THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF FINANCE

TESTING THE EFFICIENT MARKET HYPOTHESIS IN FANTASY FOOTBALL AUCTION LEAGUES

KEITH ARMINGTON
FALL 2011

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Finance
with honors in Finance

Reviewed and approved* by the following:

Dr. James Miles
Professor of Finance
Thesis Supervisor/Honors Advisor

Dr. Chris Muscarella
Professor of Finance
Thesis Reader

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

When people think of the Efficient Market Hypothesis, they're likely to think of capital markets and the pricing of securities. What they probably won't think of, however, is the market for National Football League players within fantasy football. Piggybacking on the success of the National Football League and the proliferation of the internet, fantasy football – in which participants attempt to predict which football players will perform best on Sundays – has created a whole new market. Such markets for football players are referred to as "auction-style fantasy football" and assigns cash values to players. Better performing players demand higher prices, and participants usually engage in bidding wars to buy the best players. The market for fantasy football is huge, and now with Las Vegas getting involved, payoffs for successful participants can be enormous. My thesis will focus on these auction-style markets and determine if they are "efficient." If not, obviously, the inefficiencies could be exploited for personal gain.

To determine if fantasy football auction-style markets are efficient, I focused on top quarterbacks from the past three years and attempted to produce a predicative model based largely on historical production and prices. Then, I took the best models and applied them to the 2010 football season to determine their accuracy in predicting the price of players. Finally, I compared the statistical differences in production of the quarterbacks with the difference between my model and the actual prices to see if any quarterbacks were incorrectly valued relative to their production. The result would show me if the fantasy football markets were indeed efficient or not, and if not, which players would be incorrectly valued. Hopefully, someone could use my models to determine a more accurate price for players and play the market accordingly.

TABLE OF CONTENTS

LIST OF TABLES iii

ACKNOWLEDGEMENTS iv

Chapter 1 Introduction 1

Chapter 2 Related Papers 5

Chapter 3 Methodology 12

Chapter 4 Research and Analysis 15

Chapter 5 Conclusion 17

Chapter 6 Bibliography 18

Appendix A Two Variable Regression with Prior Year Price and Prior Year Points: Top 20 QBs 19

Appendix B One Variable Regression with Prior Year Points: Top 20 QBs 20

Appendix C One Variable Regression with Prior Year Price: Top 20 QBs 21

Appendix D 2010 Price Prediction Model 22

Appendix E 2010 Price Prediction Model with 2010 Prices and Variances 23

Appendix F One-Variable Regression with 2010 Actual Price and 2010 Actual Points: Top 20 QBs 24

Appendix G One Variable Regressions with Residuals and Actual-Predicted Prices: Prior Year Points Model: Top 20 QBs 25

Appendix H One Variable Regressions with Residuals and Actual-Predicted Prices: Prior Year Price Model: Top 20 QBs 26

LIST OF TABLES

Table **3-1**: Sample One-Variable Regression Using Prior Year Points. 4

Table **3-2**: Sample Two-Variable Regression Using Prior Year Points and Prior Year Price. 6

ACKNOWLEDGEMENTS

Thank you to Dr. James Miles, who gave me great insight and guidance in my research efforts and helped me through the thesis process. Thank you also to Dr. Chris Muscarella, who graciously agreed to be my Thesis Reader and who supported me throughout my Penn State career.

Chapter 1

Introduction

There are many events people look forward to every year. Birthdays, holidays, vacations, award ceremonies, and parties are just a few of these occasions. Sports seasons on both a professional and collegiate level also stand as highly anticipated happenings. Arguably one of the most popular sports in America is football, operated exclusively by the privately held National Football League (NFL). In 2009, the league surpassed $8 billion in revenue (Bramhall). In early 2010, a record 106.5 million viewers tuned in to watch the culmination of the season: the Super Bowl (Bramhall).

Riding the NFL's wave of popularity is a different sport limited not by brute size or athletic ability, but by connection to the Internet. Fantasy football, as it's known, is a sport in which actual statistics (touchdowns, passing yards, turnovers) of NFL players – who have been selected by the fantasy team's owner – combine to create a total score. This score is then pitted against an opponents score to determine weekly winners and losers. Invented in the late 1960s as a hobby among a few sports writers, the game gradually gained popularity and saw exponential growth with the advent of the Internet (St. Amant). Now, anyone can play fantasy football online at sites like ESPN.com, Yahoo.com, and CBSSportsline.com. In fact, fantasy football is now a $1 billion million industry with over 30 million players in America and Canada (Gregory).

The average fantasy football league consists of ten teams, each owned by a separate individual (Gregory). In 2009, participants paid an average of $73 to join a league (Gregory). With most sites and venues, the overall fantasy winner receives a cash prize. In popular gambling areas like Las Vegas and Atlantic City, this prize can exceed $1 million. Clearly, there is money to be made in this game.

Competition in fantasy football is mostly limited to amateur players. The very nature of the online game – virtually no barrier to entry, low fees, high payouts, and the ability to make football games on Sunday significant – attracts a largely casual base of players with cursory knowledge of the statistics. Could someone research historical data and create a model that predicts how to win, potentially unlocking the secret to winning thousands of dollars? Perhaps if we impart knowledge of financial markets, in which – according to some – an Efficient Market Hypothesis keeps stocks and bonds fairly valued, to the markets in Fantasy Football, one could do just that.

One of the most popular ways to play Fantasy Football is "auction-style." In this game, a Fantasy Football league sets a salary cap (usually $200) of fictitious money. Each NFL player is assigned a market value based on predicted performance. Players who have scored a lot of touchdowns in previous seasons, for example, will command higher prices than rookies or those players who rarely score. In the initial set-up of the teams, NFL players are placed on the auction block. Bidding starts at their market price, and league players can bid depending on their preference. The catch, however, is that a team needs to start a certain amount of quarterbacks (QBs), wide receivers, and running backs, and with the salary cap ceiling looming overhead, a player would not want to buy the top two QBs with their $200. (This year, for example, top QBs Aaron Rodgers and Drew Brees were valued at $45 each.) Thus, players employ various strategies like paying top dollar for top talent, picking a balanced team, or trying to pick up undervalued players in their attempts to win the league.

Ultimately, though, the auction-style format gives every NFL player a market value – akin to a stock price of a company. Just as Wall Street investors try to pick undervalued stocks, one could theoretically pick undervalued players. If the Fantasy Football market is efficient, however, all players would be fairly priced at all times.

In financial markets, the Efficient Market Hypothesis states that it is impossible to "beat the market" because stock market efficiency causes share prices to incorporate and reflect all relevant information. Shares always trade at their fair value, making it impossible for investors to purchase undervalued stocks or sell overvalued stocks. In a world where news makes it to the mass public instantaneously via the Internet and millions of traders exchange thousands of shares every second, one could see how our capital markets could qualify as efficient.

At first, one might point out that the comparatively smaller market of Fantasy Football could not be efficient. The size is laughable compared to the behemoth U.S. capital markets, and player information and updates are usually contained to team locker rooms shielded from the public. Even private companies cannot gain access to player injury reports before the teams themselves publicly report them. However, the Fantasy Football market has grown at an outstanding pace since its creation and now is an industry worth upwards of $1 billion (CNN.com). As many as 26 million Americans play, and increasingly they demand the services of firms like Rotowire, which specialize in fantasy sports updates, advice, and player valuations (Rotowire.com).

The astounding increase in Fantasy Football players and service firms can be compared to individual investors and Wall Street institutions, and player values to share prices. The question then becomes: has the Fantasy Football market in auction-style play – where players are assigned a market value – become truly efficient? If not, there could be serious potential to exploit its inefficiencies and "beat the market."

To discover if the Fantasy Football market is truly efficient, I plan to use historical data to create a valuation model for NFL players in Fantasy Football leagues. The ultimate valuation metric would clearly be the number of points an NFL player posted in a season, but I'd like to delve deeper than that. I would like to create valuation models that might allow one to acquire winning players at low prices, just as an investor chooses undervalued stocks to purchase. These models will be able to predict a player's value before a season so that a participant could tell if a player was undervalued.

The foundation for such models would be historical data and performance from significant NFL players. I would focus on key players in important positions, like QBs of top franchises. The models would incorporate variables to determine a player's value, using metrics ranging from touchdowns scored in previous seasons to playing in a contract year (the NFL year preceding the end of a player's contract). I could then theoretically use these models to choose the highest valued Fantasy Football team at a price that corresponds – in the market – to a lower valuation. This would put me at an advantage in any leagues I joined and turn the odds in my favor. I would be one of the best investors outside Wall Street.

Chapter 2

Related Papers

I am not the first person with the idea to create a predictive model for sports gambling purposes. Many people before me – some at Penn State – have written articles in the fields of betting, statistical modeling, and predictive systems. Ms. Carly Kurkiewicz attempted to predict box office success in her 2008 thesis, "A Financial Analysis of Movies: Anticipating Box Office Success." I will be trying to create a similar model in order to anticipate fantasy football success. Kurkiewicz used mostly regression analysis, tallying the correlations of factors like the "Opening Week ROI vs. Total ROI," "Budget vs. Gross Profit," and "Budget vs. ROI" (Kurkiewicz 11-14). As with her thesis, correlations will be the foundation of my model.

Also in 2008, Ms. Kathleen Hayes wrote a thesis testing the efficiency of the NFL betting market. Ultimately, she discovered that "inefficiency in the NFL betting market, therefore, cannot be proved" (Hayes 26). This thesis related to betting in the NFL, but focused on traditional betting lines produced by longstanding Las Vegas firms. I will be focusing on the efficiency of the betting market within fantasy football. As a relatively new and unique gambling opportunity, the firms that produce each player's "market values" haven't really been tested in terms of efficiency. It is my hope to find a way to exploit possible inefficiencies in this market.

Mr. Jon W. Schultz authored a similar thesis to mine just two years ago, yet focused on soccer players. In his piece, entitled "Building a Winning Soccer Team," Schultz blended traditional soccer statistics like goals scored with intangible ones like contract incentives to create accurate market values for players.

Schultz noted "players are traded and bought in an open market…with no salary cap…so bidding wars [with no limits] occur over players" (Schultz 1). Though the NFL does restrict situations like this with league-imposed salary caps, fantasy football markets operate without restrictions. Players can be bought for any amount of money an owner is willing to pay for them. Schultz goes on to point out "many [soccer] teams are willing to pay high costs for players because they believe having better players will lead to more wins" (Schultz 2). This begs the question of each player's true worth, which Schultz attempts to find based on statistics that have the most correlation with "increasing wins for a team" (Schultz 2). The similarities with my thesis are very clear, as I am attempting to build a model that places a true value on a football player based solely on their contribution to a winning fantasy football team.

Schultz delves into other sports, as well, in an effort to discredit the use of traditional statistics. For example, he points out that baseball's Earned Run Average (ERA) statistic does not stand as a solid measure to evaluate a pitcher's worth and ability (Schultz 4). Instead, Schultz argues, "pitchers should be evaluated with defense independent statistics like the number of strikeouts, walks, and homeruns given up" (Schultz 4). This variance from traditional methods of evaluating a player's value will most likely be seen in my thesis, as well. Like Schultz, I will not give credence to any one statistic simply because it is traditionally used to measure a player. Statistics will be used if a correlation is present with the points they have contributed to fantasy football teams in the past.

Schultz's data mining and analysis essentially mimics what I plan to do. First, he gathered a wide variety of historic statistics on soccer players like goals scored, shots on goal, height, weight, and age. Then, he ran a Principal Component Analysis, Regression, and Correlation in a computer statistical analysis program. Schultz looked for any relationships with the market value of a player. Ultimately, he found out the correlation between each statistic and the market value of the player. From that, he could determine which statistics were the most relevant to use when valuing a player. For example, he discovered that fouls suffered had a strong correlation with market value, perhaps because the player with more fouls was putting in more effort in his game. Though Schultz did not plan to use this information for betting purposes, one could easily apply his work to the sports gambling world.

Like Schultz, Radu Tunaru, Ephraim Clark, and Howard Viney discussed the pricing of soccer players and proposed a new valuation method in their 2005 paper entitled "An Option Pricing Framework for Valuation of Football Players." Unlike Schultz, they found a way to take unexpected events like injuries into account to value a player. Ultimately, they produced a modeling approach to value any player using an Opta Index – a propriety performance statistician index – as the foundation in the equation.

The Opta Index came from a research company of the same name and, since its inception in the late 1990s, had become "the quantitative indicator of the form of the player used by the betting industry, the media, and in fantasy games" (Tunaru et. al 285). The Index used a unique approach to assign value to a player: a six game moving average of each player's Game Score, another proprietary measure that assigned points to a player based on their on-field actions. For instance, a player's Game Score would go up if successful passes, shots, tackles, and saves were made, with higher Game Score translating into higher valuations. Where the Game Score differentiated itself from prior valuation procedures was that it took into account minutes played in each game. Taking a moving average of the Game Score, as the Opta Index did, could then take into account injuries, which would lower a player's value since the Game Score's would be lower. For an individual player, Tunaru et. al assumed, the number of Opta Index points followed a geometric Brownian motion (Tunaru et. al 285). Tunaru et. al believed that player valutation should be based on real options models, which become "one of the best theoretical tools for decision making when the objects analysed [sic] are not market traded" (Tunaru et. al 294).

After player Opta Index points could be determined, Tunaru et. al applied the same formula to individual teams. Game Score became more generalized and now was based on team-performance metrics like goals-scored, wins, goals allowed, etc. The Opta Index for teams was then compared with the Opta Index for players, and the resulting correlation coefficient determined how influential a player was to a team. Tunaru et. al reasoned that doing this, each team in a soccer league could assign a value to a player based not only on the player's individual Opta Index, but on the player's contribution to the team's Opta Index, as well (Tunaru et. al 289).

Mr. David Dorey, on the other hand, addresses American fantasy football specifically. In his book Fantasy Football The Next Level: How to Build a Championship Team Every Season, he proposes a simple three-step guide to valuing football players at any position. Like me, Dorey is attempting to create a system people can use pre-draft to analyze players and determine quality draft choices.

The core of Dorey's hypothesis relies on two overarching variables in valuing a player: previous individual statistics and team performance (Dorey 46). Dorey suggests first focusing on team performance – not individual statistics – in determining a player's true value. As Dorey says, "everything that a football player does is within the context of the team – how he fits into the scheme, his role, and the offense around him" (Dorey 48). Dorey goes on to explain that "this is precisely why projecting players independently makes little sense and yields even less accurate results" (Dorey 48). Though Dorey doesn't go into detail about which team's to target, what team statistics to focus on, etc., he recommends targeting players from "the best offensive teams" of the past years (Dorey 49). This contradicts how many experts predict fantasy football values. Traditionally, a player-first approach is used – which is the method I will be researching – but Dorey's team-first method reveals a interesting avenue that could be researched further by others.

Dorey's second method in predicting player performance is a combination of individual statistics. This is the orthodox method and is used by essentially everyone in predicting player value. Dorey proposes one new wrinkle on that approach, however. Instead of the tradition of using average points per game and historic season performance, Dorey contends that a "three year [window] is needed" in determining a player's true value (Dorey 50). He claims that one cannot solely use previous season stats, nor an average of points scored in the player's career. Instead, Dorey suggests a full investigation of the prior three-year statistics, be it points scored, catches made, touchdowns thrown, etc. With a blend of three-year – not one-year or career – averages and statistics, Dorey suggests people can truly find a player's true value. Only, of course, after targeting the teams to focus on first.

Like Dorey, Mr. Kyle Thompson focuses on American family football. However, Thompson goes further into depth and attempts to analyze individual characteristics to value players. His thesis, written in 2007, explores the significance of two factors in fantasy football leagues: draft position and the use of the quarterback rating statistic to rate fantasy quarterbacks. Although I won't be focusing on draft positions in fantasy football leagues, Thompson's exploration of the quarterback rating is similar to the research I will be doing.

Thompson first hypothesizes that quarterback rating – a measure of an NFL quarterback's efficiency – could not be used to determine fantasy football success.   Validating this view is Thompson's regression analysis of total passing fantasy points versus quarterback rating in the 2006 NFL regular season. (It is key to note Thompson only uses passing fantasy points, which will be touched upon below.) Using data from all starting quarterbacks that season, he arrives at a correlation coefficient of just 0.0452, leading him to believe in practically zero predictive ability for that statistic. By limiting the data to the top twelve quarterbacks (by fantasy points score), Thompson creates a data set more representative of what an actual fantasy league would look like. He assumes that in a ten-team league, only the top twelve quarterbacks would be consistently used as starters (Thompson 10). This produced a correlation of 0.3442, still too weak to show causation.

After a couple more attempts at manipulated the data – breaking down the quarterback rating into its separate parts, choosing the highly correlated pieces, requiring that a quarterback start at least fourteen games, etc. – Thompson arrives at a relatively decent model. From further examination of what the quarterback rating really measures, Thompson reasons that it isolates a passer's efficiency. Therefore, one would need to multiple that efficiency by the number of opportunities (pass attempts per game) to create a total output number. When Thompson runs a regression using the total output of a quarterback who played fourteen or more games against fantasy points per game, the result is a correlation of 0.898. He calculates a residual standard deviation of 1.03, leaving a 95% chance that the actual passing fantasy points per game is +/- 2 away from the predicted value.

While Thompson does an excellent job of manipulating the quarterback rating to predict fantasy points per game, he leaves out one key element. Contemporary quarterbacks are increasingly running the ball, piling up yardage and touchdowns by doing the opposite of throwing the ball. Michael Vick, for example, notched 335.5 total fantasy points in the 2010 season, but 36% of those points came from rushing (fantasyfootballchallenge.com). Thompson's predictive model may be accurate for Vick's passing fantasy points, but would miss Vick's rushing points. If Vick weren't a running quarterback, one could use Thompson's model and be fine. However, since Vick does tend to run the ball a lot, Thompson's model would undervalue him compared to other quarterbacks who don't run as much. Tom Brady – who had just 3% of his total fantasy points come from rushing (fantasyfootballchallenge.com) – would be overvalued next to Michael Vick.

To Thompson's credit, he does acknowledge, "if someone created a measure of efficiency for running, fantasy owners could predict running fantasy points per game…and could then measure the total fantasy value of a quarterback by considering both predicted running and passing fantasy outputs" (Thompson 16). Therein lies an opportunity to use Thompson's model and expand upon it.

Chapter 3

Methodology

As one can see from the related papers on this topic, multiple ways of analyzing the data can be undertaken. However, underlying each analysis is basic statistical analysis. This is what I will be using to evaluate and interpret the data I will collect, focusing on regression analysis to identify relationships between the stats.

The first step in this process is obvious: to analyze data, one must first have data to analyze. Since my central thesis revolves around contemporary fantasy football, I will need historical data. Thanks to its longstanding history and tradition, stats for NFL players are available back to 1932. However, the auction-style fantasy football game has only been around for a couple of years, so historical prices of players are much harder to come by. Only a few websites provide historical prices, and even fewer go back further than a couple years. Therefore, while I can use popular websites like NFL.com and ESPN.com for football statistics, I have to search harder for fantasy statistics.

Fortunately, websites specializing in fantasy football exist. Some of these have historical prices for auction-style drafts dating back two to three years. In addition, ESPN.com actually provides similar information hidden deep in their online archives. Though this historical data is better than nothing, it represents a "least common denominator" and therefore determines how many years I can include in my analysis. Ultimately, I will be able to compile fantasy statistics from third-party websites specializing in fantasy football, filling in any blanks with data from ESPN.com archives.

After I gather the data, I will transition to breaking it down and analyzing it. As stated above, the limited fantasy data places a constraint on how many years of football data I can include; I can't, for example, use QB stats dating back to 2000 if I only have QB prices dating back to 2008. Thus, after assembling data that corresponds with each other, I can start turning to regression analysis.

The ultimate goal of the regression analysis would be to produce a historically accurate model that would predict prices of players. For example, if I started with quarterbacks – usually one of the first positions to get drafted – then I would want a model that would predict the price of significant quarterbacks in the draft. This model could be based on any number of historical data sets. Perhaps the relationship between 2008 points and 2009 price is extremely tight, suggesting a positive relationship. One could then say use the 2009 points to predict the 2010 price. See the below example for details.

Table **3-1**: Sample One-Variable Regression Using Prior Year Points

2008 fantasy points contributed (player X)	2009 price (player X)	➔ Regression analysis produces:	1. Constant (a) 2. Slope coefficient (b)
Therefore: Price2009 = a + b(points contributed 2008)			
Extrapolated: Price2010 = a + b(points contributed 2009)			

The above table simplifies the process, but it covers the essential steps of it. The key here will be to try multiple regressions with different data sets to create the best model. For instance, a two variable regression model might be a better predictor of price. The steps for this type of analysis are below.

Table **3-2**: Sample Two-Variable Regression Using Prior Year Points and Prior Year Price

2009 price (player X)	2008 fantasy points contributed (player X)	2008 price (player X)	➔ Regression analysis produces:	1. Constant (a) 2. Slope coefficient (b)
Therefore: Price2009 = a + b(points contributed 2008) + c(price 2008)				
Extrapolated: Price2010 = a + b(points contributed 2009) + c(price 2009)				

Once I have figured out the optimal model using the above regression analysis, the next step will be to produce the 2010 prices. This can be accomplished by simply plugging in the variables and calculating the output.

After I obtain the predicted 2010 prices for the top twenty players at a position, I can then compare that prediction to the actual prices. Pending a lockout in the NFL, the 2010 prices will be published sometime in the latter half of 2011. To be clear, my model will not predict the price for an individual player, but rather the prices for the top player at a position. This data could then be used to determine if a player is valued correctly.

Chapter 4

Research and Analysis

Using the top 20 quarterbacks as a foundation for my research and predictions, I first performed a two variable regression with their prior year price and prior year points. That regression was insignificant due to T-Statistics ("T-Stat") below 2. *See Appendix A for the regression output.* I then checked to see if my two variables were correlated. In fact, they had a very high correlation of 0.897. Knowing that prior year price and prior year points were indeed highly aligned, I then ran two separate regressions using each variable. Using just the prior year points, a T-Stat of over 4 was produced. *See Appendix B for the regression output.* Using prior year price, a T-Stat of 3.73 was produced. *See Appendix C for the regression output.* With these statistically-significant regressions, I then produced the predictions of 2010 prices. *See Appendix D for the 2010 price-prediction model.* Ultimately, I ended up with three models: a statistically insignificant one of both prior year price and prior year points, and two statistically significant ones using each variable separately.

After completion of the models, I had to wait for updated 2010 prices. When they came out (found on sites like ESPN.com), I inputted them into my model and checked the variances, in percentage terms, for the top 20 QBs. *The results are found in Appendix E.* The variance between actual and predicted points in 2010 was decidedly smaller when using the prior year point's model. However, it should be noted that isolating these two variances proves nothing – we need to further regress the data to see if, in fact, the models show undervalued and overvalued players. Thus, I embarked on the final steps in deciding if either model could target and predict inaccurate prices and incorrectly valued players.

The first step of these final calculations is to regress the actual 2010 season points with the actual 2010 starting prices of the top 20 quarterbacks going into the 2010 season. To be clear, I define the top 20 quarterbacks going into the 2010 season as the top point scorers in the prior season, 2009. *See Appendix F for the regression output from the regression of 2010 actual points with 2010 actual price.* The results of that regression produced 20 residuals, or estimates of the unobservable statistical error for each quarterback. The next step in this process was to compare those residuals with the difference in actual and predicted price – for each of my two final models – for the top 20 quarterbacks. A regression analysis between these two groups would show if either of my models could predict under- or overvalued players.

Thus, the final step of my research focused on regressing two items: the residuals from the 2010 actual points and actual price, and the differences between the actual points and the predicted points from my two models. To clarify, my two final models were based on prior year points and prior year prices. *See Appendix G for the 2010 actual price vs. 2010 actual points regression results using the prior year points model, and see Appendix H for the same results using the prior year price model.* Unfortunately, both regressions produced results that were insignificant due to T-Statistics below two. With a sample size of just twenty, it'd be hard to expect any significant results. One takeaway, however, is that both coefficients in the regressions were negative, so some quarterbacks who were overpriced scored fewer points than expected.

Chapter 5

Conclusion

Determining if the Efficient Market Hypothesis applies to Fantasy Football leagues, let alone capital markets, is a hard task to accomplish. A semblance of efficiency, however, can be seen in both. Like stocks, football player values fluctuate with every new piece of news, past performance, and injury updates. My one-variable regression models based on prior year points and prior year price produced predictive prices for top quarterbacks, but how successful those predictions were – and if they could be used to exploit inefficiencies in the market – proved tough to verify. A regression between residuals from the quarterbacks and their 2010 actual outcomes, and the difference between my models' predictions and actual prices, produced insignificant results due to low T-Statistics. However, as stated earlier, both outputs produced negative coefficients, showing that some overpriced quarterbacks scored lower points then expected. This points to inefficiency in the fantasy football market that one could take advantage of; though overpriced players could not be "shorted," per se, they could be bought and immediately sold or just simply passed over. Either way, players could exploit the inefficiency to their benefit. Unfortunately, the statistical insignificance means no meaningful conclusion can be drawn. My sample size of just twenty was small, as was my narrow focus on quarterbacks. Going forward, I would wonder what could be determined if this research was performed on a larger scale. Maybe then we could prove that the fantasy football market was efficient or inefficient. For now, however, the Efficient Market Hypothesis can still only be applied to capital markets.

Chapter 6

Bibliography

Bramhall, Joe. National Football League. Hoovers, Inc. 2011. http://cobrands.hoovers.com.ezaccess.libraries.psu.edu/global/cobrands/proquest/factsheet.xhtml?ID=40330.

Dorey, David. Fantasy Football The Next Level: How to Build a Championship Team Every Season. New York: Warner Books, 2007.

Gregory, Sean. “Fantasy Football: Is it going to our heads.” Aug 31, 2009. TIME.

Hayes, Kathleen Meaghan. “ NFL Expert Picks: A Comparative Analysis of Forecasting Success Across Sports Networks.” 2008.

St. Amant, Mark. Committed: Confessions of a Fantasy Football Junkie. Sep. 2004. 294p. Scribner, $24 (0-7432-6756-7). 790.332.

Kurkiewicz, Carly. “A Financial Analysis of Movies: Anticipating Box Office Success.” 2008.

“Michael Vick/Tom Brady Fantasy Points.” Fantasyfootballchallenge.com. 2011.

Schultz, Jon. “Building a Better Soccer Team.” 2009.

Thompson, Kyle. “Statistical Investigations in Fantasy Football: The Inequality of Draft Positions and Exploration of Quarterback Rating’s Purpose in Fantasy Football.” 2007.

Tunaru, Radu and Clark, Ephraim, and Viney, Howard. “An Option Pricing Framework for the Valuation of Football Players.” *Review of Financial Economics.* City University: Case Business School. 23 March 2005.

Appendix A

Two Variable Regression with Prior Year Price and Prior Year Points: Top 20 QBs

SUMMARY OUTPUT

Equation generated by regression
price 2009 = -21.95 + points2008(.12)+price2008(.24)

Regression Statistics	
Multiple R	0.696482463
R Square	0.485087821
Adjusted R Square	0.424509917
Standard Error	7.63347138
Observations	20

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	933.2119498	466.6059749	8.007669355	0.003545944
Residual	17	990.5880502	58.26988531		
Total	19	1923.8			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-21.95520595	18.1710815	-1.20824982	0.243490798	-60.29283676	16.38242486	-60.29283676	16.3824249
Points in 2008	0.123185236	0.096232563	1.280078521	0.217710148	-0.079847723	0.326218196	-0.079847723	0.3262182
Price in 2008	0.240879343	0.456382203	0.527801789	0.604458404	-0.722002938	1.203761625	-0.722002938	1.20376162

RESIDUAL OUTPUT

Observation	*Predicted Y*	*Residuals*
1	24.33187654	14.66812346
2	20.9156367	-2.915636703
3	19.92466368	-10.92466368
4	18.95565518	-1.95565518
5	15.38328332	11.61671668
6	14.90152464	-2.901524638
7	13.56296043	-2.562960426
8	9.402118037	-7.402118037
9	11.20047642	4.799523583
10	8.786191855	-8.786191855
11	10.34367089	-9.343670892
12	4.77049922	-2.77049922
13	5.840728441	-0.840728441
14	4.748534702	4.251465298
15	4.989414045	9.010585955
16	3.04590591	0.95409409
17	2.794044308	-1.794044308
18	2.183609256	-1.183609256
19	4.228338109	-3.228338109
20	1.69086831	11.30913169

Appendix B

One Variable Regression with Prior Year Points: Top 20 QBs

SUMMARY OUTPUT

Regression Statistics			
Multiple R	0.690398497		
R Square	0.476650084		Equation generated by regression
Adjusted R Squa	0.447575089		price 2009 = -29.88 + points2008(.16)
Standard Error	7.478935485		
Observations	20		

ANOVA

	df	SS	MS	F	Significance F
Regression	1	916.9794323	916.9794323	16.39381466	0.00075314
Residual	18	1006.820568	55.93447598		
Total	19	1923.8			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-29.8842249	10.01586458	-2.983688992	0.007962756	-50.926776	-8.841674243	-50.9267755	-8.8416742
Points in 2008	0.16874541	0.041676569	4.048927594	0.00075314	0.08118619	0.256304632	0.081186188	0.25630463

RESIDUAL OUTPUT

Observation	Predicted Y	Residuals
1	24.28305171	14.71694829
2	21.58312515	-3.583125148
3	19.89567105	-10.89567105
4	19.55818023	-2.558180229
5	14.66456334	12.33543666
6	14.66456334	-2.66456334
7	13.82083629	-2.82083629
8	9.770946451	-7.770946451
9	9.264710221	6.735289779
10	8.927219401	-8.927219401
11	8.420983171	-7.420983171
12	6.396038251	-4.396038251
13	5.552311201	-0.552311201
14	5.046074971	3.953925029
15	5.046074971	8.953925029
16	4.033602512	-0.033602512
17	3.358620872	-2.358620872
18	2.852384642	-1.852384642
19	2.683639232	-1.683639232
20	2.177403002	10.822597

Appendix C

One Variable Regression with Prior Year Price: Top 20 QBs

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.659891138
R Square	0.435456314
Adjusted R Squa	0.404092776
Standard Error	7.767700731
Observations	20

Equation generated by regression
price 2009 = .99 + price2008(.76)

ANOVA

	df	SS	MS	F	Significance F
Regression	1	837.7308563	837.7308563	13.88415784	0.001545957
Residual	18	1086.069144	60.33717465		
Total	19	1923.8			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.997555525	2.997401402	0.332806785	0.743124961	-5.299751145	7.29486219	-5.299751145	7.294862194
Price in 2008	0.7649113	0.205282206	3.726145171	0.001545957	0.333629389	1.19619321	0.333629389	1.196193211

RESIDUAL OUTPUT

Observation	Predicted Y	Residuals
1	22.41507194	16.58492806
2	17.82560413	0.174395865
3	18.59051544	-9.59051544
4	16.29578153	0.704218466
5	16.29578153	10.70421847
6	14.76595893	-2.76595893
7	12.47122503	-1.47122503
8	8.646668529	-6.64666853
9	15.53087023	0.469129767
10	8.646668529	-8.64666853
11	14.76595893	-13.7659589
12	1.762466825	0.237533175
13	7.116845928	-2.11684593
14	4.822112027	4.177887973
15	5.587023327	8.412976673
16	1.762466825	2.237533175
17	2.527378125	-1.52737813
18	1.762466825	-0.76246682
19	8.646668529	-7.64666853
20	1.762466825	11.23753318

Appendix D

2010 Price Prediction Model

2010 Predictions:			
	points and price	prior yr points	prior yr price
	price 2010 = -21.95 + points2009(.12)+price2009(.24)	price 2010 = -29.88 + points2009(.16)	price 2010 = .99 + price2009(.76)
QB 1	$ 30.55	$ 29.18	$ 30.83
QB 2	$ 21.18	$ 23.27	$ 14.77
QB 3	$ 16.80	$ 20.23	$ 7.88
QB 4	$ 18.36	$ 19.73	$ 14.00
QB 5	$ 20.64	$ 19.56	$ 21.65
QB 6	$ 16.66	$ 19.05	$ 10.18
QB 7	$ 15.80	$ 18.21	$ 9.41
QB 8	$ 13.51	$ 18.04	$ 2.53
QB 9	$ 16.88	$ 18.04	$ 13.24
QB 10	$ 10.20	$ 14.16	$ 1.00
QB 11	$ 9.21	$ 12.47	$ 1.76
QB 12	$ 8.83	$ 11.63	$ 2.53
QB 13	$ 9.31	$ 11.29	$ 4.82
QB 14	$ 9.65	$ 10.45	$ 7.88
QB 15	$ 9.87	$ 9.10	$ 11.71
QB 16	$ 6.36	$ 7.58	$ 4.06
QB 17	$ 5.63	$ 7.58	$ 1.76
QB 18	$ 3.05	$ 4.03	$ 1.76
QB 19	$ 0.95	$ 1.16	$ 1.76
QB 20	$ 3.84	$ 1.16	$ 10.94

Appendix E

2010 Price Prediction Model with 2010 Prices and Variance

2010 Predictions & Actuals:	prior yr points				prior yr price			
	predicted price 2010	actual 2010 price	variance ($)	variance (%)	predicted price 2010	actual 2010 price	variance ($)	variance (%)
Drew Brees, NO	$ 29.18	33	$ (3.82)	-13%	$ 30.83	33	$ (2.17)	-7%
Tom Brady, NE	$ 23.27	34	$ (10.73)	-46%	$ 14.77	34	$ (19.23)	-130%
Peyton Manning, IND	$ 20.23	24	$ (3.77)	-19%	$ 7.88	24	$ (16.12)	-205%
Kurt Warner, ARI	$ 19.73	0	$ 19.73	100%	$ 14.00	0	$ 14.00	100%
Aaron Rodgers, GB	$ 19.56	42	$ (22.44)	-115%	$ 21.65	42	$ (20.35)	-94%
Philip Rivers, SD	$ 19.05	32	$ (12.95)	-68%	$ 10.18	32	$ (21.82)	-214%
Tony Romo, DAL	$ 18.21	23	$ (4.79)	-26%	$ 9.41	23	$ (13.59)	-144%
Matt Ryan, ATL	$ 18.04	20	$ (1.96)	-11%	$ 2.53	20	$ (17.47)	-691%
Matt Schaub, HOU	$ 18.04	17	$ 1.04	6%	$ 13.24	17	$ (3.76)	-28%
Donovan McNabb, PHI	$ 14.16	3	$ 11.16	79%	$ 1.00	3	$ (2.00)	-201%
Matt Cassel, KC	$ 12.47	4	$ 8.47	68%	$ 1.76	4	$ (2.24)	-127%
Jay Cutler, CHI	$ 11.63	4	$ 7.63	66%	$ 2.53	4	$ (1.47)	-58%
Ben Roethlisberger, PIT	$ 11.29	15	$ (3.71)	-33%	$ 4.82	15	$ (10.18)	-211%
Carson Palmer, CIN	$ 10.45	0	$ 10.45	100%	$ 7.88	0	$ 7.88	100%
Eli Manning, NYG	$ 9.10	8	$ 1.10	12%	$ 11.71	8	$ 3.71	32%
Kyle Orton, DEN	$ 7.58	3	$ 4.58	60%	$ 4.06	3	$ 1.06	26%
Brett Favre, MIN	$ 7.58	0	$ 7.58	100%	$ 1.76	0	$ 1.76	100%
David Garrard, JAC	$ 4.03	1	$ 3.03	75%	$ 1.76	1	$ 0.76	43%
Matt Hasselbeck, SEA	$ 1.16	1	$ 0.16	14%	$ 1.76	1	$ 0.76	43%
Trent Edwards, BUF	$ 1.16	0	$ 1.16	100%	$ 10.94	0	$ 10.94	100%
		Average variance	$ 0.60	22%		Average variance	$ (4.48)	-78%

Appendix F

One-Variable Regression with 2010 Actual Price and 2010 Actual Points: Top 20 QBs

	pts scored in 2010	price at start of 2010
Drew Brees, NO	330	39
Tom Brady, NE	312	36
Peyton Manning, IND	335	27
Kurt Warner, ARI	0	19
Aaron Rodgers, GB	339	18
Philip Rivers, SD	330	17
Tony Romo, DAL	117	16
Matt Ryan, ATL	281	14
Matt Schaub, HOU	293	13
Donovan McNabb, PHI	225	12
Matt Cassel, KC	249	11
Jay Cutler, CHI	261	9
Ben Roethlisberger, PIT	240	9
Carson Palmer, CIN	281	8
Eli Manning, NYG	300	5
Kyle Orton, DEN	252	4
Brett Favre, MIN	159	2
David Garrard, JAC	263	2
Matt Hasselbeck, SEA	210	1
Trent Edwards, BUF	37	0

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.363582388
R Square	0.132192153
Adjusted R Square	0.083980606
Standard Error	91.27147888
Observations	20

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	22841.50856	22841.50856	2.741919	0.115076436
Residual	18	149948.6914	8330.482858		
Total	19	172790.2			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	198.9591075	32.43386825	6.13430091	8.57E-06	130.8180788	267.1001361	130.8180788	267.1001361
X Variable 1	3.186327674	1.924257249	1.655874065	0.115076	-0.856386792	7.229042139	-0.856386792	7.229042139

RESIDUAL OUTPUT

Observation	*Predicted Y*	*Residuals*
1	323.2258867	6.774113255
2	313.6669037	-1.666903725
3	284.9899547	50.01004534
4	259.4993333	-259.4993333
5	256.3130056	82.6869944
6	253.1266779	76.87332207
7	249.9403503	-132.9403503
8	243.5676949	37.43230509
9	240.3813672	52.61863277
10	237.1950396	-12.19503956
11	234.0087119	14.99128811
12	227.6360565	33.36394346
13	227.6360565	12.36394346
14	224.4497289	56.55027114
15	214.8907458	85.10925416
16	211.7044182	40.29558183
17	205.3317628	-46.33176282
18	205.3317628	57.66823718
19	202.1454351	7.85456485
20	198.9591075	-161.9591075

Appendix G

One Variable Regressions with Residuals and Actual-Predicted Prices: Prior Year Points Model: Top 20 QBs

		2010 Actual - Predicted Price	Residuals
Drew Brees,	$	9.82	6.774113255
Tom Brady, I	$	12.73	-1.666903725
Peyton Man	$	6.77	50.01004534
Kurt Warner	$	(0.73)	-259.4993333
Aaron Rodge	$	(1.56)	82.6869944
Philip Rivers,	$	(2.05)	76.87332207
Tony Romo,	$	(2.21)	-132.9403503
Matt Ryan, A	$	(4.04)	37.43230509
Matt Schaub	$	(5.04)	52.61863277
Donovan Mc	$	(2.16)	-12.19503956
Matt Cassel,	$	(1.47)	14.99128811
Jay Cutler, C	$	(2.63)	33.36394346
Ben Roethlis	$	(2.29)	12.36394346
Carson Palm	$	(2.45)	56.55027114
Eli Manning,	$	(4.10)	85.10925416
Kyle Orton, I	$	(3.58)	40.29558183
Brett Favre,	$	(5.58)	-46.33176282
David Garrar	$	(2.03)	57.66823718
Matt Hassell	$	(0.16)	7.85456485
Trent Edwar	$	(1.16)	-161.9591075

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.030951288
R Square	0.000957982
Adjusted R Square	-0.054544352
Standard Error	4.93926821
Observations	20

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.421086605	0.421086605	0.01726	0.896933477
Residual	18	439.1346682	24.39637045		
Total	19	439.5557548			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.695524478	1.104453948	-0.62974512	0.536774	-3.01589612	1.624847163	-3.01589612	1.624847163
X Variable 1	-0.00167577	0.012755317	-0.13137814	0.896933	-0.028473697	0.025122158	-0.028473697	0.025122158

RESIDUAL OUTPUT

Observation	*Predicted Y*	*Residuals*
1	-0.706876333	10.53020774
2	-0.692731131	13.42215188
3	-0.779329805	7.546167936
4	-0.260663319	-0.46626232
5	-0.834088851	-0.724091378
6	-0.824346474	-1.227597525
7	-0.472747047	-1.735469902
8	-0.758252407	-3.281219132
9	-0.783701197	-4.255770342
10	-0.675088399	-1.483238711
11	-0.720646427	-0.750226583
12	-0.751434769	-1.875711191
13	-0.716243602	-1.573411538
14	-0.790289718	-1.655638373
15	-0.838148001	-3.25781681
16	-0.7630506	-2.814205521
17	-0.617883107	-4.959373014
18	-0.792163172	-1.24143934
19	-0.708686921	0.543756379
20	-0.424118289	-0.740812253

Appendix H

One Variable Regressions with Residuals and Actual-Predicted Prices: Prior Year Price Model: Top 20 QBs

		2010	
		Actual - Predicted Price	Residuals
Drew Brees,	$	8.17	6.774113255
Tom Brady, I	$	21.23	-1.666903725
Peyton Man	$	19.12	50.01004534
Kurt Warner	$	5.00	-259.4993333
Aaron Rodge	$	(3.65)	82.6869944
Philip Rivers	$	6.82	76.87332207
Tony Romo,	$	6.59	-132.9403503
Matt Ryan, A	$	11.47	37.43230509
Matt Schaub	$	(0.24)	52.61863277
Donovan Mc	$	11.00	-12.19503956
Matt Cassel,	$	9.24	14.99128811
Jay Cutler, C	$	6.47	33.36394346
Ben Roethlis	$	4.18	12.36394346
Carson Palm	$	0.12	56.55027114
Eli Manning,	$	(6.71)	85.10925416
Kyle Orton, I	$	(0.06)	40.29558183
Brett Favre,	$	0.24	-46.33176282
David Garrar	$	0.24	57.66823718
Matt Hassell	$	(0.76)	7.85456485
Trent Edwar	$	(10.94)	-161.9591075

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.068615747
R Square	0.004708121
Adjusted R Square	-0.050585873
Standard Error	91.05636682
Observations	20

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	705.9765391	705.9765391	0.085147	0.773776104
Residual	18	149242.7149	8291.261939		
Total	19	149948.6914			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-3.369606286	23.40751517	-0.14395404	0.887136	-52.54700886	45.80779629	-52.54700886	45.80779629
X Variable 1	0.769871876	2.638357475	0.291799684	0.773776	-4.773115783	6.312859536	-4.773115783	6.312859536

RESIDUAL OUTPUT

Observation	*Predicted Y*	*Residuals*
1	2.920942722	3.853170533
2	12.97788475	-14.64478848
3	11.34899115	38.66105419
4	0.478946553	-259.9782798
5	-6.179762305	88.8667567
6	1.883621291	74.98970078
7	1.702633113	-134.6429834
8	5.462842644	31.96946245
9	-3.551401007	56.17003377
10	5.100866287	-17.29590585
11	3.742110713	11.2491774
12	1.613483262	31.7504602
13	-0.153167833	12.51711129
14	-3.278574502	59.82884564
15	-8.532608622	93.64186278
16	-3.413643516	43.70922535
17	-3.186736175	-43.14502665
18	-3.186736175	60.85497335
19	-3.956608051	11.8111729
20	-11.79308431	-150.1660232

Keith Armington Academic Vita

Keith D. Armington

School:
340 E. Beaver Ave, Apt. 225
State College, PA 16801

Keith.Armington@gmail.com
215.450.3549

Home:
207 Larrimore Lane
Erdenheim, PA 19038

EDUCATION

The Schreyer Honors College at the Pennsylvania State University — University Park, PA
Bachelor of Science in Finance, The Smeal College of Business — *Expected Graduation: Dec. 2011*
Minor in Engineering Entrepreneurship, The College of Engineering

WORK & LEADERSHIP EXPERIENCE

PricewaterhouseCoopers, LLC — Philadelphia, PA
Advisory Intern, Health Industries — *Summer 2011*

- Developed indexing model for top pharmaceutical company analyzing product data and identifying reporting errors; identified over 100 discrepancies that if corrected could generate significant financial savings (estimated savings for companies in similar situations: $3 million); presented to client with director and manager.
- Improved firm strategy for identifying potential clients by developing a standardized and transferable process to gather and evaluate company-specific indicators; produced report used by senior Health Care managing partner analyzing more than 3,000 health care companies.
- Prepared report for top partners and internal departments detailing how to more effectively communicate PwC's business strategy to lower-level employees.

Johnson and Johnson — Fort Washington, PA
Marketing Finance Co-op, McNeil Consumer Healthcare — *Spring-Summer 2010*

- Helped manage $150 million of brand marketing expenses for Tylenol brands; advised marketing and financial business partners on investments to promote long-term brand equity and drive short-term sales.
- Quantitatively analyzed 2010 product recalls to predict future effects on company's brands; presented to executive level management, prompting realignment in long-term strategy to further counter private-label (generic) competition.
- Managed trade-related recall expenses, communicating with vendors and advising management on objectives and outcomes, and successfully capped total expenses (initially unlimited) at $1.3 million.
- Given responsibilities of full-time analyst when position became vacant; lead team through projects and day-to-day tasks for one month, and then transitioned full-time replacement into the position.

UBS AG — Conshohocken, PA
Intern, Wealth Management Americas — *Summer 2010*

- Assisted investment team managing $400 million by analyzing stock market trends, assisting in transactions, and providing insight from investment ideas to economic forecasts.

Starwood Hotels and Resorts — Honolulu, HI; New York City, NY; Marshalls Creek, PA
Intern/Extern — *Summer 2008-2009*

- Interned at three Starwood hotels around the country, working in all departments.
- Prepared financial commentary for Area Controller of NYC Sheratons (largest U.S. revenue-producers).

Keith A's Service — Erdenheim, PA
Founder and President — *2002-2007*

- Started landscaping and snow removal business, hired employees, and managed twenty regular customers generating annual revenue of $7,000 and 100% customer retention.

SERVICE & EXTRACURRICULAR ACTIVITIES

Impact Thrift Organization *(a non-profit agency)* — Hatboro, PA
May-June 2011

- Prepared feasibility analysis on franchising Impact's successful thrift-store business model and helped with various store operations *(105 volunteer hours).*

Schreyer Honor College Student Council — Penn State University
2007-present

- Plan and execute informative visiting weekends and communicate with prospective students.

Highlighting Humanity Project — Penn State University
Fall 2007

- Led group conducting campus-wide program examining motivation of suicide bombers.