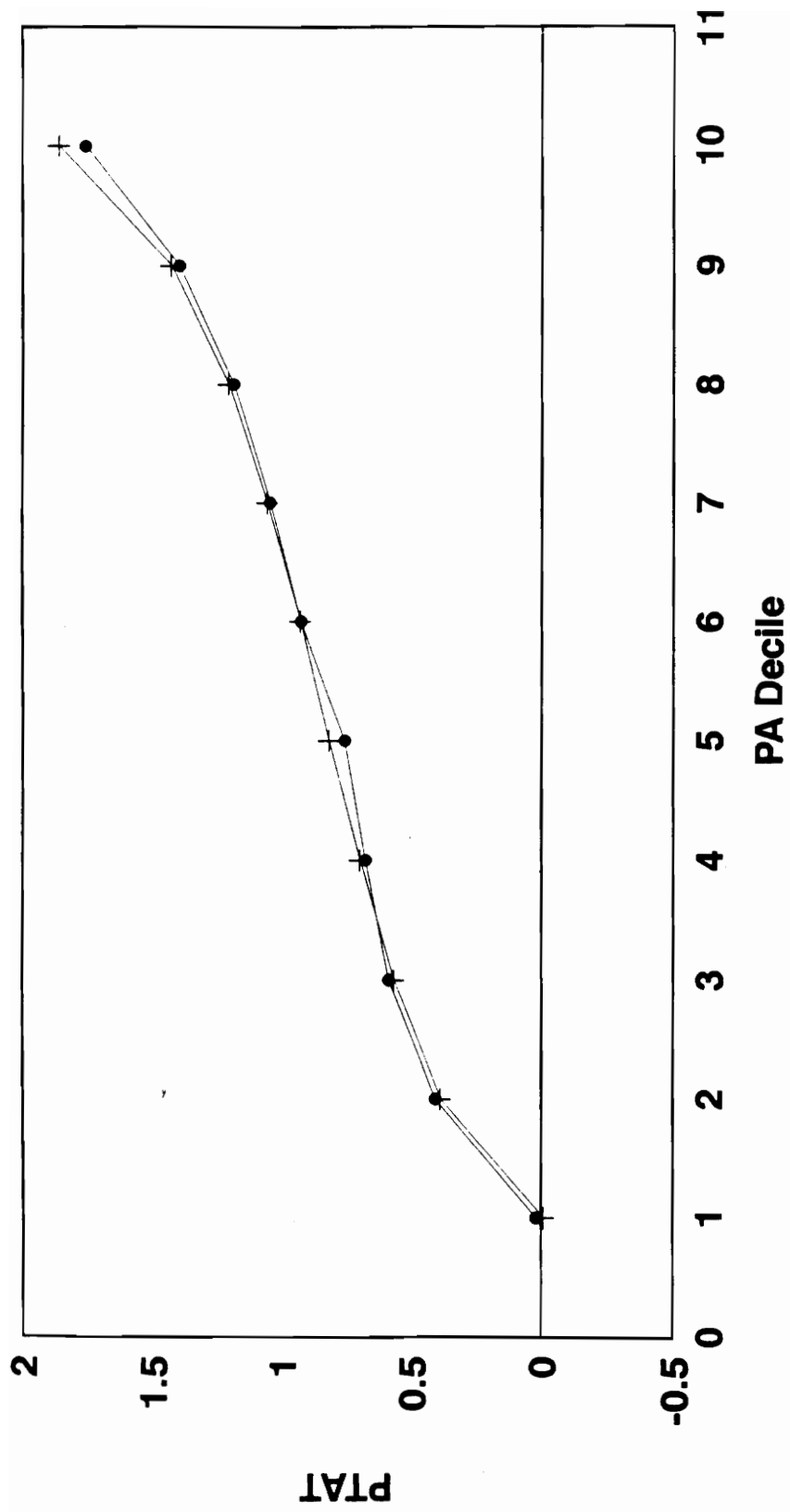


Figure 2. Average PA (+) and PTAT (.) for Bulls grouped by Decile of PA.

Impact of the Variance correction. (Objective 2)

Table 20 gives the correlations between PA, PI, PTAT_{SIRE} and PTAT_{DAM} for the W 93 and S 93 proofs by date of initial proof and overall. The overall correlations between W 93 and S 93 PA, PI, PTAT_{SIRE} and PTAT_{DAM} were .98, .99, .99, and .96 respectively. The high correlations between PA, PI, and PTAT_{SIRE} are probably a reflection of the fact that sires and maternal grand sires are already well proven and with high reliability when a bull receives his initial proof. Most bulls proven in A.I. are five years old when they receive their first proof. If the average age of young bull's sires was 6 to 8 years when the young sires were born, the sire would be 11 to 13 years old when the son receives his initial proof. The average REL of sires at this age is near 99%. Consequently, they would be expected to have daughters in herds with a wide range of within herd variances, and the within herd variance correction would have little effect on them. Because of this, the average impact of the variance adjustment on the sire and maternal grandsire will be minimal because they have been used in many types of herds.

Dams are generally 3 to 5 years old when their A.I. sampled sons are born. When a son receives his initial proof, his dam would be 8 to 10 years old. Because she has received her permanent score by this time, only her progeny can contribute to changes in her PTAT. The correlation of .96 indicates that neither the progeny contribution or the variance correction contributed greatly to changes in the PTAT_{DAM} during this 6 month period. However, it should be remembered that there are very few of her granddaughters in the son's proof at this time, and the addition of more granddaughters in later proofs might have a much larger

Table 20. Correlations between W 93 and S 93 proofs for PA, PI, PTAT_{SIRE}, and PTAT_{DAM} for date of first proof group, and for all bulls.

Pedigree	<u>Date of Initial Proof</u>					
<u>Source</u>	<u>S 91</u>	<u>W 92</u>	<u>S 92</u>	<u>W 93</u>	<u>S 93</u>	<u>All</u>
PA	.99	.99	.98	.97	.96	.98
PI	.99	.99	.99	.99	.99	.99
PTAT _{SIRE}	.99	.99	.99	.99	.99	.99
PTAT _{DAM}	.98	.98	.97	.95	.95	.96

impact.

R^2 s and regression coefficients for the regression of $(DTD_{\text{PREDICTED}} - DTD)$ on $[PTAT_{\text{DAM}} \text{ S } 93 \text{ (w Variance Adjustment - } PTAT_{\text{DAM}} \text{ W } 93 \text{ w/o Variance Adjustment)}]$ are given in Table 21 by date of the bulls initial proof. R^2 were very small but tended to increase as the date of the initial proof approached the date of the heterogeneous variance correction. However, the maximum R^2 was only .06, indicating that only accounted for 6% of the difference seen between $DTD_{\text{PREDICTED}}$ and DTD . Table 22 gives the means and SDs for $(PTAT_{\text{DAM}} \text{ S } 93 - PTAT_{\text{DAM}} \text{ W } 93)$ by date of initial proof. For bulls with initial proofs in the early evaluations in this data set, $PTAT_{\text{DAM}} \text{ S } 93 - PTAT_{\text{DAM}} \text{ W } 93$ was near zero. The trend for the regression coefficients to increase as the date of the initial proof approached that of the heterogeneous variance correction may reflect the larger change in this variable. However, it is more likely that as the date of the adjustment approached, the difference between the two dam PTATs is more likely due to adding granddaughters than to the variance correction. Thus, the later regressions were likely the result of the feed back mechanisms rather than the variance adjustment. If this is the case the variance adjustment had little impact on the accuracy of prediction.

Table 21. R^2 s, regression coefficients, and SEs for the regression of $(DTD_{ESTIMATED} - DTD)$ on $(PTAT_{DAM} S 93 - PTAT_{DAM} W 93)$ by date of proof.

Date of Initial <u>Proof</u>	<u>R²</u>	<u>B</u>	<u>SE</u>
S 91	.000	-.01	.16
W 92	.005	-.22	.15
S 92	.004	.16	.12
W 93	.06	.54	.09
S 93	.06	.43	.13

Table 22. Means and SDs for $(PTAT_{DAM} S 93 - PTAT_{DAM} W 93)$ by date of initial proof.

Date of Initial <u>Proof</u>	<u>Mean</u>	<u>SD</u>
S 91	.001	.16
W 92	-.014	.14
S 92	-.008	.16
W 93	-.010	.17
S 93	.056	.24

Impact of Granddaughter information on Dam's PTAT (Objective 3)

For this regression, only the dam's first son in the data set were included. The most frequent number of sons per dam was in this data set was 1, with 886 dams having one son. Frequencies for other numbers of sons were 223 dams with 2 sons, 81 dams with 3 sons, 43 dams with 4 sons, 21 dams with 5 sons, 5 dams with 6 sons, 5 dams with 7 sons, and one dam with 9 sons in the data set. Keeping only the first son left a total of 1040 bulls in the data set. It is unlikely that earlier granddaughter information from bulls not in this data set was impacting the dam's in this study.

The number of dams by date of evaluation ranged from 248 to 81. The variable $PTAT_{DAM} S 93 - PTAT_{DAM} W 93$ was given the name $\Delta PTAT_{DAM}$. Values for $\Delta PTAT_{DAM}$ by the date of the addition of granddaughter information are given in Table 23. Values for $\Delta PTAT_{DAM}$ ranged from $-.02$ to $.05$, with SDs from $.15$ to $.25$. The overall mean change in $PTAT_{DAM}$ with the addition of granddaughter information was $.016$, with a SD of $.174$. The maximum change was $.70$, the minimum change $-.72$. The mean change in a positive direction indicates that dam's PTATs were very slightly under estimated by pedigree information.

The R^2 s, regression coefficients, and SEs for the regression of $\Delta PTAT_{DAM}$ on the son's $DTD S 93$, $DTD S 93 - PA_{SON} W 93$, and $PTAT S 93 - PA_{SON} W 93$ are given in Table 24. R^2 s for the models using $DTD S 93 - PA_{SON} W 93$ and $PTAT S 93 - PA_{SON} W 93$ were $.54$ and $.56$ respectively. PA was removed from DTD and PTAT because PA is the estimate of transmitting ability before the animal has own records or progeny. These variables were estimates of the average Mendelian sampling of the dam's sons plus the average bias in the PA.

Samuelson (1992) found R^2 s from $.20$ to $.56$ when using

Table 23. Means and SDs for $\Delta PTAT_{DAM}$ by date of addition of granddaughter information.

<u>Date</u>	<u>Mean$\Delta PTAT_{DAM}$</u>	<u>SD</u>	<u>n</u>
S 91	.005	.16	235
W 92	-.022	.15	245
S 92	.0004	.16	248
W 93	-.004	.17	233
S 93	.05	.25	81

Table 24. R^2 s, regression coefficients, and SEs for the regression of $\Delta PTAT_{DAM}$ on the son's DTD S 93, DTD S 93 - PA_{SON} W 93, and PTAT S 93 - PA_{SON} W 93.

<u>Variable</u>	<u>R^2</u>	<u>B_1</u>	<u>SE</u>
DTD S 93	.39	.15	.02
DTD S 93 - PA_{SON} W 93	.54	.34	.03
PTAT S 93 - PA_{SON} W 93	.56	.31	.02

first DYD to predict $\Delta PTAT_{DAM}$, while the range for R^2 s to predict first DTD - PA_{SON} was .29 to .68. $PTAT_{INITIAL} - PA_{PREVIOUS}$ were highest, .33 to .72. He also found that for yield traits, $PTAT_{DAM}$ was overestimated.

One possible reason for the accuracy of $PTAT_{DAM}$ is the fact that it is not frequently used as a tool for selecting bull dams. Funk (1993) documents the use of the cow's FS, rather than her PTAT, as the criteria for selection, and suggests that PTAT would be a better choice. PTAs for yield traits are widely used in bull dam selection, and hence have incentive for preferential treatment and inflation.

Assessment of impact of testing population (Objective 4)

The ten NAAB codes in this data set were coded A through J. The number of bulls per NAAB code ranged from 473 to 14. The number of bulls per stud (by NAAB code), Means and SDs for PA and PI by NAAB code are given in Table 25. PA and PI differed significantly by NAAB code ($p < .01$).

Means and SDs for $PTAT$, $PTAT_{SIRE}$, $PTAT_{DAM}$, $PTA-PA$, and $DTD-PA$ by NAAB codes are given in Table 26. Means for $PTAT-PA$ and $PTAT-DTD$ were not significantly different ($p > .05$). Ranges for $PTAT-PA$ were from .06 to -.14. Means for $DTD-PA$ ranged from .33 to -.33. The expected value for each was zero. Positive results indicate a bull's daughters scored better than anticipated, while negative results indicate daughter's scoring lower than anticipated. The nearness of these numbers to zero would tend to indicate that on the average the PA was estimated accurately. However, 7 of the 10 NAAB codes, as well as the mean for all studs, had a negative value for $PTAT-PA$, indicative of slight inflation in PA for bulls. Five stud codes had negative values for $DTD-PA$, but the overall

mean was .04, indicating that DTD was slightly above expectation.

Regression coefficients and R^2 s for the regression of PTAT from the bull's second proof on PA and PI from the proof immediately prior to the initial proof for different NAAB codes are given in Table 27.

As was seen in earlier analysis, R^2 s for PA were higher than those for PI, due to the fact that PA explains all of a bulls genes, while PI only accounts for 75%. For PA, R^2 ranged from .38 to .69. Regression coefficients ranged from .80 to 1.04. There was no obvious relationship between the PA, PI, or number of bulls sampled and the R^2 and regression coefficients calculated. However stud I, which had the lowest R^2 also had the largest negative value for DTD - PA.

Means and SDs for PTAT at the time of the bull's second proof estimated from PA and PI of the proof immediately prior to the his initial proof by NAAB code are given in Table 28. When PI was used as the estimator, 4 studs had higher average predictions from the across stud analysis, while 5 had higher predictions from the within NAAB code analysis. One stud, B, had the same value from each analysis. However, the mean difference between the values for all the studs was .034. When PA was used as the estimator, 6 studs had higher average predictions from the across stud analysis, while 3 had higher predictions from the within NAAB code analysis. J had the same value from each analysis. The mean difference between the values for all the studs was -.025.

In Figure 3 the mean PA by NAAB code are graphed. NAAB codes are graphed on the x axis. The y axis is the range of PAs for the bulls in the decile. The range is shown by the vertical line, the mean by the X.

In Figure 4 the mean PI by NAAB code are graphed. NAAB

codes are graphed on the x axis. The y axis is the range of PIs for the bulls in the decile. The range is shown by the vertical line, the mean by the X.

Table 25. Means and SDs for PI and PA by NAAB code.

<u>NAAB Code</u>	<u>n</u>	<u>PI</u>		<u>PA</u>	
		<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
A	78	.44	.38	.90	.43
B	348	.50	.39	.94	.49
C	182	.39	.46	.74	.52
D	180	.42	.40	.86	.45
E	289	.41	.40	.82	.48
F	147	.28	.50	.66	.54
G	14	.36	.33	.94	.38
H	258	.41	.43	.78	.47
I	54	.33	.43	.75	.51
J	497	.50	.41	.97	.46
All		.44	.41	.86	.49

Table 26. Means and SD for genetic values for FS by NAAB code.

<u>NAAB Code</u>	<u>PTAT</u>		<u>PTAT_{SIRE}</u>		<u>PTAT_{DAM}</u>		<u>PTAT-PA</u>		<u>DTD-PA</u>	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
A	.76	.67	.62	.77	1.18	.60	-.01	.38	.33	.83
B	.81	.74	.72	.74	1.14	.65	-.03	.38	.08	.91
C	.64	.73	.52	.78	.96	.68	.003	.34	-.06	.72
D	.82	.63	.54	.75	1.15	.55	.06	.36	-.06	.72
E	.74	.66	.48	.72	1.01	.69	.02	.40	.23	.72
F	.54	.74	.29	.89	1.01	.69	.02	.40	.23	.72
G	.70	.51	.56	.69	1.32	.66	-.14	.31		
H	.63	.59	.48	.79	1.09	.57	-.06	.34	-.22	.68
I	.67	.67	.33	.81	1.18	.68	-.04	.54	-.33	.88
J	.86	.67	.63	.71	1.31	.62	-.02	.40	-.03	.90
All	.72	.66	.56	.76	1.16	.63	-.02	.38	.04	.83

Table 27. Regression Coefficients, R^2 s, and intercepts for regression of PTAT on PA and PI for different NAAB codes.

NAAB Code	<u>PA</u>			<u>PI</u>		
	<u>B₁</u>	<u>R²</u>	<u>a</u>	<u>B₁</u>	<u>R²</u>	<u>a</u>
A	.86	.31	-6.6	.80	.26	33.5
B	1.08	.51	-20.7	1.17	.40	20.8
C	.98	.54	-3.4	1.11	.44	21.5
D	.89	.41	5.9	.85	.29	44.4
E	.84	.40	5.2	.81	.25	38.2
F	.94	.49	-2.9	.84	.33	34.2
G	.89	.46	-14.4	.75	.19	39.0
H	.75	.37	6.5	.63	.20	36.7
I	.82	.38	5.8	.72	.21	41.8
J	.80	.32	8.3	.92	.27	38.2

Table 28. Means and SDs for PTAT at the time of the bull's second proof estimated from PA and PI of the proof immediately prior to his initial proof by NAAB code.

<u>NAAB</u>	<u>PI</u>				<u>PA</u>			
	<u>Across</u>		<u>Within</u>		<u>Across</u>		<u>Within</u>	
	<u>MN</u>	<u>SD</u>	<u>MN</u>	<u>SD</u>	<u>MN</u>	<u>SD</u>	<u>MN</u>	<u>SD</u>
A	.75	.37	.76	.34	.81	.39	.76	.39
B	.81	.37	.81	.47	.84	.44	.82	.53
C	.70	.40	.64	.48	.66	.47	.69	.51
D	.73	.35	.82	.33	.77	.40	.84	.41
E	.72	.37	.74	.35	.73	.43	.76	.42
F	.59	.44	.54	.42	.59	.48	.57	.50
G	.73	.28	.70	.23	.84	.35	.70	.35
H	.73	.28	.70	.23	.84	.35	.70	.35
I	.66	.39	.67	.29	.68	.46	.67	.40
J	.82	.35	.86	.34	.87	.42	.87	.37

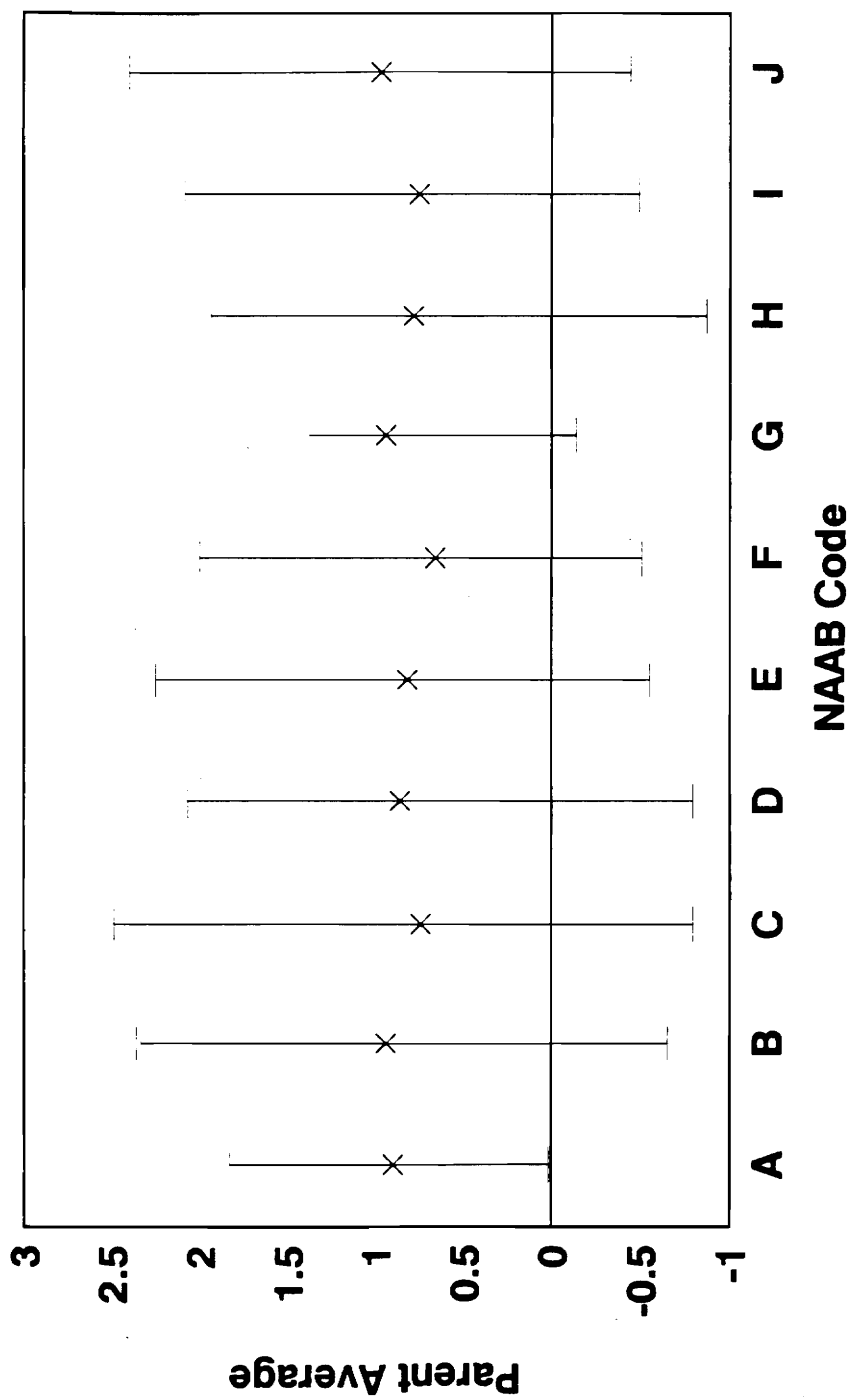


Figure 3. Average (x) and Range (I) of PA by NAAB Code.

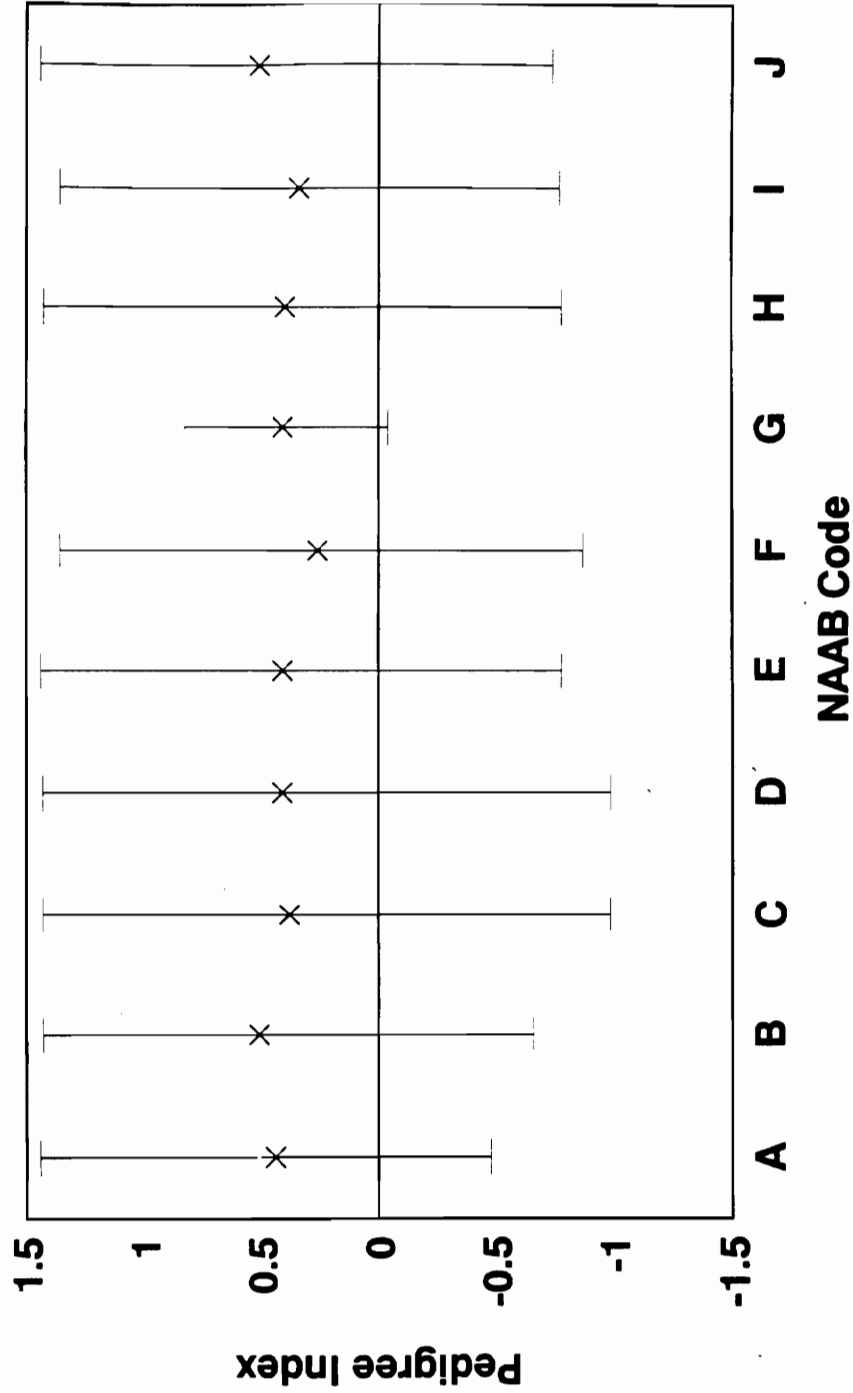


Figure 4. Average (x) and Range (l) of PI by NAAB Code.

Conclusions

PA was a better predictor of PTA and DTD for FS than PI in A.I. sampled bulls. Regression coefficients were less than one for both PA and PI, but R^2 s were higher for PA. Thus the use of $PTAT_{DAM}$ allows a more accurate estimate of genetic value for type than models using $PTAT_{MGS}$. Bulls with high PA were more likely to have high PTATs. Accuracy of prediction increased as the time of the estimation approached the date of the initial proof.

Higher R^2 s for PTAT than DTD were due to the part whole relationship between PTAT and pedigree information. PA had higher R^2 s than other pedigree sources for DTD, but R^2 s were much lower than for PTAT.

The variance correction implemented in S 93 did not appear to have had a significant impact on PTAT of bull dams.

The inclusion of first granddaughter's in the bull dam's proof resulted in a slight increase in $PTAT_{DAM}$. The impact of the change in $PTAT_{INITIAL} - PA_{PREVIOUS}$ on $PTAT_{DAM}$ was minimal. However, DTD and $DTD - PA$ were useful in predicting $\Delta PTAT_{DAM}$.

The amount of difference in PTAT that PA explained in different NAAB codes varied widely for reasons that were not apparent. However, differences between studs in PA estimated from PTAT were minimal and it was concluded that effects of testing populations were not large.

Bibliography

1. Atkeson, G. W., C. E. Meadows, and L. D. McGilliard. 1969. Weighting components of type in classifying Holsteins. J. Dairy Sci. 52:1638.
2. Bell, B. R., W. E. Vinson, J. M. White, H. D. Norman, and R. H. Kliwer. 1982. Effects of genetic merit of herd mates on sire summaries for type in Holstein cattle. J. Dairy Sci.
3. Berland, Steve. 1993. Enhanced Classification System. Holstein-Friesian Association of America. 1993. Brattleboro, VT.
4. Butcher, K. R. and J. E. Legates. 1976. Estimating son's progeny test from his pedigree information. J. Dairy Sci. 59:137
5. Cassell, B. G., J. M. White, W. E. Vinson and R. H. Kliwer. 1973. Genetic and phenotypic relationships among type traits in Holstein-Friesian cattle. J. Dairy Sci. 56:1171
6. DeLorenzo, M. A. and R. W. Everett. 1982. Relationships between milk and fat production, type, and stayability in Holstein sire evaluations. J. Dairy Sci. 65:1277
7. Funk, Dennis. Cow PTA-The Forgotten Genetic Measure. 1993. Holstein World. 90 (2): 106

8. Funk, D. C. and L. B. Hansen. 1988. Predictability of pedigree indexes for bulls from two, three, and four generations. J. Dairy Sci. 71:3148
9. Henderson, C. R. 1964. Selecting the young sire to sample in artificial insemination. J. Dairy Sci. 47:439
10. McNeill, W. W., W. E. Vinson, J. M. White, and R. H. Kliever. 1975. Predicting future type proofs of young Holstein bulls from pedigree information. J. Dairy Sci. 59:527
11. Powell, R.L., G.R. Wiggins, and H.D. Norman. Effect of Sampling Status and Adjustment for Heterogeneous Variance on Bias in Bull Evaluations. J. Dairy Sci. 77:883-890
12. Rothschild, M. F., L.W. Douglas, and R. L. Powell. 1981. Prediction of son's modified contemporary comparison from pedigree information. J. Dairy Sci. 64:331
13. Samuelson, David J. 1992. Accuracy of predicting genetic merit of A.I. sampled bulls from pedigree information and the impact of son's proof on dam's PTA. Master's Thesis. Department of Dairy Science. Virginia Polytechnic Institute and State University, Blacksburg.
14. Short, T. A. and T. J. Lawlor. 1992. Genetic parameters of conformation traits, milk yield, and herd

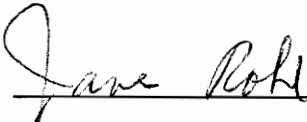
- life in Holsteins. J. Dairy Sci. 75:1987
15. Smothers, C. D., R. E. Pearson, I. Hoeschele, and D. A. Funk. 1993. Herd final score and it's relationship to genetic and environmental parameters of conformation traits of United States Holsteins. J. Dairy Sci. 76:1671
 16. Stewart, H. M., E. B. Burnside, G. B. Weaver, and M. G. Freeman. 1976. Relationship between pedigree index and progeny performance for Canadian dairy bulls. Can. J. Anim. Sci. 56:715
 17. VanRaden, P. M., E. L. Jensen, T. J. Lawlor, and D. A. Funk. 1990. Prediction of transmitting abilities for Holstein type traits. J. Dairy Sci. 73:191
 18. VanRaden, P.M., and G. R. Wiggins. 1991. Derivation, calculation, and use of national animal model information. J. Dairy Sci. 74:2737
 19. Van Vleck, L. D. 1977. Theoretical and actual genetic progress in dairy cattle. Proceedings of the International Conference on Quantitative Genetics August 16-21, 1976. Iowa State University Press. Edited by Edward Pollack, Oscar Kempthorne and Theodore B. Bailey, Jr. pp547
 20. Van Vleck, L. D. 1982. Theoretical weight's for regression of a son's genetic evaluation on his sire's and dams genetic evaluations. J. Dairy Sci. 65:164
 21. Vinson, W. E., and A. E. Freeman. 1972. Selection of

- Holstein bulls for use in Artificial insemination. J. Dairy Sci. 55:1621
22. Vinson, W. E., J. M. White, B. L. Combs and R. H. Kliewer. 1975. Sources of Variation in Holstein descriptive classification traits. J. Dairy Sci. 59:522
23. Weigel, K.A., and T. J. Lawlor. 1994. Adjustment for Heterogeneous Variance in Genetic Evaluations for Conformation of United States Holsteins. J. Dairy Sci. 77:1691-1701
24. Wiggans, G.R., and P.M. VanRaden. 1989. USDA-DHIA Animal Model Genetic Evaluations. National Cooperative Dairy Herd Improvement Program Handbook, Fact Sheet H-2.
25. Wilcox, C. J., K. O. Pfau, and J. W. Bartlett. 1957. Variations of scores among different characteristics of Holstein cows for herdlife classification. J. Dairy Sci. 40:1174

Vita

James Rohl was born September 8, 1955, and grew up on his family's farm. He attended Michigan State University, and graduated with his Bachelor of Science degree in 1992.

In August, 1992, he enrolled at Virginia Tech to pursue a Master of Science in Dairy Science, and fulfilled the degree requirements on August 29, 1994.



James Rohl