Breeders' Awards and the Gambler's Corner Solution

James Ahern and Jay E. Noel[1]

Paper presented at the Western Agricultural Economics Association Meeting Annual Meeting, Vancouver, British Columbia, June 29-July1, 2000

[1] James Ahern and Jay E. Noel are Professors in the Agribusiness Department, California Polytechnic State University, San Luis Obispo, CA. Contact person: James Ahern at jahern@calpoly.edu. Thanks are given to the California Thoroughbred Breeders Association for providing data and funding for this research.

California has had a Thoroughbred breeders awards program since 1933. The intent of the program is to encourage the production of Thoroughbred horses and provide an incentive for that production. The program rewards breeders of horses, stallion owners, and owners of registered California bred (CB) horses.

Many Thoroughbred racing states encourage the production of racing animals by incentives to producers or breeders of potential racehorses. California horse racing was reintroduced after a legal ban in 1933 by ballot initiative with a declared intent (§19401 (b)) of "encouragement of agriculture and the breeding of horses" (Deerings California Codes). At that time the state introduced a breeders' (mare owners) incentive program for winning horses at what is now a relatively low rate, equivalent to 3% of the winner's purse. State statutes say that was implemented to fulfill the encouragement part of the law (California Statutes, 1933). Today the level of California breeder incentive awards is much higher and more varied providing incentives for several horsemen's subgroups.

A concern is whether the program is stimulating, something that incites a tendency to action to improve racing stock (an incentive), or is it assistance with a grant or gift of money to an enterprise deemed publicly advantageous resulting in increased supply (a subsidy) (Smith). This paper argues that the concern should be viewed from the output side as a racing product, that is, is the program providing a higher quality racing product? Whether they are considered a subsidy or an incentive is immaterial.

The question arises as to the optimal policy format of these awards for performance, or what subgroups should be rewarded. Indeed a national public industry exchange has ensued (see Paulick). This discussion and some analysis has questioned the ability of a breeder incentive programs to enhance horse "quality," but generally agrees that it increases the supply of horses. Thus, the horse quality variable only appears to be related to the quantity or proportion of superior horses. One dilemma of that approach is that racing is a matter of competing against the existing competition or population, but if the entire breed is improving, as reflected by faster times, many horses even though faster than their predecessors, would be labeled lower quality.

The California program uses a small percentage (0.54% of handle) from pari-mutuel revenue to encourage production with incentives for mare owners or breeders-55% of the fund, stallion owners-20% of the fund, open race owners' (of CBs) bonuses-20%, and restricted CB race fund purse supplements -5%. All of these program aspects are performance based requiring animals to qualify in some way equal to or above some specified minimum, by performance achievements that is by winning or placing, finishing first through third place in a race (California Statutes, 1998).

The Problem: Is racing quality achieved by having more superior horses on a race card? Has California's breeder incentive programs increased racing quality?

Previous Work:

One of the first economic evaluations of horse quality in California racing was the 1965 Stanford Research Institute (SRI) study, which found California breds were disproportionately represented in lower quality races. The percentage of out-of-state bred horses in higher quality races exceeded that of CB's, 25% and 15% respectively. Further, SRI investigators found two-thirds of starters in California in 1964 were California breds.

So, CBs were an important source of equine input to racing; however, this treatment seems to ignore California's position as a/the leading racing state in the 60's, still true today.

California remains a net importer of racehorses from other states and countries. California offers substantial opportunity for better horses to compete in one of the top racing circuits in the world, rivaled in the 60's only by New York tracks. California remains isolated physically by great distance from other major racing and breeding states, the cost of shipping horses great distances would seem to reduce the likelihood of transporting cheaper horses and favor shipping of better horses capable of running in higher quality races (SRI, pp.18-19)**.**

Lawrence was one of the first to model horse values with annual auction sales prices of horses as a function of industry and macroeconomic variables, others have followed (Church and Bohara; Buzby and Jessup). Lansford, *et al.* modeled racing Quarter Horse prices from a single major auction sale at an Oklahoma City venue with an hedonic price function of pedigree or genetic variables. They used a pooled cross-section time series from eleven years of consignment sales. Their variables included stakes race "engagements" – paid nominations, which enhanced the market for horses eligible for specific stakes races.

Chezum and Wimmer developed an adverse selection model of racing two-year-old maidens. The adverse selection was by breeders to sell lesser quality horses, while retaining potentially better runners for their own racing stables. They used as a test the post time wagering odds on the breeder-owned versus horses sold, finding significantly lower odds on the retained home breds from a sample of 39 races.

Suits was among the first of many subsequent efforts to model the demand for horse racing, in a pooled cross-section time series (1949-1971 in 24 states) he found an elastic response to price or takeout (the dollars removed from the wager pool before winners are paid).

Neibergs and Thalheimer evaluated breeders' awards programs in three second-tier racing states utilizing a structural model of yearling supply-demand. Their data was a pooled cross-section with a thirty-year time series. Their results suggested that breeders and stallion awards were the best places to spend additional funds to enhance breeder revenue. Open race purse bonuses were not damaging on the demand side, but were relatively costly to improve at then current levels.

Smith modeled the California breeders' awards programs using two years, 1997-8, of race by race data for major tracks (*i.e.* excluding fairs) from a state database. His model utilized a double hurdle (Cragg) or tiered approach to a hedonic decision process for the consumption of horse racing. The first tier or hurdle was the decision to attend and the second decision or hurdle was whether to and how much to wager on an individual race or set of betting pools. His theoretical approach included normal demand specification, his empirical estimation assumed away many of the core demand variables, which are recognized as difficult, but not impossible to obtain (such as proprietary advertising data). His basic model appears to be 1998 weekday, distance (one mile or more), turf races of allowance horses at Santa Anita Park in the Los Angeles area.

Smith's model application suggests that CBs do not enhance the racing product; however, his estimated relationships show a negative impact of a dummy variable for CB races and a higher average number of runners in the CB races. The net effect is an increase in handle in CB races (see Appendix).

The Model: Hanemann and others (*e.g*. Deaton and Muellbauer) describe the situation of utility maximization that includes the "general" corner solutions, where some but not all the potential goods or services in a set may be consumed. Utility here is a function of vector of prices and income, but also a vector of quality characteristics associated with any good in the set. The consumer utility of several goods' quality characteristics enter the utility function through an aggregation function, whose arguments are the quality characteristics. The randomness of varying preference represents the specific individual's tastes, while recognizing that tastes vary across consumers. An aggregation of characteristics preferred and the assigned weights, which vary across individuals, is the resultant utility measure. This is presented as $\theta_i = \Sigma \gamma_k b_{ik}$, where γ_k individual weights placed on the kth characteristics by consumers and b_{ik} are the characteristics assessed by consumers (Hanemann).

Henderson and Quandt developed the logic of the transition from utility functions to linear expenditure functions whose left-hand side is expenditure, the product of price and quantity. Deaton and Muellbauer (p.28) make the transition from axioms of preference of utility to quantities as a bridge from the theoretical to the pragmatic. In this paper we use that bridge as the basis for setting expenditures on wagering, the mutual pools, as a proxy measure for racing community utility. The community utility proxy created through aggregation is probably not similar to individual utility functions. However, one might segregate the punters' philosophies into the following groupings.

- Breeder or horse enthusiasts, who see racing as revealing the best horse in a foal crop or of crops currently racing, these might signal likely prospects for the breeding shed and the improvement of the breed. Probably more interested in stakes races, allowance races, and non-claiming maiden contests.

- The "Betting Race" person, who wants a narrow range of nearly equal quality of horses with differing styles, races for aged allowance horses or claiming horses with extensive records in big fields.

- Professional gamblers who seek the highest payoff, the divergence between public odds and her own independent private valuation model of the winning probability similar to the theory of auction bidding behavior (see Milgrom and Webber).

- The occasional outing race goers, where racing is one of the set many alternative sporting action venues or recreational activities. This could include group/club outings.

- The Big Race enthusiasm – a broad appeal even to the aforementioned groups and those attracted by pre-race publicity for the major event.

- Championship or Big Event day persons with very infrequent attendance and usually on days of national sport interest, e.g. Triple Crown races or the Breeders' Cup.

The first group's perceptions dominate most of the previous work and discussion on racing quality, but one can see that each group would have varied concepts of what quality racing was depending on their own perspectives. Racing associations, the tracks, recognize and attempt to appeal to all groups that might have an interest in racing. They realize that the breeder/horse enthusiasts group helps give credibility to their racing, they none the less must package the "race cards" to fully utilize the racing inputs available, while appealing to all groups within their market. The breeder/horse enthusiasts group may not be the largest target market group. Thus the aggregation or cumulative impacts of all submarket groups will have an impact on defining what quality is in racing.

Quality Characteristics of a Race:

This unequal quality of characteristics endowment is a good representation of the situation confronting the consumer of horse racing entertainment services. Each day's racing card or cards with modern satellite wagering on intrastate and interstate racing provides a broad range of probability weighted characteristics that are subject to substantial variation between horse players. Thus players consume races on a card at varying rates of expenditure. Chezum and Wimmer suggest that horse players maximizing returns should approximate the objectives of utility maximization if bettor utility is a function of tickets cashed and payoffs.

Smith's treatment presented a linear estimation of what are really joint decision aspects of interactive product quality (see Rosen, pp.37-8). Further, his estimate of CB's value emanated solely from their impacts on handle in CB only races. Usually only one race per day is CB, yet quick calculations will show that in San Francisco over one-half the horses starting in open races are CB and in Los Angeles over one-third are CB. Their contribution as performers is ignored. Lastly, the use of quality as defined by horsemen is probably too limiting for a consumer viewpoint. This appears to be a case where "expert opinion" (Smith, see p.1-footnote) is used to define the wants or demand of the average consumer.[2] Our contention is that the typical player is more appropriately represented by a distribution of player preferences. Thus, the list of quality aspects would include caliber of horses, the absence of cheating, and field size as quality measures. A "bettable" competitive race is one that is fairly run, has a reasonable spread of odds, and enough entrants to make it an interesting betting proposition. The absence of cheating or the fair gamble is assumed, but is not unrealistic given Ali's (1977) work and the active supervision of racing in California by the regulator the California Horse Racing Board.

Our model is semi-logarithmic as suggested by Gujarati (pp.460-1), interactive in real numbered variables of the right hand side and linear in the hedonic dummy variables.

$$Y_t = \alpha X_{1t}^{\beta 1} X_{2t}^{\beta 2} X_{3t}^{\beta 3} + \Sigma_i^n(\gamma_I D_{it}) + \mu_t$$

and in log-log form the general model linear in its parameters is

$$\text{Ln } Y_t = \alpha + \beta_1 \text{ Ln } X_{1t} + \beta_2 \text{ Ln } X_{2t} + \beta_3 \text{ Ln } X_{3t} + \Sigma_i^n (\gamma_i D_{it}) + \mu_t$$

Where the X_i are interactive measurable quality factors and D_i are qualitative variables.

Specifically those right hand side variables are defined here as follows:

ThanM - total handle by region by race in $1000's
PurseM - purse or prize amounts by race in $1000's
NumCB - number of California bred runners by race
OSB - number of non-California bred runners by race
Dist - a race distance dummy variable, 1 if 1 mile or over, 0 - otherwise
RaceCB - a California bred restricted race dummy variable, 1 if CB, 0 – otherwise
dBM - a dummy variable for Bay Meadows racetrack, 1 if BM, 0 – otherwise

[2] In horse racing, "the sport of kings," industry leaders seldom "hang out" with the typical player.

dDMR - a dummy variable for Del Mar racetrack, 1 if DMR, 0 – otherwise
dGGF - a dummy variable for Golden Gate Fields racetrack, 1 if GGF, 0 – otherwise
dHP - a dummy variable for Hollywood Park racetrack, 1 if HP, 0 – otherwise
dOT - a dummy variable for Oak Tree at Santa Anita racetrack, 1 if OT, 0 – otherwise
dFairs - a dummy variable for county and state fair racetracks, 1 if Fairs, 0 – otherwise
dWknd - a dummy variable for a weekend day, 1 if a weekend day, 0 – otherwise
Claiming - a dummy variable for a claiming races, 1 if a Claiming race, 0 – otherwise
LastRace - a dummy variable for the last race of the day, 1 if Last race, 0 – otherwise
Mdn - a dummy variable for a maiden, *ie*. non-winners, 1 if a Maiden race, 0 – otherwise
Stakes - a dummy variable for a stakes race, 1 if a Stakes race, 0 – otherwise

Stakes, maiden (*Mdn*), and *Claiming*[3] races are qualitative variable race classifications that denote quality of performers, which leaves the default "allowance" races. Purses values in thousands (*PurseM*) are the prizes that horses compete to earn and are funded through the "purse" share of the handle based on the same race meeting the previous year. *NumCB* and *OSB* are the number of instate and out of state bred horses respectively by race, previous analysis did not use this very labor intensive (high cost) variable. *RaceCB* is a dummy variable for exclusively California bred horses, one but not the only aspect of state bred contribution to racing. *Dist* is a dummy variable for races greater or equal to than one mile in length, as opposed to sprint races, which are less than one mile. *Last Race* of the day and *Wkend* are dummy variables for the culminating race which is where the state regulatory agency terminates several gambling pools and thus the appearance of a large handle on this not necessarily high quality race. Wkend is a dummy variable for a weekend racing day, which like other spectator sports the attendance is higher as patrons are released from the grind of the workaday world. Lastly, *dBM*, *dGGF*, *dHP*, *dDMR*, *dOT*, and *dFairs* are dummy variables for racing at specific San Francisco, Los Angeles, and San Diego area tracks or county and state fairs.

The race by race data are taken from the California Horse Racing Information Management System (CHRIMS) for the calendar years of 1997 and 1998. Additionally, the number of CB was defined from the racing program collection of the California Thoroughbred Breeders' Association, the administrative agency for the programs.

<u>The Results</u>: Table 1 below presents the linear estimation of the relationships for comparison to our preferred interactive model. These linear estimations show similar response to Smith's outcome except attendance is not considered a characteristic explanatory variable and in key variables on CBs added to this model. Where Smith assumed that allotting all value of CBs be limited to those racing in state bred only races, the model below shows that tracking CBs in open races with the variable *NumCB* has important policy implications.

---

[3] Stakes are early closing races that require a nomination fee (the only Thoroughbred races that do require an entry fee) and are the top 4 or 5% of all races offered. Only 3% of horses win stakes races. Claiming races offer a great range of quality classifications. The term claiming refers to selling races, where horses entered can be purchased by a licensed owner. The potential owner puts in a "claim" for an upset price before the start of the race, the price is defined in the conditions of the race. Lastly, maiden races are for horses yet to win a first victory. Generally, the range of quality runs down from stakes races to allowance races to maiden to claiming.

Table 1. Regression Results of California Total Handle by Race Day on Race Characteristics, 1997-1998.

Predictor	Coef	t-value	Coef	t-value
Constant	204.76	14.76*	193.51	14.02*
LastRace	242.33	34.55*	241.63	34.34*
PurseM	3.0768	77.90*	2.9617	80.50*
Dist	-16.742	-3.19*	-17.318	3.29*
Mdn	-26.933	-4.73*	-22.865	4.02*
RaceCB	36.86	3.36*	32.810	2.99*
NumCB	68.959	43.21*	69.436	43.40*
Stake	-98.02	-7.86*		
OoSbRnr	81.178	52.45*	81.520	52.53*
Claiming	-35.160	-5.82*	-24.659	4.18*
d BM	-445.71	-50.60*	-446.19	50.50*
d DMR	56.56	5.26*	55.87	5.18*
d GGF	-438.353	-49.85*	-437.73	49.63*
d HP	-90.109	-10.57*	-89.334	10.45*
d Wkend	131.35	26.08*	127.47	25.35*
d OT	-146.60	-11.85*	-145.73	11.74*
d Fairs	-614.03	-68.19*	-617.98	68.52*

SSE = 239.6 R^2 = 75.6% R^2(adj)= 75.6% SSE=240.3 R^2 = 75.5% R^2(adj)= 75.4%
n= 9896 k = 16 $F_{(16,9879)}$ = 1913.65* n= 9896 k = 15 $F_{(15,9880)}$ = 2024.66*
dw = 1.59 dw = 1.59

Note: * - significant at 0.05 level.

This specification shows (see Table 1) substantial positive benefit and occurs as an important sign change occurs (from Smith's negative to positive) for the *RaceCB*, which is also significant in this model. This model combined many of the track-by-year dummy variables with the recognition of their very similar coefficients and standard errors apparent in Smith's results. While degrees of freedom are not a constraint here, we invoke Occam's razor to simplify the model without loss of substantial explanatory power.

Table 1 also presents a second model specification, which excludes the *Stakes* dummy variable. That variable came into this and Smith's models with an unacceptable negative sign, but it is redundant as designed. Smith's handle model attendance variables (his ON ATT and OFF ATT), weekend day (Weekend-Smith and dWknd-this paper), and the purse variable (PurseM) may be picking up the same handle phenomena as intended for Smith's dummy variable for Stake(s) races. It can be seen that explanatory power loses little from dropping what should be the most demonstrable quality variable in horse racing in the presence of what are its collinear counterparts.

Table 2. Log-Log Regression Results of California Total Handle by Race Day on Race Characteristics, 1997-1998.

Predictor	Coef	t
Constant	5.78124	159.94*
lgPurseM	0.24380	29.36*
lgNumCB	0.019078	11.30*
lgOSB	0.035811	14.78*
Dist	-0.051343	-5.77*
RaceCB	0.40633	14.57*
dBM	-0.71550	-44.78*
dDMR	0.07820	4.28*
dGGF	-0.65192	-41.26*
dHP	-0.12834	-8.87*
dWknd	0.22209	26.23*
dOT	-0.12714	-6.05*
dFairs	-1.2179	-73.31*
Claiming	0.06671	5.36*
LastRace	0.38856	32.84*
Mdn	0.044014	4.59*

SEE = 0.4067 R^2 = 68.3% R^2(adj) = 68.2%
df = 9894-15-1 F = 1463.11* dw = 1.17

Note: *-significant at 0.05 level.

A Chow test for variable inclusion supports the dropping of this variable. A better specification would be to back "stakes purses" out of general purses as done with CB runners in open races as in Table 1. Low end "stakes" purses in California typically are 40% higher or more than the highest overnight-non-stakes purses, and as most stakes races are run on weekends in the US the implication would be obvious to most horse players.

Table 2 presents our log-log specification that should capture the interactive components of race quality real valued components and the additive dummy variable influences (see Gujarati). This model maintains the same relationships established in the linear model for number of CB runners, other runners, and restricted CB races as in Table 2; however, two variables unexpectedly change sign from the linear to log transformation model. All other signs remained the same. Dummy variables for maiden (Mdn) races and "Claiming" races enter with positive coefficients. These two variables generally represent the lowest level of quality in races and should have a negative impact on the model. An explanation is that the same purse effect that rendered the dummy variable for stakes invalid may be having this effect or that their field sizes a large enough to offset the lower quality.

Summary and Conclusions: This analysis found that CB incentive programs did augment the quality of racing as reflected in the racing handle, a quantity equivalent of utility of the racing entertainment product. The double log model utilized two years of California race by race data from all Thoroughbred races run in the state. The interactive model of racing characteristics fits the data well with robust results. Both measures of races restricted to CB horses and the incremental effects of CBs in open races were positive and significant. While the majority of the data used was from the state managed racing data base, CHRIMS, a key element was the identification of numbers of CBs in all races and the separation into a separate variable. The most important explanatory variables from the set available were purse size, and dummy variables for fair racing, last race of the day, Hollywood Park race day, CB restricted races, and Del Mar racing. The dummy variables for race class, CHRIMS designation, appear to present the only aberrant behavior in the model with unanticipated positive signs on the coefficients. This is usually the lowest class or quality of horse racing.

With over half the runners in California open races being CB the contribution of CBs to viable racing programs with fuller fields appears important to the consumers' response in wagering level. This is likely owing to the larger fields of runners in the claiming and maiden race ranks. The bulk of any horse population fits into these categories larger categories and any race written for these groups would likely receive more entries.

References Cited:

Ali, Mukhtar. "Probability and Utility Estimates of Racetrack Bettors," *Journal of Political Economy* (85:4). Aug.1977. pp.803-15.

Buzby, J. and E. Jessup. "The Relative Impact of Macroeconomic and Yearling-Specific Variables in Determining Thoroughbred Yearling Price." *Applied Economics* (26), 1994, pp. 1-8.

California Statutes, selected volumes 1933-1998. Sacramento: California State Printing Office.

Chezum, Brian and B. Wimmer. "Evidence of Adverse Selection from Thoroughbred Wagering." *Southern Economic Journal* (66:3). 2000, pp. 700-714.

Cragg, John C. "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods." *Econometrica* (39:5). September 1971, pp. 829-44.

Deaton, A. and J. Muellbauer. Economics and Consumer Behavior. Cambridge : Cambridge University Press.1980, p.28-36.

Deerings California Codes. "Business and Profession Code" Volume, §19401 (b). San Francisco: Bancroft & Whitney, 1992.

Gujarati, D.N. Basic Econometrics, 2e. New York City: McGraw-Hill Book Company. 1988, pp. 460-1.

Hanemann, W. M. “Quality and Demand Analysis,” in New Directions in Econometric Modeling and Forecasting in US Agriculture, editor G. Rausser. New York City : Elsevier Science Publishing Co., Inc. 1982, pp. 81-84.

Henderson, J. and R. Quandt. Microeconomics, 3rd ed., New York City: McGraw-Hill Book Co., Inc. 19

Lansford, N.H., Jr., *et al.* “Hedonic Prices of Race-Bred Yearling Quarter Horses Produced by Quarter Horse Sires and Dams.” *Journal of Agribusiness* (16:2) Fall 1998, pp.169-86.

Lawrence, Robert G. “An Econometric Analysis of the Quarter Horse Industry.” unpublished Doctoral Dissertation, Texas A&M University, 1969.

Milgrom, P. and R. Webber. “A Theory of Auctions and Competitive Bidding.” Stanford University, Technical Paper No. 358, December 1981.

Neibergs, J.S. and R. Thalheimer. “Thoroughbred Breeder/Owner Incentive Policies – An Economic Analysis Using Selected States.” Equine Industry Program, CBPA, University of Louisville, Dec. 1998.

Paulick, Ray. “State Incentive Programs: Support or Subsidy?” *The Blood-Horse*, December 5, 1998, pp.7136-7141.

Rosen, Sherwin. “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition.” *Journal of Political Economy* (82). 1974, pp.33-55.

Smith, Martin D. “Breeding Incentive Programs and Demand for California Thoroughbred Racing: the Tradeoff Between Quantity and Quality.” Presented at the 1999 AAEA Annual Meeting, Nashville, August 1999.

Stanford Research Institute (SRI). California’s Horse Racing Industries—An Economic Analysis. SRI Project I-5329, December 1965, pp. 259.

Suits, Daniel. “The Elasticity of Demand for Gambling.” *Quarterly Journal of Economics* Feb.1969, pp.155-62.

Appendix - The Net Effect of a Cal Bred Race.

Smith's (p.9) calculation of the net effect of CB races and thus the CB program in general is curious. Instead of using the difference in "average runners in a CB race" less "the average runners in an open race (OR)," he uses the 1997-1998 difference in CB runners of 0.5. Then does the same for OR runners, a 1997-1998 difference of 0.3 and uses those time related differences as the increase in runners in a CB race, rather then a possible trend difference they may represent. The difference in CB versus OR runners should be 1.3 in 1997 and 1.1 in 1998, substantially higher. The net effect is now positive for the CB race changing his net effect from a negative to a positive (from – 4,655 to +15,659), which reverses the policy conclusion.

CB Race Effect = [(0.5 Runners * 25,392) + (1 dCB * -17351)] = -4,655 *original(Martin)*
CB Race Effect = [(1.3 Runners * 25,392) + (1 dCB * -17351)] = 15,659 *revised*