**Some Evidence of Information Aggregation in Auction Prices**

Brian Chezum
Department of Economics
St. Lawrence University
Canton, NY 13617
chezum@stlawu.edu

C. Jill Stowe
Department of Agricultural Economics
University of Kentucky
Lexington, KY 40506-0276
jill.stowe@uky.edu

***Selected Paper prepared for presentation at the Southern Agricultural Economics Association Annual Meeting, Corpus Christi, TX, February 5-8, 2011***

**Abstract**

When purchasing goods of uncertain future value, market prices should reflect the market's best estimate of the future value of the good. Further, market price should, on average, perform well in predicting the realization of future value. In this paper, we test the market's ability to aggregate information and thereby predict future value in the Thoroughbred racing industry. Using sales data from weanling, yearling, and two-year-old sales, we hypothesize that as more and/or better information becomes available, price should become a better predictor of outcomes (earnings). Results indicate that any information that becomes available between weanling and yearling sales does not improve prediction of earnings, but information on the potential quality of two-year-olds reduces the variation in predicting future racetrack earnings.

***Key Words:*** *Efficient Markets Hypothesis, Information aggregation, Thoroughbred industry*

## 1. Introduction

When purchasing goods of uncertain future value (for example, an offshore oil lease, a stock, or a Thoroughbred racehorse), market prices should reflect the market's best estimate of the future value of the good. Further, market price should, on average, perform well in predicting the realization of future value.[1] In this paper, we test, indirectly, the market's ability to aggregate information and thereby predict future value in the Thoroughbred racing industry.

In Thoroughbred racing, racehorse prospects are sold as weanlings, yearlings and two-year olds. As a horse ages, more information about the potential of that horse becomes available. For example, the conformation, temperament, and trainability of the horse become more certain in the 6 to 9 months between the weanling and yearling sales. When the horse turns two, they begin to enter training, revealing yet further information about the horse's ability. Especially relevant in this analysis is that before the two-year-old sales (which are called "two-year-old in training" sales), prospective buyers can actually watch the horse perform over a designated distance (usually 1 to 2 furlongs, or 1/8 to 1/4 mile). Consequently, prospective buyers can observe a horse's running style and speed over a short distance. In this sense, buyers have significantly more information on which to determine prices than at weanling or yearling sales.[2] We hypothesize that as more and/or better information becomes available, price should become a better predictor of outcomes.

Recently, the efficient markets hypothesis has received support from the popularity and usefulness of prediction markets, which are financial markets also referred to as "information

[1] Under the efficient markets hypothesis (Fama (1970)), if prices fully aggregate information (both private and public), then the inclusion of additional information will not improve price's ability to forecast the future value.
[2] In addition, as a horse ages, new information regarding the horse's dam and sire becomes available. This information includes quality of dam and sire progeny (i.e. new race winners, stakes race winners, etc) as well as breeding quality of siblings. This information should also lead to a more accurate assessment of a horse's ability and hence and hence future value

markets" or "event futures." In such markets, market prices are used to predict future events; one of the most well-known examples of prediction markets is the Iowa Electronic Market. Wolfers and Zitzweitz (2004) provide a useful overview of the many types and applications of prediction markets in use today. Among other things, they illustrate the power of prediction markets by using data generated in the Iowa market for the previous four presidential elections. They show that the average absolute forecast error generated by market trading as a function of days to the election is diminishing; as more information becomes available leading into an election, market forecasts, dictated by share prices in the market, improve.

In this paper, we explore fundamentally the same issue. We use auction data for prices and career racing earnings for outcomes on all horses from the 1993 thoroughbred foal crop that sold as weanlings, yearlings, and/or two-year olds. Examining the ability of price to predict earnings, we ask if the prices from two-year old sales outperform prices from weanling and yearling sales as a predictor of earnings.

Section 2 of this paper outlines the data, and Section 3 discusses the empirical methodology. We employ a permutation test to examine if the predictive power measured by (adjusted) $R^2$ for a regression of earnings on price within age group supports the hypothesis that two-year old sales prices perform better as predictors. Section 4 discusses the results and Section 5 concludes.

## 2. Data

Our data are comprised of 15,124 North American born Thoroughbred race horses from the 1993 foal crop that sold at least once as a weanling, yearling, or two-year old. In 1993, 36,455 Thoroughbreds were born in North America and registered with the Jockey Club as racing prospects.[3] From this foal crop, roughly 41% were sold as either a weanling, yearling, or two-year old.[4] Weanling sales are predominantly held in November and December of each year prior to a horse officially turning 1 year old. The majority of yearling sales are between April and September of each year, while the majority of two-year old in training sales are held between February and April.

The sample is drawn from auction summaries published by *The Blood-Horse*, a leading industry publication. In those summaries, sale prices for individual horses are available, along with information regarding each horse's pedigree, consignor, and buyer. Career racetrack earnings are recovered from the 2009 edition of the *American Produce Records*; this publication consists of the produce record (name, racing performance, and sales performance for each foal) of every mare that produced a foal in North America between 1960 and 2008. We use the 1993 foal crop to insure that the horses in our sample have completed their racing careers. Descriptions of the variables used in the study are presented in Table 1, and summary statistics for the full sample are presented in Table 2.

We observe that mean earnings rise with age at sale, including an approximately 19% increase from weanling to yearling sales horses and a roughly 26% increase from yearling to two-year old sales horses. Examining price, we observe a 22% increase in prices from weanling to yearling sales, which is near the increase in mean earnings, while mean price increases by only 6% between yearling two-year old sales, a substantially lower increase as compared to mean

[3] For a Thoroughbred to be sold as a racing prospect or appear in a race, it must be registered with the Jockey Club, which verifies that the horse is from recognized thoroughbred bloodlines.
[4] For official purposes, all Thoroughbreds are considered to have a birth date of January 1.

earnings. The change in mean prices across sales does not match the change in mean realized earnings.[5]

The final column of Table 2 presents the coefficient of variation (CV). We note that the CV for price is close 2 for each age group, with the two-year-old CV being slightly lower. The CV for earnings is consistently higher for earnings across age groups. This is unsurprising as we might expect more noise in earnings. It is difficult, for example, to forecast a career-ending injury that occurs before a horse is able to race. If anything, the summary statistics indicate considerably more noise in two-year old earnings bur less variability in prices. This would appear to indicate weaker predictive ability of price in the two-year-old sample.

At this point, we must address the issue of residual breeding value. The purchase price of a Thoroughbred reflects the net present expected value of its racetrack earnings and its residual breeding value. For horses that go on to have successful careers in the breeding shed, breeding value can be quite high. However, less than 1% of colts from any foal crop will even stand at stud for a few years, and the majority of them will be unsuccessful; for the average colt, expected breeding value is zero. Alternatively, mares are much more likely to enter the breeding shed. It is estimated that about 50% of all females will produce at least one offspring. Thus most fillies will have some positive expected value in breeding. To accommodate this potential noise in our data, we perform the test detailed in the next section on different sub-samples of the data. An exhaustive list of these subsamples is as follows: colts, fillies, non-select horses, non-select colts,

---

[5] Among the horses sold as two-year olds, were three horses that earned in excess of 3 million dollars, with these horses selling for $30,000, $90,000 and $200,000. If we drop these horses from the two year old sample, the mean total earnings for two-year olds drops 48,367 with a standard deviation of 116,545.60. We observe that the standard deviation in both prices and earnings are lower in weanling and yearling sales but, dropping the extreme earners from the two-year old sample, the standard deviation in price and earnings is quite similar between yearling and two-year old sales.

non-select fillies, select horses, select colts, and select fillies.[6] In addition, as we will discuss in Section 4, the results seem to be robust to the potential presence of residual breeding value.

## 3. Empirical Methodology

We hypothesize that because prices aggregate information, the price of a Thoroughbred at auction should predict racetrack quality as measured by career earnings; moreover, sales prices of older horses should outperform sales prices of younger horses as a predictor of earnings. To test this hypothesis, we regress the natural log of earnings (*LN_EARN*) as a function of the natural log of prices (*LN_PRICE*) for weanling, yearling and two-year old sales. A natural measure of the predictive power of a regression model is the $R^2$, which in this context is interpreted as the percent of the variation in earnings explained by the variation in sales price.[7] Defining $R^2_t$ ($t$ = 0,1,2) as the adjusted $R^2$ recovered from a regression of auction price on career earnings for weanlings, yearlings, and two-year-olds, respectively, we test the following hypotheses:

(1) $H_0$: $R^2_2 - R^2_1 = 0$
$H_1$: $R^2_2 - R^2_1 > 0$

(2) $H_0$: $R^2_2 - R^2_0 = 0$
$H_1$: $R^2_2 - R^2_0 > 0$

(3) $H_0$: $R^2_1 - R^2_0 = 0$

[6] Select sales are ones in which individuals with only the highest quality pedigrees and conformation are presented for sale; these are the individuals, particularly among the colts, most likely to have a large positive expected residual breeding value.
[7] We will use the adjusted $R^2$ in our tests as the degrees of freedom differ across the subsamples.

$$H_1: R^2{}_1 - R^2{}_0 > 0$$

There is no standard statistical method available to test these hypotheses. Fisher (1935), however, proposes an alternative testing procedure that is commonly referred to as a "randomization" or "permutation" test. A randomization test involves calculating some test statistic, *m*, based on the observed data. The distribution of the statistic is found by randomly reordering the data a sufficiently large number of times and recalculating the test statistic for each reordering. Consider, for example, asking whether Thoroughbred buyers pay higher average prices for male (colts) or female (fillies) racing prospects.[8] Using a sample of auction prices, we could calculate the mean for colts and fillies, respectively. The permutation test then involves randomly relabeling an observation as either a colt or filly (irrespective of the observations actual gender), preserving the number of horses in each group.[9] The difference in means is then calculated for each relabeling; this forms the distribution of the test statistic, which in this case is the difference in means. The significance level is then recovered as the percentage of observations that are at least as extreme as the observed value. If this probability value is small, then the observed pattern would seem unlikely if the null hypothesis were true.

The permutation test then involves three steps: (1) calculate the observed statistic, (2) randomly reorder the data many times and calculate the test statistic for each reordering, (3) find the proportion of values (the *p*-value) that are at least as extreme as the observed statistic, and (4) conclude against the null hypothesis if the *p*-value is small.

---

[8] We might expect colts to command higher average prices because the purses for races run predominantly by males are on average larger.

[9] Under Fisher's (1935) original development, we would consider every permutation but for large samples this become impractical and is unnecessary for reliable results.

Consider the first of our three hypotheses given in (1). The alternative suggests a particular pattern to the data; more specifically, it indicates that the adjusted $R^2$ in a regression of auction price on earnings for two-year old sales is greater than for yearling sales, while the null suggests that if this pattern is observed in the data, it is little more than a chance event. To test the hypothesis, we first recover the difference in adjusted $R^2$ (i.e. $R^2_2$ - $R^2_1$) from the observed data. Next, we randomly re-label the observations across the three sales, preserving the proportion of observations in each sale, and again recover the difference in adjusted $R^2$. Repeating this process 10,000 times determines the distribution of the test-statistic. From this distribution, we calculate the *p*-value for our hypothesis as the percentage of values in the constructed distribution that lie above our observed value.

## 4. Results

Regression results of *LN_EARN* on *LN_PRICE*, by age and within different subsamples, are presented in Table 3. The first set of regression results utilizes the full sample; the remaining sets or results utilize different subsets of the data.

In all but one model, *LN_PRICE* is a significant predictor of *LN_EARN* at the 5% level or better (the lone exception occurs in the model restricted to select two-year-old colts; *LN_PRICE* is nowhere near conventional levels of significance). However, one immediately notices that adjusted $R^2$ values are quite low. The highest adjusted $R^2$ occurs in the sample of two-year-old colts in which roughly 10% of the variation in earnings is explained by the variation in prices. One possible reason for this inefficiency is that this market, like prediction markets, may exhibit some of the same behavioral biases as other financial markets, such as overweighting probabilities of unlikely extreme outcome events (finding the “home run” horse) and underweighting likely extreme outcome events (finding a zero-earnings runner). Since finding a racehorse that makes substantial earnings on the track (or in the breeding shed) are small probability events, this auction market may not work as well as in other, more moderate-probability environments. Another behavioral bias relevant to the Thoroughbred racing industry is the tendency of buyers to bid and purchase according to personal preferences rather than objective probability assessments of a horse's future performance (see Rhode and Strumpf (2008) and Forsythe, Reitz, and Ross (1999)). Finally, horse racing is wrought with uncertainty even before a horse runs its first race; it is impossible to forecast a career-ending injury that may occur while a young horse romps in the paddock before training has even commenced.

Table 4 reports the test statistic and *p*-value for the permutation test in each subsample. Excluding the "select" sub-samples, we find that in the entire sample and the fillies subsample, the difference in adjusted $R^2$ between yearlings and weanlings is significant at the 10% level. Otherwise, this difference is not statistically significant, which suggests that information revealed between the weanling and yearling sales is either not fully incorporated into prices or is of little use in predicting future racing quality.

Again excluding the "select" sub-samples, all differences in adjusted $R^2$ between two-year-olds and either yearlings or weanlings are significant at the 1% level. The robustness of this result is not entirely surprising since, as mentioned earlier, prospective buyers can actually watch two-year-olds perform over a designated distance a few days before the sale. In weanling and yearling sales, prospective buyers merely analyze the horse's conformation and observe them walking up and down the shed row; their pricing decisions are based on these observations (along with pedigree). From our results, we can infer that watching two-year-olds perform at their intended task is informative, and hence prices more accurately predict the horse's future racetrack earnings.

The results are less clear in the sub-samples of select sale horses. Since there are no select sales for weanlings, we can only compare models for yearlings and two-year-olds. In the sample of select horses, the difference in adjusted $R^2$ is significant at the 5% level, while the differences are significant only at the 10% level when select colts and select fillies are analyzed separately. One possible explanation for the diminished significance of results is that horses sold at select sales have high-quality pedigrees, and hence these individuals have significant residual breeding value. This residual breeding value factors into their sale price but may not be a strong indicator of racetrack earnings, especially since horses with high residual breeding value may end their racing careers early to take advantage of their potential profitability in the breeding shed.

Taken together, we conclude that information which becomes available between the weanling and yearling sales is of little value in predicting future racetrack performance, since weanling and yearling sales prices predict earnings with a similar degree of accuracy. However, information accumulated before the two-year-old sales is valuable in determining career earnings, and moreover, prices aggregate this information and become better predictors of racetrack quality. Thus, at least to some extent, prices serve to help predict future outcomes in the Thoroughbred racing industry, but the degree of error is still quite large as evidenced by the low adjusted $R^2$ values.

## 5. Summary and Conclusion

According to the efficient markets hypothesis, market prices should reflect the market's best estimate of the future value of the good, and market price should perform well in predicting the realization of future value. In this paper, we test the market's ability to aggregate information and thereby predict future racing value in the Thoroughbred racing industry.

We find that information which becomes available between the yearling and two-year-old sales is most useful in terms of predicting future racing value; two-year-old sale prices are better predictors of earnings than either weanling or yearling sale prices. Moreover, there is little difference in the ability of weanling and yearling sales prices to predict earnings. Further, in all three age groups, price is a highly significant variable in predicting earnings, yet very little variation in earnings is explained. So, these markets are efficient in the sense that prices aggregate available information and predict earnings better, but not in the sense that prices do a satisfactory job of explaining earnings, as the highest adjusted $R^2$ was less than 0.10.

If a market is truly efficient, the market price will be the best predictor of the future event, and no other available information will aide in improving the market's predictions. While there are many potential reasons regarding why Thoroughbred auction markets are not terribly efficient, future research will include other variables to the regression models to determine what information, if any, is over- or under-weighted by the market.

## References

Fama, Eugene. 1970. “Efficient Capital Markets: A Review of Theory and Empirical Work.” *Journal of Finance* 25(2), 383-417.

Fisher, R.A. (1935) *The Design of Experiments*. Oliver and Boyd, Edinburgh.

Forsythe, Robert, Thomas Rietz, and Thomas Ross. 1999. “Wishes, Expectations and Actions: Price Formation in Election Stock Markets.” *Journal of Economic Behavior and Organization* 39(1), 83-110.

Grossman, Sanford J. and Joseph E. Stiglitz. 1976. “Information and Competitive Price Systems.” *The American Economic Review* 66(2), 246 – 53.

Grossman, Sanford J. and Joseph E. Stiglitz. 1980. “On the Impossibility of Informationally Efficient Markets.” *The American Economic Review* 70(3), 393-408.

Rhode, Paul W. and Koleman Strumpf. 2008. “Manipulating Political Stock Markets: A Field Experiment and a Century of Observational Data.” Mimeo, University of Arizona and University of Kansas.

Wolfers, Justin and Eric Zitzewitz. 2004. “Prediction Markets.” *Journal of Economic Perspectives* 18(2), 107-126.

TABLE 1 – Names, expected signs, and definitions of variables

| Variable | Expected Sign | Description |
|---|---|---|
| **Dependent Variable** | | |
| *EARN* | n/a | Career earnings |
| **Independent Variable** | | |
| *PRICE* | + | Auction price at weanling, yearling, or two-year-old sale |

TABLE 2 – Summary statistics for sales prices and earnings by age

| | Sample Size | Mean | Standard Deviation | Coefficient of Variation |
|---|---|---|---|---|
| Weanling Price<br>Weanling Earnings | 1791 | 23,121.13<br>35,014.40 | 46,765.88<br>93,456.45 | 2.023<br>2.669 |
| Yearling Price<br>Yearling Earnings | 9311 | 28,102.72<br>41,503.64 | 57,017.33<br>120,099.40 | 2.029<br>2.894 |
| Two-Year Old Price<br>Two-Year-Old Earnings | 4022 | 29,906.72<br>52,380.10 | 56823.14<br>204,215.90 | 1.900<br>3.899 |

TABLE 3 – Regression results for all subsamples of the data

| **Dependent Variable: *LN_EARN*** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **All** | | | **Colts** | | | **Fillies** | | |
| **Age** (***n***) | W (1791) | Y (9311) | 2 (4022) | W (922) | Y (4833) | 2 (2019) | W (869) | Y (4478) | 2 (2003) |
| ***LN_PRICE*** **(std error)** | 0.362*** (0.069) | 0.433*** (0.024) | 0.747*** (0.040) | 0.496*** (0.100) | 0.465*** (0.031) | 0.765*** (0.053) | 0.229** (0.095) | 0.371*** (0.037) | 0.707*** (0.060) |
| **Adj.** $R^2$ | 0.0145 | 0.0340 | 0.0815 | 0.0247 | 0.0432 | 0.0939 | 0.0055 | 0.0223 | 0.0649 |

| **Dependent Variable: *LN_EARN*** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Non-Select All** | | | **Non-Select Colts** | | | **Non-Select Fillies** | | |
| **Age** (***n***) | W (1791) | Y (7894) | 2 (3442) | W (922) | Y (4107) | 2 (1667) | W (869) | Y (3787) | 2 (1775) |
| ***LN_PRICE*** **(std error)** | 0.362*** (0.069) | 0.412*** (0.026) | 0.726*** (0.045) | 0.362*** (0.069) | 0.412*** (0.026) | 0.726*** (0.045) | 0.362*** (0.069) | 0.412*** (0.026) | 0.726*** (0.045) |
| **Adj.** $R^2$ | 0.0145 | 0.0314 | 0.0711 | 0.0145 | 0.0314 | 0.0711 | 0.0145 | 0.0314 | 0.0711 |

| **Dependent Variable: *LN_EARN*** | | | | | | |
|---|---|---|---|---|---|---|
| | **Select All** | | **Select Colts** | | **Select Fillies** | |
| **Age** (***n***) | Y (1417) | 2 (580) | Y (726) | 2 (352) | Y (691) | 2 (228) |
| ***LN_PRICE*** **(std error)** | 0.535*** (0.071) | 0.821*** (0.191) | 0.776*** (0.097) | 0.294 (0.245) | 0.226** (0.105) | 1.435*** (0.310) |
| **Adj.** $R^2$ | 0.0378 | 0.0293 | 0.0797 | 0.0013 | 0.0053 | 0.0828 |

***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively

TABLE 4 – Permutation test results on all subsamples of the data

| | Comparison | Difference | p-value | n |
|---|---|---|---|---|
| **All** | Year-Wean | 0.0194** | 0.0250 | 15124 |
| | Two-Wean | 0.0670*** | 0.0000 | |
| | Two-Year | 0.0475*** | 0.0000 | |
| **Colts** | Year-Wean | 0.0184 | 0.1214 | 7774 |
| | Two-Wean | 0.0691*** | 0.0000 | |
| | Two-Year | 0.0507*** | 0.0001 | |
| **Fillies** | Year-Wean | 0.0168* | 0.0783 | 7350 |
| | Two-Wean | 0.0594*** | 0.0000 | |
| | Two-Year | 0.0426*** | 0.0000 | |
| **Non-Select** | Year-Wean | 0.0149* | 0.0651 | 12146 |
| | Two-Wean | 0.0589*** | 0.0000 | |
| | Two-Year | 0.0440*** | 0.0000 | |
| **Non-Select Colts** | Year-Wean | 0.0143 | 0.1763 | 6172 |
| | Two-Wean | 0.0686*** | 0.0000 | |
| | Two-Year | 0.0542*** | 0.0000 | |
| **Non-Select Fillies** | Year-Wean | 0.0139 | 0.1332 | 5974 |
| | Two-Wean | 0.0495*** | 0.0000 | |
| | Two-Year | 0.0356*** | 0.0003 | |
| **Select** | Two-Year | 0.0764** | 0.0287 | 269 |
| **Select Colts** | Two-Year | 0.0203* | 0.0576 | 162 |
| **Select Fillies** | Two-Year | 0.1322* | 0.0975 | 107 |

***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively