Identifying Broodmare Prospects for an Elite Broodmare Band

Prepared for Dan Rosenberg

By Robert L. Losey, Erradame C. Morales, and Timothy Capps
Equine Industry Program
University of Louisville

June 2010

Outline

Introduction

This study provides statistically-based analyses that will be useful to breeders seeking to build an elite broodmare band. Among our goals is the elucidation of the correctness, or lack thereof, of the statement by Joe Estes, the revered editor of the Blood-Horse from 1935-1963, who argued that *"The conclusion remains unshaken that the value of a mare as a producer is in direct proportion to her class as a racer[1] ..."* Mr. Estes faced the monumental task of analyzing racing records when the best computers were less powerful than hand-held calculators of today, hence the dearth of production by racing periodicals of the Estes era of what we would now consider sophisticated statistical analysis. Had Estes been analyzing this same subject today, we strongly suspect that he would have undertaken or commissioned a sophisticated analysis of his theory, and that it would have proceeded along the lines of what we have produced in the following pages.

Questions addressed by this study include the following:.

What advantage and how much of an advantage is there to exclusively limiting one's broodmare band to GRADED STAKES WINNERS?
How well do LISTED STAKES WINNERS perform as broodmares?
How well do well-bred ALLOWANCE fillies do as broodmares?
Are there unsung broodmare types that exceed expectations?
Are there broodmare types one should shun?
HOW IMPORTANT IS SPEED?
How disadvantageous is a proven lack of speed?
Do UNRACED FILLIES make better broodmare prospects than non-winners?
How do UNRACED FILLIES perform as broodmares in comparison to allowance winners and to listed stakes winners?[2]

These and related questions are addressed in this study. In seeking the answers to these questions we use a statistical design that facilitates insights that previous studies could not provide. If the reader is willing to plow back through years of issues of the Thoroughbred Times, the Blood-Horse, or other racing/breeding periodicals, (s)he will find that there are more than a few studies that report the racing success of various breeding approaches. Many are essentially ad hoc analyses that are not scientifically defensible, while some of the better articles report useful statistics (typically in tabular form), but do not provide sophisticated analyses. Even the best of these articles invariably have suffered from serious analytical shortcomings. These shortcomings, our approaches to minimizing their effects, and the results of our (more analytically sound) approach to analyzing categories of fillies/mares that are likely to become top producers are presented in the body of this paper. (For a summary of our findings see the "Summary and Conclusions" section at the end of this paper.)

---

[1]Joseph A. Estes, The Estes Formula for Breeding Stakes Winners (Neenah Wisc., The Russell Meerdink Co.) 80.

[2] See the discussion of the question whether Toussand or Primal Force was the better broodmare in Appendix B of this paper for a practical look at the type of question the analysis in the paper seeks to address.

Shortcomings of the Best Previous Studies and an Improved Methodology

Previous studies rather consistently report the not-surprising result that high-class race fillies turn out to be better producers of quality racehorses than less accomplished race fillies. A statistician looking at these results would immediately ask the question: Have the results been "normalized" for the quality of the pedigrees and the quality of the stallions to which the high class race fillies were bred?

Said another way, the statistician is concerned that perhaps the reason the high class race fillies produced better runners is not so much that they were better race fillies, but perhaps that they had better pedigrees and that the high quality pedigrees were more important in determining their foals' racing abilities than the dams' racing class. Or maybe it was that the high-class race fillies were bred to much better stallions than the norm and that high-class stallions are the determining factor in producing high quality runners. The point is that previous studies have generally failed to sort out the relative importance of racing class, pedigrees, and sire quality. Perhaps, as statisticians might point out, race fillies with more pedestrian race records, but with similar pedigrees, and bred to similar high-quality stallions, would produce runners that are the equal of those from the better race fillies. You may argue that all three factors are important, perhaps even almost equally important—and you may very well be right-- but the point is, the evidence to support that argument has not been carefully researched.

Our study takes a significant step toward solving at least a part of this quandary by crafting a well-designed research methodology. We focus primarily on two things, the importance of the speed of the dam on the racing performance of her foals and the class of races in which the dam was successful. We also make significant progress toward solving the question of the importance of "sire power." Because the design of our study tries to hold dam pedigree relatively constant, as a result we have less to say about pedigree of the dam

Let us restate our view of the problem with previous studies in slightly different words. We argue that in previous studies:

> 1) no adjustment is typically made for the fact that top-class race mares, especially graded stakes winners, tend to be bred to much better stallions than mere stakes-winning or top allowance mares (we call this the "sire-power" problem), and

2) the standardization of, or adjustments for the quality of pedigrees of broodmare prospects that are the subjects of analyses (the pedigree problem) is typically not addressed, or when it is addressed, is done so in an inadequate fashion. We use an approach that substantially mitigates the sire-power and pedigree problems as is explained in the following pages.

Adjustments for Pedigree--Broodmare Prospect Data

Two questions of primary importance are addressed in our study:

1) What, if any difference does racing class make when choosing broodmare prospects?

2) What difference does the "speed" of a broodmare prospect make.

A third question of interest about which we find interesting results, though they are less robust than the results relative to a mare's racing class and speed, is the importance of the sire in breeding top quality runners. To address these questions we must in some way either separate out or neutralize the effects of pedigree (the pedigree problem) and the stallion to which a broodmare was bred (the sire power problem) from racing class and speed. Alternatively, we might use appropriate statistical tools (multivariate analyses such as multiple regression) that allow us to simultaneously estimate the relative importance of pedigree, sire power, and racing class/speed.

How do we manage the pedigree problem?

In this study we mitigate the effects of pedigree and sire power by choosing to analyze race fillies that are relatively homogeneous with respect to pedigrees and the sires to which they were bred. This allows us to focus on the importance of racing class and speed on broodmare productivity. We first address the question of how to "neutralize" the pedigree problem so that we may focus on the importance of speed and racing class. (A later section will discuss how we deal with the sire power question.)

We have gone a long way toward neutralizing the pedigree problem by drawing our sample of broodmare prospects from a relatively homogeneous group of fillies. The Jockey Club has provided data on 1500 named yearling fillies that were sold at the Keeneland July Sale between 1987 and 2000. We chose this source of broodmare prospects because it meant that we would be including only fillies with excellent pedigrees that met the strict requirements imposed by Keeneland Sales in limiting the July Summer Sale to top pedigrees.

In our sample we went one step further toward a more homogeneous group of broodmare prospects by asking the Jockey Club to exclude the top and bottom 10% tails based on sales prices of each year's sale from the sample that we analyzed. This means that we further narrowed the range in pedigree quality. The decision to omit the top and bottom 10% of the sales fillies was also based on the premises that the bottom 10% might contain an undue number of fillies with inferior conformation and/or pedigrees, and that the top 10% might too often be "bidding war" fillies that were not cost-effective purchases.[3]

A substantial majority of the 1500 yearling fillies (82%) subsequently raced.[4] We stratified the summer sales fillies by their racing results (see categories in Table 1 that follows). Though we had originally intended to analyze grade 1, grade 2, and grade 3 SWs by category separately, we found that the sample of graded SWs was too small to obtain meaningful results if we did this, so we considered all graded SWs as a group in the analyses.

The 1500 summer sales fillies had produced 7313 recorded foals by the cut-off date (late summer 2009) used in this study. However, the final sample of foals we analyzed was pared to a much lower number based on a number of criteria, the first being the exclusion from consideration of the 823 foals sired by foreign stallions. The decision to exclude foreign sires was based on

1) a desire to decrease the heterogeneity of the sample, and

2) a related preference to analyze foals targeted primarily for U.S. racing.

For the 6490 foals remaining in the sample, approximately 53% were winners, 23% raced but did not win, and 24% did not race. Having gone a substantial way toward basing our study on a relatively homogeneous sample of filly pedigrees, we then set out to pare the

---

[3]Note that our omission of the top 10% of the sales fillies probably biases downward the potential racing results of the foals from the sample fillies relative to what would have obtained for someone who "plays only at the top end." An investor who bought significant numbers of the top 10% fillies would almost surely find that their resulting percentage of top quality racers would be greater than an investor who spread his/her purchases across the summer sale spectrum.

[4] This 82% figure reported for the percentage of fillies that raced may have been biased upward slightly from the true percentage that raced because the Jockey Club sample was drawn from "named" fillies only. Thus the (probably very) small number of summer sale graduates that died between the July sale and the following year's naming deadline were not included in the sample of potential racers.

remaining 6490 foals down to a much smaller sample based on our efforts to solve the "sire-power" problem.

Stallion Data

In order to adjust for the "sire-power" bias inherent in other studies that have attempted to look at breeding effectiveness, we looked only at a sub-sample of the foals produced by the (former summer sale) mares that were eventually produced from matings to a group of stallions consisting of those that had historically produced at least 2.77% graded stakes winners from foals.[5] Though we initially also considered both AEI and percentage of stakes winners as potential criteria for ranking stallions, we settled on the percentage of graded stakes winners as the sole criterion for stallion effectiveness.[6] The decision to limit the analysis to the top 47 sires resulted in a dramatic decrease in the number of foals analyzed in our final sample and, coupled with the necessity of also paring the sample when speed figures were not available for racers (see the following section) this resulted in dropping the final sample to 1140 foals.

The sample would have been slightly larger had we not imposed one additional criterion: We determined that it made sense to omit any stallion that had five or fewer foals in the sample of 6490 that remained after we omitted all foreign-sired foals. Thus we omitted, for example, the outstanding sire El Gran Senor. Our decision to do this was done on the grounds that the effect on the results would likely be minimal from omitting the very small number of foals that fell into this category, while saving substantial analytical effort.

Our final sample consisted of 47 "top stallions." Storm Cat, Nureyev, A. P. Indy, Seattle Slew, and Danzig ranked 1 through 4, while Strawberry Road, Quiet American, Red Ransom, and Royal Academy were ranked 44 through 47. Number 47, Royal Academy, had produced graded stakes winners as a percentage of foals at a rate of 2.77%.[7] The top-47 list of stallions is reported as Table 2 in this document.

[5] Though the 2.77% figure (for Royal Academy) may be viewed as arbitrary, in looking at the stallion record data, we noted that there was a natural breaking point at this figure, with the next stallion exhibiting a significantly lower percentage of graded/listed stakes winners.

[6] Because percentage of graded stakes winners, overall stakes winners, and AEI are all highly correlated, we would have obtained similar rankings regardless of which criterion we had used in ranking stallions.

[7] The 2.77% figure was a natural cutoff, as the next highest figure was 2.4% for Crafty Prospector, Capote, and Forest Wildcat. When we recalculated based on the figures from the 2010 *Stallion Register*, Forestry

Combining the Mare and Stallion Data

As previously mentioned, the sample of foals was further pared before inclusion in the final analyses based on "speed" and "racing age" criteria in addition to the "sire-power" criterion.

1) So that we could focus on the importance of "speed" of the broodmare, we omitted all broodmares without reported speed ratings that were in the listed stakes winners, allowance, or "other" (mostly maiden) winners categories. (As mares in other categories were not classified by speed, the other categories were not pared down in this way).
2) Foals born after 2005 were omitted from the sample on the theory that these foals had not had sufficient time to develop a race record by the September 2009 date for which the Jockey Club provided data.

These various exclusions left us with a final sample of 1140 foals. We then focused on the ability of the (former summer sales) mares, stratified by racing class and/or speed, to produce **graded or listed** stakes winners when bred to our "top 47" sires. Our choice of focusing on the production of graded or listed stakes winners was consistent with the original intent of the study, which was to attempt to determine how a breeder might build a broodmare band that produces runners that are competitive at the highest levels.

A Brief Recap – To this point we have explained how we have developed a broodmare band that is relatively homogeneous in two very important dimensions: 1) pedigree quality, and 2) quality of the stallions to which this broodmare band is bred. We now look at the demonstrated racing abilities of the mares in this otherwise relatively homogeneous broodmare band and analyze the importance of racing class and speed.

Table 1 (see following page) reports the most important results of our analyses. Using Jockey Club classifications provided to us, we reported seven racing classifications for the summer sale graduate fillies who later became broodmares, and we listed the

and Mineshaft, who were on our "top 49" list based on 2009 Stallion Register figures, each dropped below 2.4% and were dropped from the list, leaving the "top 47" stallions.

number and percentage of graded plus listed stakes winners produced from mares from each racing category. Our sample stratified the mares as follows:

1) Graded stakes winners
2) Listed stakes winners
3) "Other stakes winners" (mostly stakes winners outside the U.S., but some restricted and non-listed stakes winners).
4) Allowance winners
5) Winners of "other races" (mostly maiden winners, though some also won starter or claiming races).
6) Non-winners, and
7) Unraced mares

We further stratified the data by partitioning groups 2, 4, and 5 above based on the top speed figure reported in the Jockey Club data. We did **not** stratify the data by speed figures for graded stakes winners for two reasons. First, we noted that the speed figures for graded SWs tended to be consistently very high and with much less differentiation than in other categories of racing class, and second and most importantly, we would have had very small split samples for graded stakes winners from which it would be difficult or impossible to infer that results were statistically significant. We also did not stratify the "Other Stakes Winners category because it contained so few mares that little could reliably be said about the results, and we did not stratify the Non-Winner category for which speed figures were quite low. Finally, as stratifying the Unraced category would only have been possible based on workout information that is not easily obtained, we did not split this sample.

TABLE 1

Frequency of Graded/Listed SW Runners from Keeneland Summer Sales Graduates Bred to "Top 47" Stallions: Dams Arranged by Racing Class and Adjusted for Sire Quality

A. Dam Category	B. No. of Foals	C. Graded SW Foals	D. Listed SW Foals	E. % Foals That Were Gr/L SWs	F. % of G/Listed SWs Ignoring Speed	G. % of G/Listed Foals Produced: Adjusted Estimate
1) Graded SW	139	14	5	Composite-->	13.67%	12.99%**
						(See Appendix)
2A) Listed SW: SR>1000*	72	3	3	8.33%	6.19%	<-Composite % for
2B) SR<1000	25	0	0	0.00%		All Listed SW Dams
3) Other SW	44	3	2	Composite-->	11.36%	
4A) Allowance: SR>1000	118	3	6	7.63%	5.86%	<-Composite % for
4B) SR<1000	104	4	0	3.85%		All Alw Dams
5A) Other Wnrs: SR>900	31	3	0	9.68%	5.99%	<-Composite % for
5B) SR<900	136	5	2	5.15%		Other Wnng Dams
6) Non-Winners	284	3	3	Composite-->	2.11%	
7) Unraced	188	6	4	Composite-->	5.32%	
Total foals	1140					

*SR is the speed rating provided by the Jockey Club. It is equal to the Equibase speed figure times ten.

**This adjusted estimate was necessary because Graded SW mares were bred to higher rated stallions than all other mare categories. An explanation of the adjustment process is contained in Appendix A to this paper.

Reading and Analyzing Table 1: An Overview

Table 1 above reports the most important results from our analyses. Refer to rows 2A and 2B as illustrative examples to explain Table 1 results. Column A lists the categories of mares (which are characterized in the written material preceding Table 1). Row 2A reports the produce records for North American (U.S. and Canadian) listed stakes-winning dams whose top speed ratings as reported by the Jockey Club were 1000

or greater. Column B reports that our sample contained 72 foals that were produced from this category of dam. Column C reports that three of these foals became graded SWs. Column D reports that three additional foals were listed (but not graded) SWs. Column E reports the calculated combined percentage of graded and listed stakes winning foals [8.33% in this case] produced from listed stakes-winning dams.

For Row 2B, Columns A-F report the same information as described above for listed SWs that had speed ratings **below** 1000.

In Row 2A, Column G reports the [6.19%] percentage of graded or listed SWs produced from all listed SWs in our sample (thus aggregating the results for both categories of speed figures reported in rows 2A and 2B).

The reader may ask why we did not report separate percentages for graded SWs produced from listed SWs. The answer relates to a "small sample" problem. For all categories but GSW dams, the number of GSW foals produced was six or less. Statisticians advise that interpreting results based on small samples should be done so with caution. Statisticians would argue that the results for Row 1, which recorded the produce records of GSWs, are definitive. GSW dams with summer sales pedigrees bred to top 47 stallions produce GSW foals at high rates and the 10%+ (14 GSWS from 139 Graded Stakes Winning mares) strike rate is a relatively reliable estimate of the results a breeder can expect.

In contrast, looking at the listed SWs, the 3 for 72 strike rate of SR>1000 listed SWs (4.17%) is a less reliable estimate of likely future results. The small sample of 72 and the small number of GSWs produced (three) should leave the interpreter of the results wondering whether two or four (or even one or five) out of 72 might not be reasonable expectations for the success rate.

A Comment on the Percentage of Stakes Winners

Because we focused on the question of producing the very highest quality runners, our focus was on breeding graded or listed SWs, hence our results report only the graded and/or listed SWs produced by our sample of mares. If this focus is not duly noted, the reader may look at Table 1 and misinterpret our results to conclude that only five to six percent stakes winners are produced from the high-quality mares in our sample. Using a representative subsample (to save time) of the runners produced by the mares in our sample, we estimate that the composite numbers would be slightly more than 50% higher had we included all stakes winners in the calculations of percentage stakes winners. Thus we estimate that the "other" winning mares, which produced 5.99% [graded + listed] SWs, produced approximately 9% stakes winners if non-graded and non-listed SWs are included. This additional 50% or more estimate probably does not apply across the board. The fillies out of the top echelons of dams are probably retired early more often rather than race until five or six in pursuit of non-listed black type. In contrast, the fillies out of lower echelon mares tend to race longer in pursuit of any kind of black-type, and thus achieve non-listed stakes-winning status more often relative to their success in graded or listed stakes, with the result that their percentage of lower-level SWs is enhanced relative to graded and listed SWs.

A Row-by-Row Analysis of Table 1

Analysis of Row 1

The results: Graded SW mares had a very high [14 GSWs from 139 foals] strike rate in producing GSWs. And, the overall stakes-winning percentage of foals of 13.67% was more than double that of any other category of mares save the "other SW" category. These results are instructive, but it is useful to look behind the numbers.

Behind the numbers: Two of the 47 stallions in our sample accounted for a remarkable nine out of the fourteen GSWs produced from GSW mares. Storm Cat, with six GSW runners from twenty-five opportunities, and A. P. Indy, with three GSWs from nine opportunities, had a combined strike rate of 26% GSWs when bred to GSWs from our sample of mares. One might argue that the sagacity of Bull Hancock's admonition to breed the best to the best could hardly have better support than these results.

But, recall that foals in our sample were all by one of our "top 47" stallions. Surprisingly to us, with the exception of the Graded SW dams in our sample, every category of mare was bred to similarly-ranked stallions on average. As reported in Table 2 below, the "sire-power[8]" rating for all categories other than graded SW dams fell in the

[8] The "sire-power" ranking is calculated by taking each individual foal that was used in our sample and assigning a rank (from 1 to 47) to it based on the rank of its sire in our sire list. The sum of all the foal rankings in each broodmare category are added up and then divided by the number of foals in a category. For instance, if half the 139 dams in the GSW category were bred

narrow range between 23.8 and 26.8. That graded SW dams were bred to substantially better sires on average, with a sire power rating of 15.6, was not a surprise. But, we were surprised that the remaining categories of mares, most notably the Listed SWs, were bred to similarly-ranked stallions.

Table 2: Mare Categories by "Sire Power" Rankings

Graded SWs	15.6
Listed Sws	24.0
Allowance	23.8
All Others	26.1
Non-Winners	26.7
Unraced	26.8

A goal of this study is to produce an analysis of breeding results that is adjusted for differences in the quality of sires to which mares are bred. As Table 2 demonstrates, with the exception of the graded SWs category, our efforts at making this adjustment have worked quite well, though with one exception. We can look at all categories other than the graded SW dams and say that there is likely to be little noise or bias in the results because the mares were bred to stallions of differing quality.

But, because the graded SW dams were bred to much better sires on average, their results are biased upwards relative to the other categories of mares that were bred to lesser stallions. This begs the question "What results would have been expected if the graded SW mares had been bred to the same group of stallions that the other categories of mares had been bred to?" Most observers would agree that the direction of the expected result is obvious: The production record of the graded SW mares would have dropped had they been bred to lesser stallions. But how much the drop would be is a question on which we would expect estimates by horsemen (and statisticians?) to disagree. The

to the no. 1 sire, and half were bred to the no. 5 sire, we would have 69.5 x 1, + 69.5 x 5, or 417 as the sum of the foal rankings. Dividing the 417 by 139, the no. of foals in the sample, produces a weighted average of 3. This indicates that on average, the mares in the category were bred to the third best sire on the list. See Table 2 for the actual "sire power" rankings of stallions bred to each category of mares.

figure reported in Row 1, Column G is an estimate by the primary author (Losey) of this report that is explained in Appendix A.

Analysis of Row 2

There are three stories to tell regarding Rows 2A & 2B. The first story is that listed stakes winners with higher speed ratings in their fastest race produced more graded and listed SWs. But, the 8.33% overall strike rate reported in column E for listed SWs with SR>1000 is **only** the fourth-highest strike rate in the Table. The surprise to us is that it is not higher.

A second "story" has to be qualified. The slower listed SWs in our sample produce fewer, in fact they produce zero, SWs. But there is a small-sample problem here. There were only 25 foals in this sample, and the zero number of SWs produced is an unreliable estimate of the results one can expect. (See a following section for more details on the statistical significance and the reliability of estimates). Nevertheless, the direction of the results are consistent with the results for the other mare categories analyzed, and we have no doubt that the speed of the dam is a significant contributor to the speed of the foal. A corollary to this that our results tend to support is that speed may be more important than racing classification.

A related third story that can infer should be told is that being a stakes winner may be over-appreciated in the marketplace by the buyers of broodmares and broodmare prospects.

Finally, we note from row 2A, column F that the combined strike rate for all (both faster and slower) listed SWs is barely higher than the combined strike rate for the allowance, other, and unraced categories.

Analysis of Row 3

The reader may have already noted that we have virtually ignored this dam category in previous discussions. Perhaps this is unfortunate; as this category of dam produced stakes winners at the second highest strike rate in Table 1. Why have we ignored this category? There are two reasons: 1) The Jockey Club uses this category as a catch-all category. "Other" SWs includes runners with widely differing abilities. Foreign graded SWs fall into this category, but so do non-listed North American SWs. Analysis of this category is further complicated by the "small sample" problem previously alluded to: The sample size of 44 makes conclusions about this category less reliable.

Analysis of Row 4

The results inferred from an analysis of the allowance winning dams and their progeny reported in row 4 is similar to that for row 2. 1) Speed is important: Faster allowance winners produce more graded and listed SWs. 2) Allowance-winning dams produce graded/listed SWs at a rate that is only slightly below the rate of SWs for listed stakes –winning dams.

Analysis of Row 5

Rows 5 and 7 provide what are the most surprising results to us. Row 5 reports progeny racing results of dams that did no more than win maiden, starter, or claiming races. For both rows 5 and 7 we have reasonably good-sized samples, thus providing reason to be confident in the results. Row 5 results reinforce the results from Rows 2 and 4 that speed of the dam is an important contributor to speed in her foals. Again we have a small sample problem for the group where SR>900, but the continuing consistency of the results reinforces the results from Rows 2 and 4. In the following section that discusses statistical significance we show that "speed" is highly significant when the three categories where speed was broken out are considered together.

Note that the speed rating breaks between "faster" and "slower" are 100 points lower in row 5 than in rows 3 or 4. We suspect that many of the dams analyzed here showed promise and ran an impressive race at ages two or three before either being injured or before their owners decided they were more valuable as broodmares than as racers. Perhaps they did not stay on the track long enough to show their true abilities.

Analysis of Row 6

The results from row 6 do little other than to suggest that one should be very wary about relying on non-winning mares as potential producers. Recall from Table 2 that the non-winning mares in our sample were bred, perhaps we should use the term "overbred" to stallions with an average "sire power" index that was similar to the sire power index for listed SW mares and allowance mares. The strike rate of 2.11% graded and listed SWs is barely 40% of the next lowest category, leaving little to commend a mare that raced but did not win as a potential broodmare.

Analysis of Row 7

As suggested in the discussion of row 5, the production records of unraced mares surprise us, but in this case the surprise is on the upside. The reported strike rate for unraced mares is only modestly below that of listed SWs. Given that this type of mare can often be bought for much less than a listed SW (see later Table), this may be a more cost-effective approach to breeding top-quality runners.

The authors of this paper have discussed the Row 7 results at some length. We suspect that there is often a "story" to explain these results. The mare may have had great workouts before being hurt. The mare's pedigree may be superior. Perhaps the owners of unraced mares choosing to breed to a top sire are, as a group, more astute judges of horseflesh than most, and are more thoughtful than usual in any variety of ways, including being careful about choosing the appropriate stallion

Further Discussion of Table 1

Table 1 reports the numbers and percentages of graded and listed SWs produced by summer sale graduates bred to our "top 47" stallions. Note that we **do not** include non-listed SWs in the reported figures for foals in Table 1, and that foreign SWs of all kinds are not reported for foals in this report. If we had, the reported percentages of SWs produced would probably have been 35-50% higher.

We find it useful to classify our comments on Table 1 results using three categories:

A. What we* were pretty sure we already knew.

B. What we suspected we knew.

C. What surprises us*.

*When the terms "we" and "us" are used above, they refer to the authors of this report.

A. **What we were pretty sure we already knew.**

1) Speed is the breeder's friend. When we stratified the listed SWs, the allowance runners, and the "other" runners by speed, we consistently found that faster dams produced graded or listed stakes-winning runners more frequently. As the next section reports, the combined results comparing speedy dams to slower dams reveals that the speed of the dam is highly statistically significant in predicting the percentage of G/LSWs.

2) GSW dams are bred to much better stallions on average than other mares.

B. **What we suspected we knew**.
3) GSW dams produce GSW runners with much more frequency than any other category of dams, even when adjustments are made for the quality of stallions to which they are bred.
4) Non-winning starters (even when bred to top-47 stallions) don't make very good broodmares on average. This supports the hypothesis that race fillies that (more or less) prove that they are slow typically make poor broodmares.

C. **What surprises us.**
5) The differences in the average "sire power" of "top 47" stallions that were bred to the summer sale graduates we analyzed that became listed SWs, allowance mares, "other" SWs, and unraced mares were not significant, with all four groups having been bred to stallions on average that rank from 24th to 27th on our list of top 47 stallions.

6) "Other" winners (winners that have won no better than a maiden, starter or claiming race) that are summer sales graduates do quite well (even relative to listed stakes-winning summer sales graduates) when bred to top-47 stallions.

7) Unraced mares bred to top-47 stallions also do quite well. We suspect that the unraced mares that breeders choose to breed to top stallions are often "story" mares that showed promise but did not make it to the races.[9]

8) Well-pedigreed listed SWs only very marginally outproduce well-pedigreed allowance mares, "other" winning mares, and unraced mares. Based on this study, we suspect that listed SWs are usually "overlays" in the sales marketplace, and that "underlays" are more likely to be found among speedy winning mares who did not win stakes.[10]

---

[9] As an example of a "story" horse, the lead author once sat with a friend who bought a broodmare in foal to Vice Regent before that stallion had any foals to race. Being quite aware that Vice Regent had a weak race record, I asked my friend why he would spend money on this mare. He told the story that he was working at the track when Vice Regent outworked stakes horses shortly before being injured. My friend's new mare produced a stake-placed Vice Regent foal, and Vice Regent became a top ten sire.
[10] To digress, it has been the experience of the lead author of this study that non-listed SW mares typically bring more relative to comparably pedigreed non-SW mares than can be justified by almost any logic. We provide some evidence on the related effect for listed SWs later in this report, but a definitive proof must wait for another study.

TABLE 1S

Frequency of Graded/Listed SW Runners from Keeneland Summer – **Tests of Significance**

Sales Graduates Bred to "Top 47" Stallions: Dams Arranged by

Racing Class and Adjusted for Sire Quality

A. Dam Category	B. No. of Foals	C. Graded SW Foals	D. Listed SW Foals	E. % Foals That Were Gr/L SWs	F. % of G/Listed SWs Ignoring SR	G. % of G/Listed Foals Produced: Adjusted Estimate
1) Graded SW	139	14	5	Composite-->	13.67%	12.99%** (See Appendix)
2A) Listed SW: SR>1000*	72	3	3	8.33%	6.19%	<-Composite % for
2B) SR<1000	25	0	0	0.00%		All Listed SW Dams
3) Other SW	44	3	2	Composite-->	11.36%	
4A) Allowance: SR>1000	118	3	6	7.63%	5.86%	<-Composite % for
4B) SR<1000	104	4	0	3.85%		All Alw Dams
5A) Other Wnrs: SR>900	31	3	0	9.68%	5.99%	<-Composite % for
5B) SR<900	136	5	2	5.15%		Other Wnng Dams
6) Non-Winners	284	3	3	Composite-->	2.11%	
7) Unraced	188	6	4	Composite-->	5.32%	
Total foals	1140					

The Statistical Significance and the Observed Results

As explained earlier in this paper, the conditions for inclusion in this study resulted in the number of foals in the final sample being limited to just 1,140. A sample size of 1140 is ample by the standards often used in statistical analyses[11]. However, because we broke the data down into ten subsamples (see the ten dam categories in Table 1S above) a number of the subsamples had a small number of foals in them. One concern when working with small samples is that the reliability of the results with respect to drawing inferences is diminished. The smallest subsample from Table 1S is Category 3,

[11] Polls for national elections often sample approximately 1100 potential voters.

"Other Stakes Winners," and it is useful to consider this category in analyzing the question of statistical significance.

Consider for example a comparison of the percentage of G/LSWs produced by the "Other SW" category (Category 3) to the percentage of G/LSWs produced by largest other category, that of "Non-Winners" in Category 6. As measured on a percentage basis, the "Other SW" category has a substantial edge in producing G/LSWs: The percentage of G/LSWs produced by Category 3 mares was 11.36%, while the percentage of G/LSWs produced by Category 6 mares was only 2.11%. But can we be sure that the difference is not caused merely by chance? Based on a standard statistical test the answer is that the difference in results is probably real[12]. The likelihood that the mares in the two categories had the same inherent ability to produce G/LSWs is very low at approximately 3%. Said another way, and framed in a more general context, the statement that "Category 3 is better/more effective at … than Cateory 6" will be a true statement 97% of the time (statisticians would say that their "confidence level" is 97%) with respect to a series of analyses where samples produce the same or analogous results to those reported in Table 1S.

However, when we compare the Other SW category to every other category we can not reliably say that the Other SW category is significantly different in its ability to produce G/LSWs. The comparison of Category 3 with Category 6 had two factors in its favor that are **not** observed in making any other comparison:

1) The Category 6 sample size was larger than any other category in Table 1S.

2) The difference in the result (percentage of SWs) was greater for the comparison of Category 3 with category 6 than for the comparison of Category 3 with any other category.

For instance, a comparison of Category 3 with Category 4A generates a "confidence level" of 75%, while the comparison of Category 3 with the combined 4 and 4A results generates a confidence level of 86%, and no comparison to any other category generates a confidence level above 86%. Though there is some arbitrariness in their

---

[12] The statistical approach we use can be found in a many texts and references. We relied primarily on a section titled "Hypotheses About Two Means" as reported on pp. 845-853 from Marketing Research 6th edition, by Gilbert Churchill.

choices, statisticians normally use 95% as the break point for the value of the confidence level that reflects "statistical significance," though they also sometimes cite 90% as an appropriate breakpoint.[13]

What are the practical implications of the absence of statistical significance when comparing the Other SW category with all but the Non-Winners category? The answer is that we cannot be highly confident that mares that were in the Other SW category are superior producers of G/LSW relative to the other categories. Can we be somewhat confident? Most statisticians would respond with a "yes, but" answer. If you are comfortable with saying "I'm confident" when you will be wrong 14% of the time (see previous paragraph), then you would feel comfortable saying "I'm confident that Other SW mares are better producers than allowance mares." But of course there is a 14% chance that you are wrong. Most statisticians would not go as far as to say "I'm confident" in such a case. Instead, they might say something like "The odds are that Other SW mares are better at producing G/LSWs than allowance-winning mares."

The bottom line with respect to the Other SW category is that, though it is likely that such mares produce G/LSWs more frequently than all other categories except graded stakes winning mares, the small sample size means that we cannot opine that that is the case if we wish to follow the guidelines (90 to 95% confidence levels) traditionally used by statisticians.

Other Category Comparisons

Does Speed Make a Difference? Before we turn to comparisons we can make "with confidence," consider one more borderline situation. Note that this study broke down three categories, listed SWs, allowance winners, and other winners, into two speed categories. Allowance winning mares produced the second largest sample size of any category with 222 total runners produced, with runners breaking the 1000 speed rating barrier numbering 118, while runners whose fastest time was below 1000 numbered 104. The test of significance comparing these two subcategories generates a confidence level of 86% that the observation that speedier mares produced G/LSWs with higher frequency than the slower mares is reliable.

[13] The implication is that anything between 95% and 90% can be defended as a reasonable cutoff for significance.

Comparisons of the other two categories broken down by speed yielded comparable results. There is evidence in all three categories that faster mares produce more G/LSWs, but the evidence in any one category is not **highly** significant using the traditional statistical breakpoint.

However, all is not lost. If we combine the produce records of all the speedy broodmares from the three categories and compare them to those of all the combined slower broodmares from the three categories, we have a larger sample and the comparison of the "strike rate" of producing G/LSWs from the speedy category to the slower category results in a confidence level of greater than 96%. The conclusion that can be drawn then is that **speedy broodmares produce more G/LSW foals**. Joe Estes would be happy.

Graded Stakes-Winning Mares produced G/LSWs at higher frequencies that were always statistically significant when compared to any of the other categories (excluding the Other SW category, which we drop from this and all further comparisons). The conclusion is rather obvious: GSW mares are more likely to produce G/LSWs than are non-GSW mares.

Listed SWs compared to Allowance winners, as well as Listed SWs compared to Other Winners or to Unraced mares provided interesting food for thought: **We do not have a sufficient statistical basis to conclude that the Listed SWs produce a higher percentage of G/LSWs than Allowance-winning, "Other Winning," or Unraced mares.** Confidence levels regarding differences in the means of these categories were always in the vicinity of 35%, a level that cannot be viewed as significant by any stretch of the imagination. Two of the comparisons were especially noteworthy.

1. Listed Stakes Winning mares were only marginally better producers of G/LSWs than any of categories 4, 5, and 7 (allowance, other winners, and unraced mares). The difference in each case was not statistically significant,
2. Unraced mares exhibited only slightly lower strike rates than listed SWs, allowance winners, and the other winners category.

We were somewhat surprised by these results, especially in the case of the listed SWs, which we had expected would be better producers than all non-stakes winning categories. We lean toward an explanation that the market tends to overvalue (and overbreed) listed SWs, and that breeders do not appropriately appreciate the importance of superior speed on the racetrack.[14]

We were less surprised by the production records of the unraced broodmares. Our expectation is that the owners of these mares viewed them as especially promising mares that were not able to show their ability in a race. Perhaps they worked fast works but had injuries. Breeders thought enough of unraced broodmares in our sample to breed them to good stallions, and the evidence suggests that the breeders confidence in these prospects were largely well founded, as they produced G/LSWs almost as frequently as Listed Stakes Winning mares.

Non-Winning Mares are inferior producers of G/LSWs. In comparing the production record (2.11% G/LSWs) of the non-winning mare category to all other groups, we found that their production record was inferior to all other categories, and the difference was always statistically significant. The confidence level for every comparison of the non-winners group with any other was 94.9% or higher in every case. Our conclusion is a corollary to our conclusion regarding speedy mares (see above). **Demonstrated lack of speed is a significant negative factor with respect to the ability of a mare to produce G/LSWs**. We have already reported that speedy winning mares are more likely to produce good runners. What the results for the non-winning mare category would seem to indicate is that the demonstrated inability to be able to run fast enough to win a race should be viewed as an indication of inability to produce G/LSWs relative to winning (or even unraced) mares. Thus we would argue that **The Non-Winning Mare category, like the Listed SW category, appears to be overvalued by the breeders in our sample, who overbred this category of mares** based on our statistical analysis.

Top-47 Stallions: Table 3

Table 3 below lists the sires bred to summer sales graduates that were utilized in this study. From the 7000+ foals produced from the summer sale graduates for which the Jockey Club provided data, we used all North American-based stallions that produced

[14] Perhaps a more subtle explanation is warranted. It is possible that breeders understand the importance of speed, but they expect that buyers will not appropriately appreciate speed. Thus breeders might feel that they can sell foals from slower stakes-winning mares at prices that are little different from the prices that foals from speedier stakes-winning mares fetch.

more than five foals from the 1500 summer sale fillies that were the potential dam pool for this study. To make the sire list a sire also had to have a graded SW % of 2.77% or greater. By limiting our usage of stallions to this category, we simultaneously decreased the pool of foals that were analyzed and, more importantly, decreased the degree of heterogeneity of the quality of the foals.

A useful way to look at what we tried to do in minimizing this heterogeneity is to say that we limited the heterogeneity

1) of the female-line pedigrees by using only dams that were in the middle 80% of the sales prices.

2) of the sire line of the foals by limiting the analysis to foals by "top-47" sires.[15]

After trying to limit the heterogeneity, we then purposefully introduced heterogeneity by considering seven categories of racing class (plus speed figures in three categories) for the dams analyzed in this study. This was in fact a primary goal of the study: Analyze the produce records of mares that differed primarily by racing class and determine the effect of racing class and speed on the race records of their foals. From this we proceeded to try to determine whether or not there might be niches where buyers of mares will be more likely to find reasonably priced racing prospects.

See Table 3 on the next page

[15] To a degree, limiting the sires of foals to a select group probably also further mitigated toward greater heterogeneity of female-line pedigrees, as most breeders will try to match female pedigree quality with sire quality.

Table 3: Top 47 Stallions Analyzed by Percentage of Graded Stakes Winners

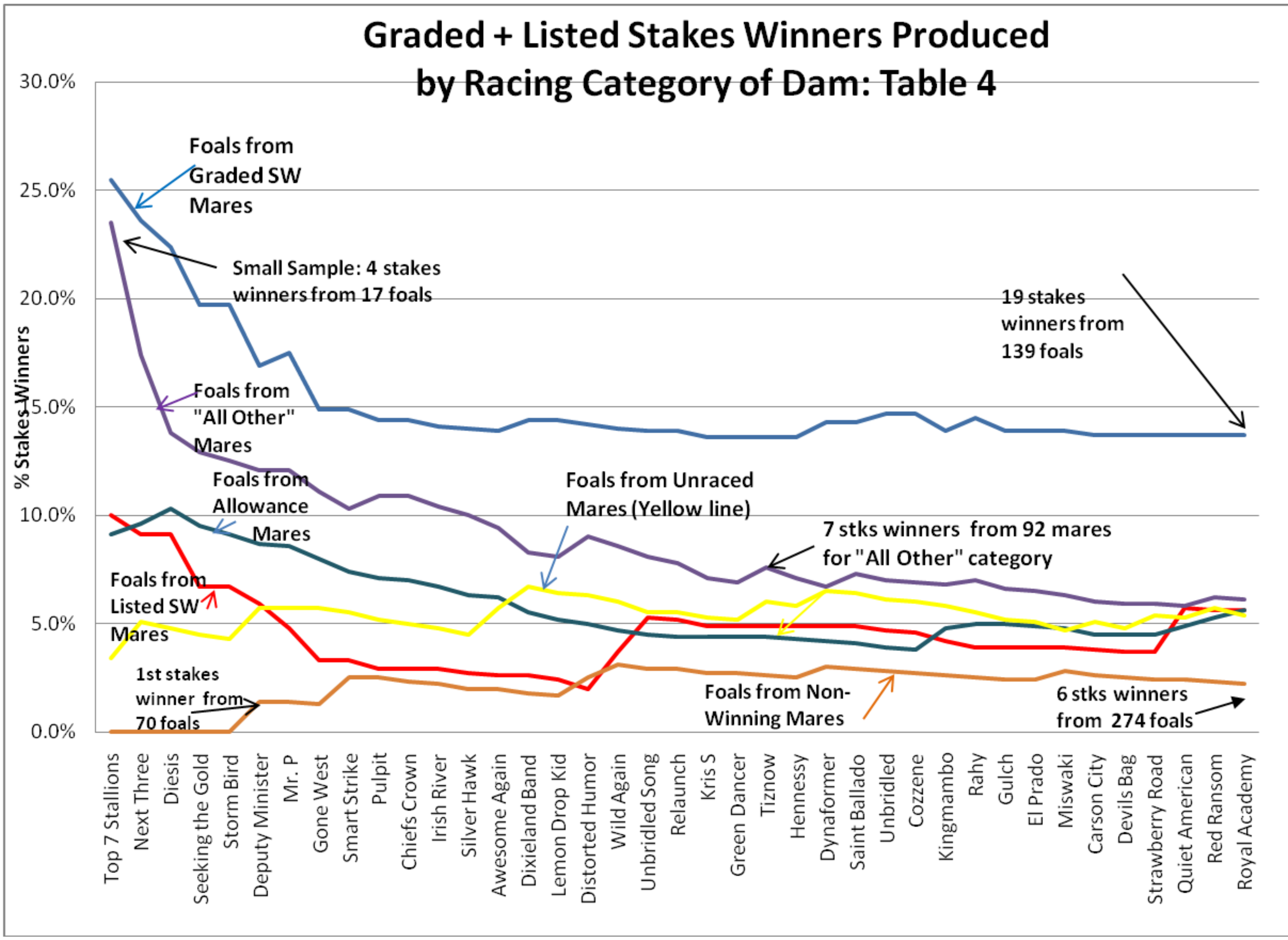
Sires	Born	AEI	CEI	SWs	Foals	GSWs	SWs/ Foals	GSWs/ Foals
Storm Cat	1983	3.7	3.7	114	782	73	0.1457	0.0933
Nureyev	1977	4.4	4.1	100	517	42	0.1934	0.0812
AP Indy	1989	3.1	3.6	127	952	77	0.1334	0.0808
Seattle Slew	1974	5.0	5.0	66	526	38	0.1254	0.0722
Danzig	1976	5.5	4.5	119	605	41	0.1966	0.0677
Affirmed	1975	3.0	2.8	58	460	29	0.1260	0.0630
Alleged	1974	2.9	3.1	69	618	37	0.1116	0.0598
Machiavellian	1987	2.3	2.1	33	302	18	0.1092	0.0596
Pleasant Colony	1978	3.8	2.6	62	437	26	0.1418	0.0595
Giant's Causeway	1997	1.9	3.3	60	678	40	0.0885	0.0590
Diesis	1980	2.2	2.4	70	574	32	0.1219	0.0557
Seeking the Gold	**1985**	2.7	3.4	71	637	34	0.1114	0.0533
Storm Bird	1978	2.7	2.8	57	568	30	0.1003	0.0528
Deputy Minister	1979	3.7	3.5	64	585	30	0.1094	0.0512
Mr. P	1970	5.1	3.2	109	628	31	0.1735	0.0493
Gone West	1984	2.2	2.5	76	772	38	0.0984	0.0492
Smart Strike	1992	2.6	2.0	68	652	32	0.1042	0.0490
Pulpit	1994	2.0	2.7	42	489	24	0.0858	0.0490
Chiefs Crown	1982	2.6	3.7	28	296	14	0.0945	0.0473
Irish River	1976	2.4	2.6	57	590	27	0.0966	0.0457
Silver Hawk	1979	2.5	2.2	53	498	22	0.1064	0.0441
Awesome Again	1994	2.4	2.2	35	502	22	0.0697	0.0438
Dixieland Band	1980	2.5	2.5	83	690	30	0.1202	0.0434
Lemon Drop Kid	1996	2.0	2.2	44	369	16	0.1192	0.0433
Distorted Humor	1993	2.3	2.1	70	578	24	0.1211	0.0415
Wild Again	1980	2.4	2.1	67	663	27	0.1010	0.0407
Unbridled Song	1993	2.0	2.5	72	765	31	0.0941	0.0405
Relaunch	1976	3.4	2.4	63	430	17	0.1465	0.0395
Kris S	1977	2.5	2.0	40	418	16	0.0956	0.0382
Green Dancer	1972	2.3	2.4	58	601	23	0.0965	0.0382
Tiznow	1997	2.2	2.0	21	341	13	0.0615	0.0381
Hennessy	1993	1.7	2.0	37	534	20	0.0692	0.0374

Dynaformer	1985	2.1	1.7	69	676	24	0.1020	0.0355
Saint Ballado	1989	1.9	1.8	27	255	9	0.1058	0.0352
Unbridled	1987	3.3	3.4	21	235	8	0.0893	0.0340
Cozzene	1980	3.0	1.8	48	443	15	0.1083	0.0338
Kingmambo	1990	2.5	2.8	84	782	26	0.1074	0.0332
Rahy	1985	2.4	2.1	66	733	24	0.0900	0.0327
Gulch	1984	2.1	2.5	54	680	22	0.0794	0.0323
El Prado	1989	2.1	1.7	63	684	22	0.0921	0.0321
Miswaki	1978	2.3	2.2	71	750	24	0.0946	0.032
Carson City	1987	2.0	2.0	71	692	22	0.1026	0.0317
Devils Bag	1981	2.1	2.9	38	665	21	0.0571	0.0315
Strawberry Road	1979	3.3	1.9	8	134	4	0.0597	0.0298
Quiet American	1986	1.7	1.8	39	514	15	0.0758	0.0291
Red Ransom	1987	1.7	2.1	63	928	27	0.0678	0.0290
Royal Academy	1987	1.6	1.7	85	1080	30	0.0787	0.0277

An Explanation of Table 4

Table 4 (See next page) is related to Table 1 in the following way. Recall that Table 1 reported the graded and listed stakes winning foals produced from mares in each category. The far right hand side of Table 4 provides precisely that same information. However, Table 4 provides the following additional information: Table 4 gives us a stallion-by-stallion record of foal racing performance, while Table 1 only gave us the composite results for each category using all 47 of the "top 47" stallions. For example, consider the start of the six lines on the left side of the diagram (The "Other SWs" category is omitted because the sample is so small). The results at the start of each line represent the average of the record for foals in our sample by the top seven sires in Table 3 above (Storm Cat through Alleged). We reported the results for the top 7 stallions so that the starting point on the graph would represent more than a tiny percentage of the foals. Reporting the top 7 stallions has two effects: one useful, and one somewhat misleading.

One effect of combining the results of the top 7 stallions is to decrease the extreme variability in the lines on the chart if the tiny sample of mares for each of the first seven stallions were plotted individually. A second effect is to treat the top 7 stallions as if they were one stallion for purposes of the diagram, thus condensing the results for the first seven stallions into one point. This, and the related decision to include three stallions in the second grouping, condenses the left side of the diagram and results in the observed drop-off in the line being more dramatic than it would have been had each of the first ten stallions been plotted separately.



The top 7 stallions produced GSWs from graded SW dams at a phenomenal 26% strike rate. Perhaps that is only somewhat surprising, but whatever your view on that score, the numbers for the second highest line are nothing if not surprising. The top seven stallions sired 17 foals from dams in the "other" mares category (recall that other mares are mostly maiden winners). Four of the 17 were graded SWs. Again, we must be wary; this is a small sample, but four out of 17 is so successful that one has to wonder how this result was achieved.

Note that the strike rate for both graded SW mares and "other" mares drops off quickly after the first ten stallions. It does bounce up with Mr. P(rospector), but next drops dramatically with Gone West, who for whatever reason had no graded or listed SWs from 41 foals in the sample).

Somewhat anomalously, and we suggest that not too much credence should be read into this because of the small initial sample, the results for non-winning mares bred to top stallions results in no graded or listed SWs among the 59 foals sired by the best 13 stallions. For this category we finally observe that a (listed) SW results from among the eleven foals sired by Dixieland Band, the fourteenth-rated sire. The remaining 200 or so foals in this category had a strike rate of one graded/listed SW every 34 foals, which still only brought the final average up to the 2.11% strike rate.

With the exception of the non-winning mares (and perhaps also for the very high strike rate for both graded SW dams and more especially "other" dams when bred to the top sires), the results we observe from Table 4 are in the ball park for what we might expect; As mares are bred to lesser stallions, their strike rates normally go down, though there is substantial variation in the data. But many breeders would argue that the variation in the data is part of what makes the breeding business challenging, interesting, and appealing to the breeder with limited means who may feel (s)he has a chance to hit the big time.

Other Data of Interest

Sale Price Analysis of Mares sold as Broodmare Prospects: Mares Sold In-Foal

Mare Category	Average Price	Median Price	Sample Size
Graded Stakes Winners	$409,000	$550,000	6 mares
Listed Stakes Winners	$301,571	$250,000	8 mares
Allowance Winners	207,950	$197,500	21 mares
All Others Winners	$204,204	$125,000	34 mares
Non-Winners	$157,576	$90,000	49 mares
Unraced	$145,292	$72,500	23 mares

The chart above was constructed by analyzing all mares from the original sample [mares that were originally sold at the Keeneland July select sale] that later resold as broodmare prospects carrying their first foal. Our original sample size was slightly over 1100 mares. From this original collection, 141 mares went on to be sold at five or six years old as broodmare prospects carrying their first foals. These results attest to the fact that graded and listed stakes winners are expensive. Allowance and "other" winners, which have strike rates approaching those of listed stakes winners, cost significantly less than listed stakes winners, and unraced mares cost much less.

Original Purchase Prices for Keeneland Summer Sales Mares by Category

Prices for Yearlings Purchased at Keeneland Summer Sale by Subsequent RacingCategory

Mare Category	Average	Median
Graded Stakes Winners	$234,835	$210,000
Listed Stakes Winners	$193,125	$170,000
Allowance Winners	$241,592	$225,000
All Other Winners	$249,061	$200,000

The prices above represent prices paid for yearling fillies at the Keeneland Summer Sales that eventually became part of the broodmare sample that produced the 1140 foals that were the subject of this study. The higher prices observed for allowance and "other" winners may appear surprising, but upon reflection, they are not. Remember that to be in our sample, a summer sale graduate had to produce a foal that was the result of being bred to a "top-47" sire. We would expect that breeders would breed almost every listed stakes winner with an original summer sales pedigree to a top-47 sire, but in contrast, would breed only the better, perhaps the top pedigreed allowance and "other" winners to a "top-47" sire. Hence only allowance winners that originally sold for higher prices would be more likely to end up in our sample. The overall average price of the fillies that eventually became allowance winners was probably below that of the average price of the fillies that became listed SWs, but the lower-priced allowance and "other" winners would **NOT** be bred to the very best stallions. Hence the prices above reflect the prices of only a fraction of the sample of foals from allowance and "other" mares produced from mares in our sample that originally went through the Keeneland Summer Sale. Summer sales graduates that were originally lower priced, and then subsequently won no more than an allowance race would tend to be bred to lesser quality stallions.

Higher-priced individuals that subsequently became allowance winners would be more likely to be bred to top stallions because the higher prices in many, perhaps most cases, reflected better conformation and/or pedigrees that resulted in them being bred to better stallions.

A Brief Summary

We have devised a research methodology that allows us to come much closer than previous studies to standardizing pedigree and the quality of sires to which a group of mares was bred. As a result, we are better able to isolate the importance of the racing class and speed of mares as it affects the ability of dams to produce high class runners. We find that speed in the dam is a very important factor in contributing to racing success if racing success is defined producing foals that win graded and listed stakes. Conversely, dams that show a lack of speed, which we consider to be demonstrated by runners that raced but could not win a race, is significantly negatively correlated with the racing records of these dams' progeny.

In addition to the importance of speed, we find especially interesting results with respect to four categories of mares: the aforementioned category of non-winners, and additionally, graded stakes winners, listed stakes winners, and unraced mares.

Graded stakes winning mares were distinguished by their ability to produce graded and listed stakes winning progeny. Even after adjusting for the quality of the sires to which they were bred, graded stakes winning mares produced graded or listed stakes winning foals at more than twice the rate of any of the following: listed stakes winners, allowance winners, other winners, and unraced mares. They produced graded or listed stakes winners at six times the frequency that comparable non-winning mares produced.

We were surprised to find that **listed stakes winning mares** performed only marginally better as broodmares than allowance winners, other winners, and unraced mares. However, our limited analysis of sales prices revealed that listed stakes winning mares sold for between 50% and 200% higher on average than these other categories of mares with similar pedigrees. We suspect there is a lesson here: Prefer to be a seller rather than a buyer of listed stakes winning mares. Our nominee for an exception to this rule would be (and here we venture more than a bit away from a well-supported statistical analysis) the especially speedy listed stakes winner who did not have a chance to prove herself against graded competition or who had bad luck when she did try these types.

Two other mare categories may have similar stories: "other winning mares," mostly maiden-only winners, and unraced mares both produced graded or listed stakes winners with almost the frequency that listed stakes winners did. We suspect that a substantial percentage of these mares, especially the unraced mares, were "story" mares that showed substantial promise but had no or little chance to fulfill that promise due to accident, injury, or temperament.

With respect to stallions, we find (see Table 4) that breeding to the top eleven to fifteen stallions on our list of the "top 47" was associated with a substantially higher frequency of production of graded and listed stakes winners.[16] As less than two-thirds of these stallions were available in a given year of our study, this implies that a breeder would have to breed to a top ten stallion each year to achieve comparable results.

For a more detailed discussion of our findings see the sections of this paper that discuss Tables 1-4.

Appendix A and Appendix B follow

Appendix A
Equalizing Sire Power:
Determining the Percentage of Graded/Listed (G/L) SWs expected if the GSW mares had been bred to sires with the same sire power as other mare categories.

In this study we were fortunate in that all categories of broodmares except the GSW mares were bred to sires with similar "sire power."[17] The sire power factor fell in a (rather narrow) range that was from 23.8 to 26.8, averaging about 25. However, the GSW mares as a group were bred to much better stallions than the other categories, their average sire power being 15.6. Because of the higher quality of sires to which the GSW mares were bred it would be expected that this would contribute toward their producing a higher percentage of G/L SWs than if they were bred to the same quality sires as the other categories of mares. A question is, "*What would the expected percentage of G/L*

[16] The non-winning mare category is an exception, but perhaps this is attributable to the small sample. Moreover, since this category of mare has such poor overall success in producing graded or listed SWs, the importance of this exception is questionable, as foals from this category of mares generate the least success of the ten categories of mares considered in this study.

[17] "Sire power" refers to the ranking of stallions based on their percentages of graded stakes winners. We ranked 47 stallions with Storm Cat having the highest percentage of graded stakes winners at 9.3% while Royal Academy ranked 47th with 2.8% graded stakes winners.

SWs have been had the GSW mares been bred to the same quality sires as the other categories of mares?" (The actual percentage of G/L SWs produced by the 140 GSWs in the sample was 13.57%.)

We determined that a reasonable approach to answering this question, which we expected would result in a downward adjustment in the percentage of G/L SWs produced for the GSW mares, could be achieved through either of the approaches described below.

Adjustment 1.
Consider a sub-sample of the stallions to which the GSWs were bred that had an average sire power similar to that of the other stallions bred to other categories of mares used in this study. This could be done by dropping some of the mares bred to highly ranked stallions from the sample. As more mares bred to highly ranked sires are dropped from the sample, the average sire power of the remaining mares in the sample will be lowered. If this method is used it is important to use an unbiased approach that does not result in dropping an undue number of either stakes-producing or non-stakes-producing mares. One unbiased approach would be to randomly choose which mares are dropped from the sample. However, by chance this approach would often change the percentage of the G/L SWs produced by the smaller group of mares bred to high ranking stallions relative to the percentage from the (larger) original group.

The purpose of decreasing the number of mares bred to high quality sires considered is that this would move the average sire power of the sires bred to the remaining mares moving toward the (lower) average sire power of sires bred to other categories of mares. If enough mares bred to high ranking sires are dropped from the original GSW mare sample that the remaining mares were bred to stallions with sire power rankings similar to those of the other mare categories, this will remove the bias arising from being bred to different quality sires. By making this adjustment we place each category of mares on an equal footing with respect to the quality of stallions to which they were bred. As we were already using a relatively homogeneous group of mares with respect to pedigree, and have now equalized the "sire power" factor, we are left with the primary difference in the mares being their racing records, and thus we can attribute the variations in their produce records primarily to their racing records.

We considered the approach described above but preferred the related approach below.

Adjustment 2
Following the same logic used for Adjustment 1 above, rather than randomly choosing mares to drop from the sample, which could result in the sub-sample exhibiting a different percentage of SWs from the original larger sample (hence biasing the results

either upwards or downwards), we decided to use the full sample of mares, but to decrease the importance of the mares bred to the higher sire power stallions in the sub-sample by factoring down the weight attached to these mares when calculating the group's sire power. In effect we would treat each mare bred to a highly ranked sire as if she were a fraction of a mare. By using the appropriate weighting for the mares bred to the high ranking sires, we could factor the reported sire power down to the point where it is equal to that of the other mare categories. In effect this approach assumes that a "fraction of a mare" is bred. Of course this is physically impossible, but from a statistical point of view it not only does not create a problem, it allows us to preserve the original ratio of G/L SWs produced by the highly ranked stallions in this sample. In effect, instead of dropping individual mares from the original sample, we are dropping a fraction of each mare from the sample.

Further Explanation of the "Adjustment 2" Approach.

There were 35 GSW mares that were bred to stallions whose "sire power" was at or below 25. The average sire power of the sires bred to these mares was 37.47. There were 104 mares whose "sire power" was above 25. Their average sire power was 8.594. One way to figure the weighted average sire power for the stallions bred to GSW mares is as follows[18]:

Sire power index = (Low sire power mares bred x their average sire power +
High sire power mares bred x their average sire power)/total mares bred

Sire power index = (33 x 37.47 + 106 x 8.594)/(33+106) = 15.45 **Equation 1A**

Equation 1A reports that on average, GSW mares were bred to a stallion ranked between the 15th and 16th highest ranked stallions on our list of the top 47 stallions.

How can we estimate what percentage of G/L SWs the GSW mares would have produced if they had been bred to sires with average sire power ratings equal to 25, the approximate mid-point of the sire power range for other categories of mares from our study? The answer requires a reconsideration of Equation 1A. The relevant question is: "What smaller number (F) of mares bred to the high ranking sires would have resulted in an average sire power of 25? The problem we have is that too many GSW mares were bred to higher ranking stallions to allow us to make straight-forward comparisons with the other groups of mares that were bred to lesser ranked stallions. We solve this

[18] The weighted average sire power can also be calculated by solving the standard equation for calculated a weighted average, which is in this is (ΣSire power rating of stallions x no. of mares bred at each sire power level)/total mares bred = 15.45

problem by asking what would happen if we decrease the prevalence of mares bred to higher ranked stallions in the GSW mare category. As explained in the "Adjustment 2" approach, we do this by recalculating the percentage of G/L SWs produced by a sub-sample of the GSW that contains fewer mares bred to higher ranked stallions. Solving the equation to achieve this (a variation of Equation 1A), is done below.

(33 x 37.47 + F x 8.594)/(33+F) = 25 Equation 1B

This equation assumes that we consider a smaller number (F) of mares bred to the higher ranked stallions so that the impact of these mares weighs less heavily in the sire power calculations. F will be a smaller number than 106 that will cause the sire power of the new GSW mare sub-sample to be 25 compared to the 15.6 sire power ranking of the original sample. Solving Equation 1B, we get

(33[19] x 37.47 +8.594F)/(33 + F) = 25, =
1236.48 +8.594F = (33 + F) x 25, =
1236.48 + 8.594F = 825 + 25F =
1236.48 – 825 = 25F – 8.594F, or
411.48 = 16.406F, or
F = 411.48/16.406 = 25.08 mares bred to the high ranking stallions.

This is to say, that if we had bred 33 mares to stallions with an average sire power of 37.47, and 25.08 mares with an average sire power of 8.594, the average sire power of this group of mares would be 25.

Translating These Results into an Estimate of the Percentage of G/L SWs Produced

How does one translate these results into a new estimate of the percentage of G/L SWs that the GSWs would produce? First we look at the numbers of G/L SWs produced by the GSWs bred to the low (37.47) sire power sires. This number was four. Then we look at the number of G/L SWs produced by the GSWs bred to the high (8.594) sire power sires. This number was 15. Then we ask the question: "How many G/L SWs would have been expected if only 25.08 mares whose sire power was 8.594 had been bred? The answer is that, since originally 106 GSW mares were bred to high sire power sires and they produced 15 G/L SWs, we would expect that 25.08 mares would produce 25.08/106 = 23.66% as many G/L SWs as 106 did, or .2366 x 15 = 3.55 G/L SWs.

[19] We use 33 for the number of mares bred to lower ranking stallions because there were 33 bred to stallions ranked 25 or below. We then ask what number of mares bred to stallions with an average sire power of 8.59 will result in the sub-sample having a sire power equal to 25.

Finally, what percentage of G/L SWs would a group of GSW mares bred to stallions with an average sire power of 25 produce? We now can easily produce that estimate by utilizing our sub-sample. We had 33 GSW mares that were bred to low sire power sires, and they produced 4 G/L SWs. We had (theoretically) 25.08 GSW mares that were bred to high sire power sires, and if they would be expected to "produce" G/L SWs in the same ratio as the 106 actual mares bred to high sire power sires, they would produce 25.08/106 = 23.66% of the number of G/L SWs that the 106 high sire power mares produced. Since the 106 mares bred to high sire power sires produced 15 G/L SWs, the smaller group of 25.08 mares bred to similar sires would produce 23.66% x 15 = 3.549 G/L SWs if the same ratio obtained.

So, from our total sample of 33 + 25.08 mares we (hypothetically) produced 4 + 3.549 G/L SWs. The number of G/L SWs produced from the sub-sample was 7.549, which is 7.549/(33 + 25.08) = 12.99 percent.

Our estimate is that GSWs from our original sample, if bred to stallions with an average sire power of 25, would produce 12.99 % G/L SWs rather than the 13.67% actually observed. Because the G/L category ignores non-listed stakes, we would expect the overall percentage of SWs produced by GSWs bred to stallions with average sire power of 25 to be significantly higher than 12.99 %.

Appendix B: Was Toussand or Primal Force the Better Broodmare?

The following discussion taken from the original proposal for this study encapsulates many of the issues this paper addresses.

In many, and more likely in most cases, superior speed and success in graded stakes go hand in hand. Broodmare of the year Toussaud (by El Gran Senor out of Image of Reality by In Reality) is one of many marvelous examples of the combination of these two traits – she was a Gr. I winner who ran a 107 Brisnet speed figure. Moreover, she passed her abilities on to her progeny: five of her ten foals have been graded stakes winners.

Another broodmare of the year, Primal Force (by Blushing Groom out of Prime Prospect by Mr. Prospector), serves as an interesting counterpoint. Primal Force ran a 109 Brisnet speed figure and won four of eight races. However, she not only was **not** a graded stakes horse, she earned no black type of any kind. And, following her racing career, her owner made what in retrospect many would consider to be a series of mistakes in breeding her to underachieving stallions (Alwuhush, Siphon, Silver Charm, and

Golden Missile). Her one early breeding to a top-class stallion, Deputy Minister, produced Awesome Again, and she later produced Macho Uno to the cover of the very useful but hardly superlative Holy Bull.

Comparing Produce Records?

Contrasting the produce records of Toussaud and Primal Force illustrates the difficulty in determining the relative quality of mares. On the face of it, many would argue that Toussaud was the better of the two based on her record of five graded stakes winners from ten foals. Compare this with Primal Force's record of two graded stakes winners (albeit both were Breeders Cup winners) from eleven foals. But, how does one adjust for the fact that Toussaud was bred to a group of stallions whose median stud fee was in the neighborhood of $175,000 while Primal Force was bred to stallions whose median stud fee was less than one tenth of that figure?

In our view Primal Force's inherent quality as a broodmare was much better than an observer might surmise by merely looking at the number and percentage of high-class horses she produced. Was she as good as Toussaud? Possibly, but to our knowledge, no one has made appropriate adjustments that would allow for a valid comparison of Toussaud's and Primal Force's true qualities as broodmares. The approach we take in this paper can be used to shed some additional light on the relative strengths of individual mare comparisons such as Toussand and Primal Force, though any comparisons of two samples of one are fraught with difficulty. More importantly, while our analysis does not provide a basis to definitively determine whether Toussaud or Primal Force was the better producer, it is our view that our approach to evaluating broodmare prospects provides a significant step forward in the statistical analysis of the categories and the characteristics of mares that identify potential as broodmares.

Addendum

After rereading and editing this paper in preparation for submitting it to the *Thoroughbred Times* I found myself wondering if our study provided a useful basis to warrant a more definitive analysis than we present relative to the stallions breeders choose. I am inclined to think that we generated some useful food for thought, but rather than expand the paper in that direction at this time I invite readers to weigh in with their thoughts on this topic (and for that matter on any topic related to this paper). At some future date we hope to provide a forum for discussion of the responses and suggestions we receive regarding our analyses.

We wish to express our great appreciation for the dual roles that Dan Rosenberg played in fostering this study. Dan developed the original idea for the study, suggesting to us that there had to be a more logical and statistically defensible approach to building an

elite broodmare band than was available to breeders in the spring of 2009 when Dan first broached this subject with us. My colleague Tim Capps and I very much agreed, and with invaluable help from Papo Morales, a truly outstanding student in the Equine Industry Program at the University of Louisville, we set out to try to prove Dan right. Dan provided financial assistance in obtaining data from the Jockey Club (to which we also express our thanks) and in financing the project so that we could devote the many hours necessary to complete the paper in a reasonable time frame.
Finally we appreciate Dan's blessing in allowing us to share this paper with others who are interested in breeding and racing.

Robert L. Losey, Lead Author. February 2011