The Dow: Risk and the Return Paradox

Anthony Peterson
Coe College

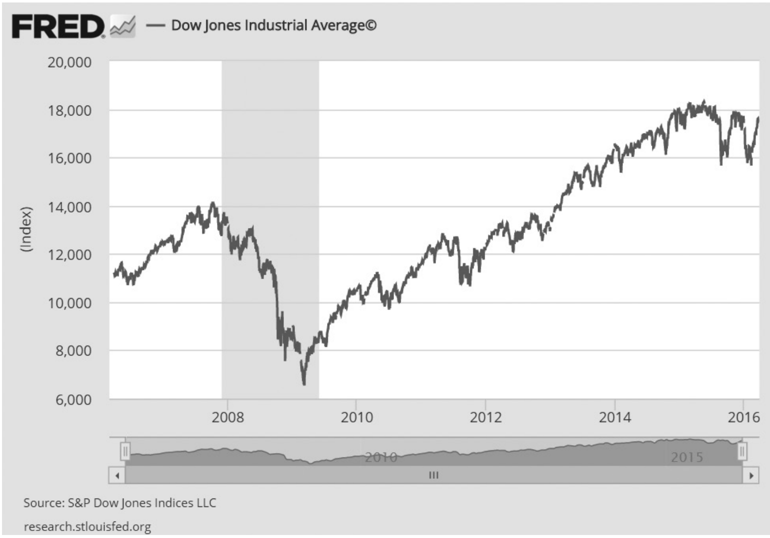
Abstract

Market indexes are important tools used to make generalizations and measure trends about the economy and certain sectors within. Among the most prominent of these indexes is the Dow Jones Industrial Average, the oldest market index still used today. Studies have shown that stocks removed from the Dow initially experience decreased returns upon their removal, but, in the long run, outperform the stocks that were added to the index. When a stock is removed from the Dow, index funds dump the shares. I hypothesize this adds investment risks to hold on to these shares, which should generate the observed higher returns in the long run. To test this hypothesis, modern stock price data (1988-2013) from the Center for Research in Security Prices was utilized to calculate returns for added and removed stock. In addition, the risk variables of both parties of stock are calculated through the framework of Modern Portfolio Theory (standard deviation) and the Capital Asset Pricing Model (beta). The study yields evidence that supports the hypothesis that risk can explain the short-term underperformance and long-term outperformance of removed stock.

Introduction

Market indexes are important tools used by investors and economists to measure trends and make generalizations about the economy and certain sectors within. Among the most prominent of these indexes is the Dow Jones Industrial Average (the Dow), which is the oldest market index used by investors today. It is a price-weighted market index composed of blue-chip stock. Specifically, these stocks are from corporations with quality reputations that are able to reliably perform profitably in both economic upturns and downturns, as well as gauge the performance of the U.S. equities markets (Dow Jones Indexes, 2011).

Figure 1.



The Dow Jones Industrial Average Index from March 28, 2006, to March 28, 2016. During the past ten years, the Dow has increased by 6,380.85 points, during which time the growth was stunted by the Great Recession, which is highlighted by the dark area at the bottom.

The index was created by Charles Dow, a co-founder of the *Wall Street Journal*, on May 26th, 1896, as a way to measure the performance of industrial stocks, which at the time held a small, but growing, segment of the market. The industrial stocks used by his average were meant to cover every industry except transportation and utilities (Dow Jones Indexes, 2011). Coupled with his first market index, the Dow Jones Railroad Average (the predecessor of the Dow Jones Transportation Average), Dow was able to track broad market trends since the work of these two segments was tied closely together; the railways would transport the goods that industrial companies made (Dow Jones Indexes, 2011).

The Dow was initially composed of twelve stocks, until October 4th, 1916, when the index expanded to twenty stocks (Dow Jones Indexes, 2011). During the time that the index was composed of either twelve or twenty stocks, the average was calculated by adding up the per-share prices of the constituent stocks and dividing by the number of constituents, essentially nothing more than a simple average. However, when the index expanded from twenty stocks to thirty on October 1st, 1928, a new way of calculating the average was needed. An adjusted divisor was created to help the Dow remain steady during events that impacted the constituents, such as a company having a 2-to-1 stock split, an event wherein a corporation doubles the outstanding shares of stock while halving the price (Dow Jones Indexes, 2011).

In a price-weighted index, each stock holds a weight proportional to its price. By setting up an index using this method, the gains of a high-priced stock can compensate forv the losses of multiple, smaller-priced stock, and vice versa (Nationwide Financial, 2013). By organizing a market index in this manner, every company has the opportunity to grow to a position of higher weight simply by a strong performance of its stock. Larger companies do not overshadow the smaller companies by the volume of stock they offer, but high-priced stock can create a top heavy index at times (Paglia, 2001).

The Dow tries for continuity among the indexes (S&P Dow Jones Indexes, 2013). In order to remain a relevant way to track the market, the index must make changes to its composition on an as-needed basis, adding stocks that are widely held by investors, showing a long history of growth, and removing those that no longer satisfy that criterion (Dow Jones Indexes, 2011). If members of the Dow Jones Averages Index Committee believe a stock is no longer meeting the criterion required to remain in the Dow at one of their privately held semi-annual meetings, the entire index is subsequently reviewed. This process can result, on occasion, in multiple changes to the stock composition of the index being instituted simultaneously.

Studies have shown that stocks removed from the Dow initially experience decreased returns upon their removal (Beneish & Gardner, 1995) but, in the long run, outperform the stocks that were added to the index (Arora, Capp, & Smith, 2005). It seems like a paradox, but I suspect that if we account for investment risk, the return behavior is not puzzling. When a stock is removed from the Dow, index funds dump the shares. I hypothesize this adds investment risks to hold on to these shares, which should generate the observed higher returns as compensation in the long run. Thus, the short run underperformance and long run outperformance can be explained.

Literature Review

In their study *Information Costs and Liquidity Effects from Changes in the Dow Jones Industrial Average List* (1995), Messod Beneish and John Gardner examine the abnormal returns of stocks from 1929 through 1987 surrounding the announcement of changes to the composition of the Dow, beginning 60 days prior to the announcement and concluding 60 days following the announcement. The average of all the abnormal returns were calculated and then cumulated to show certain trends within this 121 day window. This study was also conducted with portfolios of stock that were added or removed on the same day to account for strong correlations that may occur between these stocks. The authors found that stocks that were brought into the Dow experienced no significant change in the returns as a result of their inclusion to the index. It was found that the stocks that were removed from the index experienced significantly negative abnormal returns on the day of the announcement as a result of their exclusion. The authors concluded that these results can be explained by information costs and liquidity effects. Stocks that are included in the Dow are already well-known and actively traded, so no real change occurs as a result of their inclusion. The authors suggest that removed stocks are not traded as often and are likely not followed as closely as the included stock, making it more costly for investors to gather information on these removed stocks.

While the results of Beneish and Gardner show stock price returns experience significant abnormal movement following the announcement of inclusion into an index – specifically, the Dow – studies have looked into that same phenomenon in market indexes such as the S&P 500, which is a broader-based index that is used more widely than the Dow. It has been shown in *An Anatomy of the "S&P Game": The Effects of Changing the Rules* that after an announcement is made for changes in the S&P, trading volume increases 3.5 times normal on the day immediately following the announcement (Beneish & Whaley, 1996). Between the announcement and the day the change is implemented, trading volume increases a total of 7.2 times normal as people buy the stocks that are going to be added to the index in an attempt to profit from the price increase that will follow when the index funds rebalance adjusting to the change. The day that the change is implemented to the S&P, trading volume increases 10.6 times normal, largely as a result of index funds rebalancing. It also appears that the stocks added to the S&P continue to be traded at higher levels than before (Beneish & Whaley, 1996). This shows that stocks that are added to market indexes are in higher demand than those that are removed.

The idea that stocks added to a market index are in higher demand than their removed counterparts is further accentuated by the idea that these stocks have downward sloping demand curves. This theory is the conclusion reached in Andrei Shleifer's 1985 study *Do Demand Curves for Stocks Slope Downward?* The author states that when index funds need to rebalance after a change is made to a market index, the demand for the newly added stock increase, shifting the stock's demand curve outward. The author mentions in the paper's introduction that traditionally the demand curves for stock were thought of as being horizontal, since "several important propositions in finance rely on the ability of investors to buy and sell any amount of the firm's equity without significantly affecting the price" (Shleifer, 1985). However, the author goes on to mention that the number of index funds tracking the S&P 500 has increased "dramatically" over time (Shleifer, 1985). At the time of the study, the index funds could purchase up to 3% of added firms' equity, or their outstanding shares of stock. These large purchases lead to significantly increased abnormal returns for the added stock, meaning an index effect is occurring. These abnormal returns cannot be explained by a horizontal demand curve.

The increased demand for the added stock, coupled with these stocks having downward shifting demand curves, leads to significantly increased abnormal returns; evidence of an index effect. However, according to *The Real Dogs of the Dow* (Arora, Capp, & Smith, 2008), stocks that were removed from the Dow tended to outperform stocks that were brought in to replace them. In their study, the authors created a portfolio of the stocks that were removed from the Dow and another portfolio for the stocks that were added to replace them. Changes to these portfolios took place whenever the Dow was once again modified or if a company could no longer be found in the Center for Research in Security Prices database. These portfolios tracked the performance of the stocks included in them from January 8, 1929, when the Dow was increased to thirty stocks, to December 31, 2006. The findings were that in thirty-two of fifty cases, the deletion stocks outperformed the stocks that were brought into replace them. The subtracted stocks tended to outperform the added stocks for five years before the difference began to level off. The average daily returns, over 250 trading days, were 0.00591 for the removed portfolio and 0.00436 for the addition portfolio, which translate to annual returns of 15.9% and 11.5%, respectively.

Arora et al argue that the market overreacts to the performance of the removed stocks, and that when these stocks regress to the mean, they experience higher returns than the stocks that were added. However, we suspect that if we account for risk, the return behavior (the short-run underperformance and long -run outperformance) can be explained differently than the previous offerings. Removing stocks from the Dow makes them inherently riskier than their replacement counterparts. After the initial removal shock leading to significantly negative abnormal returns, the risk premium generates the higher long-run returns, not regression to the mean.

The idea of risk-return tradeoffs has a rich history in financial theory. This can be seen in Harry Markowitz's Modern Portfolio Theory, which measures the returns of stock against the total risk of investing in the stock (Fabozzi, Gupta, & Markowitz, 2002). Risk, in this case measured by the variable sigma (σ), is quantified through the spread of the frequency distribution of stock returns, as wider spreads are indicative of a more volatile security – the actual returns vary to a greater degree from the mean or expected return. Since the deviations between the actual returns and the expected returns can be negative, these deviations are squared making each one positive. These squares are used to calculate the variance of the returns. The square root is then taken, making it easier to interpret, which is why standard deviation is the measurement of risk utilized in the MPT model.

The standard deviation of a portfolio is generally less than the weighted average of individual securities, as long as the securities are not perfectly correlated. This allows for a mean-variance optimization to be performed, generating every possible weighted portfolio which can all be charted between the axes of expected return and standard deviation. Within this charting there lies an efficient frontier along one of the curved edges for which every portfolio along this efficient frontier "results in the greatest possible expected return for that level of risk or results in the smallest possible risk for that level to expected return" (Fabozzi, Gupta, & Markowitz, 2002). When modeling the possible portfolios for a given set of assets within the framework of MPT, it can generally be seen that higher risks lead to higher returns.

There are other ways to measure risk besides standard deviation. There are two types of risk, if stocks are viewed through the lenses of portfolio investment. Unsystematic risk affects small groups or individual assets and can be factored out through diversification within a portfolio. Systematic risk affects the market as a whole and is the level of risk every investor must accept. To measure securities' response to systematic risks, the variable of beta is used. Within the Capital Asset Pricing Model (CAPM), beta is used to measure an asset's return relative to the return on the market (Fama & French, 2004). More specifically, the beta "measures the sensitivity of the asset's return to the variation in the market return" (Fama & French, 2004). The market is thought to have a beta of one; therefore, any stock with a beta higher than one is deemed more risky than the market as a whole.

$$(The\ CAPM\ Equation) \qquad R = \beta(R_M - R_F) + R_F$$

In the equation above, the expected return (R) is equal to the beta term multiplied by the market return premium ($R_M - R_F$) added to the risk-free rate (R_F), which the interest rate the model assumes allows all parties to borrow and lend money. As the returns of stocks and the market are changing constantly, a rolling regression is used to calculate beta. The beta seen as a stock's investment risk is actually a rolling regression involving 60 prior observations, or five years' monthly return data in this study. Again, it is generally believed in the case of beta that assets riskier than the market are more volatile but can also lead to higher returns.

While the CAPM is of great importance to the field of finance in both business and academics, it should be noted that the model has fallen under criticism in the past. Eugene Fama and Kenneth French acknowledge these theoretical shortcomings and oversimplifications in their 2004 paper in the *Journal of Economic Perspectives* titled "The Capital Asset Pricing Model: Theory and Evidence." Fama and French state that the model "offers powerful and intuitively pleasing predictions about how to measure the relation between expected return and risk" (2004). The authors also note that "the empirical record of the model is poor" (Fama & French, 2004). One of the main facets of the model is comparing a securities risk to that of the market portfolio; often, the S&P 500 is used as a proxy. However, as Fama and French note "a comprehensive 'market portfolio' that in principle can include not just traded financial assets, but also consumer durables, real estate, and human capital" (2004). Another oversimplification of the model lies in the assumption that all investors have the ability to borrow and lend at a unilateral risk-free rate(Fama & French, 2004), a theoretical rate at which there is absolutely no investment risk, yet in practice the rate at which people borrow and lend will vary based on different factors and will never be completely risk-free (the Treasury Rate is the typical proxy for the risk-free rate, which while incredibly safe, is not entirely free of risk). Nevertheless, beta is still used widely as a measure of risk for stocks – displayed on any investment summary of a publicly traded company on third-party platforms like *The Wall Street Journal's* stock portal or Yahoo! Finance – and, therefore, will be a measure of risk utilized by this study.

Data & Methodology

The accuracy of stock pricing information is crucial for ensuring definitive results for both parts of this study. The pricing data that was used in this study came from the Center for Research in Security Prices (CRSP) – the first research-quality database providing comprehensive stock price and return data. CRSP data is used throughout the academic and commercial finance community, hence the choice to use it as the source for this study.

The data set obtained for this study includes pricing information for the entirety of each company's time as publicly traded corporations. The set of stocks utilized in this study include changes made to the Dow from 1991-2013 ─ the time period of changes immediately following those covered by Beneish and Gardner through the most recent changes made prior to the beginning of this study ─ and can be seen in Table 1. As explained previously, multiple changes can occur to the index simultaneously, should the Dow Jones Averages Index Committee deem it necessary. This would appear to be the modern normality, as there are only two instances in the data where a single stock is replaced from the index.

Table 1. Stocks Added and Removed from the Dow Jones Industrial Average Post-1987

Announcement Date	Date of Change	Additions		Deletions	
May 2, 1991	May 6, 1991	Caterpillar Incorporated	CAT	Navistar International	NAV
		Walt Disney Company	DIS	USX Corporation	USX
		J.P. Morgan & Company	JPM	Primerica Corporation	PRI
March 12,1997	March 17, 1997	Traveler's Group	T	Westinghouse Electric	WX
		Hewlett-Packard Company	HPQ	Texaco Incorporated	CVX
		Johnson & Johnson	JNJ	Bethlehem Steel	BHMSQ
		Wal-Mart Stores	WMT	Woolworth	Z
October 26, 1999	November 1, 1999	Microsoft	MSFT	Chevron	CVX
		Intel	INTC	Goodyear Tire & Rubber	GT
		SBC Communications	SBC	Union Carbide	UK
		Home Depot	HD	Sears, Roebuck	SHLD
April 1, 2004	April 8, 2004	American International Group	AIG	AT&T	T
		Pfizer	PFE	Eastman Kodak	KODK
		Verizon Communications	VZ	International Paper	IP
February 11, 2008	February 19, 2008	Bank of America Corporation	BAC	Altria Group, Incorporated	MO
		Chevron Corporation	CVX	Honeywell International, Incorporated	HON
September 12, 2008	September 22, 2008	Kraft Foods, Incorporated	KRFT	American International Group	AIG
June 1,2009	June 8, 2009	Cisco Systems, Incorporated	CSCO	General Motors Corporation	GM
		The Travelers Companies, Incorporated	TRV	Citigroup, Incorporated	C
September 14, 2012	September 24, 2012	UnitedHealth Group Incorporated	UNH	Kraft Foods Incorporated	KRFT
September 10, 2013	September 23, 2013	The Goldman Sachs Group Incorporated	GS	Bank of America Corporation	BAC
		Nike Incorporated	NKE	Alcoa Incorporated	AA
		Visa Incorporated	V	Hewlett-Packard Company	HPQ

A complete list of the stocks observed in this study. This list begins with the changes to the Dow occurring after the Beneish and Gardner study through the most recent change that had taken place at the beginning of this study

I first replicated the Beneish and Gardner (1995) paper with post-1987 data, looking at the immediate consequences of stock returns when added or removed from the Dow. To assure that any abnormal returns from the stock returns were not a result of larger market movement, the prediction errors of each stock in the study were calculated. This was done by subtracting the returns of a market proxy, in this case the S&P 500, by the returns of the each stock. For each change to the Dow, the prediction errors were collected in 121 day windows surrounding the announcement date, which was set at a value of 0. The average of each day within these windows could then be calculated. The variance of this series can be calculated by the equation:

$$s_{APE}^2 = \frac{1}{120} \sum_{t=1}^{121} (APE_t - \overline{APE})$$

where $\overline{APE}$ is the mean of the 121 separate average prediction errors. Then, by finding the sum of the average prediction errors – the cumulative average prediction error (CAPE), calculated with the equation:

$$CAPE_{t,t+k} = \sum_{\tau=t}^{t+k} APE_\tau$$

– different trends can be seen for various lengths of time during the 121 days. The statistical significance of these trends can be calculated through the following equation:

$$t = \frac{CAPE_{t,t+k}}{\left[k s_{APE}^2\right]^{1/2}},$$

where k represents the number of days that the average prediction error is cumulated. This process was performed on both stocks added and removed from the Dow. Similar to the Beneish and Gardner study, this methodology was applied to individual stocks whose status with the Dow was changed, along with equally weighted portfolios of stock that were added and removed from the Dow simultaneously. This portfolio-based method was conducted "because the prediction errors of firms sharing the same event date in calendar time are likely to be correlated, and the t-statistics on average abnormal performance are likely to be biased away from zero" (Beneish & Gardner, 1995).

For this study, twenty-three stock changes were made to the Dow in the observed window of time, which had occurred over nine different instances. Therefore, the CAPE for individual firms incorporates twenty-three observations, while the CAPE composed of portfolios of the simultaneous changes incorporates nine observations. In Table 2, the CAPEs are shown for each individual day relative to the announcement of a change starting ten days prior and concluding 10 days afterward. This twenty-one day window is to show significance of returns surrounding the announcement, when the returns would most likely be prone to significant movement. The CAPEs of days outside this smaller twenty-one day window can be grouped together, as the stock returns are not expected to change to any significant degree.

The second part of the research included calculating risk-adjusted returns of stocks added and removed from the Dow before and after a change occurred. Two risk measures were utilized: standard deviation and beta. Both measures of risk for each stock were measured against the monthly returns of that stock as a gauge of its performance. Individual stock's performance, as well as hypothetical portfolios' performance, was observed. The monthly returns were calculated for the stocks added and removed from the Dow from 1988-2013. To track broad trends of these stocks, four categories were created for each instance surrounding a change being made to the index ─ removed-stocks-before, removed-stocks-after, added-stocks- before, and added-stocks-after ─ which measures the average monthly return over a five year period.

The total risk associated with investing in Dow stocks is calculated through standard deviation. To measure the total risk against the calculated returns, the same four categories were used for grouping the standard deviations; wherein the standard deviations of individual added and removed stocks were averaged five years before and five years after an alteration of the index. In an effort to further comprehend the overarching relationship between total risk and returns for Dow constituents, counterfactual portfolios are gathered by condensing the four categories via a simple average. The resulting counterfactual portfolios highlight the general instances of stocks' total risk and returns five years prior and following addition or removal from the index.

The systematic risk associated with investing in Dow stocks is quantified through the variable beta, where a value of one is equal to the "market risk." The beta variable was calculated using the CAPM with the monthly returns serving as a proxy for the expected return variable in the model. Beta is calculated by a rolling regression requiring 60 observations or five years of monthly stock returns. The market risk premium and the risk free rate were gathered from Kenneth French's website. The overarching relationship between beta and returns for Dow constituents was compared following the same logic as the standard deviations. Counterfactual portfolios composed of cumulative averages of the four categories were created.

Results

Index Effect

When looking at the short-term abnormal returns for stocks added (Table 2, Panel A) and removed (Table 2, Panel B) from the Dow post-1987 using the Beneish and Gardner framework, the results of this study remain consistent with their finding. The stocks removed from the Dow experienced significantly negative abnormal returns similar to those of the previous study. This can be observed on day zero of Table 2, Panel B – the day that the Dow announced a change in the index that would remove those stocks. Both the individual and portfolio CAPEs are statistically significant, similar to the Beneish and Gardner study. On day two (and day four in Panel A), both categories of CAPEs are also statistically significant, although it is unclear why this is the case.

However, the abnormal returns for the added stocks varied from the prior findings; my results showed significant positive abnormal returns, as both CAPEs are statistically significant. I attribute these findings to the index effect. The number of index funds tracking the Dow since the Beneish and Gardner study has increased greatly – as the Dow did not license their index for investable securities until 1997 (Fredman, 1998) – which in turn requires large blocks of the added corporations' stock to be purchased as these index funds rebalance leading to the positive abnormal returns. This result was not reached in any previously published literature from what could be gathered.

Table 2. Stock Price Reaction to Announcements of Changes in the DJIA: CAPE & t-statistics

Panel A. Additions			Individual Firms (N = 23)		Portfolios (N=9)	
	Days Relative to the Event	Number of Days Cumulated	CAPE	t-statistic	CAPE	t-statistic
	-60 , -2	59	-0.00035	-0.013108	-0.01965	-0.85041
	-20 , -11	10	0.00979	0.8880025	0.00701	0.73641
	-10	1	-0.00147	-0.422064	-0.00172	-0.57321
	-9	1	-0.00228	-0.654834	-0.00125	-0.41691
	-8	1	-0.00041	-0.116771	-0.00066	-0.22000
	-7	1	0.00069	0.1964558	0.00002	0.00614
	-6	1	-0.00366	-1.049406	-0.00268	-0.89100
	-5	1	-0.00488	-1.40009	-0.00433	-1.44087
	-4	1	0.00012	0.0344297	-0.00336	-1.11527
	-3	1	0.00296	0.84774	0.00288	0.95600
	-2	1	-0.00181	-0.519193	-0.00133	-0.44079
	-1	1	0.00167	0.477609	0.00283	0.94150
	0	1	0.00884	2.5335975*	0.00820	2.72682*
	1	1	0.00489	1.4012406	0.00518	1.72147
	2	1	0.00150	0.428771	0.00140	0.46397
	3	1	0.00350	1.0045433	0.00366	1.21790
	4	1	0.00710	2.0350277*	0.00711	2.36350*
	5	1	0.00156	0.448379	0.00124	0.41139
	6	1	-0.00051	-0.145947	0.00064	0.21283
	7	1	-0.00084	-0.240373	-0.00092	-0.30551
	8	1	0.00299	0.8571921	0.00145	0.48251
	9	1	-0.00135	-0.385898	-0.00158	-0.52598
	10	1	-0.00124	-0.356153	0.00088	0.29260
	11 , 20	10	-0.01547	-1.402888	-0.01239	-1.30226
	2 , 60	59	-0.01194	-0.445836	-0.00550	-0.23821
Window Statistics						
CAPE (-1,+1)		3	0.01539	2.5475275*	0.01622	3.11180*
Standard Deviation			0.007369635		0.00630	
Percentage Positive			51.72%		48.28%	

Panel B. Deletions			Individual Firms (N = 23)		Portfolio (N = 9)	
	Days Relative to the Event	Number of Days Cumulated	CAPE	t-statistic	CAPE	t-statistic
	-60 , -2	59	0.03064	0.48130	-0.00071	-0.00608
	-20 , -11	10	-0.00332	-0.12656	-0.00589	-0.12238
	-10	1	-0.00118	-0.14194	-0.00171	-0.11197
	-9	1	0.01310	1.58049	0.01523	0.99982
	-8	1	0.00066	0.07953	0.00464	0.30451
	-7	1	-0.00174	-0.20975	-0.00299	-0.19656
	-6	1	0.00792	0.95541	0.01007	0.66142
	-5	1	0.00054	0.06490	-0.00247	-0.16250
	-4	1	-0.00655	-0.79060	-0.01529	-1.00410
	-3	1	-0.00244	-0.29409	-0.00874	-0.57364
	-2	1	-0.01008	-1.21669	-0.01047	-0.68768
	-1	1	-0.01580	-1.90704	-0.03930	-2.58071*
	0	1	-0.02601	-3.13819*	-0.06354	-4.17197*
	1	1	-0.01486	-1.79356	-0.02975	-1.95324
	2	1	-0.02193	-2.64684*	-0.05091	-3.34286*
	3	1	0.01038	1.25248	0.02933	1.92603
	4	1	0.01022	1.23342	0.03826	2.51890*
	5	1	0.00837	1.01061	0.02539	1.66722
	6	1	0.00209	0.25252	0.00814	0.53452
	7	1	-0.01376	-1.66084	-0.03618	-2.37551*
	8	1	-0.00897	-1.08283	-0.01464	-0.96124
	9	1	0.00250	0.30198	0.00704	0.46230
	10	1	-0.00285	-0.34414	-0.01132	-0.74304
	11 , 20	10	0.01584	0.60432	0.02979	0.61866
	2 , 60	59	-0.03470	-0.54508	-0.01542	-0.13183
Window Statistics						
CAPE (-1,+1)		3	-0.05667	-3.948379394*	-0.13259	-5.026367*
Standard Deviation			0.019569432		0.02579	
Percentage Positive			56.52%		37.93%	

*Significant at the 5-percent level

This table shows the significance of abnormal stock return movements surrounding the day of an announcement of a change being made to the Dow. CAPE represents the "cumulative average prediction errors," which is simply the sum all the average prediction errors. The prediction errors are calculated by subtracting the returns of Dow stock by a market proxy – the S&P 500 returns in this case – to determine if any movement in returns is abnormal or simply the result of systematic market movement. These calculations were performed twice, once for individual stocks and once for portfolios of stocks. The portfolios consist of all the stocks that were added or removed from the Dow whenever a change was made to the index. Day zero represents the date the Dow announced a change was being implemented.

Long Run Risk-Return Tradeoffs

To begin observing the risk-return tradeoffs, it is necessary to understand the long run return behavior, as this will be utilized understanding how the two measures of risk – standard deviation and beta – fit into the hypothesis. The study conducted by Arora et al (2008) found that the portfolio of removed stocks outperformed the portfolio of added stocks in thirty-two of fifty cases. Translating this measure to this study would mean comparing the individual removed-stock-after and the individual added-stock-after counterfactual portfolios, which include stocks added or removed from the Dow in the year a change is made. Upon this comparison, the result is that the counterfactual portfolio of removed stock outperforms the counterfactual portfolio of added stock in six out of the eight instances measured, which can be seen in Table 3. This would suggest that the data supports the findings of Arora et al (2008) that removed stocks typically have higher returns on average than their added counterparts. The stocks removed from the index in September 2008 and 2009 have substantially better performances between the five year average prior to a change and the five year average preceding a change. It would appear that the Great Recession only seemed to amplify the observed trend. Another substantial difference in removed stock performance occurs in the 1991 change. More research should be pursued to see if this trend is a general occurrence during times of recession.

As hypothesized at the beginning of this study, the total risk calculations follow the theoretical models as the stocks removed from the Dow had standard deviations that were approximately three times as great as their added counterparts. Removed stock also had returns that were much larger than those of their added counterparts. This demonstrates the risk-return payoff Harry Markowitz proposed in MPT, which can be seen in Figure 2. However, the reasons for the increased risk associated with the removed stock most likely stray from those proposed to describe the initial shock of negative abnormal returns in Beneish and Gardner (1995).

Beneish and Gardner (1995) suggested that removed stocks succumb to information costs as a consequence of dissociation with the index. Yet, stock information and corporate financials are currently available in quantities never previously seen, even to the most casual of investors. Therefore, information costs surrounding publicly traded companies, especially corporations with blue chip stock, are minimized relative to the effect it had on the previous study. Yet, this phenomenon cannot be ruled out completely, as fewer analysts may follow a stock once removed from the Dow. The second explanation for the observed negative return behavior is liquidity effects, which, if anything, have only been magnified by the index effect. Perhaps the most fitting explanation for the increase in total risk of removed stock is a lack of investor confidence, in combination with an increased liquidity effect. Once a stock is removed from the Dow, it is seen as flawed by many, hence the excess supply of shares from the index funds (if investors had faith in the company's performance despite their newfound position relative to the Dow, the supply of shares on the market would not be in excess). While investor behavior is hard to quantify, this would seem to be a fitting explanation of the observed and calculated results.

Table 3. Average Monthly Returns

Year	Status	5 Year Average Before	5 Year Average After
1991	Added	0.70%	0.77%
	Removed	-0.47%	8.22%
1997	Added	0.10%	0.73%
	Removed	0.20%	2.34%
1999	Added	1.44%	-0.63%
	Removed	1.05%	-0.55%
2004	Added	-0.54%	-5.09%
	Removed	-0.43%	-0.25%
Feb 2008	Added	0.17%	0.51%
	Removed	1.54%	0.18%
Sept 2008	Added	0.29%	0.28%
	Removed	-2.64%	19.06%
2009*	Added	0.09%	1.16%
	Removed	-1.21%	9.11%
2012*	Added	0.96%	2.32%
	Removed	0.46%	-1.35%
2013*	Added	1.25%	N/A
	Removed	-0.29%	N/A

* Average returns after the change encompass return data through 2013, as opposed to complete 5 year data

The monthly returns averaged for added and removed stock over a five year period before and after a change is made to the Dow. Each of the four percentages per row represents a portfolio composed of stocks that were brought in or removed from the index in a given year.

The results generated by the beta variable differed from the risk-return framework, which it was hypothesized to follow. The resulting trend can be seen in Figure 3, where the returns of each counterfactual portfolio remain the same, yet the counterfactual portfolio containing the cumulative averages of the added stocks in a five year window following their inclusion outperforms its counterpart of the removed stocks. It is unclear why the hypothesized results and calculated results differ to the extent that they do. These results do seem to be supportive of the idea that the increased risk of removed stock is driven by investor confidence, or lack thereof. The belief surrounding the strength of a single organization after a setback (in this case, being taken out of the Dow) is not consequential to the stock market as a whole. Investor uncertainty of small, specific groups of stocks is therefore not accurately measured through traditional methods of tracking systematic risk.

A comparison of how each counterfactual portfolio compares for the different measures of risk can be seen in Table 4.

Table 4. Total Risk and Systematic Risk Counterfactual Portfolios

Risk Variable	Specified Counterfactual Portfolio			
	Removed Stock Before	Removed Stock After	Added Stocks Before	Added Stocks After
Sigma	0.105	0.375	0.096	0.126
Beta	1.181	1.018	1.001	1.187

The counterfactual portfolios for total risk (sigma) and systematic risk (beta). The data supports the hypothesis in terms of total risk (sigma), as the removed stock had standard deviations that were approximately three times as great as their added counterparts. However, the results of the beta variable differed from the hypothesized result.

Discussion

There are a few limitations to the data that should be addressed. First, the set of data is limited in scope. Applying the data that was available for this study to the framework laid out in the methodology poses some ambiguity when it comes to the later individual counterfactual portfolios. When it comes to the later changes made to the Dow (2009, 2012, & 2013), data is not yet available from five years after the change has taken place. As a result, the comparisons of added and removed stock in those final three instances incorporate data through the end of 2013, but no further, as noted in the footnote of Table 2.

Another limitation that should be addressed is the method of which beta is calculated potentially leading to the unexpected results observed. As mentioned in the methodology section, beta is calculated with a rolling regression of 60 observations, which spans five years of monthly returns.

Figure 2. Total Risk Counterfactual

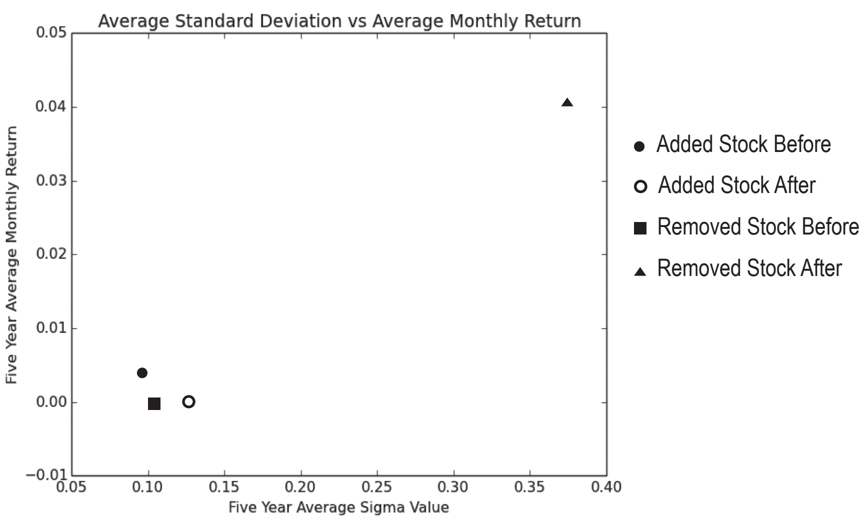
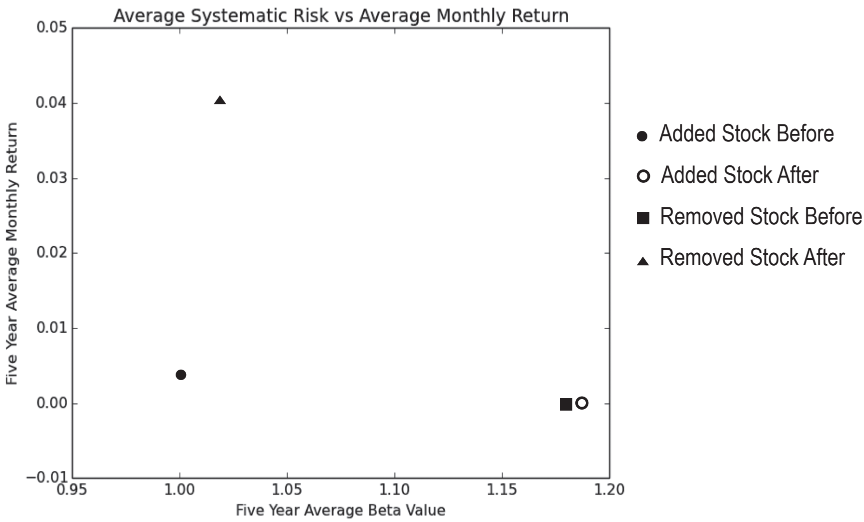


Figure 3. Systematic Risk Counterfactual



By the nature of this measurement, there are going to be points where the beta measurement incorporates data from both before and after a stock changed its position in the Dow. This could be a potential explanation as to why the resulting counterfactual portfolios defy the logic of the hypothesis. Based on the hypothesis, the removed-stock-after portfolio should be riskier in terms of beta than the removed-stock-before portfolio, and the added-stock-before portfolio should have a higher beta than the added-stock-after portfolio. However, based on Figure 3, the opposite is true. While beta is a widely used measure of risk in the market, between the shortcomings of its calculation in this framework and the logic that small groups of stock becoming associated/dissociated with the Dow does not necessarily reflect risk systematic to the market, this may not be the ideal criterion for this particular study.

When looking at returns, it cannot be overlooked that the Great Recession of 2008 did occur during the time period encompassed in the data set. While the Dow is meant to track the US equities market and the Recession was a market-wide collapse, it should be noted that some of the returns calculated during this event could be considered "outliers" from returns that occur in normal market conditions. The fact that two changes were made to the Dow within the same year, 2008, is abnormal in and of itself. Dating back to 1929, the year the Dow expanded from twenty stocks to thirty stocks, multiple changes occurring to the index within the same year only happened one other time. This other isolated incidence took place in 1930, a year that falls during the Great Depression.

Future possibilities for this research should include extending the study when data becomes available to complete the five year average for each of the counterfactual portfolios. It may also be interesting to apply this framework to changes that occurred throughout the entire *Real Dogs of the Dow* study, as this may allow for a better understanding whether risk or regression to the mean is truly the superior explanation for dissociated stocks outperforming their newly associated counterparts.

Another area of potential future research would be to apply different methods of measuring risk to see if any of those may better explain the observed behavior in this study, as beta does not seem to be the ideal measure for this study. One such model that may pose interesting results is measuring the R-squared value of added and removed stock. The R-square value "is a statistical measure that represents the percentage of a security's movements that can be explained by movements in a benchmark index" (Loth, 2007). In this case, the "benchmark index" would be the Dow. A higher R-squared value is representative of a security's movements being more closely correlated with those of the index. Theoretically, one would expect a removed stock to have a lower R-square value than its added counterpart.

Another area of future research could be conducting the decomposition of the total risk measurement. This study shows that removed stocks are, in fact, riskier when measuring the standard deviation, yet it is impossible to be completely sure in what way these stocks become riskier, unless due to. Perhaps, liquidity risks, similar to those noted in the Beneish and Gardner (1995) study and suggested in the explanation of their results. However, within the parameters of this study, a high level of skepticism remains as to what is impacting the total risk. It does not appear to be systematic risk.

Conclusions

To a large degree, the hypothesis presented at the beginning of this study was supported by the results. Stocks removed from the Dow continue to initially underperform following their removal from the Dow, only to outperform the replacement stocks in the long-run. While the initial significantly negative abnormal returns a stock experiences could be a result of information costs and liquidity effects as suggested by Beneish and Gardner, it is more likely explained as a result of index effects, where index fund managers dump the removed stocks in large quantities causing a steep drop in price, and thus a significant drop in returns, in the time period immediately after an announcement of the stocks' deletion from the Dow. The opposite effect is observed for the same reason by the added stock. An index effect is created by fund managers buying the added stock in large quantities, thus giving these stocks an abnormal increase in both price and returns.

While the measure of standard deviation supports the hypothesis of risk-return tradeoffs explaining the long-term outperformance of removed stock, the systematic risk measurement does not do so. While suggestions as to why this may be are found in the results and discussion sections, these explanations are untested as it falls beyond the scope of this study. However, it is an area that should be pursued by future research.

References

Arora, A., Capp, L., Smith, G. (2008). The real dogs of the Dow. *The Journal of Wealth Management*, Spring 2008, 64-72.

Beneish, M., Gardner, J. (March 1995). Information costs and liquidity effects from changes to the Dow Jones Industrial Average list. *The Journal of Financial and Quantitative Analysis*, Volume 30, No. 1, pp. 135-157.

Beneish, M., Whaley, R. (December 1996). An anatomy of the "S&P Game": the effects of changing the rules. *The Journal of Finance*, Vol. 51, No. 5, pp. 1909-1930.

Dow Jones Indexes (December 31, 2011). Dow Jones Industrial Average. Retrieved from http://www.djindexes.com/mdsidx/downloads/brochure_info/Dow_Jones_Industrial_Average_Brochure.pdf

Fabozzi, F., Gupta, F., Markowitz, H. (Fall 2002). Legacy of modern portfolio theory. *The Journal of Investing*, pp. 7-22.

Fama, E., French, K. (Summer 2004). The capital asset pricing model: theory and evidence. *Journal of Economic Perspective*, Volume 18, Number 3, pp. 25-46.

Fredman, A. (May 1998). A closer look at funds that focus on the Dow. *AAII Journal*, pp. 3-7.

Loth, R. (November 5, 2007). Measuring risk with alpha, beta, and Sharpe. *Forbes*. Retrieved from http://www.forbes.com/2007/11/05/risk-alpha-beta-pf-education-in_rl_11050investopedia_inl.html

Nationwide Financial (2013). Price-weighted vs. market capitalization-weighted indexes. Retrieved from http://commentary.nationwide.com/img/price-weighted-vs-market-capitalization-weighted-indexes.pdf

Paglia, John K. (2001). Has the Dow really escaped the bear? *Graziado Business Review*, Volume 4, Issue 2.

Shleifer, A. (1986). Do demand curves for stocks slope down? *The Journal of Finance*, 41(3), 579-590.

S&P Dow Jones Indices LLC (August 2013). Dow Jones Averages methodology. Retrieved from http://www.djindexes.com/mdsidx/downloads/meth_info/Dow_Jones_Industrial_Average_Methodology.pdf