Forecasting Yearling Prices: A Comparison of Alternative Time Series Models

Dr. James E. Payne, Economics, Eastern Kentucky University
Dr. Robert R. Sharp, Economics, Eastern Kentucky University
Dr. Susan A. Simmons, Economics, The Citadel

Abstract

The thoroughbred breeding industry in North America has fallen on hard times. The health of this industry is often gauged by prices obtained for yearlings at North American auctions, particularly the average prices of summer sales at Keeneland and Saratoga. We examine various exponential smoothing algorithms along with a market-based structural model, as well as an ARIMA model in generating one-step ahead forecasts. The market-based structural model outperforms the other approaches with respect to both in- and out-of-sample forecasting accuracy.

Introduction

Hard times have fallen on the thoroughbred breeding industry in North America. Declining attendance at meets, a drop in yearling prices, tracks folding, horse farms failing, competition from lotteries and simulcasting, and unfavorable tax laws are obstacles facing the thoroughbred horse racing industry. Going from a popular form of gambling up until the 1970s, competition and unfavorable economic conditions have lead to a decline in wagering and in the profitability of raising and racing thoroughbreds.

The health of the industry is often gaged by prices obtained for yearlings at North American auctions, particularly the average prices of summer sales at Keeneland and Saratoga. Reaching an industry high (in 1970 dollars) of $170,677 in 1984, the summer sales average fell 12 percent in 1985 (its first decline since 1975), then nosedived by 27 percent in 1986. Following declines the next two years, the average staged a 9 percent rally in 1989 only to fall in each of the next three years. Although summer yearling sales account for a declining proportion of total yearling sales (5.4 percent in 1980, 3.9 percent in 1992) the average price of all North American auction yearlings appears to mirror swings in the summer sales prices.

The objective of this paper is to examine several time series approaches to forecasting the percentage change in yearling prices. We will entertain several smoothing algorithms, ARIMA model, as well as a structural market based reduced-form model, in analyzing both in-sample and out-of-sample forecasting accuracy. The following section will present the data and the alternative time series models. The last section summarizes both the in-sample and out-of-sample forecasting performance of each model as well as provide concluding remarks.

Data and Methodology

We will entertain two smoothing algorithms, an ARIMA model, and a market-based reduced-form model. Let us begin by outlining the market-based reduced-form model. The reduced-form model of yearling prices is derived from interaction between the demand and supply of yearlings. The following specifications for the demand and supply of yearlings parallels the work by Simmons and Sharp (1992).

The demand function is given by the following:

$$(1)\quad P_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 TAX_t + \alpha_3 RTFN_t + \alpha_4 Q_t + u_t$$

where α_1, α_2, $\alpha_3 > 0$ and $\alpha_4 < 0$. As Simmons and Sharp (1992) have demonstrated there is a positive relationship between the percentage change in average yearling price, P_t, and the percentage change in real per capita, Y_t, income. An increase in the highest marginal tax rate, TAX_t, leads to greater tax write offs for owners (Hereth and Talbot, 1988). For instance, with a marginal tax rate of 70 percent owners could write off 70 percent of expenses. Thus, marginal rates are hypothesized to have a positive impact upon demand. The percentage change in the ratio of real assets to financial assets, $RTFN_t$, serves as a proxy for the availability of related investment assets. As the value of financial assets increase relative to real assets (horse ownership) individuals substitute towards the higher valued financial assets. Finally, there is an inverse relationship between the percentage change in average yearling prices and quantity demanded for yearlings.

The supply is defined as the quantity sold as follows:

(2) $Q_t = \beta_0 + \beta_1 P_t + \beta_3 LAGCRP_t + \beta_3 RCOPT_t + \beta_4 PPRR_t + v_t$

where β_1, β_2, >0 and β_3, β_4 <0. The quantity supplied, Q_t, is positively related to the percentage change in yearlings price, P_t. Annual number of breeding age horses (foals), $LAGCRP_t$, serves as a proxy for the production of horses and is hypothesized as having a positive relationship with quantity supplied. As race opportunities, $RCOPT_t$, increase the supply of horses should decrease. However, as the percentage change in the dollar amount of purses per race, $PPRR_t$, increase the supply of horses brought to auction declines as owners are encouraged to race not sell their stock. Table 1 provides a detailed description of the data.

Table 1
Variable Descriptions

Variable	Description
P_t	Percentage change in average yearling prices from North American auctions (excluding million dollar yearlings).
Q_t	Percentage change in the number of yearlings sold in North American auctions (excluding million dollar yearlings).
Y_t	Percentage change in real per capita income in the U.S. in 1970 dollars.
TAX_t	Highest federal marginal tax rate in percentage terms.
$RTFN_t$	Percentage change in the ratio of real assets to financial assets where farm value per acre index in 1970 dollars denotes real assets and the Dow Jones industrial index serves as proxy for financial assets.
$LAGCRP_t$	Percentage change in the annual number of "breeding age" horses (foals).
$RCOPT_t$	Percentage change in the number of racing opportunities defined as the number of races relative to the number of starters.
$PPRR_t$	Percentage change in the dollar amount of purses per race in 1970 dollars.

Given the demand and supply equations above we proceed to estimate the following reduced-form equation for the percentage change in average yearling prices. Equation (3) is derived by substituting equation (2) into equation (1) and re-arranging terms.

(3) $P_t = \Pi_0 + \Pi_1 Y_t + \Pi_2 TAX_t + \Pi_3 RTFN_t + \Pi_4 LAGCRP_t + \Pi_5 RCOPT_t + \Pi_6 PPRR_t + \epsilon_t$

where:

$\Pi_0 = (\alpha_0 + \alpha_4\beta_1) >< 0$
$\Pi_1 = \alpha_1 / (1 - \alpha_4\beta_1) > 0$
$\Pi_2 = \alpha_2 / (1 - \alpha_4\beta_1) > 0$
$\Pi_3 = \alpha_3 / (1 - \alpha_4\beta_1) > 0$
$\Pi_4 = \alpha_4\beta_2 / (1 - \alpha_4\beta_1) < 0$
$\Pi_5 = \alpha_4\beta_3 / (1 - \alpha_4\beta_1) > 0$
$\Pi_6 = \alpha_4\beta_4 / (1 - \alpha_4\beta_1) > 0$
$\epsilon_t = (\alpha_4 v_t + u_t)/(1 - \alpha_4\beta_1)$

Using annual data over the time frame 1940 to 1989 we estimated equation (3) by ordinary least squares corrected for first-order autocorrelation with t-statistics in parentheses.

(4) $P_t = -36.830 + 3.814Y_t + .2634TAX_t$
(-2.405) (5.706) (2.037)

$+ 9.616\ RTFN_t + .6847LAGCRP_t$
(3.048) (.4786)

$+ 1.942RCOPT_t + 2.814PPRR_t$
(2.101) (6.580)

Adj.R^2=.654 AR(1)=.4488 F=13.43
(3.753)

The coefficients are statistically significant at the 5 percent level except for $LAGCRP_t$. A 1 percentage increase in real per capita income leads to a 3.814 percentage increase in price. A 1 unit increase in the marginal tax rate results in a .2634 percentage increase in yearling prices. Both purses per race and the ratio of real assets to financial assets have a positive impact upon price. The F-statistic, 13.43, is statistically significant at the 1 percent level, suggesting the model has overall predictive power. Equation (4) will later be used in generating in-sample and out-of-sample forecasts. Let us proceed to examine three other univariate approaches to forecasting the percentage change in average yearling prices.

We entertain an ARIMA model for yearling prices. Initial inspection of the autocorrelation function (ACF) for yearling prices suggest that the time series is stationary. Examination of both the autocorrelation and partial autocorrelation (PACF) functions reveals spikes at the first and third lag in the PACF with the ACF displaying an exponential decay. The second lag in the PACF was insignificant. This suggests an AR(3) model, ARIMA(3,0,0). We estimated the following moving average model with t-statistics in parentheses.

$$(5)\quad P_t = \underset{(7.152)}{.4439P_{t-1}} - \underset{(-1.936)}{.1171P_{t-3}} + \epsilon_t$$

Adj. R^2 = .534 DW = 2.86 F= 52.644 Q(16)=18.909

The autoregressive coefficients fulfilled the stationarity conditions. The Box-Pierce Q-statistic of 18.909 does not exceed the critical value of 23.70 with 14 degrees of freedom based on the chi-square distribution at the 5 percent significance level. This suggests that the residuals behave as "white noise".

The next two approaches represent popular smoothing algorithms: single exponential (SES) and Holt-Winters linear trend (HW). The SES algorithm (Muth, 1960) is given by the following:

$$(6)\quad L_t = \alpha X_t + (1 - \alpha)L_{t-1} \qquad 0 < \alpha < 1$$

where L_t is the new level estimate of the time series, yearling price, which is simply a weighted average of the previous level estimate, L_{t-1} and the new observation, X_t. The smoothing parameter, α is chosen to minimize the sum of squared errors. There are several caveats associated with the SES algorithm. First, SES is appropriate for nonseasonal time series which exhibit no discernable upward or downward trend. Second, the forecasts from SES will yield constant forecasts for all future values of a time series. As seen later the HW algorithm attempts to correct for these shortfalls in SES by incorporating a localized linear trend.

The HW algorithm (Holt, 1957 and Winters, 1960) is given by the following:

$$(7a)\quad L_t = \alpha X_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \qquad 0 < \alpha < 1$$
$$(7b)\quad T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \qquad 0 < \beta < 1$$

where L_t and T_t are estimates of the level and slope at time period t along with α and β as smoothing parameters. The HW algorithm allows the level and slope to change over time as new observations become available. In equation (7a) the level, L_t, is estimated as a weighted average of X_t and $(L_{t-1} + T_{t-1})$. In equation (7b) the slope at time t is estimated as a weighted average of the most recent change in estimated levels, $(L_t - L_{t-1})$ and the previous slope estimate.

Table 2A
In-Sample Forecasting Performance
1940 to 1989
(*Lowest RMSE)

Method	Smoothing Parameters	RMSE
SES	α = .001	32.67014
HW	α = .000, β = .220	33.24063
ARIMA (3,0,0)		32.91054
Regression		17.29926*

Forecast Results and Concluding Remarks

Based upon the models presented in the previous section, we examine the accuracy of each model in terms of one-step ahead (static) forecasts for the period: 1940 to 1989 (in-sample) and 1990 to 1991 (out-of-sample). Table 2A presents the in-sample forecast results.

The reduced-form regression model yields the lowest RMSE. Given the models estimated over the period 1940 to 1989, we then undertake out-of-sample forecasts. Table 2B presents these forecasts for the years 1990 and 1991. The SES, HW, and ARIMA (3,0,0) models generate forecasts for positive percentage change

Table 2B
Out-of-Sample Forecasting Performance
1990 to 1991
(*Lowest RMSE)

Method	1990	1991	RMSE
Actual	-15.970856	-10.417804	
SES	9.441638	9.441638	22.805616
HW	5.769611	5.935224	19.236286
ARIMA (3,0,0)	1.4151608	.6482179	12.395828
Regression	-23.586214	-13.384559	5.7790705*

increases in average yearling prices. The reduced-form regression model generates negative percentage changes in average yearling prices corresponding to the actual decline in prices.

Suggestions for Future Research

The forecasting experiments presented are only preliminary. Of the four approaches only the reduced-form model incorporates some of the economic determinants in understanding yearling prices. It is hoped this investigation can generate further discussion in the modelling and forecasting of key variables associated with the thoroughbred horse industry by both the private and public sector. Such modelling efforts would help in analyzing the impact of regulatory and tax policy upon this industry.

*****References*****

1. "Auctions of 1992", *The Blood Horse*, Lexington, Kentucky, January 16, 1993.
2. *Economic Report of the President*, U.S. Government Printing Office, Washington, DC, 1990.
3. Goode, Richard, "Lessons from Seven Decades of Income Taxation", *Options for Tax Reform*, edited by Joseph A. Pechman, 1984.
4. Hereth, R. and J. Talbott, "TRA 86 - Its Impact on Thoroughbred Investments", *CPA Journal*, April, pp. 30-37, 1988.
5. Holt, C.C., "Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages", Office of Naval Research (Memorandum 52), 1957.
6. Muth, J.F., "Optimal Properties of Exponentially Weighted Forecasts", *Journal of the American Statistical Association*, 55. pp. 299-306, 1960.
7. Simmons, S.A. and R.R. Sharp, "An Econometric Model to Predict the Supply and Demand of Thoroughbred Yearlings Brought to Auction" presented at the annual meeting of National Decision Science Institute, November 1992, San Francisco, CA.
8. *Statistical Abstract of the United States*, U.S. Department of Commerce, Washington, DC, 1992.
9. *The American Racing Manual*, Daily Racing Form, Heightstown, NJ, 1992.
10. Winters, P.R., "Forecasting Sales by Exponentially Weighted Moving Averages". *Management Science*, 6, pp. 324-342, 1960.