APPLICATION OF DATA ENVELOPMENT ANALYSIS TO IDENTIFY UNDERVALUED EQUITIES ON THE DOW JONES INDUSTRIAL AVERAGE

A Senior Project submitted

In Partial Fulfillment

of the Requirements for the Degree of

Bachelor of Science in Industrial Engineering

the Faculty of California Polytechnic State University,

San Luis Obispo

by

Ryan L. Kotzebue

June 2013

Graded by:_______________ Date of Submission___________________________

Checked by:______________ Approved by:________________________________

Abstract

APPLICATION OF DATA ENVELOPMENT ANALYSIS TO
IDENTIFY UNDERVALUED EQUITIES ON THE
DOW JONES INDUSTRIAL AVERAGE

Ryan L. Kotzebue

Two factors that drive investors to and away from the stock market are reward and risk, respectively. By using a stock selection strategy that is quantitative, investors may feel more comfortable and secure with their decisions. However, there lacks a quantitative strategy that can produce increased returns with lower risk by purchasing a small number of stocks. The objective of this project was to formulate a quantitative stock trading strategy that produced exceptional returns with low risk while also fulfilling additional requirements to benefit the common investor.

By using a linear programming based operations research technique known as data envelopment analysis (DEA), a solution was generated that produced a portfolio of stocks that experienced superior performance to the Dow Jones Industrial Average over eight years. From the results, it is reasonable to conclude that data envelopment analysis is a suitable tool for generating a portfolio of stocks that is superior to the pool of stocks it was created from. It is also safe to recommend the use of data envelopment analysis to the common investor by selecting stocks exactly as shown in this project or to the institutional investor by developing DEA efficient exchange-traded-funds (ETFs).

Acknowledgements

I would like to thank Ycharts.com for allowing me access to their professional tools and resources. Also, my advisor Dr. Yang, it was his class that sparked the idea for this project and he's was very supportive and helpful throughout its duration.

Table of Contents

List of Tables

List of Figures

I. Introduction

Many people are attracted to the stock market as a way to invest their capital because of the potential to earn exceptional returns on their investments. However, deciding which stock to purchase is difficult due to the many attributes that need to be considered. There is no sure way to guard against capital depreciation in the stock market, so the careful investor must take into account many variables in every decision. Traditionally, there has been a tradeoff of risk versus reward in investing, with the goal being to generate the greatest return while keeping risk to an acceptable level.

There are thousands of publicly traded companies available to invest in, making it difficult to analyze each one in-depth. When trying to manage the many variables, it is easy for the common investor to feel overwhelmed. Thus, many investors use some sort of quantitative strategy to determine which stocks to buy. There are thousands of quantitative trading strategies that claim to outperform the stock market, some of which have been proven over decades, such as the Dogs of the Dow (*dogsofthedow.com*). The first step in most of these strategies is the same and involves screening for stocks within a specific range of financial ratios. It is simple to use a free stock screener (www.ycharts.com/stock_screener) to find stocks that have criteria that match an investment style, such as growth, value, dividend yield, etc. However, the next step of filtering through these results is much more difficult and is where most strategies experience weakness. Some investors choose to buy nearly every stock from their initial screen and create a very large portfolio, like a mutual-fund. Problems with doing this are that the transaction fees from the sheer volume of stocks that need to be purchased would deteriorate any possible gains. Also, a large portfolio is less likely to achieve superior performance since any large number of stocks is likely to mimic the market average over time. Another common strategy is to rank results from the initial screen based on a single criteria, such as price-to-earnings ratio or dividend yield, and buy an arbitrary number of stocks from the top of this list. The problem here is that the decision to purchase a stock is being based off of a single variable. Also, forcing the purchase of an arbitrary number of stocks could be very unfavorable in a bull or bear market where more or less exposure

is desired. A key weak point in both of these methods is the fact that neither takes into account the interaction between all the variables that were screened for.

Successful non-quantitative investors analyze and understand the interactions of a company's financials to determine its potential for investment. Therefore, a quantitative strategy should use a similar approach. This problem typically falls outside the scope of those with traditional finance backgrounds, which may be why there are limited solutions. But from an industrial engineering perspective, this problem could be approached with operations research techniques, which lends the possibility of generating a unique, unexplored solution.

The purpose of this project is to use company financial data to optimize the selection of stocks based on the desired investment strategy whether it is growth, value, dividend yield, etc. A successful solution would not only create a portfolio with exceptional returns with low risk, but would also satisfy the following:

- Low trading frequency
- Concentrated portfolio
- No specified number of holdings
- Reasonable formulation

The objective of this project is to design a trading strategy that will satisfy all of the requirements stated above using an industrial engineering based methodology to select stocks from an initial portfolio. The selected stocks should make efficient use of the desired variables and have increased returns as a whole when compared to the original portfolio. The solution should also be repeatable for both an individual investor as well as an institutional fund.

The method chosen to reach these objectives was an operations research technique known as data envelopment analysis (DEA). This form of analysis was selected as the optimal technique for this problem because it can effectively indicate the most desirable units from a large data set when many variables must be considered. The piecewise solution

frontier of DEA is also able to uncover unique relationships that other methodologies miss. In DEA, each unit is referred to as a decision-making-unit (DMU) and the output of desirable DMUs are referred to as efficient.

The data set that will be used for developing and testing the solution will be the stocks on the Dow Jones Industrial Average, which may be referred to as the Dow. The Dow was chosen because of its relatively small size and the ability to easily access its historical information. Also, the Dow contains stocks from a variety of industries and is similar in style to a portfolio that a typical investor may hold.

This paper will focus on providing a solution based on a value investing style. Value investing is an attempt to purchase stocks that are worth more than the current market price. This differs from growth investing where an investor buys stocks in companies with a large potential for growth, regardless of current price. Value investing was chosen due to its popularity among investors, but also because it is more in-line with the investment style that should be associated with the stocks on the Dow. The Dow is made up of mostly large-cap stocks that pay dividends, which are typical features that value investors look for. The requirement stated earlier of a *low trading frequency* is reinforced with a value oriented approach, since it may take more time for a stock to rise from its undervalued state (*Rousseu, Rensburg, 2004*).

The tasks that must be accomplished to solve this problem are as follows:

1. Determine inputs and outputs for the data envelopment analysis
2. Gather historical information
3. Perform data envelopment analysis
4. Formulate results
5. Check results for significance
6. Draw conclusions
7. Make recommendations

II. Background

Discussion

The efficient market hypothesis states that due to the competition among traders, every stock is always fairly priced based on the information available. This means, in theory, a stock should never be considered ‘undervalued.’ Therefore, it is impossible to outperform the market by strategically selecting stocks. Jensen supported the efficient market hypothesis in 1968 when he showed that mutual fund managers did not outperform a random selection of stocks.

However, numerous studies provide evidence against the efficient market hypothesis and conclude there are factors that influence the future performance of a stock. Notably, Fama and Fench in 1992 provided a theory known as the ‘size-effect’, which states that smaller firms will tend to yield higher returns. They were also able to find a correlation with high returns and a low market-to-book ratio. Similar studies were performed by Hamao and Lakonishok (1993), as well as Gaunt (2004), who concluded similar results on the Japanese and Australian stock markets, respectively. More recently in 2007, Anderson and Brooks found that small portfolios of stocks with low price-to-earnings ratios (P/E) outperformed portfolios of high P/E stocks in excess of 30% compounded annually.

The results of these prior studies are crucial for this project as they confirm that current information can be used to forecast the future performance of a stock. While this prior work made important conclusions, most lack any recommendations of how a common investor could use the information to their benefit. Another issue is that in many cases an unrealistic sample, such as the entire stock market, was used. This is unrealistic because the inclusion of micro-cap stocks that have outlier financial ratios may experience price changes of 1000% or more in a day and could skew results.

Data Envelopment Analysis

Data envelopment analysis (DEA) is an increasingly popular operations research tool used to evaluate the relative efficiency of a number of units called decision-making-units (DMUs). Since its development in the late 1970s, DEA has been adopted by management to determine the most efficient branches in the healthcare, education, banking, manufacturing, human resources and utilities sectors (*Anderson, 1996*). It is useful because it is also able to calculate the amount of resources necessary to make an inefficient unit efficient – that is, as efficient as the best unit. With efficiency being defined as:

$$Efficiency = \frac{outputs}{inputs}$$

For a DMU to be considered efficient, the outputs must be sufficient to offset the corresponding inputs (*Charnes, Cooper, Rhodes, 1978*). Linear programming is used to determine the set of weights for the inputs and outputs that provide the maximum efficiency value on a scale of zero to one. By using linear programming, DEA is a much more powerful tool then other techniques such as regression. The mathematical model for this linear program is shown below:

$$Maximize \sum_{r=1}^{s} u_r y_{ro}$$

Where:

j = decision making unit (DMU) being compared

y_{rj} = amount of output r used by DMU j

x_{ij} = amount of input i used by DMU j

i = number of inputs

r = number of outputs

u_r = coefficient assigned by DEA to output r

v_i = coefficient assigned by DEA to input i

Subject to the constraints:

$$\sum_{r=1}^{s} v_i x_{io} = 1$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, j = 1,\ldots,n$$

$$u_r, v_i \geq 0$$

The efficiency score from DEA is essentially a benchmarking of the output to input ratios for each unit. A simplified example can be explained using baseball players as the decision-making-units, as follows:

Recruiting college athletes to professional teams is a difficult process that involves many factors, not unlike the stock market. Assume a scout for the San Francisco Giants is looking for a new power-hitter and has their eye on six college athletes, players A through F. Each player had x at-bats last season, with y_1 home-runs, and y_2 RBIs. The scout decides to use DEA to make his selections of whom to draft. For their analysis, number of at-bats (x) would be the input and number of home-runs (y_1) and RBIs (y_2) would be the outputs. Figure 1 on the following page shows the plot from the scout's analysis. The axis are ratios of the outputs, y_1 and y_2, to input x. The frontier line L(y) is the piecewise line that envelops the data points. Any player who lies on the frontier should be considered efficient and drafted by the team, in this case player A, B and C. From this plot the scout can also tell how many more home-runs and RBIs players D through F would've needed to be as efficient as players A through C.

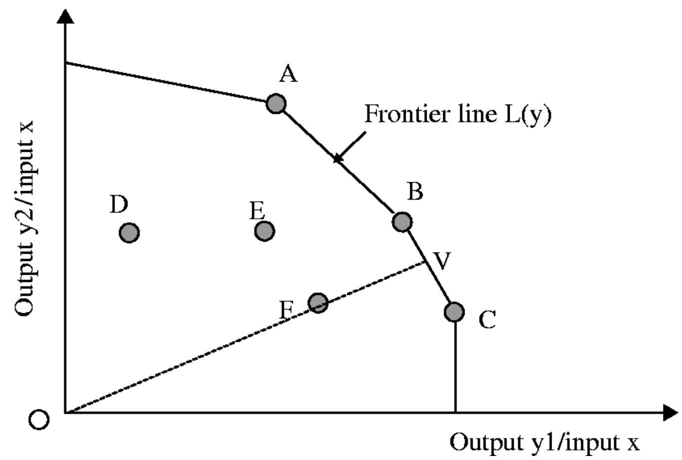


Figure 1. Graphical DEA solution

Source: "Whole life cycle performance measurement re-engineering for the UK National Health Service estate"

Ideally, DEA should be used when there are numerous DMUs with multiple inputs and outputs, where a trivial solution may not be obvious. But even in simple cases like shown above, DEA still has benefits such as determining how much improvement a DMU needs to be efficient.

The advantages of DEA can be seen from the example case just presented. First, multiple inputs and outputs can be used along with any unit of measurement. Notice that the scout used at-bats, home-runs, and RBIs, but it very well could have been pitches thrown, strikeouts, and fastball-speed if they were analyzing a pitcher. Also, DEA can unveil the sources of inefficiency for each DMU. This can be an extremely useful management tool to cut costs and increase productivity. In the baseball example, the scout could see that player F needs to be getting more RBIs and player D more home-runs, respectively. Finally, due to the piecewise frontier, DEA achieves unique results compared to linear methods. Looking back at Figure 1, if a linear regression was done on those data points the best-fit-line may start at the origin and pass through a point just under player E. If this were the case player A, D and E would be considered 'above average' and drafted by the scout, missing out on the talent in players B and C. Instead of comparing against an average, like most statistical techniques do, DEA compares against the optimums.

Of course, DEA has drawbacks as well. Dr. Tim Anderson from Portland State University lists limitations to DEA on his website, as does Sanford Berg in his 2010 book on water utility benchmarking.

1. DEA measures efficiency as a benchmark to the other DMUs but "not compared to a theoretical maximum." (Anderson 1996)
2. The number of efficient DMUs "tends to increase with the number of inputs and outputs." (Berg 2010)
3. Results depend on the chosen inputs and outputs. (Berg 2010)

Applications in the Stock Market

This project will not be the first time data envelopment analysis has been applied to the stock market. In 2000, a study was done using DEA to select efficient, large market-cap securities (Powers, McMullen). However the paper states that the goal was not to choose the best stocks to invest in, but to prove DEA as a viable technique to assist investors in multi-criteria problems. Thus, no back-testing or validation of their strategy was done. Also, all of the outputs in the study were historical return percentages over the previous 1, 3, 5, and 10 years. This approach does not take into account current financial ratios and assumes that historical returns indicate future performance. After checking the returns of the stocks they deemed efficient over the 12 months after the paper was published (beginning in fall of 2000), there was no increase in performance compared to the Dow Jones Average. Greg Gregoriou's 2006 paper uses DEA to optimize U.S. mutual funds, but he too draws conclusions without any back-testing.

Chen used data envelopment analysis in 2008 to select efficient stocks on the Taiwanese stock exchange. He concluded that DEA was an effective tool for stock selection when all stocks were in the same industry. Later, in 2010 Patari, Leivo, and Honkapuro proved similar results on the Finland stock exchange.

In spite of previous research, it is still necessary to further explore the use of data envelopment analysis in stock selection. Of the research that has been done, none fulfill all the requirements of a successful solution that were stated earlier.

III. Design

In the introduction section above, four requirements for a successful solution were stated. Those requirements will be revisited here in further detail along with the steps needed to accomplish them.

Low trading frequency

Quantitative strategies that attempt to time the market and buy and sell frequently would not be beneficial to the common investor who does not have the time or resources to accomplish this. Also, having a low trading frequency is important for the reduction of transaction fees, taxes and complexity. The goal is to reexamine the portfolio on an annual basis and make any necessary trades on this single pre-determined day.

Concentrated portfolio

A portfolio with a low number of stocks has the potential for a higher return with the tradeoff of increased variability. Another benefit is lower transaction fees, which have greater effect on smaller investment amounts.

No specified number of holdings

Setting restrictions on the minimum or maximum amount of stocks that a portfolio contains will restrict the performance during an economic upswing or downswing. For example, if an investor required a portfolio with 35 stocks they may have excessive losses in a recession when there may only be a handful of stocks worth investing in.

Reasonable formulation

Lastly, it is important for the solution to have a reasonable formulation that can be explained and only contain necessary components. A quantitative trading strategy that has no explanation for its derivation is likely to be a coincidence that worked when back-tested but has a low probability of predicting future performance. This is sometimes referred to

as data mining. Such a strategy was the once famous ‘Foolish Four’, which greatly outperformed the Dow Jones in a 25-year back-test but failed to produce results after that time (*Fool.com*).

The tasks that must be accomplished to solve this problem are as follows:

Determine inputs and outputs for the data envelopment analysis

Determining the inputs and outputs for the data envelopment analysis is a crucial step in order to achieve the desired results. The inputs and outputs can be thought of as the costs and benefits of owning a stock, respectively. As explained in the Background section, there are certain factors that must be taken into consideration when choosing inputs and outputs for any data envelopment analysis.

Gather historical information

Annual historical data on the chosen inputs and outputs for all stocks in the Dow for the past 10 years must be retrieved. The data will be supplied by YCharts.com and accessed via their Data Export feature. Any data not available will be looked up in a company’s 10-K filing with the Securities Exchange Commission. It is important to realize that the companies listed on the Dow index have changed many times over the past 10 years so attention must be paid to retrieve data on the correct companies given the respective point in time.

Perform data envelopment analysis

A data envelopment analysis will be performed annually for the past 10 years using the historical information that was gathered as inputs and outputs. This task is important because without any back-testing, there would be no way to validate the solution as effective. By back-testing 10 years, the solution would be tested over a number of economic upturns and downturns, which may give further insight to the effectiveness of the approach.

Formulate results

For each year, a portfolio will be created of the companies deemed efficient by the data envelopment analysis. The returns of this efficient portfolio will be compared to that of the Dow as a whole.

Check results for significance

The difference in the returns of the efficient portfolio and the Dow Jones Industrial Average will determine if the solution was able to achieve improved performance. Even a 1% increase in performance over the Dow Jones would be seen as significant to most investors. A hypothesis test will be needed to check if the results are statistically significant before any conclusions are made.

Draw conclusions

After using a hypothesis test to check the results, a conclusion can be made about the performance of the DEA-based efficient portfolio versus the Dow Jones Industrial Average. Economic justification of the results must be shown as well.

Make recommendations

Finally, based on the conclusions, recommendations will be made on how a DEA investment strategy may be applied to stocks outside of the Dow and also with other styles of investing. It will be important to address how both the common investor and the institutional fund could apply DEA to their portfolios.

IV. Methods

The first step of the design was to determine the inputs and outputs for data envelopment analysis. This was done by looking at the solution requirements as a whole as well as prior art on the topic to create a large list of inputs and outputs. There are over 1,500 metrics available through the Ycharts.com data export feature, the following table shows which were considered as variables for this project (listed alphabetically):

Inputs:		**Outputs:**
P/E		Dividend Yield
Beta		Earnings Yield
Book Value		EPS Growth
Long Term Debt		Free Cash Flow
Market Cap		Net Income
Operating Expense		Operating Income
P/B		Quick Ratio
PE 10		Return on Assets
PEG		Return on Equity
Price		Volume

Table 1. Original inputs / outputs for consideration

As mentioned earlier, increasing the number of inputs and outputs will increase the number of efficient units by DEA. Thus, a small list of inputs and outputs is needed to produce the desired concentrated portfolio. The list above was selectively narrowed down by looking for relationships between figures and discarding any that may have built in redundancy. For example, including both the Book Value and the Price / Book Value ratio would be redundant. Also, a visual inspection of the data was done to check for any obvious trends in any of the factors. From this it was clearly seen that some factors, such as volume, had no effect on the future returns for the stock and were therefore eliminated. Table 2 below shows the remaining inputs and outputs that were left for consideration.

Inputs:	Outputs:
P/E	Dividend Yield
Beta	Operating Income
P/B	Return on Assets
Price	Return on Equity

Table 2. Inputs / outputs for consideration after eliminating the redundant

From this list, a sensitivity analysis was be done to determine which factors had a greater impact on the outcome of the data envelopment analysis. By eliminating the factors which had less effect, the list of inputs and outputs was reduced further. The sensitivity analysis was done using the same DEA software that would later be used to produce a solution. The software will be described in greater detail in a later section on how the DEA was done.

The type of sensitivity analysis done was the "one-at-a-time" approach. In this method, a single variable was manually altered while all others remained constant, from doing this it can be seen how that single variable affects the results of the DEA. The goal was to remove any input or output variables that did not greatly change the list of efficient stocks. Reducing the number of inputs or outputs will reduce the number of stocks listed as efficient, if a factor could be removed without affecting the make-up of stocks on said list, then it was considered unnecessary. . An example of how this was done can be found in Appendix D. After analyzing each factor, it was determined that two inputs and one output would be needed. Table 3 below shows beta and price as inputs, and dividend yield as the output.

Inputs:	Outputs:
Beta	Dividend Yield
Price	

Table 3. Inputs / outputs selected for analysis

A description of each input and output follows below:

Beta

Beta is the measure of risk of a stock in relation to the market as a whole. A high beta stock will tend to move more than the market while a low beta stock will move less. Since one of the goals of this project was to create a low-risk portfolio, using beta as an input will be beneficial to achieving that goal. By inspecting the portfolios of many successful value investors, many were seen to have a low overall beta.

Price

Price is simply the current price for one share of stock. Price can be used as a valuation metric for stocks on the Dow. All companies on the Dow index are all large and well-established so a lower share price generally means more room for that price to go up.

Dividend Yield

Dividend Yield is the total amount of dividends a company has paid out annually expressed as a percentage of the share price. A high dividend is beneficial to value investors as it is generally a sign of a strong company and can help investors weather economic downturns. Dividends are another important aspect to help lower the risk of a portfolio.

After the inputs and outputs were determined, the next step was to gather the historical information on each metric. This was simple to do using the Ycharts.com data export feature. The list of stocks was entered along with the desired financial metric and dates, and then all of the data was exported to Excel spreadsheets. Since one of the requirements stated earlier was low trading frequency, portfolio selection would be done annually. The last trading day of the year was set as the recalculation date, although any date could have been used as long as it was consistent.

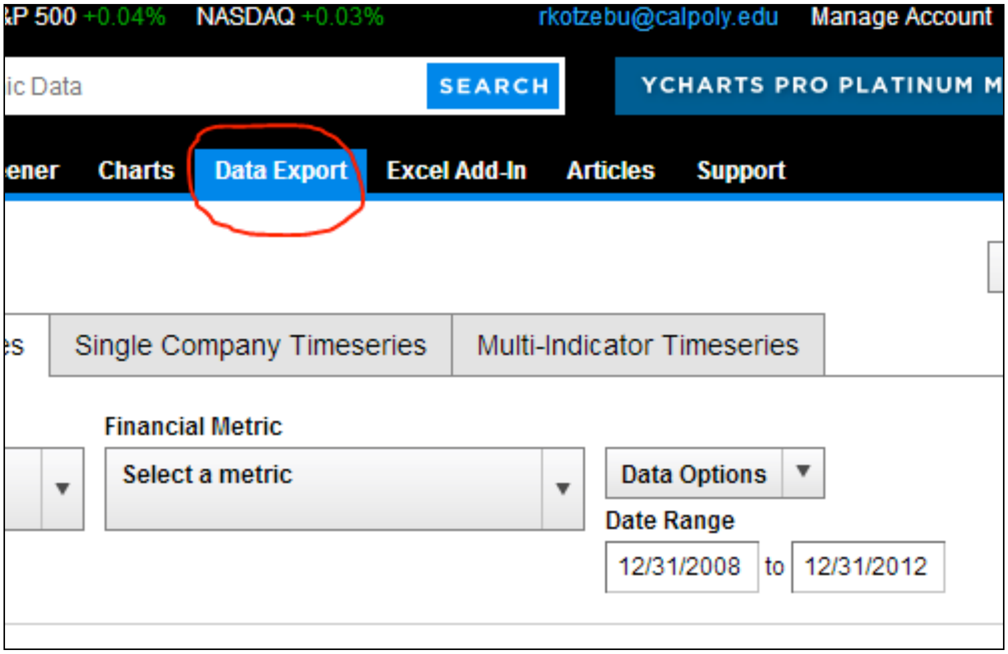


Figure 2. Screenshot highlighting the Ycharts.com data export feature

Figure 2 shows a screenshot highlighting the data export feature at Ycharts.com. The export date was set to 31-Dec for each year, but since that is a holiday, Ycharts.com automatically pulls data from the last day the stock market was open. Since the components of the Dow Jones Industrial Average have changed many times over the past decade, precautions were taken so the right companies were analyzed for the correct years so an accurate comparison could be made.

An unexpected constraint was encountered during data export. The data for stock 'beta' was only available for the previous eight years, while the design called for back-testing of ten years. Since beta is derived from a calculation and cannot be looked up in any company SEC filings, there was not much that could be done with the current resources. Thus, the project design had to shift from a back-test of ten years to eight years. An effect of reducing the sample size includes reduced confidence in the results, but since only two samples would be lost it was decided to continue on as planned.

Once the historical data was exported into Excel spreadsheets, DEA software was used to determine the efficient stocks for each year. The software used was MaxDEA version 5.2

and has a limited free version available at www.MaxDEA.cn. MaxDEA is a Microsoft Access based program that can calculate efficiencies for a large number of inputs and outputs very quickly. Although there are limitations in the free version, such as no ability to weight inputs, it was still considered fit for this project. With a price tag of $690, the full version could be used in future experiments for a more exhaustive analysis.

MaxDEA allows data to be imported from Excel spreadsheets. This was a perfect complement to exporting to Excel from Ycharts.com. As long as each input and output is in a unique column, they can easily be defined as such in MaxDEA. The basic DEA model that is available in the free version can be run with a single click, after which a results window opens and displays the efficiency score of each DMU, the results were sorted from largest to smallest to easily see which stocks had an efficiency score of one.

Appendix C illustrates an example of how MaxDEA was used to find efficient companies in 2012. This process was repeated for every year back to 2005. For each year, the efficient stocks were organized into a simulated portfolio that would be held until the next recalculation date (last trading day of the year in this case). The returns of the simulated portfolio, known as the DEA efficient portfolio, were compared to the returns of the Dow Jones Industrial Average for each year to determine if DEA provided a superior portfolio generation tool.

V. Results and Discussion

Summarized results of the DEA efficient portfolio in comparison to the Dow can be seen in Table 4 below. Returns for both the Dow index and the DEA efficient portfolio are shown from 2005 through 2012. Also shown is the number of stocks that made up the DEA efficient portfolio in each period. It should be noted that the returns for the Dow were calculated as an equally weighted portfolio of all the companies analyzed, while the real Dow Jones Industrial Average is a price weighted index. This was done to block for any differences or errors in the data, as well to provide a better benchmark for comparison.

Year	Dow*	DEA Efficient	# Efficient
2005	4.08%	15.51%	3
2006	19.68%	20.83%	3
2007	10.43%	25.05%	3
2008	-31.00%	-16.88%	3
2009	26.77%	16.46%	3
2010	11.67%	9.24%	7
2011	3.16%	-2.11%	4
2012	13.82%	28.49%	6
mean =	**7.33%**	**12.07%**	**4.0**

Table 4. Yearly returns of DEA efficient portfolio

It can be seen that the mean annual return was nearly 5% greater for the DEA efficient portfolio, with the average number of holdings being four, versus thirty on the Dow.

While these results were compelling, a hypothesis test was needed to see if these results were statistically significant. A paired t-test was used to test the null hypothesis (H_0) that the mean annual returns for the DEA efficient portfolio are less than or equal to those of the Dow, against the alternative hypothesis (H_1) that the mean annual returns for the DEA efficient portfolio are greater than the Dow.

H_0: $\mu_{efficient} \leq \mu_{dow}$

H_1: $\mu_{efficient} > \mu_{dow}$

Before any statistics were computed, the data was checked for normality. Figure 1 in Appendix A shows the probability plot concluding that the difference between the returns of the DEA efficient portfolio and the returns of the Dow are normal, with a p-value of 0.135. Since the data was normal, a paired t-test was used to test the hypothesis.

```
Paired T for DEA Efficient - Dow

                N    Mean   StDev  SE Mean
DEA Efficient   8  0.1207  0.1509   0.0533
Dow             8  0.0733  0.1731   0.0612
Difference      8  0.0475  0.1014   0.0359


95% lower bound for mean difference: -0.0205
T-Test of mean difference = 0 (vs > 0): T-Value = 1.32  P-Value = 0.114
```

The output from the paired t-test, done with MiniTab statistical software, is shown above. The p-value from this test was 0.114, which is not enough to reject the null hypothesis, H_0, with 95% confidence. However, since the p-value is still relatively small, the results should still be taken seriously and recommendations can be made, especially considering the small sample size of eight.

Another point to highlight from the MiniTab output is the standard deviation of the DEA efficient portfolio being less than the Dow. Although not statistically significant, this is still noteworthy considering the DEA efficient portfolio was comprised of an average of four stocks versus thirty.

The results from this project were in-line with the expectations stated in the project design. A quantitative trading strategy was developed that met all of the goals outlined earlier, while experiencing higher returns and lower risk. The theory that data envelopment analysis could create a superior portfolio was upheld with this experiment.

The DEA efficient portfolio was shown outperform the pool of stocks it was created from over an eight year period.

Original expectations were that DEA would create a portfolio with a higher standard deviation along with higher returns. This is simply because the lower number of stocks in the portfolio would lead to greater variance. The results however, showed the opposite. The standard deviation was in fact slightly lower on the DEA efficient portfolio. This is most likely due to the selection of inputs and outputs, namely 'beta' as an input. Because of this, the design of the experiment can be verified. However, this experiment was designed around the stocks on the Dow Jones Industrial Average and if another pool of stocks was used, some design changes may have to be made. Since back-testing was used as a means of validation, it may be the case that similar results will not be seen on future trials. Since the design was built around a reasonable formulation, it raises the probability of repeatability into the future. Based on the results, it would be possible to implement a DEA based portfolio, as shown here, with confidence.

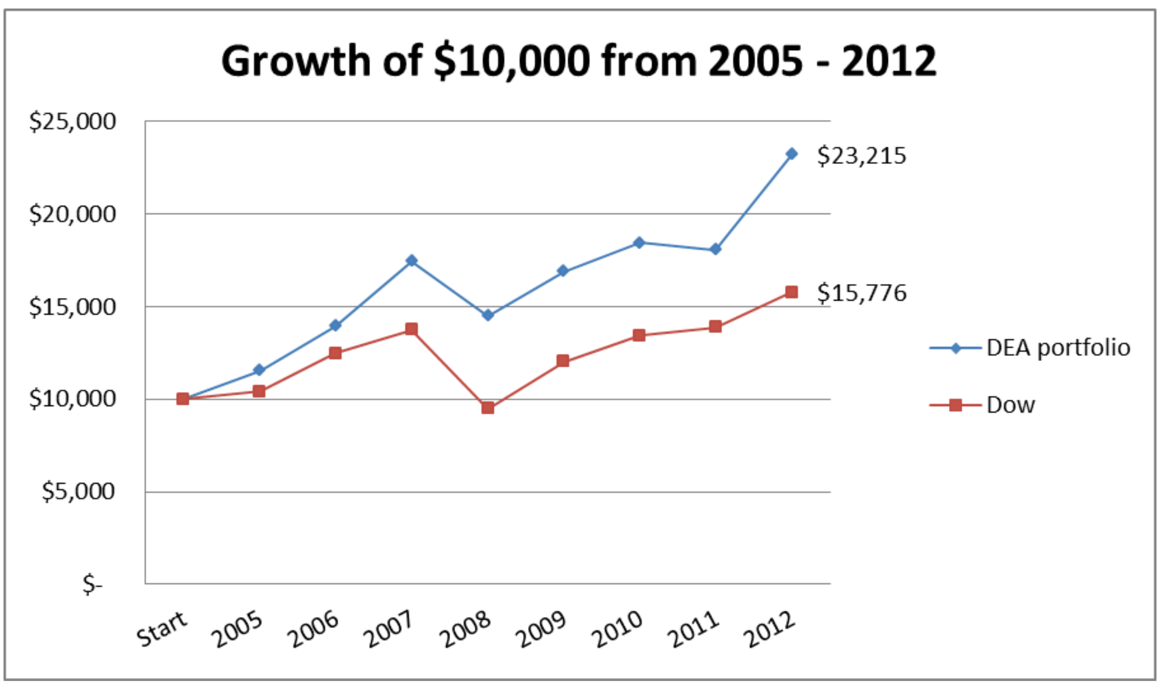


Figure 3. Growth of $10,000 from 2005 through 2012

Figure 3 on the previous page shows the theoretical growth of $10,000 invested in both the DEA portfolio and the Dow from the beginning of 2005 through 2012. Assuming no transaction fees, investing in the DEA portfolio would result $23,215 while the Dow would only net $15,776.

The initial pool of stocks that are used may limit the use of DEA as a portfolio generation tool. Since DEA is based around optimums, not averages, stocks with outlier financial ratios could skew the results. This was not a problem with the thirty companies on the Dow index, but if a much larger pool of stocks was used it could be beneficial to first remove any outliers.

VI. Conclusion

Two factors that drive investors to and away from the stock market are reward and risk, respectively. By using a stock selection strategy that is quantitative, investors may feel more comfortable and secure with their decisions. However, there lacks a quantitative strategy that can produce increased returns with lower risk by purchasing a small number of stocks. The objective of this project was to formulate a quantitative stock trading strategy that produced exceptional returns with low risk while fulfilling the following requirements:

- Low trading frequency
- Concentrated portfolio
- No specified number of holdings
- Reasonable formulation

By using a linear programming based operations research technique known as data envelopment analysis (DEA) a solution was generated to fulfill these goals. DEA was used to create a portfolio of stocks from the Dow Jones Industrial Average on an annual basis; over eight years the DEA efficient portfolio had nearly double the annualized returns of the Dow. This was done with a smaller standard deviation and less drawdown. Also, the DEA efficient portfolio was comprised of only four stocks versus thirty on the Dow. The ‘winning %’ is the percentage of stocks in the portfolio that achieve positive returns and was 78.13% and 62.74% for the DEA portfolio and Dow, respectively. The table on the following page summarizes the results from this experiment.

Metric	**DEA portfolio**	**Dow**
Annualized Returns	11.10%	5.86%
Winning %	78.13%	62.74%
Max Drawdown	16.88%	31.00%
Typical # of Holdings	4	30
Standard Deviation	15.09%	17.31%

Table 5. Highlights of DEA portfolio versus Dow

From Table 5, it can be seen that all of the goals of this project were satisfied. It is reasonable to conclude that data envelopment analysis is a suitable tool for generating a portfolio of stocks that is superior to the pool of stocks it was created from.

It is safe to recommend the use of DEA to the common investor by selecting stocks exactly as shown in this project. Every tool that was used is available to the public so it would be possible for anyone to recreate the results here. New solutions could be explored by altering one, or many, of the variables in the setup.

A unique opportunity exists for the institutional investor to implement data envelopment analysis to create efficient exchange-traded-funds (ETFs). The demand for ETFs has gone up greatly in the past decade due to the numerous advantages they have over traditional mutual-funds, including lower management fees and the ability to make intraday trades. According the Investment Company Institute, as of April 2013 the combined assets of all national ETFs was $1.406 trillion. ETFs can determine their portfolio of holdings from an index, region, industry, company size, growth-rate, or any other factor. DEA was used to determine the portfolio, it may be possible to attract much attention to the fund and provide a high return, low cost investment to the masses.

In future iterations of this project, different results could be obtained by selecting different factors as inputs and outputs for the analysis. The inputs and outputs were chosen based on a ‘value’ style of investing, and would need to be set up differently if the

goal was purely growth. Also, changing the initial pool of stocks could affect the outcome of similar experiments. Although, the Dow Jones Industrial Average was used here, there are countless options available such as by industry, country, mutual fund, or a custom screen.

V. References

A. Halim Boussabaine, Richard J. Kirkham, (2006) "Whole life cycle performance measurement re-engineering for the UK National Health Service estate", Facilities, Vol. 24 Iss: 9/10, pp.324 - 342

Anderson, K. and Brooks, C. (2007), "Extreme returns from extreme value stocks: enhancing the valuepremium", Journal of Investing, Vol. 16 No. 1, pp. 69-81.

Anderson, T. (n.d.). A DEA Homepage. *A Data Envelopment Analysis (DEA) Homepage*. Retrieved April 27, 2013, from www.emp.pdx.edu/dea/homedea.html

Berg, S. (2010). "Water Utility Benchmarking: Measurement, Methodology, and Performance Incentives." International Water Association.

Chan, L.K.C., Hamao, Y. and Lakonishok, J. (1993), "Can fundamentals predict Japanese stock returns?", Financial Analysts Journal, Vol. 49 No. 4, pp. 63-9.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2 (6), 429–444.

Chen, H. H. (2008). Stock selection using data envelopment analysis.*Industrial Management & Data Systems*, *108*(9), 1255-1268.

Fama, E.F. and Fench, K.R. (1992), “The cross-section of expected stock returns”, Journal of Finance, Vol. 47 No. 2, pp. 427-65.

Gardner, D., & Gardner, T. (2000, December 11). Farewell, Foolish Four .*Fool.com*. Retrieved April 27, 2013, from http://www.fool.com/ddow/2000/ddow0012

Gaunt, C. (2004), "Size and book to market effects and Fama French three factor asset pricing model: evidence from the Australian stock market", Accounting and Finance, Vol. 44 No. 1, pp. 27-44.

Gregoriou, G. N. (2006). Optimisation of the largest US mutual funds using data envelopment analysis. *Journal of Asset management*, *6*(6), 445-455.

Highest dividend paying stocks. (n.d.).*Dogs of the Dow*. Retrieved April 27, 2013, from http://www.dogsofthedow.com/

Jensen, M. C. (1968). The performance of mutual funds in the period 1944–1964. Journal of Finance, 23(2), 389–416.

Pätäri, E. J., Leivo, T. H., & Honkapuro, J. S. (2010). Enhancement of value portfolio performance using data envelopment analysis. *Studies in Economics and Finance*, *27*(3), 223-246.

Powers, J., & McMullen, P. (2000). Using data envelopment analysis to select efficient large market cap securities. *Journal of Business and Management Volume*, *7*(2).

Release: Exchange-Traded Fund Data, April 2013. (n.d.). *ICI* . Retrieved May 6, 2013, from http://www.ici.org/etf_resources/research/etfs_04_13

Rousseau, R., & Paul, v. R. (2004). Time and the payoff to value investing. *Journal of Asset Management, 4*(5), 318-325.

V. Appendices

A. Probability Plot of Data

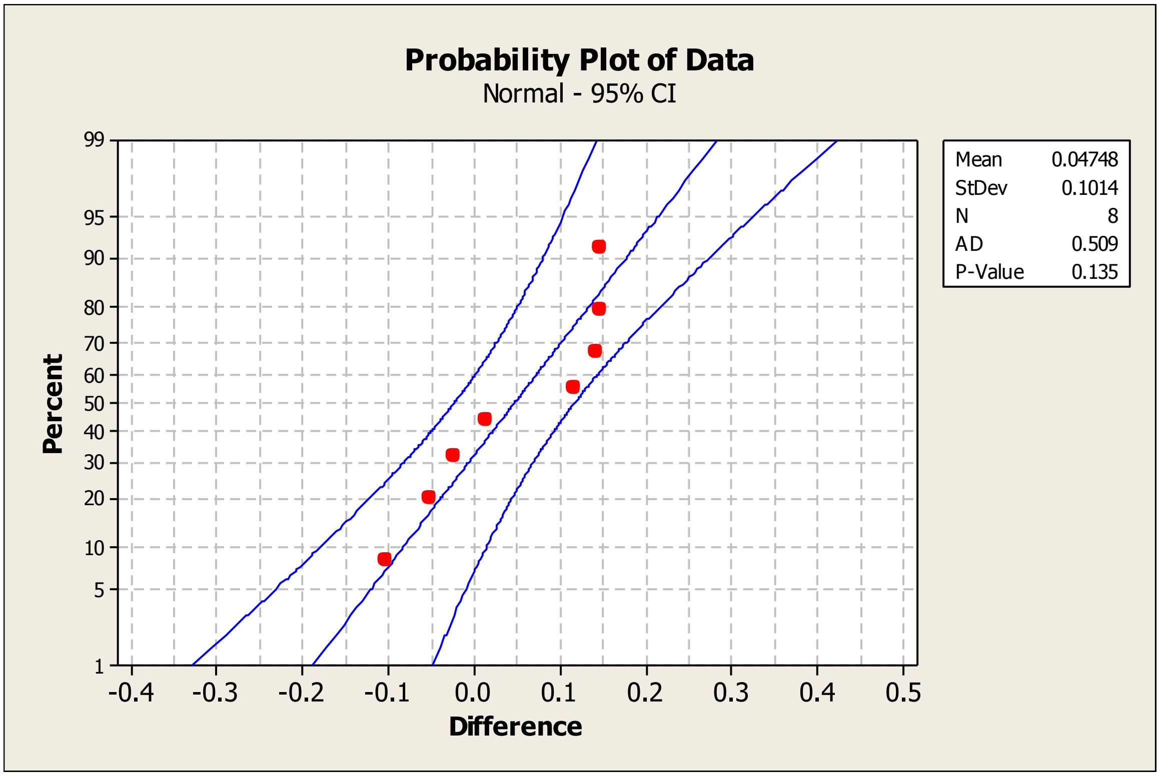


Figure 1. Probability plot of difference of DEA efficient returns and Dow returns

B. Annual Data

Red highlighting, indicates a DEA efficient stock

2012

Ticker	Buy Price	Sell Price	Next yr. Growth	Beta	Dividend
AA	8.65	8.68	0.35%	2.0567	1.39%
AXP	47.17	57.48	21.86%	1.8471	1.53%
BA	73.35	75.36	2.74%	1.2408	2.29%
BAC	5.56	11.61	108.81%	2.186	0.72%
CAT	90.6	89.6085	-1.09%	1.858	1.99%
CSCO	18.08	19.6494	8.68%	1.1696	0.00%
CVX	106.4	108.14	1.64%	0.7895	2.90%
DD	45.78	44.9785	-1.75%	1.4436	3.58%
DIS	37.5	49.79	32.77%	1.2058	1.60%
GE	17.91	20.99	17.20%	1.5825	2.62%
HD	42.04	61.85	47.12%	0.7778	2.47%
HPQ	25.76	14.25	-44.68%	1.1291	1.71%
IBM	183.88	191.55	4.17%	0.6626	1.58%
INTC	24.25	20.62	-14.97%	1.081	3.23%
JNJ	65.58	70.1	6.89%	0.5471	3.43%
JPM	33.25	43.9691	32.24%	1.2224	2.41%
KO	34.985	36.25	3.62%	0.5477	2.69%
MCD	100.33	88.21	-12.08%	0.4402	2.52%
MMM	81.73	92.85	13.61%	0.8607	2.69%
MRK	37.7	40.94	8.59%	0.6765	4.14%
MSFT	25.96	26.7097	2.89%	0.9768	2.62%
PFE	21.64	25.0793	15.89%	0.7172	3.70%
PG	66.71	67.89	1.77%	0.4586	3.08%
T	30.24	33.71	11.47%	0.619	5.69%
TRV	59.17	71.82	21.38%	0.7114	2.69%
UTX	73.09	82.01	12.20%	1.0278	2.55%
VZ	40.12	43.27	7.85%	0.5958	4.89%
WMT	59.76	68.23	14.17%	0.3577	2.44%
XOM	84.76	86.55	2.11%	0.5011	2.18%

	Return	Dividend*	Total
Mean:	11.22%	2.60%	**13.82%**
Efficient:	25.72%	2.77%	**28.49%**

**Since future dividends are not available, current dividend yield was used as an estimate*

2011

Ticker	Buy Price	Sell Price	Next yr. Growth	Beta	Dividend
AA	15.39	8.65	-43.79%	2.0838	0.78%
AXP	42.92	47.17	9.90%	1.9392	1.68%
BA	65.26	73.35	12.40%	1.2433	2.57%
BAC	13.34	5.56	-58.32%	2.196	0.30%
CAT	93.66	90.6	-3.27%	1.7493	1.84%
CSCO	20.23	18.08	-10.63%	1.2534	0.00%
CVX	91.25	106.4	16.60%	0.7363	3.11%
DD	49.88	45.78	-8.22%	1.3764	3.29%
DIS	37.51	37.5	-0.03%	1.1069	1.07%
GE	18.29	17.91	-2.08%	1.615	1.97%
HD	35.06	42.04	19.91%	0.7955	2.70%
HPQ	42.1	25.76	-38.81%	1.024	0.76%
IBM	146.76	183.88	25.29%	0.7166	1.70%
INTC	21.03	24.25	15.31%	1.0908	3.00%
JNJ	61.85	65.58	6.03%	0.5717	3.41%
JPM	42.42	33.25	-21.62%	1.1397	0.47%
KO	32.885	34.985	6.39%	0.6052	2.68%
MCD	76.76	100.33	30.71%	0.4998	2.94%
MMM	86.3	81.73	-5.30%	0.8074	2.43%
MRK	36.04	37.7	4.61%	0.732	4.22%
MSFT	27.91	25.96	-6.99%	1.0706	1.97%
PFE	17.51	21.64	23.59%	0.6982	4.11%
PG	64.33	66.71	3.70%	0.524	2.93%
T	29.38	30.24	2.93%	0.6717	5.72%
TRV	55.71	59.17	6.21%	0.6006	2.53%
UTX	78.72	73.09	-7.15%	1.0062	2.16%
VZ	35.78	40.12	12.13%	0.6859	5.35%
WMT	53.93	59.76	10.81%	0.3121	2.24%
XOM	73.12	84.76	15.92%	0.4724	2.38%

	Return	Dividend	Total
Mean:	0.56%	2.60%	**3.16%**
Efficient:	-5.25%	3.14%	**-2.11%**

2010

Ticker	Buy Price	Sell Price	Next yr. Growth	Beta	Dividend
AA	16.12	15.39	-4.53%	2.0904	1.61%
AXP	40.52	42.92	5.92%	2.1477	1.78%
BA	54.13	65.26	20.56%	1.2743	3.10%
BAC	15.06	13.34	-11.42%	2.4247	0.27%
CAT	56.99	93.66	64.34%	1.8188	2.95%
CSCO	23.94	20.23	-15.50%	1.2109	0.00%
CVX	76.99	91.25	18.52%	0.6488	3.46%
DD	33.67	49.88	48.14%	1.3994	4.87%
DIS	32.25	37.51	16.31%	1.1048	1.09%
GE	15.13	18.29	20.89%	1.5483	1.98%
HD	28.93	35.06	21.19%	0.7024	3.11%
HPQ	51.51	42.1	-18.27%	1.0283	0.62%
IBM	130.9	146.76	12.12%	0.8007	1.64%
INTC	20.4	21.03	3.09%	1.1715	2.75%
JNJ	64.41	61.85	-3.97%	0.5537	3.00%
JPM	41.67	42.42	1.80%	1.0976	1.27%
KO	28.5	32.885	15.39%	0.5993	2.88%
MCD	62.44	76.76	22.93%	0.6341	3.28%
MMM	82.67	86.3	4.39%	0.7833	2.47%
MRK	36.54	36.04	-1.37%	0.9168	4.16%
MSFT	30.48	27.91	-8.43%	0.9605	1.71%
PFE	18.19	17.51	-3.74%	0.7659	4.40%
PG	60.63	64.33	6.10%	0.5899	2.84%
T	28.03	29.38	4.82%	0.6793	5.85%
TRV	49.86	55.71	11.73%	0.6793	2.47%
UTX	69.41	78.72	13.41%	0.9684	2.22%
VZ	30.9745	35.78	15.51%	0.6126	5.99%
WMT	53.45	53.93	0.90%	0.2504	2.04%
XOM	68.19	73.12	7.23%	0.4437	2.43%

	Return	Dividend	Total
Mean:	9.24%	2.42%	**11.67%**
Efficient:	6.05%	3.19%	**9.24%**

2009

Ticker	Buy Price	Sell Price	Next yr. Growth	Beta	Dividend
AA	11.26	16.12	43.16%	1.5586	6.04%
AXP	18.55	40.52	118.44%	1.3233	3.88%
BA	42.67	54.13	26.86%	1.2046	3.75%
BAC	14.08	15.06	6.96%	0.3552	15.91%
CAT	44.67	56.99	27.58%	1.2128	3.49%
CVX	73.97	76.99	4.08%	1.1757	3.42%
DD	25.3	33.67	33.08%	0.8742	6.48%
DIS	22.69	32.25	42.13%	0.9399	1.54%
GE	16.2	15.13	-6.60%	0.4162	5.74%
HD	23.02	28.93	25.67%	0.8961	3.91%
HPQ	36.29	51.51	41.94%	1.445	0.88%
IBM	84.16	130.9	55.54%	0.92	2.26%
INTC	14.66	20.4	39.15%	1.8639	3.73%
JNJ	59.83	64.41	7.66%	0.2947	3.00%
JPM	31.53	41.67	32.16%	0.9022	4.82%
KO	22.635	28.5	25.91%	0.7806	3.36%
MCD	62.19	62.44	0.40%	1.2085	2.61%
MMM	57.54	82.67	43.67%	0.8222	3.48%
MRK	30.4	36.54	20.20%	0.8753	5.00%
MSFT	19.44	30.48	56.79%	0.9742	2.37%
PFE	17.71	18.19	2.71%	0.6359	7.23%
PG	61.82	60.63	-1.92%	0.5492	2.51%
T	28.5	28.03	-1.65%	0.8806	5.61%
UTX	53.6	69.41	29.50%	1.0305	2.51%
VZ	31.6944	30.9745	-2.27%	0.9672	5.52%
WMT	56.06	53.45	-4.66%	0.0224	1.69%
XOM	79.83	68.19	-14.58%	1.0327	1.94%

	Return	Dividend	Total
Mean:	24.14%	2.63%	**26.77%**
Efficient:	15.16%	1.31%	**16.46%**

2008

Ticker	Buy Price	Sell Price	Next yr. Growth	Beta	Dividend
AA	36.55	11.26	-69.19%	1.9276	1.86%
AIG	976.7703	26.3041	-97.31%	1.2001	1.25%
AXP	52.02	18.55	-64.34%	1.3356	1.15%
BA	87.46	42.67	-51.21%	0.7423	1.60%
C	294.4	67.1	-77.21%	1.2832	7.34%
CAT	72.56	44.67	-38.44%	1.4234	1.82%
DD	44.09	25.3	-42.62%	1.0133	3.45%
DIS	32.28	22.69	-29.71%	1.1972	1.08%
GE	37.07	16.2	-56.30%	0.7858	2.35%
HD	26.94	23.02	-14.55%	1.4993	3.34%
HON	61.57	32.83	-46.68%	1.6193	1.62%
HPQ	50.48	36.29	-28.11%	1.8902	0.63%
IBM	108.1	84.16	-22.15%	1.5658	1.39%
INTC	26.66	14.66	-45.01%	2.1175	1.69%
JNJ	66.7	59.83	-10.30%	0.3716	2.43%
JPM	43.65	31.53	-27.77%	1.7513	3.30%
KO	30.685	22.635	-26.23%	0.6326	2.22%
MCD	58.91	62.19	5.57%	1.5029	2.55%
MMM	84.32	57.54	-31.76%	0.7534	2.28%
MO	23.3098	15.06	-35.39%	0.9098	13.08%
MRK	58.11	30.4	-47.69%	0.8754	2.62%
MSFT	35.6	19.44	-45.39%	0.9906	1.15%
PFE	22.73	17.71	-22.09%	0.6588	5.10%
PG	73.42	61.82	-15.80%	0.1904	1.85%
T	41.56	28.5	-31.42%	1.4871	3.42%
UTX	76.54	53.6	-29.97%	0.6781	1.53%
VZ	40.6576	31.6944	-22.05%	1.4192	4.05%
WMT	47.53	56.06	17.95%	0.4594	1.85%
XOM	93.69	79.83	-14.79%	0.7441	1.46%

	Return	Dividend	Total
Mean:	-35.17%	4.17%	**-31.00%**
Efficient:	-24.43%	7.55%	**-16.88%**

2007

Ticker	Buy Price	Sell Price	Next yr. Growth	Beta	Dividend
AA	30.01	36.55	21.79%	1.8759	2.00%
AIG	1205.6495	976.7703	-18.98%	1.144	0.88%
AXP	60.67	52.02	-14.26%	1.2771	0.89%
BA	88.84	87.46	-1.55%	0.6844	1.35%
C	557	294.4	-47.15%	1.3473	3.52%
CAT	61.33	72.56	18.31%	1.3519	1.79%
DD	48.71	44.09	-9.48%	0.9923	3.04%
DIS	33.6214	32.28	-3.99%	1.0782	0.90%
GE	37.21	37.07	-0.38%	0.8221	2.10%
HD	40.16	26.94	-32.92%	1.4249	1.68%
HON	45.24	61.57	36.10%	1.4486	2.01%
HPQ	41.19	50.48	22.55%	1.776	0.78%
IBM	97.15	108.1	11.27%	1.664	1.13%
INTC	20.25	26.66	31.65%	2.0278	1.98%
JNJ	66.02	66.7	1.03%	0.3269	2.20%
JPM	48.3	43.65	-9.63%	1.7698	2.82%
KO	24.125	30.685	27.19%	0.5179	2.57%
MCD	44.33	58.91	32.89%	1.3961	2.26%
MMM	77.93	84.32	8.20%	0.5851	2.36%
MO	19.9981	23.3098	16.56%	0.7852	16.60%
MRK	43.6	58.11	33.28%	0.7538	3.49%
MSFT	29.86	35.6	19.22%	1.0406	1.24%
PFE	26.168	22.73	-13.14%	0.6071	3.67%
PG	64.27	73.42	14.24%	0.1681	1.88%
T	35.75	41.56	16.25%	1.5454	3.72%
UTX	62.52	76.54	22.42%	0.6337	1.62%
VZ	34.6552	40.6576	17.32%	1.4268	4.67%
WMT	46.18	47.53	2.92%	0.4822	1.45%
XOM	77.24	93.69	21.30%	0.7565	1.66%

	Return	Dividend	Total
Mean:	7.69%	2.74%	**10.43%**
Efficient:	19.33%	5.72%	**25.05%**

2006

Ticker	Buy Price	Sell Price	Next yr. Growth	Beta	Dividend
AA	29.57	30.01	1.49%	1.8931	2.03%
AIG	1143.1396	1205.6495	5.47%	1.0089	0.81%
AXP	51.46	60.67	17.90%	1.4044	0.93%
BA	70.24	88.84	26.48%	1.0344	1.42%
C	485.3	557	14.77%	1.3615	3.63%
CAT	57.77	61.33	6.16%	1.3587	1.58%
DD	42.5	48.71	14.61%	1.0134	3.44%
DIS	23.5164	33.6214	42.97%	1.3044	1.13%
GE	35.05	37.21	6.16%	0.8801	1.97%
HD	40.48	40.16	-0.79%	1.5519	0.99%
HON	37.25	45.24	21.45%	1.5588	2.22%
HPQ	28.63	41.19	43.87%	2.0082	1.12%
IBM	82.2	97.15	18.19%	1.5501	0.95%
INTC	24.96	20.25	-18.87%	2.2805	1.28%
JNJ	60.1	66.02	9.85%	0.261	2.12%
JPM	39.69	48.3	21.69%	1.7067	3.43%
KO	20.155	24.125	19.70%	0.3864	2.78%
MCD	33.72	44.33	31.47%	1.1992	1.99%
MMM	77.5	77.93	0.55%	0.6709	2.17%
MO	17.4512	19.9981	14.59%	0.6344	17.53%
MRK	31.81	43.6	37.06%	0.6071	4.78%
MSFT	26.15	29.86	14.19%	1.1267	1.22%
PFE	23.32	26.168	12.21%	0.4938	3.26%
PG	57.88	64.27	11.04%	0.1148	1.88%
T	24.49	35.75	45.98%	1.1163	5.27%
UTX	55.91	62.52	11.82%	1.0137	1.57%
VZ	27.0088	34.6552	28.31%	1.0676	5.92%
WMT	46.8	46.18	-1.32%	0.5366	1.28%
XOM	56.17	77.24	37.51%	0.6012	2.03%

	Return	Dividend	Total
Mean:	17.05%	2.63%	**19.68%**
Efficient:	15.11%	5.72%	**20.83%**

2005

Ticker	Buy Price	Sell Price	Next yr. Growth	Beta	Dividend
AA	31.42	29.57	-5.89%	1.8944	1.91%
AIG	1100.2488	1143.1396	3.90%	0.9623	0.43%
AXP	49.7723	51.46	3.39%	1.435	0.84%
BA	51.77	70.24	35.68%	1.061	1.49%
C	481.8	485.3	0.73%	1.4549	3.32%
CAT	49	57.77	17.90%	1.2527	1.59%
DD	49.05	42.5	-13.35%	1.0102	2.85%
DIS	27.2739	23.5164	-13.78%	1.3629	0.86%
GE	36.5	35.05	-3.97%	0.9216	1.70%
HD	42.74	40.48	-5.29%	1.5379	0.76%
HON	35.41	37.25	5.20%	1.5548	2.12%
HPQ	20.97	28.63	36.53%	2.0383	1.53%
IBM	98.58	82.2	-16.62%	1.508	0.71%
INTC	23.39	24.96	6.71%	2.2589	0.68%
JNJ	63.42	60.1	-5.23%	0.305	1.73%
JPM	39.01	39.69	1.74%	1.8086	3.49%
KO	20.82	20.155	-3.19%	0.3765	2.40%
MCD	32.06	33.72	5.18%	1.1336	1.72%
MMM	82.07	77.5	-5.57%	0.6654	1.75%
MO	14.251	17.4512	22.46%	0.6841	19.79%
MRK	32.14	31.81	-1.03%	0.5903	4.32%
MSFT*	26.72	26.15	-2.13%	1.1675	11.83%
PFE	26.89	23.32	-13.28%	0.4557	2.34%
PG	55.08	57.88	5.08%	0.0786	1.77%
T*	25.77	24.49	-4.97%	1.14	4.85%
UTX	51.98	55.91	7.56%	1.0476	1.35%
VZ	36.3255	27.0088	-25.65%	1.0852	4.24%
WMT	52.82	46.8	-11.40%	0.5371	0.98%
XOM	51.26	56.17	9.58%	0.5194	2.07%

**changed name from SBC Communications*

	Return	Dividend	Total
Mean:	1.29%	2.78%	**4.08%**
Efficient:	8.12%	7.40%	**15.51%**

C. *Example to Determine Efficient Companies in 2012*

The first step to determine which companies on the Dow are efficient was to import the input and output data into MaxDEA, as shown below.

Ticker	Buy Price	Beta	Dividend
AA	8.65	2.0567	1.39%
AXP	47.17	1.8471	1.53%
BA	73.35	1.2408	2.29%
BAC	5.56	2.186	0.72%
CAT	90.6	1.858	1.99%
CSCO	18.08	1.1696	0.00%
CVX	106.4	0.7895	2.90%
DD	45.78	1.4436	3.58%
DIS	37.5	1.2058	1.60%
GE	17.91	1.5825	2.62%
HD	42.04	0.7778	2.47%
HPQ	25.76	1.1291	1.71%
IBM	183.88	0.6626	1.58%
INTC	24.25	1.081	3.23%
JNJ	65.58	0.5471	3.43%
JPM	33.25	1.2224	2.41%
KO	34.985	0.5477	2.69%
MCD	100.33	0.4402	2.52%
MMM	81.73	0.8607	2.69%
MRK	37.7	0.6765	4.14%
MSFT	25.96	0.9768	2.62%
PFE	21.64	0.7172	3.70%
PG	66.71	0.4586	3.08%
T	30.24	0.619	5.69%
TRV	59.17	0.7114	2.69%
UTX	73.09	1.0278	2.55%
VZ	40.12	0.5958	4.89%
WMT	59.76	0.3577	2.44%
XOM	84.76	0.5011	2.18%

Figure 1. Input / Output data imported to MaxDEA

After the data was imported, the inputs and outputs were defined.

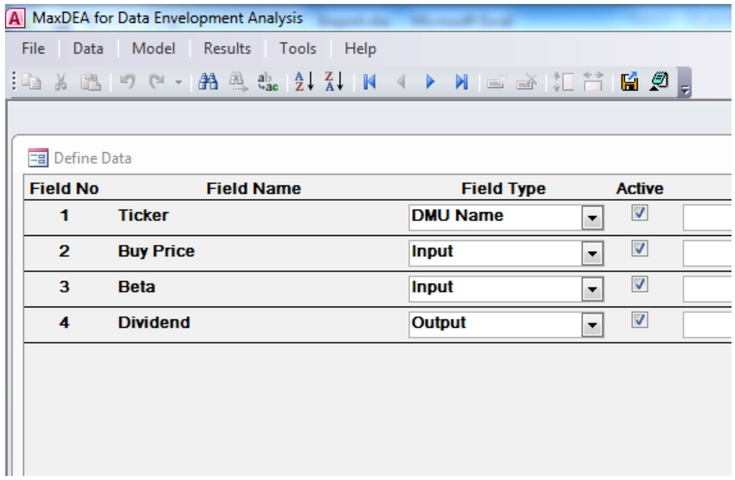
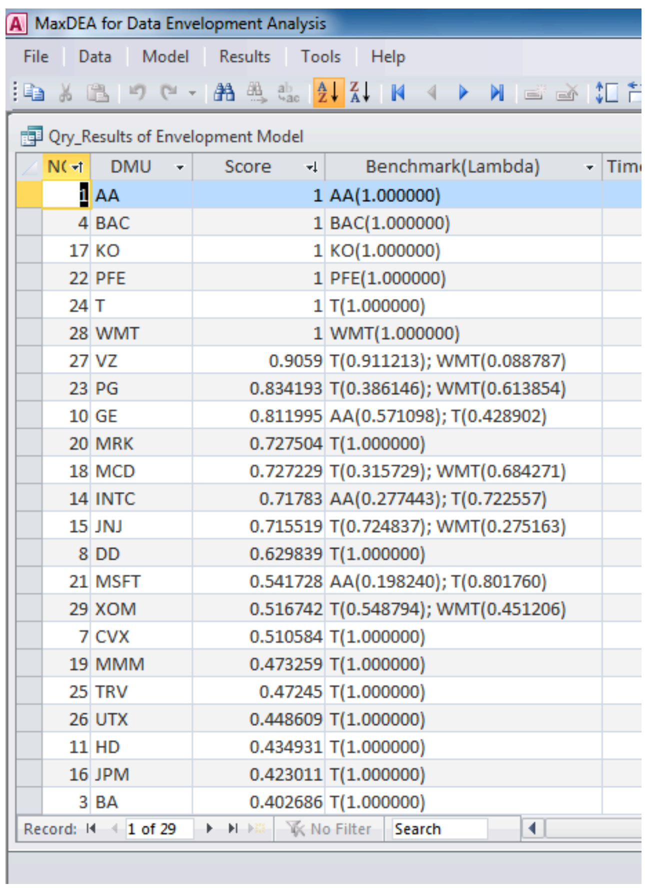


Figure 2. Defining inputs / outputs in MaxDEA

Once the inputs and outputs were defined, the DEA could be run. Results can be sorted by efficiency score, from largest to smallest, to easily see the efficient companies.



NO	DMU	Score	Benchmark(Lambda)
1	AA	1	AA(1.000000)
4	BAC	1	BAC(1.000000)
17	KO	1	KO(1.000000)
22	PFE	1	PFE(1.000000)
24	T	1	T(1.000000)
28	WMT	1	WMT(1.000000)
27	VZ	0.9059	T(0.911213); WMT(0.088787)
23	PG	0.834193	T(0.386146); WMT(0.613854)
10	GE	0.811995	AA(0.571098); T(0.428902)
20	MRK	0.727504	T(1.000000)
18	MCD	0.727229	T(0.315729); WMT(0.684271)
14	INTC	0.71783	AA(0.277443); T(0.722557)
15	JNJ	0.715519	T(0.724837); WMT(0.275163)
8	DD	0.629839	T(1.000000)
21	MSFT	0.541728	AA(0.198240); T(0.801760)
29	XOM	0.516742	T(0.548794); WMT(0.451206)
7	CVX	0.510584	T(1.000000)
19	MMM	0.473259	T(1.000000)
25	TRV	0.47245	T(1.000000)
26	UTX	0.448609	T(1.000000)
11	HD	0.434931	T(1.000000)
16	JPM	0.423011	T(1.000000)
3	BA	0.402686	T(1.000000)

Figure 3. Results from MaxDEA

It can be seen in Figure 3 on the previous page that six companies had efficiency scores of one for this year, 2012. Therefore these six companies would represent the DEA efficient portfolio for that year. This process was repeated for each year back to 2005, and a portfolio was formed for each year with the respective number of efficient companies.

D. Example of Sensitivity Analysis

The following is an example of how a manual one-at-a-time sensitivity analysis was used to determine the appropriate inputs and outputs. In this example the sensitivity of free cash flow will be examined.

Using price and beta as inputs, and dividend yield as an output the following companies on the Dow are efficient by DEA.

Ticker	Efficiency
AA	1
BAC	1
KO	1
PFE	1
T	1
WMT	1

Table 1. Efficiencies with price and beta as inputs, dividend yield as output

To determine how free cash flow affects which companies are efficient, another analysis was run with the same inputs and outputs with the addition of free cash flow as an output. The following table shows the results from that analysis.

Ticker	Efficiency
AA	1
BAC	1
JPM	1
KO	1
PFE	1
T	1
WMT	1

Table 2. Efficiencies with price and beta as inputs, dividend yield and free cash flow as outputs

It can be seen from Table 2 on the previous page that the addition of free cash flow as an output resulted in the same list of six efficient companies with the addition of JPM, for a total of seven. This increase in efficient companies is expected with an increase in inputs or outputs. However, since the addition of free cash flow as an output did not affect the

original six companies that were efficient, it can be assumed that the efficiency scores are not very sensitive to free cash flow. Because of this, free cash flow is not necessary as an output.

This process showed how a manual sensitivity analysis was done to determine which factors to use as inputs and outputs. The goal was to keep inputs and outputs to a minimum so any factor that showed little sensitivity of results, such as free cash flow, was discarded.