Jonathan Murello
October 23, 2010

Time Series- Summer 2010
Student Project

Introduction

Fantasy Football has occupied the fall and winter Sundays of a growing number of teenage and middle aged men for the last 20 years. The basic premise is to form a league, get together before the season starts for a draft of the top offensive and defensive players in the National Football League, set lineups each week, and see who comes out on top at the end of the season. Typical drafts utilize a snake-draft style of drafting: A team is assigned a draft position, their turn comes up, they draft the player with the most potential, and then wait until their turn arrives again.

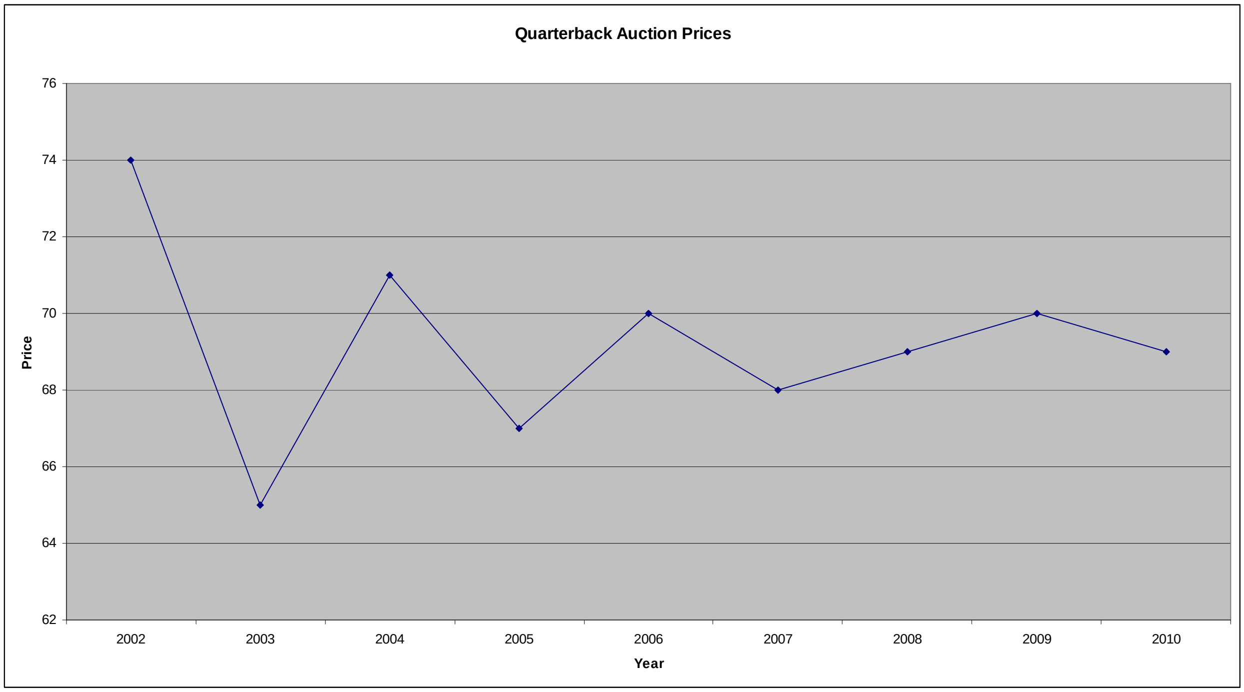
However, an increasing number of leagues have drifted away from the snake-draft style of drafting in favor of an auction draft. In an auction draft, each manager is assigned a budget for the draft. An auctioneer elects a player to be bid on, all managers take part in an open-bidding war, and the process is repeated until all rosters are filled. East Coast Fantasy League is one such league that utilizes the auction draft style with a $200 budget. Year in an year out, the top quarterbacks and running backs are bought for a large portion of each manager's budget. Never in league history has a manager ended the draft with the highest priced quarterback and running back. In the following analysis, we attempt to prove that it is possible to buy the highest priced quarterback and running back in the 2011 draft next season by developing models to predict the price of the highest valued quarterback and running back.

Data

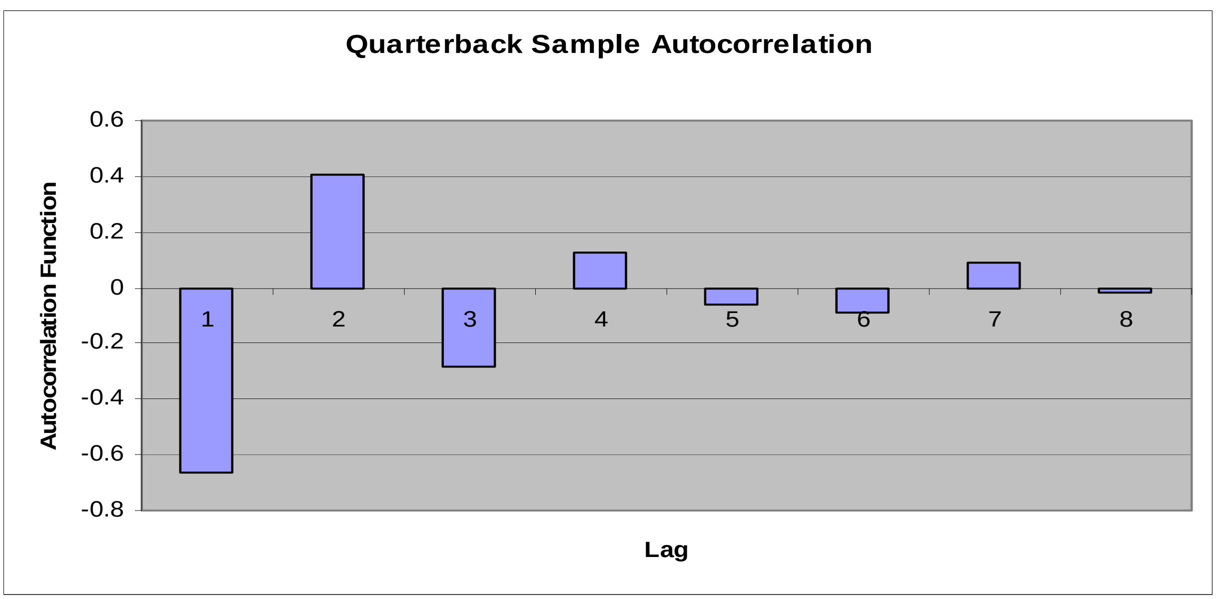
East Coast Fantasies has been in existence since 2002, and this year the 9th auction draft was completed. Every year the data is recorded in an effort to provide each manager with a tool to analyze what went well and what went poorly during any particular season's draft. Auction value data is also vital when a new manager becomes introduced to the league. Below are auction values representing the quarterback and running back that sold for the highest amount in each year from 2002-2010.

Highest Priced Quarterbacks and Running Backs from 2002 - 2010				
Year	**Quarterback**	**Price**	**Running Back**	**Price**
2002	Kurt Warner	$74	Marshall Faulk	$91
2003	Rich Gannon	$65	Ricky Williams	$78
2004	Peyton Manning	$71	Jamal Lewis	$64
2005	Peyton Manning	$67	LaDanian Tomlinson	$72
2006	Peyton Manning	$70	Shaun Alexander	$85
2007	Peyton Manning	$68	LaDanian Tomlinson	$102
2008	Tom Brady	$69	LaDanian Tomlinson	$94
2009	Drew Breese	$70	Adrian Peterson	$92
2010	Aaron Rodgers	$69	Chris Johnson	$85

Model Specification- Quarterbacks



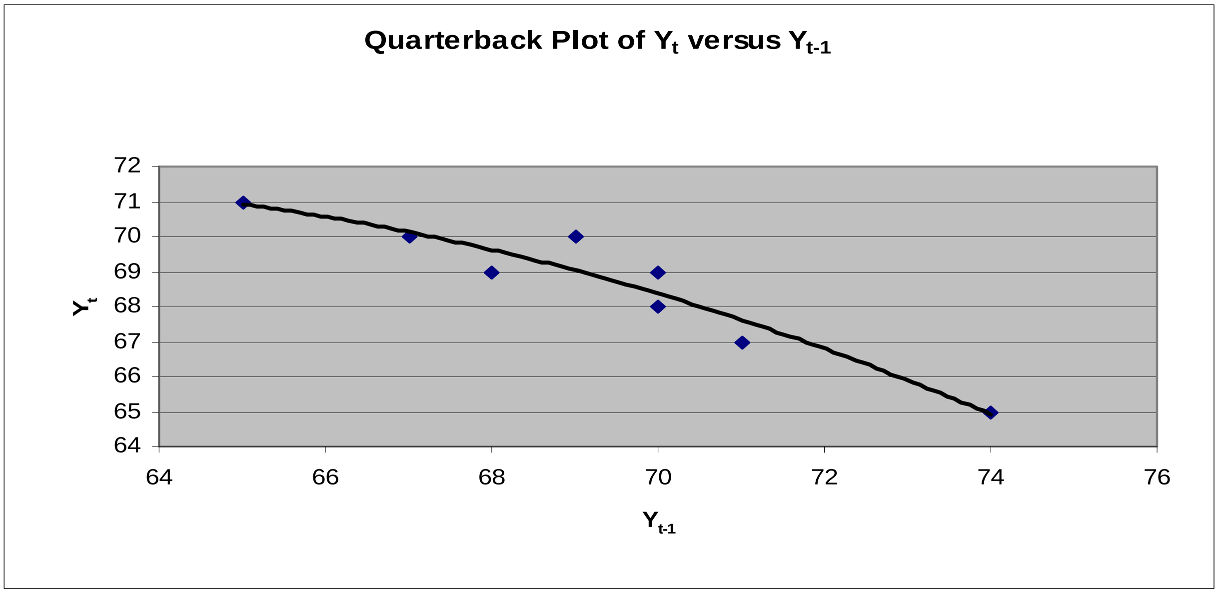
The auction prices for the highest valued quarterback exhibit negative correlation; that is, the data shows a period of increase in quarterback auction price followed by a period of decrease in quarterback auction price the next season. Also, a period of decrease is followed by a period of increase the next season.



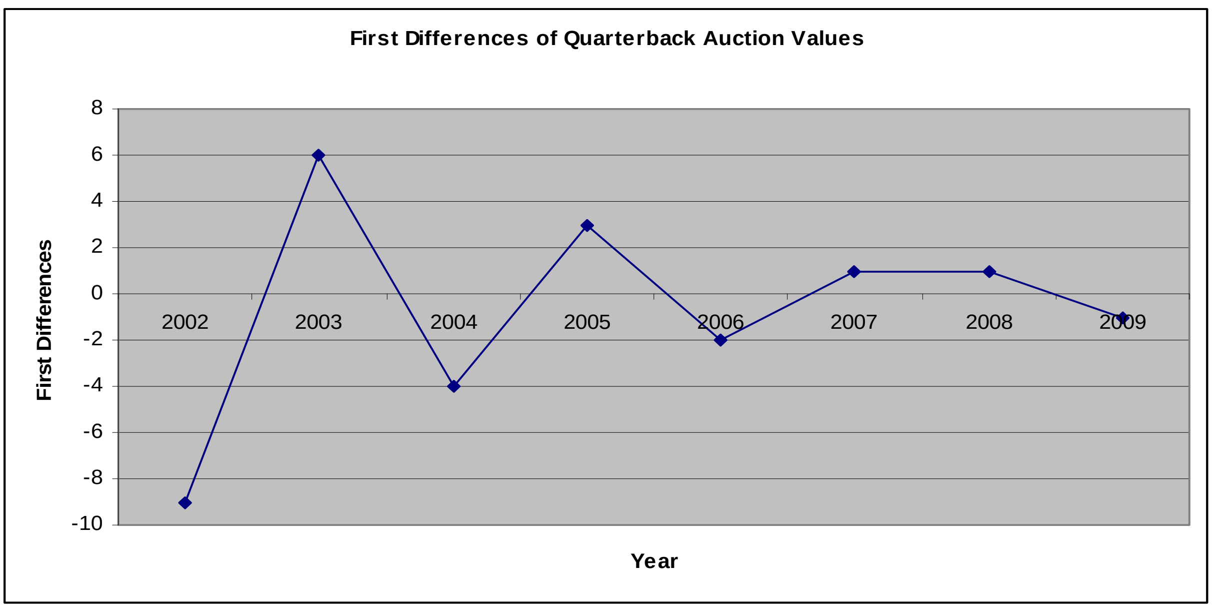
The Sample Autocorrelation values were calculated using the sample autocorrelation function,

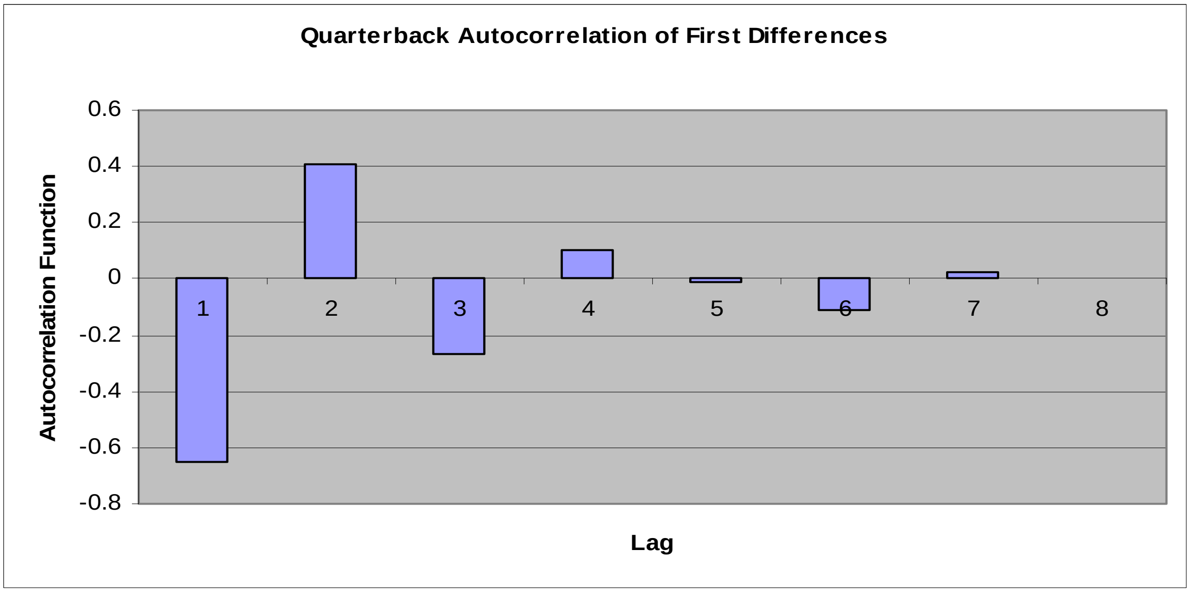
$$r_k = \frac{\sum_{t=k=1}^{n} (Y_t - \overline{Y}_t)(Y_{t-k} - \overline{Y})}{\sum_{t=1}^{n} (Y_t - \overline{Y})^2},$$

where Y_t represents the auction value in a given year, Y_{t-k} represents the auction in the t-k year, and $\overline{Y}$ represents the average of the auction prices from 2002 to 2010. The sample autocorrelations in the above chart exhibit negative correlation and move rapidly to zero as lag increases. We are inclined to state that the process is stationary, but first differences must be tested before we settle on this conclusion.



The negative correlation characteristic of quarterback auction price is further shown in the above plot of Y_t versus Y_{t-1}. The negative sloped trend line displays negative correlation at work with the quarterback auction prices. Next, we take first differences of the quarterback auction values to test stationarity.





Taking first differences, we again see negative correlation displayed by the quarterback auction values in the plot of first differences on the previous page and the plot of sample autocorrelation first difference values above. Most importantly, the sample autocorrelation of first differences show the sample autocorrelations oscillating about zero with decreasing amplitudes, indicating that the first differences are stationary.

Since the sample autocorrelations and the sample autocorrelations of first differences appear to oscillate about zero with decreasing amplitudes, we conclude that quarterback auction values are a stationary process. Also, since the autocorrelations begin at non-zero values and decline to zero over time, we conclude that quarterback auction values follow an autoregressive process. This was to be expected based on the perception that has evolved from the managers of the East Coast Fantasy League. At the inception of the league, it was hard to value quarterbacks, as shown by the disparity in the auction values during the first 4 years. But over time, the league managers have discovered that quarterback statistics are fairly consistent season to season (as evidenced by Peyton Manning being the highest valued quarterback 4 years in a row),they rarely get injured, and that there is no reason to overpay or underpay for the top quarterback. The auction price of the highest valued quarterback is heavily dependant on the highest valued quarterback in previous years, and this is a strong indication that the process is autoregressive in nature. In fact, the Sample Autocorrelation correlogram strongly supports an AR(1) process with a negative φ parameter.

Model Specification- Running Backs

Contrary to the top quarterback value, the top running back value varies greatly from year to year. Many changes that have a temporary effect on running backs affect their value in each specific season. For example, LaDanian Tomlinson set the single season running back touchdown record in 2005, resulting in him being sold for the astronomical price of $102 in 2006. This feat has never been duplicated. Furthermore, if one particular season saw a decent amount of running backs performing well, then the abundance of talent drives the price of the top running back down, as shown in the 2004 season when Jamal Lewis went for $64. Nevertheless, overwhelming league perception is that running back value is based on last season's statistical performances and last season's statistical performances only. This temporary effect is indicative of a moving average process. For simplicity, we will assume the running back value is moving average in nature.

Model Parameterization-Quarterbacks

Using the quarterback auction values that we have shown to be stationary, we fit the highest priced quarterback to the following AR(1) equation using Excel's regression tool:

$$Y_t = -.665Y_{t-1} + 114.680$$

Modeling quarterback auction pricing as an AR(1) process produces an R^2 value of 86%, suggesting that the model is able to explain 86% of the observed auction prices. Our p-value of .06% is extremely low, indicating that there is a strong probability that the above equation is representative of the quarterback auction data. Finally, $|\varphi|=.665<1$, giving further evidence that the autoregressive process is stationary.

Model Parameterization-Running Backs

Being that we are assuming the top running back auction value data follow a moving average process, in order to develop a model which we can use to forecast future auction values we will assume the that the running back auction values follow an ARIMA(0,1,1) process. Based on this assumption, the following table was developed:

Running Back Auction Values Model						
	ARIMA(0,1,1)				ARMA(0,1)	
	Year	Forecast	Actual	Residual	Forecast	Actual
	2008	93*	94	1	-	-
$\hat{Y}_{2008}(1)$	2009	98*	92	-6	4	-2
$\hat{Y}_{2009}(1)$	2010	93*	80	-13	1	-12

* Forecast values were calculated as the average of the two preceding year's actual values.

To calculate the forecast and actual values of the ARMA(0,1) process, we used the following equations:

$$\hat{ARIMA}_t(1) = ARIMA_t + \hat{ARMA}_t(1)$$

$$ARIMA_{t+1} = ARIMA_t + ARMA_{t+1}$$

Now, we set up a pair of linear equations of the form:

$$\hat{Y}_i(1) = \mu - \theta\, \chi_i$$

$$(1):\ 4 = \mu - (1)\theta$$

$$(2):1 = \mu - (-6)\theta$$

Where $\hat{Y}(1)$ represents each ARMA(0,1) forecast, μ is the mean of the ARMA(0,1) process, θ is the Moving Average parameter, and each χ_i represents the residual of the prior period. Solving for μ and θ, we arrive at μ=3.57142 and θ=-.42857, giving an ARMA(0,1) process of the form: $\hat{Y}_i(1) = 3.57142 - (-.42857)x_i$.

Model Diagnostics

We will examine the diagnostics of the quarterback data only, since regression analysis was not used for the running back data. In order to determine the goodness of fit of the quarterback auction values to the AR(1) model we developed, the Box-Pierce Q statistic was calculated for the 8 lags available using the following equation, $Q = n(\hat{r}_1^2 + \hat{r}_2^2 + \cdots + \hat{r}_K^2)$, where n represents the number of lags, and each $\hat{r}_i^2$ represents the autocorrelation of lag_i. Our calculation resulted in a Box-Pierce statistic of Q=5.8177 with seven degrees of freedom. The Box-Pierce Q statistic was then used to test the null hypothesis that the quarterback auction data followed an AR(1) process. Assuming Q is an approximate chi-square distribution with 7 degrees of freedom and interpolating a chi-square distribution table, we arrive at the approximate P-value of .36, indicating that our α (significance level) is approximately .64. This indicates that we would reject the null hypothesis at the .64 significance level and higher. Generally, we would desire our α to be less than .05.

Since our significance level is much higher than .05, we are inclined to state that the quarterback auction data does not follow an AR(1) process. However, an important requirement of the Box-Pierce statistic is that n be 'large'. The Box-Pierce statistic is intended for sets of data with many observations, and our data simply does not fit that criteria. It is difficult to apply a diagnostic test to our data since our sample size is so small. As a result, we will not reject that our quarterback data follows an AR(1) process. It would be interesting to test the diagnostics of the quarterback auction data 20 years from now when the sample size is greater.

Projections

Using our AR(1) model for the quarterback data and the 2010 auction value of $69 for Aaron Rodgers, we project that the top quarterback in 2011 will also be sold for $69. To project the 2011 auction price for the top running back, we enter in the 2010 residual of -13 and calculate

$\hat{Y}_{2010}(1) = -2$. Adding this to Chris Johnson's 2010 auction value of $80 for the ARIMA(0,1,1) process, we project the 2011 top running back to be bought for $78.

Conclusion

In the preceding analysis we obtained auction value data for the top quarterbacks and running backs sold in the East Coast Fantasy League over the last 9 years. We specified the quarterback auction data to be stationary, autoregressive in nature, and exhibit strong negative correlation. Running back data was assumed to follow a moving average process. The quarterback data was then fit to an AR(1) process while the running back data was modeled after an ARIMA(0,1,1) process. After testing the model diagnostics of the quarterback model it was determined that, although the significance level as derived using the Box-Pierce statistic was notably high, the AR(1) model representing the quarterback data would not be rejected due to small sample size. It should also be mentioned in the quarterback model parameterization section strong evidence was given to signify that the quarterback data followed an AR(1) process.

As a result of our models, we project that purchasing the top quarterback and top running back in next years auction draft would cost a manager $147 of their $200 budget. However, each team's roster must include 17 players, and the league requires teams to pay at least $1 for each of their players. Taking this into account, a manager who wishes to go after the top running back and quarterback in the 2011 draft will have a remaining budget of only $38 left to spend on their remaining team. In conclusion, it is possible for a manger to successfully bid on the top running back and top quarterback in the 2011 draft, but the price of doing so may not be worth the risk due to the possibility of injury to either of the team's star players.