1987

Genetics of racing performance in the American Quarter Horse

Samuel T. Buttram

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Biostatistics Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/11670

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.

- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.

- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17”x 23” black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6”x 9” black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.
Genetics of racing performance in the American Quarter Horse

Buttram, Samuel T., Ph.D.
Iowa State University, 1987
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages ✓
2. Colored illustrations, paper or print ✓
3. Photographs with dark background ✓
4. Illustrations are poor copy ✓
5. Pages with black marks, not original copy ✓
6. Print shows through as there is text on both sides of page ✓
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements ✓
9. Tightly bound copy with print lost in spine ✓
10. Computer printout pages with indistinct print ✓
11. Page(s) lacking when material received, and not available from school or author.
12. Page(s) seem to be missing in numbering only as text follows.
13. Two pages numbered. Text follows.
14. Curling and wrinkled pages ✓
15. Dissertation contains pages with print at a slant, filmed as received ✓
16. Other

University
Microfilms
International
Genetics of racing performance in the American Quarter Horse

by

Samuel T. Buttram

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

Department: Animal Science Major: Animal Breeding

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major/Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1987
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>SECTION I. DESCRIPTION OF THE DATA</td>
<td>3</td>
</tr>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>5</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>7</td>
</tr>
<tr>
<td>References</td>
<td>27</td>
</tr>
<tr>
<td>SECTION II. ADJUSTMENT FACTORS AND CONTEMPORARY GROUPS</td>
<td>28</td>
</tr>
<tr>
<td>Abstract</td>
<td>28</td>
</tr>
<tr>
<td>Introduction</td>
<td>29</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>30</td>
</tr>
<tr>
<td>Data</td>
<td>30</td>
</tr>
<tr>
<td>Adjustment factors</td>
<td>31</td>
</tr>
<tr>
<td>Contemporary group effects</td>
<td>32</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>33</td>
</tr>
<tr>
<td>Data</td>
<td>33</td>
</tr>
<tr>
<td>Adjustment factors</td>
<td>35</td>
</tr>
<tr>
<td>Contemporary group effects</td>
<td>45</td>
</tr>
<tr>
<td>References</td>
<td>48</td>
</tr>
</tbody>
</table>
### SECTION III. VARIANCE COMPONENT ESTIMATION

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>50</td>
</tr>
<tr>
<td>Introduction</td>
<td>51</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>53</td>
</tr>
<tr>
<td>Data</td>
<td>53</td>
</tr>
<tr>
<td>Model</td>
<td>53</td>
</tr>
<tr>
<td>Estimation procedures</td>
<td>55</td>
</tr>
<tr>
<td>Computational aspects</td>
<td>58</td>
</tr>
<tr>
<td>Heritability and repeatability</td>
<td>59</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>61</td>
</tr>
<tr>
<td>References</td>
<td>73</td>
</tr>
</tbody>
</table>

### SUMMARY

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>77</td>
</tr>
</tbody>
</table>

### REFERENCES

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>79</td>
</tr>
</tbody>
</table>

### ACKNOWLEDGEMENTS

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
</tr>
</tbody>
</table>

### APPENDIX A. TALBES AND FIGURES

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>84</td>
</tr>
</tbody>
</table>

### APPENDIX B. COMPUTER PROGRAM TO EDIT DATA, ASSIGN EQUATION NUMBERS AND SORT RECORDS

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
</tr>
</tbody>
</table>

### APPENDIX C. PL/1 PROGRAM TO ADJUST RACING TIMES FOR SEX AND AGE

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>113</td>
</tr>
</tbody>
</table>

### APPENDIX D. PL/1 PROGRAM TO ABSORB CONTEMPORARY GROUP EQUATIONS INTO SIRE AND HORSE EQUATIONS

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>116</td>
</tr>
</tbody>
</table>

### APPENDIX E. FORTRAN PROGRAM TO ESTIMATE VARIANCE COMPONENTS USING IHSIM

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>122</td>
</tr>
</tbody>
</table>

### APPENDIX F. FORTRAN PROGRAM TO ESTIMATE VARIANCE COMPONENTS USING THM

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Distributional properties of racing time by distance</td>
<td>8</td>
</tr>
<tr>
<td>Table 2</td>
<td>Frequency distributions and means of horses grouped by number of races run and distance</td>
<td>14</td>
</tr>
<tr>
<td>Table 3</td>
<td>Frequency distributions and progeny means of sires grouped by number of progeny records and distance</td>
<td>16</td>
</tr>
<tr>
<td>Table 4</td>
<td>Summary of racing time by age and distance</td>
<td>21</td>
</tr>
<tr>
<td>Table 5</td>
<td>Summary of racing time by sex and distance</td>
<td>22</td>
</tr>
<tr>
<td>Table 6</td>
<td>Summary of racing times used to estimate fixed effects by distance</td>
<td>34</td>
</tr>
<tr>
<td>Table 2</td>
<td>Means used for calculating sex adjustment factors</td>
<td>36</td>
</tr>
<tr>
<td>Table 3</td>
<td>Additive and multiplicative sex adjustment factors by distance</td>
<td>37</td>
</tr>
<tr>
<td>Table 4</td>
<td>Additive age adjustments by sex and distance</td>
<td>42</td>
</tr>
<tr>
<td>Table 5</td>
<td>Regression coefficients for adjusting racing times for differences in handicap weight by sex and distance</td>
<td>44</td>
</tr>
<tr>
<td>Table 6</td>
<td>Percentages of the total variance accounted for by tracks, years, days and individual races by distance</td>
<td>46</td>
</tr>
<tr>
<td>Table 1</td>
<td>Summary of data sets used for estimating variance components by distance</td>
<td>64</td>
</tr>
<tr>
<td>Table 2</td>
<td>Estimates of heritability for racing performance by data set from three methods</td>
<td>66</td>
</tr>
<tr>
<td>Table 3</td>
<td>Estimates of repeatability for racing performance by data set from three methods</td>
<td>67</td>
</tr>
</tbody>
</table>
Table 4. Weighted means of heritability and repeatability estimates from the prior method by age and data source 69

Table 5. Weighted means of heritability and repeatability estimates from the prior method by distance 71

Table A1. Number of racing records by year and distance from 1971-1986 85

Table A2. Means for racing time by year and distance from 1971-1986 86

Table A3. Summary of racing time by month and distance 87

Table A4. Frequency distributions for number of horses per race by distance 88

Table A5. Summary of racing times at the five largest tracks for each of five racing distances 89

Table A6. Frequency distributions for number of tracks per horse by distance 90

Table A7. Frequency distributions for number of tracks per sire by distance 91

Table A8. Handicap weight means and standard deviations by distance 92
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.</td>
<td>Frequency distribution of racing times from 320 m racing records</td>
<td>11</td>
</tr>
<tr>
<td>Figure 2.</td>
<td>Relationship between handicap weight and mean racing time from 320 m racing data (data labels are the percentages of the data represented)</td>
<td>19</td>
</tr>
<tr>
<td>Figure 3.</td>
<td>Relationship between age and mean racing time from 320 m racing records by sex</td>
<td>25</td>
</tr>
<tr>
<td>Figure 1.</td>
<td>Actual (symbols only) and predicted (regression lines) age effects relative to two-year-olds from 320 m racing records by sex</td>
<td>40</td>
</tr>
<tr>
<td>Figure A1.</td>
<td>Frequency distribution of racing times from 201 m racing records</td>
<td>93</td>
</tr>
<tr>
<td>Figure A2.</td>
<td>Frequency distribution of racing times from 366 m racing records</td>
<td>94</td>
</tr>
<tr>
<td>Figure A3.</td>
<td>Frequency distribution of racing times from 402 m racing records</td>
<td>95</td>
</tr>
<tr>
<td>Figure A4.</td>
<td>Frequency distribution of racing times from 796 m racing records</td>
<td>96</td>
</tr>
<tr>
<td>Figure A5.</td>
<td>Relationship between handicap weight and mean racing time from 201 m racing data (data labels are the percentages of the data represented)</td>
<td>97</td>
</tr>
<tr>
<td>Figure A6.</td>
<td>Relationship between handicap weight and mean racing time from 366 m racing data (data labels are the percentages of the data represented)</td>
<td>98</td>
</tr>
<tr>
<td>Figure A7.</td>
<td>Relationship between handicap weight and mean racing time from 402 m racing data (data labels are the percentages of the data represented)</td>
<td>99</td>
</tr>
</tbody>
</table>
Figure A8. Relationship between handicap weight and mean racing time from 796 m racing data (data labels are the percentages of the data represented) 100

Figure A9. Relationship between age and mean racing time from 201 m racing records by sex 101

Figure A10. Relationship between age and mean racing time from 366 m racing records by sex 102

Figure A11. Relationship between age and mean racing time from 402 m racing records by sex 103

Figure A12. Relationship between age and mean racing time from 796 m racing records by sex 104

Figure A13. Actual (symbols only) and predicted (regression lines) age effects relative to two-year-olds from 201 m racing records by sex 105

Figure A14. Actual (symbols only) and predicted (regression lines) age effects relative to two-year-olds from 366 m racing records by sex 106

Figure A15. Actual (symbols only) and predicted (regression lines) age effects relative to two-year-olds from 402 m racing records by sex 107

Figure A16. Actual (symbols only) and predicted (regression lines) age effects relative to three-year-olds from 796 m racing records by sex 108
INTRODUCTION

Quarter Horse racing began in the United States as early as 1674 in Virginia (Essary, 1983) and became popular because the short sprint races could be run down village streets. Quarter Horse races are still characterized as short races run at top speed, usually on a straight track, but today Quarter Horse racing is an industry. More than 100 race tracks that conduct Quarter Horse races are located throughout North America. The American Quarter Horse Association maintains individual racing records for all horses that race in officially sanctioned races at these tracks. These records have been maintained since 1940 and provide an extensive data base from which breeders may receive information on the performance of individual horses. These records are summarized periodically in terms of the total number of placings or money won, but a method of summarizing the data is needed that accurately evaluates the genetic potential of horses.

Mixed-model methodology is available for obtaining best linear unbiased predictions (BLUP) of breeding values for sires, dams and individual horses that have no progeny (Wilson et al., 1986b). In order to use this methodology, it is necessary to know what environmental sources may cause variation in racing performance and also to obtain estimates of heritability and repeatability.
The purpose of this study is three-fold. Because racing performance in Quarter Horses has not been previously studied, the first objective is to describe Quarter Horse racing data by using racing times supplied by the American Quarter Horse Association. The second objective is to examine the importance of various sources of environmental variation and to calculate adjustments for eventual preadjusting of the data when appropriate. The final purpose of the study is to estimate heritability and repeatability of racing time that can be used in a genetic evaluation of Quarter Horses for racing performance.
SECTION I. DESCRIPTION OF THE DATA

Abstract

Over one million racing records of American Quarter Horses were used to provide a description of Quarter Horse racing data. The data represented five racing distances (201, 320, 366, 402 and 796 m). Finish time was used to measure racing performance. Means and variances for finish time increased with length of the race, but the distributions were similar for the five distances. Each distribution was skewed to the right and more peaked than a normal distribution. Repeated records were an important source of information for improving accuracy of genetic evaluations and for comparing horses and sires across races and tracks. There was a tendency for more older horses and geldings to be found in the longer races. Two-year-old horses raced almost exclusively against other two-year-olds and most three-year-olds raced only with horses their own age. An interaction was found between sex and age that was interpreted to be the result of differential selection rates among the sexes. Further study of age and sex effects is needed to develop adjustment factors for preadjusting Quarter Horse racing data for use in genetic evaluations.
Introduction

Quarter Horse races are characterized as relatively short races run at top speed for a matter of seconds. Horses start from a starting gate and usually race the entire distance on a straightaway or, in the case where races are longer than 402 m, a track with one turn. Quarter Horse races may range in length from 201 to 796 m. The racing distances reported in this study are 201 m (220 yards), 320 m (350 yards), 366 m (400 yards), 402 m (440 yards) and 796 m (870 yards). Younger horses tend to run in the shorter races in preparation for the more prestigious and lucrative races that are run at the quarter mile (402 m) distance from which the "Quarter Horse" received its name.

The American Quarter Horse Association (AQHA) maintains individual performance records on all horses that have raced in officially sanctioned pari-mutual and nonpari-mutual Quarter Horse races in the United States, Canada and Mexico. These records comprise a large data base which can be used by Quarter Horse breeders in the selection of horses for racing performance. Records for the top young horses and progeny groups are summarized periodically in terms of total number of placings or money won. Ojala and Van Vleck (1981) found that time traits were more heritable than traits based on placings or money won and would be more useful in
evaluating sires for racing ability. The only measure of racing ability used in this study was finish time.

To the author's knowledge no previous attempt has been made to genetically evaluate Quarter Horses for racing time. Adjustment factors for important sources of variation are unknown as are the heritability and repeatability for finish time. The purpose of this paper is to provide a description of Quarter Horse racing data to aid in the further analyses and interpretation of results necessary to carry out genetic evaluations of American Quarter Horses for racing performance.

**Materials and Methods**

Individual racing records of horses that raced from Jan. 1, 1971 to Sept. 1, 1986 were obtained from the AQHA. Racing distances included 201, 320, 366, 402 and 796 meters. Information available on each horse included the registration number of the horse, year of birth, sex, sire, dam and birth year of each of the parents. In many cases multiple records were obtained for individual horses. Each record included the track and date of the race, a unique race designation, the handicap weight and finish time measured in hundredths of a second.

Handicap weights were assigned to horses by a racing secretary in an effort to create more evenly matched races.
Weight was added in .4536 kg increments in the form of a lead pad. After each race, one weight was taken of the jockey, his tack and any extra weight so that the handicap weight included the total weight carried by the horse.

The original data sets included all records of horses with finish times. Only the first horse to cross the finish line was timed in each race and times for all other horses were calculated from a photograph by converting into seconds the distance between the first horse and every other horse. Most races were electronically timed but hand-held timers were used in cases where electronic timers were not available or did not work.

The data were initially sorted into five data sets based on racing distance. Data from the 201, 320, 366, 402 and 796 m races made up data sets one (DS1), two (DS2), three (DS3), four (DS4) and five (DS5), respectively. Data from each distance were considered separately and summary statistics were calculated for each of the five data sets. The data were examined for normality and for the presence of extreme values. Summaries also were made by years, months and tracks.

The number of races per horse and progeny per sire were summarized for each data set. Horses with individual records were classified into one of three groups based on the total number of races run within each distance. The
number of horses in each group and the group means were calculated in order to provide a description of repeated records. Sires were classified similarly, based on progeny number, so that the number of progeny per sire could be summarized. Frequency distributions were calculated for the number of tracks at which individual horses and sire progeny groups raced to help determine whether comparisons could be made across tracks or racing was localized within tracks.

Summaries were made by sex and age (calculated as the difference between year of the race and birth year). Various combinations of sexes and ages within races also were observed. Plots of sex by age subclass means were used to check for the presence of interaction and the possible effects of selection were discussed. Handicap weight means and standard deviations were calculated for each data set. Mean racing times at each weight also were calculated and the relationship between handicap weight and average time was examined.

Results and Discussion

Over one million records were included in the five original data sets (Table 1). Two of data sets (DS2 and DS3) were extremely large and together contained over 82% of the total number of records. The 402 and 796 m races each
Table 1. Distributional properties of racing time by distance

<table>
<thead>
<tr>
<th>Item</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. Obs.</td>
<td>16,540</td>
<td>561,738</td>
<td>368,846</td>
<td>92,664</td>
<td>94,025</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.79</td>
<td>18.75</td>
<td>21.05</td>
<td>22.97</td>
<td>47.65</td>
</tr>
<tr>
<td>Std. Dev., s</td>
<td>.408</td>
<td>.550</td>
<td>.579</td>
<td>.643</td>
<td>1.287</td>
</tr>
<tr>
<td>Minimum, s</td>
<td>11.40</td>
<td>16.29</td>
<td>19.18</td>
<td>21.02</td>
<td>44.30</td>
</tr>
<tr>
<td>Maximum, s</td>
<td>15.70</td>
<td>24.89</td>
<td>26.73</td>
<td>27.77</td>
<td>56.30</td>
</tr>
<tr>
<td>Best time(^a), s</td>
<td>11.62</td>
<td>17.20</td>
<td>19.18</td>
<td>21.02</td>
<td>44.30</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.05</td>
<td>1.21</td>
<td>1.17</td>
<td>.81</td>
<td>.49</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.30</td>
<td>4.40</td>
<td>7.99</td>
<td>1.85</td>
<td>21.08</td>
</tr>
</tbody>
</table>

\(^a\)World record time recognized by the AQHA.
accounted for about 8% of the data and DS1 contained less than 2% of the records.

Means and standard deviations for time increased with length of the race. The linear regression of time on distance (R-square = .997) showed there was little evidence that horses were beginning to tire at these distances. In fact, the average speed (m/s) for the entire race increased as distance increased up to 402 m. This points out the importance of both a quick start and ability to sprint in determining the racing performance of a horse. Certainly there is less "holding back" and "jockeying" for position in Quarter Horse racing than is found in the longer Thoroughbred and Standardbred races. There is also possibly less "pace of the race" effect than that described by Tolley et al. (1983) in Standardbred trotters and pacers.

Except for the obvious differences in means and variances, all of the data sets had distributions that were similar. For example, the official best time for each distance was about three standard deviations below the mean in each data set. The percentage of observations greater than three standard deviations above the mean of each data set ranged from .8 to 1.2% and the percentage greater than four standard deviations ranged from .2 to .4% for the five data sets.
The distributions were similar in shape (Figure 1; Appendix A, Figures A1-A4) and when tested with the Kolomogorov D statistic (SAS, 1985), each deviated (p<.01) from a normal distribution. In each case the distribution was skewed to the right and appeared more peaked than a normal distribution. Some skewness was expected due to the physical limit of the ability of the horse to run infinitely fast. This limit, however, did not appear to be as important in causing skewness as the fact that there was essentially no limit on how slow a horse could run. As can be seen from the range in the data, times in the 320 m race were located as far as eleven standard deviations above the mean. However, relatively few extreme values were found in the very narrow upper tail of the distribution and it was assumed that these values were the result of horses that were affected by extraneous circumstances that were nongenetic. For example, horses that do not have a chance to place are sometimes allowed to "coast" through the last part of the race. This is one possible cause of skewness. At some point, however, it becomes difficult to decide between horses that were affected by nongenetic factors and those that were genetically slow. Objective criteria should be used to edit extreme values from these data.
Figure 1. Frequency distribution of racing times from 320 m racing records
Wilson et al. (1986), using a subset of the same data used in this study, reported values of skewness and kurtosis similar to the ones in Table 1 and found that a natural log transformation reduced both skewness and kurtosis. However, the product-moment and Spearman rank correlations between breeding value estimates obtained from transformed and untransformed data were .9996 and .9997 respectively. After extreme values were removed, lack of normality was not considered to be a major problem with the data in this study.

The numbers of Quarter Horse races increased steadily each year from 1971 to 1983 (Appendix A, Table A1) and appeared to stabilize after 1983. Trends toward faster times were evident in the yearly means (Appendix A, Table A2) of the shorter distances, particularly during recent years. A further study of phenotypic, environmental and genetic trends will be done following the genetic evaluation. Races were distributed throughout the year (Appendix A, Table A3) with some concentration of racing during the summer months (May through September). No trends in monthly means were found.

The number of horses per race (Appendix A, Table A4) ranged from 1 to 12 with 10 being the most common in the three medium length races and 8 the most common in DS1 and
DS5. The average race sizes, calculated from Table 1, were 6.8, 8.2, 8.3, 8.1 and 7.8 horses per race in the 201, 320, 366, 402 and 796 m races, respectively.

Most of the race tracks represented accounted for less than one percent of the records for any given data set. But in each case, there were three to five tracks (Appendix A, Table A5) that each accounted for as much as 5 to 17% of the races at that distance. There did seem to be some specialization of race tracks in that tracks which dominated racing in the medium and long distance races were not the same tracks with large numbers of 201 m races. Of the five largest tracks in each data set, about two-thirds had mean times faster than the overall data set indicating either better track management, superior horses, or both. Within track standard deviations for these tracks were always smaller than overall standard deviations.

The average numbers of records per horse were 1.4, 4.5, 4.0, 2.4 and 5.9 for the 201, 320, 366, 402 and 796 m races respectively. Horses in DS1 had relatively few repeated records (Table 2) and 72.8% of them raced only once at that distance. About 60% of the horses found in the two larger data sets (DS2 and DS3) had from 2 to 9 records and a number of horses raced more than ten times each. Horses in the 796 m race appeared to race more times than horses in other
Table 2. Frequency distributions and means of horses grouped by number of races run and distance

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>1</th>
<th>2-9</th>
<th>10&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. horses</td>
<td>Pct. horses</td>
<td>Mean time, s</td>
</tr>
<tr>
<td>201</td>
<td>8,395</td>
<td>72.8</td>
<td>12.88</td>
</tr>
<tr>
<td>320</td>
<td>33,805</td>
<td>26.9</td>
<td>19.28</td>
</tr>
<tr>
<td>366</td>
<td>29,555</td>
<td>32.1</td>
<td>21.56</td>
</tr>
<tr>
<td>402</td>
<td>19,592</td>
<td>50.4</td>
<td>23.31</td>
</tr>
<tr>
<td>796</td>
<td>6,252</td>
<td>39.0</td>
<td>48.76</td>
</tr>
</tbody>
</table>
races. Almost 15% of the horses in DS5 had more than 10 records each and as many as 19 horses raced 100 times or more at 796 m.

Horses with more records had faster overall mean times than horses with fewer records. This was probably because horses that ran fast in earlier races were selected to perform in future races, but the poorer performing horses were not allowed to have repeated records. Possibly horses improved with experience also.

The percentage of horses that raced at multiple tracks in each of the data sets (Appendix A, Table A6) was related to the percentage of horses with repeated records. Data sets with a higher percentage of horses with repeated records also had a higher percentage of horses that raced at more than one track. In DS1, 91.6% of the horses raced at only one track while the percentages were 48.5 and 54.3% for the horses in DS2 and DS3, respectively. About 5% of the horses in the 320, 366 and 796 m races raced at five or more tracks but that percentage was considerably lower for either DS1 or DS4.

A study of sire progeny groups (Table 3) indicated that a surprisingly high percentage (44.9 to 59.3%) of the sires represented in each data set had only one progeny with performance records. Means were 2.9, 8.0, 7.2, 5.2 and 2.7
Table 3. Frequency distributions and progeny means of sires grouped by number of progeny records and distance

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>No. sires</th>
<th>1</th>
<th>2-9</th>
<th>10+</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
<td></td>
<td>2,375</td>
<td>1,421</td>
<td>207</td>
</tr>
<tr>
<td></td>
<td>Pct. sires</td>
<td>59.3</td>
<td>35.5</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>Prog. mean time, s</td>
<td>12.96</td>
<td>12.84</td>
<td>12.74</td>
</tr>
<tr>
<td>320</td>
<td></td>
<td>7,021</td>
<td>6,272</td>
<td>2,338</td>
</tr>
<tr>
<td></td>
<td>Pct. sires</td>
<td>44.9</td>
<td>40.1</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td>Prog. mean time, s</td>
<td>19.41</td>
<td>19.13</td>
<td>18.82</td>
</tr>
<tr>
<td>366</td>
<td></td>
<td>6,054</td>
<td>4,990</td>
<td>1,731</td>
</tr>
<tr>
<td></td>
<td>Pct. sires</td>
<td>47.4</td>
<td>39.1</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td>Prog. mean time, s</td>
<td>21.74</td>
<td>21.43</td>
<td>21.12</td>
</tr>
<tr>
<td>402</td>
<td></td>
<td>3,856</td>
<td>2,872</td>
<td>728</td>
</tr>
<tr>
<td></td>
<td>Pct. sires</td>
<td>51.7</td>
<td>38.5</td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td>Prog. mean time, s</td>
<td>23.56</td>
<td>23.26</td>
<td>22.97</td>
</tr>
<tr>
<td>796</td>
<td></td>
<td>3,469</td>
<td>2,128</td>
<td>267</td>
</tr>
<tr>
<td></td>
<td>Pct. sires</td>
<td>59.2</td>
<td>36.3</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>Prog. mean time, s</td>
<td>48.69</td>
<td>48.19</td>
<td>47.59</td>
</tr>
</tbody>
</table>
progeny per sire for DS1 through DS5, respectively. A few sires had rather large numbers of progeny. For instance in DS2, 1% of the sires each had more than 110 progeny and more than 575 progeny records. The average number of progeny records per sire was 4.1, 35.9, 28.9, 12.4 and 16.0 for DS1 through DS5. More evidence of selection was seen in that sires with more offspring had consistently faster progeny averages than sires with small progeny groups. Again, the number of tracks at which a sire’s progeny raced (Appendix A, Table A7) increased as the number of progeny per sire increased in that particular data set. In DS2 through DS5 more than 50% of the sires had progeny that raced at more than one track and from 17.7 to 29.1% of the sires had progeny records at five tracks or more.

These findings indicate that racing performance of individual horses is not entirely localized within tracks and repeated records can help to create comparisons between individual horses across races and tracks. The amount of information contributed by repeated records should significantly increase the accuracy of a genetic evaluation of horses for racing performance. This is particularly true if racing performance is lowly to moderately heritable. Use of all records seems to have more intuitive appeal than using only best time for each year as suggested by Ojala and
Van Vleck (1981). Best annual racing time is affected by a combination of favorable environmental effects that occur simultaneously (Ojala et al., 1987). Unless these effects are removed, best time does not provide an accurate measure of overall racing ability.

Mean handicap weight was approximately 55 kg (Appendix A, Table A8) and was almost identical in all five data sets. Standard deviations were in the range of 1.1 to 1.3 kg. There was little relationship between handicap weight and average performance except that horses which ran at 55.3 kg were remarkably faster than horses that carried either more or less weight (Figure 2). This phenomenon was seen in all of the data sets except DS1 (Appendix A, Figures A5-A8). In the 402 m races, horses were unusually fast when they carried either 55.3 or 54.4 kg. Interestingly, 55.3 and 54.4 kg were the most common weights carried and were the weights at which many horses were trained. It is difficult to draw conclusions from this relationship but possibly either faster horses run at these weights more often than slower horses or Quarter Horses are extremely sensitive to weights that deviate from that at which they were trained. Another consideration is that the better jockeys tend to ride at these standard weights even when they are allowed to carry less weight (Valis, 1987).
Figure 2. Relationship between handicap weight and mean racing time from 320 m racing data (data labels are the percentages of the data represented)
Many Quarter Horse races are classified according to age, money won, number of placings, etc., with the objective being that all horses within a race are of equal ability. Race classifications were not known in this study, but more than 50% of the races at each of the three shorter distances contained horses of the same age. Many of these were races for two-year-old horses which rarely raced against older horses. A number of races for only three-year-olds were found in DS2, DS3 and DS4 but 40% or more of the three-year-olds ran in races with mixed ages. No differentiation was apparent for age groups greater than three years old or for sex with respect to race classification.

The most common racing age for Quarter Horses (Table 4) was two years in the two shortest races, three years in the 366 and 402 m races and four years in the longest race. The percentage of two-year-old records in each data set decreased from 44.2% in DS1 to 14.4% in DS4 and only a negligible number of two-year-old horses raced in the 796 m race. In addition, about 20% of the records in DS5 were by horses that were more than six years old. These observations illustrate the tendency for younger horses to race in the shorter races as previously described. There was also a trend for more geldings to be found in the longer races (Table 5). The percentage of geldings increased from
<table>
<thead>
<tr>
<th>Age</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>7,318</td>
<td>227,542</td>
<td>93,914</td>
<td>13,315</td>
<td>13</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>44.2</td>
<td>40.5</td>
<td>25.5</td>
<td>14.4</td>
<td>.01</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.92</td>
<td>18.87</td>
<td>21.13</td>
<td>22.95</td>
<td>48.11</td>
</tr>
<tr>
<td>3 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>4,607</td>
<td>188,995</td>
<td>160,131</td>
<td>38,546</td>
<td>18,609</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>27.9</td>
<td>33.6</td>
<td>43.4</td>
<td>41.6</td>
<td>19.8</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.72</td>
<td>18.70</td>
<td>21.04</td>
<td>22.97</td>
<td>48.01</td>
</tr>
<tr>
<td>4 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>2,217</td>
<td>73,231</td>
<td>58,903</td>
<td>20,210</td>
<td>25,795</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>13.4</td>
<td>13.0</td>
<td>16.0</td>
<td>21.8</td>
<td>27.4</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.66</td>
<td>18.63</td>
<td>21.02</td>
<td>22.98</td>
<td>47.64</td>
</tr>
<tr>
<td>5 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>1,019</td>
<td>33,842</td>
<td>26,790</td>
<td>9,658</td>
<td>18,567</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>6.2</td>
<td>6.0</td>
<td>7.3</td>
<td>10.4</td>
<td>19.8</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.64</td>
<td>18.58</td>
<td>20.97</td>
<td>22.97</td>
<td>47.50</td>
</tr>
<tr>
<td>6 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>567</td>
<td>17,268</td>
<td>13,860</td>
<td>5,159</td>
<td>12,720</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>3.4</td>
<td>3.1</td>
<td>3.8</td>
<td>5.6</td>
<td>13.5</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.65</td>
<td>18.55</td>
<td>20.95</td>
<td>22.96</td>
<td>47.47</td>
</tr>
<tr>
<td>7 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>357</td>
<td>9,904</td>
<td>7,570</td>
<td>2,824</td>
<td>8,250</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>2.2</td>
<td>1.8</td>
<td>2.1</td>
<td>3.1</td>
<td>8.8</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.61</td>
<td>18.58</td>
<td>20.98</td>
<td>22.98</td>
<td>47.54</td>
</tr>
<tr>
<td>8 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>210</td>
<td>5,612</td>
<td>4,185</td>
<td>1,633</td>
<td>5,007</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>1.3</td>
<td>1.0</td>
<td>1.1</td>
<td>1.8</td>
<td>5.3</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.66</td>
<td>18.60</td>
<td>21.00</td>
<td>23.03</td>
<td>47.58</td>
</tr>
<tr>
<td>9 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>124</td>
<td>3,123</td>
<td>2,117</td>
<td>785</td>
<td>2,894</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>.8</td>
<td>.6</td>
<td>.6</td>
<td>.9</td>
<td>3.1</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.59</td>
<td>18.63</td>
<td>21.05</td>
<td>23.03</td>
<td>47.66</td>
</tr>
<tr>
<td>10 yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. obs.</td>
<td>75</td>
<td>1,472</td>
<td>925</td>
<td>337</td>
<td>1,420</td>
</tr>
<tr>
<td>Pct. obs.</td>
<td>.5</td>
<td>.3</td>
<td>.3</td>
<td>.4</td>
<td>1.5</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.65</td>
<td>18.68</td>
<td>21.10</td>
<td>23.13</td>
<td>47.70</td>
</tr>
</tbody>
</table>
Table 5. Summary of racing time by sex and distance

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stallion</td>
<td>No. obs.</td>
<td>3,451</td>
<td>109,764</td>
<td>79,984</td>
<td>21,798</td>
<td>13,915</td>
</tr>
<tr>
<td></td>
<td>Pct. obs.</td>
<td>20.9</td>
<td>19.5</td>
<td>21.7</td>
<td>23.5</td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td>Mean, s</td>
<td>12.76</td>
<td>18.72</td>
<td>21.02</td>
<td>22.91</td>
<td>47.68</td>
</tr>
<tr>
<td></td>
<td>Std. Dev., s</td>
<td>.406</td>
<td>.549</td>
<td>.572</td>
<td>.639</td>
<td>1.298</td>
</tr>
<tr>
<td>Mare</td>
<td>No. obs.</td>
<td>7,793</td>
<td>240,849</td>
<td>140,504</td>
<td>29,909</td>
<td>13,511</td>
</tr>
<tr>
<td></td>
<td>Pct. obs.</td>
<td>47.1</td>
<td>42.9</td>
<td>38.1</td>
<td>32.3</td>
<td>14.4</td>
</tr>
<tr>
<td></td>
<td>Mean, s</td>
<td>12.82</td>
<td>18.80</td>
<td>21.12</td>
<td>23.03</td>
<td>48.06</td>
</tr>
<tr>
<td></td>
<td>Std. Dev., s</td>
<td>.413</td>
<td>.571</td>
<td>.603</td>
<td>.693</td>
<td>1.306</td>
</tr>
<tr>
<td>Gelding</td>
<td>No. obs.</td>
<td>5,296</td>
<td>211,125</td>
<td>148,358</td>
<td>40,957</td>
<td>66,599</td>
</tr>
<tr>
<td></td>
<td>Pct. obs.</td>
<td>32.0</td>
<td>37.6</td>
<td>40.2</td>
<td>44.2</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td>Mean, s</td>
<td>12.76</td>
<td>18.69</td>
<td>21.01</td>
<td>22.96</td>
<td>47.57</td>
</tr>
<tr>
<td></td>
<td>Std. Dev., s</td>
<td>.399</td>
<td>.518</td>
<td>.547</td>
<td>.602</td>
<td>1.220</td>
</tr>
</tbody>
</table>
32% of the records in DS1 to 44% in DS4 while the percentage of stallions remained constant and the percentage of records from mares decreased. Geldings represented 71% of the records in DS5.

These findings revealed what appeared to be differential culling rates among sexes. More specifically, mares were removed from the racing circuit at younger ages than stallions, and geldings continued to race longer than either mares or stallions. The same occurrence also was seen within data sets. One explanation for this was that mares and stallions were removed from racing because they were selected for breeding stock. Since horses were not used for breeding and racing simultaneously, the differential culling rates observed may have been the result of differential selection rates among the sexes at different ages. This might also explain why mares seemed to be removed at earlier ages than stallions. Differential selection rates have other implications which will be discussed.

Age means (Table 4) showed that racing performance improved from two to four years of age in most cases and decreased after age six. Evidence of selection bias was seen in DS4 where a small number of the better young horses were chosen to race. Horses in this data set did not appear to improve from two to four years of age. Selection over
time may have caused the means of older age groups to be
biased downward. Sex means (Table 5) indicated that
geldings were slightly faster than stallions when summarized
over all ages and mares were slowest. Standard deviations
for racing time were also higher in mares.

Plots of sex by age subclass means (Figure 3; Appendix
A, Figures A9-A12) revealed an important interaction that
was seen in each of the four shorter races. As two-year-
olds, geldings ranked between mares and stallions for
performance in the 201, 320 and 366 m races and slower than
either of the sexes in DS4. However, by four or five years
of age geldings were faster than either mares or stallions
and continued to rank fastest as age increased. No
interaction was found in DS5 as geldings were fastest at all
ages (two-year-olds not included).

The possibility of a biological explanation for the
interaction cannot be overlooked at this point. However,
further study of age effects (Buttram, 1987) showed that
stallions and geldings had very similar maturity patterns in
terms of racing performance. Rather, the interaction was
Figure 3. Relationship between age and mean racing time from 320 m racing records by sex.
Conversely, faster stallions and mares were removed at relatively young ages to be used for breeding stock.

The difference in racing performance between stallions and mares is illustrated by the lines in Figure 3. In all of the data sets the lines for mares and stallions were almost parallel from ages two to four. Some spread in the difference between the two sexes occurred after four years of age.

From these observations it was not clear how much effect selection had on age means. To increase the accuracy of estimating age effects individual horse effects could be fit simultaneously in order to remove some of the bias caused by selection. The difference in performance between mares and stallions can be approximated by the difference between their respective means. The performance of geldings relative to mares and stallions, however, is not so easily determined. Sex and age effects should be accounted for when genetically evaluating Quarter Horses for racing performance, but further study is needed to develop appropriate adjustments for these effects.
References


SECTION II. ADJUSTMENT FACTORS AND CONTEMPORARY GROUPS

Abstract

Over one million individual racing records were used to estimate the effects of sex, age and handicap weight on racing performance in American Quarter Horses. Adjustment factors for preadjusting racing times were calculated for each of five distances. At 320 m, two-year-old stallions were .060 s faster than mares and .032 s faster than geldings of the same age. Multiplicative sex adjustment factors calculated from two-year-olds ranged from .9956 to .9988 for mares and .9946 to .9995 for geldings depending on distance. Sex adjustments generally increased as distance increased. Additive age adjustments were calculated separately for males and females because young mares appeared to have larger adjustments than young stallions and geldings. Two-year-old males and females were .097 and .161 s slower than four-year-olds of the same sex in 320 m races. At the same distance, three-year-olds were .035 and .062 s slower than four-year old males and females, respectively. Racing performance declined after age six. Regression coefficients for racing time (s) on handicap weight (kg) ranged from -.0051 to .0158. A hierarchical analysis of variance was used to evaluate the relative importance of tracks, years, days and individual races as sources of
variation in racing time. Tracks alone accounted for 10.6 to 31.8% of the variation depending on the distance of the race, but individual races within tracks, years and days should be used as contemporary groups. Contemporary groups defined in this way accounted for 49.4 to 70.9% of the variation in racing time.

Introduction

Individual racing records maintained by the American Quarter Horse Association are available to Quarter Horse breeders for selecting breeding stock. Data for individual horses or progeny groups are summarized periodically without regard for environmental differences. In order to genetically evaluate horses for racing performance it is necessary to make allowances for known sources of environmental variation. Environmental effects may be included in a genetic evaluation model, or when appropriate adjustments are known, used to preadjust the data for differences that exist within contemporary groups.

Factors that have been found to influence racing performance include sex, age, handicap weight, track, year, season, class of race, post position, and pace of the race (Hintz and Van Vleck, 1978; Hintz, 1980; Tolley et al., 1983; Ojala et al., 1987). Some of these factors may not apply to Quarter Horse racing because of the nature of the
race. In the previous study (Buttram, 1987) sex, age and handicap weight were identified as possible factors for preadjusting the data. Variation was seen among tracks but the data were studied to evaluate whether contemporary groups should be defined within tracks in order to remove more environmental variation.

The purposes of this study were to estimate adjustment factors for sex, age and handicap weight and to determine how contemporary groups should be defined for genetically evaluating American Quarter Horses for racing performance.

Materials and Methods

Data

Over one million individual racing records representing five racing distances were obtained from the American Quarter Horse Association. Records from 201, 320, 366, 402 and 796 m races made up data sets one (DS1), two (DS2), three (DS3), four (DS4) and five (DS5), respectively. Data were edited by removing from each data set all records with times greater than three standard deviations above the mean or less than the official world record time for that distance. Records of horses more than ten years old also were deleted. A further description of the original data was given by Buttram (1987).
**Adjustment factors**

Fixed effects of sex, age and handicap weight were estimated for each distance to obtain correction factors for preadjusting racing times. Marginal means were computed as the average of the subclass means for each sex at two, three and four years of age in DS1 through DS4 and at three, four and five years of age in DS5. Additive sex adjustments were calculated as the average difference between the marginal means of stallions and either mares or geldings. Multiplicative sex adjustment factors were calculated by dividing the marginal mean of the stallions by that of either mares or geldings. In addition, sex adjustment factors were calculated from two-year-old records only. Both additive and multiplicative adjustments were calculated as above except that means of two-year-old horses were used instead of marginal means.

Age and handicap weight effects were estimated using a linear model that included horses and ages as fixed effects and weight was treated as a covariate. Horse equations were absorbed into the other equations and the method of least squares was used to estimate age and handicap weight effects for each sex. A restriction was imposed so that the age effect of two-year-olds was zero and the age effects were plotted. Quadratic regression lines were fit through the age effects for males (stallions and geldings) and females
at each distance. These quadratic equations were used to predict separate age adjustments for males and females. Additive age adjustments were expressed relative to a four-year-old equivalent.

**Contemporary group effects**

The following completely nested model was used to determine the relative importance of sources of variation that could be included in contemporary group effects:

\[ Y_{ijklm} = t_i + g_{ij} + d_{ijk} + r_{ijkl} + e_{ijklm} \]

where:

- \( Y_{ijklm} \) = the finish time of the \( m \)th horse in the \( l \)th race on the \( k \)th day of the \( j \)th year at the \( i \)th track,
- \( t_i \) = mean common to all observations,
- \( g_{ij} \) = effect of the \( j \)th year at the \( i \)th track,
- \( d_{ijk} \) = effect of the \( k \)th day of the \( j \)th year at the \( i \)th track,
- \( r_{ijkl} \) = effect of the \( l \)th race on the \( k \)th day of the \( j \)th year at the \( i \)th track and
- \( e_{ijklm} \) = random residual error effect.

The nested procedure of SAS (1985) was used to calculate mean squares, expectations of mean squares and variance components for the sources of variation in the
model. The amount of variation accounted for by tracks, years within tracks, dates within years and tracks and individual races within dates, years and tracks was expressed as a percentage of the total variation.

Results and Discussion

Data

The data used in this study are summarized in Table 1. The percentage of the original data edited because of extreme values ranged from .8 to 1.2% depending on the data set. Records of horses more than ten years old were deleted because numbers were too small to estimate age effects for them. The percentage of records that were deleted because of age ranged from .11 to .29% in DS1 through DS4 and was about .8% in DS5. The distributions of the edited data sets were essentially the same as those described in the previous study (Buttram, 1987) except the upper tails of the distributions were removed. Each of the first few moments decreased with the removal of extreme values except in the 796 m race where skewness increased slightly. Kurtosis was much less in the edited data than in the original data sets.
Table 1. Summary of racing times used to estimate fixed effects by distance

<table>
<thead>
<tr>
<th>Item</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. Obs.</td>
<td>16,291</td>
<td>554,692</td>
<td>364,842</td>
<td>91,633</td>
<td>92,968</td>
</tr>
<tr>
<td>Mean, s</td>
<td>12.77</td>
<td>18.72</td>
<td>21.02</td>
<td>22.95</td>
<td>47.61</td>
</tr>
<tr>
<td>Std. dev., s</td>
<td>.371</td>
<td>.499</td>
<td>.532</td>
<td>.599</td>
<td>1.198</td>
</tr>
<tr>
<td>Minimum, s</td>
<td>11.64</td>
<td>17.21</td>
<td>19.18</td>
<td>21.02</td>
<td>44.30</td>
</tr>
<tr>
<td>Maximum, s</td>
<td>14.02</td>
<td>20.40</td>
<td>22.79</td>
<td>24.90</td>
<td>51.52</td>
</tr>
<tr>
<td>Skewness</td>
<td>.60</td>
<td>.63</td>
<td>.51</td>
<td>.41</td>
<td>.60</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>.30</td>
<td>.24</td>
<td>.13</td>
<td>.03</td>
<td>.07</td>
</tr>
</tbody>
</table>
Adjustment factors

Because of possible selection bias (Buttram, 1987), sex effects were estimated from young horses. The difference in racing time between stallions and mares was consistent across age groups (Table 2) indicating that the difference in marginal means was a satisfactory estimate of the difference between the two sex effects. The difference between stallions and mares became more pronounced as racing distance increased. This was reflected in the additive sex adjustments (Table 3) which decreased at an increasing rate. Multiplicative sex adjustment factors also decreased and were similar for the two shorter distances and for the 366 and 402 m races.

Although sex adjustments were computed on young horses, the interaction between sex and age was a problem in estimating adjustment factors for geldings. Marginal means for stallions and geldings (Table 2) were similar for all distances but stallions were generally faster than geldings as two-year-olds and slower than geldings as four-year-olds. Therefore, sex adjustments for geldings calculated from these marginal means were considered to be of little value. If this interaction was caused by differential selection rates as proposed, the best adjustment for geldings should come from two-year-old records which were the least
Table 2. Means used for calculating sex adjustment factors

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>Age, years</th>
<th>Marginal Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>210</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stallions</td>
<td>12.882</td>
<td>12.674</td>
</tr>
<tr>
<td>Mares</td>
<td>12.897</td>
<td>12.715</td>
</tr>
<tr>
<td>Geldings</td>
<td>12.889</td>
<td>12.711</td>
</tr>
<tr>
<td>320</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stallions</td>
<td>18.798</td>
<td>18.660</td>
</tr>
<tr>
<td>Mares</td>
<td>18.853</td>
<td>18.709</td>
</tr>
<tr>
<td>Geldings</td>
<td>18.830</td>
<td>18.660</td>
</tr>
<tr>
<td>366</td>
<td></td>
<td></td>
</tr>
<tr>
<td>402</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stallions</td>
<td>22.848</td>
<td>22.878</td>
</tr>
<tr>
<td>Mares</td>
<td>22.942</td>
<td>22.977</td>
</tr>
<tr>
<td>Geldings</td>
<td>22.971</td>
<td>22.967</td>
</tr>
<tr>
<td>796</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stallions</td>
<td>47.915</td>
<td>47.602</td>
</tr>
<tr>
<td>Mares</td>
<td>48.227</td>
<td>47.897</td>
</tr>
<tr>
<td>Geldings</td>
<td>47.822</td>
<td>47.518</td>
</tr>
</tbody>
</table>

aAges are 3, 4, and 5 years, respectively.
Table 3. Additive and multiplicative sex adjustment factors by distance

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>Type of adjustment</th>
<th>Add.</th>
<th>Multi.</th>
<th>Add.</th>
<th>Multi.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mares</td>
<td>-.036</td>
<td>.9972</td>
<td>-.015</td>
<td>.9988</td>
</tr>
<tr>
<td></td>
<td>Geldings</td>
<td>-.024</td>
<td>.9981</td>
<td>-.007</td>
<td>.9995</td>
</tr>
<tr>
<td>201</td>
<td></td>
<td>-.056</td>
<td>.9970</td>
<td>-.060</td>
<td>.9968</td>
</tr>
<tr>
<td></td>
<td>Mares</td>
<td>+.005</td>
<td>1.0003</td>
<td>-.032</td>
<td>.9983</td>
</tr>
<tr>
<td></td>
<td>Geldings</td>
<td>-.028</td>
<td>.9987</td>
<td>-.067</td>
<td>.9964</td>
</tr>
<tr>
<td>366</td>
<td></td>
<td>-.089</td>
<td>.9958</td>
<td>-.092</td>
<td>.9956</td>
</tr>
<tr>
<td></td>
<td>Mares</td>
<td>-.028</td>
<td>.9987</td>
<td>-.067</td>
<td>.9964</td>
</tr>
<tr>
<td></td>
<td>Geldings</td>
<td>-.082</td>
<td>.9964</td>
<td>-.123</td>
<td>.9946</td>
</tr>
<tr>
<td>796</td>
<td></td>
<td>-.321</td>
<td>.9933</td>
<td>-.071</td>
<td>1.0015</td>
</tr>
</tbody>
</table>

aAdjusted to a stallion equivalent.
bCalculated from marginal means.
cCalculated from means of two-year-old horses.
selected. Therefore, geldings probably rank between stallions and mares for racing performance. This was not found in DS4, but two-year-olds that raced in 402 m races also were a highly selected group of horses.

Additive sex adjustments (Table 3) indicate the amount of time to be added to the racing time of either mares or geldings. A negative sign means those animals were slower than stallions and therefore time must be subtracted to adjust them to a stallion equivalent. Some adjustments for geldings, calculated from the marginal means, carried a positive sign meaning geldings were faster than stallions for those distances.

Records also may be adjusted to a stallion equivalent by multiplying finish time by the appropriate multiplicative adjustment factor. Either the additive or multiplicative adjustment results in the same mean, but use of the multiplicative adjustment also causes a change in variance. More specifically, the variance of the adjusted data is the product of the original variance times the square of the multiplicative adjustment factor. Since the means and variances of racing time for the three sexes were positively related, multiplicative sex adjustment factors should be used. Sex adjustments calculated from the marginal means or from the means of two-year-olds were similar for mares and
either can be used to adjust mares to a stallion equivalent. Adjustments computed from the means of two-year-olds should be used to adjust gelding records for the reasons described earlier.

To remove some of the effect of selection over time, individual horse effects were included in the model for estimating age and handicap weight effects. Some bias may have been incurred by treating horses as fixed instead of random but this was not considered as important as the bias caused by selection. Also, estimates of variance components were not available when age adjustments were computed.

Plots from the two large data sets (Figure 1; Appendix A, Figure A14) showed that differences between the age effects of young horses and four- or five-year-olds were larger in mares than in stallions and geldings. Also in a number of cases stallions and geldings reached their peak performance at four years of age while mares nearly always performed best at five years of age. Age effects for stallions and geldings were very similar for each age group and for each distance. This finding dispels credibility for a biological explanation of the fact that racing times for geldings improved with age relative to stallions and mares as described in the previous study (Buttram, 1987). That is, selection is more likely to have caused the interaction.
Figure 1. Actual (symbols only) and predicted (regression lines) age effects relative to two-year-olds from 320 m racing records by sex.
Had the interaction been caused by inherent biological differences between the sexes, one would have expected the age effects for geldings to be larger than those for stallions and mares in young horses.

Plots of age effects from the smaller data sets did not display the smooth curves found in DS2 and DS3 (Appendix A, Figures A15-A16). Estimates of age effects also were more variable for older horses because of fewer numbers and possibly more selection bias. Because of their inconsistency, age effects of nine- and ten-year-old horses were not used in computing prediction equations for the four shorter races. Also, since age effects for stallions and geldings were similar, only one curve was fit through these effects and the same equation was used to predict age adjustments for both stallions and geldings. R-squares for the regression equations ranged from .88 to .96 in DS2 and DS3 indicating that age adjustments from the prediction equations were very close to the age effects estimated from the linear model. Curves did not fit so well in the smaller data sets but the shape of the curve was consistent for the four shorter races.

Age adjustments (Table 4) for DS1 through DS4 showed that racing performance improved until age four or five and declined after six years of age. Mares tended to be fastest
Table 4. Additive age adjustments by sex and distance

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age^, yr</td>
<td>age</td>
<td>additive</td>
</tr>
<tr>
<td></td>
<td>effect^</td>
<td>adjustment^</td>
</tr>
<tr>
<td>201</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.162</td>
<td>.189</td>
</tr>
<tr>
<td>3</td>
<td>.010</td>
<td>.052</td>
</tr>
<tr>
<td>5</td>
<td>-.022</td>
<td>-.027</td>
</tr>
<tr>
<td>6</td>
<td>.001</td>
<td>-.028</td>
</tr>
<tr>
<td>7</td>
<td>.010</td>
<td>.003</td>
</tr>
<tr>
<td>8</td>
<td>.034</td>
<td>.047</td>
</tr>
<tr>
<td>320</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.127</td>
<td>.097</td>
</tr>
<tr>
<td>3</td>
<td>.019</td>
<td>.035</td>
</tr>
<tr>
<td>5</td>
<td>.009</td>
<td>-.009</td>
</tr>
<tr>
<td>6</td>
<td>.033</td>
<td>.008</td>
</tr>
<tr>
<td>7</td>
<td>.065</td>
<td>.050</td>
</tr>
<tr>
<td>8</td>
<td>.120</td>
<td>.009</td>
</tr>
<tr>
<td>366</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.073</td>
<td>.074</td>
</tr>
<tr>
<td>3</td>
<td>.009</td>
<td>.013</td>
</tr>
<tr>
<td>5</td>
<td>.031</td>
<td>.006</td>
</tr>
<tr>
<td>6</td>
<td>.049</td>
<td>.030</td>
</tr>
<tr>
<td>7</td>
<td>.095</td>
<td>.073</td>
</tr>
<tr>
<td>8</td>
<td>.138</td>
<td>.135</td>
</tr>
<tr>
<td>402</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.094</td>
<td>.066</td>
</tr>
<tr>
<td>3</td>
<td>.013</td>
<td>.021</td>
</tr>
<tr>
<td>5</td>
<td>.055</td>
<td>.005</td>
</tr>
<tr>
<td>6</td>
<td>.060</td>
<td>.034</td>
</tr>
<tr>
<td>7</td>
<td>.104</td>
<td>.089</td>
</tr>
<tr>
<td>8</td>
<td>.175</td>
<td>.170</td>
</tr>
<tr>
<td>796</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.020</td>
<td>.041</td>
</tr>
<tr>
<td>5</td>
<td>.083</td>
<td>.074</td>
</tr>
<tr>
<td>6</td>
<td>.166</td>
<td>.181</td>
</tr>
<tr>
<td>7</td>
<td>.416</td>
<td>.321</td>
</tr>
<tr>
<td>8</td>
<td>.550</td>
<td>.495</td>
</tr>
<tr>
<td>9</td>
<td>.701</td>
<td>.701</td>
</tr>
<tr>
<td>10</td>
<td>.966</td>
<td>.941</td>
</tr>
</tbody>
</table>

^Adjusted to a four-year-old equivalent.

bAge effects estimated from the linear model.

cAdjustments predicted by regression equations.
at five years of age and age effects for two- and three-year-olds were larger than in males which were usually fastest at four years of age. Adjustments for DS5 showed that horses of either sex were fastest at three years and slowed at an increasing rate through age ten.

Adjustments for handicap weight effects were in the form of regression coefficients (Table 5). Adjustments were interpreted as the number of seconds subtracted from the finish time of horses for each kilogram they carried above the average handicap weight. Adding time to horses which carried less than the average weight was not warranted. According to earlier observations horses that carried less weight did not have an advantage over horses that carried average weights. Selection again may have biased the estimates of handicap weight effects. In general, heavier weights were assigned to horses that were considered to be faster, and lighter weights were assigned to slower horses. Most regression coefficients, however, carried a positive sign meaning that time increased (horses became slower) as handicap weight increased. Mean weights were about 55 kg in each data set and standard deviations ranged from 1.1 to 1.3 kg. Given these small standard deviations and adjustments in the range of .01 to .015 s/kg it is doubtful that using
Table 5. Regression coefficients\textsuperscript{a} for adjusting racing times for differences in handicap weights by sex and distance

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>Stallion</th>
<th>Mare</th>
<th>Gelding</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
<td>-.0051</td>
<td>.0074</td>
<td>.0091</td>
</tr>
<tr>
<td>320</td>
<td>.0140</td>
<td>.0137</td>
<td>.0106</td>
</tr>
<tr>
<td>366</td>
<td>.0136</td>
<td>.0117</td>
<td>.0103</td>
</tr>
<tr>
<td>402</td>
<td>.0158</td>
<td>.0104</td>
<td>.0066</td>
</tr>
<tr>
<td>796</td>
<td>.0003</td>
<td>.0105</td>
<td>-.0018</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Expressed as change in time (s) for each kg increase in handicap weight above the mean.
handicap weight as an adjustment factor would have a significant effect on the outcome of a genetic evaluation.

Contemporary group effects

The environmental components included in this analysis played a larger role in longer races than in shorter races. This caused a decrease in the percentage of the total variance due to error variance as the length of the races increased. The magnitude of the error variance, however, remained higher in the longer races.

The percentages of the total variance (Table 6) attributable to tracks ranged from 10.6 to 31.8% and were generally larger than that due to other effects in the model. Differences in tracks could have been due to a number of environmental factors including climate, type of racing surface or track condition. Years within tracks accounted for a relatively constant and small amount (5.2 to 7.8%) of the total variance, but race days within years and tracks accounted for 7.4 to 19.5% of the variation. The amount of variation accountable to individual races within days, years and tracks was high (15.1 to 24.4%) but unlike the other sources of variation, tended to decrease as length of the race increased. It is doubtful that environmental racing conditions changed drastically from one race to another during the same day at the same track. It is more
Table 6. Percentages of the total variance accounted for by tracks, years, days and individual races by distance

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Distance, m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>201</td>
</tr>
<tr>
<td>Tracks</td>
<td>10.6</td>
</tr>
<tr>
<td>Years</td>
<td>7.0</td>
</tr>
<tr>
<td>Days</td>
<td>7.4</td>
</tr>
<tr>
<td>Races</td>
<td>24.4</td>
</tr>
<tr>
<td>Error</td>
<td>50.6</td>
</tr>
<tr>
<td>Total Variance, s²</td>
<td>.1374</td>
</tr>
</tbody>
</table>
likely that individual races were accounting for different
types of races and competition levels. For example, on any
given day, there may be at the same track, several classes
of races based on amount of money won, number of placings
and age of the horses. Much of the variation due to race
days and individual races may have been removed had daily
track conditions and race classes been included in the
model, however, this information was not available in the
original data.

Each of the factors in the hierarchical classification
described could be considered as contemporary groups in a
genetic evaluation. Removal of track effects is considered
essential and use of tracks as contemporary groups would
allow for a minimum number of fixed effects. Inclusion of
any of the other factors would increase the number of fixed
effects substantially. Contemporary groups defined as
individual races would require computation of the largest
number of fixed effects but would remove effects of all of
the factors in the hierarchy. In addition, age effects
would be partially accounted for by races since many races
include only horses of the same age. This was particularly
important for two-year-olds. The use of races as
contemporary groups may remove a portion of the genetic
variance in as much as race classes are related to the
genetic potential of horses. Also, since the average size of a race was about eight horses, contemporary groups defined in this way would make it impossible to edit the data based on sire groups and expect adequate numbers to remain for estimating race effects.

An important consideration for defining contemporary groups is whether or not comparisons can be made across the groups. The previous study showed that indeed comparisons can be made between both sires and nonparents across races when repeated records are used. In the absence of other known environmental factors, therefore, individual races should be used as contemporary groups when computationally feasible to reduce the error variance as much as possible.

References
Factors influencing best annual racing time in Finnish 
Analysis System Institute Inc., Cary, NC.
Heritability and repeatability of speed for 2- and 3- 
56:1294.
SECTION III. VARIANCE COMPONENT ESTIMATION

Abstract
An analysis of variance method (the prior method), Iterative Henderson's Simple Method (IHSM) and Tilde-hat Method (THM) were used to estimate heritability and repeatability for racing time in American Quarter Horses. Racing times were adjusted for age and sex, and individual races were defined as contemporary groups. A model including sires and horses within sires was used to represent the data since horses had repeated records. Both heritability and repeatability estimates were slightly higher using THM than when IHSM was used. Converged estimates of heritability ranged from .01 to .36 and from .025 to .38 using IHSM and THM, respectively. Repeatability estimates were between .20 and .42 using IHSM and between .348 and .51 using THM. In some cases these methods did not converge because of negative estimates. The prior method generally yielded higher heritability estimates and lower estimates of repeatability than either IHSM or THM. Heritability and repeatability estimates were data dependent and weighted means of estimates from the prior method were used to compare estimates from different data sources. Records of two-year-old horses yielded higher estimates of
heritability (.20 vs. .12) than records from older horses. The entire data set for each distance gave higher estimates of heritability (.25 vs. .07) than a subset of records from the largest track. Overall weighted means for estimates of heritability and repeatability using the prior method were .22 and .32.

Introduction

The American Quarter Horse Association maintains an extensive data base which could be used to genetically evaluate horses for racing performance. In order to perform a genetic evaluation using mixed-model procedures (Henderson, 1972), variances for any random effects in the model must be known or at least must be estimated. The most common methods for estimating variance components include Methods I, II and III (Henderson, 1953), MINQUE (Rao, 1971), maximum likelihood (Hartley and Rao, 1967) and restricted maximum likelihood (REML) (Patterson and Thompson, 1971). Brief descriptions of these methods and their properties are given by Henderson (1980), Schaeffer (1983), Henderson (1984), Skaar (1985) and VanRaden (1986). Restricted maximum likelihood has the most desirable statistical properties for most animal breeding problems but is severely limited by the computational requirements for large data
sets. In fact, none of the methods described above could be used in this study without discarding a large portion of the data or making undue assumptions. Harville (1977) suggested that possibly quadratic forms exist that can be used to approximate REML estimates but that are easier to compute. One such approximate method (HSM), proposed by Henderson (1980), is simple to compute and does not require unwarranted assumptions about the model. Henderson originally proposed HSM as a noniterative procedure. More recently the quadratic form used in HSM has been used in an iterative procedure (IHSM) by Hudson and Van Vleck (1982), Skaar (1985), Wilson et al. (1986) and Bertrand and Benyshek (1987). Other iterative procedures that have been derived to approximate REML are Schaeffer's (1986) method and the tilde-hat method (THM) (VanRaden and Jung, 1987).

A simple analysis of variance method, IHSM and THM were each used in this study to estimate variance components. The primary purpose of the study was to estimate heritability and repeatability of racing performance in Quarter Horses. A secondary objective was to make comparisons between the estimates obtained from the three procedures.
Materials and Methods

Data

The racing records used in this study were obtained from the American Quarter Horse Association and were the same records used to estimate sex and age adjustments in the previous paper (Buttram, 1987). Multiplicative sex adjustments calculated from the marginal means in the previous study were used to preadjust the records of mares to a stallion equivalent. Sex-adjusted times were adjusted to a four-year-old stallion equivalent using additive age adjustments. Data from each racing distance (201, 320, 366, 402 and 796 m) were considered as independent data sets. Within each distance, variance components were estimated for two-year-olds and for three- to ten-year-olds separately. In addition, the track containing the most records in each data set was chosen as a subset of the data from which variance component estimates for each age group also were computed. This was done to obtain data sets small enough that all three methods could be used.

Model

Variance components from either IHSM or THM were estimated assuming the following model:

\[ Y = X_0 + Z_{1s} + Z_{2w} + e, \]
where \( Y \) was a vector of preadjusted racing times, \( \varphi \) was a vector of contemporary group fixed effects, \( s \) was a vector of random sire effects, \( w \) was a vector of random horse effects within sires and \( e \) was a random residual error vector. Incidence matrices \( X, Z_1, \) and \( Z_2 \) assigned the appropriate effects to the vector of individual racing times. The vectors \( s, w \) and \( e \) were assumed to have zero means and variances \( I\sigma_s^2, I\sigma_w^2 \) and \( I\sigma_e^2 \), respectively. Covariances among \( s, w, \) and \( e \) were assumed to be zero. Contemporary groups were defined as races and the two are used interchangeably in this paper.

Absorption of race equations into sire equations and horse equations within sires resulted in the following reduced mixed model equations:

\[
\begin{bmatrix}
Z_1'MZ_1 + I k_1 & Z_1'MZ_2 \\
Z_2'MZ_1 & Z_2'MZ_2 + I k_2
\end{bmatrix}
\begin{bmatrix}
\hat{s} \\
\hat{w}
\end{bmatrix}
= \begin{bmatrix}
Z_1'MY \\
Z_2'MY
\end{bmatrix},
\]

where \( M = I - X(X'X)^{-1}X' \), \( k_1 = \sigma_e^2/\sigma_s^2 \) and \( k_2 = \sigma_e^2/\sigma_w^2 \). Solutions to these equations yield best linear unbiased predictors (BLUP) of sire and horse effects which are denoted as \( \hat{s} \) and \( \hat{w} \). Approximate solutions to the reduced equations were computed by dividing each right hand side element by the corresponding diagonal element of the
coefficient matrix. Approximate solutions were denoted as \( \hat{s} \) and \( \hat{w} \). In matrix notation the approximate solutions are

\[
\hat{s} = D_1^{-1}Z_1'MY
\]

and

\[
\hat{w} = D_2^{-1}Z_2'MY,
\]

where \( D_i \) is a diagonal matrix with diagonal elements equal to those of \((Z_i'MZ_i + I_k)\) in the reduced equations.

**Estimation procedures**

Any of the common procedures for estimating variance components involves equating the value of quadratic forms to their expectations. The basic difference between IHSM and THM is that the quadratic form for IHSM involves only the approximate solutions and the quadratic for THM involves both the approximate and actual (BLUP) solutions. The quadratic forms for IHSM and their expectations were

\[
E(\hat{s}'\hat{s}) = \text{tr}[D_1^{-2}Z_1'MZ_1Z_1'MZ_1\sigma^2_s + D_1^{-2}Z_1'MZ_2Z_2'MZ_1\sigma^2_w + D_1^{-2}Z_1'MZ_1\sigma^2_e]
\]
and

\[ E(\tilde{w}'\tilde{w}) = tr[D_2^{-2}Z_2'MZ_1'Z_1'MZ_2\sigma_s^2 + D_2^{-2}Z_2'MZ_2Z_2'MZ_2\sigma_w^2 + D_2^{-2}Z_2'MZ_2\sigma_e^2]. \]

The quadratic forms and expectations for THM (VanRaden and Jung, 1987), were

\[ E(s's) = tr[(I-C_{11}k_1)D_1^{-1}\sigma_e^2 + Z_1'MZ_1D_1^{-1}\sigma_s^2 - (I-C_{11}k_1)k_2D_1^{-1}\sigma_w^2] \]

and

\[ E(w'w) = tr[(I-C_{22}k_2)D_2^{-1}\sigma_e^2 + Z_2'MZ_2D_2^{-1}\sigma_w^2 - (I-C_{22}k_2)k_1D_2^{-1}\sigma_s^2], \]

where \( C_{ii} \) was a submatrix of the inverse of the coefficient matrix of the reduced equations. However, pseudo expectations (Schaeffer, 1986) for THM derived under the assumption that apriori values for variance ratios (\( k_1 \) and \( k_2 \)) were equal to the true values (\( \sigma_e^2/\sigma_s^2 \) and \( \sigma_e^2/\sigma_w^2 \)), reduced very simply to:

\[ E(\tilde{s}'\tilde{s}) = tr[Z_1'MZ_1D_1^{-1}\sigma_s^2] \]

and

\[ E(\tilde{w}'\tilde{w}) = tr[Z_2'MZ_2D_2^{-1}\sigma_w^2]. \]
Unbiased estimates of error variance for both IHSM and THM were obtained from the following quadratic form:

\[ \hat{\sigma}_e^2 = \frac{[Y'MY - \hat{s}'Z_1'MY - \hat{w}'Z_2'MY]}{[n - \text{rank}(X)]}, \]

where \( n \) was the total number of observations and \( \text{rank}(X) \) was equal to the number of races in each data set. This quadratic is equivalent to the more common form:

\[ \frac{[Y'Y - \hat{s}'X'Y - \hat{s}'Z_1'Y - \hat{w}'Z_2'Y]}{[n - \text{rank}(X)]}. \]

Both IHSM and THM were used iteratively. New estimates of variance components for sires and horses within sires were computed each round. This was done by equating the quadratic forms described above to their expectations for IHSM and to their pseudo expectations for THM. The use of pseudo expectations results in the same estimates of variance components, upon convergence, as when actual expectations are used. New estimates of error variance also were computed each round which required finding BLUP solutions. Because BLUP solutions were computed iteratively within each round of the two procedures, the overall iterations for IHSM and THM are referred to as primary iterations and the iterations to determine BLUP solutions
are secondary iterations. After each primary iteration, new variance ratios were computed and used as priors for the next round.

A third method used to estimate variance components was one based on the analysis of variance. First, preadjusted racing times were deviated from their contemporary group (race) means. A hierarchical model including sires, horses within sires and error was assumed and the nested procedure of SAS (1985) was used to calculate mean squares, expectations of mean squares and to estimate variance components from the deviated times. These estimates were used as initial priors for both IHSM and THM and the analysis of variance procedure is referred to as the "prior" method. Both first round and convergence estimates of variance components were reported for IHSM and THM. Relative comparisons were made among the estimates from IHSM, THM and the prior method.

Computational aspects

Data were originally sorted by race, fixed effects were absorbed and the reduced equations were set up directly as the records were read. Expectations were trivial to compute because they involved only diagonal matrices and submatrices of the coefficient matrix of the reduced equations. The most computationally demanding aspect was obtaining BLUP
solutions for sires and horses within sires. Iterative BLUP solutions were computed each round of the primary iteration using the successive overrelaxation algorithm found in ITPACK (Kincaid et al., 1982).

The criterion used to terminate primary iteration was the square root of the average squared difference in BLUP solutions from one round to the next. At convergence this criterion should be zero. Iteration was stopped when the criterion reached .001 for IHSM or .0002 for THM. Preliminary testing showed that variance component estimates changed very little after these convergence criteria were reached. The criterion was smaller for THM because the convergence rate was much slower than IHMS. The EMC algorithm of VanRaden and Freeman (1987) was used to speed up convergence of THM. This did not change either first round or convergence estimates but reduced the number of rounds of iteration by 25% in a preliminary trial.

**Heritability and repeatability**

Expectations of the variance components were

\[
E(\sigma_s^2) = \frac{1}{4}\sigma_A^2,
\]

\[
E(\sigma_w^2) = \frac{3}{4}\sigma_A^2 + \sigma_{PE}^2
\]

and

\[
E(\sigma_e^2) = \sigma_{TE}^2,
\]
where $\sigma_A^2$ was the additive genetic variance of racing performance, $\sigma_{PE}^2$ was the variance due to permanent environment effects and $\sigma_{TE}^2$ was the variance of temporary environment effects. Dominance and epistasis effects were ignored. Heritability and repeatability were calculated from the variance component estimates as follows:

$$h^2 = \frac{4\hat{\sigma}_S^2}{\left(\hat{\sigma}_S^2 + \hat{\sigma}_W^2 + \hat{\sigma}_e^2\right)}$$

and

$$r = \frac{\left(\hat{\sigma}_S^2 + \hat{\sigma}_W^2\right)}{\left(\hat{\sigma}_S^2 + \hat{\sigma}_W^2 + \hat{\sigma}_e^2\right)}.$$  

Heritability and repeatability estimates from the prior method for each data set were weighted by the number of observations in the data set and weighted means were calculated. Only prior estimates were used for comparing estimates from different data sources because estimates from the prior method could be computed for all data sets. Weighted means for heritability and repeatability were each calculated for two-year-olds and for three- to ten-year-olds from both the largest track and the entire data set by weighting the estimates from each of the five distances. Similarly, estimates from the four data sources within each distance were used to obtain weighted means for each distance.
Results and Discussion

The properties of IHSM and THM are not well defined. Henderson (1980) originally proposed HSM as a noniterative procedure which yielded unbiased, translation invariant estimates of variance components. These are the same as first round estimates of IHSM used in this study. Because pseudo expectations were derived for THM, first round estimates from this procedure are not unbiased unless priors are assumed to be true values. Hudson and Van Vleck (1982), using a sire model, found that neither HSM or IHSM differed greatly from Method III estimates. Bertrand and Benyshek (1987) reported that IHSM and REML gave similar (co)variance estimates from a sire and maternal grandsire model. Both THM and IHSM gave good approximations to REML estimates from field data for dairy health traits (VanRaden and Jung, 1987). They also showed that heritability estimates from both methods were biased downward substantially by selection in simulated data.

The prior method was similar to Method II except that race means, rather than least squares estimates of race effects, were used to adjust the data. Method II was not used because the number of levels of fixed effects in the model was too large to allow for computing the inverse necessary to calculate the expected increase in error
variance that resulted in adjusting the data. How much estimates from the prior method were biased by using race means is unknown, but estimates were expected to approximate those that would have been obtained for Method II. However, Method II does not at all account for selection and is known to approximate REML only in data sets that are balanced. The hierarchical nature of the model is completely ignored by IHSM because the quadratic forms used to estimate variance components do not consider any off-diagonal elements from the reduced equations. Off-diagonal elements are "half" accounted for by THM because the quadratic forms used include both BLUP solutions which require the use of off-diagonal elements and the approximate solutions used in IHSM. In one sense, the prior method was considered "better" than the other methods because it accounted for the hierarchical structure of the data. The prior method, like Method II, is based on the analysis of variance which is a very powerful tool for partitioning variance among different sources. This method accounted for the fact that horses were nested within sires so that all available information was used to partition the variance between sires, horses within sires and error.

Since no REML estimates were obtained, methods were not compared for superiority or inferiority of their estimates.
Even the smallest data sets did not allow for REML estimates because of the number of levels of each factor in the model. In addition, maximum likelihood estimates were not computed because of the bias that would have been caused by the number of fixed effects in the model.

Both iterative methods were well behaved and converged in every case except when negative estimates were obtained. Preliminary results showed that the prior used did not affect convergence values. The two procedures required the same computing time per round but IHSM usually required four to ten rounds for convergence and THM required 14 to 30 rounds or more. Algorithms that use pseudo expectations often converge more slowly than those that use actual expectations (VanRaden and Jung, 1987). When the number of nonzero elements in the reduced equations was greater than 800,000, the cost of obtaining BLUP solutions (about 99% of the computing costs) was prohibitive because solutions could not be computed in core. For this reason, the two iterative methods were not used to estimate heritability and repeatability for some of the larger data sets. The prior method was easiest to compute and could be used on all data sets.

The number of records, sires and horses along with the within-race standard deviation is given (Table 1) for each
Table 1. Summary of data sets used for estimating variance components by distance

<table>
<thead>
<tr>
<th>Data set</th>
<th>No. Records</th>
<th>No. Sires</th>
<th>No. Horses</th>
<th>Std. Dev., s</th>
</tr>
</thead>
<tbody>
<tr>
<td>201 m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>661</td>
<td>297</td>
<td>550</td>
<td>.0896</td>
</tr>
<tr>
<td>TRACK3</td>
<td>1,071</td>
<td>376</td>
<td>764</td>
<td>.0706</td>
</tr>
<tr>
<td>ALL2</td>
<td>7,169</td>
<td>2,425</td>
<td>5,673</td>
<td>.0832</td>
</tr>
<tr>
<td>ALL3</td>
<td>9,122</td>
<td>2,687</td>
<td>6,286</td>
<td>.0580</td>
</tr>
<tr>
<td>320 m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>28,385</td>
<td>1,453</td>
<td>9,956</td>
<td>.0587</td>
</tr>
<tr>
<td>TRACK3</td>
<td>56,521</td>
<td>1,797</td>
<td>11,311</td>
<td>.0374</td>
</tr>
<tr>
<td>ALL2</td>
<td>223,830</td>
<td>11,299</td>
<td>87,356</td>
<td>.1294</td>
</tr>
<tr>
<td>ALL3</td>
<td>331,461</td>
<td>11,975</td>
<td>79,028</td>
<td>.0839</td>
</tr>
<tr>
<td>366 m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>9,648</td>
<td>1,081</td>
<td>5,696</td>
<td>.0613</td>
</tr>
<tr>
<td>TRACK3</td>
<td>34,785</td>
<td>1,634</td>
<td>9,853</td>
<td>.0440</td>
</tr>
<tr>
<td>ALL2</td>
<td>92,823</td>
<td>7,378</td>
<td>47,480</td>
<td>.1377</td>
</tr>
<tr>
<td>ALL3</td>
<td>272,452</td>
<td>10,785</td>
<td>69,703</td>
<td>.0974</td>
</tr>
<tr>
<td>402 m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>5,340</td>
<td>627</td>
<td>3,862</td>
<td>.1555</td>
</tr>
<tr>
<td>TRACK3</td>
<td>6,791</td>
<td>695</td>
<td>3,073</td>
<td>.0919</td>
</tr>
<tr>
<td>ALL2</td>
<td>13,175</td>
<td>2,124</td>
<td>9,465</td>
<td>.1643</td>
</tr>
<tr>
<td>ALL3</td>
<td>78,461</td>
<td>6,838</td>
<td>32,735</td>
<td>.1114</td>
</tr>
<tr>
<td>796 m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK3</td>
<td>16,082</td>
<td>735</td>
<td>1,770</td>
<td>.2228</td>
</tr>
<tr>
<td>ALL3</td>
<td>93,273</td>
<td>5,731</td>
<td>15,730</td>
<td>.4609</td>
</tr>
</tbody>
</table>

\(^a\)Within race standard deviation.

\(^b\)Includes records of two-year-olds from the largest track.

\(^c\)Includes records of three- to ten-year-olds from the largest track.

\(^d\)Includes all records of two-year-olds.

\(^e\)Includes all records of three- to ten-year-olds.
First round estimates of heritability (Table 2) using IHSM were extremely variable (-.29 to .36) but repeatability estimates (Table 3) were more consistent (.20 to .48). Because of the slow convergence rate of THM, first round estimates from this method did not differ significantly from the priors. Heritability and repeatability estimates were slightly higher using THM than when IHSM was used. Converged estimates of heritability ranged from .01 to .36 and .02 to .38 using IHSM and THM, respectively. Convergence was not reached in some cases because of negative values for heritability. Repeatability estimates were between .20 and .42 using IHSM and between .34 and .51 using THM. The prior method usually gave higher estimates of heritability and lower estimates of repeatability than either IHSM or THM.

Estimates of heritability (Table 2) and repeatability (Table 3) varied considerably with the data set used. Weighted means of estimates from the prior method were used to compare heritability and repeatability estimates across different data sources because priors could be computed for every data set. In order to compute weighted means for the different data sources, the assumption is made that racing performance is a single trait, regardless of the distance of
Table 2. Estimates of heritability for racing performance by data set from three methods

<table>
<thead>
<tr>
<th>Data set</th>
<th>Prior Method</th>
<th>IHSM</th>
<th>THM</th>
</tr>
</thead>
<tbody>
<tr>
<td>201 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2\textsuperscript{a}</td>
<td>.261</td>
<td>.003</td>
<td>_b</td>
</tr>
<tr>
<td>TRACK3\textsuperscript{c}</td>
<td>.201</td>
<td>-.130</td>
<td>_b</td>
</tr>
<tr>
<td>ALL2\textsuperscript{d}</td>
<td>.333</td>
<td>.226</td>
<td>.138</td>
</tr>
<tr>
<td>ALL3\textsuperscript{e}</td>
<td>.423</td>
<td>.313</td>
<td>.050</td>
</tr>
<tr>
<td>320 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>.083</td>
<td>.031</td>
<td>.014</td>
</tr>
<tr>
<td>TRACK3</td>
<td>.007</td>
<td>.014</td>
<td>.007</td>
</tr>
<tr>
<td>ALL2</td>
<td>.266</td>
<td>.014</td>
<td>.007</td>
</tr>
<tr>
<td>ALL3</td>
<td>.220</td>
<td>.014</td>
<td>.007</td>
</tr>
<tr>
<td>366 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>.097</td>
<td>.024</td>
<td>.004</td>
</tr>
<tr>
<td>TRACK3</td>
<td>.002</td>
<td>.014</td>
<td>.007</td>
</tr>
<tr>
<td>ALL2</td>
<td>.292</td>
<td>.014</td>
<td>.007</td>
</tr>
<tr>
<td>ALL3</td>
<td>.217</td>
<td>.014</td>
<td>.007</td>
</tr>
<tr>
<td>402 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>.344</td>
<td>.286</td>
<td>.169</td>
</tr>
<tr>
<td>TRACK3</td>
<td>.097</td>
<td>-.010</td>
<td>_b</td>
</tr>
<tr>
<td>ALL2</td>
<td>.345</td>
<td>.364</td>
<td>.362</td>
</tr>
<tr>
<td>ALL3</td>
<td>.270</td>
<td>.014</td>
<td>.007</td>
</tr>
<tr>
<td>796 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK3</td>
<td>.069</td>
<td>-.293</td>
<td>_b</td>
</tr>
<tr>
<td>ALL3</td>
<td>.200</td>
<td>.014</td>
<td>.007</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Includes records of two-year-olds from the largest track.
\textsuperscript{b}Convergence estimates were not obtained because of negative estimates.
\textsuperscript{c}Includes records of three- to ten-year-olds from the largest track.
\textsuperscript{d}Includes all records of two-year-olds.
\textsuperscript{e}Includes all records of three- to ten-year-olds.
Table 3. Estimates of repeatability for racing performance by data set from three methods

<table>
<thead>
<tr>
<th>Data set</th>
<th>Prior Method</th>
<th>IHSM</th>
<th>THM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Round 1</td>
<td>Converge</td>
<td>Round 1</td>
</tr>
<tr>
<td>201 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2(^a)</td>
<td>.483</td>
<td>.479</td>
<td>- (^b)</td>
</tr>
<tr>
<td>TRACK3(^c)</td>
<td>.450</td>
<td>.449</td>
<td>- (^b)</td>
</tr>
<tr>
<td>ALL2(^d)</td>
<td>.414</td>
<td>.410</td>
<td>.422</td>
</tr>
<tr>
<td>ALL3(^e)</td>
<td>.397</td>
<td>.392</td>
<td>.421</td>
</tr>
<tr>
<td>320 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>.305</td>
<td>.322</td>
<td>.345</td>
</tr>
<tr>
<td>TRACK3</td>
<td>.212</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL2</td>
<td>.354</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL3</td>
<td>.306</td>
<td></td>
<td></td>
</tr>
<tr>
<td>366 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>.299</td>
<td>.299</td>
<td>.308</td>
</tr>
<tr>
<td>TRACK3</td>
<td>.223</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL2</td>
<td>.364</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL3</td>
<td>.306</td>
<td></td>
<td></td>
</tr>
<tr>
<td>402 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK2</td>
<td>.214</td>
<td>.196</td>
<td>.200</td>
</tr>
<tr>
<td>TRACK3</td>
<td>.363</td>
<td>.385</td>
<td>- (^b)</td>
</tr>
<tr>
<td>ALL2</td>
<td>.279</td>
<td>.270</td>
<td>.229</td>
</tr>
<tr>
<td>ALL3</td>
<td>.323</td>
<td></td>
<td></td>
</tr>
<tr>
<td>796 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRACK3</td>
<td>.253</td>
<td>.315</td>
<td>- (^b)</td>
</tr>
<tr>
<td>ALL3</td>
<td>.324</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Includes records of two-year-olds from the largest track.
\(^b\)Convergence estimates were not obtained because of negative estimates.
\(^c\)Includes records of three- to ten-year-olds from the largest track.
\(^d\)Includes all records of two-year-olds.
\(^e\)Includes all records of three- to ten-year-olds.
the race. In other words, if racing performance at the five distances were considered as five different traits, a genetic correlation of 1.0 is assumed. Further research is needed to validate this assumption.

Heritability estimates were much higher (.24 vs .05) when the entire data set for a particular distance was used than when records from a single track were used (Table 4). The weighted average for heritability using three- to ten-year-old records at a single track was .01 across the five distances. Two of these data sets (320 and 366 m) had prior estimates of heritability that were so low that IHSM and THM were not used and would surely have been negative. Repeatability estimates too, were higher (.32 vs .25) when obtained from the entire data sets than when obtained from the largest track. In addition, two-year-old records yielded higher estimates of both heritability (.25 vs .20) and repeatability (.34 vs .30) than records from older horses.

These same trends were seen in the two iterative methods. In fact, in the three cases where the entire data sets could be used (Table 2), heritability estimates from IHSM and THM upon convergence averaged .18 and .29. Sire variances were understandably lower at individual tracks because there was generally less variance (Table 1) than in
Table 4. Weighted means of heritability and repeatability estimates from the prior method by age and data source

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Age, yr.</th>
<th>2</th>
<th>3-10</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRACK&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heritability</td>
<td>.12</td>
<td>.01</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>Repeatability</td>
<td>.30</td>
<td>.23</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>ALL&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heritability</td>
<td>.27</td>
<td>.22</td>
<td>.24</td>
<td></td>
</tr>
<tr>
<td>Repeatability</td>
<td>.35</td>
<td>.31</td>
<td>.32</td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heritability</td>
<td>.25</td>
<td>.20</td>
<td>.22</td>
<td></td>
</tr>
<tr>
<td>Repeatability</td>
<td>.34</td>
<td>.30</td>
<td>.32</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Includes only records from the largest track for each distance.
<sup>b</sup>Includes all records for each distance.
the complete data sets. The same can be said for two-year-olds vs older horses. Heritability estimates for older horses also may have been biased downward by selection more than those for two-year-olds.

Higher heritability and repeatability estimates were found in 201 m records than in the other four distances (Table 5). This could have been caused by less selection in the data or by a sampling problem. The largest track in the 201 m data set contained only 1,732 records. The weighted mean for repeatability was .31 for each of the other distances and heritabilities ranged from .18 to .27. Among the data sets from which estimates for all three methods were obtained, the complete two-year-old data set for 201 m was considered to be the most reliable for estimating heritability. This data set contained a sample of sires from across the Quarter Horse racing industry and likely had less selection than any of the other data sets. Heritability estimates from this data set ranged from .14 to .33 depending upon the method.

The difference between heritability and repeatability is the proportion of the total variance attributable to permanent environment effects. Permanent environment effects on racing performance may include such things as early nutrition, injury, owner and particularly trainer.
Table 5. Weighted means of heritability and repeatability estimates from the prior method by distance

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distance, m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>201</td>
</tr>
<tr>
<td>Heritability</td>
<td>.37</td>
</tr>
<tr>
<td>Repeatability</td>
<td>.41</td>
</tr>
</tbody>
</table>
Permanent environment effects were smaller in the 201 and 402 m races than at other distances. There was also a smaller percentage of repeated records in these two data sets. The longest race showed the greatest permanent environment effect.

In a review, Hintz (1980) reported heritability of time and best time to be .15 and .23 in Thoroughbreds and .32 and .25 in Standardbred trotters. Using a sire model, Ojala and Van Vleck (1981) reported estimates of about .30 for heritability of either best time or average time in trotters, but estimates were higher for older age groups. Repeatability was estimated from a different model including fixed ages and random horses. Estimates were about .70 for either trait. Repeatability was estimated to be .43 by Hintz and Van Vleck (1978) using Method II. Tolley et al. (1983) obtained Method III estimates from a sire and repeated records model. Estimates of repeatability and heritability of racing time were .44 and .29, respectively.

Overall weighted means for heritability and repeatability in this study using the prior method were .22 and .32. Previous estimates of heritability for racing performance in Thoroughbreds and Standardbred trotters are similar to the estimates obtained by the prior method using Quarter Horse racing data. Repeatability estimates
previously reported from other breeds are generally higher than those found in this study. Possibly permanent environment effects are not as important in short races characteristic of Quarter Horse racing as in the longer Thoroughbred and Standardbred races.

References


This thesis presents an extensive study of racing performance in American Quarter Horses as measured by finish time. The goal of this project is to genetically evaluate Quarter Horses for racing performance using mixed-model methodology. This study is summarized in three parts, each of which pertains to a different aspect of this problem.

First a broad description of Quarter Horse racing data is presented including a discussion of distributions, data structure (in terms of numbers of repeated records and progeny per sire), means for age and sex classifications and possible effects of selection. In addition, numerous tables and figures are listed in Appendix A which summarize various other aspects of the data not formally presented.

The second section deals with fixed effects. Adjustments are calculated for sex, age and handicap weight. The relative importance of tracks, years, race days and individual races is examined and contemporary groups are defined. The final aspect of the study involves estimating variance components for sires and horses within sires. Three variance component estimation procedures are used and estimates are given for heritability and repeatability of racing time.
The next step is to perform the genetic evaluation. Mixed-model methodology has been developed for estimating breeding values for racing performance that uses a reduced animal model and accounts for repeated records. Results from this study should serve to enhance such an evaluation and provide a basis for further research in this area.
REFERENCES


Essary, D. 1983. Quarter Horse Racing. (booklet). Published by the American Quarter Horse Association, Amarillo, TX.


Tolley, E. A., D. R. Notter and T. J. Marlowe. 1983. Heritability and repeatability of speed for 2- and 3-


ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Richard Willham for allowing me the opportunity to study animal breeding at Iowa State University. I deeply appreciate the time and effort he has invested in my graduate career and highly value his advice and friendship.

Special thanks are given to Dr. Doyle Wilson for his help in preparing this dissertation and for his efforts on behalf of my graduation goals. The comments and suggestions of Drs. Rothschild, Berger, Freeman and Cox are appreciated. Thanks also are extended to the American Quarter Horse Association for a substantial grant that supported this research.

The interactions with fellow graduate students were both academically and personally fulfilling. The help offered by Dr. Paul VanRaden, Keith Boldman and Susan Durham was especially appreciated. Most of all, I am thankful for the support and encouragement of my wife, Susan.
APPENDIX A.

TABLES AND FIGURES
Table A1. Number of racing records by year and distance from 1971-1986

<table>
<thead>
<tr>
<th>Year</th>
<th>Distance, m</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>201</td>
<td>320</td>
<td>366</td>
<td>402</td>
<td>796</td>
</tr>
<tr>
<td>1971</td>
<td>397</td>
<td>24,696</td>
<td>15,168</td>
<td>3,222</td>
<td>4,652</td>
</tr>
<tr>
<td>1972</td>
<td>359</td>
<td>24,595</td>
<td>15,900</td>
<td>3,215</td>
<td>5,104</td>
</tr>
<tr>
<td>1973</td>
<td>537</td>
<td>27,423</td>
<td>18,612</td>
<td>4,018</td>
<td>5,647</td>
</tr>
<tr>
<td>1974</td>
<td>795</td>
<td>28,820</td>
<td>19,703</td>
<td>4,509</td>
<td>5,677</td>
</tr>
<tr>
<td>1975</td>
<td>855</td>
<td>20,833</td>
<td>30,172</td>
<td>5,176</td>
<td>6,294</td>
</tr>
<tr>
<td>1976</td>
<td>827</td>
<td>35,324</td>
<td>22,149</td>
<td>5,912</td>
<td>6,565</td>
</tr>
<tr>
<td>1977</td>
<td>977</td>
<td>38,181</td>
<td>25,362</td>
<td>6,358</td>
<td>6,087</td>
</tr>
<tr>
<td>1978</td>
<td>1,461</td>
<td>34,161</td>
<td>25,439</td>
<td>6,366</td>
<td>5,961</td>
</tr>
<tr>
<td>1979</td>
<td>973</td>
<td>35,840</td>
<td>24,244</td>
<td>6,699</td>
<td>5,788</td>
</tr>
<tr>
<td>1980</td>
<td>1,215</td>
<td>36,841</td>
<td>25,257</td>
<td>6,476</td>
<td>5,662</td>
</tr>
<tr>
<td>1981</td>
<td>1,489</td>
<td>37,376</td>
<td>26,341</td>
<td>6,221</td>
<td>5,489</td>
</tr>
<tr>
<td>1982</td>
<td>1,865</td>
<td>39,760</td>
<td>26,287</td>
<td>6,451</td>
<td>5,653</td>
</tr>
<tr>
<td>1983</td>
<td>2,014</td>
<td>45,609</td>
<td>28,718</td>
<td>8,144</td>
<td>6,213</td>
</tr>
<tr>
<td>1984</td>
<td>1,254</td>
<td>45,853</td>
<td>28,788</td>
<td>8,077</td>
<td>7,090</td>
</tr>
<tr>
<td>1985</td>
<td>922</td>
<td>41,833</td>
<td>26,700</td>
<td>7,175</td>
<td>7,009</td>
</tr>
<tr>
<td>1986(^a)</td>
<td>602</td>
<td>35,254</td>
<td>20,245</td>
<td>4,646</td>
<td>5,134</td>
</tr>
</tbody>
</table>

\(^a\)Includes records through Sept. 1, 1986.
Table A2. Means for racing time by year and distance from 1971-1986

<table>
<thead>
<tr>
<th>Year</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>12.80</td>
<td>18.91</td>
<td>21.16</td>
<td>23.11</td>
<td>47.49</td>
</tr>
<tr>
<td>1972</td>
<td>12.80</td>
<td>18.84</td>
<td>21.12</td>
<td>23.06</td>
<td>47.66</td>
</tr>
<tr>
<td>1973</td>
<td>12.93</td>
<td>18.81</td>
<td>21.07</td>
<td>23.03</td>
<td>47.42</td>
</tr>
<tr>
<td>1974</td>
<td>12.88</td>
<td>18.80</td>
<td>21.10</td>
<td>23.03</td>
<td>47.52</td>
</tr>
<tr>
<td>1975</td>
<td>12.76</td>
<td>18.77</td>
<td>21.08</td>
<td>23.03</td>
<td>47.52</td>
</tr>
<tr>
<td>1976</td>
<td>12.79</td>
<td>18.79</td>
<td>21.08</td>
<td>23.00</td>
<td>47.56</td>
</tr>
<tr>
<td>1977</td>
<td>12.88</td>
<td>18.82</td>
<td>21.12</td>
<td>23.07</td>
<td>47.85</td>
</tr>
<tr>
<td>1978</td>
<td>12.85</td>
<td>18.76</td>
<td>21.09</td>
<td>23.06</td>
<td>47.92</td>
</tr>
<tr>
<td>1979</td>
<td>12.77</td>
<td>18.75</td>
<td>21.08</td>
<td>23.09</td>
<td>47.95</td>
</tr>
<tr>
<td>1980</td>
<td>12.80</td>
<td>18.74</td>
<td>21.07</td>
<td>23.03</td>
<td>47.66</td>
</tr>
<tr>
<td>1981</td>
<td>12.78</td>
<td>18.72</td>
<td>21.01</td>
<td>22.98</td>
<td>47.57</td>
</tr>
<tr>
<td>1982</td>
<td>12.80</td>
<td>18.72</td>
<td>21.02</td>
<td>22.92</td>
<td>47.53</td>
</tr>
<tr>
<td>1983</td>
<td>12.75</td>
<td>18.69</td>
<td>21.00</td>
<td>22.86</td>
<td>47.54</td>
</tr>
<tr>
<td>1984</td>
<td>12.72</td>
<td>18.67</td>
<td>21.01</td>
<td>22.88</td>
<td>47.65</td>
</tr>
<tr>
<td>1985</td>
<td>12.71</td>
<td>18.63</td>
<td>20.94</td>
<td>22.81</td>
<td>47.71</td>
</tr>
<tr>
<td>1986a</td>
<td>12.67</td>
<td>18.67</td>
<td>21.00</td>
<td>22.79</td>
<td>47.79</td>
</tr>
</tbody>
</table>

*Includes records through Sept. 1, 1986.*
Table A3. Summary of racing time by month and distance

<table>
<thead>
<tr>
<th>Distance, m</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Month</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>680</td>
<td>18,568</td>
<td>16,876</td>
<td>3,728</td>
<td>5,868</td>
</tr>
<tr>
<td>Mean</td>
<td>12.77</td>
<td>18.65</td>
<td>21.03</td>
<td>23.08</td>
<td>47.62</td>
</tr>
<tr>
<td>Feb.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>699</td>
<td>16,893</td>
<td>11,854</td>
<td>3,076</td>
<td>4,857</td>
</tr>
<tr>
<td>Mean</td>
<td>12.86</td>
<td>18.69</td>
<td>21.09</td>
<td>23.13</td>
<td>47.77</td>
</tr>
<tr>
<td>Mar.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>890</td>
<td>32,552</td>
<td>19,404</td>
<td>5,131</td>
<td>7,295</td>
</tr>
<tr>
<td>Mean</td>
<td>12.96</td>
<td>18.71</td>
<td>21.02</td>
<td>23.07</td>
<td>47.63</td>
</tr>
<tr>
<td>Apr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,244</td>
<td>47,852</td>
<td>23,012</td>
<td>5,520</td>
<td>7,510</td>
</tr>
<tr>
<td>Mean</td>
<td>12.81</td>
<td>18.79</td>
<td>21.04</td>
<td>23.01</td>
<td>47.65</td>
</tr>
<tr>
<td>May</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,938</td>
<td>79,579</td>
<td>36,352</td>
<td>7,157</td>
<td>8,720</td>
</tr>
<tr>
<td>Mean</td>
<td>12.81</td>
<td>18.80</td>
<td>21.05</td>
<td>22.98</td>
<td>47.67</td>
</tr>
<tr>
<td>Jun.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>827</td>
<td>85,788</td>
<td>52,357</td>
<td>12,327</td>
<td>11,138</td>
</tr>
<tr>
<td>Mean</td>
<td>12.77</td>
<td>18.75</td>
<td>21.05</td>
<td>22.91</td>
<td>47.68</td>
</tr>
<tr>
<td>Jul.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,617</td>
<td>85,434</td>
<td>64,939</td>
<td>12,553</td>
<td>13,189</td>
</tr>
<tr>
<td>Mean</td>
<td>12.82</td>
<td>18.76</td>
<td>21.03</td>
<td>22.94</td>
<td>47.62</td>
</tr>
<tr>
<td>Aug.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>3,545</td>
<td>65,668</td>
<td>46,986</td>
<td>16,758</td>
<td>11,975</td>
</tr>
<tr>
<td>Mean</td>
<td>12.73</td>
<td>18.73</td>
<td>21.03</td>
<td>22.88</td>
<td>47.60</td>
</tr>
<tr>
<td>Sept.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,907</td>
<td>47,296</td>
<td>32,913</td>
<td>9,757</td>
<td>7,572</td>
</tr>
<tr>
<td>Mean</td>
<td>12.84</td>
<td>18.79</td>
<td>21.14</td>
<td>22.97</td>
<td>47.78</td>
</tr>
<tr>
<td>Oct.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,408</td>
<td>33,845</td>
<td>27,467</td>
<td>7,896</td>
<td>6,022</td>
</tr>
<tr>
<td>Mean</td>
<td>12.77</td>
<td>18.79</td>
<td>21.11</td>
<td>23.10</td>
<td>47.67</td>
</tr>
<tr>
<td>Nov.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>577</td>
<td>28,084</td>
<td>19,865</td>
<td>5,288</td>
<td>5,105</td>
</tr>
<tr>
<td>Mean</td>
<td>12.72</td>
<td>18.67</td>
<td>21.10</td>
<td>23.03</td>
<td>47.78</td>
</tr>
<tr>
<td>Dec.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>208</td>
<td>20,179</td>
<td>16,821</td>
<td>3,473</td>
<td>4,777</td>
</tr>
<tr>
<td>Mean</td>
<td>12.61</td>
<td>18.51</td>
<td>20.89</td>
<td>22.90</td>
<td>47.38</td>
</tr>
</tbody>
</table>
Table A4. Frequency distributions for number of horses per race by distance

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. horses/race</td>
<td>1</td>
<td>No races</td>
<td>6</td>
<td>55</td>
<td>28</td>
</tr>
<tr>
<td>Pct. total</td>
<td>.2</td>
<td>.1</td>
<td>.1</td>
<td>.1</td>
<td>.1</td>
</tr>
<tr>
<td>2</td>
<td>No races</td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Pct. total</td>
<td>.1</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.2</td>
</tr>
<tr>
<td>3</td>
<td>No races</td>
<td>12</td>
<td>91</td>
<td>49</td>
<td>20</td>
</tr>
<tr>
<td>Pct. total</td>
<td>.5</td>
<td>.1</td>
<td>.1</td>
<td>.2</td>
<td>.6</td>
</tr>
<tr>
<td>4</td>
<td>No races</td>
<td>215</td>
<td>1,377</td>
<td>874</td>
<td>417</td>
</tr>
<tr>
<td>Pct. total</td>
<td>8.8</td>
<td>2.0</td>
<td>2.0</td>
<td>3.6</td>
<td>2.3</td>
</tr>
<tr>
<td>5</td>
<td>No races</td>
<td>351</td>
<td>3,186</td>
<td>1,789</td>
<td>642</td>
</tr>
<tr>
<td>Pct. total</td>
<td>14.3</td>
<td>4.6</td>
<td>4.0</td>
<td>5.6</td>
<td>5.5</td>
</tr>
<tr>
<td>6</td>
<td>No races</td>
<td>456</td>
<td>6,553</td>
<td>3,910</td>
<td>1,064</td>
</tr>
<tr>
<td>Pct. total</td>
<td>18.6</td>
<td>9.6</td>
<td>8.7</td>
<td>9.3</td>
<td>11.5</td>
</tr>
<tr>
<td>7</td>
<td>No races</td>
<td>534</td>
<td>10,405</td>
<td>6,208</td>
<td>1,606</td>
</tr>
<tr>
<td>Pct. total</td>
<td>21.8</td>
<td>15.2</td>
<td>14.0</td>
<td>14.1</td>
<td>19.3</td>
</tr>
<tr>
<td>8</td>
<td>No races</td>
<td>610</td>
<td>14,269</td>
<td>8,646</td>
<td>2,210</td>
</tr>
<tr>
<td>Pct. total</td>
<td>24.9</td>
<td>20.8</td>
<td>19.5</td>
<td>19.3</td>
<td>30.2</td>
</tr>
<tr>
<td>9</td>
<td>No races</td>
<td>129</td>
<td>13,600</td>
<td>9,136</td>
<td>2,333</td>
</tr>
<tr>
<td>Pct. total</td>
<td>5.3</td>
<td>19.8</td>
<td>20.6</td>
<td>20.4</td>
<td>12.4</td>
</tr>
<tr>
<td>10</td>
<td>No races</td>
<td>133</td>
<td>18,391</td>
<td>13,319</td>
<td>3,000</td>
</tr>
<tr>
<td>Pct. total</td>
<td>5.4</td>
<td>26.8</td>
<td>30.0</td>
<td>26.3</td>
<td>17.9</td>
</tr>
<tr>
<td>11</td>
<td>No races</td>
<td>2</td>
<td>399</td>
<td>242</td>
<td>68</td>
</tr>
<tr>
<td>Pct. total</td>
<td>.1</td>
<td>.6</td>
<td>.5</td>
<td>.6</td>
<td>.0</td>
</tr>
<tr>
<td>12</td>
<td>No races</td>
<td>1</td>
<td>246</td>
<td>172</td>
<td>55</td>
</tr>
<tr>
<td>Pct. total</td>
<td>.0</td>
<td>.4</td>
<td>.4</td>
<td>.5</td>
<td>.0</td>
</tr>
</tbody>
</table>
Table A5. Summary of racing times at the five largest tracks for each of five racing distances

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>Track</th>
<th>No.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Pct. total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Obs.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>201</td>
<td>CAS</td>
<td>1,754</td>
<td>12.73</td>
<td>.366</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>SAL</td>
<td>1,726</td>
<td>12.81</td>
<td>.372</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>AKD</td>
<td>1,380</td>
<td>12.73</td>
<td>.353</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td>1,097</td>
<td>12.82</td>
<td>.408</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>LBD</td>
<td>973</td>
<td>12.93</td>
<td>.404</td>
<td>5.9</td>
</tr>
<tr>
<td>320</td>
<td>LA</td>
<td>84,935</td>
<td>18.37</td>
<td>.344</td>
<td>15.1</td>
</tr>
<tr>
<td></td>
<td>BRD</td>
<td>31,351</td>
<td>18.67</td>
<td>.472</td>
<td>5.6</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>30,367</td>
<td>18.62</td>
<td>.368</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>DED</td>
<td>28,575</td>
<td>18.86</td>
<td>.382</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>PPK</td>
<td>21,279</td>
<td>18.76</td>
<td>.466</td>
<td>3.8</td>
</tr>
<tr>
<td>366</td>
<td>LA</td>
<td>44,464</td>
<td>20.70</td>
<td>.394</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>DED</td>
<td>33,394</td>
<td>21.27</td>
<td>.416</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td>RUI</td>
<td>26,272</td>
<td>20.86</td>
<td>.481</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>SUN</td>
<td>14,422</td>
<td>20.61</td>
<td>.453</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>14,251</td>
<td>20.88</td>
<td>.371</td>
<td>3.9</td>
</tr>
<tr>
<td>402</td>
<td>RUI</td>
<td>11,540</td>
<td>22.61</td>
<td>.519</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>DED</td>
<td>10,207</td>
<td>23.19</td>
<td>.419</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>LA</td>
<td>6,504</td>
<td>22.43</td>
<td>.433</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>4,260</td>
<td>22.82</td>
<td>.407</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>LAM</td>
<td>3,742</td>
<td>22.60</td>
<td>.549</td>
<td>4.0</td>
</tr>
<tr>
<td>976</td>
<td>LA</td>
<td>16,111</td>
<td>46.95</td>
<td>.837</td>
<td>17.1</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>9,027</td>
<td>47.26</td>
<td>.817</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>SUN</td>
<td>9,020</td>
<td>47.10</td>
<td>1.011</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>RUI</td>
<td>7,083</td>
<td>47.96</td>
<td>1.178</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>DED</td>
<td>5,786</td>
<td>48.92</td>
<td>.924</td>
<td>6.2</td>
</tr>
</tbody>
</table>
Table A6. Frequency distributions for number of tracks per horse by distance

<table>
<thead>
<tr>
<th>No. tracks/horse</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 No horses</td>
<td>10,555</td>
<td>60,907</td>
<td>50,036</td>
<td>27,876</td>
<td>10,278</td>
</tr>
<tr>
<td>Pct. total</td>
<td>91.6</td>
<td>48.5</td>
<td>54.3</td>
<td>71.7</td>
<td>64.0</td>
</tr>
<tr>
<td>2 No horses</td>
<td>873</td>
<td>34,174</td>
<td>23,020</td>
<td>7,460</td>
<td>3,121</td>
</tr>
<tr>
<td>Pct. total</td>
<td>7.6</td>
<td>27.2</td>
<td>25.0</td>
<td>19.2</td>
<td>19.4</td>
</tr>
<tr>
<td>3 No horses</td>
<td>85</td>
<td>16,214</td>
<td>10,268</td>
<td>2,321</td>
<td>1,235</td>
</tr>
<tr>
<td>Pct. total</td>
<td>.7</td>
<td>12.9</td>
<td>11.1</td>
<td>6.0</td>
<td>7.7</td>
</tr>
<tr>
<td>4 No horses</td>
<td>13</td>
<td>7,625</td>
<td>4,762</td>
<td>743</td>
<td>596</td>
</tr>
<tr>
<td>Pct. total</td>
<td>.1</td>
<td>6.1</td>
<td>5.2</td>
<td>1.9</td>
<td>3.7</td>
</tr>
<tr>
<td>5+ No horses</td>
<td>2</td>
<td>6,639</td>
<td>4,059</td>
<td>470</td>
<td>817</td>
</tr>
<tr>
<td>Pct. total</td>
<td>.0</td>
<td>5.6</td>
<td>4.4</td>
<td>1.2</td>
<td>5.2</td>
</tr>
</tbody>
</table>
Table A7. Frequency distributions for number of tracks per sire by distance

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. tracks/sire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>No sires</td>
<td>2,522</td>
<td>5,371</td>
<td>4,922</td>
<td>3,472</td>
</tr>
<tr>
<td></td>
<td>Pct. total</td>
<td>63.0</td>
<td>34.4</td>
<td>38.5</td>
<td>46.6</td>
</tr>
<tr>
<td>2</td>
<td>No sires</td>
<td>730</td>
<td>2,819</td>
<td>2,244</td>
<td>1,281</td>
</tr>
<tr>
<td></td>
<td>Pct. total</td>
<td>18.2</td>
<td>18.0</td>
<td>17.6</td>
<td>17.2</td>
</tr>
<tr>
<td>3</td>
<td>No sires</td>
<td>294</td>
<td>1,764</td>
<td>1,316</td>
<td>757</td>
</tr>
<tr>
<td></td>
<td>Pct. total</td>
<td>7.3</td>
<td>11.3</td>
<td>10.3</td>
<td>10.2</td>
</tr>
<tr>
<td>4</td>
<td>No sires</td>
<td>153</td>
<td>1,137</td>
<td>803</td>
<td>446</td>
</tr>
<tr>
<td></td>
<td>Pct. total</td>
<td>3.8</td>
<td>7.3</td>
<td>6.3</td>
<td>6.0</td>
</tr>
<tr>
<td>5-9</td>
<td>No sires</td>
<td>226</td>
<td>2,381</td>
<td>1,869</td>
<td>914</td>
</tr>
<tr>
<td></td>
<td>Pct. total</td>
<td>5.7</td>
<td>15.3</td>
<td>14.6</td>
<td>12.2</td>
</tr>
<tr>
<td>10&gt;</td>
<td>No sires</td>
<td>78</td>
<td>2,159</td>
<td>1,621</td>
<td>585</td>
</tr>
<tr>
<td></td>
<td>Pct. total</td>
<td>1.9</td>
<td>13.8</td>
<td>12.7</td>
<td>7.8</td>
</tr>
</tbody>
</table>
Table A8. Handicap weight means and standard deviations by distance

<table>
<thead>
<tr>
<th>Distance, m</th>
<th>201</th>
<th>320</th>
<th>366</th>
<th>402</th>
<th>796</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handicap weight&lt;sup&gt;a&lt;/sup&gt;</td>
<td>55.1</td>
<td>54.9</td>
<td>54.8</td>
<td>54.9</td>
<td>54.5</td>
</tr>
<tr>
<td>Mean, kg</td>
<td>55.1</td>
<td>54.9</td>
<td>54.8</td>
<td>54.9</td>
<td>54.5</td>
</tr>
<tr>
<td>Std. dev., kg</td>
<td>1.29</td>
<td>1.10</td>
<td>1.10</td>
<td>1.17</td>
<td>1.17</td>
</tr>
</tbody>
</table>

<sup>a</sup>Includes total weight carried by the horse.
Figure A1. Frequency distribution of racing times from 201 m racing records
Figure A2. Frequency distribution of racing times from 366 m racing records
Figure A3. Frequency distribution of racing times from 402 m racing records
Figure A4. Frequency distribution of racing times from 796 m racing records
Figure A5. Relationship between handicap weight and mean racing time from 201 m racing data (data labels are the percentages of the data represented)
Figure A6. Relationship between handicap weight and mean racing time from 366 m racing data (data labels are the percentages of the data represented)
Figure A7. Relationship between handicap weight and mean racing time from 402 m racing data (data labels are the percentages of the data represented)
Figure A8. Relationship between handicap weight and mean racing time from 796 m racing data (data labels are the percentages of the data represented)
Figure A9. Relationship between age and mean racing time from 201 m racing records by sex
Figure A10. Relationship between age and mean racing time from 366 m racing records by sex.
Figure A11. Relationship between age and mean racing time from 402 m racing records by sex.
Figure A12. Relationship between age and mean racing time from 796 m racing records by sex.
Figure A13. Actual (symbols only) and predicted (regression lines) age effects relative to two-year-olds from 201 m racing records by sex.
Figure A14. Actual (symbols only) and predicted (regression lines) age effects relative to two-year-olds from 366 m racing records by sex.
Figure A15. Actual (symbols only) and predicted (regression lines) age effects relative to two-year-olds from 402 m racing records by sex.
Figure A16. Actual (symbols only) and predicted (regression lines) age effects relative to three-year-olds from 796 m racing records by sex.
APPENDIX B.
COMPUTER PROGRAM TO EDIT DATA,
ASSIGN EQUATION NUMBERS
AND SORT RECORDS
EDIT JOB U3825,SAM

THIS PROGRAM EDITS THE AQHA RACING DATA ACCORDING TO THE CRITERIA SET UP IN STEP 1 (LINES 60/62). THE PROGRAM CREATES EQUATION NUMBERS TO SIRES AND HORSES AND CREATES A NEW DATA FILE SORTED BY RACE AND SIRE WITHIN RACE. THE FORMAT OF THE NEW DATA SET IS AS FOLLOWS:

<table>
<thead>
<tr>
<th>ITEM</th>
<th>COL</th>
</tr>
</thead>
<tbody>
<tr>
<td>HORSE EQN</td>
<td>1-7</td>
</tr>
<tr>
<td>AGE</td>
<td>8-9</td>
</tr>
<tr>
<td>SEX</td>
<td>10</td>
</tr>
<tr>
<td>SIRE EQN</td>
<td>11-17</td>
</tr>
<tr>
<td>TRACK</td>
<td>18-20</td>
</tr>
<tr>
<td>RACE</td>
<td>21</td>
</tr>
<tr>
<td>DATE (YR MO DY)</td>
<td>22-27</td>
</tr>
<tr>
<td>TIME</td>
<td>28-31</td>
</tr>
</tbody>
</table>

S1 EXEC PLIXCLG
PLI.SYSIN DD *
AGE: PROC OPTIONS(MAIN) REORDER;
DCL SYSPRINT FILE PRINT;
DCL RECIN FILE INPUT RECORD
   ENV(TOTAL FB RECSIZE (46) BLKSIZE (32200));
DCL 1 DATA BASED (P),
   2 PAD1 CHAR(7),
   2 BIRTH_YR PIC '99',
   2 PAD2 CHAR(23),
   2 YEAR PIC '99',
   2 PAD3 CHAR(12);
DCL RECOUT FILE OUTPUT RECORD
   ENV(TOTAL FB RECSIZE (46) BLKSIZE (18998));
DCL (P) PTR;
DCL EOD BIT(1) INIT ('1'B);
ON ENDFILE (RECIN) EOD='0'B;
READ FILE (RECIN) SET (P);
DO WHILE (EOD);
BIRTH_YR=YEAR-BIRTH_YR;
WRITE FILE (RECOUT) FROM (DATA);
READ FILE (RECIN) SET (P);
END;
END AGE;
GO.RECIN DD DSN=FILE440,UNIT=(TAPE,DEFER),DISP=(OLD,KEEP),
   VOL=SER=AQHA3,LABEL=(4,SL)
GO.RECOUT DD DSN=&DATA,UNIT=SCRTCH,DISP=(NEW,PASS),
   SPACE=(TRK,(1500,150),RLSE),
   DCB=(RECFM=FB,LRECL=46,BLKSIZE=18998)
S2 EXEC SYMSORT,TRACKS=1500,REGION=1024K
//**
//*EDIT DATA AND OUTPUTS NEW FILE SORTED BY SIRE.
//**
//SORTIN DD DSN=&DATA,UNIT=SCRTCH,DISP=(OLD,DELETE)
//SORTOUT DD DSN=&DATA1,UNIT=SCRTCH,DISP=(NEW,PASS),
// SPACE=(TRK,(1000,100),RLSE),
// DCB=(RECFM=FB,LRECL=33,BLKSIZE=19008)
//SYSIN DD *
OMIT COND=(43,1,CH,EQ,'C'H',OR,
8,2,CH,LT,'C'03',OR,8,2,CH,GT,'C'10',OR,
39,4,CH,LT,'C'2102',OR,39,4,CH,GT,'C'2490')
INREC FIELDS=(C'2',1,10,C'1',11,7,29,14)
SORT FIELDS=(12,8,CH,A)
//S3 EXEC SYMSORT,TRACKS=500,REGION=1024K
//**
//*CREATES A UNIQUE LIST OF SIRES.
//**
//SORTIN DD DSN=&DATA1,UNIT=SCRTCH,DISP=(OLD,PASS)
//SORTOUT DD DSN=S<SIRES,UNIT=SCRTCH,DISP=(NEW,PASS),
// SPACE=(TRK,(200,20),RLSE),
// DCB=(RECFM=FB,LRECL=8,BLKSIZE=19016)
//SYSIN DD *
INREC FIELDS=(12,8)
SORT FIELDS=(1,8,CH,A)
SUM FIELDS=NONE
//S4 EXEC SYMSORT,TRACKS=500,REGION=1024K
//**
//*CREATES A UNIQUE LIST OF HORSES WITH INDIVIDUAL RECORDS.
//**
//SORTIN DD DSN=&DATA1,UNIT=SCRTCH,DISP=(OLD,PASS)
//SORTOUT DD DSN=&HORSES,UNIT=SCRTCH,DISP=(NEW,PASS),
// SPACE=(TRK,(200,20),RLSE),
// DCB=(RECFM=FB,LRECL=8,BLKSIZE=19016)
//SYSIN DD *
INREC FIELDS=(1,8)
SORT FIELDS=(1,8,CH,A)
SUM FIELDS=NONE
//S5 EXEC SYMSORT,TRACKS=1000,REGION=1024K
//**
//*COMBINES HORSES AND SIRES INTO ONE DATASET.
//**
//SORTIN DD DSN=&SIRES,UNIT=SCRTCH,DISP=(OLD,DELETE)
// DD DSN=&HORSES,UNIT=SCRTCH,DISP=(OLD,DELETE)
//SORTOUT DD DSN=&REGS,UNIT=SCRTCH,DISP=(NEW,PASS),
// SPACE=(TRK,(500,50),RLSE),
// DCB=(RECFM=FB,LRECL=8,BLKSIZE=19016)
//SYSIN DD *
SORT FIELDS=COPY
//S6 EXEC MATCHUP
//**
ASSIGNS EQUATION NUMBERS TO SIRES AND HORSES.
/**
/MASTERIN DD DSN=&REGS,UNIT=SCRTCH,DISP=(OLD,DELETE)
/MASTROUT DD DSN=&EQNS,UNIT=SCRTCH,DISP=(NEW,PASS),
/  SPACE=(TRK,(500,50),RLSE),
/  DCB=(RECFM=FB,LRECL=15,BLKSIZE=19020)
/SYSIN DD *
MASK(9,S,'000001,1')
/S7 EXEC MATCHUP
/**
//MATCHES RECORDS TO SIRE LIST AND INSERTS SIRE EQUATION NO.
/**
/MASTERIN DD DSN=&DATA1,UNIT=SCRTCH,DISP=(OLD,DELETE)
/SELECT DD DSN=&EQNS,UNIT=SCRTCH,DISP=(OLD,PASS)
/MASTROUT DD DSN=&DATA2,UNIT=SCRTCH,DISP=(NEW,PASS),
/  SPACE=(TRK,(1000,100),RLSE),
/  DCB=(RECFM=FB,LRECL=33,BLKSIZE=19008)
/SYSIN DD *
MASK(12,1,8//13,9,7)
/S8 EXEC SYMSORT,TRACKS=1500,REGION=1024K
/**
//SORTS RECORDS IN HORSE ORDER.
/**
/SORTIN DD DSN=&DATA2,UNIT=SCRTCH,DISP=(OLD,DELETE)
/SORTOUT DD DSN=&DATA3,UNIT=SCRTCH,DISP=(NEW,PASS),
/  SPACE=(TRK,(1000,100),RLSE),
/  DCB=(RECFM=FB,LRECL=33,BLKSIZE=19008)
  SORT FIELDS=(1,8,CH,A)
/S9 EXEC MATCHUP
/**
//MATCHES RECORDS TO HORSE LIST AND INSERTS HORSE EQUATION NO.
/**
/MASTERIN DD DSN=&DATA3,UNIT=SCRTCH,DISP=(OLD,DELETE)
/SELECT DD DSN=&EQNS,UNIT=SCRTCH,DISP=(OLD,DELETE)
/MASTROUT DD DSN=&DATA4,UNIT=SCRTCH,DISP=(NEW,PASS),
/  SPACE=(TRK,(1000,100),RLSE),
/  DCB=(RECFM=FB,LRECL=33,BLKSIZE=19008)
/SYSIN DD *
MASK(1,1,8//2,9,7)
/S10 EXEC SYMSORT,TRACKS=1500,REGION=1024K
/**
//SORTS RECORDS BY RACE AND SIRE WITHIN RACE.
/**
/SORTIN DD DSN=&DATA4,UNIT=SCRTCH,DISP=(OLD,DELETE)
/SORTOUT DD DSN=S.U3825.FILE440.EDIT,UNIT=DISK,DISP=(NEW,CATLG),
/  SPACE=(TRK,(500,50),RLSE),
/  DCB=(RECFM=FB,LRECL=31,BLKSIZE=6200)
/SYSIN DD *
  INREC FIELDS=(2,10,13,21)
  SORT FIELDS=(18,1Q,CH,A,11,7,CH,A)
APPENDIX C.

PL/I PROGRAM TO ADJUST RACING TIMES
FOR SEX AND AGE
//ADJUST JOB U3825,SAM
//**********************************************************************
//*THIS PROGRAM READS DATA FROM THE EDIT PROGRAM AND
//*CALCULATES SEX AND AGE ADJUSTMENTS FOR 220, 350,
//*400, 440 OR 870 YARD RACES. IT WILL ONLY CALCULATE
//*ADJUSTMENTS FOR ONE RACE DISTANCE AT A TIME AND THE
//*DISTANCE MUST BE SPECIFIED BY INITIALIZING THE
//*VARIABLE "LENGTH" (LINE 50) ACCORDING TO THE FOLLOWING:
//*
//* DISTANCE "LENGTH"
//*
//* 220 1
//* 350 2
//* 400 3
//* 440 4
//* 870 5
//*
//*DATA IS OUTPUT IN EXACTLY THE SAME FORMAT AS IT IS
//*READ, EXCEPT THAT ADJUSTED INSTEAD OF ACTUAL TIME
//*IS WRITTEN OUT.
//**********************************************************************
//SI EXEC PLIXCLG
//PLI.SYSIN DD *
ADJUST: PROC OPTIONS(MAIN) REORDER;
DCL SYSPRINT FILE PRINT;
DCL RECIN FILE INPUT RECORD
  ENV(TOTAL FB RECSIZE (31) BLKSIZE (31000));
DCL 1 DATA BASED (P),
  2 PAD1 CHAR(7),
  2 AGE PIC '99',
  2 SEX PIC '9',
  2 PAD2 CHAR(17),
  2 TIME PIC '99V99';
DCL RECOUT FILE OUTPUT RECORD
  ENV(TOTAL FB RECSIZE (31) BLKSIZE (31000));
DCL SEX_TABLE(5) STATIC INIT
  (.9969, .9968, .9957, .9956, .9933);
DCL AGE_TABLE(5,2,2:9) STATIC INIT
  (.189, .052, .000, -.027, -.028, .003, .047, .047,
  .174, .066, .000, -.024, -.007, .051, .151, .151,
  .097, .035, .000, -.009, .008, .050, .119, .119,
  .161, .062, .000, -.026, -.016, .030, .113, .113,
  .074, .013, .000, .006, .030, .073, .135, .135,
  .110, .042, .000, -.016, -.006, .031, .093, .093,
  .066, .021, .000, .005, .034, .089, .170, .170,
  .115, .040, .000, -.004, .028, .096, .200, .200,
  .000, .041, .000, .074, .181, .321, .495, .701,
  .000, -.051, .000, .095, .232, .413, .637, .905);
DCL ADJUSTIME FLOAT DEC(16) INIT(0);
DCL LENGTH FIXED BIN(15) INIT(4);
DCL AGE_FAC FIXED BIN(15) INIT(0);
DCL (P) PTR;
DCL EOD BIT(1) INIT(1'B);
ON ENDFILE (RECIN) EOD='0'B;
READ FILE (RECIN) SET (P);
DO WHILE (EOD);
IF AGE>8 THEN AGE_FAC=9;
ELSE AGE_FAC=AGE;
IF SEX=2 THEN
   ADJTIME=TIME*SEX_TABLE(LENGTH)-AGE_TABLE(LENGTH,2,AGE_FAC);
ELSE ADJTIME=TIME-AGE_TABLE(LENGTH,1,AGE_FAC);
   
   /* TIME IS ROUNDED TO THE NEAREST .01 SECONDS. */
   TIME=ADJTIME+.005;
WRITE FILE (RECOUT) FROM (DATA);
READ FILE (RECIN) SET (P);
END;
END ADJUST;
/*
//GO.RECIN DD DSN=S.U3825.FILE440.EDIT,UNIT=DISK,DISP=SHR
//GO.RECOUT DD DSN=FILE440.ADJUST,UNIT=(TAPE,,DEFER),DISP=(NEW,KEEP),
// VOL=SER=AGSL53,LABEL=(13,SL),
// DCB=(RECFM=FB,LRECL=31,BLKSIZE=31000)
APPENDIX D.

PL/I PROGRAM TO ABSORB CONTEMPORARY GROUP EQUATIONS
INTO SIRE AND HORSE EQUATIONS
//ABSORB JOB U3825,SAM
*JOBPARM DUPLEX=NO
*S1 EXEC PLIXCLG

//ABSORB JOB U3825,SAM
*JOBPARM DUPLEX=NO
*S1 EXEC PLIXCLG

/******************************************************************************/
/* THIS PROGRAM READS DATA OUTPUT BY THE EDIT PROGRAM */
/* AND CREATES TWO NEW FILES. RHS IS A FILE OF THE */
/* ABSORBED RHS'S SORTED BY ROW AND LHS IS A FILE OF */
/* THE UPPER DIAGONAL ELEMENTS OF THE COEFFICIENT */
/* MATRIX OF THE REDUCED EQUATIONS SORTED BY */
/* ROW AND COLUMN. THE DSN'S AND */
/* AMOUNT OF SPACE REQUIRED SHOULD BE UPDATED IN THE */
/* JCL LINES AT THE END OF THE PROGRAM BEFORE RUNNING. */
/******************************************************************************/

//PLI.SYSIN DD *

ABSORB: PROC OPTIONS (MAIN) REORDER;
DCL SYSPRINT FILE PRINT;
DCL RECIN FILE INPUT RECORD
  ENV(TOTAL FB RECSIZE (31) BLKSIZE (31000));
DCL 1 DATA BASED (P),
  2 HORSE PIC '99999999',
  2 PAD CHAR (3),
  2 SIRE PIC '99999999',
  2 RACE CHAR (10),
  2 TIME PIC '99999999';
DCL RHSOUT FILE OUTPUT RECORD
  ENV(TOTAL FB RECSIZE (20) BLKSIZE (19000));
DCL LHSOUT FILE OUTPUT RECORD
  ENV(TOTAL FB RECSIZE (20) BLKSIZE (19000));
DCL 1 EQN BASED (Q),
  2 (CODE, ROW, COL) FIXED BIN (31),
  2 COEF FLOAT DEC (16);
DCL (SIRE_COUNT, TEMP_SIRE, CHECK2, N1, N2, RDF, EDF, TDF, I, J, K)
  FIXED BIN (31) INIT (0);
DCL (TEMP_RACE, CHECK1) CHAR (10);
DCL (HORSE_COUNT, SUM, RACE_MEAN, YY, YXXY, YMY)
  FLOAT DEC (16) INIT (0);
DCL (SUM1, MEAN, RSS, ESS, TSS, RMS, EMS) FLOAT DEC (16) INIT (0);
DCL (HORSE_ARRAY, SIRE_ARRAY) (12)
  FIXED BIN (31) INIT ((12) 0);
DCL SIRE_PROG (12) FLOAT DEC (16) INIT ((12) 0);
DCL (P, Q) PTR;
DCL EOD BIT (1) INIT ('1'B);
ON ENDFILE (RECIN) EOD='0'B;
READ FILE (RECIN) SET (P);
CHECK1, TEMP_RACE = RACE;
START: DO WHILE (EOD);
CYCLE: DO WHILE ('CHECK1 = TEMP_RACE & EOD);
N1=N1+1;
SUM1=SUM1+TIME;
SUM=SUM+TIME;
YY=YY+(TIME**2);
HORSE_COUNT=HORSE_COUNT+1;
HORSE_ARRAY(HORSE_COUNT)=HORSE;
CHECK2=SIRE;
LOCATE EQN FILE (LHSOUT) SET (Q);
  CODE=5;/*** ZLMZ2 ***/
  ROW=SIRE;
  COL=HORSE;
  COEF=1.;
LOCATE EQN FILE (RHSOUT) SET (Q);
  CODE=6;/*** Z1MY ***/
  ROW=SIRE;
  COL='99999';
  COEF=TIME;
LOCATE EQN FILE (RHSOUT) SET (Q);
  CODE=7;/*** Z2MY ***/
  ROW=HORSE;
  COL='99999';
  COEF=TIME;
IF HORSE_COUNT=1 THEN DO;
  TEMP_SIRE=SIRE;
  SIRE_COUNT=1;
  SIRE_PROG(SIRE_COUNT)=1;
  SIRE_ARRAY(SIRE_COUNT)=TEMP_SIRE;
END;
ELSE DO;
  IF CHECK2=TEMP_SIRE THEN DO;
    SIRE_PROG(SIRE_COUNT)=SIRE_PROG(SIRE_COUNT)+1;
  END;
ELSE DO;
  TEMP_SIRE=CHECK2;
  SIRE_COUNT=SIRE_COUNT+1;
  SIRE_PROG(SIRE_COUNT)=1;
  SIRE_ARRAY(SIRE_COUNT)=TEMP_SIRE;
END;
END;
READ FILE (RECIN) SET (P);
CHECK1=RACE;
END;

END;

RACE_MEAN=SUM/HORSE_COUNT;
YXXY=YXXY+(HORSE_COUNT*RACE_MEAN**2);
N2=N2+1;
DO I=1 TO SIRE_COUNT;
  LOCATE EQN FILE (RHSOUT) SET (Q);
  CODE=6;/*** Z1MY ***/
ROW=SIRE_ARRAY(I);
COL=99999;
COEF=-(SIRE_PROG(I)*RACE_MEAN);
DO J=I TO SIRE_COUNT;
  LOCATE EQN FILE (LHSOUT) SET (Q);
  ROW=SIRE_ARRAY(I);
  COL=SIRE_ARRAY(J);
  IF I=J THEN DO;
    CODE=1;/* Z1MZ1 DIAG */
    COEF=SIRE_PROG(I)-(SIRE_PROG(I)**2/HORSE_COUNT);
  END;
  ELSE DO;
    CODE=2;/* Z1MZ1 OFFDIAG */
    COEF=-(SIRE_PROG(I)*SIRE_PROG(J)/HORSE_COUNT);
  END;
END;
DO K=1 TO HORSE_COUNT;
  LOCATE EQN FILE (LHSOUT) SET (Q);
  CODE=5;/* Z1MZ2 */
  ROW=SIRE_ARRAY(I);
  COL=HORSE_ARRAY(K);
  COEF=-(SIRE_PROG(I)/HORSE_COUNT);
END;
END;
DO I=1 TO HORSE_COUNT;
  LOCATE EQN FILE (LHSOUT) SET (Q);
  CODE=7;/* Z2MY */
  ROW=HORSE_ARRAY(I);
  COL=99999;
  DO J=I TO HORSE_COUNT;
    LOCATE EQN FILE (LHSOUT) SET (Q);
    IF I=J THEN DO;
      CODE=3;/* Z2MZ2 DIAG */
      ROW, COL=HORSE_ARRAY(I);
      COEF=1-(1./HORSE_COUNT);
    END;
    ELSE IF HORSE_ARRAY(I) < HORSE_ARRAY(J) THEN DO;
      CODE=4;/* Z2MZ2 OFFDIAG */
      ROW=HORSE_ARRAY(I);
      COL=HORSE_ARRAY(J);
      COEF=-(1./HORSE_COUNT);
    END;
    ELSE DO;
      CODE=4;/* Z2MZ2 OFFDIAG */
      ROW=HORSE_ARRAY(J);
      COL=HORSE_ARRAY(I);
      COEF=-(1./HORSE_COUNT);
    END;
  END;
END;
END;
TEMP_RACE=RACE;
HORSE_COUNT, SUM, SIRE_COUNT=0;
END;
/* ******************************************/
/* THIS ENDS DO LOOP LABELED 'START': */
/* THE LOOP CONTINUES UNTIL ALL */
/* RECORDS ARE READ. */
/* ******************************************/

MEAN=SUM1/N1;
TDF=N1-1;
RDF=N2-1;
EDF=N1-N2;
TSS=YY-(SUM1**2/N1);
RSS=YXXY-(SUM1**2/N1);
ESS=TSS-RSS;
RMS=RSS/RDF;
EMS=ESS/EDF;
YMY=YY-YXXY;
PUT PAGE EDIT('ANALYSIS OF VARIANCE') (SKIP(5), COL(55), A);
PUT EDIT('SOURCE', 'DF', 'SS', 'MS')
(SKIP(2), COL(35), A, X(8), A, X(11), A, X(21), A);
PUT EDIT('RACE', RDF, RSS, RMS)
(SKIP(2), COL(35), A, X(4), F(8), X(3), F(15, 4), X(8), F(12, 4));
PUT EDIT('ERROR', EDF, ESS, EMS)
(SKIP, COL(35), A, X(3), F(8), X(3), F(15, 4), X(8), F(12, 4));
PUT EDIT('TOTAL', TDF, TSS)
(SKIP, COL(35), A, X(3), F(8), X(3), F(15, 4));
PUT SKIP(5) EDIT('OVERALL MEAN=', MEAN, 'YMY=', YMY)
(COL(50), A, F(6, 3), SKIP, COL(50), A, E(30, 16));
END ABSORB;
/*
// GO.RECIN DD DSN=FILE440.ADJUST,UNIT=(TAPE,,DEFER),DISP=(OLD,KEEP),
// VOL=SER=AGSL53, LABEL=(13,SL)
// GO.RHSOUT DD DSN=&RHS,UNIT=SCRATCH,DISP=(NEW,PASS),
// SPACE=(TRK,(500,50),RLSE),
// DCB=(RECFM=FB,LRECL=20,BLKSIZE=19000)
// GO.LHSOUT DD DSN=&LHS,UNIT=SCRATCH,DISP=(NEW,PASS),
// SPACE=(TRK,(1000,100),RLSE),
// DCB=(RECFM=FB,LRECL=20,BLKSIZE=19000)
// STEP2 EXEC SYMSORT,TRACKS=1500,REGION=1024K
// SORTIN DD DSN=&RHS,UNIT=SCRATCH,DISP=(OLD,DELETE)
// SORTOUT DD DSN=FILE440.RHS,UNIT=(TAPE,,DEFER),DISP=(NEW,KEEP),
// VOL=SER=AGSL53, LABEL=(14,SL),
// DCB=(RECFM=FB,LRECL=20,BLKSIZE=31000)
// SYSIN DD *
// SORT FIELDS=(5, 4, BI,A)
SUM FIELDS=(13, 8, FL)
// STEP3 EXEC SYMSORT,TRACKS=1500,REGION=1024K
// SORTIN DD DSN=&LHS,UNIT=SCRATCH,DISP=(OLD,DELETE)
//SORTOUT DD DSN=FILE440.LHS,UNIT=(TAPE,,DEFER),DISP=(NEW,KEEP),
// VOL=SER=AGSL53,LABEL=(15,SL),
// DCB=(RECFM=FB,LRECL=20,BLKSIZE=31000)
//SYSIN DD *
SORT FIELDS=(5,8,BI,A)
SUM FIELDS=(13,8,FL)
APPENDIX E.

FORTRAN PROGRAM TO ESTIMATE VARIANCE COMPONENTS USING HSM
// This program reads output from the absorption program, sets
// up the necessary vectors and calculates the traces needed
// to estimate variance components for the sire and repeated
// records model using Henderson simple method. Estimates of
// the variance components, heritability, repeatability and
// the new alpha values are printed at the end of each round.
//-------------------------------------------------------------------------------

/ / F O R T . S Y S I N   D D *
C*****VALUES FOR THE PARAMETERS P1-P10 MUST BE SUPPLIED BEFORE
C*****THE PROGRAM IS RUN. PARAMETERS ARE DESCRIBED BELOW.
C******************************************************************************
C*****P1=NO. SIRES. (FROM EDIT PROGRAM)
C*****P2=NUMBER OF NONZERO ELEMENTS. (FROM ABSORB PROGRAM)
C*****P3=NO. SIRES + NO. HORSES. (ORDER)
C*****P4=AMOUNT OF WORKSPACE NEEDED FOR SUBROUTINE.
C*****P5=PRIOR FOR ALPHA1.
C*****P6=PRIOR FOR ALPHA2.
C*****P7=VALUE FOR Y'MY. (FROM ABSORPTION PROGRAM)
C*****P8=TOTAL NO. RECORDS - NO. RACES.
C*****P9=MAX NUMBER OF INTERNAL ITERATIONS ALLOWED.
C*****P10=MAX NUMBER OF EXTERNAL ITERATIONS ALLOWED.
C******************************************************************************

INTEGER P1,P2,P3,P4,P9,P10
REAL*8 P5,P6,P7,P8
PARAMETER (P1=627,P2=75135,P3=4489,P4=4489)
PARAMETER (P5=9.13054595D0,P6=6.11869396D0)
PARAMETER(P7=738.21497895D0,P8=4746.DO)
PARAMETER(P9=100,P10=15)
CHARACTER^20 RECIN
CHARACTER^1 REC(20)
INTEGER CODE,ROW,COL
REAL*8 COEF,DUMMY(3)
EQUIVALENCE (REC(1),DUMMY(1)),(REC(5),RECIN,CODE),
* (REC(9),ROW),(REC(13),COL),(REC(17),COEF)

C*****DUMMY IS A VECTOR USED TO CREATE MORE EFFICIENT
C*****CHARACTER ALIGNMENT. DO NOT USE FOR STORAGE!!!

INTEGER SIRES/P1/, NLNL/P2/, ORDER/P3/, NW/P4/
INTEGER BEFORE,AFTER,DIFF,IER,ITMAX1/P9/, ITMAX2/P10/
INTEGER JA(P2), IA(P3+1), IWKSP(P3*3), I Parm(12)
REAL*8 ALPHA1/P5/, ALPHA2/P6/, PALPH1/P5/, PALPH2/P6/
REAL*8 YMY/P7/, TRQ00/P8/
REAL*8 TRQ10, TRQ11, TRQ12, TRQ20, TRQ21, TRQ22
REAL*8 U111, U222, U121MY, U222MY, DALPH1, DALPH2
REAL*8 DIAG1(P3), DIAG2(P3), DIAG3(P3)
REAL*8 A(P2), RHS(P3), U(P3), PREVU(P3), WKSP(P4), RP Parm(12)
REAL*8 F(2,2), FINV(2,2), WKAREA(18)
REAL*8 ERRVAR, SIRVAR, PERVAR, H2, R, SSDIFF, CRITER

DO 10 I = 1, ORDER
    U(I) = 0.0D0
    PREVU(I) = 0.0D0
    DIAG1(I) = 0.0D0
    DIAG2(I) = 0.0D0
    DIAG3(I) = 0.0D0
10 CONTINUE

C***** READING AND STORING RHS
DO 50 I = 1, ORDER
    READ(10, 100) RECIN
    RHS(ROW) = COEF
50 CONTINUE
CLOSE (10)

C***** READING AND STORING LHS
DO 700 I = 1, NLNL
    READ(11, 100) RECIN
100 FORMAT (A20)
    GOTO (200, 300, 400, 500, 600), CODE
200 A(I) = COEF + ALPHA1
    JA(I) = COL
    IA(ROW) = I
    DIAG1(ROW) = COEF
    DIAG2(ROW) = DIAG2(ROW) + COEF**2
GOTO 700
300 A(I) = COEF
    JA(I) = COL
    DIAG2(ROW) = DIAG2(ROW) + COEF**2
    DIAG2(COL) = DIAG2(COL) + COEF**2
GOTO 700
400 A(I) = COEF + ALPHA2
    JA(I) = COL
    IA(ROW) = I
    DIAG1(ROW) = COEF
    DIAG2(ROW) = DIAG2(ROW) + COEF**2
GOTO 700
500 A(I) = COEF
    JA(I) = COL
    DIAG2(ROW) = DIAG2(ROW) + COEF**2
    DIAG2(COL) = DIAG2(COL) + COEF**2
GOTO 700
600 A(I) = COEF
    JA(I) = COL
    DIAG3(ROW) = DIAG3(ROW) + COEF**2
    DIAG3(COL) = DIAG3(COL) + COEF**2
GOTO 700
700 CONTINUE
CLOSE (11)
    IA(ORDER + 1) = NLNL + 1

C***** BEGINNING OF EXTERNAL ITERATION
DO 999 I=1,ITMAX2
WRITE(6,750) I
750 FORMAT(T1,'I',T10,'************* ITERATION NO. ', I4,'*************')
C*****SETTING PARAMETERS FOR INTERNAL ITERATION
CALL ERRSET(208,0,-1,1,0)
CALL DFAULT(IPARM,RPARM)
IPARM(1)=ITMAX1
IPARM(2)=2
C*****SOLVING EQUATIONS
CALL CLOCK(BEFORE)
CALL SOR(ORDER,IA,JA,A,RHS,U,IWKSP,NW,WKSP,IPARM,RPARM,IER)
CALL CLOCK(AFTER)
DIFF=BEFORE-AFTER
WRITE(6,800)BEFORE,AFTER,DIFF
800 FORMATIO, 'BEFORE=',I8,T35, 'AFTER=',I8,T60, 'DIFF=', 15,/) IF (IER.NE.0) GOTO 1000
U1U1=0.DO
U2U2=0.DO
U2Z2MY=0.DO
TRQ10=0.DO
TRQ11=0.DO
TRQ12=0.DO
TRQ20=0.DO
TRQ21=0.DO
TRQ22=0.DO
SSDIFF=0.DO
C*****CALCULATING SS AND EXPECTATIONS FOR QUADRATIC FORMS
DO 925 J=1,SIRES
   U1U1=U1U1+(RHS(J)/(DIAG1(J)+ALPHA1))**2
   U1Z1MY=U1Z1MY+(U(J)*RHS(J))
   TRQ10=TRQ10+(DIAG1(J)/(DIAG1(J)+ALPHA1)**2)
   TRQ11=TRQ11+(DIAG2(J)/(DIAG1(J)+ALPHA1)**2)
   TRQ12=TRQ12+(DIAG3(J)/(DIAG1(J)+ALPHA1)**2)
   SSDIFF=SSDIFF+(U(J)-PREVU(J))**2
   PREVU(J)=U(J)
925 CONTINUE
DO 950 J=SIRES+1,ORDER
   U2U2=U2U2+(RHS(J)/(DIAG1(J)+ALPHA2))**2
   U2Z2MY=U2Z2MY+(U(J)*RHS(J))
   TRQ20=TRQ20+(DIAG1(J)/(DIAG1(J)+ALPHA2)**2)
   TRQ21=TRQ21+(DIAG3(J)/(DIAG1(J)+ALPHA2)**2)
   TRQ22=TRQ22+(DIAG2(J)/(DIAG1(J)+ALPHA2)**2)
   SSDIFF=SSDIFF+(U(J)-PREVU(J))**2
   PREVU(J)=U(J)
950 CONTINUE
WRITE(6,960) U1U1,U2U2,YMY,TRQ10,TRQ11,TRQ12,
   TRQ20,TRQ21,TRQ22
960 FORMAT(T10,'U1U1=',F13.6,T35, 'U2U2=',F13.6, T55,
   'YMY=',F13.6,...
C*****SOLVING FOR THE VARIANCE COMPONENTS

F(1,1) = TRQ11
F(I,2) = TRQ12
F(2,1) = TRQ21
F(2,2) = TRQ22

IDGT = 6

CALL LINVF(F, 2, 2, FINV, IDGT, WKAREA, IER)

ERRVAR = (YMY - U1Z1MY - U2Z2MY) / TRQ00
SIRVAR = FINV(1, 1) * (U1U1 - (TRQ10 * ERRVAR))
PERVAR = FINV(2, 1) * (U1U1 - (TRQ10 * ERRVAR))

ALPHA1 = ERRVAR / SIRVAR
ALPHA2 = ERRVAR / PERVAR

H2 = (4 * SIRVAR / (SIRVAR + PERVAR + ERRVAR))
R = (SIRVAR + PERVAR) / (SIRVAR + PERVAR + ERRVAR)

CRITER = DSQRT(SSDIFF / ORDER)

WRITE(6, 975) ERRVAR, SIRVAR, PERVAR, ALPHA1, ALPHA2, H2, R, CRITER

IF (CRITER.LT.0.001) GOTO 1000

C*****CHANGING ALPHA VALUES OF LHS

DALPH1 = ALPHA1 - PALPH1
DALPH2 = ALPHA2 - PALPH2

DO 990 J = 1, SIRES
   A(IA(J)) = A(IA(J)) + DALPH1
990 CONTINUE

DO 995 J = SIRES + 1, ORDER
   A(IA(J)) = A(IA(J)) + DALPH2
995 CONTINUE

PALPH1 = ALPHA1
PALPH2 = ALPHA2

1000 CONTINUE

STOP

END

//LKED.SYSLIB DD

//
// DD
// DD DSN=SYSU.LINPACK.VSUBLIB,UNIT=DISK,DISP=SHR
// DD DSN=SYSU.ITPACK.SUBLIB,DISP=SHR
// DD DSN=SYS2.FORTV.IMSL.DOUBLE,DISP=SHR
// GO.FT10F001 DD DSN=FILE440.RHS,UNIT=(TAPE,DEFER),DISP=(OLD,KEEP),
// VOL=SER=AGSL53,LABEL=(10,SL)
// GO.FT11F001 DD DSN=FILE440.LHS,UNIT=(TAPE,DEFER),DISP=(OLD,KEEP),
// VOL=SER=AGSL53,LABEL=(11,SL)
APPENDIX F.

FORTRAN PROGRAM TO ESTIMATE VARIANCE COMPONENTS USING THM
//THM JOB U382S,SAM
//STEP1 EXEC FORTVCILG,FVPOPT=2,D=DOUBLE,TIME.GO=5,REGION.GO=3000K
//JOBPARM DUPLEX=NO
//***************************************************************************
//THIS PROGRAM READS OUTPUT FROM THE ABSORPTION PROGRAM, SETS
//UP THE NECESSARY VECTORS, CALLS AN ITERATION PROCEDURE
//TO GET TRUE SOLUTIONS AND CALCULATES THE TRACES NEEDED
//TO ESTIMATE VARIANCE COMPONENTS FOR THE SIRE AND REPEATED
//RECORDS MODEL USING THE TILDE-HAT METHOD. ESTIMATES OF
//THE VARIANCE COMPONENTS, HERITABILITY, REPEATABILITY AND
//THE NEW ALPHA VALUES ARE PRINTED AT THE END OF EACH ROUND.
//***************************************************************************
//FORT.SYSIN DD *

C*****VALUES FOR THE PARAMETERS P1-P10 MUST BE SUPPLIED BEFORE
C*****THE PROGRAM IS RUN. PARAMETERS ARE DESCRIBED BELOW.
C***************************************************************************
C*****P1=NO. SIRES. (FROM EDIT PROGRAM)
C*****P2=NUMBER OF NONZERO ELEMENTS. (FROM ABSORB PROGRAM)
C*****P3=NUMBER OF SIRE + NUMBER OF HORSES. (ORDER)
C*****P4=AMOUNT OF WORKSPACE NEEDED FOR SUBROUTINE.
C*****P5=PRIOR FOR ALPHA1.
C*****P6=PRIOR FOR ALPHA2.
C*****P7=VALUE FOR Y'MY. (FROM ABSORPTION PROGRAM)
C*****P8=TOTAL NO. RECORDS - NO. RACES.
C*****P9=MAX NUMBER OF INTERNAL ITERATIONS ALLOWED.
C*****P10=MAX NUMBER OF EXTERNAL ITERATIONS ALLOWED.
C***************************************************************************

INTEGER P1,P2,P3,P4,P9,P10
REAL P5,P6,P7,P8
PARAMETER (P1=2124,P2=180733,P3=11589,P4=11589)
PARAMETER (P5=8.35736868D0,P6=3.73735100D0)
PARAMETER (P7=1897.49249718D0,P8=11550.D0)
PARAMETER (P9=100,P10=15)
CHARACTER*20 RECIN
CHARACTER*1 REC(20)
INTEGER CODE,ROW,COL
REAL*8 C0EF,DUMMY(3)
EQUIVALENCE (REC(1),DUMMY(1)),(REC(5),RECIN,CODE),
* (REC(9),ROW),(REC(13),COL),(REC(17),COEF)
C*****DUMMY IS A VECTOR USED TO CREATE MORE EFFICIENT
C*****CHARACTER ALIGNMENT. DO NOT USE FOR STORAGE!!
INTEGER SIRES/P1/,NLNL/P2/,ORDER/P3/,NW/P4/
INTEGER BEFORE,AFTER,DIFF,IER,ITMAX1/P9/,ITMAX2/P10/
INTEGER IA(P2),JA(P3+1),INKSP(P3*3),IPARM(12)
REAL*8 A(P2),D(P3),RHS(P3),U(P3),PREVU(P3)
REAL*8 WKSP(P4),RPARM(12)
REAL*8 ALPHA1/P5/,ALPHA2/P6/,PALPH1/P5/,PALPH2/P6/
REAL*8 YMY/P7/,TRQQO/P8/,CF1/0/,CF2/0/
REAL*8 TRQ11,TRQ22,U111,U2U2,U1ZIMY,U2Z2MY,DIAG
REAL*8 ERRVAR,SIRVAR,PERVAR,H2,R,SSDIFF,CRITER
DO 10 I=1,ORDER
  U(I)=0.DO
  PREVU(I)=0.DO
10 CONTINUE
C***** READING AND STORING RHS DATA
DO 50 I=1,ORDER
  READ(10,100)RECIN
  RHS(ROW)=COEF
50 CONTINUE
CLOSE (10)
C***** READING AND STORING LHS DATA
DO 200 I=1,NLNL
  READ(11,100)RECIN
  IF(ROW.NE.COL)THEN
    JA(I)=COL
    A(I)=COEF
  ELSEIF(CODE.EQ.1)THEN
    JA(I)=COL
    A(I)=COEF+ALPHA1
    IA(ROW)=I
    D(ROW)=A(I)
    TRQ11=TRQ11+(COEF/A(I))
  ELSE
    JA(I)=COL
    A(I)=COEF+ALPHA2
    IA(ROW)=I
    D(ROW)=A(I)
    TRQ22=TRQ22+(COEF/A(I))
  END IF
200 CONTINUE
CLOSE (11)
IA(ORDER+1)=NLNL+1
C***** BEGINNING OF EXTERNAL ITERATION
DO 999 I=1,ITMAX2
  WRITE(6,750) I
  750 FORMAT(T1,'1'ITERATION NO. ',*,I4,' **********')
C***** SETTING PARAMETERS FOR INTERNAL ITERATION
  CALL ERRSET(208,0,-1,1,0)
  CALL DFAULT(IPARM,RPARM)
  IPARM(1)=ITMAX1
  IPARM(2)=2
C***** SOLVING EQUATIONS******
  CALL CLOCK(BEFORE)
  CALL SOR(ORDER,IA,JA,A,RHS,U,IWKSP,NW,WKSP,IPARM,RPARM,IER)
  CALL CLOCK(AFTER)
  DIFF=BEFORE-AFTER
  WRITE(6,800)BEFORE,AFTER,DIFF
  800 FORMAT(TIO,'BEFORE=', I8,T35,'AFTER= ', I8,T60,'DIFF=', 15,/)

IF (IER.NE.0) GOTO 1000
U1U1=0.D0
U2U2=0.D0
U1Z1MY=0.D0
U2Z2MY=0.D0
SSDIFF=0.D0
C*****CALCULATING SS AND EXPECTATIONS OF QUADRATIC FORMS.
DO 925 J=1,SIRES
   U1U1=U1U1+(U(J)*RHS(J)/D(J))
   U1Z1MY=U1Z1MY+(U(J)*RHS(J))
   SSDIFF=SSDIFF+(U(J)-PREVU(J))**2
   PREVU(J)=U(J)
925 CONTINUE
DO 950 J=SIRES+1,ORDER
   U2U2=U2U2+(U(J)*RHS(J)/D(J))
   U2Z2MY=U2Z2MY+(U(J)*RHS(J))
   SSDIFF=SSDIFF+(U(J)-PREVU(J))**2
   PREVU(J)=U(J)
950 CONTINUE
WRITE(6,960) YMY,U1Z1MY,U2Z2MY,TRQ00,U1U1,TRQ11,U2U2,TRQ22
960 FORMAT(T2,'YMY=',F13.8,T20,'SIRE SS=',F13.8,T45,'HORSE SS=',F13.8,T45,'DF=',F7.0,T2,'U1U1=',F13.8,T20,'TR(D1)=',F13.8,T2,'U2U2=',F13.8,T20,'TR(D2)=',F13.8,T2)
C*****SOLVING FOR THE VARIANCE COMPONENTS
IF (CF1*ERRVAR.GT.U1U1) CF1=CF1*.95
   SIRVAR=(U1U1-(CF1*ERRVAR))/(TRQ11-(CF1*ALPHAl))
IF (CF1*ERRVAR.GT.U2U2) CF2=CF2*.95
   PERVAR=(U2U2-(CF2*ERRVAR))/(TRQ22-(CF2*ALPHA2))
   ERRVAR=(YMY-U1Z1MY-U2Z2MY)/TRQ00
   ALPHAl=ERRVAR/SIRVAR
   ALPHA2=ERRVAR/PERVAR
   H2=(4*SIRVAR)/(SIRVAR+PERVAR+ERRVAR)
   R=(SIRVAR+PERVAR)/(SIRVAR+PERVAR+ERRVAR)
   CRITER=DSQRT(SSDIFF/ORDER)
WRITE(6,975) ERRVAR,SIRVAR,PERVAR,ALPHAl,ALPHA2,H2,R,CRITER
975 FORMAT(T10,'ERROR VAR=',F12.8,T35,'SIRE VAR=',F12.8,*T60,'PER ENV VAR=',F12.8,T10,'ALPHAl=',F12.8,*T35,'ALPHA2=',F12.8,T10,'H2=',F8.4,T35,'R=',F8.4,*T10,'ITERATION CRITERION=',F10.6,/) IF (CRITER.LT.0.0002) GOTO 1000
C*****CHANGING ALPHA VALUES OF LHS AND RECALCULATING TRACES
TRQ11=0.D0
TRQ22=0.D0
CF1=0.D0
CF2=0.D0
DO 990 J=1,SIRES
   DIAG=A(IA(J))-PALPHl
   A(IA(J))=DIAG+ALPHAl
   D(J)=A(IA(J))
   TRQ11=TRQ11+(DIAG/A(IA(J)))
CF1=CF1+(DIAG/(A(IA(J))*2))

990 CONTINUE
DO 995 J=SIRE+1,ORDER
   DIAG=A(IA(J))-PALPH2
   A(IA(J))=DIAG+ALPHA2
   D(J)=A(IA(J))
   TRQ22=TRQ22+(DIAG/A(IA(J))}
   CF2=CF2+(DIAG/(A(IA(J))*2))
995 CONTINUE
PAPH1=ALPHA1
PALPH2=ALPHA2
999 CONTINUE
1000 CONTINUE
STOP
END

//LKED.SYSLIB DD
// DD
// DD
// DD DSN=SYSU.LINPACK.VSUBLIB,UNIT=DISK,DISP=SHR
// DD DSN=SYSU.ITPACK.SUBLIB,DISP=SHR
//GO.FT10F001 DD DSN=FILE440.RHS,UNIT=(TAPE,DEFER),DISP=(OLD,KEEP),
// VOL=SER=AQHAS,LABEL=(11,SL)
//GO.FT11F001 DD DSN=FILE440.LHS,UNIT=(TAPE,DEFER),DISP=(OLD,KEEP),
// VOL=SER=AQHAS,LABEL=(12,SL)